Identifying genes associated with abiotic stress tolerance suitable for CRISPR/Cas9 editing in upland rice cultivars adapted to acid soils (version 2)

Luz Stella Barrero
Corporacion Colombiana de Investigacion Agropecuaria, AGROSAVIA, Km 14 via Mosquera, Colombia
Matthew Willmann
Plant Transformation Facility, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA. Current address: Pairwise, 807 East Main Street, Suite 4-100, Durham, NC 27701, USA
Eric Craft
USDA-ARS, Robert W. Holley Center, Tower Road, Ithaca, NY 14853-2901, USA
Kazi Manna-Akther
Plant Breeding & Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA
Sandra Harrington
Plant Breeding & Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA
Gina A. Garzon-Martinez
Corporacion Colombiana de Investigacion Agropecuaria, AGROSAVIA, Km 14 via Mosquera, Colombia
Raymond Glahn
USDA-ARS, Robert W. Holley Center, Tower Road, Ithaca, NY 14853-2901, USA
Miguel A. Piñeros
USDA-ARS, Robert W. Holley Center, Tower Road, Ithaca, NY 14853-2901, USA
Susan McCouch (srm4@cornell.edu)
Plant Breeding & Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

Research Article

Keywords: acid soils, Phosphorus use efficiency, CRISPR/Cas9, cultivar improvement

Posted Date: March 30th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1503282/v1

Identifying genes associated with abiotic stress tolerance suitable for CRISPR/Cas9 editing in upland rice cultivars adapted to acid soils

Luz Stella Barrero[1,3], Matthew Willmann[2,5], Eric Craft[4], Kazi Manna-Akther[3], Sandra Harrington[3], Gina A. Garzon-Martinez[1], Raymond Glahn[4], Miguel A. Piñeros[4], Susan McCouch[3*]

[1] Corporacion Colombiana de Investigacion Agropecuaria, AGROSAVIA, Km 14 via Mosquera, Colombia

[2] Plant Transformation Facility, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA.

[3] Plant Breeding & Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

[4] USDA-ARS, Robert W. Holley Center, Tower Road, Ithaca, NY 14853-2901, USA

[5] Current address: Pairwise, 807 East Main Street, Suite 4-100, Durham, NC 27701, USA

* Corresponding author

Abstract

Five genes of large phenotypic effect known to confer abiotic stress tolerance in rice were selected to characterize allelic variation in commercial Colombian *tropical japonica* upland rice cultivars adapted to drought-prone acid soil environments (cv. Llanura11 and Porvenir12). Allelic variants of the genes *ART1*, *DRO1*, *SUB1A, PSTOL1*, and *SPDT* were

characterized by PCR and/or Sanger sequencing in the two upland cultivars and compared with the Nipponbare and other reference genomes. Two genes were identified as possible targets for gene editing: *SUB1A*, *Submergence 1A*, to improve tolerance to flooding, and *SPDT* (*SULTR3;4*), *SULTR-like Phosphorus Distribution Transporter*, to improve phosphorus utilization efficiency and grain quality. Based on technical and regulatory considerations, *SPDT* was targeted for editing. The two upland cultivars were shown to carry the *SPDT* wild-type (non-desirable) allele based on sequencing, RNA expression, and phenotypic evaluations under hydroponic and greenhouse conditions. A gene deletion was designed using the CRISPR/*Cas9* system and specialized reagents were developed for *SPDT* editing, including vectors targeting the gene and a protoplast transfection transient assay. The desired edits were confirmed in protoplasts and serve as the basis for ongoing plant transformation experiments aiming to improve the P-use efficiency of upland rice grown in acidic soils.

Key words: acid soils, Phosphorus use efficiency, CRISPR/Cas9, cultivar improvement

Introduction

Abiotic stresses associated with rice production in non-irrigated systems include Phosphorus (P) and Nitrogen (P) deficiency, Aluminum (Al) toxicity, water deficit or surplus during the growing season, and temperature extremes associated with climate change. These abiotic challenges account for much of the gap between yield potential and actual crop productivity, particularly in developing countries (Mickelbart et al. 2015; Rao et al. 2016). These stresses often occur in combination; for example, heat and drought, or Al toxicity and Pi (P in the

47 inorganic form of orthophosphate) deficiency, or in succession, for example, flood followed
48 by drought (Mickelbart et al. 2015; Heuer et al. 2017). In highly acidic soils, Al toxicity is
49 associated with numerous other mineral deficiencies and drought, representing a primary
50 factor reducing crop yields (Kochian et al. 2015). Maintaining yield and nutritional quality
51 in variable environments will require intensive research efforts in crop breeding and
52 management (Chaturvedi et al. 2017; Dhankher and Foyer 2018). Rice (*Oryza sativa* L.) is a
53 staple food for more than 50% of the world´s population (Roy et al. 2021) and integrated
54 models of climate change and crop production predict significant yield reductions in the near
55 future, with severe consequences for global food security (Mickelbart et al. 2015; Ray et al.
56 2019).
57
58 The Llanos ecoregion and the foothills of the Colombian Orinoco are part of the great
59 savanna biome comprising half of the African continent, large parts of South America and
60 Australia, and smaller parts of North America and Eurasia (Rincón et al. 2014). The
61 Colombian savanna represents 17 million ha, with approximately 4.6 million ha localized in
62 the flat Altillanura, considered one of the largest land reserves for the expansion of crop
63 production (Amezquita et al. 2013). However, this region is characterized by fragile
64 ecosystems with low fertility, drought-prone, and highly acidic soils, prompting the
65 development of integrated strategies to improve soil productivity and generate adapted
66 cultivars (Amezquita et al. 2013). Among the rice cultivars recently released for this region,
67 two upland *tropical japonica* varieties stand out, CORPOICA Llanura11, also known as
68 Cirad 409 (Guimarães et al. 2020) and CORPOICA Porvenir12 (originated from Line 23),
69 derived from a recurrent selection population PTC11 (Grenier et al. 2015). Llanura11 is
70 prominent because of its importance to farmers and its agro-industrial (Krispies, brewery)

and direct consumption market, while Porvenir12 carries resistance to the main disease constraint *Pyricularia* (Colombian Agricultural Institute -ICA- resolutions No. 002581 of 2011 and No. 00024795 of 2018; Saito et al. 2018). Currently, there is great demand for new rice varieties with increased yield and efficient use of water and nutrients, particularly P. A feasible way to tackle this demand is through genetic improvement of target traits in the already adapted cultivars.

Genome editing using the CRISPR/*Cas* system opens opportunities to address the demand for new, improved climate-resilient cultivars (Chen et al. 2019). The system has been widely used to generate precise changes in the genomes of many organisms and is rapidly evolving as an accurate and predictable breeding technique (Chen et al. 2019; Graham et al. 2020). Applications in plant improvement are numerous and include, but are not limited to, enhancing pathogen resistance, drought tolerance, and food product quality (Wang et al. 2014; Chilcoat et al. 2017; Ansari et al. 2020). In rice, genome editing using various nucleases has demonstrated improved disease resistance, crop yield and quality, abiotic stress tolerance, herbicide tolerance, and male sterility (Zafar et al. 2020). Meanwhile, the legislation and regulation of gene-edited crops and agricultural products are evolving rapidly, with several countries classifying transgene-free genome-edited products under the 'non-genetically modified organism' (non-GMO) status, except for the EU and New Zealand (Schmidt et al. 2020).

This facilitates the use of genome editing during cultivar development and release of transgene-free edited varieties for use in farmers´ fields. In rice, detailed molecular cloning and characterization of genes conferring tolerance to numerous abiotic stresses provides the

95 foundation for genetic improvement of modern cultivars (Mickelbart et al. 2015). Loci
96 harboring genes of known function with large phenotypic effects represent potential targets
97 for gene editing. Among these are genes involved in tolerance to drought, *Deeper Rooting1*
98 (*DRO1*) (Uga et al. 2013), submergence, *Submergence 1A* (*SUB1A*) (Xu et al. 2006; Bailey-
99 Serres et al. 2010), aluminum toxicity, *Aluminum Resistant Transcription Factor*
100 (*ART1*) (Yamaji et al. 2009; Famoso et al. 2011; Arbelaez et al. 2017), P-deficiency,
101 *Phosphorus Starvation Tolerance1* (*PSTOL1*) (Gamuyao et al. 2012; Pariasca-Tanaka et al.
102 2014), *SULTR-like Phosphorus Distribution Transporter* (*SPDT / SULTR3;4*) (Yamaji et al.
103 2017) and other traits (Roy et al. 2021). Indeed, rice is a model crop for deploying genes
104 through new breeding technologies due to its amenability for transformation and its extensive
105 genetic and genomic resources (McCouch et al. 2016; Toki et al. 2006; Wang et al. 2018).
106 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated nucleases
107 such as *Cas9* and *Cas12a Cpf1*, base editors, protoplast transfection methods, as well as
108 transgenic and DNA-free transformation protocols are all part of the rice gene-editing
109 toolbox (Shan et al. 2014; Wang et al. 2017; Yin et al. 2017; Zafar et al. 2020).
110
111 In the present study, we provide the foundation for genome editing in the upland cultivars,
112 Llanura11 and Porvenir12, as well in the Nipponbare reference genome, with the goal to
113 enhance abiotic stress tolerance. Specifically, the study aims to: i) characterize allelic
114 variation in the genes *DRO1, SUB1A*, *ART1*, *PSTOL1,* and *SPDT* via PCR and sequencing
115 to determine whether the upland varieties carried alleles that might be appropriate targets for
116 gene editing; ii) evaluate gene expression and phenotypic variation associated with *SPDT*
117 alleles; iii) generate vectors for targeting *SPDT* using CRISPR/*Cas9* and determine whether
118 the gene can be deleted in protoplasts. We further examine quick and reliable methods for

evaluating *SPDT*-associated phenotypes and discuss varietal adaptation mechanisms as well as technical and regulatory implications of developing *SPDT*- edited rice varieties.

Materials and Methods

Plant material

The plant material used in the study is described in Table S1. Seed of the two cultivars Llanura11 and Porvenir12 were provided as advanced lines by the International Center for Tropical Agriculture (CIAT) to the Colombian Agricultural Research Corporation (AGROSAVIA), the national institution that evaluated and registered the two cultivars for the acid soil savannas of Colombia. The lines were developed in the CIAT- French Agricultural Research Centre for International Development (CIRAD) breeding program (Saito et al. 2018). The other materials correspond to rice reference cultivars or related species carrying desirable or undesirable alleles at each of the abiotic stress tolerance genes *ART1, DRO1, SUB1A*, *PSTOL1,* and *SPDT*. Homozygous *Tos17* mutants at *SPDT* were obtained from two lines (ND0047 and NE3502) provided by the Rice Genome Resource Center of the Institute of Crop Science, NARO (RGRC-NICS) in Japan following the provider recommendations (Miyao et al. 2003; Rice Tos17 Insertion Mutant Database 2013). The seeds were disinfected using sodium hypochlorite (2%) for 30 minutes and then washed with sterile water before germination. All seeds were amplified under greenhouse conditions at Cornell University following the procedures manual (Harrington 2016).

PCR and sequence of gene variants

Total genomic DNA was extracted from fresh foliar tissue using the DNeasy Plant Mini Kit (QIAGEN) as the template for PCR of gene regions. For *SPDT,* cDNA obtained from the shoot basal region (SBR, 0.5 cm above the union of root and shoot) as described below (qPCR) was also used as a template. The PCR was carried out in a 50µl reaction with 40 ng/µL of total DNA or cDNA. The Q5® High-Fidelity DNA Polymerase (NEB) was used following manufacturer instructions. PCR products were visualized on 1% agarose TAE gels. PCR products of the expected size were cleaned using the ExoSAP-IT enzyme (Thermo Fisher Scientific) and sequenced by Sanger. For *ART1* and *SPDT*, the complete coding sequence (CDS) was further sequenced. The *ART1* CDS was obtained using primers *ART1*-3F/4b, and the PCR product was cloned using the Zero Blunt® PCR Cloning Kit (Invitrogen). The *SPDT* CDS, as well as intronic, upstream and downstream regions, were obtained from cDNA and genomic templates using different primers (Figure S1). All primers used for PCR and sequencing are described in Table S2. Trimming of noisy sequences was followed by nucleotide or predicted amino acid sequence alignment using the Clustal W 1.83 against the GeneBank reference genes using the GeneStudio version 2.2 software.

