



RESEARCH ARTICLE OPEN ACCESS

Removal of Transgenes and Evaluation of Yield Penalties in Genome Edited Bacterial Blight Resistant Rice Varieties

Eliza P. I. Loo^{1,2} | José C. Huguet-Tapia³ | Michael Selvaraj⁴ | Melissa Stiebner^{1,2} | Britta Killing^{1,2} | Marcel Buchholzer¹ | Van Schepler-Luu^{1,5} | Thomas Hartwig^{1,2,6} | Sandra P. Valdéz Gutierrez⁴ | Madlen I. Rast-Somssich¹ | Boris Szurek⁷ | Joe Tohme⁴ | Paul Charraviaga⁴ | Frank F. White³ | Bing Yang^{8,9} | Wolf B. Frommer^{1,2,10}

¹Faculty of Mathematics and Natural Sciences, Institute for Molecular Physiology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany | ²Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University, Düsseldorf, Germany | ³Department of Plant Pathology, University of Florida Gainesville, Gainesville, Florida, USA | ⁴The Alliance of Bioversity International and International Center for Tropical Agriculture (CIAT), Cali, Colombia | ⁵International Rice Research Institute, Metro Manila, Philippines | ⁶Independent Research Group, Max Planck Institute for Plant Breeding Research, Cologne, Germany | ⁷Plant Health Institute of Montpellier (PHIM), IRD, CIRAD, INRAE, Institut Agro, Université Montpellier, Montpellier, France | ⁸Division of Plant Sciences, Bond Life Sciences Center, University of Missouri, Columbia, Missouri, USA | ⁹Donald Danforth Plant Science Center, St. Louis, Missouri, USA | ¹⁰Institute for Transformative Biomolecules, ITbM, Nagoya, Japan

Correspondence: Wolf B. Frommer (frommew@hhu.de)

Received: 26 May 2025 | **Revised:** 5 August 2025 | **Accepted:** 8 August 2025

Funding: This work was supported in part by the Gates Foundation to HHU (Investment ID INV-008733 and Investment ID INV-063189), Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2048/1—project ID 390686111 (CEPLAS), and an Alexander von Humboldt Foundation (WF).

Keywords: Africa | field trial | India | regulations | TAL effector | whole genome sequencing | *Xanthomonas oryzae* pv. *oryzae* | yield

ABSTRACT

Bacterial blight (BB) of rice, caused by *Xanthomonas oryzae* pv. *oryzae* (Xoo), is one of the major drivers of yield losses in Africa and Asia. Xoo secretes TAL-effectors (TALE) that induce host SWEET sucrose uniporter by binding to the effector binding element (EBE) of SWEET promoters, likely required for Xoo reproduction and virulence. We had multiplex edited the EBEs of three SWEET genes to prevent TALE binding, producing genome-edited (GE'd) rice mega-varieties (IR64, Ciherang-Sub1 for Asia and Komboka for Africa) that were resistant to a wide spectrum of Xoo strains. Here, we report comprehensive analyses of the GE'd lines, including evaluation of agronomic performance in multi-location multi-season experimental field plots under different fertilisation regimes and tests for the presence/absence of foreign DNA/transgene in the offspring of GE'd lines (IR64-BC1T6, Ciherang-Sub1-BC1T5, Komboka-T3). Various strategies were evaluated, including herbicide tolerance, PCR, DNA gel blotting, whole genome sequencing (WGS), and specific tests stipulated by country-specific biosafety guidelines. Different WGS technologies were evaluated and also used to identify the heritability of the edits, single-nucleotide polymorphisms (SNPs), and insertions/deletions (indels) that might have resulted from somaclonal variation and potential GE-induced off-target mutations. Complete genome reference sequences for the parental lines IR64, Ciherang-Sub1, and Komboka are provided. In the field experiments, the GE'd lines did not show performance defects. Together, the results indicate that select GE'd lines do not contain foreign DNA or transgene fragments and fulfil the requirements for treatment equivalent to classical breeding lines in countries such as India and Kenya.

Eliza P. I. Loo, José Huguet-Tapia, Michael Selvaraj, Frank F. White, Bing Yang are equally contributed.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Plant Biotechnology Journal* published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd.

1 | Introduction

In 2023, 800 million tons of rice (*Oryza sativa* L.) were produced, providing >20% of the calories consumed worldwide (www.fao.org/faostat/en/#data/QCL). Still, in the same year, approximately 760 million of the >8 billion people on the planet faced hunger (<http://www.who.int/publications/m/item/the-state-of-food-security-and-nutrition-in-the-world-2024>). With the global population projected to reach around 10 billion by 2050, global food production will need to increase by nearly 40% to ensure adequate nutrition for everyone (www.un.org/development/desa/pd/content/World-Population-Prospechs-2022). Limited resources, national conflicts, and climate change are likely to exacerbate world hunger; therefore, ensuring food security is an urgent need. Pathogens and pests reduce the yield and quality of agricultural production, causing substantial losses in addition to threatening food security and impacting a large number of small-scale food producers in Asia and Africa. Globally, yield losses for rice are estimated at 25%–41%; the highest losses are associated with food-scarce regions with endemic and emerging pests and diseases (Savary et al. 2019). Losses are severe in Southeast Asia and increasingly damaging in West African countries (Niño-Liu et al. 2006); and despite the implementation of successful breeding programs, BB causes yield losses in the range of 8% in the Indo-Gangetic plain (Savary et al. 2019). Recent outbreaks in Tanzania and Madagascar caused by the inadvertent introduction of strains of Asian Xoo (*Xoo*^{Asia}) are impacting rice production in East Africa (Raveloson et al. 2023; Schepler-Luu et al. 2023). The virulence of Xoo critically depends on Transcription Activation Like effector (TALE) proteins that are injected into host cells where they bind to effector binding elements (EBEs) and, in turn, induce transcription of at least one of three SWEET sugar uniporter genes (*SWEET11a*, *SWEET13*, or *SWEET14*; Chen et al. 2012; Chen et al. 2010; Oliva et al. 2019; Streubel et al. 2013; Wu, Eom, et al. 2022). Notably, BB resistance has been found in a diverse germplasm due to nucleotide polymorphisms in EBEs in *SWEET* promoters: *xa13* (*SWEET11a*), *xa25* (*SWEET13*; Liu et al. 2011; Zhou et al. 2015), and *xa41t* (*SWEET14*; Hutin et al. 2015; Zaka et al. 2018). The resistance can be broken by strains that target either an alternate EBE in the same SWEET or another sucrose-transporting SWEET (Eom et al. 2019).

Theoretically, marker-assisted breeding could be used to combine *xa13*, *xa25*, and *xa41t*; this would, however, require challenging crosses between diverse germplasm, and importantly, since the *SWEET14* promoter contains multiple EBEs, resistance would not cover all known strains/TALE. A more promising approach would entail multiplex editing of all known EBEs in the three *SWEET* promoters with CRISPR/Cas9 to confer broad-spectrum resistance to virtually any known rice cultivar in Asia and Africa (Bezruczyk et al. 2018; Eom et al. 2019; Ni et al. 2021; Oliva et al. 2019; Ramasamy et al. 2025; Rifhani et al. 2023; Schepler-Luu et al. 2023). Using an infection assay in growth chamber or greenhouse conditions used widely by breeders as highly predictive for resistance in the field, multiplex edited Kitaake was resistant to 96 strains from a worldwide collection (Loo et al. 2024; Oliva et al. 2019). Multiplex editing had then been used for the foundational breeding lines (mega varieties), including IR64, a rice variety widely grown in South East

Asia, Ciherang-Sub1 (Bina Dhan11), a submergence-tolerant variety used widely in Indonesia, India, and Bangladesh, and Komboka, a popular Kenyan variety (Ng'endo et al. 2022; Pramudyawardani et al. 2020; Toledo et al. 2015). The resulting genome-edited (GE'd) lines in the three elite backgrounds showed resistance to representative extant isolates and did not exhibit obvious defects in growth or habit in greenhouse or screenhouse trials (Eom et al. 2019; Oliva et al. 2019; Schepler-Luu et al. 2023). As of May 2025, twenty-one countries implemented guidelines exempting transgene-free GE'd crops from genetically modified organism (GMO) regulations (Buchholzer and Frommer 2023; www.editagenome.org). Here, we present comprehensive analyses of BB-resistant GE'd elite lines (IR64, Ciherang-Sub1, and Komboka) to lay the basis for transfer to small-scale food producers in Asia and Africa (Eom et al. 2019; Oliva et al. 2019; Schepler-Luu et al. 2023). A key prerequisite for classifying a GE'd crop as equivalent to classical breeding lines is the effective removal of foreign DNA/transgene fragments by segregation. However, proof of removal is challenging since proving absence is impossible, since all technologies used have an inherent detection limit and potential flaws that prevent detection, e.g., enzymatic fragmentation for Illumina sequencing may overlook foreign DNA fragments if the enzyme did not cleave in a fully random manner or local coverage is insufficient. After careful testing for the effective removal of foreign DNA/transgenes used for editing by segregation, GE'd lines are eligible for transfer to interested government organisations in interested countries. Moreover, we show here in multiple experimental field plots performed in the absence of disease pressure in the Philippines and Colombia that the GE'd IR64 and Ciherang-Sub1 lines did not show performance defects.

2 | Results

As a basis for further dissemination, GE'd elite lines in three genetic backgrounds were evaluated for successful outcrossing of the transgenes and for field performance.

