

# Broad-spectrum resistance to bacterial blight in rice using genome editing

Ricardo Oliva<sup>1,12\*</sup>, Chonghui Ji<sup>2,12</sup>, Genelou Atienza-Grande<sup>1,10,12</sup>, José C. Huguet-Tapia<sup>3,12</sup>, Alvaro Perez-Quintero<sup>1,4,11,12</sup>, Ting Li<sup>5</sup>, Joon-Seob Eom<sup>6</sup>, Chenhao Li<sup>2</sup>, Hanna Nguyen<sup>1</sup>, Bo Liu<sup>2</sup>, Florence Auguy<sup>4</sup>, Coline Sciallano<sup>4</sup>, Van T. Luu<sup>6</sup>, Gerbert S. Dossa<sup>7</sup>, Sébastien Cunnac<sup>4</sup>, Sarah M. Schmidt<sup>6</sup>, Inez H. Slamet-Loedin<sup>1</sup>, Casiana Vera Cruz<sup>1</sup>, Boris Szurek<sup>4</sup>, Wolf B. Frommer<sup>6,8\*</sup>, Frank F. White<sup>3</sup> and Bing Yang<sup>2,9\*</sup>

**Bacterial blight of rice is an important disease in Asia and Africa. The pathogen, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), secretes one or more of six known transcription-activator-like effectors (TALEs) that bind specific promoter sequences and induce, at minimum, one of the three host sucrose transporter genes *SWEET11*, *SWEET13* and *SWEET14*, the expression of which is required for disease susceptibility. We used CRISPR-Cas9-mediated genome editing to introduce mutations in all three *SWEET* gene promoters. Editing was further informed by sequence analyses of TALE genes in 63 *Xoo* strains, which revealed multiple TALE variants for *SWEET13* alleles. Mutations were also created in *SWEET14*, which is also targeted by two TALEs from an African *Xoo* lineage. A total of five promoter mutations were simultaneously introduced into the rice line Kitaake and the elite mega varieties IR64 and Ciherang-Sub1. Paddy trials showed that genome-edited *SWEET* promoters endow rice lines with robust, broad-spectrum resistance.**

*Xanthomonas oryzae* pv. *oryzae* (*Xoo*) is the etiological agent of bacterial blight disease in rice. The disease is most severe in southeast Asia but is increasingly damaging in west African countries, and results in substantial yield loss<sup>1</sup>. TALEs from *Xoo* are injected by a type III secretion system into plant cells and recognize effector-binding elements (EBEs) in cognate *SWEET* host gene promoters, which results in induction of *SWEET* genes and production of sugars that enable disease susceptibility in rice<sup>2,3</sup>. An array of central repeats, which are 34–35 amino acids long, are present in each TALE and interact with EBEs via two repeat variable di-residues (RVDs) at the 12th and 13th position of each repeat<sup>4,5</sup>. Aberrant repeats, longer than 35 amino acids, are hypothesized to allow looping out of the repeat to accommodate alternate sequence binding for a particular TALE<sup>6</sup>.

Bacterial blight depends on TALE-mediated induction of at least one member of a family of sugar-transporter genes. Although rice has more than 20 *SWEET* genes, only those of clade III are reported to be induced by *Xoo*<sup>7–10</sup>. Although all five of the known clade III *SWEET* genes in rice can function as susceptibility genes for bacterial blight, only three are known to be targeted in nature<sup>10</sup>. More specifically, *SWEET11* expression is induced by strains encoding the TALE PthXo1, *SWEET13* by PthXo2 and *SWEET14* by any one of several TALEs, namely AvrXa7, PthXo3, TalC and TalF (originally Tal5)<sup>7,9–15</sup> (Table 1). Effectors of *Xoo* that target clade III *SWEET* genes are referred to as major TALEs owing to their strong virulence effect.

Naturally occurring resistance has arisen as the result of nucleotide polymorphisms in EBEs of *SWEET* promoters. EBE alleles of *SWEET11* that are not recognized by PthXo1 are collectively referred to as the recessive resistance gene *xa13*. Rice varieties containing *xa13* are resistant to strains that solely depend on PthXo1 for *SWEET* induction. Most *indica* rice varieties carry a *SWEET13* allele that contains four adenines in the EBE for PthXo2, and rice lines carrying this allele are susceptible to PthXo2-dependent strains<sup>12</sup>. A rare exception is the recessive resistance allele *xa25*, which contains three adenines in the EBE for *SWEET13* in the *indica* cultivar Minghui 63, conferring resistance to strains that depend solely on PthXo2<sup>16</sup>. A similar recessive resistance allele in *japonica* rice varieties is equally effective against strains relying on PthXo2 (ref. <sup>12</sup>). Additional naturally occurring recessive EBE polymorphisms that confer resistance to strains carrying PthXo2, and the overlapping EBEs for PthXo3, TalF and AvrXa7 have subsequently been identified in the promoters of *SWEET13* and *SWEET14*, respectively, in germplasm collections<sup>17,18</sup>.

Rice susceptibility genes are good targets for genome editing for disease resistance. TALE-mediated susceptibility is particularly modifiable. For instance, transcription-activator-like effector nuclease (TALEN)-directed mutations in *SWEET14* created lines resistant to strains carrying PthXo3/AvrXa7 (refs. <sup>19,20</sup>). Importantly, characterization of TALEs and their cognate EBEs indicates that a limited set of EBEs are targeted by extant *Xoo* populations. At the

<sup>1</sup>International Rice Research Institute, Metro Manila, Philippines. <sup>2</sup>Division of Plant Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, USA. <sup>3</sup>Department of Plant Pathology, University of Florida, Gainesville, FL, USA. <sup>4</sup>IRD, CIRAD, Université Montpellier, IPME, Montpellier, France.

<sup>5</sup>Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA, USA. <sup>6</sup>Institute for Molecular Physiology and Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine Universität Düsseldorf and Max Planck Institute for Plant Breeding Research, Köln, Germany.

<sup>7</sup>Erfurt University of Applied Sciences, Erfurt, Germany. <sup>8</sup>Institute of Transformative Bio-Molecules (WPI-ITBM), Nagoya University, Aichi, Japan. <sup>9</sup>Donald Danforth Plant Science Center, St. Louis, MO, USA. <sup>10</sup>Present address: College of Agriculture and Food Science, University of the Philippines Los Baños, Los Baños, Philippines. <sup>11</sup>Present address: Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO, USA.

<sup>12</sup>These authors contributed equally: Ricardo Oliva, Chonghui Ji, Genelou Atienza-Grande, José C. Huguet-Tapia, Alvaro Perez-Quintero.

\*e-mail: [r.oliva@irri.org](mailto:r.oliva@irri.org); [frommew@hhu.de](mailto:frommew@hhu.de); [yangbi@missouri.edu](mailto:yangbi@missouri.edu)

**Table 1 | Known EBEs in *SWEET* promoters**

Host gene	TALe	EBE sequence
<i>SWEET11</i>	PthXo1	GCATCTCCCCCTACTGTACACCAC
<i>SWEET13</i>	PthXo2(A)	ATAAAGCACCACAACCTCCCTT
	PthXo2(A)	ATATAAGCACCACAACCTCCCTT
	PthXo2B	ATAAAGCACCACAACCTCCCTTC
	PthXo2C	ATAAAGCACCACAACCTCCCTTC
<i>SWEET14</i>	TalC	CATGCATGTCAGCAGCTGGTCAT
	PthXo3	ATATAAACCCCTCCAACCAGGTGCTAAG
	AvrXa7	ATAAACCCCTCCAACCAGGTGCTAA
	TalF	AAGCTCATCAAGCCTTCA

same time, relatively few strains have been assessed for TALE gene content, and optimized genome editing requires knowledge of the TALE variants present in bacterial populations. Targeting both prevalent and rare TALEs should increase the durability of resistance. In addition, the development of better tools for TALE characterization, which is a challenging and laborious procedure owing to their repetitive DNA sequences and presence of multiple copies per genome, would facilitate identification of novel TALEs and determination of gene-editing strategies<sup>21–23</sup>.

