

Xa1 Allelic *R* Genes Activate Rice Blight Resistance Suppressed by Interfering TAL Effectors

Chonghui Ji^{1,6}, Zhiyuan Ji^{2,3,6}, Bo Liu¹, He Cheng⁴, Hua Liu¹, Sanzhen Liu⁴, Bing Yang^{1,5,*} and Gongyou Chen^{2,*}

¹Division of Plant Sciences, C. Bond Life Science Center, University of Missouri, Columbia, MO 65211, USA

²School of Agriculture and Biology/State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University, Shanghai 200240, P.R. China

³National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R. China

⁴Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA

⁵Donald Danforth Plant Science Center, St. Louis, MO 63132, USA

⁶These authors contributed equally to this article.

*Correspondence: Bing Yang (yangbi@missouri.edu), Gongyou Chen (gyouchen@sjtu.edu.cn)

<https://doi.org/10.1016/j.xplc.2020.100087>

ABSTRACT

Xanthomonas oryzae pathovar *oryzae* (*Xoo*) uses transcription activator-like effectors (TALEs) to cause bacterial blight (BB) in rice. In turn, rice has evolved several mechanisms to resist BB by targeting TALEs. One mechanism involves the nucleotide-binding leucine-rich repeat (NLR) resistance gene *Xa1* and TALEs. Reciprocally, *Xoo* has evolved TALE variants, C-terminally truncated versions (interfering TALEs or iTALEs), to overcome *Xa1* resistance. However, it remains unknown to what extent the two co-adaptive mechanisms mediate *Xoo*–rice interactions. In this study, we cloned and characterized five additional *Xa1* allelic *R* genes, *Xa2*, *Xa31(t)*, *Xa14*, *CGS-Xo1₁₁*, and *Xa45(t)* from a collection of rice accessions. Sequence analysis revealed that *Xa2* and *Xa31(t)* from different rice cultivars are identical. These genes and their predicted proteins were found to be highly conserved, forming a group of *Xa1* alleles. The *XA1* alleles could be distinguished by the number of C-terminal tandem repeats consisting of 93 amino acid residues and ranged from four in *XA14* to seven in *XA45(t)*. *Xa1* allelic genes were identified in the 3000 rice genomes surveyed. On the other hand, iTALEs could suppress the resistance mediated by *Xa1* allelic *R* genes, and iTALE genes were prevalent (~95%) in Asian, but not in African *Xoo* strains. Our findings demonstrate the prominence of a defense mechanism in which rice depends on *Xa1* alleles and a counteracting mechanism in which *Xoo* relies on iTALEs for BB.

Keywords: *Xa2*, *Xa14*, *Xa45(t)*, TAL effector, iTAL effector, bacterial blight

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 Comm.* **1**, 100087.

INTRODUCTION

Major food crops suffer significant yield losses to an average of 20% due to damage caused by pathogens and pests (Savary et al., 2019). To ward off infections caused by microbial pathogens, host plants depend on defense responses that are induced upon the recognition of pathogen-derived molecules by plant immune receptors. The receptors are either intracellular nucleotide-binding leucine-rich repeat receptor proteins (NLRs) or plasma membrane-bound receptor proteins. NLRs recognize effectors inside plant cells that are delivered by pathogens, whereas membrane immune receptors recognize a wide range

of extracellular pathogen-derived signals at the plant–pathogen interface (Boller and Felix, 2009; Cook et al., 2015; Monteiro and Nishimura, 2018). The extracellular signals include general pathogen-associated molecular patterns or compounds released from plant cells during infection, and the recognition by membrane immune receptors activates basic and broad immunity to plants. To establish successful infections,

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, CAS.

Plant Communications

pathogens have evolved a diverse repertoire of effectors that are delivered into plant cells to interfere with immunity by acting as virulence factors. However, these effectors can be recognized by NLRs and the recognition triggers effector-specific immunity, including the hypersensitive response (HR) at the infection site (Boller and Felix, 2009; Cook et al., 2015). Pathogens can also use virulence mechanisms to prepare the host for pathogen infection by evading or suppressing the effector triggered immunity (Jackson et al., 1999; Feng and Zhou, 2012).

Bacterial blight (BB), caused by *Xanthomonas oryzae* pathovar *oryzae* (*Xoo*), is one of the most important diseases in rice. BB also provides a well-established model for studying host-microbe interactions (Niño-Liu et al., 2006). *Xoo* causes blight disease by colonizing and spreading along the vascular tissue of rice leaves, causing severe yield loss and damaging grain quality (Mew et al., 1993; Niño-Liu et al., 2006). Host resistance bestowed by genetically inheritable *R* genes is one of the most economically and environmentally sustainable approaches to control diseases, other than the use of chemicals and antibiotics (White and Yang, 2009). Forty-four *R* genes for BB have been identified, and several have been cloned and characterized (He et al., 2012; Triplett et al., 2016; Dilla-Ermita et al., 2017; Kim, 2018). The cloned *R* genes for BB can be classified into five groups based on their structural characteristics and mechanisms of action. They are receptor-like kinase genes (e.g., *Xa21*), NLR genes (e.g., *Xa1*), executor genes (e.g., *Xa27*), recessive *R* genes derived from the sugar transporter *SWEET* genes (e.g., *xa13*), and a variant of the transcription factor gamma subunit gene (*xa5*) (Song et al., 1995; Yoshimura et al., 1998; Iyer and McCouch, 2004; Gu et al., 2005; Chu et al., 2006). Unfortunately, an increasing number of *R* genes, such as *Xa10*, *Xa4*, and *Xa21*, when used as single major BB-resistant genes in rice breeding, were defeated in fields due to the rapidly emerging virulent strains of *Xoo* (Adhikari et al., 1999; Lee et al., 1999; Vera Cruz et al., 2000). Altering *R* gene recognizable avirulence (*avr*) genes and recruiting new virulence factors or resistance suppressors are two common mechanisms used by *Xoo* to evade and counteract rice immunity (Jiang et al., 2020).

Transcription activator-like effectors (TALEs), from *Xanthomonas* and *Ralstonia* bacteria, comprise the largest family of type III effector proteins in bacteria. TALEs are highly conserved and differ from each other by the number of central 34-amino acid (aa) repeats and the composition of two amino acids at positions 12 and 13 of each repeat, two unique features that determine the specificity of DNA binding to the promoters of host target genes (Boch et al., 2009; Moscou and Bogdanove, 2009). In addition, the C termini of TALEs contain nuclear localization signals and transcription activation domains, characteristic of eukaryotic transcription activators. A group of atypical TALE variants lacking C-terminal transcription activation domains, the so-called iTALEs (interfering TALEs [Ji et al., 2016]) or truncTALEs (truncated TALEs [Read et al., 2016]) have been identified. iTALEs exist in two forms; type A iTALEs are characterized by the C-terminal truncation of 103 aa due to a premature stop codon introduced by a C-to-T change in the coding sequence of the genes, whereas type B iTALEs are characterized by a loss of 229 aa and an addition of 10 aa due to a large deletion and frameshift of the coding sequence at the 3' end of the genes (Salzberg et al., 2008; Ji et al., 2016; Read et al., 2016).

Xa1 Allelic *R* Genes Activate Rice Blight Resistance

Members of the NLR superfamily, the critical components of the innate immune system in plants and animals, can detect a variety of microbial pathogens and protect hosts from the threat of disease (Jones et al., 2016; Benthams et al., 2017). *Xa1* is the only cloned NLR gene against BB from rice, and it can confer race-specific resistance to certain *Xoo* strains (Yoshimura et al., 1998). *Xa1* resistance is triggered by TALEs but suppressed by iTALEs (Ji et al., 2016), resulting in race specificity that is determined by the presence or absence of both TALEs and iTALEs, particularly in the *Xoo* strain (Ji et al., 2016). *Xa1* plants, including transgenic rice plants expressing *Xa1*, were resistant to *Xoo* strains lacking iTALE genes (e.g., Δ Tal3, a mutant of PXO99^A with its iTALE genes inactivated), and resistance was suppressed in the presence of *iTal3a* or *iTal3b* (Ji et al., 2016). Resistance responses include the HR (brown coloring at the inoculation site). Similar to *Xa1*, *Xo1*, which has not yet been cloned, in the Carolina Gold Select (CGS) rice variety was shown to confer resistance against several African *Xoc* (*Xanthomonas oryzae* pathovar *oryzicola*) strains in a TALE-dependent manner, and resistance was suppressed by truncated TALEs (Read et al., 2016; Triplett et al., 2016). It has been reported that *Xo1* in CGS is located within a region harboring 14 NLR *R* genes, and hypothesized that one of them (CGS-*Xo1*₁₁, allelic to *Xa1*) might be *Xo1* (Read et al., 2020).

The main feature of the predicted protein encoded by *Xa1* includes six nearly identical leucine-rich repeats at the C terminus, and the leucine-rich repeat (LRR) region is hypothesized to function in pathogenic molecule recognition and protein-protein interactions. It has been reported that all TALEs are the avirulent triggers for *Xa1* resistance, and they are widely present in *Xanthomonas* strains (Ji et al., 2016). iTALEs, probably the evolutionary derivatives of TALEs, share unique and conserved functional protein structures, namely two internal deletions at the N terminus and the lack of transcription activation domains at the C terminus. However, further studies are needed to determine how XA1 recognizes TALEs to initiate resistance and how iTALEs suppress *Xa1* resistance to *Xoo* in rice (Zuluaga et al., 2017).

We previously reported that *Xa1* is a broad-spectrum *R* gene that recognizes TALEs independent of the TALE central repeat regions. However, the function of this presumably excellent *R* gene was masked by iTALEs (Ji et al., 2016). In this study, we report the cloning of *Xa1* allelic genes from different rice varieties and show that they have high variations at their C termini, mainly in the LRR regions. Surprisingly, iTALEs have different suppression profiles compared with *Xa1* allelic *R* genes. Although the resistance of *Xa1* *R* alleles can be defeated by new virulence factors, an investigation into the structural and resistance features of *Xa1* *R* alleles can help us to understand the resistance mechanism and suppression basis of iTALEs, and potentially help us to engineer new *R* genes for an expanded spectrum of BB resistance.

RESULTS

Four Rice Varieties, Each with an *R* Gene, Display a Resistance Spectrum Similar to *Xa1*

Previous studies have mapped *Xa2* in IRBB2 (He et al., 2006), *Xa14* in IRBB14 (Bao et al., 2010), *Xa31(t)* in Zhachanglong

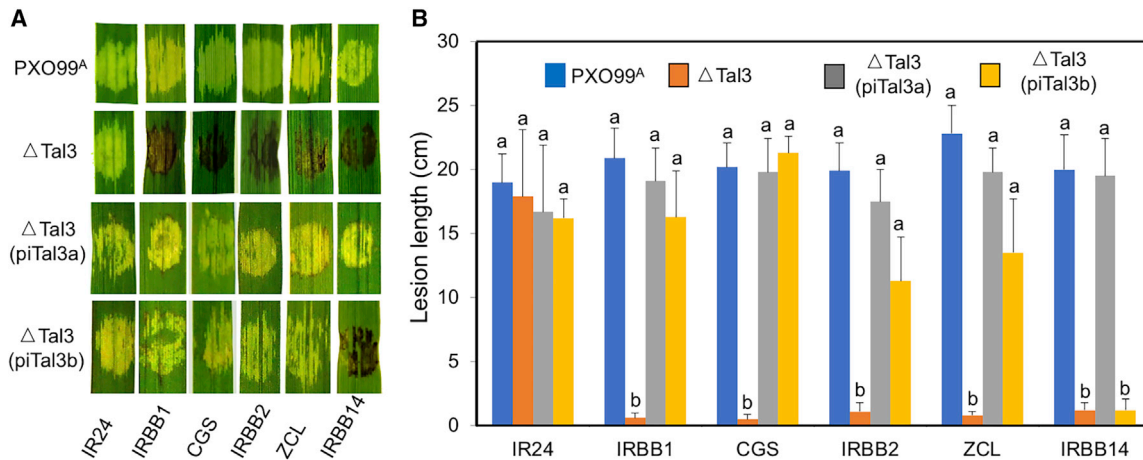


Figure 1. Several Rice Varieties Show Disease (Resistant or Susceptible) Responses to *Xoo* Strains Lacking or Carrying iTALE Genes.

(A) Hypersensitive responses (dark brown inoculation spots) or water soaking symptoms (clear inoculation spots) in different rice varieties. CGS, Carolina Gold Select; ZCL, Zhachanglong.

(B) Lesion length measurements of different rice varieties infected with *Xoo* strains as indicated. Error bars indicate standard deviations. Identical lower-case letters indicate no significant difference ($p < 0.01$). The experiments were repeated three times with similar results.

(ZCL) (Wang et al., 2009), and *Xo1* in CGS (Triplett et al., 2016) within the regions corresponding to the *Xa1* locus in IRBB1. We asked whether these loci, in which all four respective *R* genes have not yet been cloned, could confer disease responses similar or different from that of *Xa1*, i.e., susceptible to PXO99^A and resistant to its mutant strain Δ Tal3. As shown in Figure 1 and Table 1, rice varieties (IRBB2, CGS, ZCL, and IRBB14), such as IRBB1, were resistant to Δ Tal3 and susceptible to PXO99^A in terms of the HR and lesion length. Except for IRBB14, their resistance to Δ Tal3 could be suppressed by the introduction of *iTal3a* or *iTal3b*. Their resistance spectra were also examined with an additional 24 *Xoo* isolates that contained different iTALE genes as confirmed by PCR with iTALE gene-specific primers and the analysis of their available genome sequences. Five *Xoo* strains (KXO85, JW89011, T7174, Aust-2013, and Aust-R3) were avirulent to the five rice varieties (IRBB1, IRBB2, CGS, ZCL, and IRBB14), whereas a PXO86 mutant containing only a type B iTALE gene and PXO112 were incompatible with IRBB14. The African strain AXO1947, which contains no known iTALE gene, was avirulent to IRBB1 and CGS, but virulent to three rice varieties (IRBB2, ZCL, and IRBB14). The remaining *Xoo* strains were virulent to the five varieties, which exhibited a susceptible phenotype, such as IR24, a rice variety that contained no known dominant *R* gene and served as a control (Table 1). These findings indicate that the four rice varieties are similar to IRBB1 (*Xa1*) in response to TALE- and iTALE-containing *Xoo* strains, except for IRBB14 (*Xa14*).

Xa2, *Xa31(t)*, *CGS-Xo1₁₁*, and *Xa14* Confer *Xa1*-like Resistance

We hypothesized that the *R* genes in the four tested rice varieties that conferred resistance to Δ Tal3 and susceptibility to PXO99^A were allelic to *Xa1*. To test this hypothesis, we PCR-amplified the respective genomic fragments from IRBB2, IRBB14, and ZCL using primers 53120-F4b and 53120-R4b, and from CGS, we used 53120-F6b as the forward primer, which was specific to both *Xa1* in IRBB1 and *xa1* in Nipponbare, and 53120-R4b

as the reverse primer (see sequence information in Supplemental Table 1). The individual genomic fragments were sequenced and used to transform Kitaake, a *japonica* variety susceptible to most *Xoo* strains (Oliva et al., 2019). The transgenic T0 plants, as long as they were positive for transgenes as verified by PCR, were resistant to Δ Tal3, with the lesion length ranging from 0.1 to 0.7 cm, whereas the wild-type control and transgene-negative lines were susceptible to Δ Tal3, with the lesion length ranging from 11.3 to 14.7 cm (Supplemental Figure 1). Similar results were obtained by the segregation analysis of transgenes and resistance phenotypes in selected T1 and T2 families derived from more than two independent resistant T0 plants. The transgenic plants also showed resistance or susceptibility to the 26 *Xoo* strains, similar to their donor rice (Table 1). These results demonstrate that the genes from these four rice varieties are allelic to *Xa1* and function similar to *Xa1*, in that they are resistant to *Xoo* strains lacking iTALE genes.

Xa1, *Xa2*, *Xa31(t)*, *CGS-Xo1₁₁*, and *Xa14* Form a Group of *Xa1* Allelic *R* Genes

We then analyzed the sequences derived from the genomic clones that complemented Kitaake with resistance to Δ Tal3. The clone from IRBB2 contained 8997 nucleotides with a predicted gene structure similar to that of *Xa1*, consisting of two introns and three exons. The predicted amino acid sequences of XA2 consisted of an N-terminal zinc finger BED (zFBED) domain, an NB-ARC domain, and strikingly, the C-terminal, a highly conserved tandem repeat of 93-aa residues within the LRR region (Supplemental Figure 2). XA2 had five 93-aa repeats, whereas XA1 contained six 93-aa repeats (Figure 2). In addition to the repeat regions, the remaining 1244 aa were identical between XA2 and XA1 (Supplemental Figure 2). By comparison, *Xa31(t)* encoded a protein almost identical to XA2 (with only a 1-aa difference in the entire 1709-aa predicted protein), including the five 93-aa tandem repeats. In addition, the CGS clone contained a gene encoding five 93-aa repeats. It was 99.6% identical to

Xoo strain	iTALE type A	iTALE type B	IR24	IRBB1	CGS	IRBB2	ZCL	IRBB14	Kit	Kit-Xa1	Kit-CGS-Xo1 ₁₁	Kit-Xa2	Kit-Xa31(t)	Kit-Xa14
PXO99 ^A	+	+	S	S	S	S	S	S	S	S	S	S	S	S
ΔTal3	–	–	S	R	R	R	R	R	S	R	R	R	R	R
ΔTal3/iTAL3a	+	–	S	S	S	S	S	S	S	S	S	S	S	S
ΔTal3/iTAL3b	–	+	S	S	S	S	S	R	S	S	S	S	S	R
PXO86	+	+	S	S	S	S	S	S	S	S	S	S	S	S
PXO83ΔTal3	–	+	S	S	S	S	S	R	S	S	S	S	S	R
PXO83ΔTal6	+	–	S	S	S	S	S	S	S	S	S	S	S	S
PXO61	+	+	S	S	S	S	S	S	S	S	S	S	S	S
PXO79	+	+	S	S	S	S	S	S	S	S	S	S	S	S
PXO112	–	+	S	S	S	S	S	R	S	S	S	S	S	S
KXO85	–	–	S	R	R	R	R	R	S	R	R	R	R	R
JW89011	–	+	S	R	R	R	R	R	S	R	R	R	R	R
K202	+	+	S	S	S	S	S	S	S	S	S	S	S	S
T7174	–	+	S	R	R	R	R	R	S	R	R	R	R	R
H75373	+	+	S	S	S	S	S	S	S	S	S	S	S	S
Xoo2	+	–	S	S	S	S	S	S	S	S	S	S	S	S
A3842	+	+	S	S	S	S	S	S	S	S	S	S	S	S
PbXo7	+	–	S	S	S	S	S	S	S	S	S	S	S	S
IXO56	+	+	S	S	S	S	S	S	S	S	S	S	S	S
NXO260	+	+	S	S	S	S	S	S	S	S	S	S	S	S
CIAT1185	+	+	S	S	S	S	S	S	S	S	S	S	S	S
ZHE173	+	–	S	S	S	S	S	S	S	S	S	S	S	S
C1	–	+	S	S	S	S	S	S	S	S	S	S	S	S
GD1358	+	+	S	S	S	S	S	S	S	S	S	S	S	S
HB21	–	+	S	S	S	S	S	S	S	S	S	S	S	S
Aust-2013	–	+	S	R	R	R	R	R	S	R	R	R	R	R
Aust-R3	–	+	S	R	R	R	R	R	S	R	R	R	R	R
AXO1947	–	–	S	R	R	S	S	S	S	R	R	S	S	S

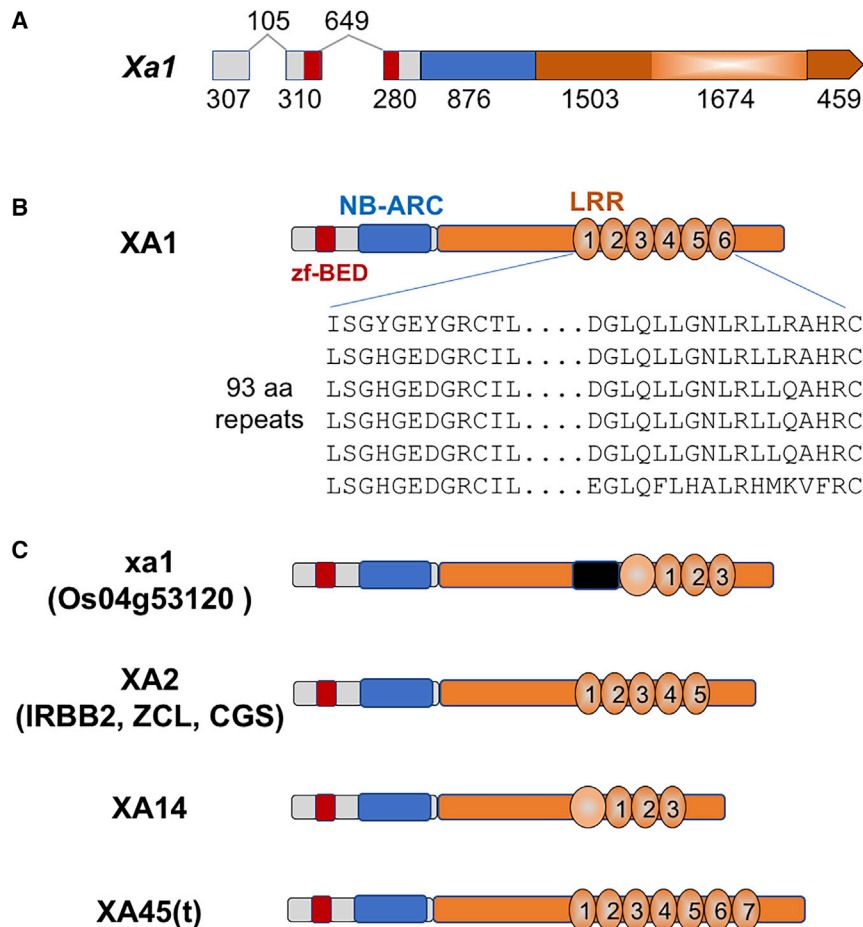
Table 1. Disease Spectrum of Cultivars and Transgenic Plants Containing the Genomic Clones of Xa1 Allelic R Genes.

