The bHLH transcription factor OsPRI1 activates the *Setaria viridis PEPC1* promoter in rice

Pedro Carvalho[1] ORCID: 0000-0002-0760-323X
Célia Gomes[1] ORCID: 0000-0003-1159-0555
Ivan Gonçalves [1] ORCID:
Tiago F. Lourenço [1] ORCID: 0000-0001-5085-861X
Daniela Vlad [2] ORCID: 0000-0002-1341-5868
Jane A. Langdale[2] ORCID: 0000-0001-7648-3924
Nelson J.M. Saibo[1,3] ORCID: 0000-0002-6679-3811

[1] Instituto de Tecnologia Química e Biológica António Xavier da Universidade Nova de Lisboa (ITQB NOVA), Oeiras, Portugal

[2] Department of Biology, University of Oxford, South Parks Rd, OX1 3RB United Kingdom

[3] Instituto de Biologia Experimental e Tecnológica (iBET), 2780-157, Oeiras, Portugal.

Corresponding Author: Nelson J.M. Saibo, saibo@itqb.unl.pt

Keywords: C_4 photosynthesis; phosphoenolpyruvate carboxylase; cell-specific gene expression; rice; transcriptional regulation

Summary

The efficiency of photosynthesis is reduced by the dual role of Rubisco, which can act either as a carboxylase or as an oxygenase, the latter leading to photorespiration. C_4 photosynthesis evolved as a carbon concentrating mechanism to reduce photorespiration. To engineer C_4 into a C_3 plant such as rice, it is essential to understand how C_4 genes, like that which encodes the *phosphoenolpyruvate carboxylase* (*PEPC1)*, are expressed at high levels and cell-specific manner. The *PEPC1* promoter from several C_4 grasses has been shown to have cell-specific activity in rice, but the regulators are largely unknown. Here we show that OsPRI1, a rice bHLH transcription factor that plays a role in iron homeostasis, activates the *Setaria viridis PEPC1* promoter. We show that OsPRI1 binds to an N-box present in the proximal promoter and that GUS activity is highly reduced in *Ospri1* loss of function rice lines harbouring *SvPEPC1*$_{pro}$*::GUS* constructs. Cross species comparisons showed that the SvPRI1 ortholog binds to the *SvPEPC1* promoter but the maize ZmPRI1 does not bind to the *ZmPEPC1* promoter. Collectively, these results suggest that elements of the iron homeostasis pathway were co-opted to regulate *PEPC1* gene expression during the evolution of some but not all C_4 species.

Introduction

Rice is the staple food crop for more than half of the world's population (Maclean *et al.*, 2013). In a climate change scenario, regions with warm and dry seasons are predicted to increase, creating a new constraint on agro-industrial systems. Combined with an increasing population, this presents a huge challenge to plant research (Godfray *et al.*, 2010) because crop production must be improved under adverse environmental conditions. C_4 photosynthesising plants are higher yielding in warm dry environments than their C_3 counterparts because the C_4 pathway generates a higher carbon concentration around the central carbon fixing enzyme, Ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco). This concentrating mechanism suppresses the enzyme's oxygenase activity and consequently eliminates the energetically wasteful process of photorespiration (Sage *et al.*, 2012). Understanding the molecular mechanisms underlying the regulation of the C_4 carbon concentrating shuttle is crucial for engineering C_4 metabolism into C_3 plants such as rice. Predictions indicate that engineering C_4 photosynthesis into rice could lead to a 50% increase in photosynthetic efficiency, which should translate into a substantial increase in grain yield (von Caemmerer *et al.*, 2012).

Phosphoenolpyruvate carboxylase (PEPC) is a critical enzyme in the C_4 cycle, because it is responsible for the initial fixation of CO_2 (Svensson et al., 2003). In comparison to PEPC enzymes in C_3 plants, C_4 pathway PEPCs have altered enzyme kinetics and accumulate specifically in leaf mesophyll cells. It is this mesophyll-specific accumulation profile that enables an initial carbon fixation step to occur in the absence of Rubisco, which accumulates specifically in the bundle sheath cells of C_4 leaves (Hatch & Slack, 1966). Notably, all of the enzymes required for C_4 metabolism can be found in ancestral C_3 species (Sage *et al.*, 2012). The evolution of the C_4 pathway occurred in multiple independent occasions (Sage, 2004). This required the innovation or co-option of mechanisms to ensure that genes, encoding C_4 enzymes, are either transcriptionally or post-transcriptionally regulated in the appropriate cell-type.

Zea mays (maize) and *Setaria viridis* (setaria) represent two independent origins of the C_4 pathway, with both species belonging to the so-called PACMAD clade of grasses, in which there are multiple origins of C_4 (Grass Phylogeny Working Group II, 2012). By contrast there are no C_4 species in the sister (BEP) clade, which includes *Oryza sativa* (rice) (Grass Phylogeny Working Group II, 2012). It has been shown that a number of *PEPC* promoters from C_4 grass species drive mesophyll-specific gene expression in rice, indicating that the trans-regulatory network required for cell-specificity is present in rice (Matsuoka *et al.*, 1994; Ku *et al.*, 1999; Gupta *et al.*, 2020). Although *cis*-elements in a number of different C_4 *PEPC* promoters have been characterized (Stockhaus *et al.*, 1997; Gowik *et al.*, 2004; Akyildiz *et al.*, 2007; Gupta *et al.*, 2020), the knowledge of transcription factors that bind C_4 *PEPC1* promoters is restricted to maize, where five have been identified (Górska et al., 2019; Górska et al., 2021; Yanagisawa and Sheen, 1998). The only known transcription factor from rice that has been shown to bind to C_4 *PEPC1* promoter is OsbHLH112, an homologue of ZmbHLH80 and ZmbHLH90, which have been reported as regulators of *ZmPEPC1*$_{pro}$ in maize (Górska *et al.*, 2019).

The value of C_4 *PEPC1* promoters for biotechnological approaches in rice has already been demonstrated in the C_4 Rice project, where different *PEPC* promoters were used to drive expression of genes encoding the mesophyll specific enzymes of the C_4 cycle (Ermakova *et al.*, 2021). Although mesophyll-specificity was achieved, the high expression level that characterizes C_4 *PEPC* promoter activity in C_4 species was not. It is thus important to understand the trans factors that regulate C_4 promoters in rice to facilitate a better manipulation of gene activity for biotechnological applications.

Here we used yeast one-hybrid assays to identify a bHLH transcription factor from rice that activates expression of the *Setaria viridis PEPC1* gene promoter. Activation was confirmed *in planta* in rice using promoter-reporter fusions. The known role of this bHLH in C_3 plants, and the fact that its setaria homolog binds to the *SvPEPC1* gene promoter, allows to hypothesize a link between iron sensing/homeostasis and the C_4 photosynthetic metabolism in setaria.

Materials and methods

Setaria virdis genomic DNA purification

Setaria virdis (var. ME034V) genomic DNA (gDNA) was isolated using a sodium dodecyl sulfate (SDS) extraction protocol. Briefly, gDNA was isolated from two-week-old seedling material ground in liquid nitrogen, 200 µl of extraction buffer (250 mM Tris-HCl (pH 7.5), 25 mM Ethylenediaminetetraacetic acid (EDTA), 250 mM NaCl, 1% (v/v) SDS) was added to 25 mg of tissue followed by centrifugation at 12 000 x *g* for 10 minutes. The aqueous phase was then transferred to a new tube and 0.8 volumes of isopropanol were added. The mixture was incubated at room temperature for 10 minutes followed by centrifugation at 12 000 x *g* for 10 minutes. The supernatant was discarded and the pellet was washed twice with 200 µl of 70% ethanol, and allowed to dry completely. gDNA was resuspended in 50 µl of ddH_2O with 10 µg/ml RNase A.

Cloning the *Setaria virdis PEPC1* promoter

The *SvPEPC1* (Sevir.4G143500) promoter, -1190bp from the ATG, was amplified by PCR using primers with attB adaptors (Table S1). After analysis by gel electrophoresis, the PCR products were cloned into the pJet1.2/blunt cloning vector, using a CloneJet PCR cloning Kit (Thermo Fisher), and then the inserted DNA was sequenced. The insert was then sub-cloned into the pDONR221 entry vector using the Gateway cloning system

(Thermo Fisher). Subsequently, the insert was sub-cloned into a modified pHGWFS7 vector to create reporter constructs expressing the *β-glucuronidase* (*GUS*) reporter gene under the control of the *SvPEPC1* promoter. Vector pHGWFS7 (Karimi et al. (2002)) was modified so that the *hptII* gene is expressed under the control of the rice *Actin* promoter isolated from pANIC6B expression vector (Mann et al. (2011)).

Rice transformation

Rice (*Oryza sativa* cv. Kitaake) genetic transformation was performed following the protocol described by the Langdale lab (https://langdalelab.files.wordpress.com/2018/06/kitaake-rice-transformation.pdf). Briefly, dehulled and surface disinfected rice seeds were placed in callus initiation media (R1 - 4.3 g/L MS salts & vitamins, 30 g/L Sucrose, 0.5 g/L MES, 300 mg/L Casamino acid, 2.8 g/L L-Proline, 2 mg/L 2,4-D, 4 g/L Phytagel, pH 5.8) for two weeks in continuous light at 32°C, to initiate callus formation. Afterwards, calli were sub-cultured to fresh R1 for three days followed by incubation with *Agrobacterium tumefaciens* (strain EHA105 (Wise et al. (2006))), harbouring the construct of interest. Following Agrobacterium transformation, calli were grown in co-cultivation media (R2 - 4.3 g/L MS salts & vitamins, 30 g/L Sucrose, 0.5 g/L MES, 10 g/L Glucose, 300 mg/L Casamino acid, 2 mg/L 2,4-D, 4 g/L Phytagel, 20 mg/L Acetosyringone, pH 5.2) on top of sterile filter paper for 3 days in the dark at room temperature. Hygromycin resistant calli were selected on selection medium (R3 - 4.3 g/L MS salts & vitamins, 30 g/L Sucrose, 0.5 g/L MES, 300 mg/L Casamino acid, 2.8 g/L L-Proline, 2 mg/L 2,4-D, 4 g/L Phytagel, 200 mg/L Timentin, 30 mg/L Hygromycin, pH 5.8) for two weeks in continuous light at 32 °C, followed by a second round of selection in the same conditions. Plant regeneration was achieved by growing hygromycin resistant calli on regeneration medium (R4 - 4.3 g/L MS salts & vitamins, 30 g/L Sucrose, 0.5 g/L MES, 2 g/L Casamino acid, 30 g/L Sorbitol, 2 mg/L Kinetin, 1 mg/L NAA, 200 mg/L Timentin, 20mg/L Hygromycin, 4 g/L Phytagel, pH 5.8) in rounds of two-week culture in continuous light at 32 °C. Regenerated plants were assessed by PCR for the presence of the hygromycin resistance gene (*hptII*). Positive plants were transferred into soil and grown in a growth chamber at 28 °C with a 12 h/12 h light/dark photoperiod and light intensity of 400 µmol m^{-2} s^{-1}.