Here we report a detailed characterization of 63 *Xoo* strains from diverse geographic areas which includes genome sequence, the TALEs present and virulence assays. Building on the understanding of TALEs and their *SWEET* gene target EBEs, we developed a strategy to use CRISPR–Cas9 genome editing to engineer broad and durable resistance to bacterial blight in rice. An accompanying manuscript<sup>24</sup> reports the development of a disease management diagnostic tool kit for customized deployment of the resistant lines we report here.

### Diversity of TALEs in 63 *Xoo* strains

An assessment of the diversity of the major TALEs in extant strains of *Xoo* was obtained by analyzing *Xoo* genome sequence data and virulence assays on rice lines with *SWEET* promoter polymorphisms. TALE gene content was derived from an analysis of complete genome sequences of 33 strains from Asia (*Xoo*<sup>S</sup>) and 30 strains from Africa (*Xoo*<sup>F</sup>), which were either in databases or newly sequenced (Supplementary Table 1). Whole-genome, single nucleotide polymorphism (SNP)-based parsimony trees clearly separated *Xoo*<sup>S</sup> from *Xoo*<sup>F</sup> genomes, revealing two distinct evolutionary lineages (Fig. 1). Trees based on alignments of repeat regions of 856 TALEs revealed two major clusters (Supplementary Fig. 1). Expansion and diversification of TALE loci were more frequent in *Xoo*<sup>S</sup> (18–21 TALEs per genome) than in *Xoo*<sup>F</sup> genomes (nine TALEs per genome). Expansion and diversification of TALE loci is also reflected by the number of *SWEET*-gene-inducing TALEs, of which there are 13 variants (unique RVD sequences) in *Xoo*<sup>S</sup> strains and three variants in *Xoo*<sup>F</sup> strains (Supplementary Fig. 2).

TALEs predicted to induce *SWEET* genes (major TALEs) were present in all *Xoo* strains, regardless of origin (Fig. 1). Indeed, all strains possessed one or more close homologs of the known major TALE groups (PthXo1, PthXo2 and PthXo3/AvrXa7). PthXo3/AvrXa7 is treated as a single class owing to the closely overlapping EBEs in the *SWEET14* promoter. Asian strains had approximately equal numbers of PthXo2 (targeting *SWEET13*) and PthXo3/AvrXa7 (*SWEET14*) (24 versus 23, respectively; Fig. 1). Most Asian strains had multiple major TALEs with a combination of PthXo2 with either PthXo3 or AvrXa7 (Fig. 1). None of the strains contained both PthXo3 and AvrXa7 (Fig. 1). In African lineages, the two known major TALEs of *Xoo* are TalC and TalF, which each

induce *SWEET14* using non-overlapping EBEs<sup>10,13</sup>. Half of the analyzed African strains had both TalC and TalF, while none had TalF alone (Fig. 1). Overall, we found six non-overlapping EBEs in three *SWEET* promoters, if we treat the EBEs of PthXo2 variants that occur in some rice varieties as a single EBE (Table 1). These results support the hypothesis that activation of *SWEET* genes was a critical evolutionary step that occurred independently in different geographic regions during *Xoo* adaptation to rice.

### Diversity of TALEs from infection trials

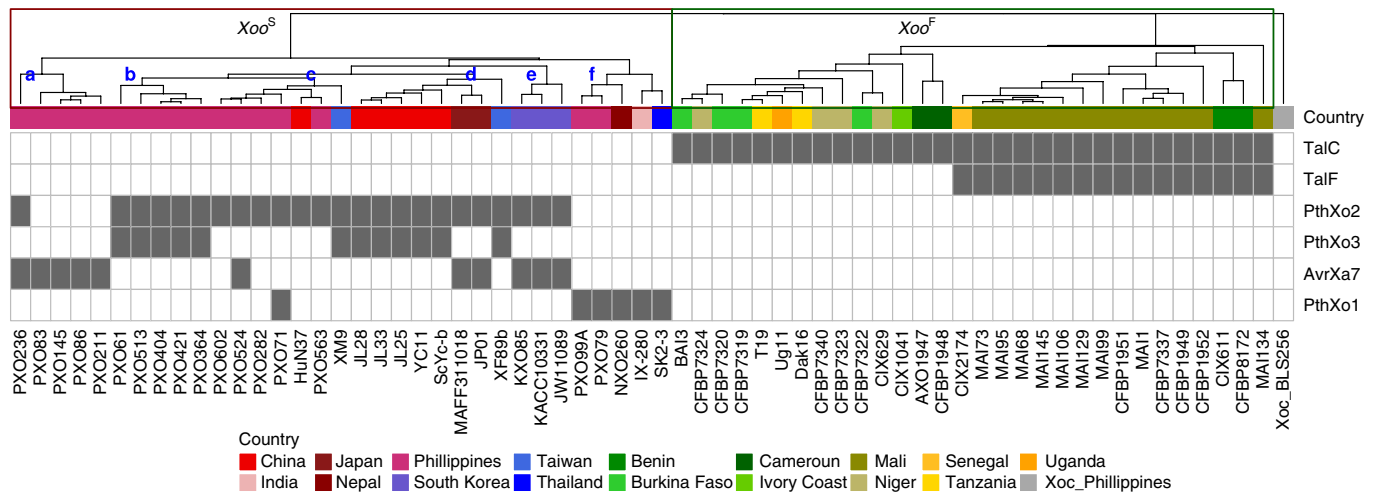
To estimate the prevalence and variation of TALEs in geographically diverse *Xoo* strains, a collection of 105 *Xoo* strains was screened against the *japonica* rice variety Kitaake and derivative Kitaake lines carrying mutations in the three *SWEET* genes normally targeted by TALEs. Kitaake contains a natural variant allele of *SWEET13* that is not recognized by PthXo2<sup>12</sup>. Strains were also screened against lines with alternate alleles of *SWEET11* and *SWEET14* generated by TAL effector nucleases (TALEN) both singly and in combination (Fig. 2 and Supplementary Table 2).

Ten of 105 *Xoo* strains were not virulent on the *japonica* reference line Kitaake. These ten strains were virulent on an *indica* reference line IR24 (Supplementary Fig. 3). IR24 contains few known resistance genes and serves as a recurrent susceptible parent line for breeding diverse near-isogenic resistance gene lines against bacterial blight<sup>25</sup>. These strains may be solely dependent on PthXo2 for induction of the IR24 *indica* allele of *SWEET13*, as the EBEs in the other *SWEET* gene promoters are the same in Kitaake and IR24. Alternatively, Kitaake may have a resistance gene that is effective against the ten strains. These ten strains were excluded from further screening.