2.1 | Monitoring the Presence of Vector Sequences in GE'd IR64 and Ciherang-Sub1 Lines by PCR

Progenies from eleven GE'd IR64 lines (BC1T4) and eight GE'd Ciherang-Sub1 lines (BC1T3) resistant to Asian and African Xoo strains, previously validated to not carry the Cas9 gene, were transferred from IRRI to CIAT (Eom et al. 2019; Oliva et al. 2019). The GE'd IR64 and Ciherang-Sub1 lines underwent a multilayered pipeline for transgene analysis of GE'd plants (Figure 1). PCR was performed to screen for the presence of T-DNA in the GE'd lines using primer pairs specific to the CaMV35S promoter, gRNAs, ZmUBI10 promoter, and parts of *Cas9* (Table S1). PCR amplicons were not detected in any of the GE'd IR64 and Ciherang-Sub1 lines. Eleven GE'd IR64 and eight GE'd Ciherang-Sub1 lines were advanced two generations before they were subjected to DNA gel blotting and short-read WGS (Figure S1a, described in detail below). Concurrently, a parental T2 GE'd IR64-C3 line was transferred from IRRI to the University of Missouri and advanced independently of the lines at CIAT. Since transformation could lead to the insertion

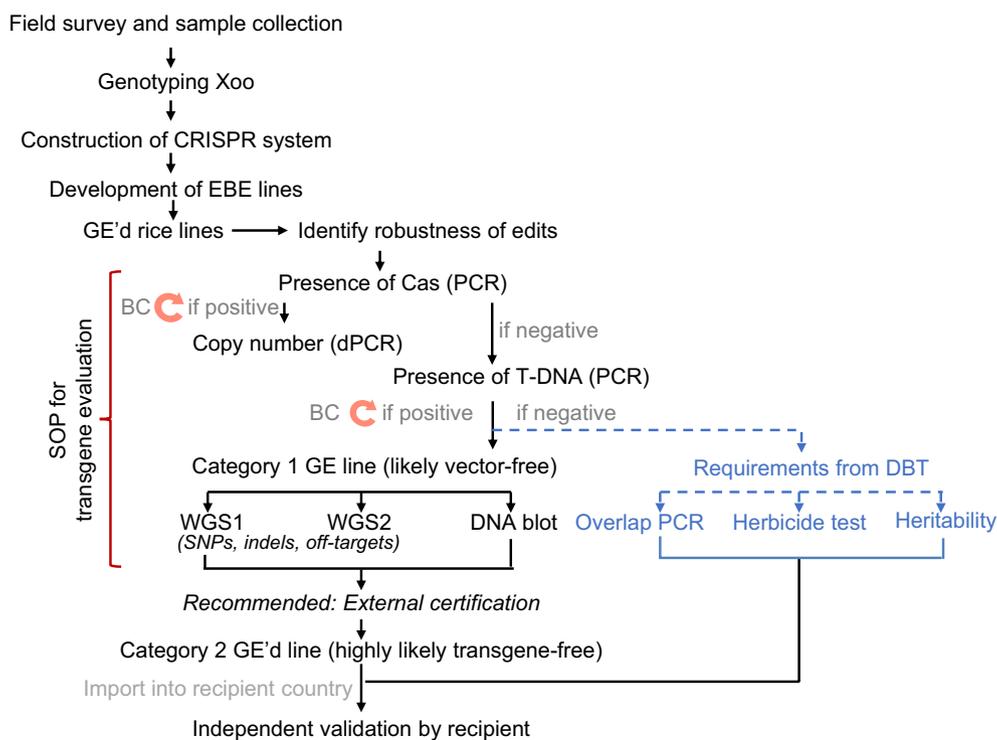


FIGURE 1 | Transgene evaluation pipeline for validation of elimination of transgenes. After identifying lines with an optimal combination of edits, F2 lines are analysed for the presence of transgenes. In the first step, lines are tested for the presence of Cas9 (or sequences close to the right border of the T-DNA) using RT-PCR and validated primers (including negative and positive controls). In parallel, it is useful to determine the copy number, e.g., by digital PCR (Schepler-Luu et al. 2023). Lines that do not appear to contain Cas9 are subjected to DNA gel blots using fragments covering the vector, overlap PCR (max amplicon size 500bp, minimal overlap 50 bp), as well as tests for the loss of the selectable marker (here hygromycin resistance) used for transformation. Lines are subjected to WGS in two rounds using independent enzymatic and mechanical fragmentation (provided Illumina or technologies that require fragmentation are used). Alternatively, and preferably, WGS is performed using long-read WGS technology with adequate coverage. It may be useful to obtain external validation and certification that the lines are transgene free based on certified GM tests. If in one of the step indications for the presence of a transgene are identified, lines will be backcrossed (BC) and subsequently have to re-enter the analysis pipeline.

of vector backbone fragments, the PCR screens for T-DNA were not considered fully conclusive.

2.2 | DNA Hybridization Analyses for Detecting Vector Sequences

To detect potential vector fragment integration, DNA gel blots were performed using eight overlapping DNA fragments as probes for detecting IRS1132 in six GE'd IR64 and two GE'd Ciherang-Sub1 lines (Figure S2, Table S1). Wild-type (WT) plants of the respective cultivars served as negative controls. Parental lines carrying vector integration (IR64-136b for IR64 lines and CS-3d for Ciherang-Sub1 lines) and WT plant DNA spiked with vector DNA of various copy number reconstructions were used as positive controls. In controls, DNA gel blots using all eight probes detected the presence of vector spiked into the genomic DNA of WT plants at the sensitivity of ≤ 0.5 copies of vector per genome (Figure S2). The probes did not detect vector sequences in the eight GE'd lines, providing support for the hypothesis of the absence of vectors at the levels of 0.5 copies per genome (Figure S2). Notably, however, we assume that the DNA gel blots cannot detect small insertions

(< 30 bp), and the sensitivity may depend on the length of the hybridising sequence.

2.3 | Whole Genome Sequencing of GE'd IR64 and Ciherang-Sub1 Lines

As a potentially more reliable alternative to detecting vector integration, WGS may be suitable for identifying transgene fragments. Genomic DNA of 11 GE'd IR64 lines and eight GE'd Ciherang-Sub1 lines was analysed by short-read paired-end WGS (Figure S1a). The first protocol used enzymatic digestion of the genomic DNA in library preparation for sequencing. Since enzymatic digestion of genomic DNA exhibits restriction site bias and may introduce indels, resulting in incomplete coverage (Knierim et al. 2011; Sato et al. 2019), a second round of WGS was performed using mechanical shearing via sonication. Leaf samples were pooled from two individual plants as a biological replicate. Genome sequences of GE'd IR64 and Ciherang-Sub1 lines were mapped to the vector pBY02UbiOs-Cas9_gSWEET11_13_14 (hereafter, IRS1132). We applied stringent inspection criteria to classify reads mapped to the vector as 'true' genome integration if:

- i. Mapped reads were not singletons/orphan reads, i.e., mapped reads must be paired. Since pair-end WGS was employed, each sequenced fragment should produce two reads. Singletons were defined here as sequencing fragments with only one of the reads successfully aligned to the vector.
- ii. For chimeric reads, soft-clipped reads had to map to the rice genome. For discordant read pairs, one pair had to map to the genome; the other pair had to map to the vector.
- iii. Mapped reads did not correspond to the native rice-derived U3 promoter (on the vector driving gRNA) or EBE (on the vector as gRNA).
- iv. Mapped reads were consistently found in all biological replicates in both WGS protocols.

Based on the criteria listed above, vector integration was not detected in six GE'd IR64 and two Ciherang-Sub1 lines, namely, IR64-5d, IR64-9a, IR64-134b, IR64-136a, IR64-7b, IR64-C3, CH-3b, and CH-4a (see Supplemental Methods for details, Figures 2 and S3, Table S2). An insertion of 273 bp derived from *Cas9*, i.e., a partial insertion without adjacent T-DNA sequences, was found on chromosome 12 of IR64-7a. To validate that singletons were not caused by small insertions, the genomes of a representative IR64 line, i.e., IR64-C3 (BC2T5), that had been subjected to short-read and long-read (without barcoding/multiplexing) WGS were compared. Notably, none of the singletons in IR64-C3 sequenced via

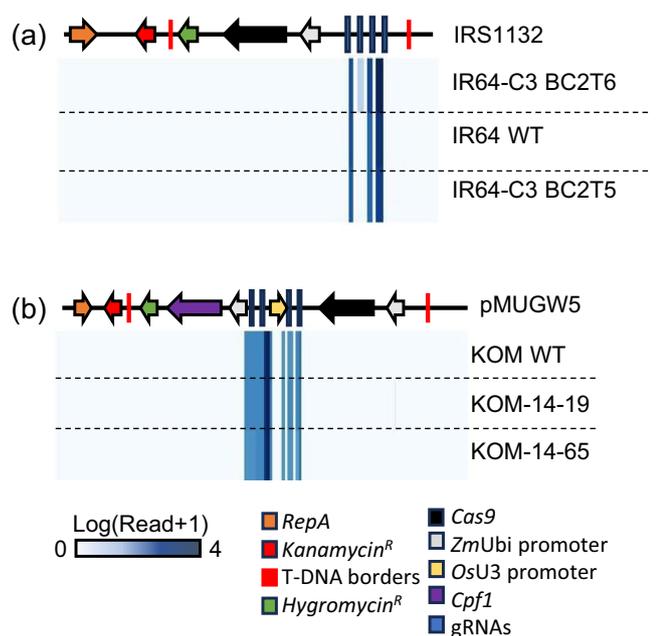


FIGURE 2 | Analyses for vector integration in GE'd rice lines. (a) Analysis for the presence of vector sequences by WGS. (a) Alignment of GE'd IR64 and Ciherang-Sub1 and (b) Komboka against vector sequences used for genome editing (IRS1132 for IR64 and Ciherang-Sub1, pMUGW5 for Komboka). Linearized IRS1132 and pMUGW5 indicated on the top of respective genome alignments. Legends indicate components in vectors IRS1132 and pMUGW5. The blue scale bar indicates number of reads in the sequenced genome of GE'd lines mapped to respective vectors.

short-read sequencing that mapped to the vector IRS1132 were found in the genome derived from long-read sequencing (Figures 2 and S3, Table S2). Since samples were pooled prior to short-read sequencing, we posit that reads mapped to the vector may be caused by index hopping, i.e., an artefact generated by misassignment of reads during demultiplexing (Kircher et al. 2012).

2.4 | Identification of SNPs, Indels, and Off-Target Mutations in GE'd IR64 and Ciherang-Sub1 Lines

Agrobacterium-mediated transformation of rice involves extended tissue culture phases known to induce somaclonal variation (Miyao et al. 2012). In addition, new 'natural' mutations occur in the progeny. To evaluate new mutations, the genomes of the GE'd lines were inspected for SNPs and indels. To avoid assessing SNPs/indels generated due to sequencing errors, only indels/SNPs consistently detected in all replicates from both WGS protocols (i.e., consensus indels/SNPs) were considered (Liu et al. 2021). Indels/SNPs in a representative GE'd line, IR64-C3, were mapped across three generations, including IR64-C3 T1, the T3 generation backcrossed once with WT IR64 (IR64-C3 BC1T3), and a fifth-generation progeny from two backcrosses (IR64-C3 BC2T5). In the T1 generation, 303 indels/SNPs were evenly distributed across all chromosomes (Figures 3 and S4a). In IR64-C3 BC1T3, the number of indels/SNPs was reduced to 131 and further reduced to 120 in IR64-C3 BC2T5, indicating successful reduction of somaclonal variation. The natural mutation rate was estimated to be 5.4×10^{-8} per rice generation (Tang et al. 2018). The GE'd IR64 and Ciherang-Sub1 lines have been cultivated for seven and six generations, respectively, after *Agrobacterium*-mediated transformation. Therefore, the theoretical number of 'natural' SNP/indel expected after seven generations of cultivation was 163 and 139 after six generations. The observed number of SNPs/indels in IR64-9a BC1T6 was 242, 270 in IR64-136a BC1T6, and 226 in IR64-7b, whereas 48 SNPs/indels in CS-3a BC1T5 and 74 SNPs/indels in CS4 BC1T5 were within the expected range (Figure S4b,c).