Note: CGS, Carolina Gold Select; ZCL, Zachanglong; Kit, Kitaake. The disease scores were determined by lesion length, followed by leaf tip clipping. R represents the disease reaction with a lesion length shorter than 5 cm, whereas S indicates the reaction with a lesion length longer than 5 cm (usually longer than 10 cm).

Xa2 and 99.2% identical to XA2, and very similar to CGS-Xo1₁₁ as reported by Read et al. (2020). Our CGS-Xo1₁₁ differed from that of Read et al., 2020 in 17 nucleotides, which corresponded to 7 aa all occurring within the repeat region, probably due to sequencing errors. For example, the CGS-Xo1₁₁ of Read et al., 2020 had a frameshift “A” deletion at position 78 in the fourth repeat. Due to the high similarity, we refer to the gene in our genomic clone as CGS-Xo1₁₁ (Supplemental Figure 2). It is obvious that Xa2 and Xa31(t) are identical, whereas CGS-Xo1₁₁ is not. Furthermore, Xa14 encodes a protein that is slightly different from XA1 in terms of structure (Figure 2). XA14 contains an imperfect first repeat of 135 aa, followed by three 93-aa repeats. Beyond their repeat regions, XA14 and XA1 are 96% (1199/1245 aa) identical (Supplemental Figure 2).

Prevalence of Xa1 Allelic R Genes in Rice

Based on the conserved sequences between xa1 (susceptible) and Xa1 (resistant) alleles in the database, as well as our newly cloned Xa1 allelic R genes, the significant difference between the resistant and susceptible alleles was an insertion of 1305 bp preceding the first 279-bp (for 93-aa) repeat in xa1 (Figure 2C). We therefore chose a 36-bp sequence (5'-tctctgccaccttccgca/atcagtggtatggagaa-3') that spanned the junction of the first 279-bp repeat and its preceding region as the feature sequence for the functional Xa1 allelic R genes. The 36-bp feature sequence was used to scan 3000 rice genome sequences to identify rice accessions that contained the feature sequence. A total of 463 accessions were found to contain at least ten reads with perfect



matches to the feature sequence, suggesting that approximately 15.4% of the accessions contained functional *Xa1* allelic *R* genes. There were 994 (33.1%) accessions that contained *Xa1* alleles, that is, if the reads of the matches were extended to include at least one hit (Supplemental Table 2).

We next screened our collection of rice accessions using a PCR approach with two pairs of primers that specifically detect the presence of the *Xa1* feature sequence (XaL-F1 and XaL-R1). We also determined the number of repeat sequences (XaL-F2 and XaL-R2) (Figure 3A). As shown in Supplemental Table 3, 16 out of 87 accessions were positive for the *Xa1* signature sequence. Among them, the repeats ranged from four to seven, with the majority of accessions having five or six repeats (Figure 3B). To determine whether alleles confer BB resistance similar to *Xa1*, 22 accessions, including several accessions lacking the *Xa1* signature sequence, were inoculated with PXO99^A, Δ Tal3, and the complemented strains of Δ Tal3 with either *iTal3a* or *iTal3b*. Indeed, accessions that contained *Xa1* allelic genes exhibited the HR phenotype in response to Δ Tal3, and the disease phenotype (water soaking) after the inoculation of both PXO99^A and Δ Tal3 with either *iTal3a* or *iTal3b* (Figure 3C and 3D and Supplemental Table 3), indicating that the resistance by *Xa1* allelic *R* genes was specifically suppressed by iTALEs.

To further confirm whether the gene predicted to encode the seven 93-aa LRR repeats was indeed functional, the genomic fragment

Figure 2. Gene and Protein Domain Structures of XA1 Allelic Members.

(A) *Xa1* consists of three exons (bars) and two introns (lines). Sequences encoding zf-BED, NB-ARC, and LRR are denoted as red, blue, and orange bars, respectively. Nucleotide numbers are shown below the bars or above the lines.

(B) Domains of XA1 are presented as zf-BED, NB-ARC, and LRR. Partial sequences of 99 amino acids are denoted below the numbered circles.

(C) Domains of XA1 alleles similar to XA1 are presented. *xa1* from Nipponbare contains a sequence (black bar) upstream of the first imperfect repeat.

was PCR-amplified with 53120-F4b and 53120-R4b from one accession (*Oryza nivara*-1, IRGC 102463), similar to the approach used to clone the other *Xa1* allelic genes. The genomic clone in pCAMBIA1300 was introduced into Kitaake and the primary transgenic plants were assessed for resistance. As expected, the transgenic plants (T0 and T1 generations) that were positive for the transgene were resistant to Δ Tal3 but susceptible to PXO99^A (Figure 4). The genomic clone was Sanger sequenced and found to contain a gene allelic to *Xa1* (Supplemental Figure 2). We named the new gene *Xa45(t)*.

Prevalence of iTALE Genes in *Xanthomonas oryzae* Pathovars

The wide spread of *Xa1* allelic *R* genes in nature might create significant selection pressure for the adaptation of *Xanthomonas oryzae* (*Xo*) pathogens, particularly *Xoo* populations in rice fields. It is believed that cognate iTALE genes exist in the two *Xo* pathovars. To survey the prevalence of iTALE genes among *Xo* isolates that represent the geographic and genetic relatedness of the pathogens (Oliva et al., 2019), we surveyed the database for the presence of iTALE genes among the *Xo* genome sequences deposited in NCBI (<https://www.ncbi.nlm.nih.gov>, Supplemental Table 5 for a list of strains with a prevalence of iTALE genes). Out of 42 *Xoo* strains from Asia and Australia, 40 (~95%) strains contained type A, type B, or both types of iTALE genes, whereas none of the 33 African *Xoo* genomes contained iTALE genes (Supplemental Table 5). On the other hand, ten (100%) *Xoc* genomes from Asia contained either type B or both A and B iTALE genes, whereas two out of three African *Xoc* strains contained type B iTALE genes (Supplemental Table 5). These findings reveal the prevalence of both *Xa1* allelic *R* genes in Asian rice and iTALE genes in *Xanthomonas oryzae* populations in Asia, as well as the co-adaptation of pathogens and host plants in BB and leaf streak diseases in Asia.

DISCUSSION

In this study, we identified several rice varieties that were resistant to the *Xoo* strain (Δ Tal3) lacking iTALE genes, and the resistance was suppressed by *Xoo* strains carrying

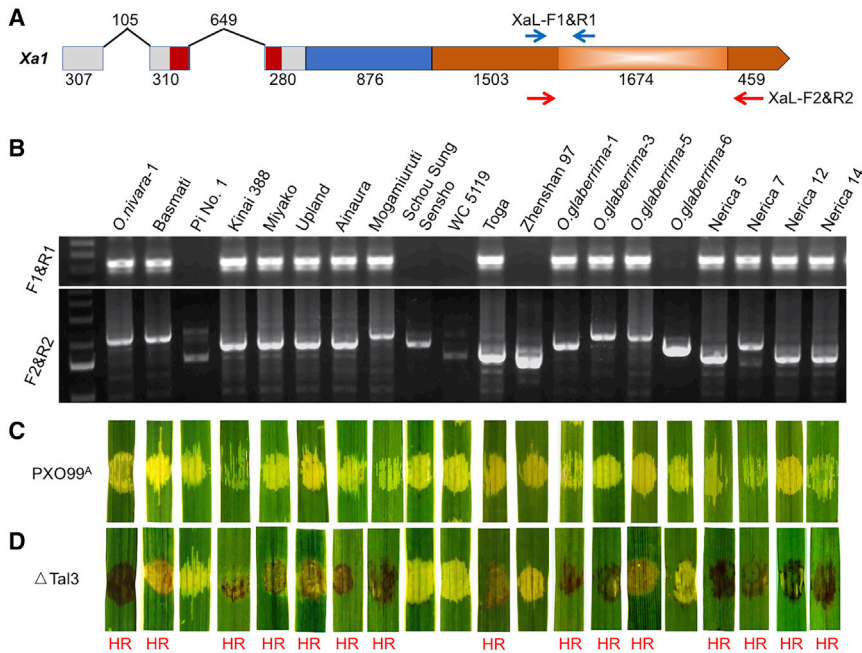


Figure 3. Prevalence of Xa1 Allelic Genes in Representative Rice Accessions.

(A) Gene structure of Xa1 in which two pairs of primers were used to detect the signature sequences of Xa1 alleles in 20 rice varieties.

(B) Gel images of PCR amplicons of the junction of the first repeat of 279 nucleotides and its upstream region derived from primers XaL-F1 and XaL-R1 (upper panel), as well as the whole repeat region derived from primers XaL-F2 and XaL-R2 (lower panel).

(C) Water soaking (susceptible) phenotypes of 20 rice varieties in response to PXO99^A inoculation through the syringe infiltration of leaves.

(D) Hypersensitive reaction (HR) and water soaking phenotypes of 20 rice varieties in response to ΔTal3 inoculation through the syringe infiltration of leaves.

iTALE genes. The resistance was mediated by several previously identified but as-yet uncloned R genes allelic to Xa1. These genes are highly conserved on both nucleotide and predicted amino acid levels, and the most distinguishable feature is the variable repeat number of the C-terminal 93-aa residues, with the repeats ranging from four in XA14 to seven in XA45(t). Based on the signature sequence of Xa1 allelic members, a search of 3000 rice genome sequences revealed that approximately 15% of the accessions contained the Xa1 allelic R genes. Reciprocally, most Asian Xoo and Xoc strains (>95%) contained iTALE genes, probably due to co-adaptation or arms races between the host plant and its pathogens.

The cloning of a gene usually involves mapping of the gene and functional complementation of the lines otherwise lacking the gene or the functional allele. For Xa2, Xa31(t), and Xa14, all three R genes were mapped in their respective donors within the regions, wherein Xa1 in IRBB1 was cloned (He et al., 2006; Wang et al., 2009; Bao et al., 2010). The transgenic Kitaake plants containing putative Xa2, Xa31(t), and Xa14 genomic clones were resistant to Xoo strains lacking iTALEs and were susceptible to Xoo strains carrying iTALE genes. The phenotypes of resistance and susceptibility of the transgenic plants to the tested Xoo strains were identical to the donors, IRBB2, Zhachanglong, and IRBB14, respectively. These data evidence the functional identities of the three R genes. As Xa2 and Xa31(t) are identical, we refer to them as Xa2. However, it remains to be determined whether CGS-Xo1₁₁ is Xo1, which confers resistance to Xoc. We currently do not have the avirulent Xoc strains used by Triplett et al. (2016) to determine the Xoc resistance spectrum of the Kitaake CGS-Xo1₁₁ transgenic plants. However, CGS-Xo1₁₁ is at least an NLR gene resistant to Xoo, which is supported by the notion that the Kitaake transgenic plants with CGS-Xo1₁₁ had the same resistance spectrum against the tested Xoo strains as the donor CGS.

features of the zfBED domain flanked by two nuclear localization signals and variable 93-aa tandem repeats in the LRR region. The XA1 allelic members are highly conserved and can be distinguished by the number of 93-aa repeats. It remains to be determined whether repeat number plays a role in the functionality and specificity of the members. It is also unknown whether the 93-aa repeats play a role in the recognition of TALEs for the initiation of resistance to BB. As TALEs and iTALEs contain variable numbers of 34-aa repeats, it is reasonable to speculate that there are physical interactions between XA1 allelic members and TALEs for resistance and the interference of the interaction by iTALEs as a decoy for susceptibility.

The prevalence of both Xa1 allelic R genes in rice and TALE genes in Xoo creates an extraordinary challenge for Xoo to proliferate in rice, which is the only known host. Therefore, it is important to further investigate the prevalence of iTALE genes in Xoo, particularly in Asia where conditions are more conducive to BB and where epidemics of blight disease frequently occur (Niño-Liu et al., 2006). To understand how iTALEs interfere with the resistance activated by the recognition of TALEs by Xa1 or its allelic members will provide the basis for the production of broad-spectrum resistance against iTALE-containing Xoo pathogens and mitigate the yield loss caused by the pathogens in rice.

METHODS

Plant Materials, Bacterial Strains, Medium, and Growth Conditions

The seeds of several rice varieties (Supplemental Table 3) used in this research were kindly provided by the International Rice Research Institute and the U.S. National Small Grains Collection.

All plants were grown in a greenhouse and growth chamber at 30°C with a 12-h light period and at 28°C with a 12-h dark period and relative humidity of 60%–75%. Escherichia coli strains were grown in Luria-Bertani medium supplemented with appropriate antibiotics at 37°C. Agrobacterium

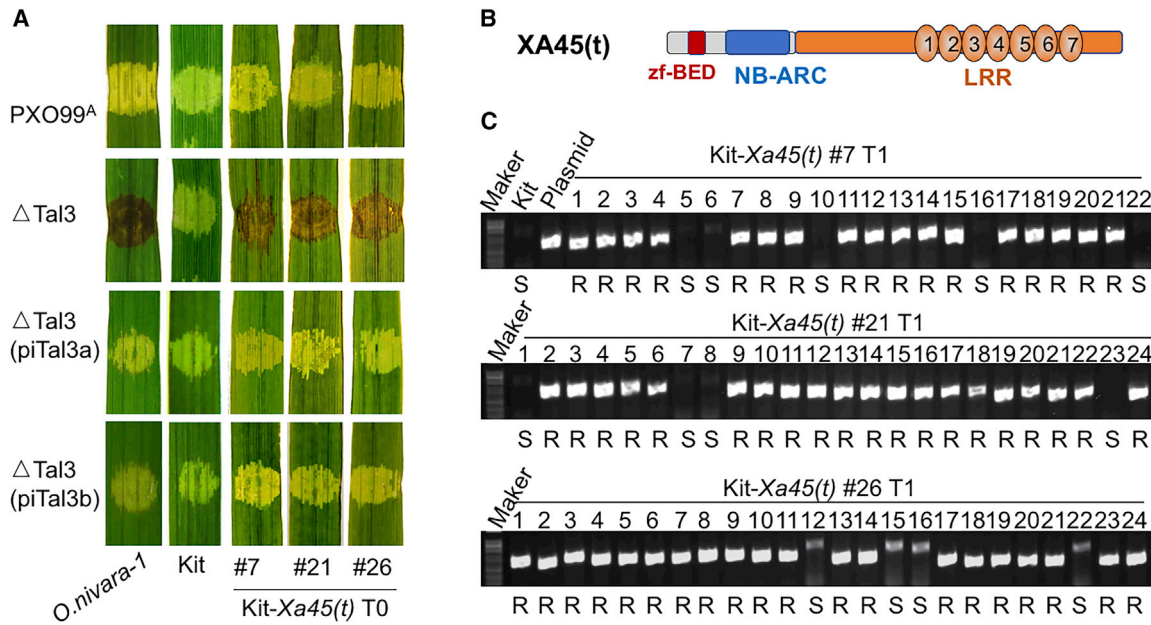


Figure 4. Functional Complementation of *Xa45(t)*.

(A) Three transgenic Kitaake (Kit) lines of the first generation (T0) in response to the inoculation of different *Xoo* strains as indicated on the left side of each panel. The hypersensitive response (dark brown inoculation spots) or water soaking (clear inoculation spots) symptoms were scored 3 days after inoculation.

(B) Predicted domains of *Xa45(t)* as indicated by zf-BED, NB-ARC, and LRR with seven 93-aa repeats.

(C) Association of genotypes and phenotypes of *Xa45(t)* transgenic Kitaake plants in the T1 generation. Gel images of PCR amplicons specific for *Xa45(t)* of individual progeny plants from three lines. Disease reactions are characterized as “S” for susceptibility to bacterial infection when lesion lengths were >10 cm and “R” for resistance when lesion lengths were <5 cm 12 days after leaf tip clipping inoculation with Δ Tal3.

tumefaciens strains were grown at 30°C in the dark. All *Xoo* strains were grown at 28°C in TSA medium (10 g/l tryptone, 10 g/l sucrose, 1 g/l glutamic acid). Antibiotics were used at the following concentrations if required: 100 μ g/ml ampicillin, 10 μ g/ml cephalixin, 25 μ g/ml chloramphenicol, 25 μ g/ml kanamycin, 100 μ g/ml spectinomycin, and 10 μ g/ml tetracycline. Bacterial strains and DNA plasmids used in this study are listed in Supplemental Table 4.

Rice Transformation with the Genomic Fragments of R Genes

For the functional complementation of R genes, with *Xa2* as an example, an 8997-bp genomic sequence containing a 1596-bp promoter sequence, the entire coding region (5885 bp) of *Xa2*, and a 1516-bp downstream sequence, were PCR-amplified using primers 53120-F4b and 53120-R4b (Supplemental Table 1), cloned into the binary pCambia1300 vector, and linearized with *EcoRI* and *HindIII* using the Gibson Assembly Master Mix (New England BioLabs, Ipswich, MA, USA). After Sanger sequencing, the pCambia1300-*Xa2* plasmid was electroporated into the *Agrobacterium tumefaciens* EHA105 strain and introduced into calli that were induced from the immature embryos of the *japonica* variety, Kitaake. The rice transformation method was used as described previously (Hiei et al., 1997). Transgenic plants were genotyped with primers BB2-F1 and M13F (Supplemental Table 1). The genomic fragments of the other R genes (*Xa31(t)*, CGS-*Xo1₁₁*, *Xa14*, and *Xa45(t)*) were PCR-amplified and introduced into Kitaake by *Agrobacterium*-mediated transformation as indicated.

iTal3 and *iTal6* Deletion in PXO86

Based on the PXO86 genomic sequence (NCBI accession, CP007166.1), two pairs of primers, 86TalKDF1/86TalKDR1 and 86TalKDF2/86TalKDR2 (Supplemental Table 1), were used to amplify the upstream and downstream regions flanking *iTal3* (a type A iTALE gene) and *iTal6* (a type B iTALE gene) using PXO86 genomic DNA as the template. A pair of primers, KD13-F and KD13-R, was used to amplify the whole expres-

sion cassette of the kanamycin resistance gene from plasmid pKD13 (NCBI accession, AY048744.1). The three fragments were fused by overlapping PCR, and the resulting fragment was inserted into the pBluscript SK vector at the *EcoRV* site. The plasmid was introduced into the competent cells of PXO86 by electroporation, and the transformants were grown on TSA plates supplemented with kanamycin. Single colonies with deletions were identified by PCR using two pairs of primers, 86TAL3-F2/R2 and 86TAL6-F2/R2.

Disease Assays

The assay for the HR through leaf infiltration using a needleless syringe and the assay for the lesion length through leaf tip clipping were performed as described previously (Yang and Bogdanove, 2013). In brief, *Xoo* stocks, preserved in a -80°C freezer, were streaked on TSA plates supplemented with appropriate antibiotics and incubated at 28°C for 2–4 days. Cells were harvested from the plates, suspended in sterilized water and washed twice, re-suspended in water, and adjusted to an optical density of 0.5 at 600 nm. Bacterial cells in suspension were infiltrated into the leaves of 3-week-old rice plants through a blunt (needleless) syringe pressed against the underside of the leaf. The inoculation spot turned brown, and bacterial spread was restricted in resistant plants 2–3 days after infiltration, whereas the symptoms in susceptible plants were much stronger, with water soaking extending in both directions and bacterial exudate appearing at the inoculation spot. Fully expanded leaves of rice plants (6–8 weeks old) were inoculated using the leaf tip clipping method. Scissor blades were immersed in *Xoo* suspension and used to clip approximately 1–2 cm from the leaf tip. The lesion length was measured 12–14 days after inoculation. Three replications with approximately ten leaves from two to five plants per replicate were inoculated per strain. One-way analysis of variance (ANOVA) was conducted on all measurements. The Tukey honestly significant difference test was used for post-ANOVA pairwise tests for significance, which was set at 5% ($p < 0.05$).