SPDT phenotyping at early vegetative development

The *SPDT* phenotype was evaluated following the method described for Nipponbare (Yamaji et al. 2017) with modifications in the hydroponic nutrient solution and developmental stages targeted for evaluation. One-hundred seeds per cultivar (Llanura11, Porvenir12, and Nipponbare) and 40-50 seeds per *Tos-17 spdt* line (Table S1) were dried at 30°C for 3 days followed by 42-45°C for one day. All seed was treated with fungicide (Captan400, Trilex, and Allegiance), while the *Tos-17 spdt* seed was further treated with gibberellic acid 10 µM for 24 hours at room temperature in the dark to promote germination. Next, seeds were placed

166 on germination paper under Milli-Q water between 30/26°C day/night for 5 days. Up to eight
167 homogeneous seedlings per cultivar/line were transferred to Magnavaca nutrient solution
168 (Famoso et al. 2010; Table S3) at pH 5.6 and 0 μM phosphorous (P) in 30-liter containers.
169 Ten days after germination, the seedlings were transplanted to the same fresh solution, this
170 time with 90μM P in 180-liter containers, and growth in chambers at 30/26°C - day/night -
171 12 hours each, light intensity of 450 mmol photons m^{-2} s^{-1} and continuous aeration. The
172 accuracy of the final concentration of elements in the nutrient solution was confirmed by
173 Inductively Coupled Plasma Emission Spectroscopy (ICP-ES). Seedlings were sampled at 5
174 and 8-leaf stages and were first washed with a 5 mM $CaCl_2$ solution and then MilliQ water.
175 Samples were taken from roots (R), shoot basal region (SBR), the old leaf (L2), and the
176 youngest leaf (L5 or L8) at the corresponding leaf stage (Figure S2). Follow-up of
177 phenological stages and leaf tracking was carried out as previously reported (Counce et al.
178 2000; Xing and Zhang 2010) for each cultivar (Figure S3). All samples were oven-dried at
179 65-70°C for a week, followed by determination of P by ICP-ES with 3 technical replicates.
180
181 **Real-time quantitative PCR (RT-qPCR) for *SPDT* at early vegetative development**
182 cDNA was obtained from the SBR of Llanura11, Porvenir12, and Nipponbare seedlings
183 grown under hydroponic conditions as described above. Briefly, 120 seeds per cultivar were
184 treated and germinated in water for 6 days, and 40 homogenous seedlings per cultivar were
185 transferred to Magnavaca nutrient solution, pH 5.6 and 0 μM P in 30-liter containers. Twelve
186 days after germination, half of the seedlings were transferred to a fresh solution without
187 phosphorus (0 μM P, treatment) and the other half to a solution containing P (90 μM P,
188 control experiment) for one week according to Yamaji et al. (2017). On day 19 after
189 germination, SBR samples were collected, frozen in liquid N2, and stored at -80°C. Total

190 RNA extraction, on-column DNA digestion, and synthesis of single-stranded cDNA were
191 performed using the RNeasy Plant Mini Kit (QIAGEN), the RNase-Free DNase Set
192 (QIAGEN), and the High-Capacity RNA to cDNA kit (Applied Biosystems), respectively,
193 following the manufacturer instructions. RT-qPCR was carried out using the Power
194 SybrGreen Mastermix (Applied Biosystems) and the primers listed in Table S2 according to
195 the manufacturer's instructions. Relative expression levels (at the 90µM P control) were
196 normalized against the endogenous actin control and calculated using the ΔΔCt method with
197 the RQ-Manager software (Life Technologies). Four independent biological replicates with
198 three technical replicates each were analyzed. Each biological sample was represented by 20
199 SBRs that were pooled to obtain the amount of RNA required for cDNA synthesis.
200
201 ***SPDT* phenotyping at seed maturity**
202 Seeds were dried (as described above) and sown in germination trays. Seedlings were
203 transplanted to pots and randomly distributed in paddy tanks under greenhouse conditions
204 (Harrington 2016). Plants were grown to maturity, and samples were taken from the flag leaf
205 blade (FL) (of the main culm/stem), three green leaf blades (GL) (of culms harboring panicles
206 from three different positions), and from the panicles. Fifteen to 20 grams of dried seeds per
207 sample were de-husked using the TR200 Rice Husker (Kett) and polished using the Pearlest
208 Grain Polisher (Kett) to obtain brown seed (BS) and polished seed (PS). All samples were
209 homogenized and subjected to ICP-ES analyses to determine P, Fe, and Zn. The PS was
210 further analyzed for P and phytic acid concentrations using the Megazyme Phytic Acid Assay
211 Kit (McKie and McCleary 2016). Fe bioavailability was determined by Fe uptake in Caco-
212 2 cells by ferritin formation in response to exposure to a digest of the PS (Glahn et al. 2002).
213 The molar ratios of phytic acid to Fe and Zn were calculated as the micromoles of phytic acid

per gram of PS divided by the micromoles of the element per gram of PS. Six biological replicates per genotype with three technical replicates each were analyzed.

Statistical analyses

For hydroponic and greenhouse experiments, the response variables (concentration of P and other elements) were analyzed using a linear model. Residual analysis was carried out to ensure that model assumptions related to normal distribution and homogeneity of variances were met. Multiple comparison of means was performed using the Tukey method. Means were declared significantly different if p values were >0.05 or 0.01. The RStudio package Version 3.6.2 was used for the analyses.

Selection and verification of guide RNAs (gRNAs)

The best gRNAs for targeted deletion of *SPDT* were selected using CRISPR-P v.2.0 (Bioinfo 2016) using the following parameters: Protospacer Adjacent Motif (PAM): NGG *Cas9* of *Streptococcus pyogenes*, nucleotide length: 20, genome: *Oryza sativa* (RAP-DB) against the *SPDT* gene and 5', 3' regions upstream and downstream of the CDS. The selection criteria included high on-target scores (> 0.5), low off-target scores (<0.5), low off-target numbers (<40), high number of MisMatches (MM) for off-targets (> 2MM), low number of off-targets in genes (<1). Using the primers described in Table S2, we generated nucleotide sequences of the gRNA target regions in the cultivars Llanura11 and Porvenir12 and compared them with corresponding regions of the Nipponbare reference genome.

Design and construction of CRISPR/Cas9 vectors

Base vectors from three modules A (pMOD_A1511), B (pMOD_B2112), and C (pMOD_C000) were used to target the *SPDT* gene. For B module vectors, a previous assembly of two gRNAs in pair-wise combinations driven by one PvUbi promoter was carried out. The A, B, C vectors were assembled in pTRANS_240, suitable for *Agrobacterium* plant transformation. We followed the protocols used for assembly in module B vectors and in the plant transformation vector described previously (Čermák et al. 2017). Primers used for vector assembly are described in Table S2. Constructs from positive *E. coli Top10* cells (Thermofisher) were confirmed by Sanger sequence.

Protoplast transfection and evaluation

Nipponbare protoplasts were obtained and transformed using polyethylene glycol (PEG) as previously described (Lowder et al. 2015) using the plant vector assembled with the best gRNA pair. Total genomic DNA was extracted by the Cetyl Trimethylammonium Bromide (CTAB) method (Fulton et al. 1995). PCR and Sanger sequencing was performed with the primers described in Table S2 using Q5® High-Fidelity DNA Polymerase (NEB) following the manufacturer's instructions to confirm the *SPDT* deletion.

Results

Allelic variation at abiotic stress tolerance genes

Five abiotic stress tolerance genes, *ART1*, *DRO1*, *SUB1*, *PSTOL1* and *SPDT,* previously demonstrated to have large phenotypic effects in rice under field conditions, were examined for allelic variation in two *tropical japonica* upland cultivars, Llanura11 and Porvenir12.

Four of the genes are associated with Quantitative trait loci (QTL) variation, while *SPDT* was identified using a *Tos17* mutant screen. Table 1 describes each gene and provides information about desirable and undesirable alleles harbored by reference cultivars. The PCR and sequence analysis of allelic variants in Llanura11 and Porvenir12 identified favorable alleles at *ART1*, *DRO1*, and *PSTOL1*, and null or undesirable alleles at *SUB1A* and *SPDT*.

At the *ART1* locus, PCR products of identical size (1.5 Kilobases (Kb) were generated using the primers described in Supplementary Table S2 in both *tropical japonica* Colombian cultivars as well as in the two *japonica* reference genomes, Nipponbare and Azucena. Sanger sequencing of the PCR products confirmed that the predicted amino acid sequences of the ART1 protein were identical in the two Colombian cultivars and 100% identical to Nipponbare, but they differed by a single non-synonymous substitution at position 436 compared to Azucena (Figure 1). When ART1 sequences from these four *japonica* genomes were compared to those from the *indica* (IR64) and *aus* (Kasalath) reference genomes, they differed by 4 amino acid (aa) substitutions, six small InDels, and an 8-aa InDel at position 387-395 (Figure 1), consistent with a previous report (Arbelaez et al. 2017).

At the *DRO1* locus, PCR was used to amplify exons 1 and 2 (667 base pairs (bp) and exons 3, 4 and 5 (516 bp) using primers described in Supplementary Table S2. Sequencing of the PCR products showed that the Colombian cultivars were again identical to each other and carried the desirable deep-rooting allele found in the *tropical japonica* reference variety, *Kinandang* Patong, and the *temperate japonica* reference variety, Nipponbare (Figure 2a). The shallow-rooting allele discovered in the *indica* variety IR64 differs from the deep-rooting allele by a one bp deletion in exon 4 at nucleotide position 943 from start of the 5′

284 untranslated region (UTR) in the transcribed region. This single bp change causes a
285 premature stop that produces a truncated protein (Uga et al. 2013).
286
287 At the *PSTOL1* locus, a single PCR product of 258 bp was strongly amplified in *Oryza*
288 *glaberrima* (cv CG14) and in the two Colombian cultivars using primers K46-
289 CGsp2fw/K46-1 (Figure 2b; Suppl Table S2), suggesting that they harbor the same P-
290 deficiency tolerant allele (Pariasca-Tanaka et al. 2014). Amplification of the 258 bp allele
291 was weak in Nipponbare, IR64, and Kasalath. Using primers K46-K1F/K46-Ksp3rv, a 342
292 bp amplicon was strongly amplified in the P- deficiency tolerant *aus* reference cultivar,
293 Kasalath, while amplification of the same amplicon was weak in Llanura11, Porvenir12 and
294 *O. glaberrima* and was completely lacking in the P-deficiency intolerant varieties, IR64 and
295 Nipponbare (Figure 2b) (Gamuyao et al. 2012; Pariasca-Tanaka et al. 2014). Sequencing of
296 the 342 bp amplicon confirmed its identity in Kasalath but amplification was too weak in
297 Llanura11, Porvenir12 and CG14 to allow sequencing. However, sequencing of the stronger
298 258 pb amplicon in these three materials indicated that the overlapping region in the 258 and
299 342 amplicons (from positions 451 to 526) between primers K46-CGsp2 (the forward primer
300 of the 258 pb amplicon) and K46-Ksp3 (the reverse primer of the 342 pb amplicon) harbors
301 three SNPs at each of these primers, confirming that the Colombian cultivars are identical to
302 the CG14 allele (Figure 2b).
303
304 At the *SUB1A* locus, a single strong PCR product was observed in the submergence-
305 tolerant *aus* variety FR13A (Xu et al. 2006) using primers 1F/1R (1015 bp product) or 1F
306 (GN15f)/1/2R (OFR1) (825 bp product) (Suppl. Table S2) while no amplification was
307 observed in the Nipponbare reference nor in either of the Colombian cultivars (Figure 2c).

This indicated that the *SUB1A-1* allele conferring submergence tolerance was missing from the *japonica* varieties, consistent with expectations (Xu et al. 2006; Singh et al. 2010; Bailey-Serres et al. 2010).