To obtain loss of function *OsPRI1* alleles, single guide RNAs (sgRNA) were designed using the web-based tool CRISPR-P (http://crispr.hzau.edu.cn/CRISPR2/) with default settings. For rice transformation, sgRNAs were synthetized with *Bsa*I compatible overhangs (Table S1) and cloned into a modified CRISPR-Cas9 expression vector (Miao *et al.*, 2013) under the control of a U3 promoter. The modified CRISPR-Cas9 expression vector contained the *neomycin phosphotransferase* (*nptII*) gene, allowing for the selection of rice calli using G418 (150mg/l). In order to investigate the function of OsPRI1 regulating *SvPEPC1*$_{pro}$ in rice, the rice reporter line *SvPEPC1*$_{pro}$::*GUS* (previously obtained as described above in the cv. Kitaake) was used as background for the OsPRI1 knockout, using the CRISPR/Cas9 technology. Alternativelly, we could have crossed a rice reporter line *SvPEPC1*$_{pro}$::*GUS* (cv. Kitaake) with the *ospri1* mutants (cv. Nipponbare) already reported (Zhang *et al.*, 2017), but since they have different backgrounds this would make the cross more complicated.

To evaluate putative Cas9 edits (indels), the targeted region of the *OsPRI1* gene was amplified by PCR using primers listed in Table S1 and sequenced. DNA sequence comparisons allowed the identification of mutations resulting from Cas9 activity. DNA sequence deconvolution was performed using the TIDE software (http://tide.nki.nl) with default settings. Transgenic plant lines with indels causing frameshift mutations were propagated until homozygous mutations were obtained in transgene (CRISPR-Cas9 cassette)-free lines.

Yeast one-Hybrid

Yeast bait strains containing three different overlapping fragments of the *SvPEPC1* promoter (F1:-1 to -300 bp; F2:-200 to -600 bp; F3: -500 to -1190 bp, from the ATG) were constructed. The *SvPEPC1* promoter fragments were amplified from the promoter by PCR, using primers listed in Table S1. Using *Not*I (*Nar*I for F2) and *Spe*I restriction sites included in the primer sequence, promoter fragments were cloned into the pINT/HIS vector system (Ouwerkerk and Meijer, 2001) and integrated into the Y187 yeast strain (Takara Bio, USA). Titration with 3-amino triazole (3-AT), a competitive inhibitor of the *HIS3* reporter gene product, was performed for all bait strains to determine the concentration needed to counter any non-specific expression of the reporter gene. All experiments were performed in Synthetically Defined medium (SD) (Takara Bio, USA) with experimentally defined concentrations of 3-AT.

Bait strains were transformed with a salt-induced rice cDNA expression library available in the host lab (Almeida et al., 2017). Yeast transformation was performed as described in the YEASTMAKER™ protocol (Takara Bio, USA). Briefly, cells were cultivated overnight at 30 °C in YPDA medium and collected by centrifugation at 300 x *g* for 10 min. Pellets were washed twice with sterile water and competent cells were obtained using 1.1 x TE/LiAc solution (10 mM Tris-HCl, 1 mM EDTA, 100 mM Lithium acetate, pH 7.5). Competent cells were incubated at 30 °C with 800 ng of library plasmid DNA and 100 ng of carrier DNA (Takara Bio, USA) for 30 min, followed by the addition of PEG/LiAc solution (40 % (w/v) polyethylene glycol (PEG), 100 mM LiAc), and incubated at 42 °C for 15 min. Transformed cells were plated on SD lacking leucine and histidine, with previously determined concentrations of 3-AT, and incubated for 3 days at 30 °C. To calculate transformation efficiency, 1:100 and 1:1000 dilutions were plated (SD/-L) and colony-forming units (cfu) were counted after 3 days. For each bait strain over one million clones were screened (F1- 1.83×10^6 cfu; F2- 3.75×10^6 cfu; F3- 1.05×10^6 cfu). To confirm OsPRI1 interactions, the pDEST22 expression vector containing *OsPRI1* cDNA, empty vector, or a non-related transcription factor (TF) (the bHLH TF OsPIF14 (Cordeiro et al. (2016)) was used as the negative control.), were each transformed into the yeast bait strain. pDEST22 expression vectors were created using the Gateway technology.

Isolation of *PEPC* promoters and identification of TF homologues

Two overlapping fragments (spanning -1000bp from the ATG) of the *PEPC* promoter from *Dichanthelium oligosanthes* (Do021545) and *Oryza sativa* (*OsPEPC1*- LOC_Os02g14770 and *OsPEPC4*- LOC_Os01g11054) were synthesized by Twist Biosciences according to the available sequences, with overhangs compatible for cloning into the pINT/HIS vector system as described before, using *Not*I and *Spe*I restriction enzymes. Homologues for OsPRI1 in maize (GRMZM2G017586 and GRMZM2G093744), setaria (Sevir.6G068300), and dichanthelium (OEL18192) were identified using OrthoFinder (Emms and Kelly, (2019)). The identified orthologues were synthesized according to available DNA sequences. TF DNA sequences were cloned into pDONR221 and sub-cloned into pGADT7, using the Gateway system.

Recombinant protein production

For recombinant protein purification, OsPRI1 was fused to the C-terminal of the Maltose Binding Protein (MBP), by sub-cloning into the pMAL-C2X expression vector, adapted for gateway cloning, using the LR gateway technology. *Escherichia coli* strain Rosetta (DE3) pLysS was used for the expression of recombinant proteins. Cells transformed with pMAL::*MBP-OsPRI1* were grown at 37 °C overnight, in Luria-Bertani medium (LB) to an OD_{600}≈0.4. Induction of MBP-OsPRI1 expression was performed using 0.8 mM IPTG, after which cultures were grown for 4 h at 16 °C.

Cells were collected by centrifugation at 4 000 x *g* for 10 min and then lysis buffer (100 mM $KHPO_4$, 0.25 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1X Protease Inhibitor Complete (Hoffmann–La Roche AG, Switzerland), 0.1 mg/ml Lysozyme, 0.1 mg/ml DNAse) was added to the pellet on ice. Cell debris was removed by centrifugation at 14 000 x *g* for 1 h at 4 °C. MBP and MBP-OsPRI1 recombinant proteins were purified by affinity chromatography using MBPTrap HP (GE Healthcare) and HiLoad 16/600Superdex 200 pg columns (GE Healthcare). Recombinant protein purification was assessed by SDS-PAGE.

Radioactive electrophoretic mobility shift assay

DNA probes were designed to include putative bHLH binding sites (*cis*-elements) plus flanking regions, defined here as 13 nucleotides to either side of the element. Putative bHLH binding sites were identified by scanning the promoter fragment for “CANNTG”, “GTNNAC”, CACGGC”, “CGGCAC”, “CACGAG”, and “GACGAC”. Probes used in this study are described in Table S2.

Single-stranded oligos were annealed in labelling buffer (10 mM Tris–HCl pH 8, 50 mM NaCl and 1 mM Na_2EDTA pH 8) by incubation at 95 °C for 5 min, and reactions were allowed to cool down to room temperature. DNA probes were radiolabelled with ATP [γ-32P] (Perkin-Elmer) using T4 Polynucleotide kinase (PNK) (Sigma-Aldrich). Ten pmol of double-strand DNA probe were incubated with T4 PNK at 37 °C for 60 min, and the reaction was stopped by addition of 0.5 M EDTA followed by 15 min incubation at 75

°C. Unincorporated ATP was removed using Sephadex G-25 columns following the manufacturer’s instructions (Sigma-Aldrich). Radiolabelling was confirmed using a Geiger counter.

Protein-DNA incubation was conducted in Binding Buffer (1 mM HEPES pH 7.9, 4 mM KCl, 0.1 mM EDTA, 1 mM DTT, 50 ng herring sperm DNA, 0.05% (w/v) BSA, 10% (w/v) glycerol) at 30 °C for 1 h, using 1 µg of recombinant protein and 5 pmol of labelled probe. For competition assays, unlabelled probe (“cold”) was used in 600 x molar excess.

Reactions were loaded onto NativePAGE 4 to 16 % pre-cast gels (Invitrogen) and electrophoresis was run at 50 V for 3 h at 4 °C in a X-Cell SureLock™ Mini-Cell system. Radioactivity detection was performed using a phosphor screen, imaged with Fuji TLA-5100.

Isolation and transformation of rice protoplasts

Protoplasts were obtained as described in Górska et al. (2021). Rice protoplasts were isolated from rice cell suspension culture or etiolated seedlings. Rice suspension cell cultures were grown for four days, and cells were collected by centrifugation in a “swing-out” rotor at 150 x *g* for 5 min. The digestion solution (0.4 M mannitol, 10 mM MES pH 5.7, 1 mM $CaCl_2$, 0.1% BSA, 50 mg L-1 ampicillin, 5 mM β-mercaptoethanol, 2.25 % (w/v) Cellulase R10 (Duchefa), 0.75 % (w/v) Macerozyme R10 (Duchefa)) was added to the cell pellet or to the etiolated seedling pieces, and the samples were vacuum infiltrated. Digestion was performed for 5 h with 80 RPM orbital agitation at room temperature (RT). The enzyme solution containing protoplasts was filtered through a 100 µm mesh, washed with 1 volume of wash solution (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 2 mM MES pH 5.7) and filtered again through a 50 µm filter. Protoplasts were harvested by centrifugation in a “swing-out” rotor (150 x *g*, 5 min) and resuspended in 200 µl MMg solution (0.4 M mannitol, 4 mM MES pH 5.7, 15 mM $MgCl_2$). Afterwards, protoplasts were diluted to a 1 x 10^6 cells ml^{-1} concentration and permeabilized with polyethylene glycol 4000 (PEG 4000). Protoplasts were transformed by gently mixing 200 µl of protoplast solution with 10 µl of plasmid mix (2 µg of effector plasmid and 3 µg of reporter plasmid) and 220 µl PEG solution (PEG 4000 40%, 0.4 M mannitol, 0.1 M $CaCl_2$). Protoplasts were incubated at RT in the dark for 20 min before being diluted with 3 volumes of wash solution, harvested by centrifugation in a “swing-out” rotor (150 x *g*, 5 min) and resuspended in 750 µl incubation solution (0.4 M mannitol, 4 mM MES pH 5.7, 20 mM KCl). After this, protoplasts were incubated for 15-16 h at RT in the dark. Finally, protoplasts were collected by centrifugation at 150 x *g* for 3 min at RT and the pellet was flashed frozen in liquid nitrogen.

Transactivation assay in rice cultured-derived protoplasts

Protoplasts were co-transformed with an effector plasmid (p2GW::*OsPRI1* or p2GW::Empty) and a reporter plasmid (pGreenII-LUC:: $\textit{SvPEPC1}_{pro}$ *_F3*, *F2*, *F1* or pGreenII-LUC::Empty). Transformed protoplasts were lysed by the addition of 100 µl of Passive lysis buffer (Promega), followed by two freeze-thaw cycles. Luciferase activity assays were performed using the Dual-Luciferase Reporter System (Promega), with a modified protocol. Briefly, 40 µl of LarII solution was added to 20 µl of protoplast lysate and *Luciferase* (LUC) activity was measured for 15 s. Immediately after, 40 µl of Stop&Glow solution was added into the same sample and *Renilla* (REN) activity was measured for 15 s. Measurements were performed using FluoStar fluorometer (BMG Labtech). Transcriptional activity was defined as LUC/REN ratio.

Plant growth conditions and gas exchange measurements

Plants were grown in hydroponics using Yoshida medium [1.4 mM NH_4NO_3, 370 µm NaH_2PO_4, 512 µM K_2SO_4, 998 µM $CaCl_2$, 1.6 mM $MgSO_4$, 9.5 µM $MnCl_2$, 0.075 µM $(NH_4)_6Mo_7O_{24}$, 0.152 µM $ZnSO_4$, 19 µM H_3BO_3, 0.16 µM $CuSO_4$, 100 µM FeNaEDTA, 70.7 µM citric acid, pH 5.4]. Growth conditions were set at constant 28 °C with a 12 h/12 h (light/dark) photoperiod with 800 $\mu molm^{-2}s^{-1}$ light intensity. Plants were grown for 2 weeks in Yoshida medium, with a medium change every week.