Plant Communications

Xa1 Feature Sequence searching in 3000 Rice Genomes

The raw read data of 3000 rice genomes were downloaded from the Sequence Read Archive (Wang et al., 2018). Reads were subjected to adaptor clipping and quality trimming with Trimmomatic software (v0.36), keeping trimmed reads longer than 50 bp (Bolger et al., 2014). The resistant *Xa1* allele-associated sequence 5'-tctctgccacctccgca/atcagtggtatggagaa-3' (36 bp) was used to identify rice accessions potentially harboring resistant alleles. Reads carrying perfectly matched sequences were counted and normalized to the number of total reads. The sequence was absent in the reference genome (IRGSP-1.0) that carried the susceptible allele. Reads were also directly aligned to the 279-bp repeat sequence that encoded the first 93-aa LRR repeat of XA1. Alignments were conducted with the mem module of bwa (0.7.12-r1039) (Li and Durbin, 2009). Read alignments required matches of at least 50 bp with identities of 95%. Aligned reads were counted and normalized to the number of total reads.

Sequence Analysis of iTALE Genes in Sequenced *Xo* Genomes

The complete or partial genome sequences of *Xoo* and *Xoc* strains were obtained from the NCBI (Supplemental Table 5) via BLASTn using a 690-bp sequence at the 5' coding region of *iTal3a* from PXO99^A. Distinctive C-terminal sequences of iTALE genes (*iTal3a* and *iTal3b*) of PXO99^A were used to categorize the iTALE genes as type A and type B. All annotations were performed through NCBI Blast and further confirmed using SnapGene software.

ACCESSION NUMBERS

The NCBI GenBank accession numbers for *Xa2*, *Xa31(t)*, *CGS-Xo1₁₁*, *Xa14*, and *Xa45(t)* are as follows: MT395899, MT395900, MT395901, MT395902, and MT395903, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Plant Communications Online*.

FUNDING

This work was supported by the National Natural Science Foundation of China (31830072 to G.C.), the United States Department of Agriculture National Institute of Agriculture and Food (2017-67 013-26 521 to B.Y.), the Young Elite Scientist Sponsorship of the China Association for Science and Technology (2017QNRC001 to Z.J.), and the Chinese National Transgenic Major Program (2016ZX08001-002 to G.C.).

AUTHOR CONTRIBUTIONS

C.J., S.L., G.C., and B.Y. conceived the experiments. C.J., Z.J., H.C., B.L., and H.L. performed the experiments. C.J., Z.J., and B.Y. wrote the manuscript with input from all other co-authors.

ACKNOWLEDGMENTS

We are grateful to the U.S. National Small Grains Collection (NSGC), Dr. Zhikang Li (Institute of Crop Sciences, Chinese Academy of Agricultural Sciences), and Dr. Bin Han (Institute of Plant Physiology and Ecology, Chinese Academy of Sciences) for providing rice accessions. No conflict of interest is declared.

Received: February 23, 2020

Revised: June 3, 2020

Accepted: June 8, 2020

Published: June 20, 2020

REFERENCES

- Adhikari, T.B., Basnyat, R.C., and Mew, T.W. (1999). Virulence of *Xanthomonas oryzae* pv. *oryzae* on rice lines containing single resistance genes and gene combinations. *Plant Dis.* **83** (1):46–50.
- Bao, S., Tan, M., and Lin, X. (2010). Genetic mapping of a bacterial blight resistance gene *Xa14* in rice. *Acta Agronomica Sin.* **36**:6.
- Bentham, A., Burdett, H., Anderson, P.A., Williams, S.J., and Kobe, B. (2017). Animal NLRs provide structural insights into plant NLR function. *Ann. Bot.* **119**, 827–702.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**:1509–1512.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114–2120.
- Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* **60**:379–406.
- Chu, Z., Yuan, M., Yao, J., Ge, X., Yuan, B., Xu, C., Li, X., Fu, B., Li, Z., Bennetzen, J.L., et al. (2006). Promoter mutations of an essential gene for pollen development result in disease resistance in rice. *Genes Dev.* **20**:1250–1255.
- Cook, D.E., Mesarich, C.H., and Thomma, B.P. (2015). Understanding plant immunity as a surveillance system to detect invasion. *Annu. Rev. Phytopathol.* **53**:541–563.
- Dilla-Ermita, C.J., Tandayu, E., Juanillas, V.M., Detras, J., Lozada, D.N., Dwiyantri, M.S., Vera Cruz, C., Mbanjo, E.G.N., Ardales, E., Diaz, M.G., et al. (2017). Genome-wide association analysis tracks bacterial leaf blight resistance loci in rice diverse germplasm. *Rice (N Y)* **10**:8.
- Feng, F., and Zhou, J.M. (2012). Plant-bacterial pathogen interactions mediated by type III effectors. *Curr. Opin. Plant Biol.* **15**:469–476.
- Gu, K., Yang, B., Tian, D., Wu, L., Wang, D., Sreekala, C., Yang, F., Chu, Z., Wang, G.L., White, F.F., et al. (2005). R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**:1122–1125.
- He, Q., Li, D., Zhu, Y., Tan, M., Zhang, D., and Lin, X. (2006). Fine mapping of *Xa2*, a bacterial blight resistance gene in rice. *Mol. Breed.* **17**:6.
- He, W., Huang, D., Li, R., et al. (2012). Identification of a resistance gene *bls1* to bacterial leaf streak in wild rice *Oryza rufipogon* Griff. *J. Integr. Agric.* **11**:8.
- Hiei, Y., Komari, T., and Kubo, T. (1997). Transformation of rice mediated by *Agrobacterium tumefaciens*. *Plant Mol. Biol.* **35**:205–218.
- Iyer, A.S., and McCouch, S.R. (2004). The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. *Mol. Plant Microbe Interact* **17**:1348–1354.
- Jackson, R.W., Athanassopoulos, E., Tsiamis, G., Mansfield, J.W., Sesma, A., Arnold, D.L., Gibbon, M.J., Murillo, J., Taylor, J.D., and Vivian, A. (1999). Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola*. *Proc. Natl. Acad. Sci. U S A* **96**:10875–10880.
- 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, N., Yan, J., Liang, Y., Shi, Y., He, Z., Wu, Y., Zeng, Q., Liu, X., and Peng, J. (2020). Resistance genes and their interactions with bacterial blight/leaf streak pathogens (*Xanthomonas oryzae*) in rice (*Oryza sativa* L.)—an updated review. *Rice (N Y)* **13**:3.
- Jones, J.D., Vance, R.E., and Dangl, J.L. (2016). Intracellular innate immune surveillance devices in plants and animals. *Science* **354**:aaf6395.
- Kim, S.M. (2018). Identification of novel recessive gene *xa44(t)* conferring resistance to bacterial blight races in rice by QTL linkage analysis using a SNP chip. *Theor. Appl. Genet.* **131**:2733–2743.

- Lee, S.W., Choi, S.H., Han, S.S., Lee, D.G., and Lee, B.Y. (1999). Distribution of *Xanthomonas oryzae* pv. *oryzae* strains virulent to Xa21 in Korea. *Phytopathology* **89**:928–933.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754–1760.
- Mew, T.W., Alvarez, A.M., Leach, J.E., and Swing, J. (1993). Focus on bacterial blight of rice. *Plant Dis.* **77**:8.
- Monteiro, F., and Nishimura, M.T. (2018). Structural, functional, and genomic diversity of plant NLR proteins: an evolved resource for rational engineering of plant immunity. *Annu. Rev. Phytopathol* **56**:243–267.
- Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* **326**:1501.
- 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.
- Oliva, R., Ji, C., Atienza-Grande, G., Huguet-Tapia, J.C., Perez-Quintero, A., Li, T., Eom, J.S., Li, C., Nguyen, H., Liu, B., et al. (2019). Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* **37**:1344–1350.
- Read, A.C., Moscou, M.J., Zimin, A.V., Perte, G., Meyer, R.S., Purugganan, M.D., Leach, J.E., Triplett, L.R., Salzberg, S.L., and Bogdanove, A.J. (2020). Genome assembly and characterization of a complex zFBED-NLR gene-containing disease resistance locus in Carolina Gold Select rice with Nanopore sequencing. *Plos Genet.* **16**:e1008571.
- Read, A.C., Rinaldi, F.C., Hutin, M., He, Y.Q., Triplett, L.R., and Bogdanove, A.J. (2016). Suppression of Xa1-mediated disease resistance in rice by a truncated, non-DNA-binding TAL effector of *Xanthomonas oryzae*. *Front. Plant Sci.* **7**:1516.
- 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* PXO99^A. *BMC Genomics* **9**:204.
- Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* **3**:430–439.
- 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.
- Triplett, L.R., Cohen, S.P., Heffelfinger, C., Schmidt, C.L., Huerta, A.I., Tekete, C., Verdier, V., Bogdanove, A.J., and Leach, J.E. (2016). A resistance locus in the American heirloom rice variety Carolina Gold Select is triggered by TAL effectors with diverse predicted targets and is effective against African strains of *Xanthomonas oryzae* pv. *oryzicola*. *Plant J.* **87**:472–483.
- Vera Cruz, C.M., Bai, J., Ona, I., Leung, H., Nelson, R.J., Mew, T.W., and Leach, J.E. (2000). Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proc. Natl. Acad. Sci. U S A* **97**:13500–13505.
- Wang, C., Wen, G., Lin, X., Liu, X., and Zhang, D. (2009). Identification and fine mapping of the new bacterial blight resistance gene, Xa31(t), in rice. *Eur. J. Plant Pathol.* **123**:6.
- 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.
- White, F.F., and Yang, B. (2009). Host and pathogen factors controlling the rice-*Xanthomonas oryzae* interaction. *Plant Physiol.* **150**:1677–1686.
- Yang, B., and Bogdanove, A. (2013). Inoculation and virulence assay for bacterial blight and bacterial leaf streak of rice. *Methods Mol. Biol.* **956**:249–255.
- Yoshimura, S., Yamanouchi, U., Katayose, Y., et al. (1998). Expression of Xa7, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. *Proc. Natl. Acad. Sci. U S A* **95**:1663–1668.
- Zuluaga, P., Szurek, B., Koebnik, R., Kroj, T., and Morel, J.B. (2017). Effector mimics and integrated decoys, the never-ending arms race between rice and *Xanthomonas oryzae*. *Front. Plant Sci.* **8**:431.

Plant Communications, Volume 1

Supplemental Information

***Xa1* Allelic *R* Genes Activate Rice Blight Resistance Suppressed by Interfering TAL Effectors**

Chonghui Ji, Zhiyuan Ji, Bo Liu, He Cheng, Hua Liu, Sanzhen Liu, Bing Yang, and Gongyou Chen

SUPPLEMENTAL INFORMATION

***Xa1* allelic *R* genes activate rice blight resistance suppressed by interfering TAL effectors**

Chonghui Ji^{1,6}, Zhiyuan Ji^{2,3,6}, Bo Liu¹, He Cheng⁴, Hua Liu¹, Sanzhen Liu⁴, Bing Yang^{1,5,*} and Gongyou Chen^{2,*}

¹Division of Plant Sciences, C. Bond Life Science Center, University of Missouri, Columbia, MO 65211, USA

²School of Agriculture and Biology/State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University, Shanghai 200240, P.R. China.

³National Key facility for Crop Gene Resources and Genetic Improvement (NFCRI), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R. China

⁴Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA

⁵Donald Danforth Plant Science Center, St. Louis, MO 63132, USA

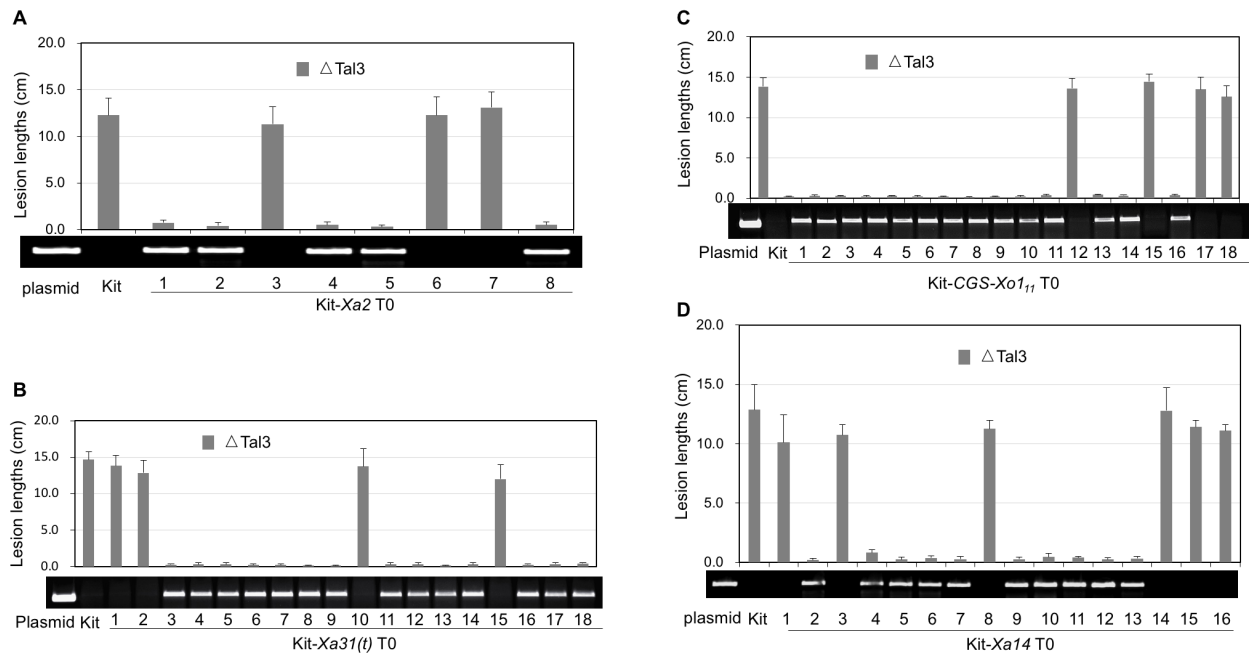
⁶These authors contributed equally to this article.

***Correspondence: Bing Yang (yangbi@missouri.edu), Gongyou Chen (gyouchen@sjtu.edu.cn)**

Contain:

Supplemental Figure 1

Supplemental Table 1 to 5.



Supplemental Figure 1. Association of genotypes and phenotypes of the *Xa2*, *Xa31(t)*, *CGS-Xo11* and *Xa14* transgenic Kitaake plants in T0 generation. Lesion lengths (centimeter) derived from inoculation of Δ Tal3 are presented above the gel images of PCR-amplicons from individual plants with gene specific primers.

Supplemental Figure 2 Nucleotide sequences of and predicted amino acid sequences encoded by *Xa2*, *Xa14*, *Xa31(t)*, *CGS-Xo11* and *Xa45(t)*. The genomic fragments cloned into pCAMBIA1300 were Sanger sequenced. The promoter sequences are in green lower letters. The sequence for start and stop codons are red and shaded in yellow; between them are introns in blue lower letters shaded in grey, the sequences encoding for 93 amino acids repeats are alternating red and blue upper letters. The terminator sequences are in black lower letters. The predicted amino acids sequences follow the DNA sequences of individual genes.

>*Xa2*

```

ttcaggaagacgggactgaatacatatcttcttaattccaaaggcattcaggatgtttatgaaagataattactgttg
tttttaggaaacatattacatcggcgcttgtcaggtaagctgacatagagtgatgttcatggtttcaaagtgaa
tatataattcctcgcgatggacaactttactatgcttctgtcaaggagatgtaaggcaaacagtaacagggattttt
gccaaactacaaactaggaagagatagacctaccgaaaatattatgtataaagggttcttgggaatattagaaaa
atataatgacaccgtagcgttagcacgggcatattactagtagaaciaaaagaaacacttatgcatttatacccaaaa
aaaaaacgtaattaatggaccgcatttctgactgtgtageggcccgctcgaagcataaactcaaaccttttaagccagag
atgggctgtggccaacatcgacataaactggaataagtccacttgccttccctcaactttacatggagtggtgatga
tggccctaattcccaataccagatgtcgaccctcaagtcttcaaaacctgcaagtagttcctcgaggggataga
ccatatttccctggtttccattgatgtggcatatatacagggcccactacctgaaccatcgccacgtccagtccttt
cggcctgttcgcatgacgagctcggccaccacctcaacatcgtccatgctgcgggcatggcgaggagacagctggg
acggtgaacctcatagctcctgtctgttgcaacatcctctcgccattctcacgcccacgacacccttgtcctcctg
ctcctgtagcagcctcgcaagctcacgcccggcctcatcgaggtcccctgtccccgagccgcttctagcatcgatga
gctgaggagctccccactgcccaggcgatgacgatttggcagggagggtcgcattcgcccactccaattcgagcgcg
tcacacctccccctccccctcctattcatccaccaccaacgcggataggagaagggatagaggaaagtgtgctgga
cacttacatattggttccatataatcttcttactcggatgccacgtcagataaaaacaaccgtcgggtcaacatttttc
gtagtggggattagggagctcatacatgactcgacgcagagctgagggatggcaagtgaacttattcccataaactt
ccaaaagtgtggaatgctcaacttgggcttctcttccggatgctgaagtgctaattgggccaataaaccaactagccca
cagcccacgcaagtggaaagccgacgagtcacccgtcaacgcggcagtagtaatccgagaaaaacgcggggtcgcgcgaa
gcaatcgtatcatcatcagacgattaatccacgcgaccaggagcgtgccccgagccacacgcgccacgcactggt
gaatcttctctcttccactagttttctcactctcctctcctactactactctgcttaccatcctcctgctt
gttgagccctcttgacacgccttggccacagtcacaaggactgctcctgttccgATGAGGAGGTGGAAGCCGGT
TGGCTGGAGGGCGGGATCAGGTGGCTGGCGGAGACCATCCTGGATAACCTGGACGCCGACAAGCTGGATGAATGGAT
TCGCCAGATTAGGCTCGCCGCTGACACCGAGAAGCTACGGGCTGAGATCGAGAAGGTGGATGGGGTGGTGGCTGCCG
TGAAGGGGAGGGCGATCGGGAACAGGTGCTGGCCCGATCGCTCGGCCGTCTCAGGGGGTTGCTGTACGACGCCGAC
GATGCGGTGACGAGCTCGACTACTTCAGGCTCCAGCAGCAGGTGAGGGAGGAGgtactgtcctttgcatatataccg
tgcttttaattaagtgtgcaagctgcgttgcctgcaacaatggtgtattggcgtcagtttccaatccatgcttgtg
ctacagTTACTACACGGTTTTGAGGCTGAAGAGACGGTCCGAGATGGAGCAGAGGACGAGGACGATATTCGATGGAC
AATACTGATGTACCGGAGGCAGTGGCGGCAGGCAGCAGCAAGAAACGGTCCAAGGCATGGGAACACTTTACTACCGT
AGAGTTCAGTCTGACGGGAAGGATTCTAAAGCACGGTGAAGTACTGCCACAAGGACCTATGTTGCACATCTAAGA
ACGGGACATCAGCTTTGCGCAACCATCTCAATGTTTGAAGAGGAAACGTGTAACAAGTACTGACCAACCCGGTAAAT
CCATCAAGgtaatgctaattggagttctgaatcttagtgtaaatccggttgaagtgtaaatttggcccgttacatctgct
caagatctcattctgtctcctaattcttctaataagccaactcatggtcatttttttctctaataatataagtagccggtgat
ggtgcaccaaagtgaattagatgcaaggaacaaaagtgaacaattgtatatacaaatataattatatactaaaaca
tgagtgtgtatcaaatccaattctttcaaaaatctactatgcaaaaattgagtgacaaaatctgctgcctttttttt
tttacagaaagcaaccaattaatataagtcaaatataaaaacgcttctgtagtctccaataaaaatagctcattgtttc
gtttatacttatgtttataaatttaaatttaaaacttaattttggagttgattttgtggttttcttttctcctatt
ttattttacaacatttgattttgaatagtttaagaatcgtatataaaaaattttaccataagttatttttaaatgt
ttaataaactcgaaggaataatcaagataagtgaaacgattctcctcctcatctacttaagatttgcgttataattgc
tgacctttctaatacgcctaaccacgatcacatgctcttccagTGCCGGTGAGGGTGCATCAAATGCAACTGGTAATT
CAGTTGGCAGAAAAGGATGAGAATGGATGGGACTTCAACACACCACGAGGCAGTTAGCACGCACCCTTGGAAACAAG
GCTGAACTTTCCAACAGGATCCAATGCATGACTCATCAGTTAGAAGAGGCTGTAAATGAGGTTATGAGGCTATGTCTG
ATCCTCAAGTTCAAACCAGAGTCGACAGGGTACACCACCGCCACAAATGCAACAACATCGTCTTATCTTCCGGAGC
CCATAGTGTATGGGAGGGCTGCAGAGATGGAAACCATCAAACAGCTGATCATGAGCAATAGATCTAATGGCATAACC
GTCCTGCCAATTGTAGGCAATGGAGGGATAGGAAAAACCACTTTGGCGCAACTGGTCTGCAAAGATCTGGTAATTAA
AAGTCAGTTTTAATGTTAAGATATGGGTGTATGTATCTGATAAATTTGATGTAGTTAAGATTACAAGGCAGATTTTGG
ATCATGTCTCCAACCAGAGCCACGAAGGAATAAGCAACCTTGATACGCTTCAGCAGGATCTTGAGGAACAAATGAAA