At the *SPDT (SULTR3;4)* locus, sequence analysis confirmed that the 670 amino acid protein predicted in the two Colombian cultivars was identical to that predicted in the reference variety, Nipponbare (Figure 3a; Yamaji et al. 2017). This wild-type allele differed from the allele observed in the P-use efficient *Tos17* mutant lines (ND0047 and NE3502) (Yamaji et al. 2017; Figure 3b). Considering technical and country-level regulatory feasibility for the development of gene editing products, we selected *SPDT* for further analyses.

SPDT phenotyping and RNA expression at early vegetative development

The *SPDT* phenotype was studied in the two Colombian cultivars (both carried the *SPDT* allele), Nipponbare (*SPDT*), and the two *Tos17* mutant lines (*spdt*) during the early stages of vegetative development grown in hydroponic conditions. At the 5 and 8-leaf stages, in the presence of sufficient Pi (90 μM), P concentrations significantly increased in older leaves (L2) and decreased in the SBR of the mutants as compared to the other three varieties (Figure 4a-b). These differences are consistent with previous observations for Nipponbare at the 8-leaf stage (Yamaji et al. 2017). No differences in P concentration were detected in the youngest leaves at the 5 and 8-leaf stages (Figure 4a-b). P concentrations were similar in Nipponbare, Llanura11 (Figure 4; Figure S4), and Porvenir12 (Figure S4). Expression of the *SPDT* gene was induced in the SBR of both Colombian cultivars and in Nipponbare under P-deficiency conditions. In the present study, a 4.4, 5.5, and 2.3-fold increase in *SPDT*

331 expression was detected in Llanura11, Porvenir12, and Nipponbare, respectively, when 0 μM
332 P treatments were compared to the P-sufficient (90 μM Pi) control condition (Figure 4c).
333
334 ***SPDT* phenotyping at seed maturity**
335 The *SPDT* phenotype was also evaluated at seed maturity under greenhouse conditions in the
336 Colombian cultivars using Nipponbare *(SPDT wt)* and *Tos17* mutants as references. Higher
337 P concentrations were observed in leaf tissue, specifically in the FL and GL, while lower P
338 concentrations were observed in the BS of the mutants compared to Nipponbare (reduced by
339 about 15-19%) (Figure 5a), consistent with the study (Yamaji et al. 2017). Llanura11 and
340 Porvenir12 showed P concentrations similar to Nipponbare for both straw and seed (Figure
341 5a). No significant differences were observed in P and phytic acid concentrations in the PS
342 for any lines (Figure 5b-c).
343
344 Fe concentrations in the BS of the mutant lines were higher than in Nipponbare or the
345 Colombian cultivars (Figure 5d), while there was no clear correspondence between Fe
346 concentration (Figure 5d) and Fe-bioavailability in the PS (Figure 5e). The latter measured
347 as Fe uptake based on the Caco-2 cell ferritin formation assay showed no significant
348 differences among any of the lines (Figure 5e). As with Fe, Zn concentrations were higher in
349 the mutants than in Nipponbare, in both the BS and the PS (Figure 5f). Fe and Zn
350 concentration differences trend between mutants and wild type in BS is in agreement with
351 the previous study (Yamaji et al. 2017). Zn concentrations were the highest in Llanura11 and
352 Porvenir12 (Figure 5f). The molar ratios of phytic acid:Fe and phytic acid:Zn of PS were not
353 significantly different among cultivars, with the phytic acid:Fe ranging from 60:1 in the

mutant to 71:1 in Nipponbare and the phytic acid to Zn ranging from 5:1 in Porvenir12 to 8:1 in Nipponbare (Figure S5).

Generation of specialized reagents for *SPDT* CRISPR/Cas9 editing

The *SPDT* genotype, phenotype, and RNA expression analyses all indicate that the Colombian cultivars carry the wild-type (less desirable) allele. Based on this conclusion, specialized reagents were designed to target the deletion of the *SULTR3;4* gene via CRISPR/*Cas9*. This included construction of editing vectors and confirmation of vector gRNA editing by protoplast transfection in Nipponbare (as the reference genome).

Using the genome sequence information from the Nipponbare as the reference genome, five gRNAs targeting the *SPDT* (*SULTR3;4*) gene were designed (1U, 1D, 3, 4, and 1F) (Table 2, Figure 6a). The gRNAs met the following criteria: high on-target scores (0.59 to 0.90), relatively low number of off-targets (5 to 19, except gRNA 3 with 40) preferentially located in non-genic regions, and if located in genes, low off-target scores (0.22 to 0.44) with a high number of mismatches (3-4). These selection criteria were selected to give a low probability of off-target editing by CRISPR/*Cas9* in rice (Tang et al. 2018). Any of the four possible pair-wise combinations of selected gRNAs (1U or 3 at the 5'end with 4 or 1D at the 3'-end) are expected to produce a deletion of the gene. The 1F gRNA designed to produce a frameshift mutation was considered as a back-up editing strategy, to be used only if any of the four possible pair-wise combinations (described above) failed to work and if 1F produced a desirable early stop codon.

The 1U-1D gRNA pair was selected for further work based on the low number of predicted off-target sites in the Nipponbare genome (12 and 5, respectively Table 2). Moreover, this pair is expected to delete the entire gene using a template-free approach with the 1U-gRNA targeting a sequence upstream of the transcriptional start site and the 1D-gRNA targeting a sequence downstream of the stop codon (Figure 6a). Most importantly, the gRNA regions in the target *SPDT* gene showed conserved nucleotide identity among the Colombian cultivars and the reference Nipponbare (Figure 6b). Deletion of the gene is advantageous because it minimizes country-level regulatory concerns (Chilcoat et al. 2017; Schmidt et al. 2020). Therefore, we next transfected protoplasts with a vector containing the 1U-1D gRNA pair. The PCR using protoplast genomic DNA produced a band of the expected size (0.37 Kb), indicating that the desired *SPDT* deletion (about 7.3 Kb) had occurred. This was confirmed by Sanger sequencing (Figure 6c).

Discussion

Abiotic stress alleles harbored by Colombian cultivars shed light on adaptation mechanisms and needs for further genetic improvement

The rice cultivars Llanura11 and Porvenir12 are cultivated as *tropical japonica* upland varieties in the savannas of Colombia which form part of the great Savanna biome (Amezquita et al. 2013; Rincón et al. 2014; Saito et al. 2018). These ecosystems are characterized by highly acidic, low fertility soils associated with Al toxicity, P deficiency, and drought alternating with high precipitation. To investigate the potential for targeted genetic improvement, we characterized alleles found in these varieties at five major-effect

400 genes known to confer Al toxicity tolerance (*ART1*), drought avoidance (*DRO1*),
401 submergence tolerance (*SUB1A*), P-acquisition (*PSTOL1*), and P-use efficiency (*SPDT*) via
402 PCR and sequencing. Desirable alleles were found at *ART1*, *DRO1,* and *PSTOL1,* suggesting
403 that adaptative mechanisms associated with physiological response to soil pH and ionic
404 composition, including root growth angle and overall root system architecture are conferred
405 by these loci in response to drought-prone acidic soil environments. In fact, Llanura11, also
406 known as Cirad 409 (Grenier C., personal communication), is characterized by a strong root
407 system under irrigated conditions and an increased total root volume under drought
408 conditions (Guimarães et al. 2020).

409 The fact that the Colombian cultivars carry Al tolerant alleles at *ART1*, a deep rooting allele
410 at *DRO1,* and a likely P-deficiency tolerant allele at *PSTOL1* is consistent with their
411 adaptation to acid soils. In rice varieties that are susceptible to Al toxicity (a ubiquitous
412 feature of low pH soils), root growth is severely limited (Famoso et al. 2011; Kochian et al.
413 2015), making it difficult for roots to explore the soil and reach deeper horizons to find water
414 (Uga et al. 2013) and/or to find an immobile nutrient such as Pi (Heuer et al. 2017). Thus,
415 favorable alleles at these three loci would be mutually complementary in promoting root
416 growth in acid soils.

417 The present study suggests that the desirable *PSTOL1* allele found in the Colombian cultivars
418 shares ancestry with the *PSTOL1* allele found in CG14, a cultivar of African rice, *O.*
419 *glaberrima*. This allele is different from the allele originally cloned from the *aus* variety
420 Kalalath (Gamuyao et al. 2012; Pariasca-Tanaka et al. 2014). It shows a 35 base-pair
421 substitution when aligned to the Kasalath allele, which facilitated the development of the
422 Kasalath (342-bp amplicon) and CG14 (258-bp amplicon) allele-specific markers (Pariasca-

Tanaka et al.2014). We observed that the 258-bp amplicon in the two Colombian cultivars and in *O. glaberrima* carried six nucleotide substitutions as previously reported (Figure 2, Pariasca-Tanaka et al.2014). The fact that we also observed a faint 342-bp band in these three cultivars and a faint 258-bp band in Kasalath may be explained by weak primer annealing due to nucleotide substitutions between the Kasalath and CG14 alleles. In fact, the K46-K1F / K46-Ksp3rv primer pair specific for the 342-bp Kasalath allele has two polymorphisms in the forward primer and three in the reverse primer, while the K46-CGsp2fw / K46-1 specific for the 258-bp *O. glaberrima* CG14 allele has three polymorphisms in the forward primer (Pariasca et al., 2004).

The Colombian cultivar Llanura 11 was developed by conventional cross breeding with various progenitors (IRAT 146/Oryzica Sabana10//CT10035-43-4-M-3) (ICA resolution No. 002581 of 2011). This line was developed as part of a recurrent selection population where identification of promising lines adapted to the Colombian savannah eventually gave rise to Porvenir12 (PCT-11\0\0\2,Bo\2\1>133-M-5-1-4-3-M) (ICA resolution No. 00024795 of 2018; Grenier et al. 2015). Tracing the CG14 allele back through the ancestral lines may clarify its origin in the Colombian cultivars. On the other hand, while it appears to be favorable, it remains to be determined whether the CG14 allele confers a similar P-uptake advantage as the Kasalath allele in the cultivars under investigation.

In contrast to *ART1*, *DRO1* and *PSTOL1*, adaptive alleles at *SUB1A* and *SPDT* were lacking, suggesting that targeted introgression and/or gene editing of favorable alleles at these loci could enhance the performance of Llanura11 and Porvenir12. The *SUB1* locus is known to harbor a gene family consisting of 2-3 tandemly arrayed members on chromosome 9 (Fukao

et al. 2006), with the *SUB1A* family member conferring tolerance to complete submergence in the *aus* variety FR13A for up to 14 days (Xu et al. 2006; Singh et al. 2020). It is also associated with enhanced tolerance to other abiotic stresses, including drought (Fukao et al. 2011; Bin Rahman and Zhang 2016). The *SUB1A* gene is absent in the Nipponbare reference genome and other *japonica* varieties due to an inversion and deletion (Singh et al. 2010). Consistent with this, we found that the gene was absent from the *tropical japonica* Llanura11 and Porvenir12 cultivars (Figure 2). Although the Colombian savannas are characterized by a strong dry season, there is also an important water surplus due to precipitation and even floods (Rincón et al. 2014), suggesting that marker-assisted introgression of *SUB1A* may be the most efficient way to generate climate-resilient varieties for these environments.

SPDT, a gene involved in internal P use efficiency (PUE), was cloned in cv. Nipponbare by transposon *Tos17* tagging (Yamaji et al. 2017). The loss-of-function desirable allele causes less P to be allocated to the grain (by about 20%) and consequently a potential increase of bioavailability of essential nutrients (Fe, Zn) without affecting seed germination and grain yield (Yamaji et al. 2017). At the same time, more P is assigned to the straw, which remains in the field after harvest and could be used as a fertilizer (Rose et al. 2013; Yamaji et al. 2017). The sequence of this gene locus in the upland cultivars under investigation showed that the wild-type allele was present. Thus, further cultivar improvement by knocking out *SPDT* offers an interesting possibility to generate high PUE and grain quality cultivars.