Photosynthetic measurements were performed using a Li6800 (LiCOR Biosciences, USA) with the following ambient settings: $[CO_2]_R$ at 400ppm, light intensity 800 PAR, RH at 50-60%. The last fully expanded leaf of each plant was allowed to acclimate to chamber conditions before measurements were recorded. Data analysis was performed using GraphPad Prism 8.

Analysis of gene expression in *Ospri1* mutant lines

Samples were collected at ZT4 (4 h from light initiation) which corresponds to the *SvPEPC1$_{pro}$::GUS* activity peak (Fig. S1). Three biological replicates were collected for each line, each biological sample representing a pool of three plants. Samples were macerated in liquid nitrogen and RNA was extracted using Direct-zol Miniprep (Zymo research), following the manufacturer's instruction. gDNA was removed using a Turbo DNA-*free* kit (Invitrogen), following the manufacturer's instructions. RNA was quantified by spectrophotometry and integrity verified by agarose electrophoresis. 500 ng of total RNA was used for reverse transcription. cDNA was synthesized using Transcriptor first-strand cDNA synthesis kit (Roche), following the manufacturer's instructions and an anchored Oligo $(dT)_{18}$ primer. qPCR was performed using LightCycler SYBR Green I Master mix (Roche) and gene-specific primers (listed in Table S1). Peroxiredoxin-2C (PRXIIc, LOC_Os01g4842) was used as housekeeping gene. Data analysis was performed using GraphPad Prism 8.

Histological GUS detection

To analyse the cell-specific activity of *SvPEPC1$_{pro}$::GUS* in the reporter lines (with or without *OsPRI1* mutated), fully expanded leaves from two-week-old seedlings were hand sectioned to obtain transverse leaf sections. GUS activity was detected using the 5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid cyclohexylammonium salt (X-Gluc) cleavage assay. Samples were incubated in 90 % ice-cold acetone for 1 h at -20 °C followed by a wash with 100 mM phosphate buffer ($NaPO_4$ (pH 7.6)) and then incubated with staining solution (2mM X-Gluc, 100 mM phosphate buffer, 10 mM EDTA, 6 mM Ferrocyanide, 6 mM Ferricyanide) at 37 °C for 1 h. A vacuum was applied to ensure uniform infiltration of the staining solution. After staining, samples were fixed by incubating in ethanol:acetic acid (3:1) solution for 1 h and then stored in 70 % ethanol. Images were obtained using a Leica DM 6000B optical microscope (Leica) and analysed using Image J software.

Results

The rice bHLH transcription factor OsPRI1 binds to a N-box in the *Setaria viridis PEPC1* promoter.

The main goal of this work was to identify new rice transcription factors regulating a C_4 *PEPC* promoter in rice and to characterize their biological function in rice and in a C_4 plant. Since the *PEPC1* promoter from *Setaria viridis* is known to drive cell-specific gene expression in rice (Gupta *et al.*, 2020), and *S. viridis* is a C_4 plant with a short live cycle and relatively easy to transform, we selected the *PEPC1* promoter from setaria for this study. To identify potential regulators of the *Setaria viridis PEPC1* gene promoter (*SvPEPC1*$_{pro}$) in rice, fragments of the promoter were used as bait in a Yeast one-Hybrid (Y1H) screen with a rice cDNA expression library. Among the eighteen putative interactors identified (Table S3), only one was a transcription factor, identified by sequence analysis as OsPRI1. Direct Y1H assays suggested that OsPRI1 binds to three separate regions (F1, F2, F3) of the *SvPEPC1*$_{pro}$ sequence (Fig. 1A, B). Transcription factors from the bHLH family are known to bind both E-boxes and N-boxes (Li *et al.*, 2006), but little is known regarding the nucleotide preference of each protein, with new softwares emerging to predict TF targets (e.g. Yan *et al.*, 2022; Cheng *et al.*, 2023). The *Setaria viridis PEPC1* promoter sequence contains seven different E-Boxes and two different N-Boxes variants, making a total of seventeen putative bHLH binding sites (Fig. S2). Binding of OsPRI1 to probes for each of these seventeen sites was tested in Electrophoretic Mobility Shift Assays (EMSA) (Fig. S2). No binding was observed for the probes present in the F2 and F3 regions but shifts were observed with four probes in the F1 region (probes 11-14) (Fig. S2). Further assays using mutated and non-labelled probes showed that, among these 4 probes, only probe 14 is specifically bound by OsPRI1 (Fig. 1C and S2). We thus conclude that OsPRI1 interacts with *SvPEPC1*$_{pro}$ by binding to a N-Box (CACGAG) positioned at -153bp before the ATG in the *SvPEPC1* sequence, altought other *cis*-elements may also play a role in the regulation of *SvPEPC1*$_{pro}$ by OsPRI1.

OsPRI1 activates *SvPEPC1* promoter activity in transient protoplast assays.

To test whether the binding of OsPRI1 to the *SvPEPC1* promoter is functionally relevant, the three fragments where OsPRI1 was shown to bind in Y1H assays were cloned upstream of a *LUC* reporter gene and rice protoplasts were co-transformed with reporter (F1-3::LUC) and effector (OsPRI1) constructs. Fig. 2 shows that OsPRI1 acts as an activator when the *LUC* reporter gene is driven by the F1 and F2 fragments but not by the F3 fragment. It is not clear whether activation via the F1 region is through the N box identified by EMSA but it is notable that higher levels of activation are observed with the F2 region (3-fold versus 2-fold with F1), in which no specific binding site could be identified.

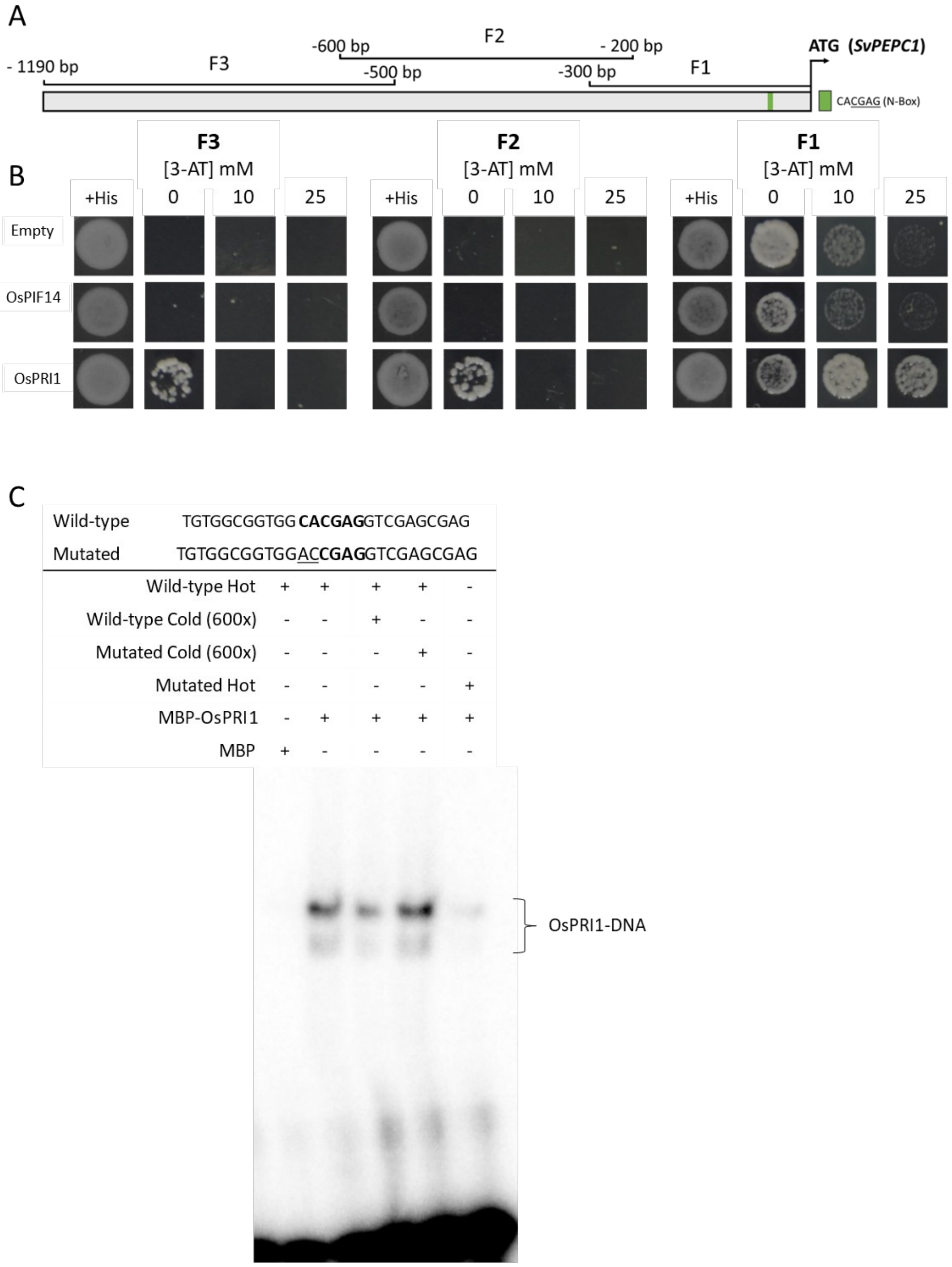
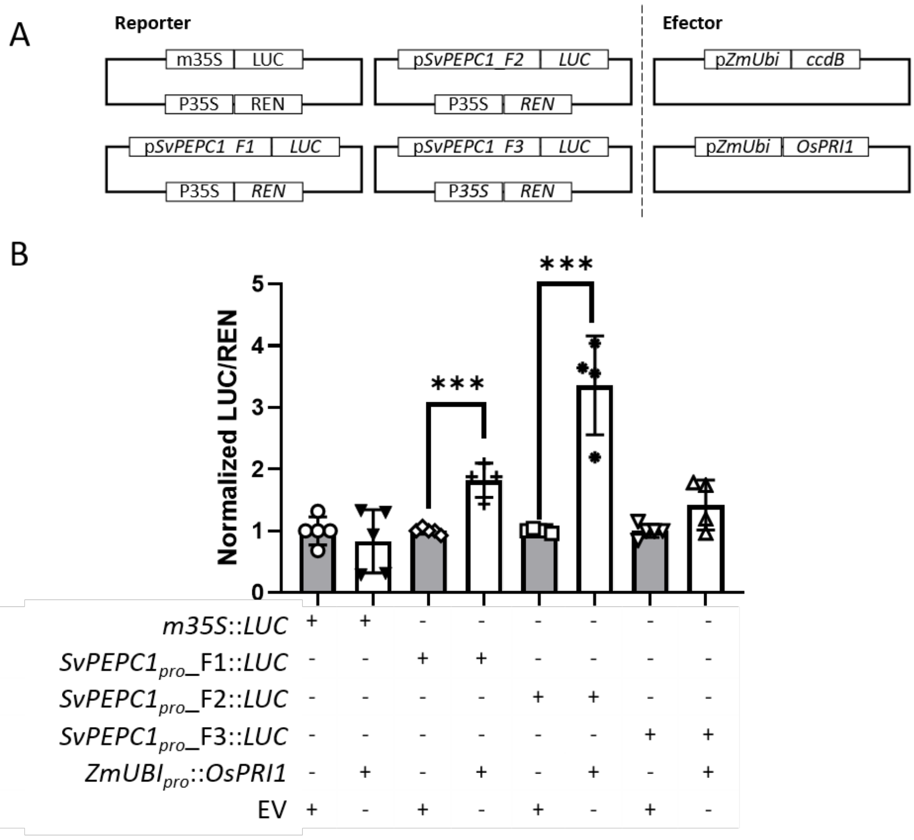