```

TCTAAGAAGTTTCCTCATTGTCTTAGATGATGTGTGGGAAATCCGTACAGATGACTGGAAAAACTACTGGCTCCTTT
AAGACCTAATGATCAGGTGAATTCATCACAGGAAGAGGCAACAGGTAATATGATAATTTTGACAACCTCGTATACAGA
GTATTGCCAAAAGTCTTGGAACAGTACAATCAATTAAGTTAGAAGCTCTGAAAGATGACGATATATGGTCACTATTT
AAAGTGCATGCTTTTGGTAATGATAAACATGATAGTAGTCCAGGCTTACAGGTTCTTGGGAAGCAAATGCTAGCGA
GCTAAAAGGCAACCCACTGGCAGCAAAAACCTGTGGGTTCACTATTAGGAACGAATCTTACCATCGATCATTGGGATA
GCATTATAAAGAGTGAAGAATGGAAATCCCTGCAACAAGCTTATGGCATCATGCAAGCGCTGAAGTTGAGCTATGAT
CATCTATCCAACCCCTTACAGCAATGCGTCTCTTATTGTTCTCTTTTCCCAAGGGTATTCTTTTTCAGCAAAGCACA
ACTAATACAAATATGGATTGCTCAAGGATTTGTGGAAAGAATCCAGTGAGAAGTTGGAGCAGAAAGGATGGAAATATC
TAGCTGAGTTGGTAAATTCGGGTTTTCCTTACAGCAAGTTGAAAGCACACGGTTTTTCATCAGAATATTTTGTATGCAC
GATCTTATGCATGATTTAGCGCAAAAAGGTTTACAAAACAGAATATGCAACTATAGATGGCTCAGAGTGCACAGAGTT
AGCCCCAAGTATACGCCATTTGTCAATAGTAAGTATTGTTTGCATACCAGCAAGGAGAAATATAGAAACATATCTCGTA
ATGAGGTGTTTTGAGAAAAGGTTGATGAAAGTTAAGTCAAGGAGTAAGTTGAGGTCACTGGTATTAATTTGGCAATAT
GATTCTCATTTTTTTAAATATTTCAAAGATGCTTTCAAGGAAGCACAACATCTGCGACTGCTGCAGATCACTGCAAC
TTATGCTGATTCTGATTCAATTTCTCTCCAGTTTGGTAAATTTACACATCTCCGGTATCTGAAAATTTGTGACCGAAG
AATCCGGCAGAACTTTGCCCGATCTCTAAGGAAGTATTACCATCTTCAAGTACTAGATATTGGCTATAGATTTGGA
ATTTCCCGTATATCTAATGATATAAATAATCTTCTCAGCCTGCGGCATCTTGTTCATATGATGAAGTGTGTTCTTC
CATTGCTAACATTGGTAAAATGACCTCACTTACAGGAAGTGGCAATTTTATTGTTTCAAGAAATTTAAGTGGTTTTG
AGGTGACACAATTGAAATCCATGAACAAGCTTGTACAACCTTAGTGTGTCTCAGCTTGAAGTGTAGAACTCAGGAG
GAGGCATGTGGGGCAAAACTGAAAGACAAAACAACACTTAGAAAAGCTACATTTGTCCTGGAAGGATGCATGGAATGG
ATATGACAGTGCAGAAAGCTATGAAGATGAATACGGCAGTGATATGAATATAGAAACAGAAGGGGAGGAAGTGTGAG
TTGGTGATGCCAATGGTGCACAAAGCTTACAACATCACAGTAATATAAGCTCTGAACTTGCTTCAAGTGAAGTGTGCTC
GAAGGTCTTGAACCACATCACGGCCTCAAGTATCTACGGATATCTGGGTATAATGGATCTACCTCCCAACTTGGCT
TCCTTCTTCACTTACCTGTCTGCAACACTTTCATCTAGAAAAATGTGGAAAATGGCAAATACTTCTTTTAGAAAGGC
TAGGGTTACTTGTAAAGCTCGTGTGATCAAAATGAGGAATGCAACAGAAGTCTCAATCCCTTCACTGGAGGAGCTT
GTGTTAATTGCATTGCCAAGCTTGAACACATGCTCCTGCACCTCCATCAGGAAGTGAAGTCCAGTTTAAAGGTTCT
GAAAATTAAGAATTGCCCTGTACTGAAGGATTTCCCTTGTGTTGAGATTTCCAGAAAATTTGAAATCGAGCGGACGT
CGCTCATGGTTGCCCCATCTTAGCAAGCTTACCATCTATAATTTGTCCTCTTTCTGTGTGCACAGTTCTCTGCCACCT
TCCGCAATCAGTGGTTATGGGAATATGGAAGGTGTACCCTTCCGCAATCACTTGAGGAAGTTCATCCATGAGTATGTA
TTCTCAAGAAACTCTGCAGCCCTGCTTTTTCAGGGAACCTCACTCTCCTGAGAAAATTTACATGTACTGGGAAACTCAA
ATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTTCAAAGCTGTGAGTCTCTTAGTTCT
CTGGATGGCTTGCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTTCATGGAGAAGATGG
AAGGTGTATCCTTCCGCAATCACTTGAGGAAGTTCATCCATGAGTATTCTCAAGAAACTCTGCAGCCCTGCTTTT
CAGGGAACCTCACTCTCCTGAGAAAATTTACATGTAATGGGAAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGC
ACAGCACTCGAAGAGTTGATAATTTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAATCT
CAGGTTGCTGCGGGCACATAGATGCCTCAGTGGTTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGG
AACTTTACATCCATGAGTATTCTCAAGAAACTCTGCAGCCCTGCTTTTTCAGGGAACCTCACTCTCCTGAGAAAATTA
CATGTACTGGAAAATCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTTCAAAG
CTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCA
GTGGTTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAAGTTCATCCATGAGTATTCTCAAGAA
ACTCTGCAGCCCTGCTTTTTCAGGGAACCTCACTCTCCTGAGAAAATTTACATGTACTGGGAAACTCAAATTTAGTGTCT
TCTGCAGCTCCATTCTGCACATCACTCGAAGAGTTGAAAATTTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCT
TGCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTTCATGGAGAAGATGGAAGGTGTATC
CTTCCGCAATCACTTGAGGAAGTTCATCAGTGAAGTATTCTCTAGAAAATTTGCAGCCCTGCTTCTGACGAATCT
CACCTGCTTAAAACAATTTAGAGGTATCAGGCACCACAAGTTTAAAATCTCTAGAACTGCAATCATGCAGTGCAGTGC
AACATTTGAAGATTTCAAGGTTGTGCGTCTGCTTGTACATTGGAGGGGTTGCAATTTCTCCACGCCCTCAGGCATATG
AAAGTATTCAGATGCCTGGCTTGGCTCCATATTTGGGAGTTTCGTGAGAGCAGGCTATGAGCTATGCCCACGACT
GGAAAGGCTCGACATCGATGACCCCTCTATCTTACCACGTCGTTCTGCAAGCACCTCACCTCCCTCCAACGCCTAG
AGCTTAACTATTGCGGAAGTGAAGTGGCAAGACTAACGGATGAGCAAGAGAGAGCGCTTCACTCCTCACGTCCCTG
CAAGAGCTCCGGTTTTAAGTATTGCTACAATCTCATAGATCTTCTGCGGGGCTCCACAGCCTTCCCTCCCTCAAGAG
GTTGGAGATCCGGAGTTGCAGGAGCATCGCGAGGCTGCCGGAGAAGGGCCTCCACCTTTCGTTTGAAGAAGTGGATA
TCATCGCTTGCAGTAATGAGCTAGCTCAGCAGTGCAGAACTCTAGCAAGCACTCTGAAGGTCAAATTAATGGGGGA
TATGTGAAC TGAattactcgggtggtggttaggegcacetttttccacctgcccactggeggtgggctcggttcaggcg
ttcaagctgctgtaaatccattgcccgaatgacgaccttcagaaccgttacacaatacaaaggacatatgatggct
agatcaactgtcgcagagctagctttgggttcacctgaaaacataaggccaaacgcggtggttcttttaatacagaagta
tcaaaaattgggttttgggttttaacgaaaagatggttagaaattgggttttaactgagcagattctgtaccagaattct
gtattttgctccgtaattctgtgctacgactgatggtgatttaactgatagcaatcgaccggcaaaaccaatccgc
tccaatgcattgtcacatacaaattctataattctatctactgaaaaggatgattgctggtcattttcgtgattaca

gaaggaactgtgtatatcgtgctatTTTTGGCATTCCCCATTGGTCCCCAGGTTGTGCTCGGATCAGCCAATTTCCGGT
GGTTAATTCACCTTTGTTGTGTCCTCTGTGCCTCAAGAATTGTTCCGCACATGCAATTCAGTTGAGCATCGCACTACG
CAAGTTTTTTTTGTTAACAAGGAGTGGTTGACAGCATTGCAGGTTGCTAAAACCGTGTCAAAAATTTGCTAACAAA
AGAATTCCTCCTCAGAAATAGTATGAAAATAAATGCCACATAAGTAATCTGAGTAAAACAGACAGAATCACTAGAGAA
GACTAACGACAAAATCTTCTTTTCAATTTCTGTGGTGAAAACTCACTAACTCCGCATATAATCTCATGTATGCAT
TGGAGGTGCATACCATTTGATCATGTATGCATTTAGCTGAACATTATCAAGAAGCAAATTTATTGCACGGGCAAAC
ATAAGCTAAACTAATTTCTCAAGCAATTTAATAATCACAAAACAGTCTGCAATACCAAGATATATAAACTTTTGCCA
GCATGGAGCAACAACACCGAAACACTACAATCAAGACAGCTCCAGAAAAGATTATATGTTGCTACTTCAATTGCCA
CAGTTACAATCACCCATAATGACACATCAAATTACATAAGAGTATTAAGAGTATATGACAGTTGTAACAAAACA
GCATTGAGAAGTAGAACAGACAGCCACATGAACCATATTCACTTCTTCAGATCAGAAAGCCTGGTGA
TACCAATTGCTACCGTAGTCTCCATTTTCACTTATATGTACATAGGACACCTGCGCTATCAGAATAGCCACAAAGC
CAGTTCCTCGCTACAAATTAATCTTGACTAAATCCAACAGATACCAACTCCTCTTAATCAACTTACAAAACACCT
ACATTTTGGTGTCTCTACACCACCAATCACACAAAAGGAATTAATCATCCACCAACTCGTATCCTACTGTTAGAA
TGGAGGCCTTACACCAGAGAATGGGCATTGCGAGATTTCTGAGACACTGTCCTTTGAGCTCCTGA

>XA2

MEEVEAGWLEGGIRWLAETILDNLDADKLEWIRQIRLAADTEKLRAEIEKVDGVVAAVKGRAIGNRSLARSLGRRLR
GLLYDADDAVDELDFRLQQQVEGGVTRFEAEETVGDGAEDDDIPMDNTDVPEAVAAGSSKKRASKAWEHFTTVEF
TADGKDSKARCKYCHKDLCCSTKNGTSALRNHLNVCKRKRVTSTDQPVNPSSAGEGASNATGNSVGRKMRMDGTST
HHEAVSTHPWNKAELSNRIQCMTHQLEEAVNEVMRLCRSSSNQSRQGTTPPATNATSSYLPEPIVYGRAEMETIK
QLIMSNRSNGITVLPVIVGNGGIGKTTLAQLVCKDLVIKSQFNVKIWVYVSDKFDVVKITRQILDHVSNSQHEGINS
DTLQQDLEEQMKSKKFLIVLDDVWEIRTDWKKLLAPLRPNQVNSQEEATGNMIILTTRIQSIKSLGTVQSIK
EALKDDDIWSLFBKVFHAFGNDKHDSSPGLQVLGKQIASELKGNPLAAKTGVSLLGTNLTIDHWDSIIKSEEWKSLQQA
YGIMQALKLSYDHLNPLQQCVSYCSLFPKGYSFKAQLIQIWIQAQGFVEESSEKLEQKWKYLAELVNSGFLQOVE
STRFSSEYFVMHDLMDLAQKVSQTEYATIDGSECTELAPSIHLSIVTDSAYRKEKYRNI SRNEVF EKRLMKVKS
SKLRSVLVIGQYDSSHFFKYFKDAFKEAQHLRLLQITATYADSDSFLSSLVNSTHLRYLKIIVTEESGRTLPRSLRKY
HLQVLDIGYRFGIPRISNDINLLSLRHLVAYDEVCSIANIGKMTSLQELGNFIVQNNLSGFVETQLKSMNKLVLQ
SVSQLENVRTQEEACGAKLKDKQHLEKLHLSWKDAWNGYDSESYEDEYGSMDNIETEGEELSVGDANGAQLQHH
NISSELASSEVLEGLEPHHGLKYLRI SGYNGSTSPWLPSSLTCLQTLHLEKCGKWQILPLERLGLLVKLVLIKMRN
ATELSIPSLEELVLIALPSLNTCSCTSIRNLNSSLKVLKIKNCPVLKVFPLFEISQKFEIERTSSWLPKSLTIYN
CPLSCVHSSLPPSA
ISGYGEYGRCTLPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLQAHRC
LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVMGNSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLRAHRC
LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLENSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLQAHRC
LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTSLEELKIQSCESLSSLD
GLQLLGNLRLRLQAHRC
LSGHGEDGRCILPQSLEELFISEYSLETLQPCFLTNTLCLKQLEVSGETTSLKSLELQSCCTALEHLKIQCASLATLE
GLQFLHALRHMKVFR
PGLPPYLGSSSEQGYELCPRLERLDIDDP SILTTSFCKHLTSLQRLELNYCGSEVARLTDEQERALQLLTSLQELRF
KYCYNLIDLPAGLHSLPSLKRLEIRSCRSIARLPEKGLPPSFEELDI IACSNELAQQCRTLASTLKVKINGGYVN

>Xa31(t)

ttcaggaagacgggactgaatacatatTTTCTTAATTCCAAAGGCATTcaggatgtttatgaaagataaactactgttg
TTTTAGGAAACATATTACATCGGCGTTGTcaggtaaagctaacaatagagtgatgatgttcatggtttcaaagtagaa
tataatattcctcgcacatggacaactttactatgctttgtcaaggagtagtaaggcaaacagtaacagggatTTTT
GCCAAACTACAAACTAGGAAAGAGATAGACCTTACCAGAAAATATTATGTATAAAGGTTTCTTGGAAATATTAGAAAA
ATATAATGACACCGTAGCGTTAGCACGGGCATATTACTAGTAGAACAAAAGAAACACTTATGCATTTATACCCAAAA
AAAACCGTAATTAATGACCGCATTTCTGACTGTGTAGCGGCGCGTGAAGCATAACTCAAATTTTAAAGCCAGAGA
TGGCCTAGGGCCAACATCGACATAAACTGGAATAAGTCCACTTGCCTTCCCTCACTTTACATCGAGTGTGTATGAT
GGCCCTAATTCCTAATCAGATGTCGACCCCTCAAGTCTTCAAACCATGCAAAGTAGTCTCCTCGAGGGTATAGAC
CATATTTCCCTGGTTTCCATTGATGTGGCATAATCAGGGCCACTACCTGAACCATCGCCAGTCCAGTCTCTT
GGCCTGTTCCGATGACGAGCTCGGCCACCACCTCAACATCGTCCATGCTCGGGCGATGGCGAGGAGACAGCTGGGA

cggtgaacctcatagctccgtctgttgcaacatcctctcgccatgctcacgccgccacgacacccttgtcctcctgc
tccgtagcagcctcgcaagctcacgccgcgccctcatcgagggtccctgtccccgagccgcttctagcatcgatgag
ctgaggagctccccactgccgaggcgatgacgatttggcagggagggtcgattcgcccactccaattcgcagcgt
cacacctccccctccccctcctattcatccaccaccaacgcggataggagaagggatagaggaaagtgtgcgtggac
acttacatattggttccatatatTTTTTTactcggatgccacgtcagataaaacaaccgtcgggtcaacatTTTTc
tagtggggattagggacgtcatacatgactcgacgcagagctgagggatggcaagtgaacttattcccataaaacttc
caaaagtttggaatgctcaacttgggtcttctctcgatgctgaagtgctaattgggcaaaataaccaactagcccac
agcccatcgaagtggagccgacgagtcaccggtcaacgcggcagtagtaatccgagaaaacgcgggtcgcgcgaag
caatcgatcatcatcagacgattaatccacgacgaccaggagcgtgccccgagccacacgcccacacgcactgggtg
aatcttctcctttccacctagttttcactctgcagtctcctctcctactactctgcttaccatcctcgttgg
ttgcagccctcttgacacgcccattggccacagtcceaaggactgctcctgttccgATGGAGGAGGTGGAAGCCGGTT
GGCTGGAGGGCGGGATCAGGTGGCTGGCGGAGACCATCCTGGATAACCTGGACGCCGACAAGCTGGATGAATGGATT
CGCCAGATTAGGCTCGCCGCTGACACCGAGAAGCTACGGGCTGAGATCGAGAAGGTGGATGGGGTGGTGGCTGCCGT
GAAGGGGAGGGCGATCGGGAACAGGTCGCTGGCCGATCGCTCGGCCGTCTCAGGGGGTGTGTACGACGCCGACG
ATGCGGTGACGAGCTCGACTACTTCAAGCTCCAGCAGCAGGTGAGGGAGGAGGTACTgtctttgcatatatccgt
gccttttaattaagtttgaagctgcggttgcctgcaacaatggtgtattggcgtcagtttccaatccatgcttgtgc
tacagttactACACGGTTTGAGGCTGAAGAGACGGTCGGAGATGGAGCAGAGGACGAGGACGATATTCCGATGGACA
ATACTGATGTACCGGAGGCAGTGGCGGCAGGCAGCAGCAAGAAACGGTCCAAGGCATGGGAACACTTTACTACCGTA
GAGTTCCTGCTGACGGGAAGGATTCTAAAGCACGGTGCAAGTACTGCCACAAGGACCTATGTTGCACATCTAAGAA
CGGGACATCAGCTTTGCGCAACCATCTCAATGTTTGAAGAGGAAACGTGTAACAAGTACTGACCAACCGGTAAATC
CATCAAGgtaatgctaattggagttctgaatttagtgtaaatccgttgaagtgtaaatttggcccgttacatctgctc
aagatctcattctgtctctaattctctaataagccaactcatgggtcatttttttccctaataatagtagccggtgatg
gtgcaccaaagtgaattagatgcaaggaacaaaagtgaacaattgtatatataaaatataattatctaaaacat
gagtagtgatcaaatccaattctttcaaaaatctactatgcaaaaattgagtgaacaaaatctgctgcctttttttt
ttacagaaaagcaaccaattaataaagtcaaatataaaaacgcttggtagtctccaataaaaatagctcattgtttc
ttatacttatgtttataaaatttaaaatttaaaacttaatttggagttgatttgggttttcttttcatcctattt
tattttacaacatttgatttgaatagttagaatgcgatataaaaaatttaccataagttattttttaaattgt
taataaatcgtaaggataatcataagtaataagtgaaacgattcgtctctcacttaagattgcggttatattgct
gacctttctaatecgcctaaccacgatcacatgctcttccagTGCCGGTGAGGGTGCATCAAATGCAACTGGTAATTC
AGTTGGCAGAAAAAGGATGAGAATGGATGGGACTTCAACACACCACGAGGCAGTTAGCACGCACCCTTGAACAAGG
CTGAACTTTCCAACAGGATCCAATGCATGACTCATCAGTTAGAAGAGGCTGTAAATGAGGTTATGAGGCTATGTCGA
TCCTCAAGTTCAAACCAGAGTCGACAGGGTACACCACCGGCCACAAATGCAACAACATCGTCTTATCTTCCGGAGCC
CATAGTGTATGGGAGGGCTGCAGAGATGGAAACCATCAAACAGCTGATCATGAGCAATAGATCTAATGGCATAACCG
TCCTGCCAATTGTAGGCAATGGAGGGATAGGAAAAACCCTTTGGCGCAACTGGTCTGCAAAGATCTGGTAATTTAA
AGTCAGTTTAAATGTTAAGATATGGGTGTATGTATCTGATAAATTTGATGTAGTTAAGATTACAAGGCAGATTTTGG
TCATGTCTCCAACCAGAGCCACGAAGGAATAAGCAACCTTGATACGCTTCCAGCAGGATCTTGAGGAACAAATGAAAT
CTAAGAAGTTCCTCATTGTCTTAGATGATGTGTGGGAAATCCGTACAGATGACTGGAAAAACTACTGGCTCCTTTA
AGACCTAATGATCAGGTGAATTCATCACAGGAAGAGGCAACAGGTAATATGATAATTTTGACAACCTCGTATACAGAG
TATTGCCAAAAGTCTTGAACAGTACAATCAATTAAGTTAGAAGCTCTGAAAGATGACGATATATGGTCACTATTTA
AAGTGCATGCTTTTGGTAATGATAAACATGATAGTAGTCCAGGCTTACAGGTTCTTGGGAAGCAAATGCTAGCGAG
CTAAAAGGCAACCCACTGGCAGCAAAAACCTGTGGGTTCACTATTAGGAACGAATCTTACCATCGATCATTGGGATAG
CATTATAAAGAGTGAAGAATGGAAATCCCTGCAACAAGCTTATGGCATCATGCAAGCGCTGAAGTTGAGCTATGATC
ATCTATCCAACCCCTTACAGCAATGCGTCTCTTATTGTTCTCTTTTCCCAAGGGTTATTCTTTTCCAGCAAAGCACA
CTAATACAAATATGGATTGCTCAAGGATTTGTGGAAGAATCCAGTGAGAAGTTGGAGCAGAAAGGATGGAAATATCT
AGCTGAGTTGGTAAATTCGGGTTTCTTCCAGCAAGTTGAAAGCACAGGTTTTTATCAGAATATTTGTTATGCACG
ATCTTATGCATGATTTAGCGCAAAAAGTTTCAAAACAGAATATGCAACTATAGATGGCTCAGAGTGCACAGAGTTA
GCCCCAAGTATACGCCATTTGTCAATAGTAACTGATTTCTGCATACCAGCAAGGAGAAATATAGAAACATATCTCGTAA
TGAGGTGTTTGAAGAAAGTTGATGAAAGTTAAGTCAAGGAGTAAAGTTGAGGTCACTGGTATTAATTTGGCAATATG
ATTCTCATTFTTTTAAATATTTCAAAGATGCTTTCAAGGAAGCACAACATCTGCGACTGCTGCAGATCACTGCAACT
TATGCTGATTCTGATTCATTTCTCTCCAGTTTGGTAAATTTCTACACATCTCCGGTATCTGAAAATTTGTGACCGAAGA
ATCCGGCAGAACTTTGCCCGATCTCTAAGGAAGTATTACCATCTTCAAGTACTAGATATTGGCTATAGATTTGGAA
TTCCCGTATATCTAATGATATAAATAATCTTCTCAGCCTGCGGCATCTTGTGTCATATGATGAAGTGTGTTCTTCC
ATTGCTAACATTTGGTAAATGACCTCACTTCCAGGAAGTGGCAATTTTATTGTTTCAAGAAATTTAAGTGGTTTTGA
GGTGACACAATTGAAATCCATGAACAAGCTTGTACAACCTTAGTGTGTCTCAGCTTGAAGTGTGAAACTCAGGAGG
AGGCATGTGGGGCAAAACTGAAAGACAAACAACACTTAGAAAAGCTACATTTGTCCTGGAAGGATGCATGGAATGGA
TATGACAGTGACGAAAGCTATGAAGATGAATACGGCAGTGATATGAATATAGAAACAGAAGGGGAGGAACTGTCAGT
TGGTGTGCCAATGGTGCCTCAAGCTTACAACATCACAGTAATATAAGCTCTGAACTTGCTTCAAGTGAAGTGTCTCG