A range between 10 and 56 indels and 38 and 214 SNPs were identified in the genomes across all GE'd lines (Figure S4b). Indels and SNPs were traced to their respective genes/chromosomal locations. Between 31 and 176 SNPs were found in intergenic regions (> 75%); 5–30 non-intergenic SNPs were present in introns (3%–24%), and < 20 SNPs fell into exons (< 10.5%). Exonic SNPs were further inspected for potential disruption of gene function. Between two and five exonic SNPs were silent mutations; 1–16 were missense mutations. Between 9 and 42 indels were in intergenic regions, 0–13 indels in introns, and one indel in the exon of LOC_Os02g32570.1 (NF2 family N-terminal domain-containing protein), resulting in an in-frame deletion in IR64-136a (Table S3). Lines IR64-7b, CS-3b, and CS-4a carried one indel each that resulted in a frameshift in g9701 (in IR64-7b; new gene of unknown function found by reannotation), LOC_Os01g13760.1 (in CS-3b; heat shock protein 40; Sarkar et al. 2013), and LOC_Os1g50320.2 (in CS-4; suppressor-of-white-apricot (SURP) module family protein; Fan et al. 2024; Nameki et al. 2022; Figure S3b, Table S3).

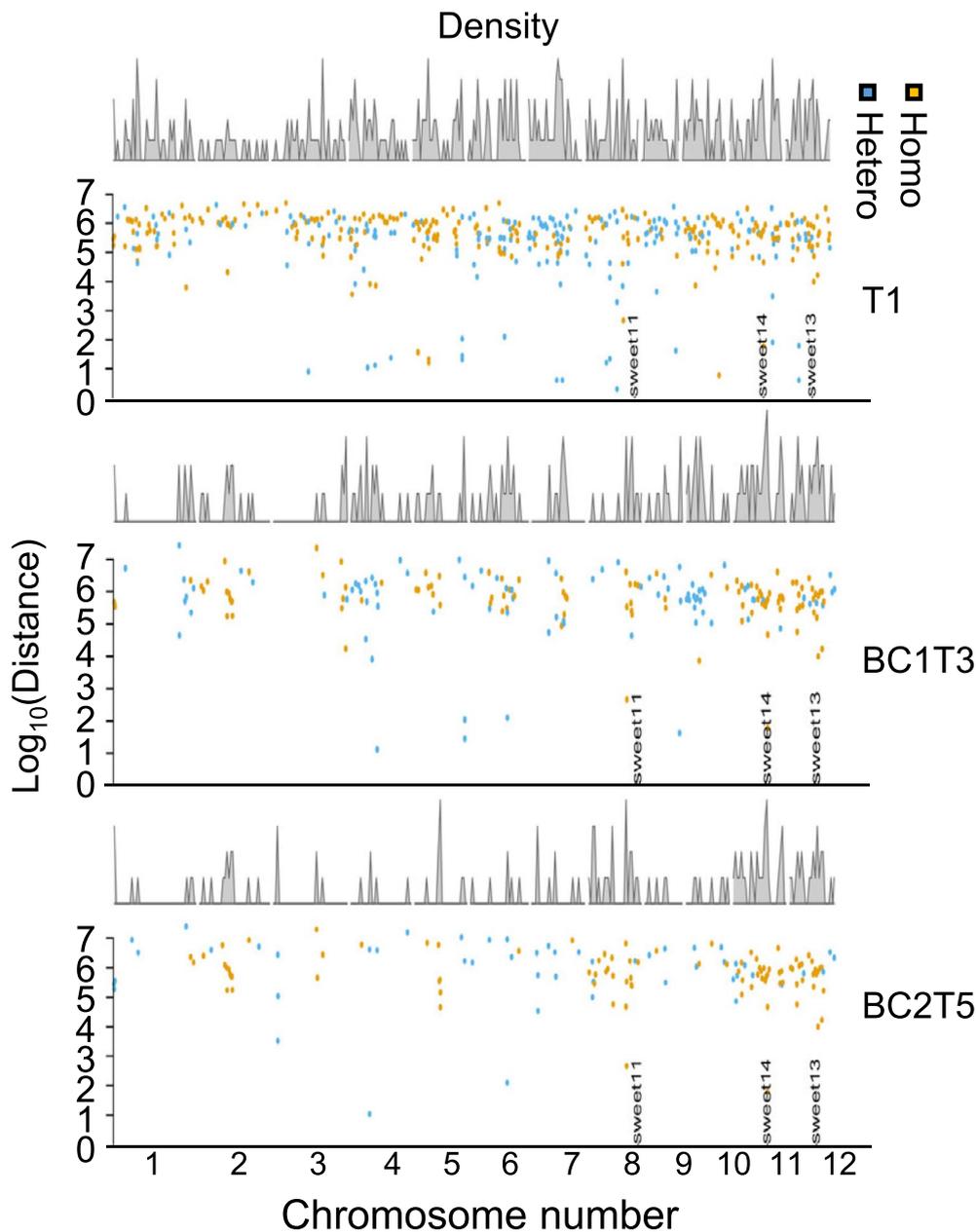


FIGURE 3 | Analysis of SNPs, indels, and potential off-target mutations in GE'd rice lines. Density and distribution of SNPs detected in each chromosome of IR64 C3 (BC2T5) and its parental genomes (T1 and BC1T1). Colours of dots indicate the type of SNPs, i.e., Homo, homozygous SNP; Hetero, heterozygous SNP. Location of SNPs, i.e., chromosome number indicated on the X-axis. 'sweet1' indicates mutation introduced in EBE of PthXo1, 'sweet1' indicates mutation introduced in EBEs of PthXo3, AvrXa7, and TalF, 'sweet1' indicates mutation introduced in EBE of PthXo2.

2.5 | Potential Off-Target Mutations Detected by WGS in GE'd IR64 and Ciherang-Sub1 Lines

Cas9 tolerates up to three mismatches to the gRNA and can cause off-target mutations (Fu et al. 2021; Hsu et al. 2013; Wang et al. 2016). We analysed the genome sequences of the GE'd lines for potential off-target mutations. *Bona fide* off-target mutations were considered only for mismatches for the sequence corresponding to the guide RNA plus five bp up- or downstream (Hsu et al. 2013). All four gRNAs were randomly (*in silico*) mutated with 1–5 mismatches, and the genomes of GE'd lines were subsequently screened for potential mutated gRNA targets. Two potential Cas9 off-targets were detected in the genome of IR64-136a, and one potential off-target in the genomes of IR64-9a and IR64-7b. The

potential off-target mutations, i.e., one-bp insertion in chromosome 1 position 6314428 in IR64-9a, IR64-136a, and IR64-7b, and eight-bp deletion in chromosome 5 position 15429580 in IR64-136a, are likely only in the event where the gRNA targeting the EBE of PthXo2 undergoes three random mutations (Table S4). All potential off-targets were traced to intergenic regions.

2.6 | Analysis of Vector Integration, SNPs/Indels, and Off-Target Mutations for GE'd Komboka Lines

Progenies of GE'd Komboka lines previously validated to be resistant to Asian and African Xoo (Schepler-Luu et al. 2023), i.e., eight T3 lines and four T2 lines, were subjected to our

multilayer pipeline for transgene analysis (Figure 1). Of the eight T3 lines and five T2 lines screened via PCR using primers specific to the *Cas9* fragment, amplicons were not detected in all eight T3 lines and three T2 lines (Figure S1b). KOM-14-19 and KOM14-65 were advanced to T3 and T4, respectively, for WGS validation for the absence of vector integration. To avoid the possibility of index hopping, GE'd KOM14-19 and KOM14-65, along with WT plants, were axenically grown and whole-genome sequenced using long-read WGS without multiplexing. Vector integration was not detected in GE'd Komboka lines based on the criteria used for evaluating GE'd IR64 and Ciherang-Sub1 lines (Figure 2b). SNP and indel analyses indicated that KOM-14-19 carried one indel that results in a frame-shift in LOC_Os03g57720.1, encoding a putative aldehyde oxidase 2 (Wu, Kamanga, et al. 2022; Figure S4b,c, Table S3). *Cas9* off-target analysis did not reveal potential off-targets in GE'd Komboka lines (Table S4).

2.7 | Tests to Fulfil Specific Requirements Set by the Department of Biotechnology (DBT), India

To classify GE'd crops as conventional breeding materials, and as a prerequisite for import, the DBT guidelines stipulate two essential tests for the presence of exogenously introduced DNA in the genome of GE'd plants: (i) successful removal of the selection/screenable markers and (ii) overlap PCR on genomic DNA for vector sequences used for generating the DNA modification. T-DNA-containing T3 parents, i.e., IR64-136b for IR64, CS-3d for Ciherang-Sub1, and T2 Komboka parent KOM-18-68, were used as positive parental controls to test for herbicide resistance.

In contrast to controls containing the T-DNA, WT and GE'd lines without detectable vector integration, as determined by WGS analyses, did not grow on media supplemented with hygromycin (Figure S5a). Accordingly, overlap PCR using 49 (IRS1132 in Ciherang-Sub1), 50 (IRS1132 in IR64), and 59 (pMUGW5 in Komboka) primer pairs (Table S1) did not produce amplicons in WT and GE'd lines, but expected amplicons (length \leq 500 bp; \geq 50 bp overlap) were observed in two positive controls, i.e., (i) a segregating parental line harbouring full or part of the vector in its genome and (ii) genomic DNA of the WT plant spiked with the vectors diluted 1000-fold (Figure S5c–e). In addition to foreign DNA/transgene analysis, DBT also requires evidence that the edited traits are heritable. The EBEs from two generations of each GE'd line were sequenced at least twice and aligned to ensure that the EBEs were inherited. The edited EBEs were identical across all generations of the GE'd line tested (Table S5), indicating the heritability of the edited trait. The information generated for transfer to India was sufficient to allow for the import of the edited lines into Kenya (KALRO; Table 1).

2.8 | Genome-Edited Rice Lines Did Not Exhibit Yield Penalties in Experimental Field Plots (EFP)

To evaluate whether the *Agrobacterium*-mediated editing process impacted growth or yield under open-field conditions, EFPs were performed in the absence of BB disease pressure in the Philippines and Colombia. EFPs in the absence of disease pressure can reveal potential yield penalties that can not be observed in the presence of disease due to the yield loss of the susceptible parents. Opportunities for EFP were limited due to

TABLE 1 | Summary of test results for transgene evaluation of GE'd lines.