Figure 1. **OsPRI1 binds to the *SvPEPC1* promoter**. (A) Schematic representation of three overlapping fragments of the *SvPEPC1* promoter used as baits in a Yeast one-hybrid (Y1H) screening of a rice cDNA expression library. (B) Direct Y1H assay to test the interaction of OsPRI1 with the three *SvPEPC1* promoter fragments. Increasing 3-AT concentrations were used to eliminate yeast growth caused by non-specific interactions. Empty pGAD vector and pGAD

containing an unrelated bHLH (OsPIF14) were used as negative controls. (C) Electrophoretic Mobility Shift Assay performed using radiolabelled probe (“Hot”) containing a N-Box motif (CACGAG), positioned as represented in panel A (probe 14 in Fig. S2), and a mutated form of the motif. Purified MBP-OsPRI1 and MBP proteins alone were used to test the interaction. Non-labelled DNA (“Cold”) was used for competition assays to confirm the specificity of the interaction.



m35S::*LUC*	+	+	-	-	-	-	-	-
SvPEPC1$_{pro}$_F1::*LUC*	-	-	+	+	-	-	-	-
SvPEPC1$_{pro}$_F2::*LUC*	-	-	-	-	+	+	-	-
SvPEPC1$_{pro}$_F3::*LUC*	-	-	-	-	-	-	+	+
ZmUBI$_{pro}$::*OsPRI1*	-	+	-	+	-	+	-	+
EV	+	-	+	-	+	-	+	-

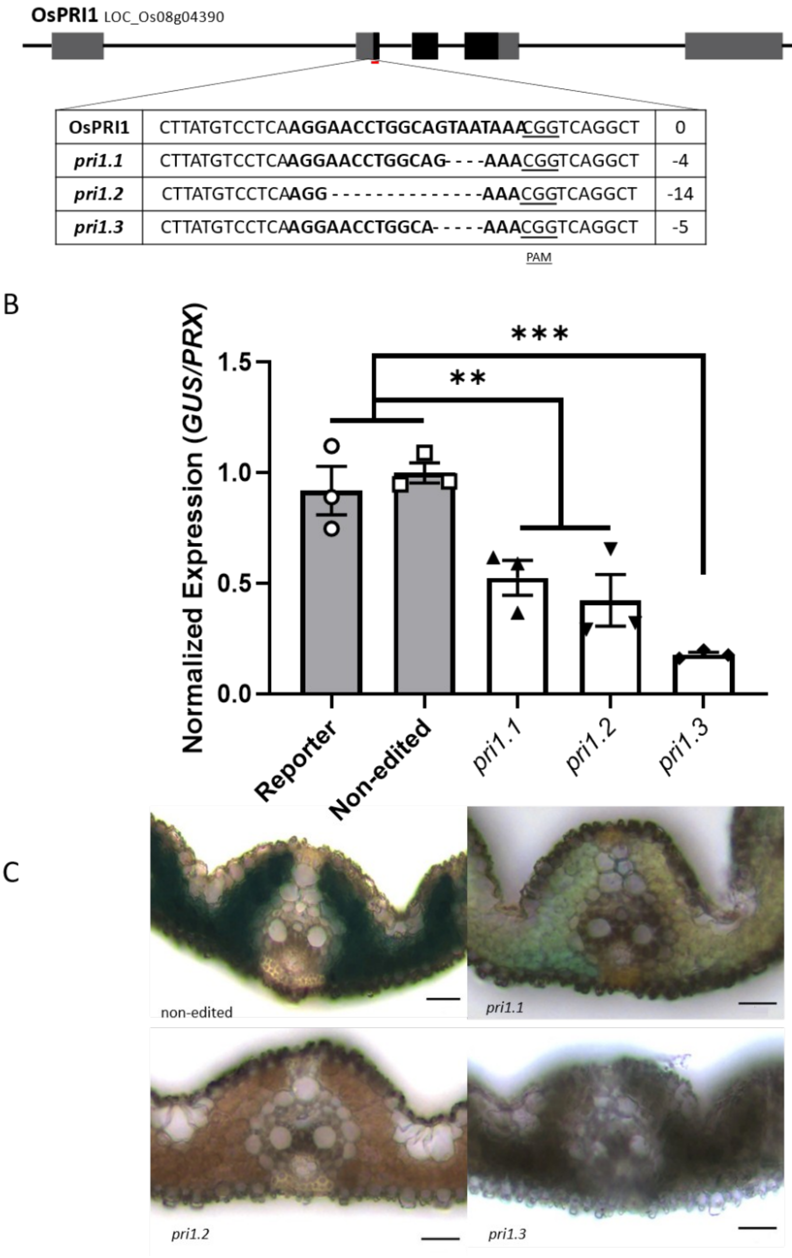
Figure 2. **OsPRI1 activates the *SvPEPC1* promoter in rice protoplasts.** (A) Schematics of constructs transformed into rice protoplasts. Each of the *SvPEPC*$_{pro}$ fragments (F1 to F3) were fused to a minimal 35S promoter to drive expression of the *Luciferase* (*LUC*) gene. The reporter constructs also contained the *Renilla* (*REN*) gene under the control of a complete 35S promoter. Reporter constructs containing just the minimal 35S promoter upstream of *LUC* were used as controls. Effector plasmids contained *OsPRI1* driven by the maize ubiquitin promoter (*ZmUbi*$_{pro}$) or empty vector (EV) as a control. (B) Constructs were co-transformed into rice protoplasts in the combinations indicated. Transactivation is shown as the ratio of LUC/REN activity for each protoplast transformation. Data was normalized for each reporter construct transformed with the empty vector. Statistical significance was tested using unpaired t-test (***:$p<0.001$) (n= 4-5).

OsPRI1 regulates the *SvPEPC1* promoter *in planta*

Given that *in vitro* and transient assays were indicative of a functional interaction between OsPRI1 and *SvPEPC*$_{pro}$, we next sought to determine whether the interaction occurred *in planta*. It is notable that transcription factors have previously been identified as binding to different C_4 *PEPC1* promoters, but their function has never been validated *in planta* (Hibberd & Covshoff, 2010). In order to investigate the function of OsPRI1 regulating *SvPEPC1*$_{pro}$ activity in rice, we have used a rice reporter line for the *Setaria viridis PEPC1* promoter (*SvPEPC1*$_{pro}$*::GUS*) to edit OsPRI1. CRISPR-Cas9 was then used to obtain *Ospri1* loss of function mutants in the *SvPEPC1*$_{pro}$*::GUS* background. Three independent *ospri1* mutations were obtained, each causing frameshifts that are predicted to form truncated proteins lacking the bHLH domain (Fig 3A and S3). To determine the impact of loss of OsPRI1 function on *SvPEPC*$_{pro}$ activity, *GUS* reporter gene expression was assayed by quantitative RT-PCR. Fig. 3B shows that *GUS* transcript levels were reduced by half relative to wild-type in all of the *Ospri1* mutant lines. Substantially less GUS activity was also observed in histological assays of leaf sections (Fig.3C). However, some enzyme activity was detectable in mesophyll cells of *Ospri1* mutant plants (Fig. S4). When photosynthetic efficiency was assessed, only one *Ospri1* mutant line showed changes as compared to wild type, indicating that the trait is not related with the *Ospri1* mutation but rather with the transgene insertion or with somaclonal variation resulting from the tissue-culture based rice transformation (Fig. S3). This result reveals that OsPRI1 is not involved in rice photosynthetic efficiency. Altogether, our results show that OsPRI1 is an important activator of *SvPEPC1*$_{pro}$ *in planta*, but loss of function is not sufficient to completely disrupt *SvPEPC1*$_{pro}$ activation or to impair cell-specificity, and OsPRI is not necessary for the C_3 photosynthesis.

OsPRI1 does not interact with the rice *PEPC* promoter

To assess how OsPRI1 and the *cis*-elements where it binds in the *SvPEPC1* promoter have been modified during the evolution of C_4 photosynthesis, we first identified *PEPC1* promoter sequences from two C_4 plants with independent C_4 origins (*Zea mays* and *Setaria viridis*), a C_3 member of the PACMAD clade (*Dichanthelium oligosanthes*), and rice, a member of the BEP clade. For rice, the promoter of both *OsPEPC1* and *OsPEPC4* were included in the analysis because *OsPEPC1* has been identified as the closest in sequence to C_4 PEPCs (Yamamoto *et al.*, 2022) and *OsPEPC4* is specifically expressed in mesophyll-cells (Masumoto *et al.*, 2010). We also identified OsPRI1 homologues in *Z. mays* (GRMZM2G017586 and GRMZM2G093744), *S. viridis* (Sevir.6G068300) and *D. oligosanthes* (OEL18192) (Fig. S5). Y1H assays were then carried out using fragments of each *PEPC* promoter as bait (Fig. 4A) in combination with each of the PRI1 proteins. Fig. 4B shows that OsPRI1 does not interact with the *OsPEPC* or *DoPEPC* promoter fragments, suggesting that the interaction between OsPRI1 and $SvPEPC_{pro}$ resulted from a C_4 innovation associated with *Setaria viridis*, and possibly with other C_4 species in the same evolutionary clade, that changed *cis*-elements in the promoter. On the basis that *cis*-elements in the $SvPEPC_{pro}$ are a C_4 innovation, binding of OsPRI1, SvPRI1 and ZmPRI proteins to those sequences suggests little variation in the transcription factor activity as C_4 plants evolved. However, lack of $SvPEPC_{pro}$ binding by the DoPRI1 protein is an anomaly because *Setaria viridis* is more closely related to *Dichanthelium oligosanthes* than rice, and the SvPRI1 protein binds to the $DoPEPC_{pro}$. Finally, failure of any of the PRI proteins to bind the $ZmPEPC_{pro}$, suggests that *cis*-changes in the promoter occurred in a lineage specific manner.



OsPRI1	CTTATGTCCTCA**AGGAACCTGGCAGTAATAAA**CGGTCAGGCT	0
pri1.1	CTTATGTCCTCA**AGGAACCTGGCAG- - - -AAA**CGGTCAGGCT	-4
pri1.2	CTTATGTCCTCA**AGG - - - - - - - - - - - - - -AAA**CGGTCAGGCT	-14
pri1.3	CTTATGTCCTCA**AGGAACCTGGCA- - - - -AAA**CGGTCAGGCT	-5

Figure 3. OsPRI1 activates the *SvPEPC1* promoter in rice. (A) Schematic representation of *OsPRI1* gene showing exons (grey boxes), the bHLH motif (black boxes) and three independent mutations with indels that create frameshift mutations. Single guide RNAs are highlighted in bold. (B) *GUS* transcript levels were analysed by RT-qPCR as a proxy for *SvPEPC1* promoter activity.Reporter indicates the *SvPEPC1*$_{pro}$::*GUS* reporter line, which was used as the background to generate the *ospri1* mutants, Non-edited indicates *SvPEPC1*$_{pro}$::*GUS* lines that were transformed with CRISPR-Cas9 but that were not edited by Cas9. Statistical significance was tested using an unpaired t test (** p<0.01; *** p<0.001) (n = 3). (C) Transverse leaf sections from different rice lines (non-edited and *Ospri1* mutants) showing GUS activity in the mesophyll cells. Leaf sections were incubated for 1 h at 37 °C in staining solution and cleared before imaging.