AAGGTCTTGAACCACATCACGGCCTCAAGTATCTACGGATATCTGGGTATAATGGATCTACCTCCCCAACTTGGCTT
CCTTCTTCACTTACCTGTCTGCAAACACTTCATCTAGAAAAATGTGGAAAATGGCAAATACTTCTTTAGAAAGGCT
AGGGTTACTTGTAAAGCTCGTGTGATCAAAATGAGGAATGCAACAGAACTCTCAATCCCTTCACTGGAGGAGCTTG
TGTTAATTGCATTGCCAAGCTTGAACACATGCTCCTGCACTTCCATCAGGAACTTGAACCTCCAGTTTAAAGGTTCTG
AAAATTAAGAATTGCCCTGTACTGAAGGTATTTCCCTTGTGTTGAGATTTCCAGAAAATTTGAAATCGAGCGGACGTC
GTCATGGTTGCCCATCTTAGCAAGCTTACCATCTATAATTGTCTCTTTCTGTGTGCACAGTTCTCTGCCACCTT
CCGCAATCAGTGGTTATGGAGAATATGGAAGGTGTACCCTTCCGCAATCACTTGAGGAACTTTACATCCATGAGTAT
TCTCAAGAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAACTCAA
TTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTC
TGGATGGCTTGAATTCAGTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGA
AGGTGTATCTCTCCGCAATCACTTGAGGAACTTTACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTC
AGGGAACCTCACTCTCCTGAGAAAATTACATGTAATGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCA
CAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGAATTGCTTGGCAATCTC
AGGTTGCTGCGGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGA
ACTTTACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTAC
ATGTACTGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGC
TGTGAGTCTCTTAGTTCTCTGGATGGCTTGAATTCAGTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAG
TGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACTTTACATCCATGAGTATTCTCAAGAAA
CTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAACTCAAATTTAGTGTCT
CTGCAGCTCCATTCTGCACATCACTCGAAGAGTTGAAAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTT
GCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCC
TTCCGCAATCACTTGAGGAACTTTTCATCAGTGGTATTCTCTAGAAAATTCAGGAACTCTGCAGCCCTGCTTCTGACGAATCTC
ACCTGCTTAAACAATTAGAGGTATCAGGCACCACAAGTTTAAAATCTCTAGAATGCAATCATGCACTGCACCTCGA
ACATTTGAAGATTCAAGGTTGTGCGTCGCTTGCTACATTGGAGGGGTTGCAATTCCTCCACGCCCTCAGGCATATGA
AAGTATTCAGATGCCTGGCTTGCCTCCATATTTGGGGAGTTCGTCAGAGCAGGGCTATGAGCTATGCCACGACTG
GAAAGGCTCGACATCGATGACCCCTCTATCCTTACCACGTCGTTCTGCAAGCACCTCACCTCCCTCCAACGCCTAGA
GCTTAACTATTGCGGAAGTGAAGTGGCAAGACTAACGGATGAGCAAGAGAGAGCGCTTCAGCTCCTCACGTCCCTGC
AAGAGCTCCGGTTAAGTATTGCTACAATCTCATAGATCTTCTGCGGGCTCCACAGCCTTCCCTCCCTCAAGAGG
TTGGAGATCCGGAGTTGCAGGAGCATCGCGAGGCTGCCGGAGAAGGGCCTCCACCTTCGTTTCAAGAACTGGATAT
CATCGCTTGCAGTAATGAGCTAGCTCAGCAGTGCAGAACTTAGCAAGCACTCTGAAGGTCAAATTAATGGGGAT
ATGTGAAC TGA ttactcgggtggccttggttaggcgcaccttttccacctgccaactggcggtgggctcggttcaggcgt
tcaagctgctgtaaattccattggcgcaatgacgaccttcagaaccgttacacaatacaaaggacatatgatggcta
gatcaactgtcgcagagctagctttggttcacctgaaaacataaggccaaacgcgtgggttcttttaatacagaagtat
caaaaattggttttggtttttaacgaaaagatggttagaaaattggttttaattgtgccagattctgtaccagaattctg
tattttgctccgtaattctgtgctacgactgatggtgtatttaactgatagcaatcgaccggcaaaaaccaatccgct
ccaatgcattgtcacatacaaaattctataattctatctactgaaaaggatgattgctgttcattttctgtgattacag
aaggaaactgtgtatatcgtgctatttttgcattcacattgtgtcccagggtgtgctcggatcagccaatctcgggtgg
ttaattcactttggtgtgctcctctgtgcaactcaagaattgttcgcacatgcaattcagttgagcatcgcactacgca
agtttttttggtaacaaaggagtggttgacagcattgcagggtgtgtaaaaccgtgtcaaaaatttgcatacaaaaag
aattctcttcagaaatagtatgaaaataaatgccacataagtaactctgagtaaaacagacagaatcactagagaaga
ctaacgacaaaatcttctttttcaatttctgtggtgaaaaactcactaactccgcatataatctcatgtatgcattg
gaggtgcataccatttcgatcatgtatgcatttagctgaacattatcaagaagcaaaatttattgcacgggcaaacat
aagctaaactaattctcaagcaatttaataatcaaaaacagctctgcaataccaagatatataaaacttttgccagc
atggagcaacaacaccgaaacactacaatcaagacagctcccagaaaagattatatgttgctacttcaattgccaca
gttacaatcaccataatgacacatcaaattacataagagtattaagagtatatgacagttgtaacaaaacaactgc
attgagaactagaacagacagccacatgaaccatattcacttcttcagatcagaaagcctggtgaactaccagtcta
ccaattgctaccgtagtctccatttttcaattatgtacataggacacctgctgatcagaatagccacaaagcca
gttcaactcgctacaaatcaattcttgactaaatccaacagataccaactcctcttaatacaacttcaaaaacacctac
acttttgggtgctctacaccaccaatcacacaaaaggaaatcaatcatccaccaactcgtatcctactggttagaatg
gaggccttacaccagagaatgggcattggccagattctgagacactgtcctttgagctccctga

>XA31(t)

MEEVEAGWLEGGIRWLAETILDNLDADKLDIEWIRQIRLAADTEKLRAEIEKVDGVVAAVKGRAIGNRSLARSLGRLR
GLLYDADDAVDELDFRLQQQVEGGGTRFEAEETVGDGAEDDDI PMDNTDVPEAVAAGSSKKRSKAWEHFTTVEF
TADGKDSKARCKYCHKDLCCSTSKNGTSALRNHLNVCKRKRVTSTDQPVNPSSAGEGASNATGNSVGRKMRMDGTST
HHEAVSTHPWNKAELSNRIQCMTHQLEEAVNEVMRLCRSSSNQSRQTPPATNATSSYLPEPIVYGRAAEMETIK

QLIMSNRSNGITVLPVINGGGIGKTTLAQLVCKDLVIKSQFNVKIWVYVSDKFDVVKITRQILDHVSNOQSHEGISNL
DTLQQDLEEQMKSKKFLIVLDDVWEIRTDWKKLLAPLRPNQVNSSQEEATGNMIILTTRIQSIKSLGTVQSIKL
EALKDDDIWSLKFVHAFGNDKHDSSPGLQVLGKQIASSELKGNPLAAKTVGSLLGTNLTIDHWDSIIKSEEWKSLQQA
YGIMQALKLSYDHLNPLQOCVSYCSLFPKGYFSKAQLIQIWIQAQGFVEESSEKLEQKGWKYLAELVNSGFLQQVE
STRFSSEYFVMHDLMDLAQKVSQTEYATIDGSECTELAPSIRHLSIVTDSAYRKEKYRNI SRNEVFEEKRLMKVKS
SKLRSVLVIGQYDSHFYFKDAFKEAQLRLLQITATYADSDSFLSSLVNSTHLRYLKIIVTEESGRTLPRSLRKY
HLQVLDIGYRFGIPRISNDINLLSLRHLVAYDEVCSIANIGKMTSLQELGNFIVQNNLSGFEVTQLKSMNKLVQL
SVSQLENVRTQEEACGAKLKDKQHLEKLHLSWKDAWNGYDSDESIEDEYGSMDNIETEGEELSVGDANGAQLQHHS
NISSELASSEVLEGLEPHHGLKYLRI SGYNGSTSPTWLPSSLTCLQTLHLEKCGKWQILPLERLGLLVKLVLIKMRN
ATELSIPSLEELVLIALPSLNTCSCTSIRNLNSSLKVLKIKNCPVLKVFPLFEISQKFEIERTSSWLPKSLTIYN
CPLSCVHSSLPESA
ISGYGEYGRCTLPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLQAHRC
LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVMGNSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLRAHRC
LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLQAHRC
LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
GLQLLGNLRLRLQAHRC
LSGHGEDGRCILPQSLEELFISEYSLETLPFCFLTNLTCLKQLEVSGETTSLKSLELQSCCTALEHLKIQGCASLATLE
GLQFLHALRHMKVFR
PGLPPYLGSSSEQGYELCPRLERLDIDDPSILTTSFCKHLTSLQRLELNYCGSEVARLTDEQERALQLLTSLQELRF
KYCYNLIDLPAGLHSLPSLKRLEIRSCRSIARLPEKGLPPSFEELDI IACSNELAQQCRTLASTLKVKINGGYVN

>CGS-Xo1₁₁

tcctcgagggtatagaccatatttcccctggtttccattgatgtggcatatatcagggccactacctcaaccatca
ccacgtccagtcctttcggcctgttcgcgatgacgagctcggccaccacctcaacatcgtccatgctgcggcgatgg
cgaggagacagctgggacggtgaacctcatagctccgtctgttgcaacatcctctcgccatgctcacgtcgccacga
cacccttgtcctcctgctccgtagcagcctcgcaagctcacgcgcggcctcgtcgaggctcctctgtccccgagccg
cttctagcatcgatgagctgaggaactccccactgcccagggcgatgacgatttgccagggagggtcgattcgccca
ctccaattcgcagcttcacacctccccctccccctcctattcatccaccaccaacgcggataggagaagggatagagg
aaggtgtgcgtggacacttacatgttgggtcccatatatttttttactcggatgccacgtcagataaaaacaaccgtc
ggtcaacatttttcgtattgaggattaggaacgtcattcatgactcgacgcagagctgagggatggcaagtgaactt
attcccataaaaacttccaaaagtttggaatgctcaacttgggcttctcttcggatgctgaagtgctaattgggcaaaa
taaccaactagcccacagcccacatcgaagtggaaagccgacgagtcacacgtcaacgcggcagtagtaatccgagaaaa
cgcggtgcgcggaagcaatcgatcatcatcagacgattaatccacgacgaccaggagcgtgccccgagccacacg
cccacacgcactggtgaatcttctcctttccacctagttttcactctgcagctcctcctcctagctactactctgcc
ttcccctcctcgtctgttgtagccctcttgcaacagcctattggccacagctccaaggactgctcctgttccgATGAG
GAGGTGAAGCCGGTTTGGCTGGAGGGCGGGATCAGGTGGCTGGCGGAGACCATCTGGATAACCTGGACGGCGGACAA
GCTGGATGAGTGGATTGCCAGATTAGGCTCGCCGCTGACACCGAGAAGCTACGGGCTGAGATCGAGAAGGTGGATG
GGGTGGTCGCTGCCGTGAAGGGGAGGGCGATCGGGAACAGGTGCTGGCCCGATCTCTCGGCCGTCTCAGGGGGTTG
CTGTACGACGCCGACGATGCGGTGACGAGCTCGACTACTTCAGGCTCCAGCAGCAGGTCGAGGGAGGAGgtactgt
ctttgcatatatccgtgccttttaattaagtttgcaagctgcggttgccctgcaacaatggcgtattggcgtcagtttc
caatccatgcttgcctacagTTACTACACGGTTTGGAGCTGAAGAGACGGTCGGAGATGGAGCAGAGGACGAGGAC
GATATTCCGATGGACAATACTGATGTACCGGAGGCAGTGGCGGCAGGCAGCAGCAAGAAACGGTCCAAGGCATGGGA
ACACTTTACTACCGTAGAGTTCACTGCTGACGGGAAGGATTCTAAAGCACGGTGCAAGTACTGCCACAAGGACCTAT
GTTGCACATCTAAGAACGGGACATCAGCTTTGCGCAACCATCTCAATGTTTGCAAGAGGAAACGTGTAACAAGTACT
GACCAACCGGTAAATCCATCAAGgtaatgctaattggagttctgaatttagtgtaaatccggtgaagtgtaaatttgg
cccgttacatctgcttaagatctcattctgtctcctaatcttctaatagccaactcatgggtcatttttttctcctaata
tatagtaccggtgatgggtgcaccaaataatgtaattagatgcaaggaaacaaaagtgaacaattgtatataataata
attatatacaaaacatgagtagtgatcaaatccaattctttcaaaaatctactatgcaaaaattgagtgaacaaatc
tgctgccttttttttacagaaagcaaccaattaataaagtcaaatataaaaaacgctttgtagtctccaataaaa
tagctcattgtttcgtttatacttatgtttataaatttaaatttaaaacttaattttggagttgattttgtggtttt
cttttcacctattttattttacaacatttgattttgaatagtttaagaatgcttatataaaaaattttaccataagt
tatttttaaaattgttaataaaatcgtaaggataatcataagtataaagtgaacgattcgtcttcatctacttaaga
ttcggttatattgctgacctttctaatacgcctaaccacgatcacatgctcttccagTGCCGGTGAGGGTGCATCAA

TGCAACTGGTAATTCAGTTGGCAGAAAAAGGATGAGAATGGATGGGACTTCAACACACCACGAGGCAGTTAGCACGC
ACCCTTGGAAACAAGGCTGAACTTTCCAACAGGATCCAATGCATGACTCATCAGTTAGAAGAGGCTGTAAATGAGGTT
ATGAGGCTATGTCGATCCTCAAGTTCAAACCAGAGTCGACAGGGTACACCACCGGCCACAAATGCAACAACATCGTC
TTATCTTCCGGAGCCCATAGTGTATGGGAGGGCTGCAGAGATGGAACCATCAAACAGCTGATCATGAGCAATAGAT
CTAATGGCATAACCGTCTGCCAATTGTAGGCAATGGAGGGATAGGAAAAACCCTTTGGCGCAACTGGTCTGCAAA
GATCTGGTAATTTAAAGTCAGTTTAAATGTTAAGATATGGGTGTATGTATCTGATAAATTTGATGTAGTTAAGATTAC
AAGGCAGATTTTGGATCATGTCTCCAACCAGAGCCACGAAGGAATAAGCAACCTTGATACGCTTCAGCAGGATCTTG
AGGAACAAATGAAATCTAAGAAGTTCCTCATTGTCTTAGATGATGTGTGGGAAATCCGTACAGATGACTGGAAAAA
CTACTGGCTCCTTTAAGACCTAAAGATCAGGTGAATTCGTACAGGAAGAGGCAACAGGTAATATGATAATTTTGAC
AACTCGTATACAGAGTATTGCCAAAAGTCTTGGAACAGTACAATCAATTAAGTTAGAAGCTCGAAAGATGACGATA
TATGGTCACTATTTAAAGTGCATGCTTTTGGTAATGATAAACATGATAGTAGTCCAGGCTTACAGGTTCTTGGGAAG
CAAATTTGTCAGCGAGCTAAAAGGCCAACCCACTGGCAGCAAAAACTGTGGGTTCACTATTAGGAACGAATCTTACCAT
CGATCATTGGGATAGCATTATAAAGAGTGAAGAATGAAATCCCTGCAACAAGCTTATGGCATCATGCAAGCGCTGA
AGTTGAGCTATGATCATCTATCCAACCCCTTACAGCAATGCGTCTCTTATTGTTCTCTTTTCCCAAGGGTTATTCT
TTCAGCAAAGCACAATAATAAATATGGATTGCTCAAGGATTTGTGGAAGAATCCAGTGAGAAGTTGGAGCAGAA
AGGATGGAATATCTAGCTGAGTTGGTAAATTCGGGTTTCTTCAGCAAGTTGAAAGCACACGGTTTTTCATCAGAAT
ATTTTGTATGCACGATCTTATGCATGATTTAGCGCAAAAGGTTTCAAAAACAGAATATGCAACTATAGATGGCTCA
GAGTGCACAGAGTTAGCCCCAAGTATACGCCATTTGTCAATAGTAACTGATTCTGCATACCGCAAGGAGAAATATAG
AAACATATCTCGTAATGAGGTGTTTGGAGAAAAGGTTGATGAAAGTTAAGTCAAGGAGTAAGTTGAGGTCAGTGGTAT
TAATTTGGGCAATATGATTCTCATTTTTTAAATATTTCAAAGATGCTTTCAAGGAAGCACAACATCTGCGACTGCTG
CAGATCACTGCAACTTATGCTGATTCTGATTCAATTTCTCTCCAGTTTGGTAAATTTACACATCTCCGGTATCTGAA
AATTTGTGACCGAAGAATCCGGCAGAACTTTGCCCGATCTCTAAGGAAGTATTACCATCTTCAAGTACTAGATATTG
GCTATAGATTTGGAATTTCCCGTATATCTAATGATATAAATAATCTTCTCAGCCTGCGGCATCTTGTTCATATGAT
GAAGTGTGTTCTTCCATTGCTAACATTGGTAAAATGACCTCACTTCAGGAAGTGGCAATTTTATTGTTTTCAGAAATA
TTAAGTGGTTTTGAGGTGACACAATTGAAATCCATGAACAAGCTTGTACAACCTTAGTGTGTCTCAGCTTGAAAATG
TTAGAACTCAGGAGGAGGCATGTGGGGCAAACTGAAAGACAAAACAACACTTAGAAAAGCTACATTTGTCTGGAAG
GATGCATGGAATGGATATGACAGTGACGAAAGCTATGAAGATGAATACGGCAGTGATATGAATATAGAAACAGAAGG
GGAGGAATGTCAAGTTGGTGTGCCAATGGTGCACAAAGCTTACAACATCACAGTAATATAAGCTCTGAACTTGCTT
CAAGTGAGGTGCTCGAAGGTCTTGAACCACATCAGGCCTCAAGTATCTACGGATATCTGGGTATAATGGATCTACC
TCCCCAACTTGGCTTCTTCTTCACTTACCTGTCTGCAAACTTCACTCTAGAAAAATGTGGAAAATGGCAAACTACT
TCCTTTAGAAAAGGCTAGGGTTACTTGTAAAGCTCGTGTTGATCAAAATGAGGAATGCAACAGAAGCTCTCAATCCCTT
CACTGGAGGAGCTTGTGTTAATTGCATTGCCAAGCTTGAACACATGCTCCTGCACCTCCATCAGGAAGTGAAGTCC
AGTTTAAAGGTTCTGAAAATTAAGAATTGCCCTGTACTGAAGGATTTCCCTTGTGTTGAGATTTCCAGAAAATTTGA
AATCGAGCGGACGTCGTCATGGTTGCCCATCTTAGCAAGCTTACCATCTATAATTATCCTCTTTTCTGTGTGCACA
GTTCTCTGCCACCTTCCGCAATCAGTGGTTATGGAGAATATGGAAGGTGTACCCTTCCGCAATCACTTGGGAACTT
TACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGT
ACTGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTG
AGTCTCTTAGTCTCTGGATGGCTTGAATTTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGT
CATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGGGAACTTTACATCCATGAGTATTCTCAAGAACTCT
GCAGCCCTGCTTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAACTCAAATTTAGTGTCTCTGC
AGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTCTCTGGATGGCTTGCAA
TTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTTCATGGAGAAGATGGAAGGTGTATCCTTCC
GCAATCACTTGGGAACTTTACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTTCAGGGAACCTCACTC
TCCTGAGAAAATTACATGTACTGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGTG
TTGATAATTCAAAGCTGTGAGTCTCTTAGTCTCTGGATGGCTTGAATTTGCTTGGCAACCTCAGGTTGCTGCAGGC
ACATAGATGCCTCAGTGGTTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGGGAACTTTACATCCATG
AGTATTTCAAGAACTCTGCAGCCCTGCTTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAACT
TCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGTGTGATAATTCAAAGCTGTGAGTCTCTTAG
TTCTCTGGATGGCTTGAATTTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTTCATGGAGAAG
ATGGAAGGTGTATCCTTCCGCAATCACTTGGGAACTTTTTCATCAGTGTAGTATTCTCTAGAAAATCTGCAGCCCTGC
TTCCTGACGAATCTCACCTGCTTAAAACAATTAAGGGTATCAGGCACCACAAGTTTCAAATCTCTAGAACTGCAATC
ATGCACTGCACTCGAACATTTGAAGATTCAAGGTTGTGCGTCTGCTTGTACATTGGAGGGGTTGCAATTTCTCCACG
CCCTCAGGCATATGGAAGTATTAGATGCCTGGCTTGCTCCATATTTGGGGAGTTTCGTGAGAGCAGGGCTATGAG
CTATGCCACGACTGGAAAGGCTCGACATCGATGACCCCTCTATCCTTACCACGTCGTTCTGCAAGCACCTCACCTC
CCTCCAACGCCTAGAGCTTAACTATTGCGGAAGTGAAGTGGCAAGACTAACGGATGAGCAAGAGAGAGCGCTTACGC
TCCTCACGTCCCTGCAAGAGCTCCGGTTAAGTATTGCTACAATCTCATAGATCTTCTGCGGGGCTCCACAGCCTT
CCCTCCCTCGAGAGGTTGGAGATCCGGAGTTGCAGGAGCATCGCGAGGCTGCCGGAGAAGGGCCTCCACCTTCTGTT