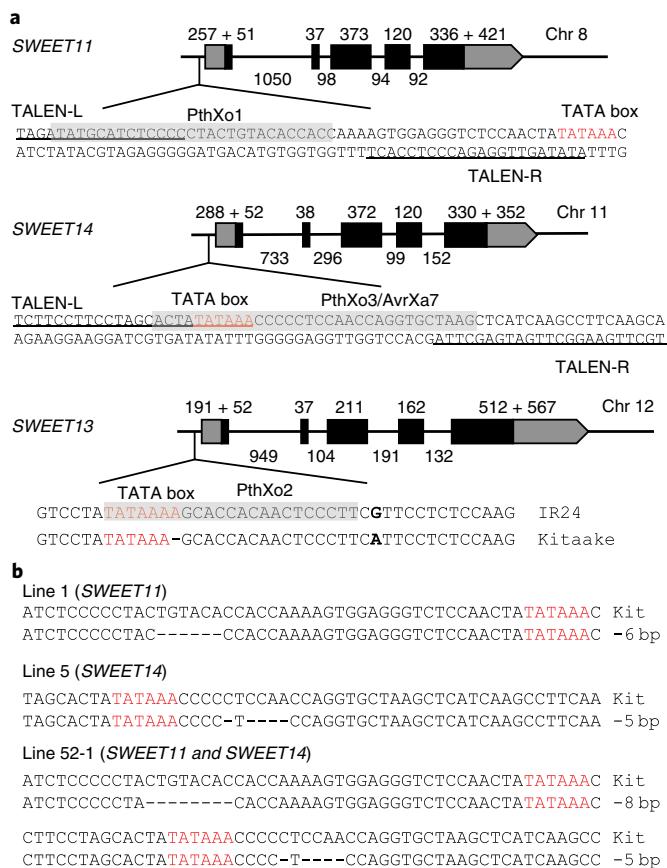
Of the remaining 95 *Xoo* strains, 18 were not virulent on edited rice line 1, which is defective in the EBE targeted by PthXo1 in *SWEET11*, and 69 strains were not virulent on line 5, which is defective in the EBEs of AvrXa7 and PthXo3 within *SWEET14* (Fig. 2 and Supplementary Table 2). Eighty-seven strains (18+69) were also not virulent on line 52-1, which harbors mutant alleles of both *SWEET11* and *SWEET14*, while the remaining eight strains (95 – 87) were virulent on line 52-1 (Fig. 2 and Supplementary Table 2). These findings indicate that 87 strains from the screen rely on major TALEs that target known EBEs of *SWEET11* and *SWEET14*, whereas eight ‘deviant’ strains have different virulence abilities and, presumably, different TALEs. These eight strains originated from the Philippines (PXO61, PXO364, PXO404, PXO421 and PXO513), the Korean peninsula (JW89011 and KXO85) and the African continent (AXO1947) (Supplementary Table 2). The virulence assay could not distinguish the deviant strains that, individually, carry both *avrXa7/pthXo3*-like and *pthXo2*-like or *pthXo1* and *pthXo2*-like TALE genes.

### Characterization of deviant *Xoo* strains

Genome sequencing and TALE content analysis of the seven deviant Asian strains identified *pthXo3* in the five Philippine strains (PXO61, PXO364, PXO404, PXO421 and PXO513) and *avrXa7* in the two Korean strains (JW89011 and KXO85) as judged by the similarity of RVDs (Fig. 1, Supplementary Fig. 2 and Supplementary Table 3). *pthXo1* was not found in any of the deviant strains. However, on the basis of similarity of the RVDs, all seven genomes had one of two variant *pthXo2*-like genes (Fig. 1 and Supplementary Table 3). One variant, named PthXo2B, is present in all five strains from the Philippines, while PthXo2C is present in the two Korean strains. All three PthXo2-related effectors have repetitive domains of 22 RVDs. PthXo2B differs in eight RVDs, while PthXo2C contains seven variant RVDs as compared to PthXo2 (Supplementary Fig. 4a). Both PthXo2B and PthXo2C were predicted to bind to the same EBE (ATAAAGCACCACAACCTCCCTTC) within the *SWEET13* promoter of Kitaake and other *japonica* varieties.



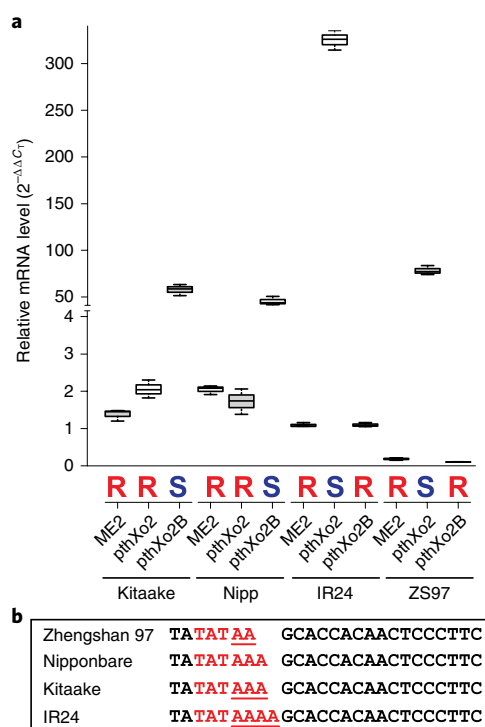
**Fig. 1 | Distribution of SWEET-inducing TALEs in *Xoo* strains.** Heat map indicates presence (gray) or absence of TALEs shown or predicted to bind to SWEET promoters in fully sequenced *Xoo* genomes. Origin is indicated. TALEs were grouped by DisTAL (Supplementary Fig. 1). Six groups containing previously characterized SWEET-inducing TALEs are shown. A parsimony tree (top) that is based on whole-genome SNPs obtained using kSNP is shown, highlighting the two main *Xoo* lineages: *Xoo*<sup>S</sup> (Asia) and *Xoo*<sup>F</sup> (Africa). Six sublineages were defined for *Xoo*<sup>S</sup>, and are indicated in bold blue letters (a-f). Seventy-three percent of branches had support values over 0.9 as determined by FasttreeMP implemented in kSNP3.



**Fig. 2 | TALEN-induced variants and natural SWEET variants to assess prevalence of major TALEs in 105 *Xoo* strains.** **a**, Schematic gene structures of SWEET11 and SWEET14 with promoter sequences targeted by respective TALEs shaded and with TALEN-bound sites underlined. Two natural variants of SWEET13 in two rice varieties are shown. The EBE for the respective TALE is highlighted in gray. **b**, Genotypes of three Kitaake-derived lines for disease assays. All lines have the Kitaake allele for SWEET13.

The predicted EBEs of PthXo2B and PthXo2C differ from the EBE for PthXo2 of SWEET13 in IR24, which is representative of most *indica* alleles (Supplementary Fig. 4b). Realigning RVDs with their respective EBEs and accounting for the shift in sequence owing to the single 'A' variation, the differences in RVD order between PthXo2 and PthXo2B or PthXo2C are reduced to four and five variant RVDs, respectively (Supplementary Fig. 4b). In addition to differences in the RVDs, PthXo2B and PthXo2C have 36-amino-acid repeats at RVD 9 and 12 relative to PthXo2. Atypical repeats have been found to accommodate alternative EBE binding in the major TALE AvrXa7 (ref. 6). The differences between PthXo2B and PthXo2C and the presence of PthXo3 or AvrXa7 reflect the overall separation of the sublineages from the Philippines and Korea (Fig. 1; sublineages b and e, respectively).

*Xoo* strains with PthXo2 alone are incompatible on Kitaake-derived lines owing to the inability to induce the *japonica* allele of SWEET13<sup>12</sup>. The presence of *pthXo2B* or *pthXo2C* could explain the virulence phenotypes of the seven deviant strains on Kitaake and the double-edited Kitaake line 52-1. Both effectors are predicted to enable induction of the *japonica* allele of SWEET13. To test this possibility, the candidate gene for PthXo2B from PXO61 was cloned and introduced into strain PXO99<sup>A</sup>ME2 (hereafter ME2), which is not pathogenic on any of the tested rice lines owing to a null mutation in the sole major TALE gene *pthXo1* of PXO99<sup>A</sup> (ref. 9). The *japonica* allele of SWEET13 was induced in Kitaake and Nipponbare only upon inoculation with ME2 carrying *pthXo2B*, while relative expression in the *indica* allele in variety IR24 was low upon inoculation with ME2 or ME2 carrying PthXo2B (Fig. 3a). The *indica* allele of SWEET13 in IR24 was induced upon inoculation with ME2 harboring PthXo2 (Fig. 3a). Lines in which SWEET13 was induced scored as susceptible based on lesion length, and lines without SWEET13 expression scored as resistant (Supplementary Fig. 5). An additional rice line (Zhengshan 97) carrying a different polymorphism in the EBE region for PthXo2-related effectors, namely a two A nucleotide deletion as compared to the *indica* allele, was also tested (Fig. 3b). The SWEET13 allele in Zhengshan 97 behaved similarly to the IR24 allele and scored as susceptible to ME2 with PthXo2 and resistant to ME2 alone or with PthXo2B (Fig. 3 and Supplementary Fig. 5).