Discussion

OsPRI, a TF regulating the C_4 *SvPEPC* promoter activity in rice

Recent attempts to engineer C_4 metabolism into rice demonstrate that C_4 *PEPC1* promoters are an important biotechnological tool, since they drive mesophyll-specific expression of adjacent genes. Nevertheless, the expression levels obtained with these promoters was lower than that found in C_4 plants (Ermakova *et al.*, 2021). This observation highlights the importance of understanding the gene regulatory networks upstream of promoter function, in order to better manipulate expression levels in transgenic plants. To date, a single rice transcription factor (TF), OsbHLH112, has been shown to regulate a C_4 *PEPC1* promoter (Górska et al., 2019). OsbHLH112 was shown to bind an E-box motif in the *ZmPEPC1* promoter and to activate its activity. In this work, we were able to identify a new rice TF, OsPRI1, which binds to a C_4 *PEPC1* promoter. We have shown that OsPRI1 binds the the *SvPEPC1* promoter through a N-box motif and that it works as an activator (Fig.1-2), albeit not on its own. Both, the transcript level of the reporter gene as well as the corresponding GUS activity in *Ospri1* mutants, indicate that other activators may also modulate *SvPEPC1* promoter activity (Fig. 3; Fig. S4). We thus hypothesize that OsPRI1 is required to promote high levels of *SvPEPC1* promoter activity in rice rather than being a regulator of cell-specific gene expression. It is worth to note that elevated *PEPC* expression level is one of the key traits in C_4 plants, but it is considered to have preceded C_4 evolution (Sage, 2004).

The PRI1 network may have been co-opted during evolution to regulate C_4 photosynthesis in *Setaria viridis*

C_4 gene expression patterns are proposed to have evolved using pre-existing gene networks, with changes in promoter sequences underpinning innovations rather than changes in the pre-existing gene networks (Matsuoka *et al.*, 1994; Gupta *et al.*, 2020). We tested whether this hypothesis was true for the evolution of PRI1-*PEPC1* promoter interactions. We observed that PRI1-DNA interactions were dependent on both *trans* and *cis* sequence composition (Fig.4). Moreover, our results indicate that PRI1 has been co-opted as a *PEPC1* regulator in the *Setaria viridis* lineage but not in the *Zea mays* lineage. Since the N-box where OsPRI1 binds in the *SvPEPC1* promoter is conserved in the *PEPC1* promoter from *Urochloa maxima* (Fig. S6), which is known to drive mesophyll-specific gene expression in rice (Gupta *et al.*, 2020), we hypothesize that OsPRI1 can bind the *UmPEPC1* promoter. However, the *Panicum miliceum PEPC1* promoter, which also drives mesophyll-specific expression in rice, does not contain the same N-box (Fig. S6). What more, OsPRI1 does not bind to the *OsPEPC4* promoter, which is also known to drive mesophyll-specific expression in rice (Masumoto *et al.*, 2010). Together these results suggest that the evolution of high levels of mesophyll cell-specific *PEPC* expression is likely to have occurred through the co-option of different *cis* and *trans* factors in the independent C_4 evolution events.

OsPRI1 was described as a major player in iron response networks in rice, regulating several iron response genes and being regulated by a major iron sensor (Zhang *et al.*, 2017). Notably, iron levels are important regulators of proteins involved in the electron transport chain of photosystems in C_3 plants and ILR3, the OsPRI1 homologue in Arabidopsis, regulates levels of photosystem proteins in the Cytb6f and PSI complexes (Li *et al.*, 2019). We show that loss of OsPRI1 function has no influence on the carbon assimilation capacity or on photosystem efficiency in rice (Fig. S3). In addition, we observed that iron levels do not affect $\textit{SvPEPC1}_{\textit{pro}}$ activity in rice (Fig. S3). We thus conclude that the relationship between ILR3 and photosystem function is not conserved in rice, and that if any relationship exists between PRI1 and photosynthesis it might be a C_4 novelty.

It would now be interesting to understand how the mis-expression of PRI1 homologues in C_4 plants affects the function of the C_4 cycle. We have attempted to obtain PRI1 loss-of-function mutants in *Setaria viridis*, using the CRISPR/cas9 system, but none of the transgenic plants obtained were edited. To overcome this issue and to get a better understanding of the gene regulatory networks where PRI1 is involved, it is important to overexpress PRI1 in both the reporter rice line ($\textit{SvPEPC1}_{\textit{pro}}\textit{::GUS}$) and in *Setaria viridis*. These tools would be extremelly valuable to unveil the molecular mechanisms underlying the evolution of gene regulatory network from C_3 to C_4.

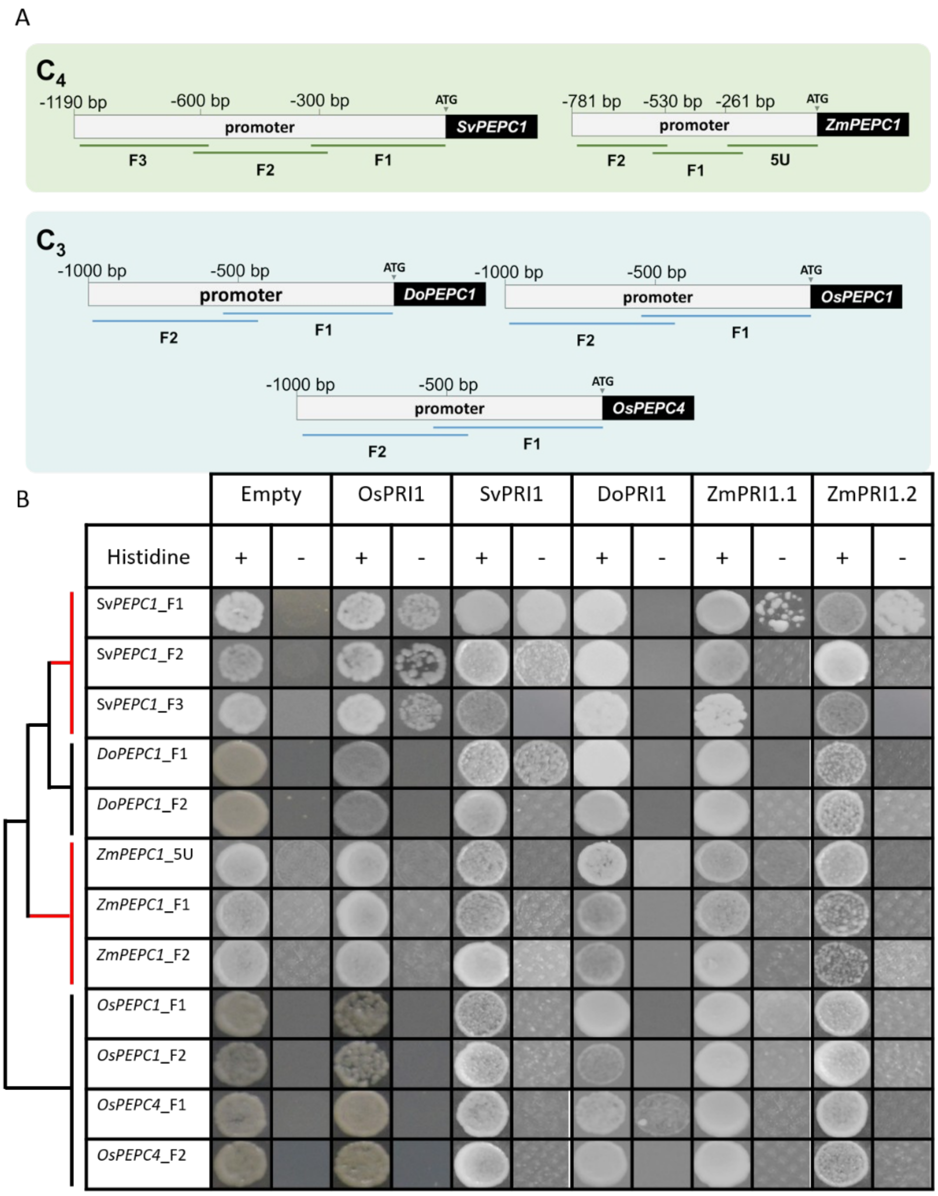


Figure 4. The interaction between PRI1 and *PEPC* promoters was not pre-established in C_3 ancestors. (A) Schematic representation of the *PEPC* promoters in C_4 (*Setaria viridis* and *Zea mays*) and C_3 (*Oryza sativa* and *Dichanthelium oligosanthes*) grasses used in this study. (B) Interaction between the *PEPC* promoters and PRI1 homologues was tested using a direct Yeast one-hybrid approach. Results are shown in the presence of histidine (His +) and in the absence of histidine (His -) with the addition of 3-AT (5mM for *DoPEPC1*_F1, *ZmPEPC1*_5U; 10mM for *OsPEPC4*_F1) when required to deplete non-specific interactions. Red lines in the cladogram indicate C_4 branches.

Acknowledgments

The authors would like to acknowledge Dr Steve Kelly and Dr David Emms for help using OrthoFinder. The authors would like to thank Dr Bruno Catarino for insights into the work. The authors would like to acknowledge the Protein purification services of the ITQB for assistance in recombinant protein purification. This work was funded by Fundação para a Ciência e a Tecnologia (Portugal) through the R&D unit “GREEN-IT Bioresources for Sustainability” (UIDB/04551/2020 & UIDP/04551/2020) and the fellowships to PC (PD/BD/128403/2017) and CG (2020.08000.BD), and CEEC contract to TFL (through Scientific Employment Stimulus contract (CEECIND/03641/2017)). It was also was supported by the European Union project 3to4 (Grant agreement no: 289582).

Competing interests

The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

Authors contributions

PC, JAL, and NJMS conceived and designed the study; PC conducted the experiments; IG and CG assisted in the Y1H experiments; TL assisted in the rice protoplast isolation and transformation; DV assisted in the rice genetic transformation and microscopy; PC, JAL, and NJMS analysed the data; PC wrote the manuscript; JAL and NJMS reviewed the manuscript. All authors read and approved the final manuscript.

References

Akyildiz M, Gowik U, Engelmann S, Koczor M, Streubel M, Westhoff P. **2007**. Evolution and function of a cis-regulatory module for mesophyll-specific gene expression in the C4 dicot Flaveria trinervia. *Plant Cell* **19**: 3391–3402.

von Caemmerer S, Quick WP, Furbank RT. **2012**. The development of C4 rice: Current progress and future challenges. *Science* **336**: 1671–1672.

Cheng H, Liu L, Zhou Y, Deng K, Ge Y, Hu X. **2023**. TSPTFBS 2.0: trans-species prediction of transcription factor binding sites and identification of their core motifs in plants. *Frontiers in Plant Science* **14**.

Ermakova M, Arrivault S, Giuliani R, Danila F, Alonso-Cantabrana H, Vlad D, Ishihara H, Feil R, Guenther M, Borghi GL, *et al.* **2021**. Installation of C$_4$ photosynthetic pathway enzymes in rice using a single construct. *Plant Biotechnology Journal* **19**: 575–588.

Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. **2010**. Food Security: The Challenge of Feeding 9 Billion People. *Science* **327**: 812–818.