Given that directed DNA insertion is both technically challenging and presents regulatory hurdles (because it would be subjected to the restrictions imposed on genetically modified

469 organisms (GMOs) (Georges and Ray 2017; Wang et al. 2017), it would be advisable to
470 introgress the *SUB1A* gene into the two Colombian varieties using marker-assisted selection
471 (MAS). This approach has successfully introduced submergence tolerance into varieties
472 grown on thousands of hectares in flood-prone regions of Asia and Africa (Mickelbart et al.
473 2015). On the other hand, editing the *SPDT* gene through a targeted gene deletion offers a
474 practical solution, given that a DNA-free editing strategy or an *Agrobacterium* mediated
475 transformation editing strategy producing transgene-free products could be used, posing
476 fewer technical and regulatory constraints (Chilcoat et al. 2017; Georges and Ray 2017).
477
478 The SPDT phenotype was next examined in the Colombian cultivars and compared with
479 Nipponbare (SPDT *wt*) and the two *Tos17* mutant lines (*spdt*). This was done by evaluating
480 gene expression (using qPCR) and organ P concentration at the early vegetative and seed
481 maturity stages. For qPCR, the SBR was targeted for analysis based on a report showing that
482 this tissue produced the most significant *SPDT* differential expression using 0 and 90 μM P
483 treatments under hydroponic conditions (Yamaji et al. 2017). The induction of *SPDT*
484 expression under 0 μM P in the reference Nipponbare was 2.3-fold higher compared to the
485 control condition of 90 μM P (figure 4c). This represents a smaller induction of gene
486 expression as compared to that observed in the previous study for Nipponbare (6.5-fold
487 increase) (Yamaji et al. 2017). There was also no difference in P concentration in the
488 youngest leaf 8 at the 8-leaf stage (Figure 4b) as opposed to the reported in the previous study
489 (Yamaji et al. 2017), where lower P concentrations were detected in the *sdpt* mutants as
490 compared to wild type Nipponbare. The differences in gene expression and P concentrations
491 in the 8th leaf in the two studies are likely due to differences in the nutrient solutions used

(Table S3; Magnavaca in the present study vs. Kimura in Yamaji et al. (2017)) and/or small differences in experimental conditions.

Llanura11 and Porvenir12 were most similar to Nipponbare in terms of induced RNA expression under P-deficiency conditions, increased P concentration in older leaves during early vegetative development (5 and 8- leaf stages) and in flag and green leaves at seed maturity, and reduced P concentrations in BS (Figures 4-5). Based on these results, we concluded that the two Colombian cultivars carry the wild-type *SPDT* allele and are therefore likely to benefit from deletion of the *SULTR3;4* gene, mimicking the results observed in the ND0047 and NE3502 mutant lines.

Early phenotyping for fast screening of *spdt* genotypes

The *SPDT* phenotype was previously evaluated at the 8-leaf stage using Kimura B solution at half-strength (Yamaji et al. 2017). One objective of this study was to determine whether we could phenotypically differentiate the *SPDT* wild-type from the *spdt* knock-out lines at earlier stages of development, i.e., at the 5-leaf stage in plants growin in Magnavaca solution (Famoso et al. 2011), which has 7.5 times greater ionic strength than the Kimura B½ solution (Table S3). We also grew plants in a growth chamber rather than in a greenhouse which helped to minimize environmental variation, providing greater consistency of temperature and luminosity, regardless of the time of year.

Our results indicate that the earlier 5- leaf stage of development and higher ionic strength solution under growth chamber conditions can be used to reliably evaluate the phenotype under hydroponic conditions, saving time and resources (Figure 4). The original method

516 required 32-42 days to evaluate the Nipponbare phenotype at the 8- or 9-leaf stage (Yamaji
517 et al. 2017), while the modified method required 22 days for Nipponbare, and 25 days for the
518 Colombian varieties to reach the 5-leaf stage. Nipponbare takes 33 days and the Colombian
519 varieties 42 days to develop to the 8-leaf stage in the Magnavaca solution under growth
520 chamber conditions (Figure S3). The higher ionic strength of the growth solution, a high
521 volume ratio (about 10 liters of solution/plant), and shorter growth periods resulted in
522 minimal nutrient depletion from the solutions, as indicated by routinary ICP analysis of the
523 Magnavaca solution throughout the experiment. These modifications reduce water usage,
524 minimal replacement of microelements, and overall less labor. Therefore, this modified
525 protocol represents an improved technique to quickly screen for desirable *spdt* mutants at the
526 seedling stage.
527
528 **Low P brown seed as a potential source of bioavailable nutrients**
529 Varieties that transfer less P to developing grains offer environmental and nutritional benefits
530 (Rose et al. 2013). Phytic acid salt (phytate) is the main form of P in cereal grains where it
531 acts as an antinutrient, decreasing the bioavailability of other essential nutrients, such as Fe
532 and Zn (Perera et al., 2018). Phytate accumulates in the bran or aleurone layer of brown rice
533 seed where it can chelate Fe, K, and Ca, while Zn is found broadly distributed from the
534 aleurone layer to the inner endosperm, often bound to phytic acid but also found in another
535 storage form (Iwai et al. 2012). In this study, we investigated the hypothesis that reducing
536 this chelant might increase nutrient bioavailability and nutritional benefits by comparing P,
537 Fe, and Zn concentrations of BS and PS in *spdt* mutants and wild-type rice.
538

539 We confirmed that higher P concentrations were detected in BS in *SPDT* wild-type genotypes
540 (Nipponbare and Colombian cultivars) compared with *spdt* mutants, while no differences
541 were found in PS for either P or phytate concentrations, nor for Fe bioavailability (Figure 5).
542 This lack of correspondence between Fe concentration and Fe uptake determined by the caco-
543 2 cell bioavailability assay has been previously reported for brown rice (Glahn et al. 2002).
544
545 The molar ratios of Phytic acid:Fe were very high (Table S5), indicating that not much Fe in
546 the PS is contributing from a nutritional perspective. This is in agreement with the fact that
547 only 4.335 ng of ferritin/mg of protein was formed in the control Nipponbare with no
548 significant differences among genotypes (Figure 5e), therefore, not much Fe is likely to be
549 delivered from the PS. On the other hand, the molar ratios of phytate:Zn were much lower
550 (Figure S5). Typical Zn levels in polished rice are low (8-12 ppm) but there is a wide genetic
551 variability in brown (7.3 to 52.7 ppm) and polished (8 to 38 ppm) rice (Babu et al. 2020).
552 The primary inhibitor of Zn is phytate (Lönnerdal 2000) and more Zn, as observed in the
553 mutans and Colombian genotypes of this study (Figure 5f), usually means more absorbed Zn.
554 However, to our knowledge there is not a good method for assessing Zn bioavailability.
555
556 Brown or unmilled rice is known to have higher vitamin and mineral content compared to
557 milled rice (Muthayya et al. 2014). The present study confirms higher Fe and Zn
558 concentrations in BS compared to PS. However, many of the nutritional benefits of brown
559 rice cannot be not realized if nutrient bioavailability is compromised by high levels of phytate
560 in BS. Thus, low P *spdt* mutants that accumulate less phytate in the bran offer a possible
561 solution for breeders interested in developing cereal varieties with high concentrations of
562 bioavailable Fe and Zn as part of the human diet and also as a component of animal feed.

563 Low phytate grain would be particularly useful in non-ruminant livestock feeds, including
564 poultry, swine, and fish feeds, where P and inositol in phytic acid are generally not
565 bioavailable, and P-deficiency is a problem, while at the same time it may help reduce P
566 excretion that contributes to environmental problems, such as eutrophication of waterways
567 (Perera et al. 2018).
568
569 **Technical, regulatory, and societal implications of editing the *SPDT* gene**
570 P is an essential micronutrient, and P-deficiency is a major constraint for crop yield; thus, to
571 obtain high yields, regular applications of P- fertilizer are needed, but the supply of phosphate
572 rock is limited (Rose et al. 2013). P efficiency (PE), the capacity of plants to tolerate stress
573 caused by P-deficiency, can be achieved by increasing P acquisition efficiency (PAE) or PUE
574 traits (Rose et al. 2013; Heuer et al. 2017). The latter has prompted more recent attention and
575 has been envisaged that breeding crops by lowering grain P concentration is one approach to
576 increase PUE in cereal systems since less P removed from the field could lower fertilizer
577 requirements, saving production costs to farmers (Rose et al. 2013). For example, the low
578 phytic acid (*lpa*) recessive mutation *lpa1-1* in barley, harbored by the US commercial cultivar
579 "Herald", reduces total grain P by 10-20% and lowers phytate levels without a penalty on
580 subsequent crop yields (Bregitzer et al. 2007; Ye et al. 2011).
581
582 In rice, the genetic architecture of PUE was investigated using GWAS with a rice diversity
583 panel grown in a hydroponic system to ensure uniform access to P (Wissuwa et al. 2015).
584 The study identified loci associated with PUE on four chromosomes, with the chromosome
585 1 haplotype showing high priority based on association with candidate genes of potential
586 utility in plant breeding. Alternatively, novel variation can be generated by mutation breeding

(Rose et al. 2013; Heuer et al. 2017). In fact, transporters involved in delivering phosphate to developing seeds and synthesis of phytic acid have been discovered through mutation screens, identifying key genes belonging to the Sulfate Transporter (SULTR) family. One of them, *OsSULTR3;3,* was discovered in two *lpa* mutants developed to improve the nutritional value of rice grains. Disruption of *OsSULTR3;3* in these mutants leads to reduced concentrations of total grain P (19–28%) and grain phytate (35–45%) (Zhao et al. 2016), with the rice *SULTR3;3* gene is closely related to the barley sulfate transporter *LPA1* (Ye et al. 2011; Zhao et al. 2016). More recently, *OsSULTR3;4* (referred to as *SPDT* in this study), the first characterized transporter for inorganic P in this family, was shown to be involved in P allocation to rice grain, with the desired mutation reducing P (by 20%) as well as reducing phytate concentration. These discoveries point to a potential role in improving PUE in rice cropping systems such that disruption of the Pi transporter leads to retention of P in the straw, that it can be easily returned to the field after harvest (as a form of mulch) to help fertilize the next season's crop (Yamaji et al. 2017).

Here we aim to leverage knowledge about *SPDT* to enhance PUE in two upland Colombian cultivars using CRISPR-associated endonuclease as a targeted form of mutagenesis. This system has great precision and minimal risk of introducing *off-target* variation in the genome compared to historical mutagenesis techniques (Graham et al. 2020). Whether this will be the case for the Colombian cultivars awaits confirmation based on whole genome sequencing of edited vs wild type genotypes.