GE line	PCR Cas9	PCR vector	WGS	DNA blot	Overlap PCR	Hyg ^R	Transfer
IR64-5d	Nd	Nd	Nd	Nd	Nd	R	
IR64-7a	Nd	Nd	273 bp <i>Cas9</i>	Nt	Nt	Nt	
IR64-9a	Nd	Nd	Nd	Nd	Nd	R	Kenya
IR64-134a	Nd	Nd	Recurring singletons	Nt	Nt	Nt	
IR64-134b	Nd	Nd	Nd	Nd	Nd	R	
IR64-136a	Nd	Nd	Nd	Nd	Nd	R	Kenya
IR64-7b	Nd	Nd	Nd	Nd	Nd	R	
IR64-134d	Nd	Nd	Recurring singletons	Nt	Nt	Nt	
IR64-C3	Nd	Nd	Nd	Nd	Nd	R	Kenya
CS-3b	Nd	Nd	Nd	Nd	Nd	R	
CS-1 h	Nd	Nd	Recurring singletons	Nt	Nt	Nt	
CS-4	Nd	Nd	Nd	Nd	Nd	R	
CS-1b	Nd	Nd	Recurring singletons	Nd	Nt	Nt	
CS-6d	Nd	Nd	Recurring singletons	Nd	Nt	Nt	
KOM-14-19	Nd	Nd	Nd	Nt	Nd	R	Kenya
KOM-14-65	Nd	Nd	Nd	Nt	Nd	R	Kenya

Note: Singletons: Sequenced fragments with only one read aligned to the vector. Recurring singletons: Singletons found in more than one biological replicate. Abbreviations: Nd, not detected (in the respective assay used); Nt, not tested.

regulations for genetically modified organisms. In Colombia, EFPs were performed at CIAT in Palmira, Valle de Cauca, Colombia, and in Los Baños, The Philippines based on the availability of permitted and suitable infrastructure and the absence of BB. Eight GE'd IR64 and six Ciherang-Sub1 lines qualified for EFPs after quarantine measures were advanced to BC1T5 for agronomic performance testing (Figure S1a). Unmanned aerial vehicle (drone)-based multispectral analyses were performed in EFPs across two planting seasons. The first trial took place during the 2020 dry season (June 2020–October 2020), and the second trial during the 2021 wet season (January 2021–May 2021). Normalised Difference Vegetation Indices (NDVI) and Normalised Difference Red Edge Indices (NDRE) were used as quantitative measurement proxies for canopy development and chlorophyll levels (Figure 4a; Huang et al. 2021). NDVI and NDRE values range between 1 and -1 , with higher values indicating optimal vegetation scores. Both the WT and GE'd lines showed high NDVI value ranges of 0.8–0.91 in the dry season, indicating high photosynthetic capacity (Tables S6 and S7). NDVI values for GE'd and WT plants in the wet season of 2021 were lower compared to NDVI values from the dry season of 2020, ranging from 0.33 to 0.83. The NDRE values for the GE'd and WT lines ranged from 0.29 to 0.55 during the dry season and between 0.2 and 0.5 during the wet season (Tables S6 and S7). The lower NDVI and NDRE values obtained in the wet season may be a consequence of the overall reduced sunlight (Figure S6a). NDVI and NDRE indices showed no significant differences in the overall photosynthetic capacity and canopy

development between GE'd and WT plants across both planting seasons (Figure 4a, Tables S6 and S7). Additionally, eleven agronomic traits were quantified in three EFPs, including two trials in Colombia (dry season from June 2020 to October 2020 and wet season from January 2021 to May 2021), and one trial in the Philippines (wet season from December 2020 to March 2021; Tables 2 and S8). Across the three EFPs, agronomic trait evaluation indicated that the GE'd IR64 and Ciherang-Sub1 cultivars did not exhibit penalties in agronomic traits and total yield (Tables 1 and S8). A reduced 1000-grain weight was observed from GE'd IR64 lines, an average of 27.3 g compared to 29 g from WT plants in the dry season, and an average of 26.4 g compared to 28.7 g from WT plants in the wet season. However, bulk grain weight, grain yield per m^2 , single plant grain yield, and yield per hectare of GE'd IR64 lines were not different relative to WT IR64 controls (Table S8). Principal component analysis clustered the rice lines based on geographical and seasonal harvest, and not by genotypic differences (Figure 4b). Drone-based multispectral and agronomic performance analyses were performed to further evaluate whether stress conditions might impact the performance of the GE'd lines with reduced nitrogen. Neither NDVI nor NDRE at 1- to 2-week intervals nor agronomic traits of plants grown with or without N-fertiliser supplementation (0, 90, and 180 $kg N ha^{-1}$; basal nitrogen content of 60 $kg ha^{-1}$) had differences between GE'd and WT plants (Figure S6b, Tables S9 and S10). PCA analysis did not separate the different GE'd lines, nor between GE'd and respective WT lines (Figure S6b).

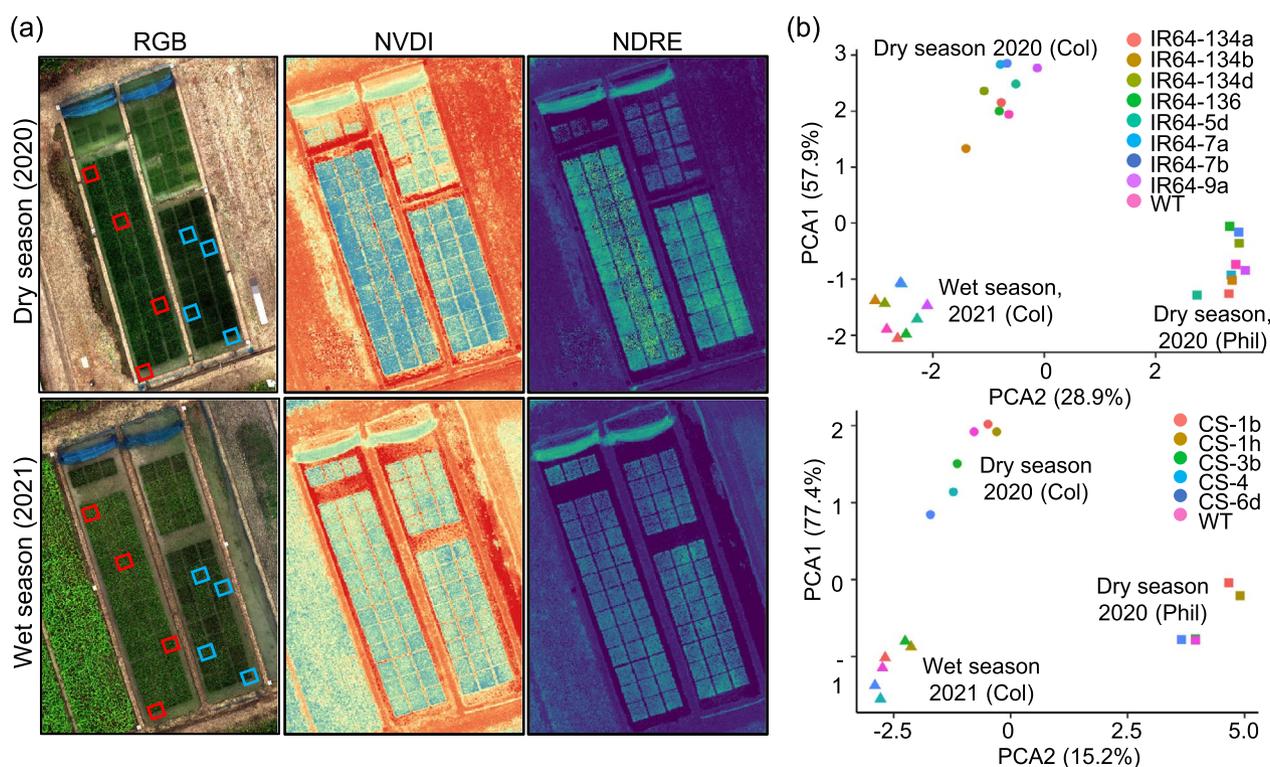


FIGURE 4 | Agronomic performance of GE'd IR64 and Ciherang-Sub1 lines is not significantly different from that of their parental WT plants. (a) Red-Green-Blue (RGB) and multispectral images of the EFPs conducted in Colombia in the dry season of 2020 and wet season of 2021. Boxed in red indicates WT IR64 plants, and boxed in blue indicates WT Ciherang-Sub1 plants. NDRE, Normalised Difference Red Edge Indices; NVDI, Normalised Difference Vegetation Indices. (b) Principal component analysis of WT and GE'd IR64 (top) and Ciherang-Sub1 (bottom) lines from three EFPs in two geographically separated fields and two harvest seasons. Colours indicate independent GE'd or WT lines, shapes indicate the location of EFP: Col, Colombia; Phil, The Philippines.

TABLE 2 | Agronomic traits of GE'd and WT plants of IR64 and Cihorang-Sub1 cultivars during the 2021 wet season in Palmira, Valle del Cauca, Colombia.

Genotype	100% flowering (DAP)		Bulk grain weight (g)	Dry biomass (g)	Grain yield (kg/m ²)	Harvest Index	Panicle length (cm)	Panicle number	Plant height (cm)	Single plant grain yield (g)	Yield (Ton/Ha)
	75.5 ± 0.5	79.5 ± 8.5									
IR64-134c	75.5 ± 0.5	79.5 ± 8.5	1658.75 ± 77.7	30.78 ± 1	0.83 ± 0.04	0.52 ± 0.02	24.31 ± 0.1	12.18 ± 0.2	82.43 ± 0.9	35.49 ± 1.7	4.15 ± 0.3
IR64-7b	79.5 ± 8.5	71.75 ± 0.9	1561.5 ± 125.2	32.78 ± 4.1	0.98 ± 0.1	0.55 ± 0.06	24.34 ± 0.4	12.68 ± 0.7	76.79 ± 1.1	39.07 ± 4.1	3.9 ± 0.2
IR64-136a	71.75 ± 0.9	73.25 ± 0.75	1485 ± 64.9	35.24 ± 3.2	0.89 ± 0.13	0.5 ± 0.08	24.1 ± 0.3	14 ± 0.6	75.89 ± 1.1	33.62 ± 5.1	3.71 ± 0.4
IR64-134b	73.25 ± 0.75	70.25 ± 0.25	1458.5 ± 151.3	28.55 ± 5.18	0.84 ± 0.07	0.54 ± 0.05	23.98 ± 0.3	12.75 ± 1.2	76.71 ± 1.1	28.32 ± 2.4	3.65 ± 0.2
IR64-134a	71.25 ± 0.75	72 ± 0.8	1396.25 ± 71.6	32 ± 4.7	0.71 ± 0.04	0.48 ± 0.03	24.41 ± 0.4	11.14 ± 0.7	79.64 ± 0.9	29.44 ± 1.8	3.49 ± 0.3
IR64-9a	70.25 ± 0.25	74.75 ± 1.4	1440.25 ± 110.8	30.79 ± 3.3	0.74 ± 0.08	0.49 ± 0.05	25.05 ± 0.6	12.86 ± 0.8	82 ± 0.1	32.35 ± 3.4	3.6 ± 0.4
IR64-7a	72 ± 0.8	78 ± 0.7	1475 ± 173.8	28.77 ± 2.5	0.81 ± 0.05	0.53 ± 0.03	24.63 ± 0.3	11.43 ± 0.9	82.71 ± 1	29.56 ± 1.8	3.69 ± 0.1
IR64-5d	71 ± 1	79 ± 1.1	1517.5 ± 48.1	32.53 ± 1.6	0.74 ± 0.04	0.48 ± 0.02	24.91 ± 0.2	11.86 ± 0.1	81.54 ± 0.9	29.46 ± 1.5	3.79 ± 0.2
IR64 WT	74.75 ± 1.4	77.75 ± 1.2	1403 ± 70.3	30.35 ± 3	0.74 ± 0.01	0.49 ± 0.01	24.41 ± 0.3	12.61 ± 0.6	81.61 ± 1.3	33.21 ± 0.4	3.51 ± 0.2
CS-6d	78 ± 0.7	77 ± 0.6	1486.75 ± 62.8	24.25 ± 2.1	0.92 ± 0.09	0.6 ± 0.06	22.82 ± 0.4	9.5 ± 0.6	77.82 ± 1	36.61 ± 3.5	3.72 ± 0.2
CS-1b	79 ± 1.1	80 ± 0.4	1895.25 ± 55.8	24.56 ± 0.8	0.88 ± 0.02	0.59 ± 0.01	24.63 ± 0.2	9.39 ± 0.3	75 ± 2.7	35.31 ± 0.7	4.74 ± 0.1
CS-4	77.75 ± 1.2	76.75 ± 1.1	1500.25 ± 56.3	25 ± 1.5	0.9 ± 0.06	0.59 ± 0.04	23.29 ± 0.3	10.21 ± 0.4	74.04 ± 0.7	35.96 ± 2.4	3.75 ± 0.1
CS-3b	77 ± 0.6	76.75 ± 1.1	1664.25 ± 74.3	27 ± 1.1	0.99 ± 0.05	0.59 ± 0.03	23.63 ± 0.3	10.36 ± 0.2	77.86 ± 1.2	39.56 ± 2	4.16 ± 0.2
CS-1h	80 ± 0.4	76.75 ± 1.1	1631.25 ± 131.5	28.85 ± 2.9	0.94 ± 0.09	0.57 ± 0.05	24.47 ± 0.3	9.93 ± 0.6	81.79 ± 1.1	37.6 ± 3.6	4.08 ± 0.3
CS WT	76.75 ± 1.1	76.75 ± 1.1	1760 ± 78.6	24.28 ± 1.2	0.9 ± 0.05	0.6 ± 0.03	23.16 ± 0.4	10.04 ± 0.8	76.1 ± 1.1	36.2 ± 1.9	4.4 ± 0.2