Górska AM, Gouveia P, Borba AR, Zimmermann A, Serra TS, Carvalho P, Lourenço TF, Oliveira MM, Peterhänsel C, Saibo NJM. **2021**. ZmOrphan94 Transcription Factor Downregulates ZmPEPC1 Gene Expression in Maize Bundle Sheath Cells. *Frontiers in plant science* **12**: 559967.

Górska AM, Gouveia P, Borba AR, Zimmermann A, Serra TS, Lourenço TF, Margarida Oliveira M, Peterhänsel C, Saibo NJM. **2019**. ZmbHLH80 and ZmbHLH90 transcription factors act antagonistically and contribute to regulate PEPC1 cell-specific gene expression in maize. *Plant Journal* **99**: 270–285.

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P. **2004**. cis-regulatory elements for mesophyll-specific gene expression in the C4 plant Flaveria trinervia, the promoter of the C4 phosphoenolpyruvate carboxylase gene. *Plant Cell* **16**: 1077–1090.

Grass Phylogeny Working Group II. **2012**. New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist* **193**: 304–312.

Gupta S Das, Levey M, Schulze S, Karki S, Emmerling J, Streubel M, Gowik U, Paul Quick W, Westhoff P. **2020**. The C4Ppc promoters of many C4 grass species share a common regulatory mechanism for gene expression in the mesophyll cell. *Plant Journal* **101**: 204–216.

Hatch MD, Slack CR. **1966**. Photosynthesis by sugar-cane leaves. A new carboxylation reaction and the pathway of sugar formation. *The Biochemical journal* **101**: 103–111.

Hibberd JM, Covshoff S. **2010**. The Regulation of Gene Expression Required for C 4 Photosynthesis. *Annual Review of Plant Biology* **61**: 181–207.

Ku MSB, Agarie S, Nomura M, Fukayama H, Tsuchida H, Ono K, Hirose S, Toki S, Miyao M, Matsuoka M. **1999**. High-level expression of maize phosphoenolpyruvate carboxylase in transgenie rice plants. *Nature Biotechnology* **17**: 76–81.

Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, *et al.* **2006**. Genome-Wide Analysis of Basic/Helix-Loop-Helix Transcription Factor Family in Rice and Arabidopsis. *Plant Physiology* **141**: 1167–1184.

Li M, Zhang J, Jin H, Feng D, Wang J, Wang H-B, Liu B. **2019**. The Iron Deficiency Response Regulators IAA-LEUCINE RESISTANT3 and bHLH104 Possess Different Targets and Have Distinct Effects on Photosynthesis in Arabidopsis. *Journal of Plant Biology* **62**: 109–119.

Maclean J, Hardy B, Hettel G. **2013**. *Rice Almanac: Source book for one of the most important economic activities on earth*. IRRI.

Masumoto C, Miyazawa SI, Ohkawa H, Fukuda T, Taniguchi Y, Murayama S, Kusano M, Saito K, Fukayama H, Miyao M. **2010**. Phosphoenolpyruvate carboxylase intrinsically located in the chloroplast of rice plays a crucial role in ammonium assimilation. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 5226–5231.

Matsuoka M, Kyozuka J, Shimamoto K, Kano‐Murakami Y. **1994**. The promoters of two carboxylases in a C4 plant (maize) direct cell‐specific, light‐regulated expression in a C3 plant (rice). *The Plant Journal* **6**: 311–319.

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ. **2013**. Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research* **23**: 1233–1236.

Sage RF. **2004**. The evolution of C_4 photosynthesis. *New Phytologist* **161**: 341–370.

Sage RF, Sage TL, Kocacinar F. **2012**. Photorespiration and the Evolution of C 4 Photosynthesis . *Annual Review of Plant Biology* **63**: 19–47.

Stockhaus J, Schlue U, Koczor M, Chitty JA, Taylor WC, Westhoff P. **1997**. The promoter of the gene encoding the C4 form of phosphoenolpyruvate carboxylase directs mesophyll-specific expression in transgenic C4 Flaveria spp. *Plant Cell* **9**: 479–489.

Svensson P, Bläsing OE, Westhoff P. **2003**. Evolution of C4 phosphoenolpyruvate carboxylase. *Archives of Biochemistry and Biophysics* **414**: 180–188.

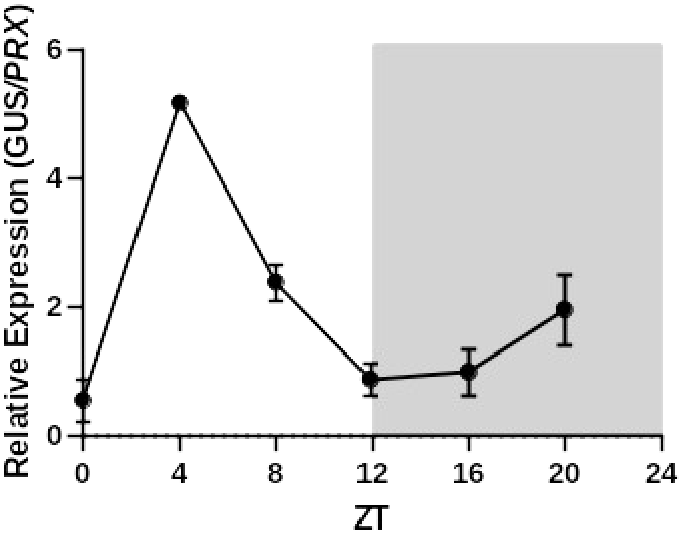
Yamamoto N, Tong W, Lv B, Peng Z, Yang Z. **2022**. The Original Form of C4-Photosynthetic Phosphoenolpyruvate Carboxylase Is Retained in Pooids but Lost in Rice. *Frontiers in Plant Science* **13**.

Yan W, Li Z, Pian C, Wu Y. **2022**. PlantBind: an attention-based multi-label neural network for predicting plant transcription factor binding sites. *Briefings in Bioinformatics* **23**.

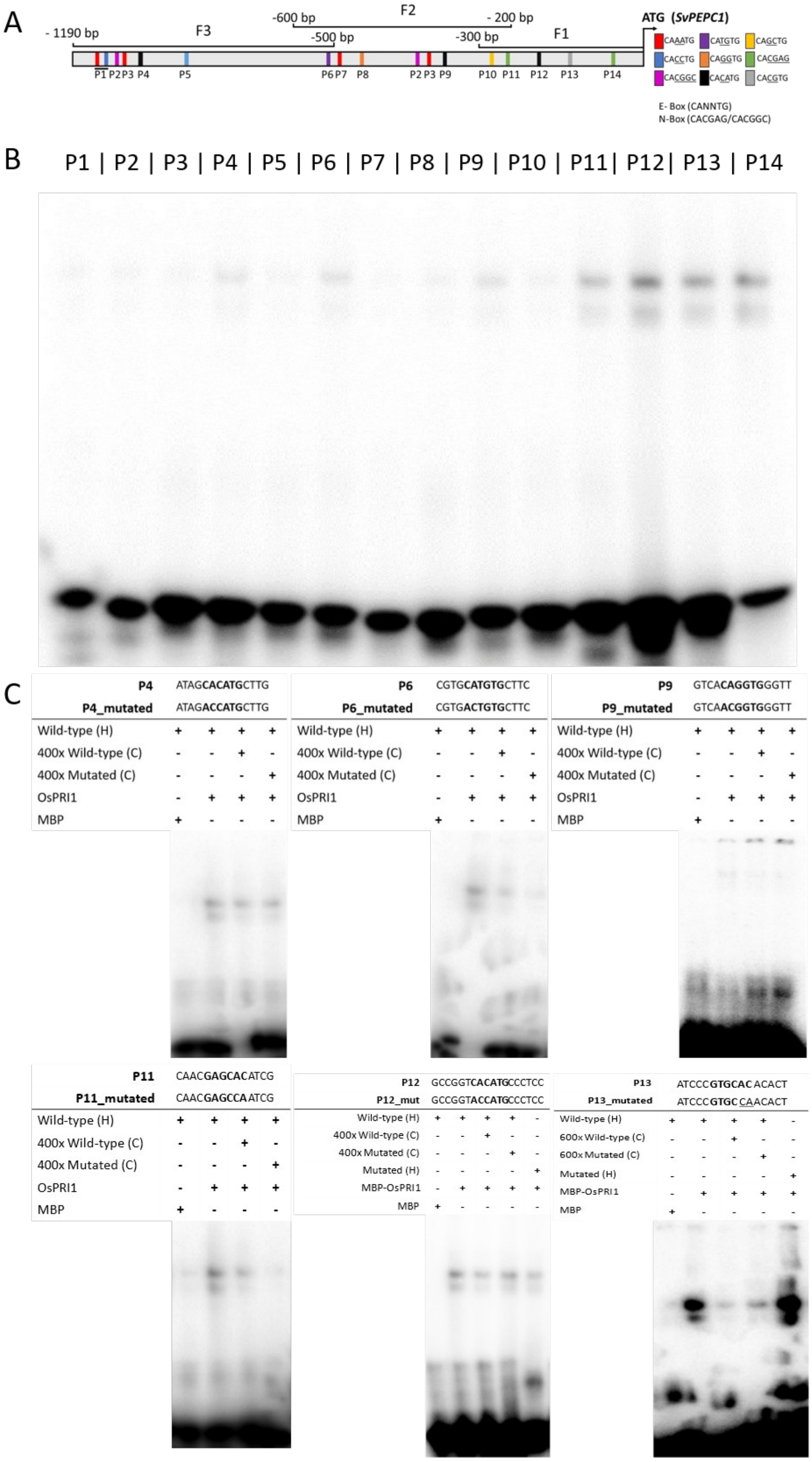
Yanagisawa S, Sheen J. **1998**. *Involvement of Maize Dof Zinc Finger Proteins in Tissue-Specific and Light-Regulated Gene Expression*.

Zhang HM, Li Y, Yao XN, Liang G, Yu D. **2017**. POSITIVE REGULATOR OF IRON HOMEOSTASIS 1 (OsPRI1) positively regulates iron homeostasis in rice. *Plant Physiology*.