CGAAGAACTGGATATCATCGCTTGCAGTAATGAGCTAGCTCAGCAGTGCAGAACTCTAGCAAGCACTCTGAAGGTCA
AAATTAATGGGGGATATGTGAAC TGAattactcgggtggcttggtaggcgcacctttttccacctgcccaactggcgtg
ggctcgttcaggcggtcaagctgctgtaaattccattgccgcaatgacgaccttcagaaccgttacacaatacaaag
gacatatgatggctagatcaactgtcgcagagctagctttgggtcacctgaaaacataaggccaaacgcgtgggttct
ttaatcagaagtatcaaaaattgggttttgggttttaacgaaaagatgtagaaattgggttttaatgtgccagattc
tgtaccagaattctgtattttgctccgtaattctgtgctacgactgatggttatttaactgatagcaatcgaccgg
caaaaccaatccgctccaatgcattgtcacatacaaattctataattctatctactgaaaaggatgattgctgttca
ttttcgtgattacagaaggaactgtgtatatcgtgctatttttgcatcaccattgtgtcccagggtgtgctcggatc
agccaatttcgggtggtaattcactttgttgtgctcctctgtgcaactcaagaattgttcgcacatgcaattcagttga
gcatcgcactacgcaagtttttttggtaacaaaggagtgggtgacagcattgcagggttgctaaaaccgtgtcaaaa
at ttgctaacaaaagaattctcttcagaaatagtatgaaaataaatgccacataagtaattctgagtaaaacagacag
aatcactagagaagactaacgacaaaatctctttttcaatttctgtggtgaaaaactcactaactccgcataataat
ctcatgtatgcattggaggtgcataccatttcgcatcatgtatgcatttagctgaacattatcaagaagcaaatttat
tgcacgggcaaacataagctaactaattctcaagcaatttaataatcacaaaacagctctgcaataccaagatataat
aaaacttttgccagcatggagcaacaacaccgaaacactacaatcaagacagctcccagaaaagattatatgttgct
acttcaattgccacagttacaatcaccataatgacacatcaaattacataagagtattaagagtatatgacagttg
taacaaaacaactgcattgagaactagaacagacagccacatgaaccatattcacttcttcagatcagaaagcctgg
tgaactaccagttaccaattgctaccgtagtctccatttttcaacttatgtacataggacacctgcgctatcaga
atagccacaaagccagttcactcgctacaaattaatcttgactaaatccaaacagataccaactcctcttaatacaac
ttacaaaacacctacacttttgggtgctctacaccaccaatcacacaaaaggaaattaatcatccaccaactcgtat
cctactgttagaatggaggccttacaccagagaatgggcattgccagattctgagacactgtcctttgagctccctg
a

>CGS-XO1₁₁

MEEVEAGLLEGGIRWLAETILDNLDADKLEWIRQIRLAADTEKLRAEIEKVDGVVAAVKGRAIGNRSLARSLGRLR
GLLYDADDAVDELDFRLQQQVEGGVTRFEAEETVGDGAEDDDI PMDNTDVPEAVAAGSSKKRKSWEHFTTVEF
TADGKDSKARCKYCHKDLCCSTKNGTSALRNHLNVCKRKRVTSTDQPVNPS SAGEGASNATGNSVGRKRMRMDGTST
HHEAVSTHPWNKAELSNIQCMTHQLEEAVNEVMRLCRSSSNQSRQGTTPPATNATTSSYLPEPIVYGRAEMETIK
QLIMSNRSNGITVLPVIVGNGGIGKTTLAQLVCKDLVIKIQFNVIWVYVSDKFDVVKITRQILDHVSNSQ SHEGISNL
DTLQQDLEEQMKS KFLIVLDDVWEIRTDDWKLLAPLRPKDQVNSSQEEATGNMI ILTTRIQSI AKSLGTVQSIKL
EALKDDDIWLSLFKVHAFGNDKHDSSPGLQVLGKQIASSELKGNPLAAKTVGSLGTLNLTIDHWDSI IKSEEWKSLQQA
YGIMQALKLSYDHLNPLQOCVSYCSLFPKGYSFKAQLIQIWIQAQGFVEESSEKLEQKGWKYLAELVNSGFLOQVE
STRFSSEYFVMHDLMDLAQKVSQTEYATIDGSECTELAPSI RHLSIVTDSAYRKEKYRNI SRNEVF EKRLMKVKS
SKLRLSLVLIQYD SHFFKYFKDAFKEA QHLRLLQITATYADSDSFLSSLVNSTHLRYLKI VTEESGRTLPRSLRKY
HLQVLDIGYRFGIPRISNDINLLSLRHLVAYDEVCSIANIGKMTSLQELGNFIVQNNLSGFVETQLKSMNKLVLQ
SVSQLENVRTQEEACGAKLKDKQHLEKLHL SWKDAWNGYDSDESYEDEYGS DMNIETE GEELSVGDANGA QSLQHHS
NISSELASSEVLEGLEPHHGLKYLRI SGYNGSTSPTWLPSLTLCLQTLHLEKCGKWQILPLERLGLLVKLVLIKMRN
ATELSIPSLLEELVLIAPSLNTCSCTSIRNLNSSLKVLKIKNCPVLKVFLFEISQKFEIERTSSWLP HLSKLT IYN
YPLSCVHSSLPPSA
ISGYGEYGRCTLPQSLEELYI HEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELI IQSCELSLSSLD
GLQLLGNLRLRLQAHR
LSGHGEDGRCILPQSLEELYI HEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELI IQSCELSLSSLD
GLQLLGNLRLRLQAHR
LSGHGEDGRCILPQSLEELYI HEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELI IQSCELSLSSLD
GLQLLGNLRLRLQAHR
LSGHGEDGRCILPQSLEELFISEYSLETLQPCFLTNTLCLKQLRVSGTTSFKSLELQSCCTALEHLKI QGCASLATLE
GLQFLHALRHMEVFRC
PGLPPYLGSSSEQGYELCPRLERLDIDDPSILTTSFCKHLTSLQRLELNYCGSEVARLTDEQERALQLLTSLQELRF
KYCYNLIDLPAGLHSLPSLERLEIRSCRSIARLPEKGLPPSFEELDI IACSNELAQQCRTLASTLKV KINGGYVN

>Xa14

ttcaggaagacgggactgaatacatatcttcttaattccaaaggcatttaggatgtttatgaaagataattactgttg
tttttaggaaacatattacatcggcggtgtcaggttaagctgacatagagtgatgatgttcatgggttcaaagagaa
tatatatattcctcgcacgacaaactttactatgctttgtcaaggagatgtaaggcaaacagtaacagggattttt
gccaaactacaaactaggaagagatagacctaccgaaaatattatgtataaagggttcttggaatattagaaaa
atataatgacaccgtagcgttagcacgggcatattactagtagaacaagaacacttatgcatttatacccaaaa
aaaaacgtaattaatggacggcatttcgtactgtgtagcggcccgtcgaagcataaactcaaacttttaagccagaga
taggcctgggccaacatcgacataaactggaataagtccacttgccttccctcaactttacatggagtggtgatgat
ggcctaattccaataaccagatgtcgaccctcaagtcttcaaaacctgcaaagtagttcctcgagggtatagac
catatctccctgggttccattgatgtggcatatcagggcccactacctgaacctcgccacgtccagtcctttc
ggcctgttcgcgatgacgagctcggccaccacctcaacatcgtccatgctgcccgcgatggcgaggagacagctggga
cgggtgaacctcatagctccgtctgttgcaacatcctctcgccatgctcagcgcgccacgacaccttgtcctcctgc
tccgtagcagcctcgcaagctcagcgcgggctcatcgaggtccctgtccccgagccgcttctagcatcgatgag
ctgaggagctccccactgcccaggcgatgacgatttggcagggagggtcgattcgcccactccaattcgacagct
cacacctccccctccccctcctattcatccaccacaacgcggataggagaaggatagaggaaggtgtgcgtggac
acttacatattggttccatataatcttttttttactcggatgccacgtcagataaaaacaaccgtcggccaacatctt
gtagtggggattagggacgtcatacatgactcgacgcagagctgagggatggcaagtgaacttattcccataaactt
ccaaaagtttggaaatgctcaacttgggcttctctcctcgatgctgaagtgtcaatgggccaataaccaactagccca
cagcccacgaaagtgaagccgacgagctccaccgtcaacgcggcagtagtaatccgagaaaaacgcccgtcgcgcgaa
gcaatcgatcatcatcagacgattaatccacgacgaccaggagcgtgccccgagccacacgcccacacgcactggt
gaatcttctcctttccacctagtttctactctgcagctcctctcctagctactactcgtccttcccatcctcgctt
gttcgacccctcttgcacacgccattggccacagctccaaggactgctcctgttccgATGGAGGAGGTGGAAGCCGGT
TTGCTGGAGGGCGGGATCAGGTGGCTGGCGGAGACCATCCTGGATAACCTGGATGCCGACAAGCTGGATGAGTGGAT
TCGCCAGATTAGGCTCGCCGCTGACACCGAGAAGCTACGGGCTGAGATCGAGAAGGTGGATGGGGTGGTTCGCTGCCG
TGAAGGGGAGGGCGATCGGGAACAGGTGCTGGCCCCGATCTCTCGGCCGTCTCAGGGGGTGTGTACGACGCCGAC
GATGCGGTGACGAGCTCGACTACTTCAGGCTCCAGCAGCAGGTGAGGGGAGGAGgtactgtctttgcatatccgtg
ccatcttaattaagtgtgcaagctgctgctgcaacaatggcgtcagttccaacctgcttgtgctacagTTAT
TACACGGTTTGGAGCTGAAGACACGGTTGGAGATGGAGCAGAGGACGAAGGCATATCCGATGGACAATACTGATG
TACCGCGGGCGGTGCGGCAGGCCGACGACGAAGAAACGGTCCAAGGCATGGGAACACTTTACTCCCCTAGAGTTC
ACTGCTGATGGGAAGGCTTCTAAAGCACGGTGCAGACTGCCACAAGGACCTATGTTGCACATAAGAACGGGAC
ATCAGCTTTGCGCAACCATTCTCAATGTTTGAAGAGGAAACGTGTAACAAGTACTGACCAACCGTAAATCCATCAA
Ggtaatgctaattggagttctgaatlttagtgtaaatccgttgaagtgtaaatlttggcccgttacatctgcttaagatc
tcattctgtctctaatcttctaatagccaactcatgggtcatttttttctctaatatatagtaccggtgatggtgcac
caaatgtaattagatgcaaggaaacaaaagtgaacaattgtatataataataataataataataataataataata
tgtatcaaatccaattctttcaaaaatctactatgcaaaaattgagtgaacaaaatctgctgccttttttttacagaa
agcaaccaattaataaagtcaaatataaaaacgcttggtagtctccaataaaaatagctcattgtttcgtttataact
tatgtttataaaatlttaaaatlttaaaacttaatlttggagttgattlttgggttttcttttctcctatttttattttac
aacatlttgatlttgaatagtttaagaatgcttatataaaaatlttaccataagttatlttttaaaatgttaataaat
cgtaaggataatcataagtataagtgaacgatcctcctctacttaagattgctgttatattgctgaccttcc
taatcgcctaaccacgatcacatgctcttccagTGCCGGTGGAGGTGCATCAAATGCAACTGGTAATTCAGTAGGCA
GAAAAAGGATGAGAATGGATGGGACTTCAACACACCACGAGGCAGTTAGCACGCACCCTTGAACAAGGCTGAACTT
TCCAACAGGATCCAATGCATGACTCATCAGTTAGAAGAGGCTGTAATGAGGTTATGAGGCTATGTCGATCCTCAAG
TTCAAACCAGAGTGCACAGGGTACATCACCGGCCACAAATGCAACAACATCGTCTTATCTTCCGGAGCCCATAGTGT
ATGGGAGGGCTGCAGAGATGGAAACCATCAAACAGCTGATCATGAGCAATAGATCTAATGGCATAACCGTCTGCCA
ATTGTAGGCAATGGAGGGATAGGAAAAACACTTTGGCGCAACTGGTCTGCAAAGATCTGGTAATTTAAAGTCAGTT
TAATGTTAAGATATGGGTGTATGTATCTGATAAATTTGATGTAGTTAAGATTACAAGGCAGATTTTGGATCATGTCT
CCAACCAGAACCACGAAGGAATAAGCAACCTTGATACGCTTCAGCAGGATCTTGAGGAACAAATGAAATCTAAGAAG
TTCCTCATTGTCTTAGATGATGTGTGGGAAATCCGTACAGATGACTGGAAAAAATACTGGCTCCTTTAAGACCTAA
TGATCAGGTGAATTCGTACAGGAAGAGGCAACAGGTAATATGATAATTTGACAACCTCGTATACAGAGTATTGCCA
AAAGTCTTGAACAGTACAATCAATTAAGTTAGAAGCTCTGAAAGATGACGATATATGGTCACTATTTAAAGTGCAT
GCTTTTGGTAATGATAAATCATGATAGTAGTCCAGGCTTACAGGTTCTTGGGAAGCAAATGCTAGCGAGCTAAAAGG
CAACCCACTGGCAGCAAAAATACTGTGGGTTCACTATTAGGAACGAATCTTACCATCGATCATTGGGATAGCATTATAA
AGAGTGAAGAATGGAAATCCCTGCAACAAGCTTATGGCATCATGCAAGCGCTGAAGTTGAGCTATGATCATCTATCC
AACCCCTTACAGCAATGCGTCTCTTATTGTTCTCTTTTCCCAAGGGTTATTCTTTTCCAGCAAAGCACAATAATACA
AATATGGATTGCTCAAGGATTTGTGGAAGAATCCAGTGAGAAGTTGGAGCAGAAAGGATGGAAATATCTAGCTGAGT
TGGTAAATTCGGGTTTCTTCCAGCAAGTTGAAAGCACACGGTTTTTTCATCAGAATATTTTGTATGCACGATCTTATG
CATGATTTAGCGCAAAAGGTTTTCACAAACAGAATATGCAACTATAGATGGCTCAGAGTGCACAGAGTTAGCCCCAAG
TATACGCCATTTGTCAATAGTAACTGATTCTGCATACCGCAAGGAGAAATATAGAAACATATCTCGTAATGAGGTGT