Note: The values for each parameter represent least squares (LS) means compared to the respective wild-type variety. Cells highlighted in orange indicate values significantly lower than the wild type ($p < 0.05$), while those in green indicate values significantly higher.

3 | Discussion

We generated broad-spectrum resistance using multiplex editing of SWEET promoter elements in four rice varieties: Kitaake, IR64, Ciherang-Sub1, and Komboka. Here, we report the next critical steps, i.e., a comprehensive laboratory and field analysis of GE'd lines before transferring the GE'd lines to interested partners from countries that have established suitable regulations to exclude GE'd products from their national GMO regulations (www.editagenome.org; Buchholzer and Frommer 2023). We tested for the presence of foreign DNA and transgenes (used in the T0 generation to introduce edits) after segregation and subsequently evaluated the performance of GE'd lines in experimental field plots under government-approved permits. We provide a set of GE'd lines in three elite varieties that are compatible with the criteria set by government agencies in India and Kenya, required for import, further characterisation, and possibly variety registration under regulations that would consider these lines as equivalent to classical breeding lines.

3.1 | Segregation of Foreign DNA/Transgenes to Obtain GE'd Elite Lines That Can Be Exempted From Country-Specific GM Regulations

Stable *Agrobacterium*-mediated transformation was used to introduce editing enzymes and gRNAs; thus, the original transformants (before segregation) are classified as GMOs. T-DNA insertions can be eliminated by segregation. Since *Agrobacterium*-mediated transformation can lead to multiple T-DNA integrations, partial T-DNA insertions, as well as vector fragment insertions, it can be challenging to eliminate the T-DNA and to test for effective outcrossing (Norris et al. 2020). For instance, up to six Cas9 fragments were detected in the GE'd Komboka lines, requiring large populations and multiple generations to identify lines in which Cas9 was not detected anymore (Figure S7a). Until March 2025, 21 countries had implemented regulations that exempt GE'd plants that do not contain foreign DNA from respective national GM regulations, e.g., India and Kenya (www.editagenome.org). In 2023, the Department of Biotechnology (DBT), India, published guidelines with detailed requirements for the evaluation of the absence of 'exogenous' DNA from GE'd plants (National Academy of Agricultural Sciences 2020). For deregulation and import of GE'd seeds into India, the DBT guidelines require validation on the successful removal of the selection markers and overlap PCR on genomic DNA for vector sequences used for generating the DNA modification. In Kenya, GE'd crops are not subject to GMO regulation if the genetic modification was introduced using SDN1 or SDN2 approaches and if no foreign DNA or recombinant sequences are present in the final product. According to the Kenyan biosafety regulations (www.biosafetykenya.go.ke/biosafety-regulations/), applicants must submit a notification dossier to the National Biosafety Authority (NBA), including a scientific description of the editing process and sufficient evidence demonstrating the absence of foreign DNA. The data generated for import into India for overlap PCR and loss of herbicide resistance were accepted by the NBA in Kenya as sufficient for import.

DBT regulations stipulate that assays need to be performed side-by-side with a segregant line that retained the T-DNA as positive controls and untransformed parental lines as negative controls. Our analysis demonstrates that while the T-DNA containing parental line is herbicide-tolerant, the negative control and the segregant edited lines are sensitive. We note, however, that we had to analyse multiple parental lines since some of the T-DNA containing lines had lost herbicide tolerance likely due to gene silencing (Figure S5b). The absence of herbicide tolerance, therefore, cannot serve as reliable evidence for the absence of the T-DNA. The results also do not exclude potential vector backbone integration into the genome (De Buck et al. 2000; Lange et al. 2006). Overlap PCR with respective positive and negative controls did not detect vector fragments (Figures S2 and S7b), although together with the herbicide tolerance tests, the data comply with the regulations. However, trying to learn from the identification of a 3.9kb fragment derived from vector backbone in the genomes of GE'd hornless cows by WGS, previously undetected via diagnostic PCR (Carlson et al. 2016; Young et al. 2020), we intimate that fragments, either from partial T-DNA insertions or vector backbone, can only be detected if both primers anneal to the inserted fragment (Figure S6b). We also note that the design and testing of the primer set is challenging and costly, since there are currently no suitable tools for overlap primer design. Each primer pair has to be hand-curated and validated. Here, 49, 50, and 59 primer pairs were designed to amplify IRS1132 in Ciherang-Sub1, IR64, and pMUGW5, respectively, and the total cost, excluding failed primers, was €870. As an additional and complementary test, DNA gel blot using overlapping probes was performed. The assay could detect ≤ 0.25 copies of vector fragment per haploid genome in positive controls but not detected in GE'd lines (Figure S2). We can, however, not exclude that the sensitivity for detecting small insertions is sufficient. Since none of the assays performed were 100% reliable, we explored whether WGS might be a superior tool for detecting foreign DNA/transgenes.

3.2 | WGS for Monitoring Segregation and Detection of Transgenes

Initially, we used Illumina technology for WGS. The standard approach uses enzymatic fragmentation. Since enzymatic digestion of genomic DNA is not perfectly unspecific, it is conceivable that insertions will not be detected for a fragment that remains undigested. We therefore performed two independent sets of WGS using different DNA fragmentation methods (enzymatic digestion and mechanical shearing via sonication). Genome assembly and screening for vector backbone with the sequences from the two independent sets of WGS detected *Cas9* fragment insertion in IR64-7a and recurring singletons in IR64-134a, IR64-134d, CS-1b, CS-1h, and CS-6d. These lines were subsequently not considered as potential vector-free GE'd lines for export to target countries. The two protocols for short-read WGS also detected false positive hits for vector integration, potentially due to index-hopping or from contaminations, e.g., pollen from transgene-containing plants grown in parallel (Li et al. 2019). Reads mapping to the vector were underrepresented and not consistently detected in all biological samples, supporting an origin from index hopping or contamination. Thus, as an improved measure, we grew plants individually in sterile culture

and performed long-read WGS without barcoding, providing us with an ultra-sensitive alignment (seed alignment as short as 13 bp). Using long-read WGS, foreign DNA/transgene insertions could not be detected, further supporting the hypothesis that the low copy reads for the vector were derived from index hopping or contaminations (e.g., pollen from other plants in the greenhouse). It may be useful to evaluate both coverage depth and breadth (i.e., uniformity of genome sequence coverage) in WGS, since focusing only on the average sequencing depth can obscure blind spots in coverage. Among the tested library preparations, HiFi-PacBio has the highest and most consistent breadth coverage even at moderate sequencing depths (~30×). Standard Illumina sequencing and NanoBall sequencing of enzymatically digested or mechanically sheared DNA samples exhibited minimal differences in sequencing breadth coverage (Figure S8). One may thus consider using long-read PacBio HiFi sequencing (cost ~750€ per rice line) combined with sensitive mapping parameters that allow the detection of short sequence matches. Long-read sequencing provided the highest confidence for detecting T-DNA or vector backbone fragment insertions.

3.3 | SNPs and Off-Target Mutations

An additional advantage of WGS is the ability to screen for potential mutations (naturally occurring mutations that manifest in each generation, potential off-target and somaclonal mutations) and to predict whether such mutations might impact performance. Here, we set a threshold of five mismatches for a stringent screen for potential off-target mutations in the GE'd lines. If we allow three mismatches, the single base insertion in position 6314414 on chromosome 1 (in IR64-9a, IR64-136a, and IR64-7b) and position 15429591 in chromosome 5 (in IR64-136a) could be potential off-target mutations that might have been caused by Cas9 editing gRNA targeting the EBE for PthXo2 in three independent lines: IR64-9a, IR64-136a, and IR64-7b (Table S4). While it cannot be excluded that the SNP traced to intergenic regions may have epigenetic effects, it is unlikely to directly affect the function of a gene. We traced SNPs across five generations in three GE'd lineages by comparing their genomes to the WT genome (Figures 3 and S4). The sum of SNPs was in the range of 100–300 for GE'd lines in the respective elite backgrounds, which is crudely in the range of the expected natural mutation rate after 5–7 generations. Notably, more than 75% of the SNPs found in GE'd lines were intergenic and up to 24% were located in the intron, with less than 10% of the SNPs in exons. The presence of high-density SNPs in chromosomes 8, 11, and 12 where the EBEs were located across three generations of the GE'd line suggests linkage drag. However, severe traits linked to the edits introduced in the GE'd line were not observed. The lack of agronomic penalties from the GE'd line pointed to the non-detrimental effects of the exonic SNPs and indels (Figure 4, Tables S6–S10). GE'd lines that underwent and passed the approaches used for transgene analysis have been transferred to KALRO, Kenya for seed amplification, field trials in the presence of BB, and potential registration as a new breeding material (Figure 1). The transfer required the execution of a Material Transfer Agreement (MTA) and a plant health inspection by the German Chamber of Agriculture to provide a plant health certificate that declares the plants to be free from pests (www.landwirtschaftskammer.de/landwirtschaft/pflanzenschutz/psd/index.htm). Once suitable agreements are in place for ICAR, the edited lines

that underwent and passed the transgene analysis protocol could be transferred to India. Transfer of the lines to farmers may require freedom to operate (evaluation of possible infringement of intellectual property). Transfer of the lines was strictly for humanitarian purposes without requiring royalties from KALRO or ICAR. The lines can also be transferred to other countries that have appropriate regulations for GE'd crops after the execution of the MTA.