**Fig. 3 | Functional analysis of PthXo2 homologs in four rice varieties.**

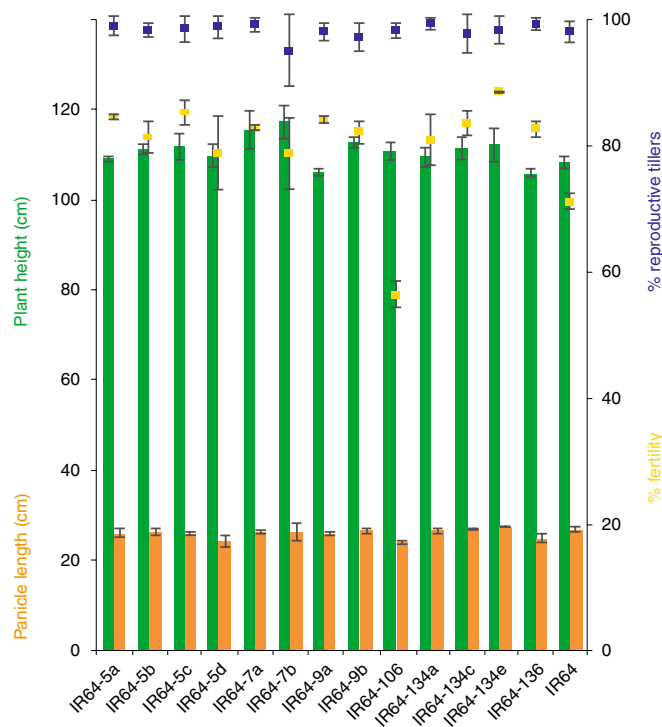
The genes *pthXo2* and *pthXo2B* were introduced into ME2 (a non-virulent derivative of PXO99<sup>A</sup>, lacking *pthXo1*). **a**, *SWEET13* induction in four rice varieties by *Xoo* strains (rice actin 1 as control). Data are plotted using BoxPlotR (see Methods); center lines show medians, box limits indicate the 25th and 75th percentiles. The experiment was repeated three times independently with comparable results. Bold letters above strains indicate susceptibility score (red R, resistance; blue S, susceptibility). Actual data for resistance or susceptibility are shown in Supplementary Fig. 5. ZS, Zhengshan. **b**, Nucleotide variations (underlined) in EBE regions for PthXo2-related TALEs in selected rice varieties. The predicted TATA box is highlighted in red.

### Stacking edited EBEs for broad-spectrum resistance to *Xoo*

We tested whether stacking multiple mutations in the EBEs of three *SWEET* promoters known to be targeted by major TALEs would provide resistance to most, if not all, *Xoo* strains in a single rice line. Using a multiplex CRISPR editing approach (Supplementary Figs. 6–8), a series of mutant lines were generated from Kitaake (Supplementary Table 4). The EBE mutant lines were first challenged with ME2 transformants, each carrying a gene for either PthXo1, PthXo2B, AvrXa7, TalC or TalF. The lines with the respective promoter mutations were resistant to ME2 derivatives carrying the corresponding TALEs (Supplementary Table 5). When challenged with eleven field strains, including the seven RVD-variant *Xoo* strains that can induce *SWEET13*, one edited Kitaake line (11.1.3) was resistant to all strains (Supplementary Table 5). The 11.1.3 line conferred moderate resistance to the *Xoo*<sup>F</sup> lineage strain AXO1947, which, of the known TALEs, only contains TalC, suggesting the presence of another unknown effector. A similar observation was obtained upon inoculation of *Xoo*<sup>F</sup> strain MA11, which has both TalF and TalC (Fig. 1). Parallel tests using *SWEET* knockout mutants showed that AXO1947 likely depends on both *SWEET13* and *SWEET14* (ref. <sup>24</sup>).

### Genome-edited *Xoo*-resistant mega varieties

Rice mega varieties, lines cultivated over one million hectares, are semidwarf and high-yielding lines that usually perform uniformly

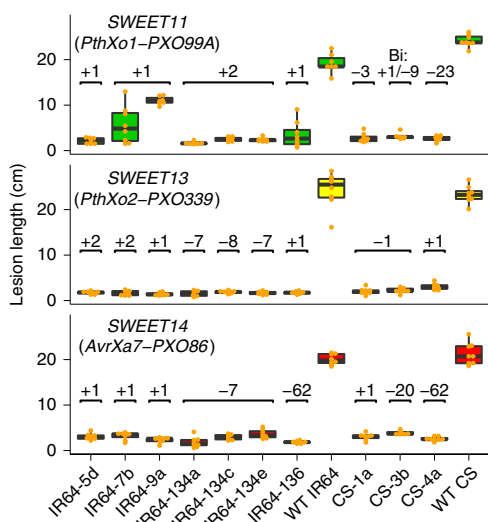


**Fig. 4 | Agronomic traits in selected genome-edited mega variety lines as compared to the parental controls.**

Performance of edited IR64-IRS1132 lines in the T3 generation (left axis: plant height in green, panicle length in orange; right axis: percentage fertility in yellow and percentage of reproductive tillers in blue) relative to parental controls (IR64) ( $n=15$ , no significant difference between IR64 and edited lines except fertility for line IR64-106,  $P < 0.01$ ). Microfield experiments in screenhouses for agronomic trait assessments were conducted using randomized complete block design with three replicates. Trials were done in a single season. Data are presented in Supplementary Fig. 10.

across multiple environments<sup>26</sup>. If the combined *SWEET* editing events tested here do not affect the normal physiological function of the genes, we predict that the same combinations would not impair agronomic traits in rice mega varieties, and would thus be useful for rice breeding programs. Genome editing by CRISPR–Cas9 was used to target *SWEET* promoters in two rice mega varieties, IR64 and Ciherang-Sub1 (Supplementary Fig. 9). Key agronomic traits were measured in confined paddy experiments. Before trials, the CRISPR-edited *SWEET* promoters of all mutant lines were sequenced, and lines were selected on the basis of the diversity of mutations in the respective EBEs (Supplementary Table 6). Overall, 13 IR64-IRS1132 and 18 Ciherang-Sub1-IRS1132 *Cas9*-free lines, spanning 30 combinations of EBE mutations in *SWEET11*, *SWEET13* and *SWEET14* were tested (Supplementary Table 7). Agronomic assessments and multivariate analysis of height, panicle length, number of reproductive tillers and fertility rate indicated that most lines performed similarly to wild-type parents (Fig. 4 and Supplementary Fig. 10). Only line IR64-106 showed phenotypic differences regarding yield, panicle length and fertility (reduced by 10% under paddy conditions). A range of lines with different insertions, deletions and substitutions in *SWEET11* ( $n=6$ ), *SWEET13* ( $n=8$ ) and *SWEET14* ( $n=7$ ) were selected, assuming that larger EBE differences will be less likely to match TALE variation in pathogen populations. Distinct combinations of these 21 *SWEET* variants were distributed across nine rice lines. The nine lines were challenged with tester *Xoo* strains harboring single major TALEs. All nine lines were resistant to all *Xoo* strains tested<sup>24</sup> (Figs. 5 and 6). Seven variants of *SWEET13*, ranging from