We generated specialized reagents to produce an *SPDT* deletion using CRISPR/Cas9 and tested our approach in the *temperate japonica* variety, Nipponbare before applying it to the

611 *tropical japonica* upland cultivars, Llanura11 and Porvenir12. As a first step, we confirmed
612 that editing occurs as expected in Nipponbare protoplasts. This is important because
613 Nipponbare is the reference genome from which the gRNAs were designed (as described in
614 the methods) and in which gene function was originally studied (Yamaji et al., 2017). We
615 further sequenced the target *SPDT* region in Llanura11 and Porvenir12 and confirmed that it
616 matched the Nipponbare reference (Figures 3a and 6b). Therefore, we conclude that it is
617 reasonable to expect that the desired edit in the *SPDT* gene may also occur in the Colombian
618 varieties using the selected gRNAs. Plant transformation in both the Colombian and the
619 Nipponbare varieties is currently underway using both transgenic (via *Agrobacterium*) and
620 non-transgenic (via Ribonucleoprotein complexes) methods.
621
622 The reagents included plant transformation vectors containing the *Cas9* and gRNAs to target
623 identical *SPDT* regions in the three cultivars. We confirmed the intended deletion of 7.3 Kb
624 at the *SPDT* locus by transfecting Nipponbare protoplasts with a vector targeting the 5' and
625 3' up- and downstream gene ends (Figure 6). These are essential tools for proof-of-concept
626 to determine whether *SPDT* can be subjected to a targeted deletion in cultivars of interest;
627 whether desired edits, if obtained in plants, produce the improved phenotype to enhance PUE
628 and grain quality under acid soils; and whether individual effects of P reduction in the grain-
629 based on deletions of *SULTR3;3* and *SULTR3;4* can be leveraged by pyramiding mutations
630 in both genes in a single cultivar.
631
632 Our work establishes the basis for targeted deletion of *SPDT* in upland rice cultivars of
633 interest to breeders. Gene deletions are preferred over gene insertions because they tend to
634 present fewer societal and regulatory concerns. Currently, genome edits involving deletions

635 are usually classified into the Site Directed Nuclease -1 (SDN-1) category, as long as there
636 is no addition of foreign DNA (Schmidt et al. 2020). SDN-1 edits follow the standards of
637 conventional mutagenesis with categorization based on the product and are considered non-
638 regulated as GMOs in most countries except the European Union and New Zealand (Schmidt
639 et al. 2020). The first *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) genome-edited resistant rice,
640 where promoter elements of the sucrose transporter genes *SWEET* were targeted through
641 CRISPR/Cas9 (Oliva et al. 2019) was declared transgene-free, non-regulated, and equivalent
642 to what could be accomplished with conventional breeding in Colombia and the USA
643 (Agdaily 2020). This opens the path for approval of other DNA-free editing products such as
644 those proposed in this study for the Colombian savannas.
645

646 ## Conclusions

647 Allelic variation at five major-effect abiotic stress tolerance genes (*ART1, DRO1, PSTOL1,*
648 *SUB1A, SPDT*) was studied in two upland Colombian rice cultivars (Llanura11 and
649 Porvenir12) and compared to a variety of reference genomes using PCR and sequence
650 analysis. The Colombian cultivars carried desirable alleles at three of the loci and non-
651 desirable alleles at *SUB1A* and *SPDT*, providing targets for further genetic improvement.
652 Based on technical and regulatory criteria, the *SPDT* gene, involved in P use efficiency (PUE)
653 and nutritional quality of the grain, was targeted for deletion via a gene knock-out strategy
654 using CRISPR/Cas9. An improved phenotyping pipeline was established to evaluate P
655 concentrations at the early vegetative stage and used in combination with evaluation of P
656 concentrations in the grain at seed maturity to establish the foundation for *SPDT* editing in
657 the Colombian cultivars. An editing vector containing a pair-wise combination of the most

suitable gRNAs was developed for targeted *SPDT* deletion and successfully used in experiments involving protoplasts to obtain the desired deletion. This study provides an essential foundation for applying CRISPR/*Cas9* editing to improve PUE and grain nutritional quality in upland *tropical japonica* cultivars.

References

Agdaily (2020) Regulators classify gene-edited rice varieties equivalent to conventional. In: AGDAILY. https://www.agdaily.com/crops/regulators-gene-edited-rice-varieties/. Accessed 15 May 2021

Amezquita E, Rao I, Rivera M, et al (2013) Sistemas agropastoriles: un enfoque integrado para el manejo sostenible de oxisoles de los Llanos Orientales de Colombia. Documento de trabajo CIAT No.223. 288p. International Center for Tropical Agriculture (CIAT), Cali, Colombia

Ansari WA, Chandanshive SU, Bhatt V, et al (2020) Genome editing in cereals: Approaches, applications and challenges. Int J Mol Sci 21:1–32

Arbelaez JD, Maron LG, Jobe TO, et al (2017) *ALUMINUM RESISTANCE TRANSCRIPTION FACTOR 1 (ART 1)* contributes to natural variation in aluminum resistance in diverse genetic backgrounds of rice (*O. sativa*). Plant Direct 1:1–19. https://doi.org/10.1002/pld3.14

Babu PM, Neeraja CN, Rathod S, et al (2020) Stable SNP Allele Associations With High Grain Zinc Content in Polished Rice (*Oryza sativa* L.) Identified Based on ddRAD Sequencing. Front Genet. https://doi.org/10.3389/fgene.2020.00763

Bailey-Serres J, Fukao T, Ronald P, et al (2010) Submergence Tolerant Rice: *SUB1*'s

681 Journey from Landrace to Modern Cultivar. Rice 3:138–147.
682 https://doi.org/10.1007/s12284-010-9048-5
683 Bin Rahman ANMR, Zhang J (2016) Flood and drought tolerance in rice: opposite but may
684 coexist. Food Energy Secur 5:76–88. https://doi.org/10.1002/fes3.79
685 Bioinfo (2016) CRISPR-P 2.0: An improved CRISPR/Cas9 tool for genome editing in
686 plants. http://crispr.hzau.edu.cn/CRISPR2/. Accessed 13 May 2021
687 Bregitzer P, Raboy V, Obert DE, et al (2007) Registration of 'Herald' Barley. Crop Sci
688 47:441–442. https://doi.org/10.2135/cropsci2006.07.0480
689 Čermák T, Curtin SJ, Gil-Humanes J, et al (2017) A Multipurpose Toolkit to Enable
690 Advanced Genome Engineering in Plants. Plant Cell 29:1196–1217.
691 https://doi.org/10.1105/tpc.16.00922
692 Chaturvedi AK, Bahuguna RN, Pal M, et al (2017) Elevated CO 2 and heat stress
693 interactions affect grain yield, quality and mineral nutrient composition in rice under
694 field conditions. F Crop Res 206:149–157. https://doi.org/10.1016/j.fcr.2017.02.018
695 Chen K, Wang Y, Zhang R, et al (2019) CRISPR/Cas Genome Editing and Precision Plant
696 Breeding in Agriculture. Annu Rev Plant Biol 70:667–697.
697 https://doi.org/10.1146/annurev-arplant-050718-100049
698 Chilcoat D, Liu Z-B, Sander J (2017) Use of CRISPR/Cas9 for Crop Improvement in
699 Maize and Soybean. In: Weeks DP, Yang B (ed) Progress in Molecular Biology and
700 Translational Science, vol 149. Academic Press, pp 27–46
701 Counce PA, Keisling TC, Mitchell AJ (2000) A Uniform, Objective, and Adaptive System
702 for Expressing Rice Development. Crop Sci 40:436–443.
703 https://doi.org/10.2135/cropsci2000.402436x
704 Dhankher OP, Foyer CH (2018) Climate resilient crops for improving global food security

and safety. Plant Cell Environ 41:877–884. https://doi.org/10.1111/pce.13207

Famoso AN, Clark RT, Shaff JE, et al (2010) Development of a Novel Aluminum Tolerance Phenotyping Platform Used for Comparisons of Cereal Aluminum Tolerance and Investigations into Rice Aluminum Tolerance Mechanisms. Plant Physiol 153:1678–1691. https://doi.org/10.1104/pp.110.156794

Famoso AN, Zhao K, Clark RT, et al (2011) Genetic Architecture of Aluminum Tolerance in Rice (*Oryza sativa*) Determined through Genome-Wide Association Analysis and QTL Mapping. PLoS Genet 7:1–16. https://doi.org/10.1371/journal.pgen.1002221

Fukao T, Xu K, Ronald PC, Bailey-Serres J (2006) A Variable Cluster of Ethylene Response Factor–Like Genes Regulates Metabolic and Developmental Acclimation Responses to Submergence in Rice. Plant Cell 18:2021–2034. https://doi.org/10.1105/tpc.106.043000

Fukao T, Yeung E, Bailey-Serres J (2011) The Submergence Tolerance Regulator SUB1A Mediates Crosstalk between Submergence and Drought Tolerance in Rice. Plant Cell 23:412–427. https://doi.org/10.1105/tpc.110.080325

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Report 13:207–209. https://doi.org/10.1007/BF02670897

Gamuyao R, Chin JH, Pariasca-Tanaka J, et al (2012) The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature 488:535–539. https://doi.org/10.1038/nature11346

Georges F, Ray H (2017) Genome editing of crops: A renewed opportunity for food security. GM Crop Food 8:1–12

Glahn RP, Cheng Z, Welch RM (2002) Comparison of iron bioavailability from 15 rice

729 genotypes: Studies using an in vitro digestion/Caco-2 cell culture model. J Agric Food
730 Chem 50:3586–3591. https://doi.org/10.1021/jf0116496
731 Graham N, Patil GB, Bubeck DM, et al (2020) Plant genome editing and the relevance of
732 off-target changes. Plant Physiol 183:1453–1471. https://doi.org/10.1104/pp.19.01194
733 Grenier C, Cao T-V, Ospina Y, et al (2015) Accuracy of Genomic Selection in a Rice
734 Synthetic Population Developed for Recurrent Selection Breeding. PLoS One
735 10:e0136594. https://doi.org/10.1371/journal.pone.0136594
736 Guimarães PHR, de Lima IP, de Castro AP, et al (2020) Phenotyping Root Systems in a Set
737 of Japonica Rice Accessions: Can Structural Traits Predict the Response to Drought?
738 Rice 13:1–19. https://doi.org/10.1186/s12284-020-00404-5
739 Harrington S (2016) Greenhouse Procedures Manual. Growing rice under controlled
740 conditions. In: McCouch RiceLab.
741 http://ricelab.plbr.cornell.edu/greenhouse_procedures_manual. Accessed 13 May 2021
742 Heuer S, Gaxiola R, Schilling R, et al (2017) Improving phosphorus use efficiency: a
743 complex trait with emerging opportunities. Plant J 90:868–885.
744 https://doi.org/10.1111/tpj.13423
745 Iwai T, Takahashi M, Oda K, et al (2012) Dynamic Changes in the Distribution of Minerals
746 in Relation to Phytic Acid Accumulation during Rice Seed Development. Plant
747 Physiol 160:2007–2014. https://doi.org/10.1104/pp.112.206573
748 Kochian L V., Piñeros MA, Liu J, Magalhaes J V. (2015) Plant Adaptation to Acid Soils:
749 The Molecular Basis for Crop Aluminum Resistance. Annu Rev Plant Biol 66:571–
750 598. https://doi.org/10.1146/annurev-arplant-043014-114822
751 Lönnerdal B (2000) Dietary Factors Influencing Zinc Absorption. J Nutr 130:1378S-1383S.
752 https://doi.org/10.1093/jn/130.5.1378S

753 Lowder LG, Zhang D, Baltes NJ, et al (2015) A CRISPR/Cas9 Toolbox for Multiplexed
754 Plant Genome Editing and Transcriptional Regulation. Plant Physiol 169:971–985.
755 https://doi.org/10.1104/pp.15.00636
756 McCouch SR, Wright MH, Tung C-W, et al (2016) Open access resources for genome-
757 wide association mapping in rice. Nat Commun 7:1–13.
758 https://doi.org/10.1038/ncomms10532
759 McKie VA, McCleary B V (2016) A Novel and Rapid Colorimetric Method for Measuring
760 Total Phosphorus and Phytic Acid in Foods and Animal Feeds. J AOAC Int 99:738–
761 743. https://doi.org/10.5740/jaoacint.16-0029
762 Mickelbart M V., Hasegawa PM, Bailey-Serres J (2015) Genetic mechanisms of abiotic
763 stress tolerance that translate to crop yield stability. Nat Rev Genet 16:237–251
764 Miyao A, Tanaka K, Murata K, et al (2003) Target Site Specificity of the Tos17
765 Retrotransposon Shows a Preference for Insertion within Genes and against Insertion
766 in Retrotransposon-Rich Regions of the Genome. Plant Cell 15:1771–1780.
767 https://doi.org/10.1105/tpc.012559
768 Muthayya S, Sugimoto JD, Montgomery S, Maberly GF (2014) An overview of global rice
769 production, supply, trade, and consumption. Ann N Y Acad Sci 1324:7–14.
770 https://doi.org/10.1111/nyas.12540
771 Oliva R, Ji C, Atienza-Grande G, et al (2019) Broad-spectrum resistance to bacterial blight
772 in rice using genome editing. Nat Biotechnol 37:1344–1350.
773 https://doi.org/10.1038/s41587-019-0267-z
774 Pariasca-Tanaka J, Chin JH, Dramé KN, et al (2014) A novel allele of the P-starvation
775 tolerance gene *OsPSTOL1* from African rice (*Oryza glaberrima* Steud) and its
776 distribution in the genus *Oryza*. Theor Appl Genet 127:1387–1398.

https://doi.org/10.1007/s00122-014-2306-y

Perera I, Seneweera S, Hirotsu N (2018) Manipulating the Phytic Acid Content of Rice Grain Toward Improving Micronutrient Bioavailability. Rice 11:1–13