3.4 | Results for GE'd Lines in Multi-Season and Multi-Location EFPs

To evaluate whether the promoter edits, off-target or other unintended mutations might affect the performance of the GE'd IR64 and Ciherang-Sub1 lines, we performed multi-year multi-site EFPs in The Philippines and Colombia. Notably, EFTs in the absence of disease allowed us to evaluate whether the editing process might cause performance defects, evaluations that can not be done under disease pressure. EFP conditions tested included wet and dry harvest seasons from 2020 to 2023 and different nitrogen regimes. Multispectral imaging and individual agronomic traits analyses indicate that the GE'd lines behaved similarly to unedited parental controls; in particular, the most important trait, i.e., yield, was unaffected. It may be noteworthy that GE'd IR64 lines exhibited a reduced 1000-grain weight (reduction of 4%–8%) compared to unedited IR64 plants, while the 1000-grain weight of GE'd Ciherang-Sub1 lines was not impacted. Similarly, the 1000-grain weight of other GE'd Zhonghua elite rice variety was not impacted (Zeng et al. 2020). Compensatory increases in fertility and/or spikelet number may explain why total yield was unaffected. Since the edits introduced in the EBEs of GE'd Ciherang-Sub1 and IR64 lines were comparable, it is unlikely that the effect on 1000-grain weight in the IR64 lines was due to the edits in the promoters or off-target mutations (Tables S6–S10). The GE'd IR64 lines tested in this study were independent transformants; thus, second site mutations are also not likely causes. The most parsimonious explanation might be that the IR64 control acquired natural mutation(s) that impacted 1000-grain weight. Since total yield, which is the trait breeders and farmers are interested in, was unchanged, the reduced 1000-grain weight was not considered a concern. In sum, large-scale multi-location and multi-season EFPs with different nitrogen regimes validated that GE'd IR64 and Ciherang-Sub1 cultivars were not affected in agronomic performance compared to respective WT plants. Of note, besides EFPs in disease-free conditions as performed here, the lines will be evaluated in field trials in the target countries after import of transgene-free seeds, a process that has already been initiated in Kenya and will be pursued in India after import. Since Kauffman's clipping assay has been widely used over many decades by breeders as highly predictive of resistance/susceptibility in the field [as discussed in (Loo et al. 2024)], and since the resistance generated by multiplex editing showed resistance against all 96 strains from our worldwide collection, one may predict that the edited lines will also be resistant in the multi-site multi-season field trials (Oliva et al. 2019).

3.5 | Summary

After careful analysis for transgenes/foreign DNA and multi-location multi-year field trials, we identified nine elite lines belonging to three elite varieties suitable for transfer to

rice-growing countries that have instituted regulations for GE'd crops.

4 | Materials and Methods

4.1 | DNA Gel Blots for Detecting Transgenes

Probes were generated using primers listed in Table S1 and labelled using the PCR DIG probe synthesis kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. For each sample, 1 mg of genomic DNA was digested with 20 units *BmiI* at 37°C for 1 h followed by enzyme inactivation at 65°C for 20 min. Digested genomic DNA was separated via electrophoresis at 90 V for 90 min using 1% (w/v) UltraPure Agarose (Thermo Fischer Scientific, Waltham, Massachusetts, USA) before blotting onto the nylon membrane overnight. DNA crosslinking was performed using UV Stratalyner followed by probe hybridization using the DIG Easy Hybridization kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. The hybridised membrane was washed with low stringency buffer for 5 min at room temperature followed by washing with high stringency buffer for 15 min at 70°C. The washed membrane was treated with a blocking solution for 30 min with shaking at 60 rpm and an Anti-Digoxigenin-AP antibody for 30 min with shaking at 60 rpm. The labelled membrane was washed and treated with CDP-Star (Roche, Mannheim, Germany) for signal detection using Imager's LAS4000 at high-binning sensitivity. For analysing the sensitivity of the probes, a 359 bp Cpf1 probe labelled with digoxigenin (DIG) was used for detecting Cpf1 fragments in genomic DNA of WT Komboka spiked with 0.1–3.5 copies of digested pMUGW5 (Figure S7c, Table S1). The copy number of T-DNA insertions was calculated by the molar ratio in the genome of Komboka according to <https://brcf.medicine.umich.edu/cores/transgenic-animal-model/training-education/protocols/spike/>.

4.2 | Whole Genome Sequencing of GE'd Lines

The genomes of the GE'd lines were subjected to nanoball sequencing using the T7 platform (DNBSEQ-T7 PE150; BGI Genomics; Drmanac et al. 2010). 100 mg tissues were harvested from young leaves of plants grown in greenhouse (GE'd IR64 and Ciherang-Sub1 lines) or MS medium under sterile conditions (GE'd Komboka lines) to avoid cross-contamination from parental lines. Leaves from two plants were pooled as one sample per GE'd line. DNA was extracted from selected lines using DNeasy Plant Mini Kits (Qiagen, Hilde, Germany). The DNA quality was evaluated using a Qubit 4 fluorimeter (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Two independent DNA libraries were prepared, i.e., enzymatic and mechanical fragmentation for each line. For enzymatic fragmentation, libraries were prepared with NEBNext Ultra II FS DNA Library Prep with Sample Purification Beads. For mechanical fragmentation, DNA was sonicated with an LE220 Focused Ultrasonicator (Covaris LLC, US) using default parameters for 400–500 bp fragments. DNA fragment size was validated with a DNF-474 high sensitivity NGS fragment analysis kit. An NEBNext Ultra II DNA Library Prep Kit (Germany) for Illumina was used to generate the

libraries. Barcoding was done using NEBNext Multiplex Oligos for Illumina (96 Unique Dual Index Primer Pairs).

4.3 | WGS for Reference Genome From IR64, Ciherang-Sub1, and Komboka

While assembled reference genomes for the rice variety IR64 are readily available (GCA_009914875 and GCA_011764405 submitted by the Arizona State University and Institute of Crop Science-NARO, respectively), 142456 SNPs were identified between the two IR64 assemblies. To ensure accurate analysis for the EBE edited lines, we performed whole-genome sequencing and assembled three reference genomes, i.e., IR64, Ciherang-Sub1, and Komboka. The genomes were sequenced using Oxford Nanopore Technology (ONT) R9 and R10. Read calling was performed with the dorado tool V0.5 (<https://github.com/nanoporetech/dorado>). Adapters were removed using porechop V0.2 (<https://github.com/rrwick/Porechop>). Filtering and cleaning processes were conducted with filtlong V0.2 (<https://github.com/rrwick/Filtlong>). Filtered and corrected long reads were assembled with the NECAT V0.01 assembler (Chen, Nie, et al. 2021). The resulting contigs were inspected and corrected using the Inspector tool (Chen, Zhang, et al. 2021), then polished with high-quality (PHRED score > 30) Illumina reads using NextPolish v1.4 (<https://github.com/Nextomics/NextPolish>). Finally, the contigs were ordered and scaffolded with the Ragtag tool V2.1 (Alonge et al. 2022), guided by the indica cultivar IR64 (IRRI) assembly (GenBank accession GCA_009914875). This process yielded a pseudomolecule representing the chromosomes of IR64, Ciherang-Sub1, and Komboka. The assembly for IR64 yielded 1508 contigs with a total length of 368972140 bp and an N_{50} of 415664 bp, whereas 124 contigs with a total length of 393489778 bp and N_{50} of 13319595 bp for Ciherang-sub-1. We provide the full assembly of IR64, Ciherang-Sub1, and Komboka with the accession numbers in Table S11. Analysis of assembly completeness using BUSCO V1.0 (doi: <https://doi.org/10.1002/cpz1.323>. PMID: 34936221) indicated that more than 97% of the core genes in *Poa*les and between 95% and 98% from GCA_009914875 were recovered in the genome assembly of IR64 and Ciherang-Sub1, indicating good completeness (Wingett and Andrews 2018; Martin 2011). The pipeline for genome assembly is available (<https://zenodo.org/records/15716853>).

4.4 | Detection of Vector Integration in GE'd Lines

For each sequence sample, filtered and cleaned reads were mapped (aligned) against the vector sequence reference IRS11345 using BWA V0.7 (<https://doi.org/10.1093/bioinformatics/btp324>) for IR64 and Ciherang-Sub1 samples. The samtools V1.19 and picardTools packages V3.2 (Danecek et al. 2021; <https://broadinstitute.github.io/picard/>) were used to post-process aligned reads, sort for optical duplicates, sort by coordinate, and recall mapped reads. A seed sequence of 19 bases was used for screening of the vector in lines sequenced with the Illumina (short-read) platform and a seed sequence of 13 for genomes sequenced with the PacBio (long-read) platform. For HIFI Pacbio, Minimap V2.3 (<https://doi.org/10.1093/bioinformatics/bty191>) was used to align reads against the vector sequences. We also conducted sequence alignments of IRS1132 with the WT rice lines to identify identical regions between the vectors

and the WT lines that could cause confounding results due to the use of native rice promoters in gene expression constructs in the vectors. The alignment tool *blastn* (doi: <https://doi.org/10.1093/nar/25.17.3389>, PMID: 9254694; PMCID: PMC146917) was used to identify identical regions between vector sequences and rice genome sequences. The Artemis genome browser V18.2 (Carver et al. 2008) was used to: (1) inspect visually the vector sequence and the bam file containing the aligned reads; (2) load the areas between the vector and the genomes of the rice lines; and (3) inspect the covered regions of the vector and detect integration events. The pipeline for bioinformatics analysis of vector integration is available (<https://zenodo.org/records/15717347>).

4.5 | Detection of Off-Target Mutations in GE'd Lines

Detection of potential off-target mutation was conducted following previously reported protocols with some modifications (Liu et al. 2021). Briefly, reads were filtered by Phred score using TrimGalore V0.6. Alignments were performed using BWA V0.7 (<https://doi.org/10.1093/bioinformatics/btp324>), followed by an Indel Realignment with GATK3 V3.8. SNVs were detected using in-house reference genome assemblies with the sequence variation callers, Loqfreq V2.1, Varscan V2.3, Mutect2-GATK V4.0 and pindel V0.2 (Chen et al. 2020; Koboldt et al. 2009; Wilm et al. 2012; Ye et al. 2009). Consensus SNV calls were obtained using bedtools V2.30 (doi: <https://doi.org/10.1093/bioinformatics/btq033>), generating vcf files for the agreement of the three SNV callers. For HIFI-Pacbio reads, variant SNP and indel were aligned using pbmm2 (v1.17) (<https://github.com/PacificBiosciences/pbmm2>) and called using DeepVariant v1.9 (<https://github.com/google/deepvariant>). Potential off-target predictions were conducted using the Cas-OFFinder tool V2.4 (Bae et al. 2014), allowing for up to 5 mismatches. The Artemis genome browser V18.2 was employed to visually inspect predicted off-targets. Additionally, SNVs were recorded and annotated using SnpEff V5.2 (Cingolani et al. 2012).