TTGAGAAAAGGTTGATGAAAGTTAAGTCAAGGAGTAAGTTGAGGTCACCTGGTATTAATTGGGCAATATGATTCTCAT
TTTTTTAAATATTTCAAAGATGCTTTCAAGGAAGCACAACATCTGCGACTGCTGCAGATCACTGCAACTTATGCTGA
TTCTGATTCAATTTCTCTCCAGTTTGGTAAATTCTACACATCTCCGGTATCTGAAAATTGTGACCGAAGAATCCGGCA
GAACTTTGCCCGATCTCTAAGGAAGTATTACCATCTTCAAGTACTAGATATTGGCTATAGATTTGGAATCCCCGT
ATATCTAATGATATAAATAATCTTCTCAGCCTGCGGCATCTTGTTCATATGATGAAGTGTGTTCTTCCATTGCTAA
CATTGGTAAAATGACCTCACTTCAGGAAGTGGCAATTTTATTGTTTCAAGATAATTTAAGTGGTTTTGAGGTGACAC
AATTGAAATCCATGAACAAGCTTGTACAACCTTAGTGTGTCTCAACTTGGAAATGTTAGAAGTCAAGGAGGAGGCATGT
GGGGCAAACTGAAAGACAAACAACACTTAGAAAAGCTACATTTGTCTTGGAAAGGATGCATGGAATGGATATGACAG
TGACGAAAGCTATGAAGATGAATACGGCAGTGTATGAATATAGAAACAGAAGGGGAGGAAGTGTGAGTTGGTGTG
CCAATTGGTGCCCAAAGCTTACAACATCACAGTAATATAAGTCTGAAGTGTCAAGTGTGAGTGTGCTCAAGGTTCT
GAACCATCACGGCTCAAGTATCTACGGATATCTGGGTATAATGGATCTACCTCCCCAAGTGGCTTCTTCTTCTT
ACTTACCTGTCTGCAAAACACTTTCATCTAGAAAAATCTGGAAAATGGCAAACTTCTTCTTAGAAAGGCTAGGGTTAC
TTGTAAAGCTCGTGTGATCAAAATGAGGAATGCAACAGAAGTCTCAATCCCTTCACTGGAGGAGCTTGTGTTAATT
GCATTGCCAAGCTTGAACACATGCTCCTGCACTTCCATCAGGAAGTGAAGTCCAGTTTAAAGGTTCTGAAAATTAA
GAATTGCCCTGTACTGAAGGATTTCCCTTGTGTTGAGATTTGCCAGAAATTTGAAATCGAGCGGACGTCGTCATGGT
TGCCCCATCTTAGCAAGCTTACCATCTATAATTGTCCTCTTCTGTGTGCACAGTTCTCTGCCACCTTCATCTATT
ACATATTGCTATGATTTAACATTCTACGGCAGCAAGGTAGATTTGCTGGATTTACCTCCCTTGAGGAGTTAGTGAT
TTCACGATGCCCCAAGCTGGTGTGTTCTTGGCGCATAACGACGGAAATGATGAACAGTCAATGGAAGATGGCTCC
TACCGCTATCACTTGGAAAAGTGGATTAAGTATGTTGATTCCCTAAAACGCTGCAGCTCTGCTTCCGGGGAAC
CTCACCCGCTGAAAAAAGTGTAGTGTGGGAAACCAAGTTTAAACATCTCTGCAGCTCCATTCTGCACAGCACT
CCAAGAGTTGATAATTCGAAGCTGTGAGTGCCTTAATTCTCTGGAAGGCTTGAATTTGCTCGGCAATCTCAGGTTGC
TGTGTGCACACAGATGCCTCAGCGGCCATGAAGAAGATGGAATGTGTATCCTTCCGCAATCACTTGAGGAAATTGAC
ATCTGGGAGTACTCTCAAGAAACGCTGCAGCCCTGCTTCCAGGGAACTCACCAGCTTGAAAAAAGTGTAGTACA
GGGAAGCCAAAAGTTAATATCTCTGCAGCTGTATTCCTGCACAGCACTCCAAGAGTTGATGATTGAAAGTTGTGTGT
CGCTTAATTCTCTGGAAGGCCTGCAATGGCTCGTCAACCTCAGGTTGCTGCGGGCACACAGATGCCTCAGTGGTTAT
GGAGAAAATGGAAGGTGTATCCTTCCACAATCACTTGGGAAGTTTACATCAGAGAGTATTCTCAAGAAACGCTGCA
GCCCTGCTTCCAGGGAACCTCACCAGCTTGAAAAAAGTGTAGTACAGGGAAGCCAAAAGTTAATATCTCTGCAGC
TGTATTCTGCACAGCACTCAAGAGTTGATGATTGGAAGTTGTGTGTGCTTAATTCTCTGGAAGGCCTGCAATGG
CTCGTCAACCTCAGGTTGCTGCGGGCACACAGATGCCTCAGTGGTTATGGAGAAAATGGAAGGTGTATCTTCCACA
ATCACTTGAAGGTTCTTACATCAGAGAGTATTCTCAAGAAATTTACAGCCCTGCTTCCAGCAAGTATCTCACTTGT
TAAAAAGATTAGAGGTATCAGGCAGTGAAGTTTCAAATCTCTGAAGTTGCAATCATGCAGTGCAGTGCAGCAATTTG
AGGATTGAATGGTGTGCGTGCCTTGTACATTAGAGGGCTTGAATTTCCCTCCATGCCCTTAAGCATTGGAAGTATT
CAGATGCCCTGGCTTGCTCCATATTTGGGGAGTTTATCAGGGCAGGGCTATGAGCTATGCCCACTACTGGAAAGGC
TCGAGATCGATGACCCCTCTATCCTTACCACGTCGTTTTGCAAGAACCTCACCTCTCTCCAATACCTAGAGCTTTGC
AGTCACGGATTGGAATGGAAGACTAACGGACGAGGAAGAGAGAGCGCTTCAACTCCTCACTTCCCTGCAAGAGCT
CCGATTTAACTGTTGCTACAATCTCGTAGATCTTCCACAGGGCTCCACAACCTTCTCTCCCTCAAGAGGTTGGAGA
TCTGGAATGCGGGAGCATCGCGAGGCCGCTGGAAGGGTCTCCACCTTCTGTTGGAAGAAGTGGATATCTCGGT
TGTAGTAAAGAGCTAGTTTCAAGCAGTGCAGAAGTCTAGCAAGCAAGCTGAAGGTCGAAATTAATGGGGAATATGTGAA
CTGAattactcgggtggcttggtagggcgacaccttttccacctgtccaactggcggtgggctcggttcaggcggtcaagct
gctgtaaattccattgcccgaatgacgaccttcagaaccggtacacaatacaaaggacatatgatggctagatcaac
tgtcgcagagctagctttgggtccacctgaaaacataaggccaaacgcggtgggttcttttaatacagaagtatcaaaaat
tggttttggtttttaacgaaaagatggttagaaattggtttttaatgtgccagattctgtaccagaattctgtattttg
ctccgtaattctgtgctacgactgatggtgtatattaactgatagcaatcgaccggcaaaaaccaatccgctccaatgc
attgtcacatacaaaattctataattctatctactgaaggaactgtgtatatcgtgctatttttgcatcacttgctg
ttccagggtgtgctcggatcagccaatttcgggtggtaattcactttgctgtgctcctctgtgcaagtgtgcaactcaag
aattggtcgcacatgcaatcagttgagcatcgcactacgcaagttttttttttttgttaacaaaggagtggttga
cagcattgcaggttgctaaaacagtggtcaaaaatttgctaacaaaagaattctcttcagaaatagtatgaaaataaa
tgccacataaagtaactctgagtataaacagacagaatcactagagaagactaacgacaaaactctctttttctatttct
gtggtgaaaaaactcagcatataatctcatgtatgcatgtgaggtgcataaccatttcgatcatgtatgcattttagctg
aacattatcaagaagcaaatatttgacagggcaaccataagctaaactaattctcaagcaatthaataatcaaaa
acagctctgcaataccaagatatataaaacttttgccagcatggagcaacaacactgaaacactacaatcaagacagc
tcccagaaaagattatatgttgetacttcaattgcccacagttacaatcaccataatgacacatcaaatcacataag
agtattaagagtatatgacagttgtaacaaaacaactgcattgagaactagaagaacagacagccacatgaaccata
ttcacttcttcagatcagaaagcctgggtgaactaccagttaccaattgctaccgtagtctccatttttctacttata
tgtacataggacacctgcgctatcagaatagcaacaaagccagttcactcgctacaaattaatcttgactaaatcca
aatagataccaactcctcttaatacaacttcaaaaacacctacacttttggttgtctctacaccaccaatcacacaaaa

ggaaattaatcatccacactactcgatcctactgtagaatggaggccttacaccagagaatgggcattgccagatt
ctgagacactgtcctttgagctccctga

>XA14

MEEVEAGLLEGGIRWLAETILDNLDADKLDDEWIRQIRLAADTEKLRAEIEKVDGVVAAVKGRAIGNRSLARSLGRLR
GLLYDADDAVDELDFRLQQQVEGGVITRFEAEDTVGDGAEDEGDI PMDNTDVPAAAAAGRSSKRSKAWEHFTPVE
FTADGKASKARCKYCHKDLCCSTSKNGTSALRNHLNVCKRKRVTSTDQPVNPSSAGEGASNATGNSVGRKMRMDGTS
THHEAVSTHPWNKAELSNRIQCMTHQLEEA VNEVMRLCRSSSSNQSRQGTSPATNATTSSYLPEPIVYGRAAEMETI
KQLIMSNRNSGITVLP IVGNNGGIGKTTLAQLVCKDLVIKSQFNVKIWVYVSDKFDVVKITRQILDHVSQNHEG I SN
LDTLQODLEEOMKSKKFLIVLDDVWEIRTDWKKLLAPLRPNQVNSSQEEATGNMI ILTTRIQSI AKSLGTVQS I K
LEALKDDDIWLSL FKVHAFGNDKHDSSPGLQVLGKQIASELKG NPLAAKT VGSLLGTNLTIDHWDSI IKSEEWKSLQQ
AYGIMQALKLSYDHL SNPLQQCVSYCSLFPKGY SFSKAQLIQIWI AQGFVEESSEKLEQKGWKYLAELVNSGFLQOV
ESTRFSSEYFVMHDL MHDLAQKVSQTEYATIDGSECTELAPSIRHLSIVTDSAYRKEKYRNI SRNEVF EKRLMKVKS
RSKLRSLVLI GQYDSHFFKYFKDAFKEA QHLRLLQITATYADSDSFLSSLVNSTHLRYLKI VTEESGRTLPRSLRKY
YHLQVLDIGYRFGI PRISNDINNLLSLRHLVAYDEVCS SIANIGKMTSLQELGNFIVQNNLSGFVETQLKSMNKL VQ
LSVSQLGNVRTQEEACGAKLKDKQHLEKHLHLSWKDAWNGYDSDESYEDEYGS DMNIETEGEELS VGDANGA QSLQHH
SNISSELASSEVLEGLEPHHGLKYLRI SGYNGSTSPTWLPSSLTCLQTLHLEKCGKWQILPLERLGLLVKLVLIKMR
NATELSIPSLEELVLI ALPSLNTCSCTSIRNLNSSLKVLKIKNCPVLKVFPLFEICQKFEIERTSSWLP HLSKLTIIY
NCPLSCVHSSLPPSS
ITYCYDLTFYGSKVDFAGFTSLEELVI SRCPKLV SFLAHNDGNDEQSNGRWLLPLSLGKLE INYVDSLKTLQLCFPG
NLTRLKKLVLGNQSLTSLQLHSCALQELI IRSCELSNSLEGLQLLGNLRLLLCAHRC
LSGHEEDGMCILPQSLEEIDIWEYSQETLQPCFPGNLTSLKKL VVQGSQKLI SLQLYSCTALQELMIESCVSLNSLE
GLQWLVNLRLLRAHRC
LSGYGENGRCILPQSLEELYIREYSQETLQPCFPGNLTSLKKL VVQGSQKLI SLQLYSCTALQELMIGSCVSLNSLE
GLQWLVNLRLLRAHRC
LSGYGENGRCILPQSLEGLYIREYSQEILQPCFQTNLTCLKRLEVSGTGSFKSLKLQ SCTALEHLRIEW CASLATLE
GLQFLHALKHLEVFRC
PGLPPYLGSLSGQGYELCPLLERLEIDDPSILTTSFCKNLTSLOYLELCSHGLEMERLTDEEERALQLLTSLOELRF
NCCYNLVDLPTGLHNL LSLKRLEIWNCGS IARPLEKGLPPSLEELDILGCSKELVQQCRTLASKL KVEINGEYVN

>Xa45(t)

ttcaggaagacgggactgaatacatatcttcttaattccaaaggcattcaggatgtttatgaaagataattactgttg
tttttaggaacatattacatcggcggtgtcaggtaagctgacatagagtgatgatgttcattggtttcaaagtgagaa
tatataattcctcgcgatggacaactttactatgctttgtcaaggagtagtaaggcaaacagtaacagggatTTTT
gccaaactacaaactaggaaagagatagaccttaccgaaaatattatgtataaaggTTTTcttggaatattagaaaa
atataatgacaccgtagcgttagcacgggcatattactagtagaaciaaaagaaacacttatgcattttatacccaaaa
aaaaacgtaattaatggaccgcatttctgactgtgtagcggcccgtcgaagcataactcaaacttttaagccagaga
tgggcctgggccaacatcgacataaaactggaataagtccactttccttccctcaactttacatggagtggtgatgat
ggccctaattccaataaccagatgtcgaccctcaagtcttcaaaaccatgcaaagtagttcctcgagggtatagac
catatctccctgggtttccattgatgtggcatatatcagggcccactacctgaaccatcgccacgtccagtcctttc
ggcctgttcgcatgacgagctcggccaccacctcaacatcgtccatgctcggcgatggcgaggagacagctggga
cgggtgaacctcatagctccgtctgttgcaacatcctctcgccatgctcacgcccacgacacccttgtcctcctgc
tccgtagcagcctcgcaagctcacgcccgggtctcatcgagggtcccctgtccccgagccgcttctagcatcgatgag
ctgaggagctccccactgccgaggcgatgacgatttggcagggagggtcgcattcggccactccaattcgcagcgt
cacacctccccctccccctcctattcatccaccaccaacgcggataggagaaggatagaggaaagtgtgcgtggac
acttacatattggttccatataTTTTTTTTactcggatgccacgtcagataaaaacaaccgtcgggtcaacatttttc
gtagtggggattagggacgtcatacatgactcgacgcagagctgagggatggcaagtgaacttattcccataaaactt
ccaaaagtttggaaatgctcaacttgggcttctctcggatgctgaagtgctaattgggccaataaccaactagccca
cagcccactcgaagtggaagccgacgagttccaccgtcaacgcggcagtagtaatccgagaaaacgcgggtcgcgcgaa
gcaatcgatcatcatcagacgattaatccacgcagcaccaggagcgtgccccgagccacgcgccacgcactggt
gaaatctctcctttccacctagttttctcactctcagctcctcctcctagctactactctgcttccatcctcgtctt
gttgcagccctcttgcacacgccattggccacagtccaaggactgctcctgttccgATG GAGGAGGTGGAAGCCGGT
TTGCTGGAGGGCGGGATCAGGTGGCTGGCGGAGACCATCCTGGATAACCTGGACGCCGACAAGCTGGATGAGTGGAT
TCGCCAGATTAGGCTCGCCGCTGACACCGAGAAGCTACGGGCTGAGATCGAGAAGGTGGATGGGGTGGTTCGCTGCCG

TGAAGGGGAGGGCGATCGGGAACAGGTCGCTGGCCCGATCTCTCGGCCGTCTCAGGGGGTTGCTGTACGACGCCGAC
GATGCGGTCGACGAGCTCGACTACTTCAGGCTCCAGCAGCAGGTCGAGGGAGGAGgtactgtctttgcatatatccg
tgccttttaattaagtttgcaagctgcgcttgcctgcaacaatggcgtattggcgtcagttccaatccatgcttgtg
ctacagTTACTACACGGTTTTGAGGCTGAAGAGACGGTCGGAGATGGAGCAGAGGACGAGGACGATATTCGGATGGAC
AATACTGATGTACCGGAGGCAGTGGCGGCAGGCAGCAGCAAGAAACGGTCCAAGGCATGGGAACACTTTACTACCGT
AGAGTTCACTGCTGACGGGAAGGATTCTAAAGCACGGTGCAAGTACTGCCACAAGGACCTATGTTGCACATCTAAGA
ACGGGACATCAGCTTTGCGCAACCATCTCAATGTTTGCAAGAGGAAACGTGTAACAAGTACTGACCAACCGGTAAAT
CCATCAAGgtaatgctaattggagttctgaattagtgtaaatccggtggaagtgtaaatttggcccgttacatctgct
taagatctcattctgtctctctaatcttctaatagccaactcatggtcatttttttctctaatatataagtaccggtgat
ggtgcaccaaatgtaattagatgcaaggaacaaaagtgaacaattgtatataatcaaatataattatatactaaaaca
tgagtagtgatcaaatccaattctttcaaaaatctactatgcaaaattgagtgacaaaatctgctgcctttttttt
ttacagaagcaaccaattaatataagtcacaataaaaaacgctttgtagctctcaataaaaatagctcattgtttcg
ttataacttatgtttataaatttaaaatttaaaacttaatttggagttgattttgtggttttcttttcatcctattt
tattttacaacatttgaatttgaatagtttaagaatgcgtatataaaaaattttaccataagttatttttaaaattgt
taataaatcgtaaggataatcataagtataagtgaacgattcgctcttcatctacttaagattgcggttatattgct
gacctttctaactgcctaaccacgatcacatgctcttccagTGCCGGTGAGGGTGCATCAAATGCAACTGGTAATTC
AGTTGGCAGAAAAAGGATGAGAATGGATGGGACTTCAACACACCACGAGGCAGTTAGCACGCACCCTTGGAACAAGG
CTGAACCTTTCCAACAGGATCCAATGCATGACTCATCAGTTAGAAGAGGCTGTAAATGAGGTTATGAGGCTATGTCGA
TCCTCAAGTTCAAACCAGAGTCGACAGGGTACACCACCGGCCACAAATGCAACAACATCGTCTTATCTCCGGAGCC
CATAGTGTATGGGAGGGCTGCAGAGATGGAAACCATCAAACAGCTGATCATGAGCAATAGATCTAATGGCATAACCG
TCCTGCCAATTGTAGGCAATGGAGGGATAGGAAAAACCCTTTGGCGCAACTGGTCTGCAAAGATCTGGTAATTA
AGTCAGTTTAAATGTTAAGATATGGGTGTATGTATCTGATAAATTTGATGTAGTTAAGATTACAAGGCAGATTTTGG
TCATGTCTCCAACCAGAGCCACGAAGGAATAAGCAACCTTGATACGCTTCAGCAGGATCTTGAGGAACAAATGAAAT
CTAAGAAGTTCCTCATTGTCTTAGATGATGTGTGGGAAATCCGTACAGATGACTGGAAAAACTACTGGCTCCTTTA
AGACCTAATGATCAGGTGAATTCGTACAGGAAGAGGCAACAGGTAATATGATAATTTTGACAACCTCGTATACAGAG
TATTGCCAAAAGTCTTGGAACAGTACAATCAATTAAGTTAGAAGCTCTGAAAGATGACGATATATGGTCACTATTTA
AAGTGCATGCTTTTGGTAATGATAAACATGATAGTAGTCCAGGCTTACAGGTTCTTGGGAAGCAAATTCGTAGCGAG
CTAAAAGGCAACCCTGAGCAGCAAAAACCTGTGGGTTCACTATTAGGAACGAATCTTACCATCGATCATTGGGATAG
CATTATAAAGAGTGAAGAATGGAAATCCCTGCAACAAGCTTATGGCATCATGCAAGCGCTGAAGTTGAGTCTGATC
ATCTATCCAACCCTTACAGCAATGCGTCTCTTATTGTTCTCTTTTCCCAAGGGTTATTCTTTTCCAGCAAAAGCACA
CTAATACAAATATGGATTGCTCAAGGATTTGTGGAAGAATCCAGTGAGAAGTTGGAGCAGAAAGGATGGAAATATCT
AGCTGAGTTGGTAAATTCGGGTTTTCTTCCAGCAAGTTGAAAGCACACGGTTTTTCATCAGAATATTTTGTATGCACG
ATCTTATGCATGATTTAGCGCAAAAGGTTTTCAAAACAGAATATGCAACTATAGATGGCTCAGAGTGCACAGAGTTA
GCCCAAGTATACGCCATTTGTCAATAGTAACTGATTCTGCATACCGCAAGGAGAAATATAGAAACATATCTCGTAA
TGAGGTGTTTGGAAAAGGTTGATGAAAGTTAAGTCAAGGAGTAAGTTGAGGTCACTGGTATTAATTTGGGCAATATG
ATTCTCATTTTTTTAAATATTTCAAAGATGCTTTCAAGGAAGCACAACATCTGCGACTGCTGCAGATCACTGCAACT
TATGCTGATTCTGATTCAATTTCTCTCCAGTTTGGTAAATCTACACATCTCCGGTATCTGAAAATTGTGACCGAAGA
ATCCGGCAGAACTTTGCCCGATCTCTAAGGAAGTATTACCATCTTCAAGTACTAGATATTGGCTATAGATTTGGAA
TTCCCGTATATCTAATGATATAAATAATCTTCTCAGCCTGCGGCATCTTGTTCATATGATGAAGTGTGTTCTTCC
ATTGCTAACATTGGTAAAATGACCTCACTTCCAGGAAGTGGCAATTTTATTGTTTCCAGAAATTTAAGTGGTTTTGA
GGTGACACAATTGAAATCCATGAACAAGCTTGTACAACCTTAGTGTGTCTCAGCTTGAAAATGTTAGAACTCAGGAGG
AGGCATGTGGGGCAAACTGAAAGACAAACAACACTTAGAAAAGCTACATTTGTCTGGAAGGATGCATGGAATGGA
TATGACAGTGACGAAAGCTATGAAGATGAATACGGCAGTGATATGAATATAGAAACAGAAGGGGAGGAACTGTCAGT
TGGTGTATGCCAATGGTGGCCAAAGCTTACAACATCACAGTAATATAAGCTCTGAACTTGCTTCAAGTGAAGTGTCTG
AAGGTCTTGAACCACATCACGGCCTCAAGTATCTACGGATATCTGGGTATAATGGATCTACCTCCCCAACTTGGCTT
CCTTCTTCACTTACCTGTCTGCAAACTTCACTTAGAAAAATGTGGAAAATGGCAAATACTTCTTTAGAAAAGGCT
AGGGTTACTTGTAAAGCTCGTGTGATCAAAATGAGGAATGCAACAGAACTCTCAATCCCTTCACTGGAGGAGCTTG
TGTTAATTGCATTGCCAAGCTTGAACACATGCTCCTGCACTTCCATCAGGAACCTTGAACCTCCAGTTTAAAGGTTCTG
AAAATTAAGAATTGCCCTGACTGAAGGTATTTCCCTTGTGTTGAGATTTCCAGAAAATTTGAAATCGAGCGGACGTC
GTCATGGTTGCCCATCTTAGCAAGCTTACCATCTATAATTATCCTCTTTCTGTGTGCACAGTTCTCTGCCACCTT
CCGCAATCAGTGGTTATGGAGAATATGGAAGGTGTACCCTTCCGCAATCACTTGAGGAACTTTACATCCATGAGTAT
TCTCAAGAACTCTGCAGCCCTGCTTTTCCAGGAACTCACTCTCTGAGAAAATTACATGTACTGGGAACTCAAA
TTTGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTC
TGGATGGCTTGAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGA
AGGTGTATCCTTCCGCAATCACTTGGAGAACTTTACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTC
AGGGAACCCCACTCTCCTGAGAAAATTACATGTACTGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCA
CAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGAATTGCTTGGCAATCTC

AGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGA
ACTTTACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTAC
ATGTACTGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGC
TGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAG
TGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACCTTTACATCCATGAGTATTCTCAAGAAA
CTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGGAACTCAAATTTAGTGTCT
CTGCAGCTCTATTCTGCACAGCACTCGAAGAGTTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTT
GCAATTGCTTGGCAATCTCAGGTTGCTGCAGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCC
TTCCGCAATCACTTGAGGAACCTTTACATCCATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTCAGGGAACCTC
ACTCTCTGAGAAAATTACATGTACTGGGAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGA
AGTGTGATAATTCAAAGCTGTGAGTCTCTTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAACCTCAGGTTGCTGC
AGGCACATAGATGCCTCAGTGGTCATGGAGAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACCTTTACATC
CATGAGTATTCTCAAGAACTCTGCAGCCCTGCTTTTCAGGGAACCTCACTCTCCTGAGAAAATTACATGTACTGGG
AAACTCAAATTTAGTGTCTCTGCAGCTCCATTCTGCACAGCACTCGAAGTGTGATAATTCAAAGCTGTGAGTCTC
TTAGTTCTCTGGATGGCTTGCAATTGCTTGGCAATCTCAGGTTGCTGCCGGCACATAGATGCCTCAGTGGTCATGGA
GAAGATGGAAGGTGTATCCTTCCGCAATCACTTGAGGAACCTTTTCATCAGTGTGAGTATTCTCTAGAAACTCTGCAGCC
CTGCTTCTGACGAATCTCACCTGCTTAAAACAATTAAGGGTATCAGGCACCACAAGTTTCAAATCTCTAGAAGTGC
AATCATGCACTGCACTCGAACATTTGAAGATTCAAGGTTGTGCGTCTGCTACATTGGAGGGGTTGCAATTCCTC
CACGCCCTCAGGCATATGGAAGTATTAGATGCCTGGCTTGCTCCATATTTGGGGAGTTTCGTGAGAGCAGGGCTA
TGAGCTATGCCCACGACTGGAAAGGCTCGACATCGATGACCCATCTATCCTTACCACGTCTGTTCTGTAAGCACCTCA
CCTCCCTCCAACGCCTAGAGCTTAACTATTGCGGAAGTGAAGTGGCAAGACTAACGGATGAGCAAGAGAGAGCGCTT
CAGCTCCTCACGTCCCTGCAAGAGCTCCGGTTAAGTATTGCTACAATCTCATAGATCTTCTGCGGGGCTCCACAG
CCTTCCCTCCCTCGAGAGGTTGGAGATCCGGAGTTGCAGGAGCATCGCGAGGCTGCCGGAGAAGGGCCTCCCACCTT
CGTTCGAAGAACTGGATATCATCGCTTGAGTAATGAGCTAGCTCAGCAGTGCAGAACTCTAGCAAGCACTCTGAAG
GTCAAATTAATGGGGGATATGTGAAC TGAattactcggctggcttggtaggcgcacctttttccacctgcccaactgg
cgtgggctcgttcaggcgttcaagctgctgtaaattccattgccgcaatgacgaccttcagaaccgttacacaatac
aaaggacatatgatggctagatcaactgtcgcagagctagcttgggttcacctgaaaacataaggccaaacgcgctgg
ttcttttaatacagaagtatcaaaaattgggttttgggttttaacgaaaagatgtagaaaattgggttttaatgtgccag
attctgtaccagaattctgtattttggctccgtaattctgtgctacgactgatgttatttaactgatagcaatcga
ccggcaaaaccaatccgctccaatgcattgtcacatacaaatctataattctatctactgaaaaggatgattgctg
ttcattttcgtgattacagaaggaactgtgtatatacgtgctattttggcattccccattgggtccccagggttggctc
ggatcagccaatttcgggtgggttaattcactttgtgtgctcctctgtgcaactcaagaattggttcgcacatgcaattca
ggtgagcatcgcactacgcaagtttttttggtaacaaaggagtggttgacagcattgcagggttgctaaaaccgtgt
caaaaatttgctaacaaaagaattctcttcagaaatagtatgaaaataaatgccacataagtaattctgagtaaaaaca
gacagaatcactagagaagactaacgacaaaatcttctttttcaatttctgtggtgaaaaactcactaactccgcat
ataatctcatgtatgcattggagggtgcataaccatttcgatcatgtatgcatttagctgaacattatcaagaagcaaa
tttattgcacgggcaaacataagctaaactaattctcaagcaatttaataatcaaaaacagctctgcaataccaaga
tatataaaaacttttggcagcatggagcaacaacaccgaaacactacaatcaagacagctcccagaaaagattatag
ttgctacttcaattgccacagttacaatcaccataatgacacatcaaattacataagagattaaagagtatatgac
agttgtaacaaaacaactgcattgagaactagaacagacagccacatgaacatattcacttcttcagatcagaaaag
cctggtgaactaccagctaccaattgctaccgtagtctccatttttacttatatgtacataggacacctgcgcta
tcagaatagccacaaagccagttcactcgtacaaaattaatcttgactaaatccaaacagataccaactcctcttaa
tcaacttacaaaacacctacacttttgggtgtctctacaccaccaatcacacaaaaggaaattaatcatccaccaact
cgtatcctactgttagaatggaggccttacaccagagaatgggcattgccagattctgagacactgtcctttgagct
ccctga

>XA45(t)

MEEVEAGLLEGGIRWLAETILDNLNDADKLDDEWIRQIRLAADTEKLRAEIEKVDGVVAAVKGRAIGNRSLARSLRRLR
GLLYDADDAVDELDFRLQQQVEGGVTRFEAEETVGDGAEDDDIPMDNTDVPEAVAAGSSKKRSKAWEHFTTVEF
TADGKDSKARCKYCHKDLCCSTKNGTSALRNHLNVCKRKRVTSTDQPVNPSSAGEGASNATGNSVGRKMRMDGTST
HHEAVSTHPWNKAELSNRIQCMTHQLEEAVNEVMRLCRSSSNQSRQGTTPPATNATTSSYLPEPIVYGRAEMETIK
QLIMSNRSNGITVLPVIVGNGGIGKTTLAQLVCKDLVIKSQFNVKIWVYVSDKFDVVKITRQILDHVSNSQSHEGISNL
DTLQQDLEEQMKSCKFLIVLDDVWEIRTDWKKLLAPLRPNQVNSQEEATGNMIILTTRIQSIAKSLGTVQS IKL
EALKDDDIWSLKFVHAFGNDKHDSSPGLQVLGKQIASLKNPLAAKTVGSLLGTNLTIDHWDSI IKSEEWKSLQQA
YGIMQALKLSYDHLNPLQCVSYCSLFPKGYSFSAQLIQIWIQAQGFVEESSEKLEQKGWKYLAELVNSGFLQQVE
STRFSSEYFVMHDLMDHLAQLKVSQTEYATIDGSECTELAPSI RHLSIVTDSAYRKEKYRNI SRNEVFEKRLMKVKS

SKLRSLVLIGQYDSHFFKYFKDAFKEAQLRLLQITATYADSDSFLSSLVNSTHLRYLKIIVTEESGRTLPRSLRKY
 HLQVLDIGYRFGIPRISNDINLLSLRHLVAYDEVCSIANIGKMTSLQELGNFIVQNNLSGFVETQLKSMNKLQV
 SVSQLENVRTQEEACGAKLKDKQHLEKLHLSWKDAWNGYDSDES YEDEYGS DMNIETE GEEL SVGDANGA QSLQH
 NISSELASSEVLEGLEPHGLKYLRI SGYNGSTSP TWLPS SLTCLQTLHLEKCGKWQILPLERLGLLVKLVLIKMRN
 ATELSIPSLEELVLIAPSLNNTCSCTSIRNLNSSLKVLKIKNCPVLKVFPLFEISQKFEIERTSSWLP HLSKLT IYN
 YPLSCVHSSSLPPSA
 ISGYGEYGRCTLPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
 GLQLLGNLRLRLQAHR
 LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNPTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
 GLQLLGNLRLRLQAHR
 LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
 GLQLLGNLRLRLQAHR
 LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLYSCTALEELIIQSCESLSSLD
 GLQLLGNLRLRLQAHR
 LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
 GLQLLGNLRLRLQAHR
 LSGHGEDGRCILPQSLEELYIHEYSQETLQPCFSGNLTLLRKLHVLGNSNLVSLQLHSCTALEELIIQSCESLSSLD
 GLQLLGNLRLRLPAHR
 LSGHGEDGRCILPQSLEELFISEYSLETLPFCFLTNLTCLKQLRVSGTTSFKSLELQSCCTALEHLKIQGCASLATLE
 GLQFLHALRHMEVFR
 PGLPPYLGSSSEQGYELCPRLERLDIDDPSILTTSFCKHLTSLQRLELNYCGSEVARLTDEQERALQLLTSLQELRF
 KYCYNLIDLPAGLHSLPSLERLEIRSCRSIARLPEKGLPPSFEELDIACSNELAQQCRTLASTLKVKINGGYVN

Supplemental Table 1. List of PCR primers used and their applications

Primer	Sequence (5' to 3')	Usage
53120-F4b	AGCTATGACATGATTACGAATT CAGGAAGACGGGACTGA	Amplification of genomic regions (promoter, coding, terminator sequences) of <i>Xal</i> allelic members
53120-R4b	AACGACGGCCAGTGCCAAGCT TCAGGGAGCTCAAAGGACAG	
53120-F6b	AGCTATGACATGATTACGAATT CCTCGAGGGTATAGACCAT	Along with primer 53120-R4b to amplify genomic regions of <i>CGS-Xo1₁₁</i>
BB2-F1	TGCCACAGTTACAATCACCCA	Along with primer M13F to genotype transgenic plants of <i>Xal</i> allelic genes
XaL-F1	ATCAGGAACTTGA ACTCCAG	To screen <i>Xal</i> alleles in a collection of rice accessions
XaL-R1	AACCACTGATTGCGGAAGG	
XaL-F2	CCTTCGCAATCAGTGGTT	To amplify the LRR region of <i>Xal</i> alleles
XaL-R2	AATGTAGCAAGCGACGCACA	
Xa14L-R1	AGGGAATCAACATAGTTAATC	Along with primer XaL-F1 to screen <i>Xa14</i> in a collection of rice accessions
Cambia-F	TGAGCGGATAACAATTCACACAG	Sequence the whole genomic regions of <i>Xal</i> allelic <i>R</i> genes (<i>Xa2</i> , <i>Xa14</i> , <i>Xa31(t)</i> , <i>CGS-Xo1₁₁</i> and <i>Xa45(t)</i>) cloned in pCAMBIA1300.
BB2-F2b	ACACCCTTGTCCTCCTGCT	
Xa1-F2	CCCTCTGCACACGCCATTGG	
Xa1upT-F	CGTGTAACAAGTACTGACCA	
Xa14-F4	GAGGCTATGTGCATCCTCA	
gXa1BE-R4	CATTGTCTTAGATGATGTGT	
Xa1C-F	TCCAGTGAGAAGTTGGAGCA	
Xa14-R4	TGTGCACACAGGAAAGAGGA	
Xa1L-F	TCAGGAACTTGA ACTCCAG	

Xa1L-R4	TCGAGTGCAGTGCATGATTG	
Xa37d-R1	ACGGTTCTGAAGGTCGTCAT	
BB2-R2	GAAATTGGCTGATCCGAG	
M13F	TGTAAAACGACGGCCAGT	
86TalKDF1	TTCGGGGAAAGTAGCACTTGA G	Amplification of the upstream regions flanking sequences of <i>iTal3</i> and <i>iTal6</i> from PXO86
86TalKDR1	AATATCCGGGTAGGCGCAATC ACTCTACCGATGCGTGTCGTAA G	
86TalKDF2	AGCCTACACAATCGCTCAAGAC GTACGTGAACGAACGCCTCTAC	Amplification of the downstream regions flanking sequences of <i>iTal3</i> and <i>iTal6</i> from PXO86
86TalKDR2	TGACCTGTTCCAGCAGAGCTTG	
KD13-F	ACGTCTTGAGCGATTGTGTAGG CT	Amplification of kanamycin resistance gene
KD13-R	AGTGATTGCGCCTACCCGGATA TT	
86TAL3-F2	CAGACGTAAACGGTCCT	detection of the deletion of <i>iTal3</i> in PXO86 mutant
86TAL3-R2	ACGCTGCCAGGTCGGCAACC	
86TAL6-F2	GCATTCAAGGAAGAGGAAA	detection of the deletion of <i>iTal6</i> in PXO86 mutant
86TAL6-R2	AGTGTCTACGGAACCCTG	

Supplemental Table 2. Prevalence of *Xa1* allelic *R* genes in 3k rice

In a separate Excel file.

Supplemental Table 3. Prevalence of *Xa1* allelic *R* genes and disease reactions in 87 rice accessions.

Rice line	Source	Accession	Presence of <i>Xa1</i> marker	Number of 93 aa	PXO99 ^A	Δ Tal3	Δ Tal3 (piTal3a)	Δ Tal3 (piTal3b)
Kitaake			-		S	S	S	S
IRBB1	IRGC	418736	+	6	S	R	S	S
IRBB2	IRGC		+	5	S	R	S	S
Basmati	NSGC	Clor 12524	+	7	S	R		
Chikanari 2	NSGC	Clor 12452	-					
Taichung 65	NSGC	PI 275428	-					
Toyohikari	NSGC	Clor 12449	-					
Nagate Eikou	NSGC	Clor 12447	-					
Pi No. 1	NSGC	Clor 12445	-		S	S	S	S
Somewake	NSGC	Clor 12423	-					
Shinriki 1	NSGC	Clor 12401	-					
Palman	NSGC	Clor 12219	-					
Safed	NSGC	Clor 12218	-					
Jhona	NSGC	Clor 12217	-					
Mushkan	NSGC	Clor 12216	-					
Ziri	NSGC	Clor 12214	-					
Kinai 388	NSGC	Clor 12199	+	6	S	R		
Miyako	NSGC	Clor 12066	+	6	S	R		
Upland	NSGC	Clor 12001	+	6	S	R	S	S

Kinai Early No.70	NSGC	Clor 12197	-					
Aikoku Kinai	NSGC	Clor 12196	-					
Vansi	NSGC	Clor 12222	-					
Ainaura	NSGC	Clor 12252	+	6	S	R		
Mogamiuruti	NSGC	Clor 12254	+	7	S	R	S	S
Santaro	NSGC	Clor 12285	-					
Guneki	NSGC	Clor 12274	-					
Schou Sung Sensho	NSGC	Clor 12286	-		S	S		
Maratelli	NSGC	Clor 12294	-					
Dosan 50	NSGC	Clor 12398	-					
WC 5119	NSGC	Clor 12399	-		S	S		
Norin 23	NSGC	Clor 12400	-					
Shiokari	USDA	GSOR 300031	-					
Murasaki Daikoku	USDA	GSOR 310168	-					
P 79	USDA	GSOR 310380	-					
Toga	USDA	GSOR 310788	+	5	S	R	S	S
Zhongyu No.6	USDA	GSOR 311449	-					
Minghui 63	USDA	GSOR 311461	-					
NERICA 1	NSGC	WAB0007739	+					
NERICA 2	NSGC	WAB0018053	-					
NERICA 3	NSGC	WAB0037218	-					
NERICA 4	NSGC	WAB001802	-					
NERICA 5	NSGC	WAB0018042	+	5	S	R		
NERICA 6	NSGC	WAB0018092	-					
NERICA 7	NSGC	WAB0018095	+	6	S	R	S	S
NERICA 8	NSGC	WAB0019617	-					
NERICA 9	NSGC	WAB0019618	-					
NERICA 10	NSGC	WAB0019619	-					
NERICA 11	NSGC	WAB0019620	-					
NERICA 12	NSGC	WAB0019621	+	5	S	R		
NERICA 13	NSGC	WAB0019622	-					
NERICA 14	NSGC	WAB0019624	+	5	S	R		
NERICA 15	NSGC	WAB0019623	-					
<i>O.glaberrima-1</i>	IRGC	100854	-					
<i>O.glaberrima-2</i>	IRGC	100983	-					
<i>O.glaberrima-3</i>	IRGC	102277	-					
<i>O.glaberrima-4</i>	IRGC	102500	-					
<i>O.glaberrima-5</i>	IRGC	103700	-					
<i>O.glaberrima-6</i>	IRGC	103445	-					
<i>O.glaberrima-7</i>	IRGC	104033	-					
<i>O.glaberrima-8</i>	IRGC	104200	-					
<i>O.glaberrima-9</i>	NPGS	PI 369455	+	6	S	R		
<i>O.glaberrima-10</i>	NPGS	PI 450252	-					
<i>O.glaberrima-11</i>	NPGS	PI 450298	+	7	S	R	S	S
<i>O.glaberrima-12</i>	NPGS	PI 450353	-					
<i>O.glaberrima-13</i>	NPGS	PI 450363	-					

<i>O.glaberrima</i> -14	NPGS	PI 450396	-		S	S	S	S
<i>O.glaberrima</i> -15	NPGS	PI 450438	-					
<i>O.nivara</i> -1	IRGC	102463	+	7	S	R	S	S
<i>O.nivara</i> -2	IRGC	105722	-					
<i>O.nivara</i> -3	IRGC	105732	-					
<i>O.nivara</i> -4	IRGC	81825	-					
<i>O.barthii</i> -1	IRGC	100119	-					
<i>O.barthii</i> -2	IRGC	101248	-					
<i>O.barthii</i> -3	IRGC	106291	-					
<i>O.barthii</i> -4	IRGC	101937	-					
<i>O.officinalis</i> -1	IRGC	105081	-					
<i>O.officinalis</i> -2	IRGC	105080	-					
<i>O.officinalis</i> -3	IRGC	105223	-					
<i>O.officinalis</i> -4	IRGC	102386	-					
<i>O.officinalis</i> -5	IRGC	81972	-					
<i>O.officinalis</i> -6	IRGC	80730	-					
<i>O.minuta</i> -1	IRGC	101133	-					
<i>O.minuta</i> -2	IRGC	105132	-					
<i>O.minuta</i> -3	IRGC	105128	-					
<i>O.minuta</i> -4	IRGC	101097	-					
<i>O.rhizomatis</i> -1	IRGC	103421	-					
<i>O.rhizomatis</i> -2	IRGC	105949	-					

Supplemental Table 4. Bacterial strains used in this study

Strains or Plasmids	Relevant characteristics	Reference/Source
Strains		
<i>Escherichia coli</i>		
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacIⁿ ZAM15 Tn10</i> (Tetr)]	Stratagene
Trans1-T1	F- ϕ 80(<i>lacZ</i>) Δ M15 Δ <i>lacX74</i> hsdR(r_K^- , m_K^+) Δ <i>recA1398 endA1 tonA</i>	TransGen Biothch
<i>Agrobacterium tumefaciens</i>		
EHA105	C58 (<i>rif^R</i>) Ti pEHA105 (pTiBo542DT-DNA) Succinamopine	(Hood et al., 1993)
<i>Xanthomonas oryzae</i> pv. <i>oryzae</i>		
PXO99 ^A	Philippine race 6	This study
Δ Tal3	<i>iTal3a</i> and <i>iTal3b</i> knock-out mutant of PXO99 ^A	(Ji et al., 2016)
PXO86	Philippine race 2	This study
PXO86 Δ iTal3	iTALE A type gene knock-out mutant of PXO86	This study
PXO86 Δ iTal6	iTALE B type gene knock-out mutant of PXO86	This study
PXO61	Philippine strain	This study
PXO79	Philippine strain	This study
PXO112	Philippine strain	This study

KXO85	Korean strain	This study
JW89011	Korean strain	This study
K202	Korean strain	This study
T7174	Japanese strain	This study
H75373	Japanese strain	This study
Xoo2	Thai strain	This study
A3842	Indian strain	This study
PbXO7	Indian strain	This study
IXO56	Indonesian strain	This study
NXO 260	Nepalese strain	This study
CIAT1185	Colombian strain	This study
ZHE 173	Chinese strain	This study
C1	Chinese strain	This study
GD1358	Chinese strain	This study
HB21	Chinese strain	This study
Aust-2013	Australian strain	This study
Aust-R3	Australian strain	This study
AXO1947	Cameroon strain	This study
Plasmids		
pCAMBIA1300	Binary T-DNA vector for rice transformation	CAMBIA
piTal3a	pHM1 expressing <i>iTal3a</i> under <i>lacZ</i> promoter with FLAG in pZW	This study
piTal3b	pHM1 expressing <i>iTal3b</i> under <i>lacZ</i> promoter with FLAG tag in pZW	This study
piTal3aFb	<i>iTal3a</i> variant containing <i>iTal3b</i> C-terminus in pHZW	This study
piTal3bFa	<i>iTal3b</i> variant containing <i>iTal3a</i> C-terminus in pHZW	This study

Hood, E.E., Gelvin, S.B., Melchers, L.S., and Hoekema, A. (1993). New *Agrobacterium* helper plasmids for gene transfer to plants. *Transgenic research* 2:208-218.

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. *Nature communications* 7:13435.

Supplemental Table 5. Prevalence of iTALE genes in Xoo and Xoc genomes.

In a separate Excel file.