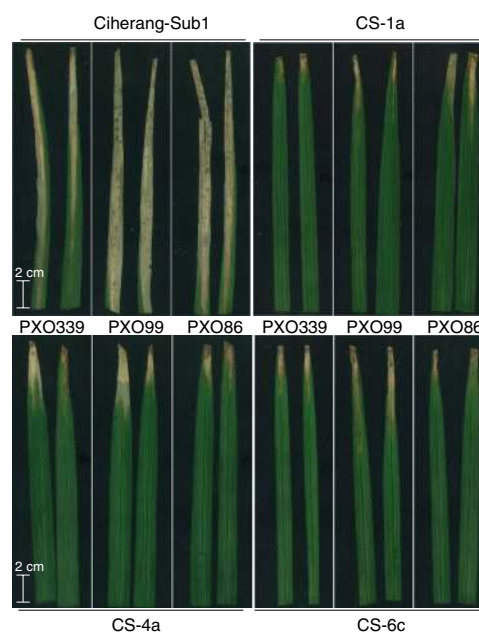


**Fig. 5 | Resistance of *SWEET* promoter edited IR64 and Ciherang-Sub1 lines.** Reactions of selected mutant lines of T3 IR64-IRS1132 and T2 Ciherang-Sub1-IRS1132 lines as compared to parental IR64 and Ciherang-Sub1 controls to infections with *pthXo1*-, *pthXo2*- and *avrXa7*-dependent *Xoo* strains (PXO99, PXO339 and PXO86, respectively). Different sequence alterations in the target *SWEET* promoters of edited lines resulted in varying levels of resistance to corresponding *Xoo* strains. Lesion lengths (cm) were measured 14 d after infection of specified plants. Types of mutations are indicated. Phenotyping experiments were conducted with four replicates per strain and two plants per replicate ( $n=8$ ), and between three and six inoculated leaf samples were scored per plant. All tested lines carried the same mutations in both alleles, except those marked as ‘Bi’ (biallelic, that is, two different mutations in EBE). Center lines show medians; box limits indicate 25th and 75th percentiles determined in the R software package; whiskers extend 1.5 times the interquartile range from 25th and 75th percentiles. Means for all edited lines are significantly different from controls (Dunnett’s test;  $P < 2 \times 10^{-16}$ ). A two-sided test was used.

two-nucleotide insertions (+2) to a seven-nucleotide deletion (−7), conferred resistance. Line IR64-9a carried only SNPs in *SWEET11*, which were insufficient to abrogate PthXo1 function (Fig. 5).

## Discussion

We set out to evaluate our hypothesis that editing known *SWEET* EBEs will endow rice with broad resistance to bacterial blight. The success of our approach depended on extant *Xoo* strains fitting the profile of the relatively few *Xoo* strains that have been genetically characterized. Indeed, analysis of a broad collection of 63 *Xoo* strains found that all strains targeted, or were predicted to target, *SWEET11*, *SWEET13* and *SWEET14*, as previously reported<sup>8,10,27</sup>. Furthermore, most *Xoo* strains, if not all, target some combination of six short EBE sequences present in the promoters of the three *SWEET* genes. We applied multiplexed CRISPR–Cas9 genome editing to systematically interfere with *SWEET* gene induction at all known major TALE EBEs and engineer Kitaake rice resistant to all currently known strains of *Xoo*. Kitaake is an excellent cultivar for proof-of-concept owing to high regeneration and rapid flowering cycle. The lines generated here will serve as diagnostic tools for rapid evaluation of the virulence of novel *Xoo* isolates<sup>24</sup>. Although Kitaake can be used for breeding resistance genes into Japanese and Chinese varieties, the *japonica* cultivar is not optimal as breeding material in large parts of southeast Asia and Africa, which use mainly *indica* varieties. Therefore, we also applied the same editing approach in two mega varieties, IR64<sup>26</sup> and Ciherang-Sub1<sup>28</sup>. The edited mega variety lines grew normally, without yield suppression,



**Fig. 6 | Resistance of three genome-edited Ciherang-Sub1 lines to three representative *Xoo* strains.** Appearance of lesions resulting from clipping leaf tips using scissors that had been dipped into bacterial cultures of three different *Xoo* strains (PXO339, PXO99 and PXO86, as indicated), for three genome-edited Ciherang-Sub1 mutant lines CS-1a, CS-4a and CS-6c (T2 generation) as compared to the parental control (Ciherang-Sub1). The experiment was repeated twice independently.

and were resistant to three representative strains of bacterial blight carrying the known major TALEs.

To devise an efficient genome editing strategy it was important to identify a panoply of TALEs and cognate EBEs to guide engineering of broadly applicable bacterial blight resistance. Sequence analysis of the seven deviant *Xoo* strains from Asia and relatedness of TALEs with close, but not identical, RVDs led us to propose that variants of PthXo2 target variant EBEs in *SWEET13*. These seven Asian strains retained virulence on the first-generation of genome-edited EBE lines in the Kitaake background, indicating that the variant strains either carry another major TALE targeting an undefined susceptibility gene or a variant major TALE targeting *SWEET11*, *SWEET13* or *SWEET14* at an alternate EBE. The two newly identified *pthXo2* members (*pthXo2B* and *pthXo2C*) were predicted to target the *SWEET13* promoters found in many *japonica* varieties, including Kitaake and Nipponbare. Indeed, PthXo2B conferred virulence to ME2, a strain lacking major TALE genes, on both Kitaake and Nipponbare, concomitantly with *SWEET13* induction. Thus, the presence of PthXo2B or PthXo2C in the deviant strains is sufficient to explain their virulence on the TALEN-edited lines. Ultimately, several lines with CRISPR-mediated edits at the PthXo2B-cognate EBE of Kitaake were resistant to *pthXo2B*-containing strains, indicating that the Kitaake allele for *SWEET13* was responsible for the susceptibility of line 52-1. Recent testing of all TALE genes of the strain PXO61 did not reveal any new major TALE other than PthXo3 and PthXo2B<sup>29</sup>. *Xoo*<sup>s</sup> strains carrying PthXo2B were obtained in the Philippines, where several bacterial blight outbreaks occurred. This sublineage is characterized as local races 1 and 9d, which have not been reported in other countries. At the same time, the Philippine archipelago may offer a safe geographic reservoir for less prevalent strains. Both strains from the Korean peninsula possess PthXo2C. The sampling of Korean strains was limited, and we are unable to state whether genes for PthXo2C are the rule or the exception. What is clear is that different variants of PthXo2 family are present in the

global *Xoo* population, suggesting an ongoing adaptation to nucleotide variations in the *SWEET13* promoter<sup>12,16,17</sup>.

African strains were shown to possess less TALE diversity and fall into two sublineages. All strains harbor the major virulence TALE, TalC, which targets *SWEET14*. Members of the second sublineage have TalC and TalF, the latter also targeting *SWEET14*. Another feature of the African strains is that loss of TalC function results in significantly reduced lesion length<sup>13</sup>. *Xoo* strains relying solely on TalC retained full virulence on Kitaake lines mutated in the TalC EBE within *SWEET14* (ref. <sup>19</sup>).

We show here that mutation of the TalC EBE in the promoter of *SWEET14* in Kitaake led to only moderate resistance to the African strain AXO1947. By contrast, double knock out of *SWEET13* and *SWEET14* resulted in complete resistance<sup>24</sup>. Consistent with previous studies<sup>19,30</sup>, the dependence of African *Xoo* strains on *SWEET* gene expression to cause symptoms appears more complicated than anticipated. Alternatively, lesion length may not be strictly correlated with ecological fitness or disease severity, and variation in lesion length, particularly at lower value ranges, may be due to variant host defense responses to diverse pathogen lineages.