Rao NKS, Laxman RH, Shivashankara KS (2016) Physiological and Morphological Responses of Horticultural Crops to Abiotic Stresses. In: Rao N, Shivashankara K, Laxman R (eds) Abiotic Stress Physiology of Horticultural Crops. Springer India, New Delhi

Ray DK, West PC, Clark M, et al (2019) Climate change has likely already affected global food production. PLoS One 14:1–18. https://doi.org/10.1371/journal.pone.0217148

Rice Tos17 Insertion Mutant Database (2013) National Agriculture of Food Research Organization, Japan. https://tos.nias.affrc.go.jp/. Accessed 13 May 2021

Rincón SA, Suárez C, Romero-Ruiz M, et al (2014) Identifying Highly Biodiverse Savannas based on the European Union Renewable Energy Directive (SuLu Map) Conceptual background and technical guidance. WWF, Colombia

Rose TJ, Liu L, Wissuwa M (2013) Improving phosphorus efficiency in cereal crops: Is breeding for reduced grain phosphorus concentration part of the solution? Front Plant Sci 4:1–6. https://doi.org/10.3389/fpls.2013.00444

Roy S, Kushwaha NK, Ram H, Soni P (2021) Genome Editing for Improving Abiotic Stress Tolerance in Rice. In: Upadhyay SK (ed) Genome Engineering for Crop Improvement. Wiley, New York

Saito K, Asai H, Zhao D, et al (2018) Progress in varietal improvement for increasing upland rice productivity in the tropics. Plant Prod Sci 21:145–158

Schmidt SM, Belisle M, Frommer WB (2020) The evolving landscape around genome editing in agriculture. EMBO Rep 21:1–4. https://doi.org/10.15252/embr.202050680

801 Shan Q, Wang Y, Li J, Gao C (2014) Genome editing in rice and wheat using the
802 CRISPR/Cas system. Nat Protoc 9:2395–2410. https://doi.org/10.1038/nprot.2014.157
803 Singh A, Singh Y, Mahato AK, et al (2020) Allelic sequence variation in the *Sub1A*, *Sub1B*
804 and *Sub1C* genes among diverse rice cultivars and its association with submergence
805 tolerance. Sci Rep 10:1–18. https://doi.org/10.1038/s41598-020-65588-8
806 Singh N, Dang TTM, Vergara G V., et al (2010) Molecular marker survey and expression
807 analyses of the rice submergence-tolerance gene *SUB1A*. Theor Appl Genet
808 121:1441–1453. https://doi.org/10.1007/s00122-010-1400-z
809 Tang X, Liu G, Zhou J, et al (2018) A large-scale whole-genome sequencing analysis
810 reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in
811 rice. Genome Biol 19:1–13. https://doi.org/10.1186/s13059-018-1458-5
812 Toki S, Hara N, Ono K, et al (2006) Early infection of scutellum tissue with Agrobacterium
813 allows high-speed transformation of rice. Plant J 47:969–976.
814 https://doi.org/10.1111/j.1365-313X.2006.02836.x
815 Uga Y, Sugimoto K, Ogawa S, et al (2013) Control of root system architecture by *DEEPER*
816 *ROOTING 1* increases rice yield under drought conditions. Nat Genet 45:1097–1102.
817 https://doi.org/10.1038/ng.2725
818 Wang DR, Agosto-Pérez FJ, Chebotarov D, et al (2018) An imputation platform to enhance
819 integration of rice genetic resources. Nat Commun 9:. https://doi.org/10.1038/s41467-
820 018-05538-1
821 Wang M, Lu Y, Botella JR, et al (2017) Gene Targeting by Homology-Directed Repair in
822 Rice Using a Geminivirus-Based CRISPR/Cas9 System. Mol Plant 10:1007–1010
823 Wang Y, Cheng X, Shan Q, et al (2014) Simultaneous editing of three homoeoalleles in
824 hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotechnol

825 32:947–951. https://doi.org/10.1038/nbt.2969
826 Wissuwa M, Kondo K, Fukuda T, et al (2015) Unmasking novel loci for internal
827 phosphorus utilization efficiency in rice germplasm through genome-wide association
828 analysis. PLoS One 10:1–21. https://doi.org/10.1371/journal.pone.0124215
829 Xing Y, Zhang Q (2010) Genetic and Molecular Bases of Rice Yield. Annu Rev Plant Biol
830 61:421–442. https://doi.org/10.1146/annurev-arplant-042809-112209
831 Xu K, Xu X, Fukao T, et al (2006) *Sub1A* is an ethylene-response-factor-like gene that
832 confers submergence tolerance to rice. Nature 442:705–708.
833 https://doi.org/10.1038/nature04920
834 Yamaji N, Huang CF, Nagao S, et al (2009) A Zinc Finger Transcription Factor ART1
835 Regulates Multiple Genes Implicated in Aluminum Tolerance in Rice. Plant Cell
836 21:3339–3349. https://doi.org/10.1105/tpc.109.070771
837 Yamaji N, Takemoto Y, Miyaji T, et al (2017) Reducing phosphorus accumulation in rice
838 grains with an impaired transporter in the node. Nature 541:92–95.
839 https://doi.org/10.1038/nature20610
840 Ye H, Zhang XQ, Broughton S, et al (2011) A nonsense mutation in a putative sulphate
841 transporter gene results in low phytic acid in barley. Funct Integr Genomics 11:103–
842 110. https://doi.org/10.1007/s10142-011-0209-4
843 Yin K, Gao C, Qiu J-L (2017) Progress and prospects in plant genome editing. Nat Plants.
844 https://doi.org/10.1038/nplants.2017.107
845 Zafar K, Sedeek KEM, Rao GS, et al (2020) Genome Editing Technologies for Rice
846 Improvement: Progress, Prospects, and Safety Concerns. Front Genome Ed 2:1–16.
847 https://doi.org/10.3389/fgeed.2020.00005
848 Zhao H, Frank T, Tan Y, et al (2016) Disruption of *OsSULTR3;3* reduces phytate and

849 phosphorus concentrations and alters the metabolite profile in rice grains. New Phytol
850 211:926–939. https://doi.org/10.1111/nph.13969

851

852 **Tables**

853

Table 1. Description of genes conferring abiotic stress tolerance used in the study

Gene abbreviation	**Gene name**	**Codes for**	**Major QTL**	**Phenotypic effect**	**Undesirable allelephenotype**	**Reference cultivar (subpopulation) with undesirable allele**	**Desirable allele phenotype**	**Reference cultivar (subpopulation or species) with desirableallele**	**Reference**
ART1	*Aluminum Resistant Transcription Factor*	C2H2 zinc-finger transcription factor	*Alt12.1 (ART1)*	Aluminum (Al) tolerance	Al susceptibility	IR64 *(indica);* Kasalath *(aus)*	Al tolerance	Azucena *(tropical japonica;* Nipponbare *(temperate japonica)*	Yamaji et al., 2009; Famoso et al., 2011; Arbelaez *et al.* , 2017
DRO1	*Deeper Rooting1*	N-myristoylation sites associated with a membrane protein	*DRO1*	Drought avoidance; controls root growth angle	Shallow-rooting	IR64 *(indica)*	Deep-rooting	Kinandang Patong *(tropical japonica*); Nipponbare (*temperatejaponica*)	Uga *et al.* , 2013
SUB1A	*Submergence1*	Ethylene-response-factor-like gene	*SUB1*	Submergence tolerance	Submergence intolerant:Death after 1 week of inundation	Nipponbare *(temperate japonica* : *SUB1A* absent)*;*IR64 *(indica;* S*UB1A-1* absent or allele *SUB1A-2*)	Submergence tolerant: Survival for more than2 weeks under inundation	FR13A, Dhalputtia *(aus)* allele *SUB1A-1* present	Xu *et al* ., 2006
PSTOL1	*Phosphorus Starvation Tolerance1*	Receptor-like cytoplasmic kinase	*PUP1 (PSTOL1)*	Phosphorus (P) acquisition efficiency (PAE); root biomass	Pi deficiency-susceptible	Nipponbare *(temperate japonica*); IR64, IR74 (*Indica*)	Pi deficiency-tolerant	Kasalath *(aus*); CG14 (*Oryza glaberrima)*	Gamuyao *et al.* , 2012; Pariasca *et al.* , 2014
SPDT (SULTR3;4)	*SULTR-like Phosphorus Distribution Transporter*	Transporter for inorganic P (Pi)	No QTL; isolated by *Tos17* tagging	P use efficiency (PUE); grain quality	High grain-P; low strawP	Nipponbare *(temperate japonica*)	Grain-P reduced and straw-P increased (by 20–30%)	Nipponbare *(temperate japonica* : *Tos17* loss-of-function allele at *SPDT*)	Yamaji *et al.* , 2017

Table 2. Description of gRNAs selected for targeting the *SPDT* gene based on CRISPR-P v.2.0

gRNA	**gRNA sequence (PAM site underlined)**	**Position in *SPDT* gene**	***On-target* Score**	**Number of *off-targets***	**Number of *off-targets* in genes**	**Localization in off-target gene**	***Off-target* gen score**	**Number of mistmatches in *off-target* gene**
1U	TTAAATCGACCTTTTCCTGG**TGG**	906 nt upstream of start codon	0.77	12	0	n.a.	n.a.	n.a.
1D	TGTTTCTCACGAGTCGCACA**GGG**	749 nt downstream of stop codon	0.61	5	1	UTR	0.44	4
3	GGATTAAGAGGCGAGTTGGG**GGG**	190 nt upstream of start codon	0.65	40	0	n.a.	n.a.	n.a.
4	GCACATCTCTGCTCTTGTCA**GGG**	At stop codon	0.59	19	1	Intron	0.22	4
1F	TCAAGAACCAGTCGTCCGCG**CGG**	In exon 1	0.90	9	1	CDS	0.26	3

n.a. = non-applicable

Acknowledgments

We gratefully acknowledge Jose Euripides Baquero, Jaime Humberto Bernal and Samuel Caicedo (r.i.p) (AGROSAVIA) for advice on key traits to further improve rice cultivars of importance to the acid soil Colombian savannas, Lyza Maron for advice on potential genes for characterization, Nathaniel Graham and Daniel Voytas (University of Minnesota) for training LSB in final vector assemblies and protoplast transfection, Jon E. *Shaff* for performing strength comparisons of Magnavaca vs. Kimura solutions, Mary Bodis for technical assistance on P and phytate (Megazyme) measurements and bioavailability assays. We acknowledge constructive input from anonymous reviewers whose comments helped us improve the manuscript.

This work was performed as part of an agreement between AGROSAVIA and Cornell University and was funded by the U.S. National Science Foundation Plant Genome Research Program (**Award #1444511 Quantitative Trait Locus Editing for Crop Improvement through S. McCouch), early vegetative and seed phenotyping was supported by USDA-ARS, and salary support for** LSB was provided by AGROSAVIA through the Colombian Ministry of Agriculture, Colombia.

Authors' contributions

LSB conceived, designed, and performed the experiments, analyzed the data, wrote the paper. MRW contributed to conception and design of gene editing. EC contributed to the phenotyping experiments under hydroponic conditions. KMA contributed to genotyping, DNA extractions, PCRs, sequencing, and preparation of seed samples for phenotyping. SH maintained, amplified, and tracked plants and seed stocks. GGM analyzed the data, prepared figures. RG conceived and

advised on nutrient evaluation of seed. MAP conceived and advised on early development phenotyping. SM conceived and supervised the overall research at Cornell University and jointly wrote the paper with LSB. All authors read and approved the final manuscript.

Conflict of Interest Statement

The authors declare that they have no conflicts of interests.

Supplemental Data Files

Figure S1. The *SPDT* gene (Os06g0143700) showing the position of primers used for Sanger sequencing.

Figure S2. Representative sampled organs used for hydroponic experiments.

Figure S3. Follow-up of leaf developmental stages per cultivar under hydroponics P-sufficiency (90µM) conditions.

Figure S4. P concentration shown for the 5- leaf developmental stage in leaf 2 (L2) and shoot basal región (SBR) under hydroponics P-sufficiency conditions (90µM).