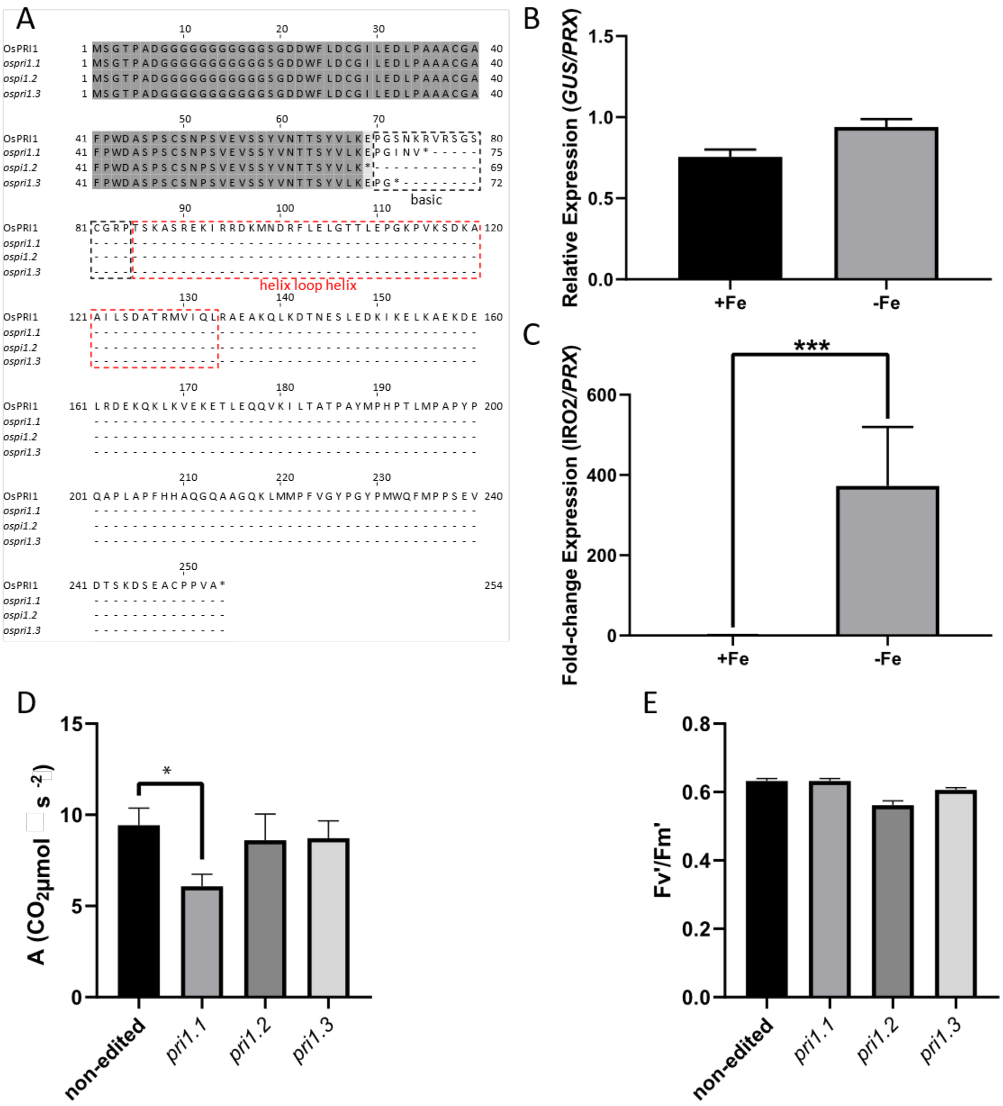
Supporting Information



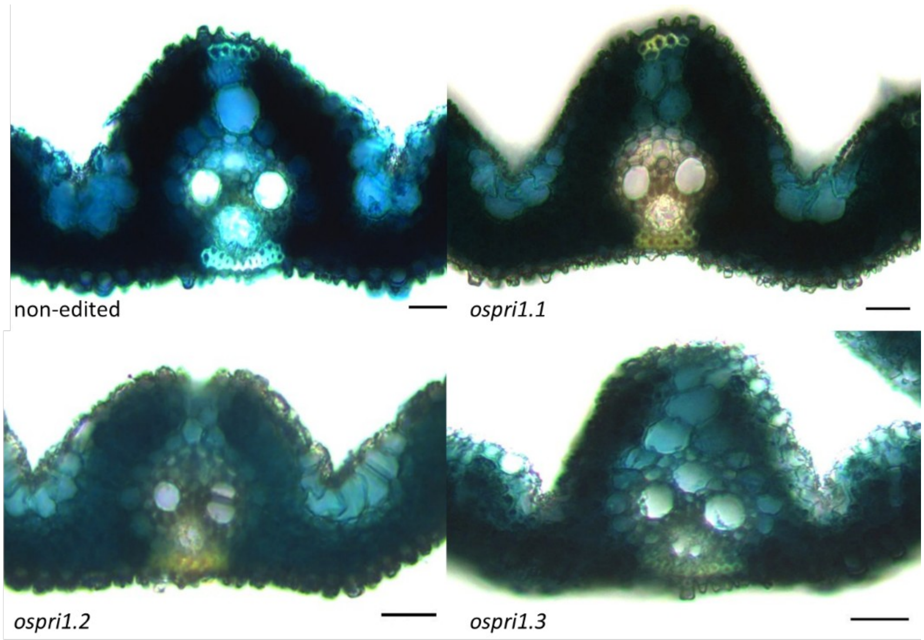
Supplementary Figure 1. *SvPEPC$_{pro}$* activity over a 24h time period. *GUS* transcript levels were analysed by RT-qPCR as a proxy for *SvPEPC1* promoter activity. Samples were collected every 4h after light initiation.



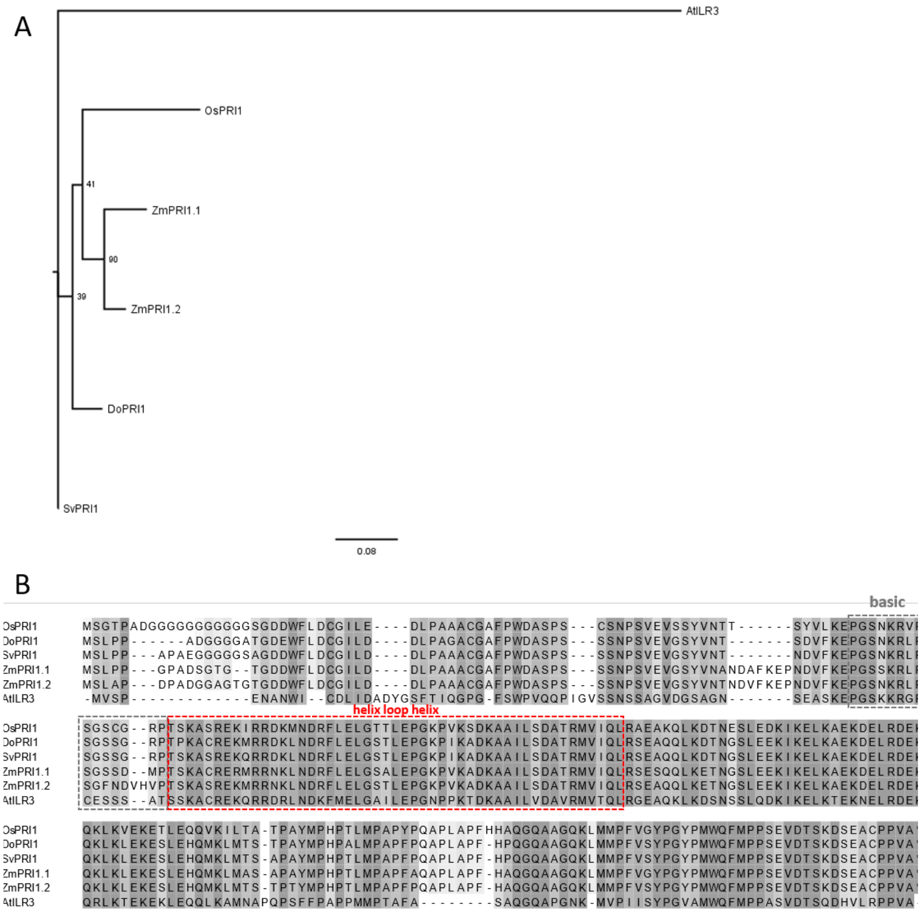
Supplementary Figure 2. Electrophoretic Mobility Shift Assay (EMSA) to analyse the interaction of OsPRI1 with the canonical bHLH binding motifs in *SvPEPC1*$_{pro}$. (A) Schematic showing 17 canonical bHLH binding sites, 7 different E-boxes and 2 different N-boxes in the *SvPEPC1*$_{pro}$ promoter. When the cis-element and flanking regions are the same, the probe is the same. (B) EMSA assays with 14 different probes that include all of the identified elements in (A). Probes were radiolabelled and tested using recombinant MBP-OsPRI1. (C) EMSAs with probes 4, 6, 9, and 11-13 non-labelled and / or mutated. Labelled probe (H - Hot) and unlabelled probe (C - Cold) were used to assess the specificity of the binding.



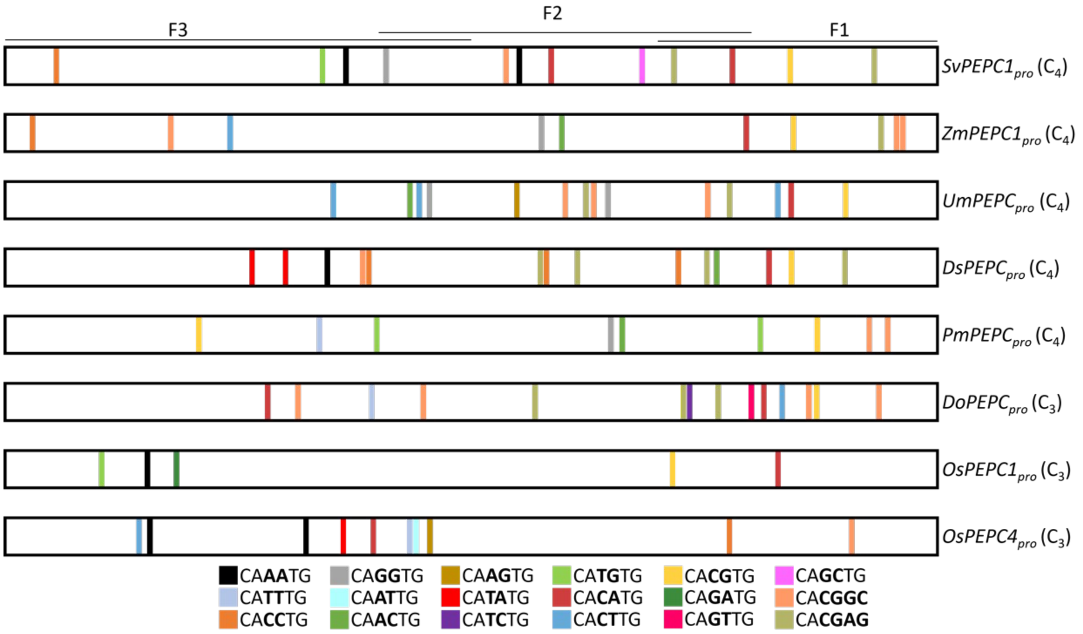
Supplementary Figure 3- OsPRI1 does not alter the photosynthetic efficiency of rice and *SvPEPC1*$_{pro}$ is not affected by iron levels in rice. (A) Alignment of proteins resulting from CRISPR-Cas9 mutagenesis showing absence of a functional bHLH motif in each case. (B-C) RT-qPCR of *GUS* expression (as a proxy for *SvPEPC1*$_{pro}$ activity) and *IRO2* in *SvPEPC1*$_{pro}$::*GUS* reporter lines under control and iron deficiency conditions. *IRO2* was used as a control for the iron deficiency response. (D-E) Photosynthetic measurements performed on *Ospri1* mutants under control conditions. For each plant, apparent photosynthesis (A) and the operating efficiency of PSII (Fv'/Fm') were measured (n=5)(* $p<0.05$).



Supplementary Figure 4. Loss of OsPRI1 function is not sufficient to completely disrupt *SvPEPC1*$_{pro}$ mesophyll-cell specificity. Transverse leaf sections of *SvPEPC1*$_{pro}$*::GUS; Ospri1* mutant lines stained for GUS activity. Sections were incubated in staining solution overnight at 37 °C and cleared before imaging.



Supplementary Figure 5. (A) Phylogenetic tree generated using the alignment in (B). The tree was created using IQtree tree with a JTT+G4 model and Maximum likelihood. (A) Sequence alignment of OsPRI1 homologues from *Arabidopsis thaliana*, *Setaria viridis*, *Zea mays* and *Dichanthelium oligosanthes*, generated with the MUSCLE algorithm. Basic domain is highlighted with a grey box and helix-loop-helix is highlighted with a red box.