4.6 | Overlap PCR

Genomic DNA from rice leaves and the plasmid vector propagated in *E. coli* were extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and NucleoSpin Plasmid mini kit for plasmid DNA (Machery Nagel, Düren, Germany), respectively, according to the manufacturer's instructions. PCR was performed with GoTaq G2 Green Master Mix (Promega, Madison, Wisconsin, USA) using primers with their corresponding annealing temperatures as listed in Table S1. The thermocycler was set to 98°C for 5 min; 25 cycles of 94°C for 30 s, annealing for 30 s, and 72°C for 30 s and one cycle of 72°C for 2 min. Electrophoresis was performed on 1% (w/v) agarose in TBA buffer. Sizes of amplicons were determined using a GeneRuler 1 kb Plus DNA ladder (Thermo Fischer Scientific, Waltham, Massachusetts, USA).

4.7 | Agronomic Traits From Multi-Environmental Experimental Field Plots

For India, IR64-21 (single plant selection from IR64, IRGC 117268) and Ciherang-Sub1 (IRGC 141193) served as both

control and parental lines for editing (Mackill and Khush 2018). The edited rice lines had been described (Eom et al. 2019; Oliva et al. 2019). For Kenya, Komboka (IRGC 141192) was used both as control and parental line for editing (Kitilu et al. 2019). The edited Komboka lines had been described (Schepler-Luu et al. 2023). To evaluate the agronomic performance of IR64 and Ciherang-Sub1 with edited EBE lines, three experimental field plots (EFPs) were conducted in two geographically separated environments (CIAT in Colombia and IRRI in the Philippines). EFTs in the Philippines were conducted in confined fields designated for transgenic materials. The fields were carefully managed to minimise pest and disease pressure, and no disease symptoms were observed during the experimental period. Two EFPs were conducted at CIAT over a dry season from June 2020 to October 2020 and a wet season from January 2021 to May 2021. A second set of EFPs was conducted at IRRI during the wet season from December 2020 to March 2021. Each EFP was performed using a complete randomised block design with three to four replicates. The plots were 3.96 m² with 20 cm between hills and 25 cm between rows. At the four-leaf stage (~21 days after sowing), each seedling was transplanted from soil trays onto a hill. In total, 120 plants were planted in each plot. The plants were treated with 60 kg/ha of nitrogen 2, 10, and 30 days post-transplant. Integrated agronomic practices were adopted to control pests and weeds throughout the trials. Seven plants per replicate from each EFP were sampled for agronomic performance. Plant height, tiller number, panicle number, and panicle length were quantified. Tissues were dried at 60°C for 4 days to obtain dry biomass. Grain moisture was adjusted to 14% prior to quantification of single plant grain weight and 100 or 1000 grain weight. Field grain yield (kg/ha) was calculated based on the yield from a single plant multiplied by plant density in the plot. All statistical analyses were performed using the Statistical Analysis System SAS 9.4 (SAS Institute Inc., Cary, NC, USA). ANOVA was performed using a linear mixed model and further verified with Dunnett's Multiple Comparison. The genotypes were treated as a fixed effect, while block (replication) and environment (season) were treated as random effects to account for spatial and temporal variability. Least squares means (LS-means) were calculated for each trait, and statistical significance among genotypes was evaluated using analysis of variance (ANOVA). Where significant differences were observed, Dunnett's multiple comparison test was applied to determine whether the edited lines significantly deviated from the wild-type (WT) control (Figure S9). For nitrogen use efficiency (NUE) evaluation, the experimental design followed Selvaraj et al. (2020), using a randomised complete block design with three replicates per treatment. Three nitrogen application rates were tested: 0 kg N ha⁻¹ (0%), 90 kg N ha⁻¹ (50%), and 180 kg N ha⁻¹ (100%). Grain yield and related traits under different nitrogen treatments were analysed using the same linear mixed-effects model framework, with genotype, nitrogen level, and their interaction as fixed effects and block as a random effect. Drone-based crop monitoring in paddy field experiments was conducted using vegetation indices (VIs) according to protocols by Selvaraj et al. (2020). A DJI MATRICE 600 drone equipped with a MicaSense Altum multispectral camera (MicaSense, Seattle, Washington, USA) captured aerial images. Ground control points (GCPs) were placed at the edges of each plot for accurate georeferencing. Regions of interest (ROIs) were identified and marked using QGIS software (<https://www.qgis.org>).

The vegetation indices, NDVI and NDRE, were analysed using the CIAT's Pheno-i platform (<http://pheno-i.ciat.cgiar.org>). NDVI and NDRE were measured in one- to two-week intervals between 37 and 89 days after planting (DAP).

Author Contributions

W.B.F., B.Y., F.F.W., B.S. developed the concept. P.C., M.S., J.T. performed EFPs; J.T., T.H., V.S.-L. performed WGS analyses; E.P.I.L., M.St., B.K., S.V.G. performed transgene analysis; E.P.I.L., M.I.R.-S. compiled data; E.P.I.L., W.B.F., F.F.W., B.Y., M.B. have written the manuscript. All authors have given approval to the final version of the manuscript.

Acknowledgements

This work was supported in part by the Gates Foundation to HHU (Investment ID INV-008733 and Investment ID INV-063189), Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2048/1—project ID 390686111 (CEPLAS), and an Alexander von Humboldt Professorship (WF). We dedicate this work to the late Umar Singh, who dedicated his career to helping small-scale producers by developing improved varieties and who supported us tremendously as a member of the SAB of Healthy Crops.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All supplementary data for genomic section, including the pseudomolecule construction, genome assemblies, annotation files, and analysis code, can be accessed at <https://zenodo.org/records/15716853>; <https://zenodo.org/records/15717347>. Raw and meta data are available at doi. Accession numbers for whole genome sequences of WT and GE'd lines are provided in Table S11.

References

Alonge, M., L. Lebeigle, M. Kirsche, et al. 2022. "Automated Assembly Scaffolding Using RagTag Elevates a New Tomato System for High-Throughput Genome Editing." *Genome Biology* 23: 258.

Bae, S., J. Park, and J.-S. Kim. 2014. "Cas-OFFinder: A Fast and Versatile Algorithm That Searches for Potential Off-Target Sites of Cas9 RNA-Guided Endonucleases." *Bioinformatics* 30: 1473–1475.

Bezruczyk, M., J. Yang, J.-S. Eom, et al. 2018. "Sugar Flux and Signaling in Plant-Microbe Interactions." *Plant Journal* 93: 675–685.

Buchholzer, M., and W. B. Frommer. 2023. "An Increasing Number of Countries Regulate Genome Editing in Crops." *New Phytologist* 237: 12–15.

Carlson, D. F., C. A. Lancto, B. Zang, et al. 2016. "Production of Hornless Dairy Cattle From Genome-Edited Cell Lines." *Nature Biotechnology* 34: 479–481.

Carver, T., M. Berriman, A. Tivey, et al. 2008. "Artemis and ACT: Viewing, Annotating and Comparing Sequences Stored in a Relational Database." *Bioinformatics* 24: 2672–2676.

Chen, L. Q., B.-H. Hou, S. Lalonde, et al. 2010. "Sugar Transporters for Intercellular Exchange and Nutrition of Pathogens." *Nature* 468: 527–532.

Chen, L. Q., X.-Q. Qu, B.-H. Hou, et al. 2012. "Sucrose Efflux Mediated by SWEET Proteins as a Key Step for Phloem Transport." *Science* 335: 207–211.

Chen, Y., F. Nie, S.-Q. Xie, et al. 2021. "Efficient Assembly of Nanopore Reads via Highly Accurate and Intact Error Correction." *Nature Communications* 12: 60.

Chen, Y., Y. Zhang, A. Y. Wang, M. Gao, and Z. Chong. 2021. "Accurate Long-Read De Novo Assembly Evaluation With Inspector." *Genome Biology* 22: 312.

Chen, Z., Y. Yuan, X. Chen, et al. 2020. "Systematic Comparison of Somatic Variant Calling Performance Among Different Sequencing Depth and Mutation Frequency." *Scientific Reports* 10: 3501.

Cingolani, P., A. Platts, L. L. Wang, et al. 2012. "A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of *Drosophila melanogaster* Strain w1118; Iso-2; Iso-3." *Fly (Austin)* 6: 80–92.

Danecek, P., J. K. Bonfield, J. Liddle, et al. 2021. "Twelve Years of SAMtools and BCFtools." *GigaScience* 10: giab008.

De Buck, S., C. De Wilde, M. Van Montagu, and A. Depicker. 2000. "T-DNA Vector Backbone Sequences Are Frequently Integrated Into the Genome of Transgenic Plants Obtained by *Agrobacterium*-Mediated Transformation." *Molecular Breeding* 6: 459–468.

Drmanac, R., A. B. Sparks, M. J. Callow, et al. 2010. "Human Genome Sequencing Using Unchained Base Reads on Self-Assembling DNA Nanoarrays." *Science* 327, no. 5961: 78–81. <https://doi.org/10.1126/science.1181498>.

Eom, J.-S., D. Luo, G. Atienza-Grande, et al. 2019. "Diagnostic Kit for Rice Blight Resistance." *Nature Biotechnology* 37: 1372–1379.

Fan, S., Y. Zhang, S. Zhu, and L. Shen. 2024. "Plant RNA-Binding Proteins: Phase Separation Dynamics and Functional Mechanisms Underlying Plant Development and Stress Responses." *Molecular Plant* 17: 531–551.

Fu, Y.-W., X.-Y. Dai, W.-T. Wang, et al. 2021. "Dynamics and Competition of CRISPR-Cas9 Ribonucleoproteins and AAV Donor-Mediated NHEJ, MMEJ and HDR Editing." *Nucleic Acids Research* 49: 969–985.

Hsu, P. D., D. A. Scott, J. A. Weinstein, et al. 2013. "DNA Targeting Specificity of RNA-Guided Cas9 Nucleases." *Nature Biotechnology* 31: 827–832.

Huang, S., L. Tang, J. P. Hupy, Y. Wang, and G. Shao. 2021. "A Commentary Review on the Use of Normalized Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing." *Journal of Forest Research* 32: 1–6.

Hutin, M., F. Sabot, A. Ghesquière, R. Koebnik, and B. Szurek. 2015. "A Knowledge-Based Molecular Screen Uncovers a Broad-Spectrum OsSWEET14 Resistance Allele to Bacterial Blight From Wild Rice." *Plant Journal* 84: 694–703.