Figure S5. Molar ratios of phytic acid (PA) to iron (Fe) and zinc (Fe) of polished seed (PS).

Table S1. Plant material used for the study

Table S2. Primers used for the study

Table S3. Magnavaca and Kimura ½ nutrient solution compositions and ionic strengths

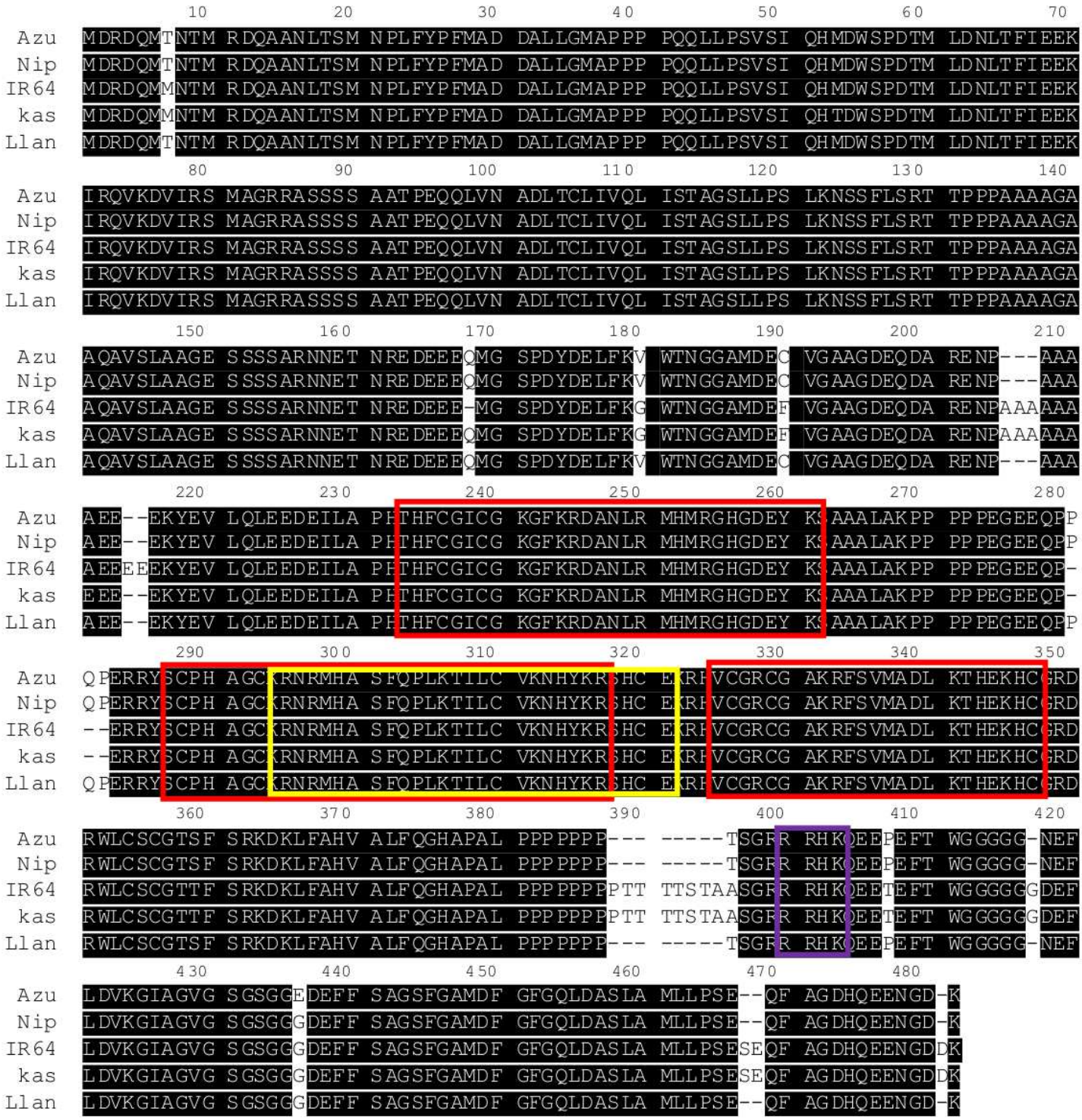

Figure 1. Annealing of the predicted aa sequence of the ART1 protein. In Llanura 11 (Llan) and reference *japonica* Azucena (Azu) Genebank accession No. ATU81899.1, Nipponbare (Nip) Genebank accession No. ATU81901.1, *indica* IR64 Genebank accession No. ATU81900.1, and *aus* Kasalath (kas) Genebank accession No. ATU81902.1. Llan is shown as a representative of upland varieties under investigation since Porvenir 12 was identical. Red squares represent the C2H2 zinc finger domains and yellow and purple squares the monopartite and bipartite nuclear localization domains according to Arbelaez *et al.* (2017).

Figure 2. Allelic variation in PCR products and sequences of three genes. A) *DRO1* nucleotide alignment at exon 4 in Llanura 11 (Llan), the reference *japonica* Kinandang Patong (KP) Genebank accession (Gb) No. AB689741.1, Nipponbare Gb No. AP005570, and the *indica* IR64 Gb No. AB68742.1, showing the InDel 943 from start of the 5´UTR in the transcribed region according to Uga *et al.* (2013). B) *PSTOL1* PCRs using primers indicated above each gel (Table S2). Blue arrows show the expected PCR product in base pairs (bp): the 258 pb indicative of the *O. glaberrima* (Og) CG14 allele and the 342 pb indicative of the Kasalath (Kas) allele. The nucleotide alignment shows six SNPs (Grey boxes) at primers K46-CGsp2 (red square) and K46-Ksp3 (yellow square) between the reference Kas Gb No. AB458444.1 and Og according to Pariasca-Tanaka *et al.* (2014). C. *SUB1A-1* PCRs using primers indicated above each gel (Table S2). Blue arrows show expected PCR products: the 1015 (left) and 825 pb product (right) as confirmed by sequence corresponded to the FR13A allele Gb No. DQ011598.1 (Xu et al., 2006). The Llan sequence is shown as representative of upland cultivars since both Llan and Porvenir 12 (Por) were identical.

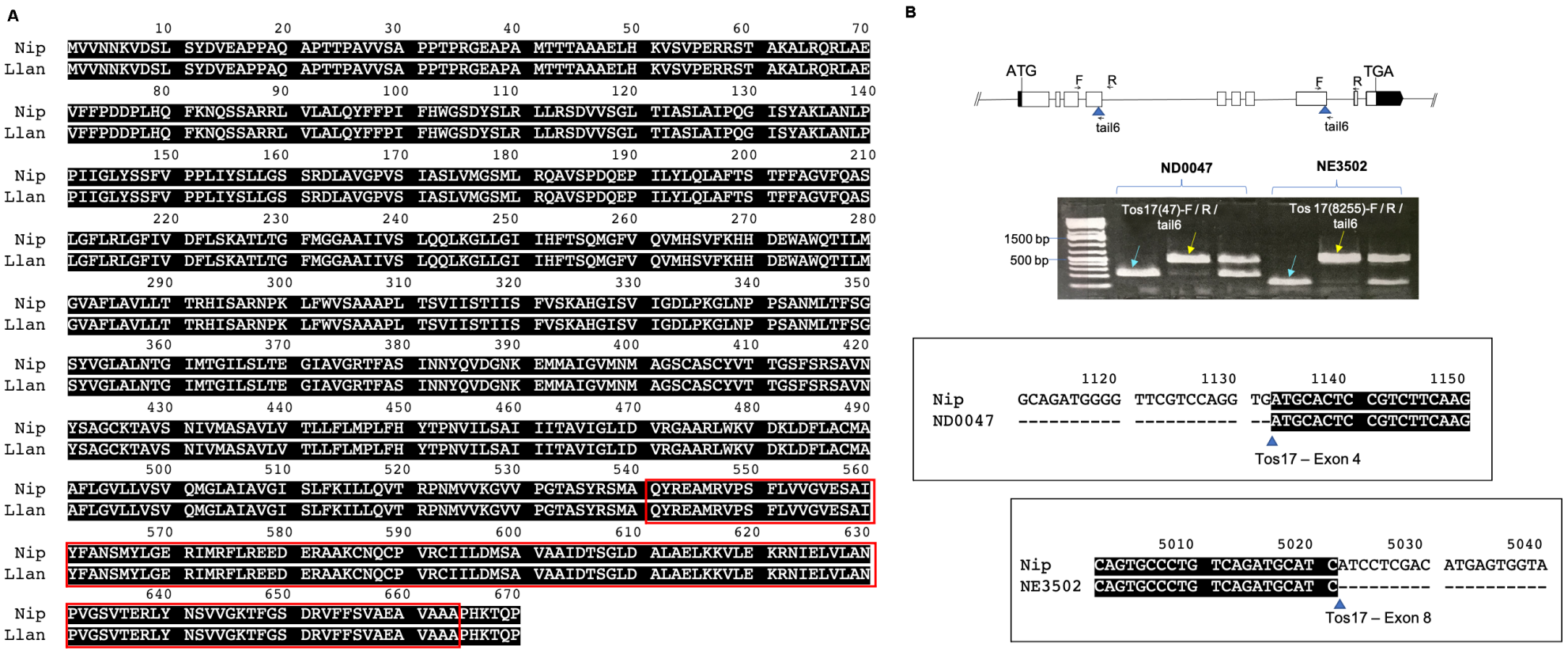

Figure 3. Allelic variation in sequences and PCR products of SPDT. A) Annealing of the predicted aa sequence of SPDT in Nipponbare (Nip) (Os06g0143700) and Llanura 11 (Llan), as a representative of upland cultivars. The sulfate transporter (SULTR) anti-sigma STAS domain is shown as predicted by InterPro-EMBL-EBI (red box). STAS is conserved in the SULTR family though SPDT is the first characterized transporter for Pi in this family (Yamaji *et al.*, 2017). B) Structure of the *SPDT* gene showing the position of the Tos17 insertions (blue triangles) targeting exons 4 and 8 (white boxes) in lines ND0047 and NE3502, respectively, and the position of the 3-primers (F, R and tail 6) used for amplification of the wild-type and mutant alleles (Miyao *et al.*, 2003; Yamaji *et al.*, 2017; Table S2). The gel shows mutant homozygous amplicons (blue arrows) of ~250 bp (in ND0047) and ~150 pb (in NE3502), and wild type homozygous amplicons (yellow arrows) of 476 bp (ND0047) and 521 bp (NE3502) along with heterozygous plants. The nucleotide annealing of wild type vs. mutant alleles shows the Tos17 insertions (Genebank accessions AG25438.1 and AG214407.1).