Broad resistance to bacterial blight at the *SWEET* promoters will not prevent adaptation of the pathogen, and the durability of this approach will depend on the ability of *Xoo* populations to adapt to recessive resistance alleles. As noted, TALE genes have limited diversity, and, when diversity is present, it comprises minor variations. We propose that creating polymorphisms that are larger than a single nucleotide change (ideally modifying the whole EBE) in the EBEs of *SWEET* genes is advisable. These changes will still capture minor TALE variants of the major, for example, for PthXo2 and AvrXa7/PthXo3. The promoter editing will also likely thwart TALE adaptation, assuming that ease of adaptation to new binding sites is inversely related to the number of novel nucleotides in the target sequence. EBEs must meet certain structural requirements for TALE binding within the available sequences of the susceptibility gene promoters. Another consideration is that breeders will likely combine these recessive resistance alleles with other locally effective resistance genes and these combinations will likely severely reduce disease pressure, further delaying strain adaptation.

A diagnostic kit that can both identify the promoter variants best suited to defeat the local pathogen population and monitor variant strain emergence is presented in an accompanying publication<sup>24</sup>. Breeders and pathologists can work together to provide selected lines to farmers, with an overall strategy of deploying multiple lines in a region to disfavor the evolution of novel strains.

Multiplex targeting using CRISPR–Cas9 has created the ability to simultaneously edit all EBEs present in any single rice line. However, this method is limited owing to the constraint of possible edits and problems with the delivery of editing complexes. For example, the most efficient genome editing is via DNA-mediated delivery, which confers regulatory concerns on edited plants. Genome editors such as CRISPR–Cpf1 (also called Cas12a, which allows for larger mutations), potent ribonucleoprotein (RNP) delivery and transgene-free genome editing may mitigate these limitations and improve the durability of disease resistance.

In conclusion, by using a combination of systematic analyses of diverse *Xoo* strains, an understanding of *SWEET* genes and genome editing, we were able to engineer broad-spectrum resistance in Kitaake and two mega varieties IR64 and Ciherang-Sub1. Genome editing can result in off-target effects and the transformation process often leads to mutations and somaclonal variation<sup>31,32</sup>, but on- and off-target analysis of rice genomes in NCBI with the Cas9 guide RNAs revealed that all the predicted off-target sites lack PAM sequences (an essential component for Cas9 recognition) and contain mismatches larger than 4 nucleotides to the guide sequences within the single guide RNA (sgRNA) genes (Supplementary Table 8). Furthermore, off-target mutations (if they are present), and

other mutations derived from tissue culture and the regeneration process, will be eliminated in crosses during breeding. Full genome sequencing will be needed to identify possible off-target mutations and to ensure that the lines do not contain any T-DNA. Finally, our preliminary analysis of agronomic traits, while promising, is insufficient for direct deployment. Extensive field trials will be necessary to ensure performance of these edited resistant lines.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at <https://doi.org/10.1038/s41587-019-0267-z>.

Received: 27 February 2019; Accepted: 28 August 2019;

Published online: 28 October 2019

### References

- Niño-Liu, D. O., Ronald, P. C. & Bogdanove, A. J. *Xanthomonas oryzae* pathogens: model pathogens of a model crop. *Mol. Plant Pathol.* **7**, 303–324 (2006).
- White, F. F., Potnis, N., Jones, J. B. & Koebnik, R. The type III effectors of *Xanthomonas*. *Mol. Plant Pathol.* **10**, 749–766 (2009).
- Bezruczyk, M. et al. Sugar flux and signaling in plant-microbe interactions. *Plant J.* **93**, 675–685 (2018).
- Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
- Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
- Richter, A. et al. A TAL effector repeat architecture for frameshift binding. *Nat. Commun.* **5**, 3447 (2014).
- Antony, G. et al. Rice xa13 recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* **22**, 3864–3876 (2010).
- Chen, L. Q. et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527–532 (2010).
- Yang, B., Sugio, A. & White, F. F. *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc. Natl Acad. Sci. USA* **103**, 10503–10508 (2006).
- Streubel, J. et al. Five phylogenetically close rice *SWEET* genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **200**, 808–819 (2013).
- Yang, B. & White, F. F. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Mol. Plant Microbe Interact.* **17**, 1192–1200 (2004).
- Zhou, J. et al. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* **82**, 632–643 (2015).
- Yu, Y. et al. Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 *Os11N3* gene. *Mol. Plant Microbe Interact.* **24**, 1102–1113 (2011).
- Tran, T. T. et al. Functional analysis of African *Xanthomonas oryzae* pv. *oryzae* TALomes reveals a new susceptibility gene in bacterial leaf blight of rice. *PLoS Pathog.* **14**, e1007092 (2018).
- Doucouré, H. et al. Functional and genome dequence-driven characterization of TAL effector gene repertoires reveals novel variants with altered specificities in closely related malian *Xanthomonas oryzae* pv. *oryzae* strains. *Front Microbiol.* **9**, 1657 (2018).
- Liu, Q. et al. A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant Cell Environ.* **34**, 1958–1969 (2011).
- Zaka, A. et al. Natural variations in the promoter of *OsSWEET13* and *OsSWEET14* expand the range of resistance against *Xanthomonas oryzae* pv. *oryzae*. *PLoS ONE* **13**, e0203711 (2018).
- Hutin, M., Sabot, F., Ghesquière, A., Koebnik, R. & Szurek, B. A knowledge-based molecular screen uncovers a broad-spectrum *OsSWEET14* resistance allele to bacterial blight from wild rice. *Plant J.* **84**, 694–703 (2015).
- Blanvillain-Baufumé, S. et al. Targeted promoter editing for rice resistance to *Xanthomonas oryzae* pv. *oryzae* reveals differential activities for *SWEET14*-inducing TAL effectors. *Plant Biotechnol. J.* **15**, 306–317 (2017).
- Li, T., Liu, B., Spalding, M. H., Weeks, D. P. & Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **30**, 390–392 (2012).
- Booher, N. J. et al. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic structure and complex TAL (transcription activator-like) effector gene relationships. *Micro. Genom.* **1**, e000032 (2015).

22. Peng, Z. et al. Long read and single molecule DNA sequencing simplifies genome assembly and TAL effector gene analysis of *Xanthomonas translucens*. *BMC Genomics* **17**, 21 (2016).
23. Gochez, A. M. et al. Pacbio sequencing of copper-tolerant *Xanthomonas citri* reveals presence of a chimeric plasmid structure and provides insights into reassortment and shuffling of transcription activator-like effectors among *X. citri* strains. *BMC Genomics* **19**, 16 (2018).
24. Eom, J.-S. et al. Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-019-0268-y> (2019).
25. Ogawa, T., Yamamoto, T., Khush, G. S. & Mew, T.-W. Breeding of near-isogenic lines of rice with single genes for resistance to bacterial blight pathogen (*Xanthomonas campestris* pv. *oryzae*). *Jpn. J. Breed.* **41**, 523–529 (1991).
26. Mackill, D. J. & Khush, G. S. IR64: a high-quality and high-yielding mega variety. *Rice* **11**, 18 (2018).
27. Chen, L. Q. et al. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**, 207–211 (2012).
28. Toledo, A. M. U. et al. Development of improved Ciherang-Sub1 having tolerance to anaerobic germination conditions. *Plant Breed. Biotech.* **3**, 77–87 (2015).
29. Li, C. et al. An efficient method to clone TAL effector genes from *Xanthomonas oryzae* using Gibson assembly. *Mol. Plant Pathol.* <https://doi.org/10.1111/mpp.12820> (2019).
30. Carpenter, S. C. D. et al. A strain of an emerging Indian *Xanthomonas oryzae* pv. *oryzae* pathotype defeats the rice bacterial blight resistance gene *xa13* without inducing a clade III SWEET gene and is nearly identical to a recent Thai isolate. *Front. Microbiol.* **9**, 2703 (2018).
31. Hahn, F. & Nekrasov, V. CRISPR/Cas precision: do we need to worry about off-targeting in plants? *Plant Cell Rep.* **38**, 437–441 (2019).
32. Tang, X. et al. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol.* **19**, 84 (2018).