Kircher, M., S. Sawyer, and M. Meyer. 2012. "Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform." *Nucleic Acids Research* 40: e3.

Kitilu, M. J. F., A. M. S. Nyomora, and J. Charles. 2019. "Growth and Yield Performance of Selected Upland and Lowland Rainfed Rice Varieties Grown in Farmers and Researchers Managed Fields at Ifakara, Tanzania." *African Journal of Agricultural Research* 14: 197–208.

Knierim, E., B. Lucke, J. M. Schwarz, M. Schuelke, and D. Seelow. 2011. "Systematic Comparison of Three Methods for Fragmentation of Long-Range PCR Products for Next Generation Sequencing." *PLoS One* 6: e28240.

Koboldt, D. C., K. Chen, T. Wylie, et al. 2009. "VarScan: Variant Detection in Massively Parallel Sequencing of Individual and Pooled Samples." *Bioinformatics* 25: 2283–2285.

Lange, M., E. Vincze, M. G. Møller, and P. B. Holm. 2006. "Molecular Analysis of Transgene and Vector Backbone Integration Into the Barley

- Genome Following *Agrobacterium*-Mediated Transformation." *Plant Cell Reports* 25: 815–820.
- Li, Q., X. Zhao, W. Zhang, et al. 2019. "Reliable Multiplex Sequencing With Rare Index Mis-Assignment on DNB-Based NGS Platform." *BMC Genomics* 20: 215.
- Liu, G., Y. Qi, and T. Zhang. 2021. "Analysis of Off-Target Mutations in CRISPR-Edited Rice Plants Using Whole-Genome Sequencing." *Methods in Molecular Biology* 2238: 145–172.
- Liu, Q., M. Yuan, Y. Zhou, X. Li, J. Xiao, and S. Wang. 2011. "A Paralog of the MtN3/Saliva Family Recessively Confers Race-Specific Resistance to *Xanthomonas oryzae* in Rice." *Plant, Cell and Environment* 34: 1958–1969.
- Loo, E. P. I., B. Szurek, Y. Arra, et al. 2024. "Closing the Information Gap Between the Field and Scientific Literature for Improved Disease Management- With a Focus on Rice and Bacterial Blight." *Molecular Plant-Microbe Interactions* 38: 134–141.
- Mackill, D. J., and G. S. Khush. 2018. "IR64: A High-Quality and High-Yielding Mega Variety." *Rice* 11: 18.
- Martin, M. 2011. "Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads." *EMBnet.Journal* 17: 10–12.
- Miyao, A., M. Nakagome, T. Ohnuma, et al. 2012. "Molecular Spectrum of Somaclonal Variation in Regenerated Rice Revealed by Whole-Genome Sequencing." *Plant and Cell Physiology* 53: 256–264.
- Nameki, N., M. Takizawa, T. Suzuki, et al. 2022. "Structural Basis for the Interaction Between the First SURP Domain of the SF3A1 Subunit in U2 snRNP and the Human Splicing Factor SF1." *Protein Science* 31: e4437.
- National Academy of Agricultural Sciences. 2020. "Genome Edited Plants: Accelerating the Pace and Precision of Plant Breeding." Policy Brief No. 7. New Delhi.
- Ng'endo, M., M. Kinyua, L. Chebet, et al. 2022. "The Importance of Market Signals in Crop Varietal Development: Lessons From Komboka Rice Variety." *CABI Agriculture and Bioscience* 3: 57.
- Ni, Z., Y. Cao, X. Jin, et al. 2021. "Engineering Resistance to Bacterial Blight and Bacterial Leaf Streak in Rice." *Rice* 14: 38.
- Niño-Liu, D. O., P. C. Ronald, and A. J. Bogdanove. 2006. "*Xanthomonas oryzae* Pathovars: Model Pathogens of a Model Crop." *Molecular Plant Pathology* 7: 303–324.
- Norris, A. L., S. S. Lee, K. J. Greenlees, D. A. Tadesse, M. F. Miller, and H. A. Lombardi. 2020. "Template Plasmid Integration in Germline Genome-Edited Cattle." *Nature Biotechnology* 38: 163–164.
- Oliva, R., C. Ji, G. Atienza-Grande, et al. 2019. "Broad-Spectrum Resistance to Bacterial Blight in Rice Using Genome-Editing." *Nature Biotechnology* 37: 1344–1350.
- Pramudyawardani, E. F., H. Aswidinnoor, B. S. Purwoko, et al. 2020. "Accelerating the Development of New Submergence Tolerant Rice Varieties: The Case of Ciherang-Sub1 and PSB Rc18-Sub1." In *IOP Conference Series: Earth and Environmental Science*, 012032. IOP Publishing Ltd.
- Ramasamy, S. P., S. A. Rathnasamy, S. Manickam, et al. 2025. "Genome Editing for Improving Disease Resistance in Rice (*Oryza sativa* L.) - Scope, Achievements, Limitations and Future Needs." *Physiological and Molecular Plant Pathology* 138: 102640.
- Raveloson, H., R. Rabekijana, N. M. Rakotonanahary, et al. 2023. "First Report of Bacterial Leaf Blight Disease of Rice Caused by *Xanthomonas oryzae* pv. *oryzae* in Madagascar." *Plant Disease* 107: 2510.
- Rifhani, N. F., A. Apriana, A. Sisharmini, et al. 2023. "Construction of the CRISPR/Cas9 Module and Genetic Transformation of Aromatic Rice cv. Mentik Wangi for Developing Bacterial Leaf Blight Resistance." *Biodiversitas, Journal of Biological Diversity* 24: 3258–3268.
- Sarkar, N. K., U. Thapar, P. Kundnani, P. Panwar, and A. Grover. 2013. "Functional Relevance of J-Protein Family of Rice (*Oryza sativa*)." *Cell Stress and Chaperones* 18: 321–331.
- Sato, M., S. Hosoya, S. Yoshikawa, et al. 2019. "A Highly Flexible and Repeatable Genotyping Method for Aquaculture Studies Based on Target Amplicon Sequencing Using Next-Generation Sequencing Technology." *Scientific Reports* 9: 6904.
- Savary, S., L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson. 2019. "The Global Burden of Pathogens and Pests on Major Food Crops." *Nature Ecology and Evolution* 3: 430–439.
- Schepler-Luu, V., C. Sciallano, M. Stiebner, et al. 2023. "Genome Editing of an African Elite Rice Variety Confers Resistance Against Endemic and Emerging *Xanthomonas oryzae* pv. *oryzae* Strains." *eLife* 12: e84864.
- Selvaraj, M. G., M. Valderrama, D. Guzman, et al. 2020. "Machine Learning for High-Throughput Field Phenotyping and Image Processing Provides Insight into the Association of Above and Below-Ground Traits in Cassava (*Manihot esculenta* Crantz)." *Plant Methods* 16: 87.
- Streubel, J., C. Pesce, M. Hutin, R. Koebnik, J. Boch, and B. Szurek. 2013. "Five Phylogenetically Close Rice SWEET Genes Confer TAL Effector-Mediated Susceptibility to *Xanthomonas oryzae* pv. *oryzae*." *New Phytologist* 200: 808–819.
- Tang, X., G. Liu, J. Zhou, et al. 2018. "A Large-Scale Whole-Genome Sequencing Analysis Reveals Highly Specific Genome Editing by Both Cas9 and Cpf1 (Cas12a) Nucleases in Rice." *Genome Biology* 19: 84.
- Toledo, A. M. U., J. C. I. Ignacio, C. Casal, Z. J. Gonzaga, M. S. Mendiolo, and E. M. Septiningsih. 2015. "Development of Improved Ciherang-Sub1 Having Tolerance to Anaerobic Germination Conditions." *Plant Breeding and Biotechnology* 3: 77–87.
- Wang, H., M. L. Russa, and L. S. Qi. 2016. "CRISPR/Cas9 in Genome Editing and Beyond." *Annual Review of Biochemistry* 85: 227–264.
- Wilm, A., P. P. K. Aw, D. Bertrand, et al. 2012. "LoFreq: A Sequence-Quality Aware, Ultra-Sensitive Variant Caller for Uncovering Cell-Population Heterogeneity From High-Throughput Sequencing Datasets." *Nucleic Acids Research* 40: 11189–11201.
- Wingett, S. W., and S. Andrews. 2018. "FastQ Screen: A Tool for Multi-Genome Mapping and Quality Control." *F1000Research* 7: 1338.
- Wu, J., B. M. Kamanga, W. Zhang, Y. Xu, and L. Xu. 2022. "Research Progress of Aldehyde Oxidases in Plants." *PeerJ* 10: e13119.
- Wu, L. B., J. S. Eom, R. Isoda, et al. 2022. "OsSWEET11b, a Potential Sixth Leaf Blight Susceptibility Gene Involved in Sugar Transport-Dependent Male Fertility." *New Phytologist* 234: 975–989.
- Ye, K., M. H. Schulz, Q. Long, R. Apweiler, and Z. Ning. 2009. "Pindel: A Pattern Growth Approach to Detect Break Points of Large Deletions and Medium Sized Insertions From Paired-End Short Reads." *Bioinformatics* 25: 2865–2871.
- Young, A. E., T. A. Mansour, B. R. McNabb, et al. 2020. "Genomic and Phenotypic Analyses of Six Offspring of a Genome-Edited Hornless Bull." *Nature Biotechnology* 38: 225–232.
- Zaka, A., G. Grande, T. Coronejo, et al. 2018. "Natural Variations in the Promoter of OsSWEET13 and OsSWEET14 Expand the Range of Resistance Against *Xanthomonas oryzae* pv. *oryzae*." *PLoS One* 13: e0203711.
- Zeng, X., Y. Luo, N. T. Q. Vu, S. Shen, K. Xia, and M. Zhang. 2020. "CRISPR/Cas9-Mediated Mutation of OsSWEET14 in Rice cv. Zhonghua11 Confers Resistance to *Xanthomonas oryzae* pv. *oryzae* Without Yield Penalty." *BMC Plant Biology* 20: 313.

Zhou, J., Z. Peng, J. Long, et al. 2015. "Gene Targeting by the TAL Effector PthXo2 Reveals Cryptic Resistance Gene for Bacterial Blight of Rice." *Plant Journal* 82: 632–643.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** pbi70332-sup-0001-FigureS1.pdf. **Figure S2:** pbi70332-sup-0002-FigureS2.pdf. **Figure S3:** pbi70332-sup-0003-FigureS3.pdf. **Figure S4:** pbi70332-sup-0004-FigureS4.pdf. **Figure S5:** pbi70332-sup-0005-FigureS5.pdf. **Figure S6:** pbi70332-sup-0006-FigureS6.pdf. **Figure S7:** pbi70332-sup-0007-FigureS7.pdf. **Figure S8:** pbi70332-sup-0008-FigureS8.pdf. **Tables S1-S12:** pbi70332-sup-0001-Tables.xlsx.