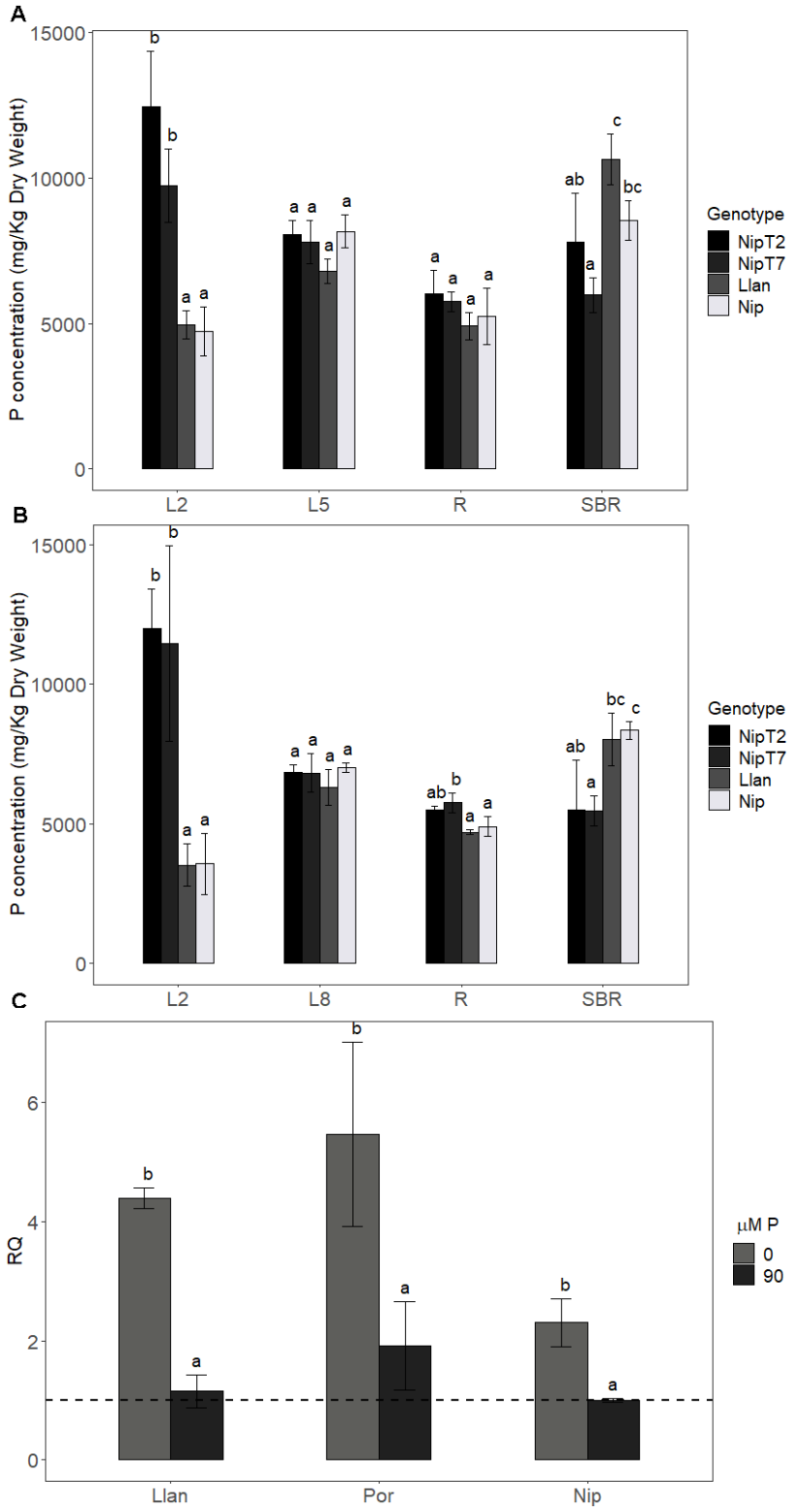

Figure 4. Phenotyping and *SPDT* expression at early vegetative development under hydroponic conditions. A) and B) P concentrations shown for the 5 and 8- leaf developmental stages, respectively, in leaf 2 (L2), leaf 5 (L5), leaf 8 (L8), root (R), and shoot basal region (SBR) under P-sufficiency conditions (90µM). C) Relative quantification (RQ= 2-ΔΔCt) in the SBR of seedlings exposed to P-deficiency (0 µM) and P-sufficiency (90µM) relative to the control condition (90µM P, dashed line) using Actin as internal standard. Significant differences are shown by different letters at $p<0.01$ and standard deviations are shown after Tukey test. NipT2= *spdt* mutant NE3502; NipT7= *spdt* mutant ND0047; Llan= Llanura 11; Nip= *spdt* wild-type Nipponbare.

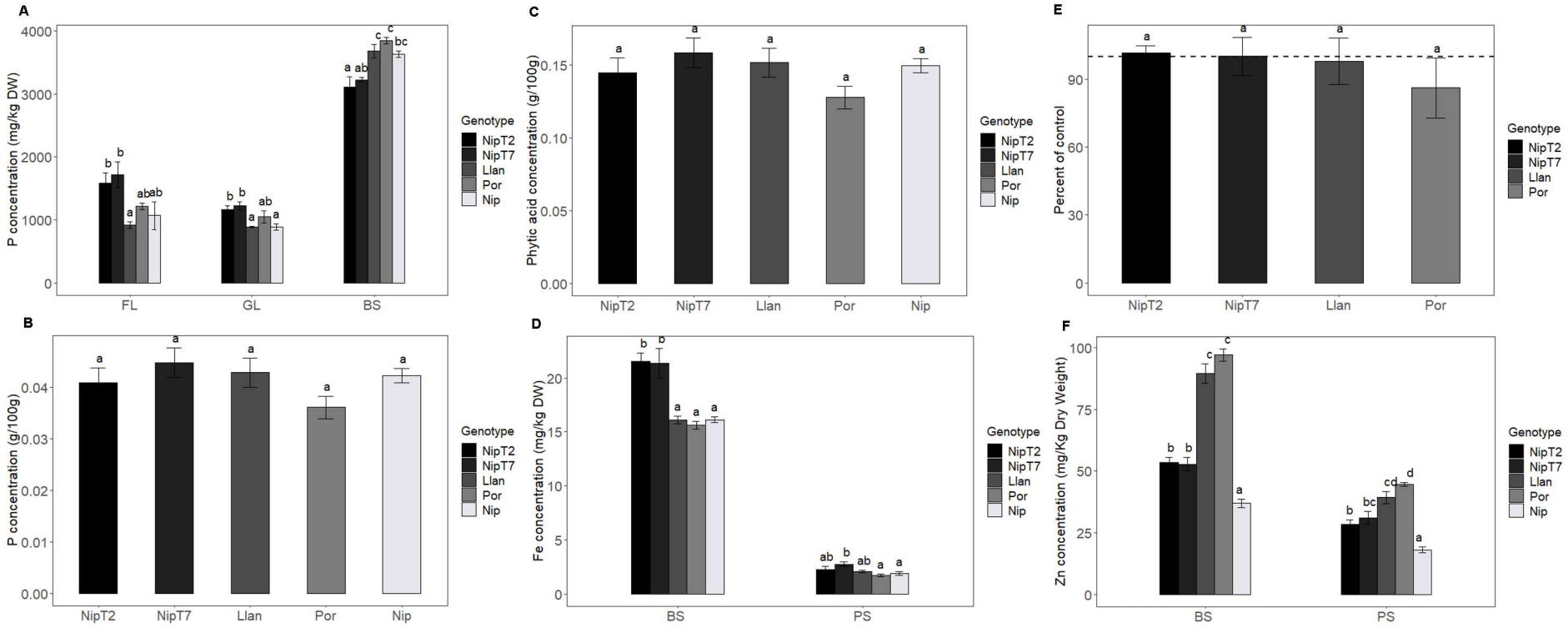

Figure 5. Phenotyping at seed maturity stage under greenhouse conditions. A) P concentration in flag leaf (FL), green leaves (GL) and Brown seed (BS) determined by ICP. B) and C) P and Phytic acid concentrations in polished seed (PS) determined by Megazyme. D) Fe and F) Zn concentrations in BS and PS determined by ICP. E) Fe-bioavailability in PS relative to the control Nipponbare (4,335 ng ferritin/mg protein,100% represented by the dashed line) was measured by Ferritin formation in Caco-2 cells *in vitro* assay. Significant differences are shown by different letters at $p<0.05$ and standard errors are shown after Tukey test. NipT2= *spdt* mutant NE3502; NipT7= *spdt* mutant ND0047; Llan= Llanura 11; Por= Porvenir 12; Nip= *spdt* wild-type Nipponbare.

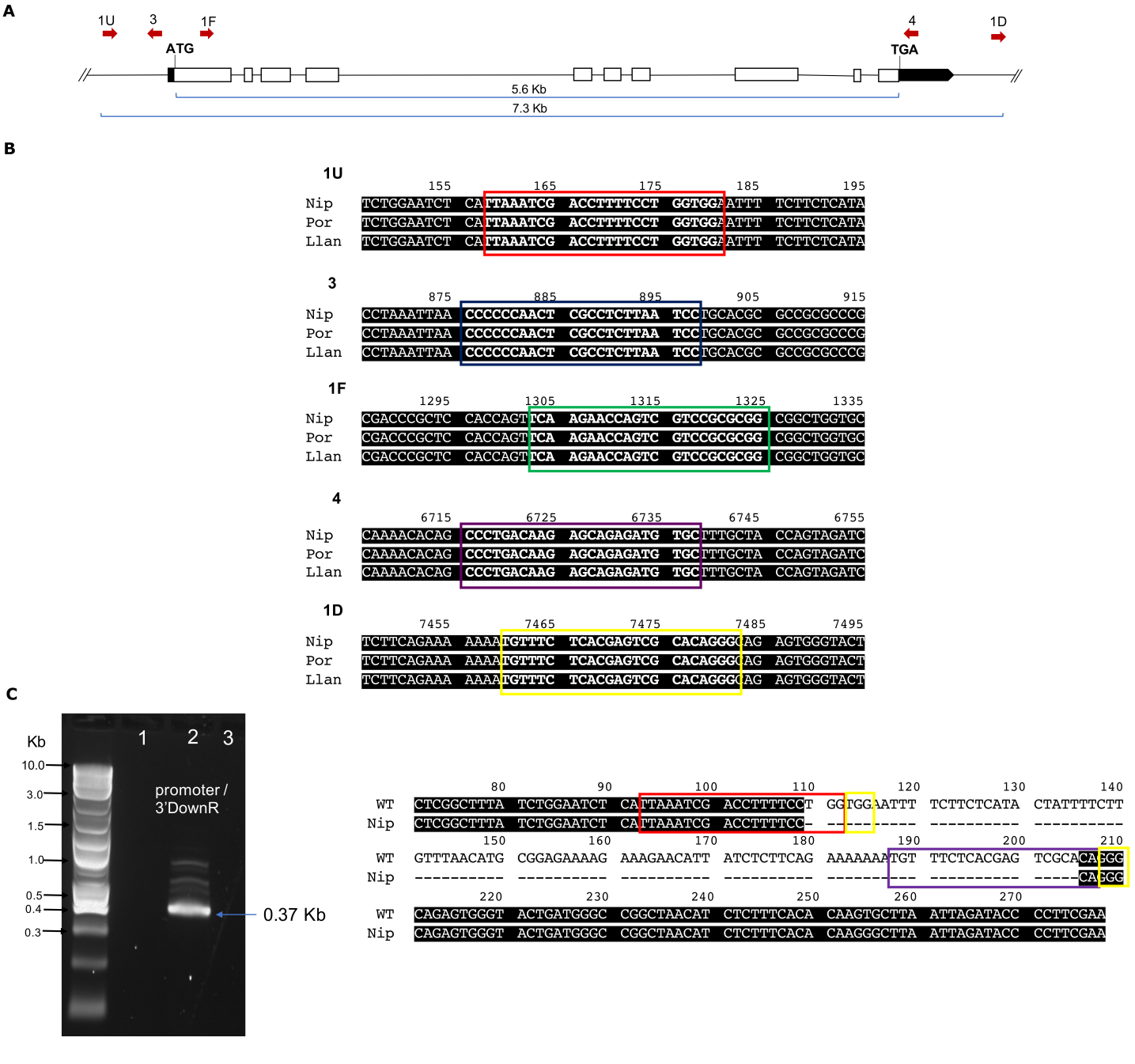

Figure 6. Effectiveness of the *SDPT* gRNAs in cv. Nipponbare protoplasts. A) Position of the selected gRNAs for deletion of the *SPDT* gene (red arrows). The exons (white boxes), UTR region (black boxes), and the start and termination codons (adapted from Yamaji *et al.*, 2017) are indicated. The size of the gene from start to stop codon (5.6 Kb) and the size of the expected deletion using 1U and 1D gRNAs (7.3 Kb) are underlined. B) Nucleotide sequences corresponding to the regions of gRNA annealing in the two Colombian cultivars Por= Porvenir12, Llan=Llanura11 and the reference Nip= Nipponbare used to target the *SPDT* gene. gRNAs are indicated in colored boxes. C) PCR from DNA protoplasts transformed with vector containing the 1U-1D gRNA pair using the primers indicated above the gel (Table S2). The sequence of the expected 370 bp band indicative of the *SPDT* deletion (line 2) was aligned with the wild-type (WT) gene. The 1U (red box), 1D (Purple box) gRNAs, and the PAM sites (Yellow boxes) are indicated. To facilitate visualization, most of the WT 7.3 kb deleted region is omitted and only the 5' and 3'-ends are shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTablesandFigures.zip