## Acknowledgements

We are grateful to A. Crepcia-Pevzner and J. Long for technical assistance; J. Jacobs for help in sequencing *Xoo* genomes; and G. Onaga, K. Wydra, H. Tall and A. von Tiedemann for supporting pathogen collection. We gratefully acknowledge funding from the Bill and Melinda Gates Foundation to Heinrich Heine University, Düsseldorf, with subawards to Iowa State University/University of Missouri, University of Florida, Institut

de recherche pour le développement and the International Rice Research Institute. Characterization of SWEET promoters was supported by a joint grant from the National Science Foundation to B.Y., F.F.W. and W.B.F. (IOS-1258018 and IOS-1258028). W.B.F. was supported by an Alexander-von-Humboldt Professorship.

## Author contributions

R.O., C.V.C., B.S., F.F.W., B.Y. and W.B.F. conceived and designed experiments. C.J., G.A.-G., J.C.H.-T., A.P.-Q., T.L., J.-S.E., C.L., V.T.L., C.S. and H.N. performed experiments. B.L. and I.H.S.-L. performed rice transformation. G.S.D. provided materials. J.S.E., C.J., J.C.H.-T., G.A.-G., A.P.-Q., F.A., S.C., B.Y. and W.B.F. analyzed the data. R.O., F.F.W., S.M.S., J.C.H.-T., B.S., B.Y. and W.B.F. wrote the manuscript, and S.M.S., J.S.E., G.A.-G. and A.P.Q. helped with revisions.

## Competing interests

W.B.F., J.-S.E., F.W., B.Y. and R.O. are inventors on a provisional patent application 62832300 that covers Kitaake, IR64 and Ciherang-Sub1 EBE-edited lines.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41587-019-0267-z>.

**Correspondence and requests for materials** should be addressed to R.O., W.B.F. or B.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2019



## Methods

**Plant material, bacterial strains, medium and growth conditions.** Rice varieties used here were *Oryza sativa* L. ssp. *japonica* Kitaake<sup>12</sup>, and *Oryza sativa* L. ssp. *indica* IR24 (ref. <sup>12</sup>), IR64 (ref. <sup>26</sup>) and Ciherang-Sub1 (ref. <sup>28</sup>). Rice plants were grown in growth chambers at 30°C for a 12-h light period and at 28°C for a 12-h dark period, with 60–75% relative humidity in the Yang laboratory, or under small-scale field environment in a screenhouse (28°C ± 7°C day and 23°C ± 4°C night; 80–85% relative humidity) at IRRI. *Escherichia coli* strains were grown in LB medium supplemented with appropriate antibiotics at 37°C. *Agrobacterium tumefaciens* strains were grown at 30°C in the dark. All *Xoo* strains were grown at 28°C in TSA (10 g l<sup>-1</sup> tryptone, 10 g l<sup>-1</sup> sucrose and 1 g l<sup>-1</sup> glutamic acid). Antibiotics were used at the following concentrations if required: 100 µg ml<sup>-1</sup> ampicillin; 10 µg ml<sup>-1</sup> cephalixin; 25 µg ml<sup>-1</sup> chloramphenicol; 25 µg ml<sup>-1</sup> kanamycin; 100 µg ml<sup>-1</sup> spectinomycin; and 10 µg ml<sup>-1</sup> tetracycline.

***Xoo* genome sequencing and assembly.** Total genomic DNA of *Xoo* strains was isolated using MagAttract HMW DNA Mini Kit (QIAGEN). DNA samples were sequenced using a single molecule real-time (SMRT-Pacbio) platform, P4-C6 chemistry. For each strain, at least two SMRT cells were used, generating around 180× coverage per genome. De novo assembly was conducted using the hierarchical assembly pipeline (HGAP) implemented in Canu v.1.5 software<sup>33</sup>. Raw PacBio reads were mapped against the resulting contigs using blasR aligner (<https://github.com/PacificBiosciences/blasr>) and corrections were conducted with the variant-caller software utilizing the arrow algorithm (<https://github.com/PacificBiosciences/GenomicConsensus>).

Genome annotation and gene prediction were conducted using the Prokka annotation pipeline<sup>34</sup> and the NCBI Prokaryotic Genome Annotation (PGAP)<sup>35</sup>. Genome sequences were deposited in the GenBank under BioProjects PRJNA497307 and PRJNA497605. Additional reported and publicly available genomes in GenBank were retrieved and used for comparisons (Supplementary Table 1).

**Phylogenetic and comparative genomic analyses of *Xoo* strains.** Phylogenetic analysis was conducted by identifying pan-genome SNPs in oligonucleotides of length  $k = 31$  and generating parsimony trees on the basis of these SNPs, as implemented in the program kSNP (<https://doi.org/10.1093/bioinformatics/btv271>).

**TALE content analysis in *Xoo*.** TALE annotation was conducted using the AnnoTALE software<sup>36</sup>. Additional annotation and comparisons were performed using Artemis genome browser and Artemis Comparative Tool<sup>37,38</sup>. A neighbor-joining tree based on the alignments of repeat regions of TALEs was obtained using DisTAL<sup>39</sup>. Target prediction for TALEs was conducted with Talvez<sup>40</sup>; the target used was sequences in the promoter regions (1 kb upstream from the translation start site) of the rice *japonica* Nipponbare genome version deposited in Phytozome database V11 and of the *indica* variety IR64 (v.CSHL1.0)<sup>41</sup>.

**CRISPR–Cas9 design and constructs.** For gene editing in Kitaake, the polycistronic tRNA–guide RNA (gRNA) system was used to generate multiple sgRNAs with different target sequences by flanking the sgRNAs with a tRNA precursor sequence<sup>42</sup>. Six intermediate vectors were constructed, pTLN-tgRNA-1 to T6, for six individual tRNA–gRNA units. A double-stranded DNA oligonucleotide for each site was produced by annealing two complementary oligonucleotides (24–25 nt) (Supplementary Table 9). Six double-stranded oligonucleotides were individually inserted into the *BsmBI*-digested pTLN-tgRNA-1 to T6 (Supplementary Fig. 7). The DNA sequence of positive clones was confirmed by Sanger sequencing (Supplementary Table 10). All six tRNA–gRNA units were transferred into another intermediate vector named pENTR4-U6.1P-cdB using the Golden Gate ligation method (Supplementary Fig. 8). The gRNA cassettes were finally mobilized to pBY02-ZmUbiP-OsCas9 (Supplementary Fig. 9) by using Gateway LR Clonase (Thermo Fisher Scientific) as described<sup>43</sup>. For gene editing in IR64 and Ciherang-Sub1, four gRNA genes (IRS1132; Supplementary Fig. 9), targeting PthXo1 EBE in *SWEET11*, PthXo2 EBE in *SWEET13*, or TalC EBE or AvrXa7 EBE in *SWEET14*, were constructed as described<sup>44</sup>.