Supplementary Figure 6. Analysis of canonical bHLH binding sites in PEPC promoters from C3 and C4 species. Canonical bHLH binding motifs were identified in the C4 PEPC promoters of *Setaria viridis* (*SvPEPC1*$_{pro}$), *Zea mays* (*ZmPEPC1*$_{pro}$), *Urochloa maxima* (*UmPEPC*$_{pro}$), *Digitaria sanguinalis* (*DsPEPC*$_{pro}$), and *Panicum miliaceum* (*PmPEPC*$_{pro}$). PEPC C3 promoters from *Dichanthelium oligosanthes* (*DoPEPC*$_{pro}$) and *Oryza sativa* (*OsPEPC1*$_{pro}$; *OsPEPC4*$_{pro}$) were also screened for the presence of canonical bHLH binding sites. C_4 promoters were selected based on the know activity in rice, as described in Gupta *et al.* (2020).

Supplementary Table 1- List of primers used in this study. Sequences in bold were added for recognition of either BP recombinase or restriction enzymes, as indicated in the main text.

Name	Sequence	Purpose	Gene identifier
promSvPEPC_Fw	**GGGGACAAGTTTGTACAAAAAAGCAGGCTTC**GGGGTTTGTGGCTGGC	Clone *SvPEPC1*$_{pro}$	Sevir.4G143500
promSvPEPC_Rv	**GGGGACCACTTTGTACAAGAAAGCTGGGTT**TGGCGTGGTGGGAAGCGAAG		
promSvPEPC_F1_Fw	**ATATGCGGCCGC**AGGCGATTGAGCAGTGACAA	Creating Y1H baits	Sevir.4G143500
promSvPEPC_F1_Rv	**ATATACTAGT**TGGCGTGGTGGGAAGCGAAG		
promSvPEPC_F2_Fw	**ATATGGCGCC**TTTTTCACAGTCAGTGAGCCA		
promSvPEPC_F2_Rv	**ATATACTAGT**ATGTGCTCGTTGCCTGGTTT		
promSvPEPC_F3_Fw	**ATATGCGGCCGC**GGGGTTTGTGGCTGGC		
promSvPEPC_F3_Rv	**ATATACTAGT**AACTCTGGAAGCACATGCAC		
m35S_Fw	**ATATGCGGCCGC**TGTGACATCTCCACTGACGT	Creating Dual Luciferase Reporter	
m35S_Rv	**ATATCCATGG**GGTGGTGGGCTGTCCTCTC		
promSvPEPC_F1_Fw	**ATATGGTACC**AGGCGATTGAGCAGTGACAA		Sevir.4G143500
promSvPEPC_F1_Rv	**ATATACTAGT**TGGCGTGGTGGGAAGCGAAG		
promSvPEPC_F2_Fw	**ATATGGTACC**TTTTTCACAGTCAGTGAGCCA		
promSvPEPC_F2_Rv	**ATATACTAGT**ATGTGCTCGTTGCCTGGTTT		
promSvPEPC_F3_Fw	**ATATGGTACC**GGGGTTTGTGGCTGGC		
promSvPEPC_F3_Rv	**ATATACTAGT**AACTCTGGAAGCACATGCAC		

OsPRI1_Fw	**GGGGACAAGTTTGTACAAAAAAGCAGGCTTC**ATGTCCGGTACCCCGGCGGAC	Cloning CDS into pDONR	LOC_Os08g04390
OsPRI1_Rv	**GGGGACCACTTTGTACAAGAAAGCTGGGTT**ACGCGACAGGCGGGCACGCT		
promOsPEPC1_F1_Fw	**ATATGCGGCCGC**TAGTTAACTATAGAAATATACTACAC	Creating Y1H baits	LOC_Os02g14770
promOsPEPC1_F1_Rv	**ATATACTAGT**TAATGTTAGAAGCGTGGAG		
promOsPEPC1_F2_Fw	**ATATGCGGCCGC**CGGTAGTGCACCTTAGGCC		
promOsPEPC1_F2_Rv	**ATATACTAGT**CTCCCTCTCTCTCGCTTACA		
promOsPEPC4_F1_Fw	**ATATGCGGCCGC**TAAAAATGCTGCTAATCAGTTTAAG		LOC_Os01g11054
promOsPEPC4_F1_Rv	**ATATACTAGT**GAGAGGAAGACAATTTGCTTCA		
promOsPEPC4_F2_Fw	**ATATGCGGCCGC**TGGGAACTGGGAAATGGTCAT		
promOsPEPC4_F2_Rv	**ATATACTAGT**GTCTCATCTCCGCTGTGG		
promDoPEPC1_F1_Fw	**ATATGCGGCCGC**GCACGTCGGCTGCCCATTAG		Do021545
promDoPEPC1_F1_Rv	**ATATACTAGT**TAGACGGTGCTCACGATGGG		
promDoPEPC1_F2_Fw	**ATATGCGGCCGC**CTTCCACATTTGCAGTAGGGAG		
promDoPEPC1_F2_Rv	**ATATACTAGT**GGCTTGGCGGGAAGAGAAG		
GUS_RT-PCR_Fw	GCCCTGATGCTCCATCACTT (Ta=58 °C)	RT-PCR	
GUS_RT-PCR_Rv	ATACCGAAAGGTTGGGCAGG	RT-PCR	
PRXIIc_RT-PCR_Fw	AAGCCAAGGGTGTAGACGAC (Ta=58 °C)	RT-PCR	LOC_Os01g4842
PRXIIc_RT-PCR_Rv	TATGACTTTGCCCACGCCTT	RT-PCR	
OsPRI1_sgRNA_Fw	**GGCA**AGGAACCTGGCAGTAATAAA	sgRNA for CRISPR-Cas9	LOC_Os08g04390
OsPRI1_sgRNA_Rv	**AAC**TTTATTACTGCCAGGTTCCT	sgRNA for CRISPR-Cas9	
OsPRI1_Indel_Fw	TGCTCAAGCAAGTAGCTACGAGGAGT	Sequence CRISPR mutation	LOC_Os08g04390
OsPRI1_Indel_Fw_Seq	GGTATCTTATTGTGCCATTTTTTGAG	Sequence CRISPR mutation	
OsPRI1_Indel_Rv	TGCTCGCAACATCCCAACCA	Sequence CRISPR mutation	
OsIRO2_RT-PCR_Fw	TCATCGACTTCCTCCGGATT (Ta=60 °C)	RT-PCR	LOC_Os01g72370
OsIRO2_RT-PCR_Rv	TTTCACAAAATGCTGGGCAC	RT-PCR	

Supplementary Table 2- List of probes used for the EMSA assays. Sequences in bold highlight the bHLH *cis*-element motif. Underlined sequences indicate mutations in the *cis*-element.

Name	Sequence
P1_Fw	CCGTTGCG**CAAATG**GA**CACCTG**GAGCGATA**G**
P1_Rv	TATCGCTC**CAGGTG**TC**CATTTG**CGCAACGG**C**
P2_Fw	ATGTGATCCTC**CACGGC**GTTGCCACAAA**G**
P2_Rv	**C**TTTGTGGCAAC**GCCGTG**GAGGATCACAT**C**
P3_Fw	ACGGCGTTGCCA**CAAATG**TCACTGATCATC**G**
P3_Rv	GATGATCAGTGA**CATTTG**TGGCAACGCCGT**C**
P4_Fw	GTTGTCCTATAG**CACATG**CTTGAGTCCAAG**G**

P4_Rv	CTTGGACTCAAG**CATGTG**CTATAGGACAAC**C**
P5_Fw	GTGACAAAGTAT**GTCCAC**AGCACAATCCGT**G**
P5_Rv	ACGGATTGTGCT**GTGGAC**ATACTTTGTCAC**C**
P6_Fw	GACAGCACCGTG**CATGTG**CTTCCAGAGTTG**G**
P6_Rv	CAACTCTGGAAG**CACATG**CACGGTGCTGTC**C**
P7_Fw	GTTGCCAACGG**CAAATG**AGCAGCAGTGTG**G**
P7_Rv	CACACTGCTGCT**CATTTG**CCGTTGGCAAC**C**
P8_Fw	TGTCCTATATAG**CACATG**CTTGAGTCCAAG**G**
P8_Rv	**C**CTTGGACTCAAG**CATGTG**CTATATAGGACA**C**
P9_Fw	GGGCTCGGTCA**CAGGTG**GGTTTCGTCTCC**G**
P9_Rv	GGAGACGAAACC**CACCTG**TGACCGAGCCC**C**
P10_Fw	AGTGACAAGTAC**GTCGAC**AGCACAACCCAG**G**
P10_Rv	CTGGGTTGTGCT**GTCGAC**GTACTTGTCACT**C**
P11_Fw	AAACCAGGCAAC**GAGCAC**ATCGCTCCGCAG**G**
P11_Rv	CTGCGGAGCGAT**GTGCTC**GTTGCCTGGTTT**C**
P12_Fw	TTCGCCGCCGGT**CACATG**CCCTCCGCTTGA**G**
P12_Rv	TCAAGCGGAGGG**CATGTG**ACCGGCGGCGAA**C**
P13_Fw	GCGTGCAATCCC**GTGCAC**ACACTCGCCGAC**G**
P13_Rv	GTCGGCGAGTGT**GTGCAC**GGGATTGCACGC**C**
P14_Fw	GCCTCGCTCGAC**CTCGTG**CCACCGCCACAGC**G**
P14_Rv	GCTGTGGCGGTGG**CACGAG**GTCGAGCGAGGC**C**
P11_Fw_mut	ACGTGAAACCAGGCAAC**GAGC**CAATCGCTCCGCAGG
P11_Rv_mut	ATCACCTGCGGAGCGATTGGCTCGTTGCCTGGTTTC
P12_Fw_mut	ACGTGTTCGCCGCCGGTAC**CATG**CCCTCCGCTTGAG
P12_Rv_mut	ATCACTCAAGCGGAGGGCATGGTACCGGCGGCGAAC
P13_Fw_mut	ACGTGGCGTGCAATCCC**GTGC**CAACACTCGCCGACG
P13_Rv_mut	ATCACGTCGGCGAGTGTTGGCACGGGATTGCACGCC
P14_Fw_mut	ACGTGGCCTCGCTCGAC**CTCG**GTCCACCGCCACAGCG
P14_Rv_mut	ATCACGCTGTGGCGGTGGACCGAGGTCGAGCGAGGCC

Supplementary Table 3- List of interactors identified In the Y1H screening.

Loci	**Predicted function**	**Number of identified clones**
LOC_Os03g39610	chlorophyll A-B binding protein	1
LOC_Os07g06550	Expressed protein	1
LOC_Os06g50380	protein phosphatase 2C	1
LOC_Os05g41060	ADP-ribosylation factor	1
LOC_Os02g06640	ubiquitin family protein	1
LOC_Os01g46140	Expressed protein	2
LOC_Os04g41680	Chitinase family protein precursor	1

LOC_Os07g47580	ribosomal protein	2
LOC_Os12g12880	Expressed protein	2
LOC_Os04g53620	ubiquitin family protein	1
LOC_Os03g14690	vacuolar ATP synthase 98 kDa subunit	1
LOC_Os01g56490	ubiquitin carboxyl-terminal hydrolase	1
LOC_Os05g51590	N-rich protein	1
LOC_Os01g09540	HAD superfamily phosphatase	1
LOC_Os04g50960	Expressed protein	1
LOC_Os06g08840	RNA recognition motif containing protein	2
LOC_Os05g41640	phosphoglycerate kinase protein	1
LOC_Os08g04390	bHLH transcription factor (OsPRI1)	16