**Generation of rice mutant lines.** *Agrobacterium* strains containing the respective CRISPR constructs were used for genome editing in Kitaake, IR64 and Ciherang-Sub1. Individual transformants were selected, propagated and regenerated into whole plants (T0) as described<sup>44,45</sup>. Leaf tissues collected from individual samples were homogenized in liquid nitrogen, and genomic DNA was isolated following the CTAB method. Initial screening for transformants (T0) was done via PCR of the phosphotransferase transgene (*hpt*) with primers HptF and HptR. Nuclease surveyor assay using Cel1 enzyme was employed in T0 plants to preselect lines with possible sequence alterations in the target regions. We selected 20 mono- and biallelic events for IR64 and sequenced the EBE regions using site-specific PCR amplicons obtained with primer pair 8NKp1-F5 and 8N3-R for *SWEET11*, primer pair 12N3-F1 and 12N3-R1 for *SWEET13* and primer pair 11N3-F2 and 11N3-R for *SWEET14*. To detect the CRISPR transgenes, paired primers OsCas9-F and OsCas9-R were used for Cas9, primers g8N3-F and g12N3-R for gRNA genes

targeting *SWEET11* and *SWEET13*, and primers g11N3-F and g11N3-R for gRNA genes targeting *SWEET14*. These primer pairs were also used for selection of candidate mutant lines in the advanced generations (T1 to T2). Primer information is provided in Supplementary Table 9. Nineteen gRNA-gene-free and Cas9-free IR64 T1 plants were selected and further analyzed. Thirty plants for each of the 19 IR64 T2 lines were phenotyped for resistance to *Xoo* strains PXO339, PXO99 and PXO86, and target promoter regions of the candidate lines were sequenced to confirm mutations. Seeds of T2 plants were bulked (from 30 plants per line), phenotyped for *Xoo* resistance and analyzed for agronomic traits as described below. Seeds from 13 individual T3 plants were advanced to T4 on the basis of three criteria: (1) mutation type, (2) consistent resistance to the three *Xoo* strains and (3) seed count. For Ciherang-Sub1, we obtained seven independent transformation events. Eighteen Cas9-free T1 plants were further analyzed for resistance and amplicon sequencing as described above up to the T3 generation.

**Evaluation for agronomic traits of the rice mutant lines.** We expected that mutant lines would show potential changes in agronomic traits owing to disruptive expression of the *SWEET* genes, off-target mutations across the genome or somaclonal variation. To assess the performance of the selected mutant lines, four agronomic characters in IR64-IRS1132 and Ciherang-Sub1-IRS1132 lines were assessed under paddy screenhouse conditions in a randomized block design experiment with three replications. At maturity, the plant height, panicle length, percentage of reproductive tillers (number of tillers with panicle/total number of panicles per plant) and percentage fertility (number of filled grains/total number of grains) were measured.

**Disease assays.** Fully expanded leaves of rice plants (6–8 weeks old) were inoculated using a leaf-tip clipping method. *Xoo* stocks, preserved in a –80°C freezer, were streaked out on TSA plates supplemented with appropriate antibiotics and incubated at 28°C for 2–4 d. The cells were harvested from the plates and resuspended in sterilized distilled water (optical density at 600 nm of 0.5, ~10<sup>8</sup> colony-forming units per milliliter). Scissor blades were immersed in *Xoo* suspension and used to clip about 2 cm from the leaf tip. The lesion lengths were measured 14 d after inoculation. Lesions were measured from each test plant. Lesion length measurements ≤ 5 cm were scored as resistant (R), 6–10 cm as moderately resistant (MR), 11–14 cm as moderately susceptible (MS) and ≥ 15 cm as susceptible (S). Three replications with about ten leaves from two to five plants per replicate were inoculated per strain.

**Statistical analysis.** Data are plotted using BoxPlotR (<http://shiny.chemgrid.org/boxplotr/>) or are presented as mean ± s.e.m. Numbers of measurements are specified in figure legends or in graphs. One-way analysis of variance (ANOVA) was conducted on all measurements. Tukey's honestly significant difference test was used for post-ANOVA pairwise tests for significance (set at  $P < 0.05$ ), or, alternatively, the two-sided Dunnett's test was used. Exact  $P$  values, statistical tests used and sample numbers ( $n$ ) can be found in figure legends or in graphs.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Links to genome sequencing data are available in Supplementary Table 3 and at BioProjects PRJNA497605 and PRJNA497307. Materials will be made available for non-profit research under a material-transfer agreement (Supplementary Note 1) and a standard material-transfer agreement (SMTA) that complies with the International Treaty on Plant Genetic Resources for Food and Agriculture (IT PGRFA) for IR64 or Ciherang-Sub1 lines (Supplementary Note 2). We aim at obtaining freedom-to-operate for use by low-income farmers and will work with breeders to make the materials available to subsistence farmers. For commercial applications, license agreements will be negotiated that are compliant with national and international law regarding biosafety, import and export, exchange of biological resources, intellectual property and seed systems. Profits will be used to support dissemination to subsistence farmers. Edited IR64- and Ciherang-Sub1-based materials can be obtained from R.O., edited Kitaake lines from B.Y., translational reporter lines from W.B.F. Distribution of *Xoo* strains may be restricted, owing to regulations on *X. oryzae* as a Select Agent by the US government, because of the Nagoya protocol, or because some strains were donated from other groups; these groups should be contacted directly (Supplementary Table 11).

## References

- Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive  $k$ -mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
- Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
- Haft, D. H. et al. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res.* **46**, D851–D860 (2018).



36. Grau, J. et al. Computational predictions provide insights into the biology of TAL effector target sites. *PLoS Comput. Biol.* **9**, e1002962 (2013).
37. Carver, T. J. et al. ACT: the Artemis comparison tool. *Bioinformatics* **21**, 3422–3423 (2005).
38. Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**, 464–469 (2012).
39. Pérez-Quintero, A. L. et al. QueTAL: a suite of tools to classify and compare TAL effectors functionally and phylogenetically. *Front. Plant Sci.* **6**, 545 (2015).
40. Pérez-Quintero, A. L. et al. An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. *PLoS One* **8**, e68464 (2013).
41. Schatz, M. C. et al. Whole genome de novo assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of *aus* and *indica*. *Genome Biol.* **15**, 506 (2014).
42. Xie, K., Minkenberg, B. & Yang, Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl Acad. Sci. USA* **112**, 3570–3575 (2015).
43. Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H. & Yang, B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucl. Acid Res* **42**, 10903–10914 (2014).
44. Hiei, Y. & Komari, T. *Agrobacterium*-mediated transformation of rice using immature embryos or calli induced from mature seed. *Nat. Protoc.* **3**, 824–834 (2008).
45. Slamet-Loedin, I. H., Chadha-Mohanty, P. & Torrizo, L. *Agrobacterium*-mediated transformation: rice transformation. *Methods Mol. Biol.* **1099**, 261–271 (2014).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

The promoter sequences of rice genes in rice varieties Nipponbare and IR64 were acquired from Phytozome database v11. Genome sequences of published Xoo strains were obtained from NCBI with genome accession numbers indicated in Supplementary Table 1.

Data analysis

Canu V1.5, BlasR aligner, Varent-Caller, Prokka annotation pipeline, and NCBI Prokaryotic Genome Annotation were used for de novo assembly, annotation and gene prediction of Xoo genomes; AnnoTALE, Artemis genome browser and Artemis Comparative Tool were used for annotation and comparison of TALEs; Talvez was used for prediction of TALEs. BoxPlotR (<http://shiny.chemgird.org/boxplotr/>) was used to plotted data; One-way analysis of variance (ANOVA) was conducted on all measurements; the Tukey honestly significant different test was used for post-ANOVA pair-wise tests for significance.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Xoo genome sequences were deposited in GenBank under Bioprojects PRJNA497307 and PRJNA497605.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="No sample size calculation was performed."/>
Data exclusions	<input type="text" value="No data were excluded in this study."/>
Replication	<input type="text" value="Most experiments were repeated at least three times independently. Where explicitly mentioned, two independent experiments are shown."/>
Randomization	<input type="text" value="N/A"/>
Blinding	<input type="text" value="N/A"/>

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging