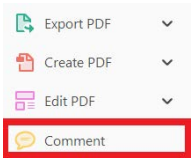
WILEY

USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

Required software to e-Annotate PDFs: Adobe Acrobat Professional or Adobe Reader (version 11 or above). (Note that this document uses screenshots from Adobe Reader DC.)
The latest version of Acrobat Reader can be downloaded for free at: http://get.adobe.com/reader/

Once you have Acrobat Reader open on your computer, click on the Comment tab (right-hand panel or under the Tools menu).



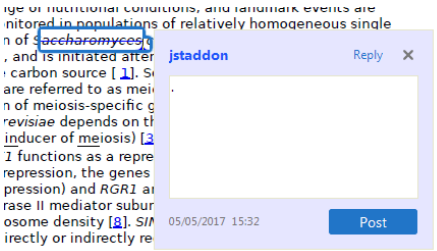
This will open up a ribbon panel at the top of the document. Using a tool will place a comment in the right-hand panel. The tools you will use for annotating your proof are shown below:

1. Replace (Ins) Tool – for replacing text.

Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it:

- Highlight a word or sentence.
- Click on .
- Type the replacement text into the blue box that appears.

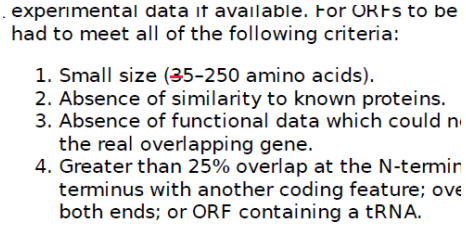


2. Strikethrough (Del) Tool – for deleting text.

Strikes a red line through text that is to be deleted.

How to use it:

- Highlight a word or sentence.
- Click on .
- The text will be struck out in red.

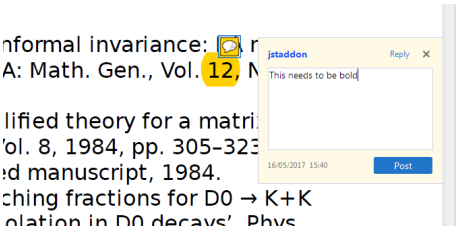


3. Commenting Tool – for highlighting a section to be changed to bold or italic or for general comments.

Use these 2 tools to highlight the text where a comment is then made.

How to use it:

- Click on .
- Click and drag over the text you need to highlight for the comment you will add.
- Click on .
- Click close to the text you just highlighted.
- Type any instructions regarding the text to be altered into the box that appears.

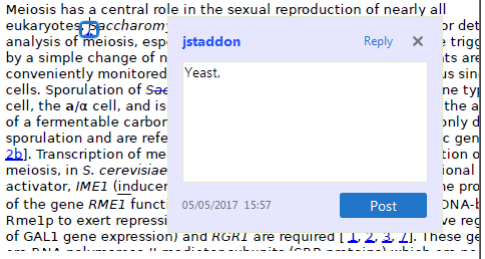


4. Insert Tool – for inserting missing text at specific points in the text.

Marks an insertion point in the text and opens up a text box where comments can be entered.

How to use it:

- Click on .
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the box that appears.



USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

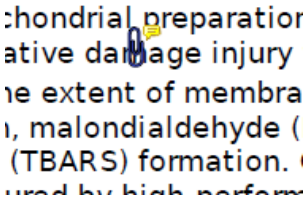
5. Attach File Tool – for inserting large amounts of text or replacement figures.

Inserts an icon linking to the attached file in the appropriate place in the text.

How to use it:

- Click on  .
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.

The attachment appears in the right-hand panel.

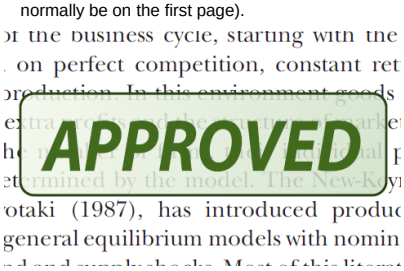


6. Add stamp Tool – for approving a proof if no corrections are required.

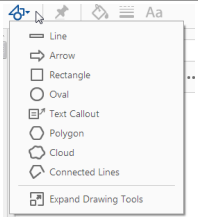
Inserts a selected stamp onto an appropriate place in the proof.

How to use it:

- Click on  .
- Select the stamp you want to use. (The Approved stamp is usually available directly in the menu that appears. Others are shown under *Dynamic*, *Sign Here*, *Standard Business*).
- Fill in any details and then click on the proof where you'd like the stamp to appear. (Where a proof is to be approved as it is, this would normally be on the first page).

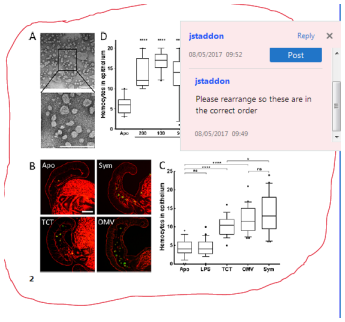


Drawing tools available on comment ribbon



7. Drawing Markups Tools – for drawing shapes, lines, and freeform annotations on proofs and commenting on these marks.

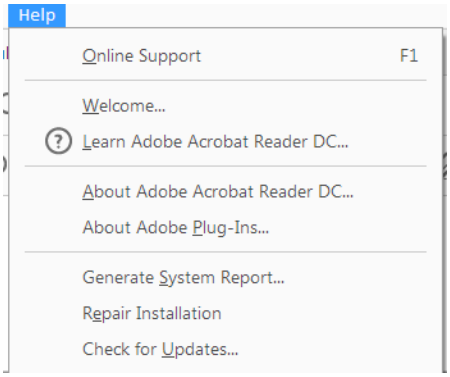
Allows shapes, lines, and freeform annotations to be drawn on proofs and for comments to be made on these marks.



How to use it:

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, right-click on shape and select *Open Pop-up Note*.
- Type any text in the red box that appears.

For further information on how to annotate proofs, click on the Help menu to reveal a list of further options:



WILEY

Author Query Form

Journal: NPH

Article: 15614/2018-27953

Dear Author,

During the copyediting of your manuscript the following queries arose.

Please refer to the query reference callout numbers in the page proofs and respond to each by marking the necessary comments using the PDF annotation tools.

Please remember illegible or unclear comments and corrections may delay publication.

Many thanks for your assistance.

Query reference	Query	Remarks
1	**AUTHOR: Please verify that the linked ORCID identifiers are correct for each author.**	
2	**AUTHOR: Please confirm that given names (blue) and surnames/family names (vermilion) have been identified correctly.**	
3	**AUTHOR: Please check that authors and their affiliations are correct.**	
4	**AUTHOR: "Quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR)" – are the changes to the text okay for sense?**	
5	**AUTHOR: "GUS activity was expressed in pmol MU min^{-1} μg^{-1} soluble protein" – are the changes to the text okay for sense?**	
6	**AUTHOR: "synthetic defined medium lacking leucine (SD/–Leu)" – are the changes to the text okay for sense?**	
7	**AUTHOR: "Quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR)" – are the changes to the text okay for sense?**	
8	**AUTHOR: "MU, 4 methylumbelliferone" – are the changes to the text okay for sense?**	
9	**AUTHOR: Hwang et al. (2004) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.**	
10	**AUTHOR: Figure 1 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	
11	**AUTHOR: Figure 2 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	
12	**AUTHOR: Figure 3 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	
13	**AUTHOR: Figure 4 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	
14	**AUTHOR: Figure 5 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	

15	**AUTHOR: Figure 6 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	
16	**AUTHOR: Figure 7 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	
17	**AUTHOR: Figure 8 is of poor quality. Please check required artwork specifications at https://authorservices.wiley.com/asset/photos/electronic_artwork_guidelines.pdf**	

Funding Info Query Form

Please confirm that the funding sponsor list below was correctly extracted from your article: that it includes all funders and that the text has been matched to the correct FundRef Registry organization names. If a name was not found in the FundRef registry, it may not be the canonical name form, it may be a program name rather than an organization name, or it may be an organization not yet included in FundRef Registry. If you know of another name form or a parent organization name for a "not found" item on this list below, please share that information.

FundRef name	**FundRef Organization Name**
MINECO	MINECO
National Natural Science Foundation of China	National Natural Science Foundation of China

New Phytologist

Research

ZmPBF and ZmGAMYB transcription factors independently transactivate the promoter of the maize (*Zea mays*) β-carotene hydroxylase 2 gene

Xin Jin[1]*, Chao Bai[1,2]*, Ludovic Bassie[1], Carmina Nogareda[3], Ignacio Romagosa[1], Richard M. Twyman[4], Paul Christou[1,5] and Changfu Zhu[1,6]

[1]Department of Plant Production and Forestry Science, University of Lleida-Agrotecnio Center, Av. Alcalde Rovira Roure, 191, Lleida 25198, Spain; [2]Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing10081, China; [3]Department of Animal Science, ETSEA, University of Lleida-Agrotecnio Center, Lleida 25198, Spain; [4]TRM Ltd, PO Box 493, Scarborough YO11 9FJ, UK; [5]ICREA, Catalan Institute for Research and Advanced Studies, Passeig Lluís Companys 23, Barcelona 08010, Spain; [6]School of Life Sciences, Changchun Normal University, Changchun 130032, China

Author for correspondence:
Changfu Zhu
Tel: +34 973702694
Email: zhu@pvcf.udl.cat

Received: *3 September 2018*
Accepted: *21 November 2018*

New Phytologist (2018)
doi: 10.1111/nph.15614

Key words: 5′-untranslated region, AACA motif, carotenoids, *cis*-regulatory elements, prolamin box, reporter gene expression, β-glucuronidase.

Summary

- The maize (*Zea mays*) enzyme β-carotene hydroxylase 2 (ZmBCH2) controls key steps in the conversion of β-carotene to zeaxanthin in the endosperm. The *ZmBCH2* has an endosperm-preferred and developmentally regulated expression profile, but the detailed regulatory mechanism is unknown.
- To gain insight into the regulation of *ZmBCH2*, we isolated 2036 bp of the 5′-flanking region ~~which contains~~ the 263 bp 5′-untranslated region (5′-UTR) including the first intron. We linked this to the β-glucuronidase reporter gene *gusA*.
- We found that high-level expression of *gusA* in rice seeds requires the 5′-UTR for enhanced activation. Truncated variants of the *ZmBCH2* promoter retained their seed-preferred expression profile as long as a prolamin box and AACA motif were present.
- We identified candidate genes encoding the corresponding transcription factors (ZmPBF and ZmGAMYB) and confirmed that their spatiotemporal expression profiles are similar to *ZmBCH2*. Both ZmPBF and ZmGAMYB can transactivate *ZmBCH2* expression in maize endosperm. To eliminate potential confounding effects in maize, we characterized the regulation of the minimal promoter region of *ZmBCH2* in transgenic rice. This revealed that ZmPBF and ZmGAMYB independently transactivate the *ZmBCH2* promoter. The mechanism that underpins our data provides an exciting new strategy for the control of target gene expression in engineered plants.

Introduction

Carotenoids play a fundamental role in human nutrition as antioxidants and precursors of vitamin A (Fraser & Bramley, 2004). Plants are good sources of dietary carotenoids, which accumulate in various plastids such as chloroplasts, amyloplasts and chromoplasts (Bai *et al.*, 2016; Giuliano, 2017). Genes encoding carotenogenic enzymes have therefore been used for metabolic engineering to increase the nutritional value and health-promoting properties of plants by improving the carotenoid content and composition (Farré *et al.*, 2014; Giuliano, 2017). However, the regulation of endogenous carotenogenic gene expression in higher plants is poorly understood, restricting the extent to which the impact of metabolic engineering in staple crops can be predicted (Giuliano *et al.*, 2008; Farré *et al.*, 2014).

Maize (*Zea mays* L.) kernels contain low concentrations of provitamin A (β-carotene, α-carotene and β-cryptoxanthin), reaching only 1–10% of the abundance of the nonprovitamin A carotenoids zeaxanthin and lutein (Harjes *et al.*, 2008). The carotenoid biosynthesis pathway is well characterized in maize endosperm. Maize phytoene synthase 1 (ZmPSY1) controls the flux of precursors into the carotenoid pathway by catalyzing the first committed step (Zhu *et al.*, 2008). At a later branch point, lycopene ε-cyclase (LCYE) directs flux towards α-carotene, whereas lycopene β-cyclase (LCYB) directs flux towards β-carotene, the most active form of provitamin A (Harjes *et al.*, 2008). The concentrations of β-carotene in maize ~~depend~~ not only on the synthesis steps but also on the activity of nonheme di-iron β-carotene hydroxylase 2 (ZmBCH2, also known as ZmcrtRB1 and ZmHYD3), which is preferentially expressed in the endosperm (Li *et al.*, 2010) and converts β-carotene first to β-cryptoxanthin and then to zeaxanthin (Vallabhaneni *et al.*,

*These authors contributed equally to this work.

Journal Code: N P H
Manuscript No.: 15614 / 2018-27953
WILEY
No. of pages: 12
Dispatch: 20.12.18
PE: Vigneshwari
CE: Raja
Color:

2009; Li *et al.*, 2010; Yan *et al.*, 2010). *ZmBCH2* alleles associated with reduced expression correlate with higher β-carotene concentrations in maize endosperm (Vallabhaneni *et al.*, 2009; Yan *et al.*, 2010). Quantitative trait locus (QTL) mapping, genome-wide association studies (GWAS) and metabolic data sorting show that ZmBCH2 controls key steps in the conversion of β-carotene to zeaxanthin in maize endosperm (Vallabhaneni *et al.*, 2009; Yan *et al.*, 2010). Furthermore, we recently demonstrated that the β-carotene content of maize endosperm increased substantially in hybrids in which *ZmBCH2* was silenced by RNA interference (RNAi), whether or not *ZmBCH1* (also known as *ZmcrtRB3* and *ZmHYD4*) was silenced simultaneously (Berman *et al.*, 2017).

The regulation of *ZmBCH2* expression in the endosperm is not well understood but it is important to unravel the molecular mechanisms responsible for the endosperm-preferred expression profile in order to develop rational strategies to engineer carotenoid biosynthesis in a more predictable manner. Several consensus regulatory sequences have been reported in storage protein genes, which are also expressed in an endosperm-preferred manner, including the prolamin box (P-box), GCN4 motif and AACA motif located *c.* 300 bp upstream of the transcriptional start site (Colot *et al.*, 1987; Mena *et al.*, 1998; Washida *et al.*, 1999; Wu *et al.*, 2000). The P-box is recognized by the P-box binding factor (PBF), which regulates many storage-protein genes in cereals (Vicente-Carbajosa *et al.*, 1997; Marzábal *et al.*, 2008). Maize PBF orthologs have been isolated from wheat (*Triticum aestivum*; WPBF), barley (*Hordeum vulgare*; BPBF) and rice (*Oryza sativa*; RPBF), and all are transcriptional activators of endosperm gene expression (Mena *et al.*, 1998; Diaz *et al.*, 2002, 2005; Yamamoto *et al.*, 2006). The AACA motif is recognized by GAMYB, an R2R3-class MYB transcription factor expressed in cereal aleurone cells and endosperm in response to gibberellin (GA; Suzuki *et al.*, 1998; Diaz *et al.*, 2002, 2005). The two AACA motifs in rice glutelin promoters are recognized by the OsMYB5 protein *in vitro*, suggesting this is a transcriptional regulator of glutelin expression *in planta* (Suzuki *et al.*, 1998). Similarly, the barley transcription factor GAMYB binds AACA motifs to transactivate the native genes encoding B-hordein, the trypsin inhibitor BTI-CMe, and α-amylase in aleurone cells (Gubler *et al.*, 1995, 1999).

To investigate the regulation of *ZmBCH2* expression in detail, we determined whether the 5′-untranslated region (5′-UTR) enhanced β-glucuronidase (GUS) reporter gene (*gusA*) expression in transgenic rice and identified minimal *ZmBCH2* promoter regions responsible for endosperm-preferred expression. To specifically determine the effects of putative *cis*-regulatory elements in the minimal *ZmBCH2* promoter, such as the P-box and AACA motif present in the first 300 bp, we overexpressed *ZmPBF* and/or *ZmGAMYB* in maize (transient expression assays in endosperm and embryo) and transgenic rice plants together with *gusA* under the control of the minimal *ZmBCH2* promoter, and investigated the impact of expression of *ZmPBF* and/or *ZmGAMYB* on *ZmBCH2* promoter activity in the leaf and seed.

Materials and Methods

Plant materials

Maize (*Zea mays* L.) inbred line B73, wild-type rice (*Oryza sativa* L. cv. EYI105) and transgenic rice plants were grown in a glasshouse or controlled growth chamber at 28°C : 20°C day : night temperature with a 10 h photoperiod and 60–90% relative humidity for the first 50 d, followed by maintenance at 21°C : 18°C day : night temperature with a 16 h photoperiod thereafter (the best conditions for rice and maize cultivation in our geographical location). Plants were self-pollinated to obtain seeds. Mature leaf, seed, embryo and endosperm tissues were frozen in liquid nitrogen and stored at −80°C.

Nucleic acid isolation and cDNA synthesis

Genomic DNA was extracted from 5 g of leaf tissue as described by Sambrook *et al.* (1989). Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen) and DNA was removed with DNase I (RNase-Free DNase Set; Qiagen). Total RNA was quantified using a Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and 1 μg of total RNA was used as the template for first-strand cDNA synthesis with Ominiscript Reverse Transcriptase in a 20 μl total reaction volume according to the manufacturer's recommendations (Qiagen).

Cloning the *ZmBCH2* promoter sequence

The *ZmBCH2* cDNA sequence (GenBank AY844958) was used as a BLASTN query against the MaizeGDB (Maize Genetics and Genomics Database) B73 RefGen_v2 (MGSC). The 5′-flanking region of *ZmBCH2* was cloned from maize B73 leaf genomic DNA by genomic PCR using forward primer 5′-GCT TAT GGT GGT ACG TGC CTG ATG ACT GAC GAT G-3′ (positions −1773 to −1740, +1 assigned to the first nucleotide of the longest *ZmBCH2* cDNA) and reverse primer 5′-CTT GGC GAC GAA GCT GGT CAT CGC GGC GGC CAT-3′ (positions 264 to 296). The genomic PCR products were transferred to vector PCRII TOPO (Invitrogen) for sequencing using the Big Dye Terminator v3.1 cycle sequencing kit on a 3130x1 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Cloning and sequencing the *ZmPBF* and *ZmGAMYB* cDNAs

The maize *PBF* cDNA sequence (GenBank ZMU82230) was used as a query to search the B73 RefGen_v2 and NCBI GenBank databases, allowing us to retrieve a cDNA sequence containing the corresponding full-length open reading frame (ORF) representing maize inbred line B73 (GenBank JQ723698). The full-length *ZmPBF* cDNA was then isolated from a maize B73 endosperm sample at mixed developmental time points of 15, 20, 25 and 30 d after pollination (DAP). Similarly, the barley *GAMYB* cDNA sequence (GenBank X87690)

was used to retrieve full-length *ZmGAMYB* cDNA (GenBank BT062692) from the same tissue. The amplified cDNAs were transferred to vector PCR II TOPO as described earlier for sequencing.

RNA blot analysis

Total RNA was extracted from maize endosperm, embryo and leaf tissues, and mRNA blotting was carried out as described by Li *et al.* (2010). The forward and reverse primers used to generate the probes are listed in Supporting Information Table S1.

Vector construction for rice transformation

Vectors pBI101 and pBI121 (Clontech Laboratories, Mountain View, CA, USA) were digested with *Hind*III and *Eco*RI, and then the *gusA*-TNos fragment from pBI101 and the CaMV35SPro-*gusA*-TNos fragment from pBI121 were individually transferred to pUC8 to generate pUC8-*gusA*-TNos and pUC8-CaMV35S-*gusA*-TNos, respectively. The 5′-flanking region of *ZmBCH2* was transcriptionally fused to the *gusA* reporter gene in vector pUC8-*gusA*-TNos by amplifying each 5′-flanking region using primers containing additional sequences to provide appropriate restriction sites (Table S2). Forward primer BCH2F1 and reverse primer BCH2R1 were used to amplify the 2036 bp *ZmBCH2* 5′-flanking region containing a 1773 bp promoter fragment (abbreviated as PF in the vector names) and the 263 bp 5′-UTR. The resulting 2036 bp DNA fragment was then transferred to the PCRII TOPO vector using the Invitrogen TA Cloning kit to yield intermediate vector pCR-ZmBCH2PF-5′UTR. The pCR-ZmBCH2PF-5′UTR and pUC8-*gusA*-TNos vectors were digested with *Hind*III and *Bam*HI, allowing the ZmBCH2PF-5′UTR fragment to be inserted upstream of *gusA*, yielding the final construct pUC8-ZmBCH2PF-5′UTR-*gusA* (PF-5′UTR-*gusA*, −1773 to +263). The construct containing the *ZmBCH2* promoter fragment (1773 bp) upstream of the longest cDNA was similarly constructed using the specific forward primer BCH2F1 and reverse primer BCH2R2 (positions −27 to −1) (Table S2) to generate pUC8-ZmBCH2PF-*gusA* (PF-*gusA*, −1773 to −1). Four 5′ deletions of the *ZmBCH2* promoter region were also generated by PCR, containing upstream promoter fragments and the 5′-UTR. Forward primers BCH2FD1 (positions −1271 to −1247), BCH2FD2 (positions −1264 to −1241), BCH2FD3 (positions −518 to −494) and finally BCH2FD4 (positions −511 to −488) (Table S2) were combined with reverse primer BCH2R1 to generate the four stepwise deletion constructs with *Hind*III and *Bam*HI restriction sites. These plasmids were named pUC8-ZmBCH2PFΔ1-5′UTR-*gusA* (PFΔ1-5′UTR-*gusA*, −1271 to +263), pUC8-ZmBCH2PFΔ2-5′UTR-g*usA* (PFΔ2-5′UTR-*gusA*, −1264 to +263), pUC8-ZmBCH2PFΔ3-5′UTR-*gusA* (PFΔ3-5′UTR-*gusA*, −518 to +263) and pUC8-ZmBCH2Δ4PF-5′UTR-*gusA* (PFΔ4-5′UTR-*gus*A, −511 to +263), respectively. The synthesized ZmBCH2PFΔ4-5′UTR fragments (−511 to +263) with a mutated P-box (mP, 5′-TGAAAAG-3′ to 5′-TGgAgAG-3′) or mutated AACA motif (mA, 5′-AACAAAC-3′ to 5′-gAgAAcC-3′) were transferred to the pUC8-*gusA*-TNos vector at the *Hind*III and *Bam*HI sites to generate pUC8-ZmBCH2PFΔ4-mP-5′UTR-*gusA* (PFΔ4-mP-5′UTR-*gusA*) and pUC8-ZmBCH2PFΔ4-mA-5′UTR-*gusA* (PFΔ4-mA-5′UTR-*gusA*), respectively. The *gusA* reporter gene driven by the Cauliflower mosaic virus 35S (CaMV35S) promoter in pUC8 (pUC8-CaMV35SPro-*gusA*-TNos, CaMV35S-Pro-*gusA*) was also introduced into rice as a positive control plasmid. The integrity of all intermediate and final constructs was confirmed by sequencing.

Rice transformation

Seven-day-old mature zygotic rice embryos were transferred to MS osmoticum medium (4.4 g l^{-1} Murashige-Skoog salts (Sigma-Aldrich) supplemented with 0.3 g l^{-1} casein hydrolysate, 0.5 g l^{-1} proline, 72.8 g l^{-1} mannitol and 30 g l^{-1} sucrose) 4 h before transformation and then bombarded with 0.4 mg gold particles coated with the transgene construct(s) and a plasmid containing hygromycin phosphotransferase (*hpt*) selectable maker gene at a molar ratio of 3 : 1 (for one transgene), 3 : 3 : 1 (for two transgenes) or 3 : 3 : 3 : 1 (for three transgenes) (Christou *et al.*, 1991). The embryos were returned to osmoticum medium for 12 h before selection on MS medium (as before but without mannitol) supplemented with 30 mg l^{-1} hygromycin and 2.5 mg l^{-1} 2,4-dichlorophenoxyacetic acid (2,4-D) in the dark for 2–3 wk. Transgenic plantlets were regenerated and hardened off in soil.

Yeast one-hybrid analysis

The bait sequences from pUC8-ZmBCH2PFΔ4-5′UTR-*gusA*, pUC8-ZmBCH2PFΔ4-mP-5′UTR-*gusA*, and pUC8-ZmBCH2PFΔ4-mA-5′UTR-*gusA* were inserted into the pAbAi vector at the *Hind*III and *Xma*I sites to form the bait vectors, which were then introduced into yeast strain Y1HGold using Yeastmaker Yeast Transformation System 2 (Clontech). To determine the minimal inhibitory concentration of aureobasidin A (AbA), the bait strains were separately screened on synthetic defined (SD) medium lacking uracil and containing different concentrations of AbA. The coding sequences of *ZmPBF* and *ZmGAMYB* were introduced into vector pGADT7 at the *Eco*RI and *Bam*HI sites, respectively, to generate the prey vectors pGADT7-*ZmPBF* and pGAT7-*ZmGAMYB*. The prey vector was introduced into the bait strains and grown on SD medium lacking leucine and containing 200 ng ml^{-1} AbA for 72 h at 30°C. The integrity of all intermediate and final constructs was confirmed by sequencing.

Vectors for transient expression in maize endosperm and embryo tissue

The full-length ORF sequences for *ZmPBF* and *ZmGAMYB* were amplified by PCR using the primer combinations listed in Table S3. After sequence verification, both cDNAs were transferred to the expression vector pAL76, containing the maize *ubiquitin-1* promoter plus first intron and *nos*

transcriptional terminator. The final constructs were verified by sequencing.

Maize B73 kernels at 10 and 25 DAP were sterilized and the endosperm was separated from the embryo. Tangential endosperm sections were prepared to expose a large area of subaleurone endosperm cells. After dissection, the tissues were placed in sterile Petri dishes on N6 osmoticum medium (Duchefa, Haarlem, the Netherlands) and were transformed by particle bombardment as previously described (Zhu *et al.*, 2008) using 0.70 μg of pAHC25 (Ubi-*gusA*), 0.63 μg of pAL76-*ZmPBF* (Ubi-*ZmPBF*) and/or 0.70 μg of pAL76-*ZmGAMYB* (Ubi-*ZmGAMYB*). The pAHC25 plasmid carrying *gusA* driven by the maize *ubiquitin-1* promoter (Ubi-*gusA*) was used as an internal control. All samples were bombarded twice and subsequently returned to osmoticum medium for 24 h at 25°C in the dark to express *ZmPBF*, *ZmGAMYB* and *gusA* before freezing in liquid nitrogen.

Gene expression analysis by quantitative real-time PCR

First-strand cDNA was synthesized from 1 μg total RNA using Ominiscript Reverse Transcriptase in a 20 μl total reaction volume according to the manufacturer's recommendations (Qiagen). Quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR) was performed on a CFX96 system (Bio-Rad) using a 25 μl mixture containing 10 ng cDNA, 1× iQ SYBR Green Supermix (Bio-Rad) and 0.2 μM of each primer. To calculate relative expression levels, serial dilutions (0.2–125 ng) were used to produce standard curves for each gene. Triplicate PCRs in 96-well optical reaction plates were carried out with the following profile: a heating step for 3 min at 95°C was followed by 40 cycles of 95°C for 10 s, 59°C for 30 s, and 72°C for 20 s. Amplification specificity was confirmed by melt curve analysis of the final PCR products in the temperature range 50–90°C with fluorescence acquired after each 0.5°C increment. The fluorescence threshold value and gene expression data were calculated using CFX96 system software. Primer combinations are listed in Table S4.

Histochemical and fluorimetric GUS assays

Histochemical GUS assays were carried out as described by Jefferson *et al.* (1987) with minor modifications. Leaves and sectioned seeds were incubated at 37°C overnight (12 h) in the dark in 1 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-gluc) in 100 mM sodium phosphate (pH 7.0), 10 mM EDTA, 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 0.3% (v/v) Triton X-100 and 20% (v/v) methanol to eliminate endogenous GUS activity (Kosugi *et al.*, 1990). After staining, tissues were destained in an ethanol series (50%, 70%, 80% and 95%) to remove Chl, and then stored in 70% (v/v) ethanol, and photographed with a digital camera.

Fluorimetric GUS assays were carried out as described by Jefferson *et al.* (1987) with minor modifications. Plant tissues (100 mg) were ground to powder under liquid nitrogen, dispersed in 0.5 ml extraction buffer (50 mM Na_2HPO_4, pH 7.0, 10 mM EDTA, 0.1% (v/v) sodium dodecanoyl(methyl)aminoacetate, 10 mM 2-mercaptoethanol and 0.1% (v/v) Triton X-100) and centrifuged at 15 300 ***g*** for 20 min at 4°C. The supernatant (250 μl) was mixed with 250 μl 2 mM 4-methylumbelliferyl-β-D-glucuronide on ice. Four 100 μl aliquots were then transferred to a fresh tube, and 1.9 ml of GUS stop buffer (0.2 M Na_2CO_3) was immediately added to one aliquot as a control. The remaining aliquots were incubated at 37°C for 1 h before adding 1.9 ml of GUS stop buffer to each. The released fluorescent product 4-methylumbelliferone (MU) was measured on a QuantiFluor hand-held fluorimeter (Promega) with excitation and emission wavelengths of 365 and 455 nm, respectively. The protein content was determined using the Bradford Reagent Kit (Sigma-Aldrich) with bovine serum albumin as a standard. GUS activity was expressed in pmol MU min^{-1} μg^{-1} soluble protein. Three biological replicates and three technical replicates were performed for GUS activity analysis.

Statistical analysis

Significant differences in quantitative dependent variables for any two alternative treatments were calculated using Student's *t*-test, with *P*-values of 0.05 and 0.01 considered significant and highly significant, respectively. GUS activity and normalized *gusA* expression in rice tissues were analyzed by standard two-way factorial analysis of variance (ANOVA), using presence or absence of the transgenes encoding the two transcription factors as fixed factors. Analysis of covariance (ANCOVA) was also carried out, using as independent covariates the product of the expression of the transcription factors, rather than their presence or absence.

Results

Cloning and bioinformatic analysis of the *ZmBCH2* 5′-flanking region

The *ZmBCH2* cDNA sequence was used as a blastn query against sequence database B73 RefGen_v2 (MGSC), revealing that the endogenous *ZmBCH2* gene is located on chromosome 10. The 2036 bp 5′-flanking region (GenBank KF941345) was cloned from the leaf genomic DNA of maize inbred line B73 (Fig. 1). Because *ZmBCH2* is preferentially transcribed in the endosperm (Li *et al.*, 2010), we searched the PLACE and PlantCARE databases (Higo *et al.*, 1999; Lescot *et al.*, 2002) to identify *cis*-regulatory elements potentially responsible for this expression profile. This revealed two P-boxes (5′-TG(T/A/C)AAA(G/A)-3′, positions −198 to −192 and −518 to −512), one AACA motif (5′-AACAAAC-3′, position −253 to −247), one reverse and complementary P-box (5′-CTTTACA-3′, position −1271 to −1265), one GCN4 motif (5′-CAAGCCA-3′, position −1283 to −1277), one ACGT motif (ACGTOSGLUB1 motif, 5′-GTACGTG-3′, position −1763 to −1757) and three Skn-1 motifs (5′-GTCAT-3′, positions −393 to −389, −434 to −430 and −1560 to −1556). These motifs are known to be responsible for the endosperm-specific expression of other genes in monocots (Vicente-Carbajosa *et al.*, 1997; Washida *et al.*, 1999; Wu *et al.*,

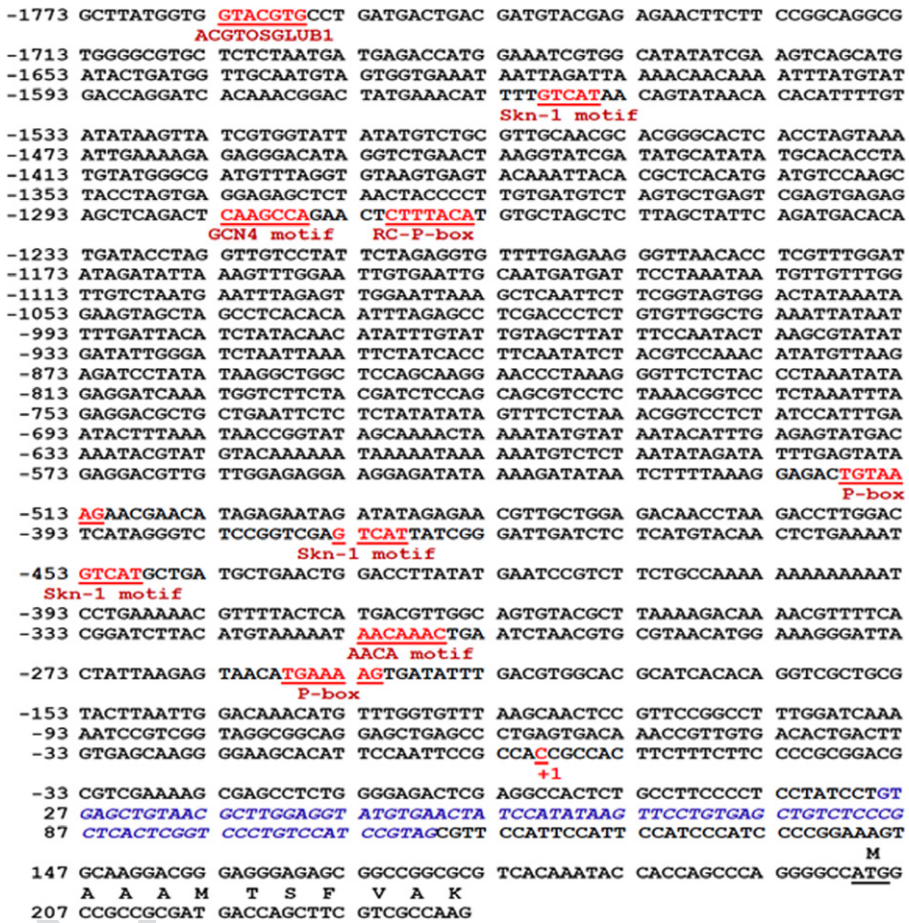


Fig. 1 Sequence of the maize (*Zea mays*, inbred line B73) β-carotene hydroxylase 2 gene (*ZmBCH2*) promoter and part of the cDNA region (GenBank KF941345). Position +1 is the first nucleotide of the longest *ZmBCH2* cDNA (GenBank AY844958). The first intron (positions 86–173) is in blue italics. The P-box, reverse and complementary P-box (RC-P-box), AACA motif, GCN4 motif, Skn-1 motif and ACGTOSGLUB1 motif are in red and underlined.

2000; Yamamoto *et al.*, 2006) and we therefore considered them likely to be involved in the endosperm-preferred expression of *ZmBCH2*.

Characterization of the 5′-flanking region of *ZmBCH2* in transgenic rice

We prepared an expression construct in which the full-length *ZmBCH2* 5′-flanking region was fused to the reporter gene *gusA* (PF-5′UTR-*gusA*) (Fig. 2a). This included 1773 bp of upstream promoter sequence and the 263 bp 5′-UTR containing the first intron. The PF-5′UTR-*gusA* construct was introduced into rice, and at least six independent transformants were analyzed by histochemical staining for GUS activity in the leaves and hand-cut half cross-sections of seeds. The full-length construct was sufficient to drive strong GUS activity in the seeds but not in the leaves of the transgenic rice plants (Fig. 2b), and the expression profile of *gusA* was similar to that of the native *ZmBCH2* in maize (Li *et al.*, 2010). These data indicated that the *ZmBCH2* 5′-flanking region in transgenic rice is regulated in a similar manner to the endogenous *ZmBCH2* promoter in maize.

The 5′-UTR of plant genes, particularly the first intron, is often necessary to achieve high-level gene expression (Hernandez-Garcia & Finer, 2014; Gallegos & Rose, 2015). To determine whether the 5′-UTR plays a similar role in the *ZmBCH2* gene, the 1773 bp promoter region was linked to *gusA* following the removal of the 263 bp 5′-UTR (PF-*gusA*). Analysis of the corresponding transgenic rice plants revealed that the 1773 bp promoter fragment alone was sufficient to confer seed-specific GUS activity (Fig. 2), but the level of GUS activity determined using a fluorimetric GUS assay was 8.4-fold lower than that of the construct containing the 5′-UTR (Fig. S1). This confirmed that the 5′-UTR, containing the first intron, is necessary for high-level *ZmBCH2* expression.

To identify the *cis*-regulatory elements responsible for the endosperm-preferred expression of *ZmBCH2*, the 5′-flanking region was progressively deleted from the 5′-end, and the nested set of constructs was tested in transgenic rice (Fig. 2). Deletions that removed the 5′ sequence beyond position −1271 removed the ACGTOSGLUB1 (ACGT motif, −1763 to −1757), a Skn-1 motif (−1560 to −1556) and a GCN4 motif at position −1283 to −1277 (Fig. 2a). The GUS activity was similar to that of the full-length 5′-flanking region with the 5′-UTR (Fig. 2b). Further progressive deletion to positions −1264, −518 and −511 removed the reverse and complementary P-box (−1271 to −1265) and P-box (−518 to −512) but had little further impact on GUS activity (Fig. 2b). All constructs produced similar GUS profiles (Fig. 2b) and similar levels of GUS activity in transgenic rice seeds (Fig. S1) compared with the full-length construct PF-5′UTR-*gusA*. None of the constructs were active in leaves, but

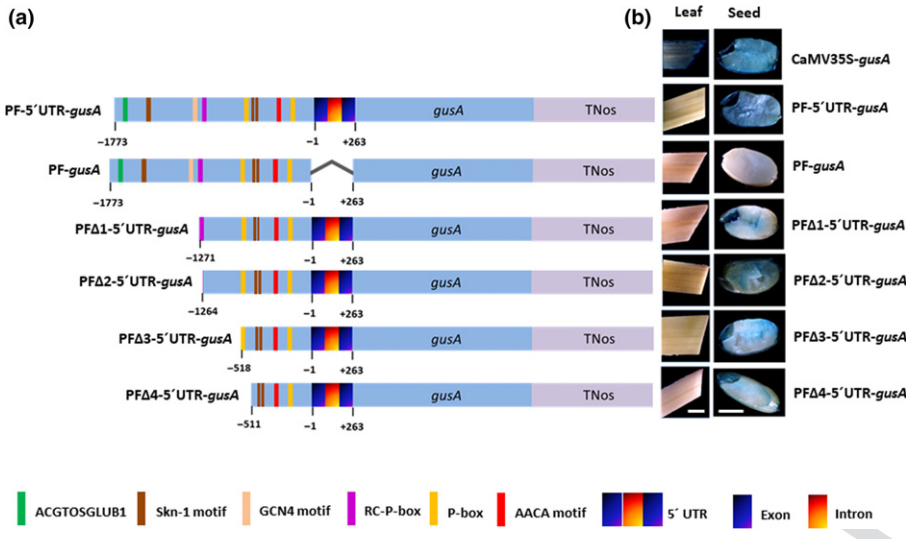


Fig. 2 Schematic representation of *gusA*-fusion constructs and histochemical β-glucuronidase (GUS) assays of representative transgenic rice tissues expressing *gusA* driven by the constitutive CaMV35S promoter (CaMV35SPro-*gusA*), and various 5′ upstream regions of maize *ZmBCH2*. (a) Schematic representation of *gusA*-fusion constructs driven by various 5′ upstream regions of the *ZmBCH2* gene. (b) Histochemical GUS staining of different tissues in transgenic rice plants (T1 mature leaves and T2 seeds at 25 d after pollination (DAP)) expressing *gusA*-fusion constructs driven by various 5′ upstream regions of the *ZmBCH2* gene. Rice mature leaf tissues and cross-sectioned seeds at 25 DAP were incubated in 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-gluc) solution. Bars, 2 mm.

even the shortest, with 774 bp of 5′-flanking sequence (PFΔ4-5′-UTR, −511 to +263), contained all the *cis*-regulatory elements necessary for high GUS activity in transgenic rice seeds. As shown in Fig. 1, this 774 bp region contains a single P-box, a single AACA motif and two Skn-1 motifs, suggesting these are the key *cis*-regulatory elements responsible for seed-preferred expression, whereas the distal ACGT-motif, reverse and complementary P-box, GCN4 and Skn-1 motifs are dispensable.

Expression patterns of *ZmBCH2*, *ZmPBF* and *ZmGAMYB* in maize B73 leaf, endosperm and embryo tissues

Our analysis of the *ZmBCH2* 5′-flanking region in transgenic rice revealed that a 774 bp flanking region (−511 to +263) containing one AACA motif, one P-box and two Skn-1 motifs was sufficient to confer preferential expression in the endosperm. To identify the corresponding transcription factors, we used the barley *GAMYB* cDNA sequence (GenBank X87690) to search MaizeGDB and NCBI GenBank, revealing one maize ortholog, *ZmGAMYB* (GenBank BT062692). The *ZmGAMYB* gene is located on chromosome 3 and has four exons. The deduced ZmGAMYB amino acid sequence has 80.1% identity and 84.9% similarity with barley GAMYB (Fig. S2). Similarly, by using the *ZmPBF* cDNA sequence (GenBank ZMU82230) to search the MaizeGDB, we isolated the *ZmPBF* full-length ORF cDNA from inbred line B73 (GenBank JQ723698). *ZmPBF* is located on chromosome 2 and has two exons.

RNA gel blots were then prepared from 20 μg of total RNA and were hybridized with gene-specific probes for *ZmBCH2*, *ZmPBF* and *ZmGAMYB*. *ZmBCH2* expression levels gradually increased in the endosperm from 15 to 30 DAP (Fig. 3) and both *ZmPBF* and *ZmGAMYB* expression levels gradually increased in the endosperm from 15 to 25 DAP, before declining at 30 DAP. Furthermore, *ZmBCH2* and *ZmPBF* expression levels gradually increased in the embryo from 20 to 30 DAP, whereas *ZmGAMYB* expression was only detected in the embryo at 20 DAP. No transcripts for any of the genes were detected in leaves.

ZmPBF/P-box and ZmGAMYB/AACA motif interactions

Next we used the yeast one-hybrid system to determine whether ZmPBF and ZmGAMYB interact with the P-box and AACA motifs, respectively, in the minimal *ZmBCH2* promoter. The bait plasmids pAbAi-ZmBCH2PFΔ4-5′UTR (wild-type sequence), pAbAi-ZmBCH2PFΔ4-mP-5′UTR (mutated P-box) and pAbAi-ZmBCH2PFΔ4-mA-5′UTR (mutated AACA motif) were independently introduced into Y1HGold yeast cells, which can grow on medium lacking uracil but not on medium containing 200 ng ml^{-1} AbA. We therefore chose 200 ng ml^{-1} as the minimal inhibitory concentration of AbA for the bait plasmids. The prey plasmid pGADT7-*ZmPBF* was separately introduced into the bait (pAbAi-ZmBCH2PFΔ4-5′UTR), mutated P-box bait (pAbAi-ZmBCH2PFΔ4-mP-5′UTR) and void bait (pAbAi) yeast strains. The transformed yeast cells with the wild-type promoter were able to grow on medium lacking leucine and containing 200 ng ml^{-1} AbA, but the transformed mutated P-box bait and void bait yeast cells could not (Fig. 4a), indicating that ZmPBF can specifically bind to the P-box element in the

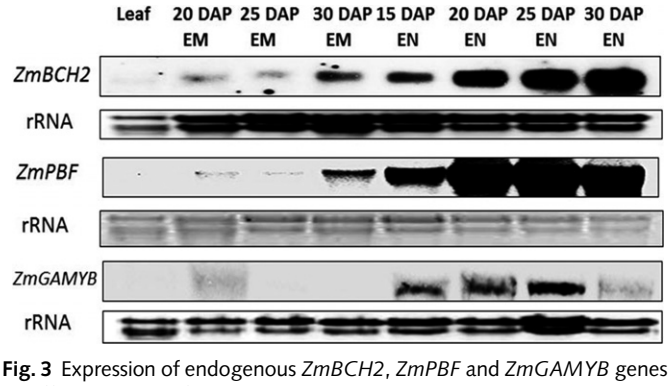


Fig. 3 Expression of endogenous *ZmBCH2*, *ZmPBF* and *ZmGAMYB* genes in different tissues of maize (*Zea mays*, inbred line B73). We loaded 20 μg of total RNA for each sample. DAP, d after pollination; EM, embryo; EN, endosperm.

truncated *ZmBCH2* promoter in yeast. Likewise, we demonstrated that ZmGAMYB can specifically bind to the AACA motif in the *ZmBCH2* promoter in yeast (Fig. 4b). Thus, the yeast one-hybrid system indicated that ZmPBF and ZmGAMYB interact with the P-box and AACA motif, respectively, in the proximal region of the truncated *ZmBCH2* promoter.

Transient expression of *ZmPBF* and/or *ZmGAMYB* and the impact on *ZmBCH2* and *ZmPSY1* expression in maize endosperm and embryo tissues

A transient expression assay was carried out to determine the effect of ZmPBF and ZmGAMYB on *ZmBCH2* promoter activity. Maize seeds were bombarded with the expression vectors pAHC25 (Ubi-*gusA*), pAL76-*ZmPBF* (Ubi-*ZmPBF*) and/or pAL76-*ZmGAMYB* (Ubi-*ZmGAMYB*) (Fig. S3) and the effect on *ZmBCH2* expression was determined by qRT-PCR. In maize endosperm and embryo cells at 10 and 25 DAP, the overexpression of either *ZmPBF* or *ZmGAMYB* alone significantly increased endogenous *ZmBCH2* mRNA levels (Fig. 5). However, the overexpression of both transcription factors did not result in additive transactivation of the *ZmBCH2* promoter in the endosperm and embryo. These results indicate that ZmPBF and ZmGAMYB can act independently as effective stimulators of the *ZmBCH2* promoter in maize endosperm and embryo.

Bioinformatic analysis of the *ZmPSY1* 5′-flanking region revealed the presence of one AACA motif (−827 to −821), two P-boxes (−1221 to −1215 and −1341 to −1335) and one reverse complementary P-box (−1600 to −1594) in the distal region (Fig. S4). However, P-box and AACA motifs were not found in the promoters of other carotenoid biosynthesis genes up to 2 kb upstream from the transcriptional start site (Table S5). RNA gel blots revealed that *ZmPSY1* is abundantly transcribed in maize endosperm from 20 DAP (Fig. S5). We therefore investigated endogenous *ZmPSY1* expression using the transient expression strategy described earlier. The transient expression of *ZmPBF* and/or *ZmGAMYB* caused a slight but nonsignificant increase in *ZmPSY1* expression in maize embryo and endosperm tissue (Fig. S6).

Overexpression of *ZmPBF* and/or *ZmGAMYB* in transgenic rice plants and the impact on *gusA* expression and GUS activity in leaves and seeds

We transformed 7-d-old mature zygotic rice embryos with the PFΔ4-5′UTR-*gusA* construct with or without pAL76-*ZmPBF* and/or pAL76-*ZmGAMYB*, in which the *ZmPBF* and *ZmGAMYB* cDNAs are driven by the constitutive maize *ubiquitin-1* promoter (Fig. S3). In all cases, an additional construct was introduced containing the selectable marker *hpt* (encoding hygromycin phosphotransferase), also driven by the maize *ubiquitin-1* promoter. We measured transcript abundance in the leaf and seed tissues of the resulting transgenic rice plants (at least 10 independent transformants for each combination of transgenes) by qRT-PCR and selected the top three in each case with the highest expression levels of *ZmPBF* and/or *ZmGAMYB* for the analysis of *gusA* expression and GUS activity in the T2 homozygous lines. Looking at the lines expressing the PFΔ4-5′UTR-*gusA* construct plus one of the two transcription factors, we found that *gusA* expression levels and GUS activity were consistently higher than in the control line expressing the PFΔ4-5′UTR-*gusA* construct alone (Figs 6, 7, S7, S8), confirming that the reporter construct was transactivated by both ectopic transcription factors acting independently. In the seed, the *gusA* expression level was 2.6 (normalized against the rice *actin* gene) in the absence of any transcription factors, rising to 7.0 in the presence of *ZmPBF* or 15.6 in the presence of *ZmGAMYB*, and to 27.3 when both transcription factors were present (Fig. 6). In the leaves, the corresponding values were 0.6 for no transcription factors, rising to 1.9 in the presence of *ZmPBF* or 2.8 in the presence of *ZmGAMYB*, and to 4.5 in the presence of both (Fig. 6). The model that incorporated the quantitative amount of transcription factor mRNA as independent covariates better explained the *gusA* expression compared with the model based on solely the presence or absence of the transgenes ($R^2 = 90.4$ vs 82.6% in the rice seed and 96.2% vs 90.2% in the leaf) (Table S6). ZmGAMYB had a stronger effect on *gusA* expression ($P = 0.0006$ and 0.0004 for seed and leaf, respectively) than did ZmPBF ($P = 0.0298$ and 0.0070, respectively) (Table S6). No significant statistical

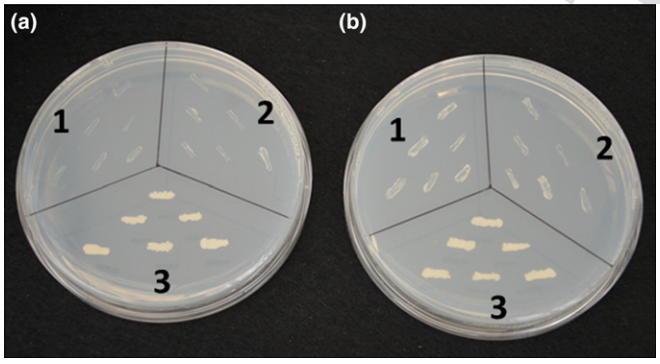


Fig. 4 Identification and characterization of P-box and AACA motifs in the maize *ZmBCH2* basal promoter region which could be recognized and bound by the ZmPBF and ZmGAMYB transcription factors, respectively. (a) Plasmids pAbAi (a1), pAbAi-ZmBCH2PFΔ4-mP-5′UTR (mP, mutated P-box in 774 bp ZmBCH2PFΔ4-5′UTR fragment) (a2) and pAbAi-ZmBCH2PFΔ4-5′UTR (a3) were introduced into the yeast strain Y1HGold. These strains were retransformed with pGADT7-*ZmPBF*. The resulting strains could grow on synthetic defined medium lacking leucine (SD/–Leu). Once these yeast strains were transferred onto SD/–Leu/aureobasidin A (AbAi) medium (200 ng ml^{-1}), only strain (a3) transformed with pAbAi-ZmBCH2PFΔ4-5′UTR containing an intact P-box motif could grow, whereas strains transformed with pAbAi (a1) and pAbAi-ZmBCH2PFΔ4-mP-5′UTR (a2) could not. (b) Plasmids pAbAi (b1), pAbAi-ZmBCH2PFΔ4-mA-5′UTR (mA, mutated AACA motif in 774 bp ZmBCH2PFΔ4-5′UTR fragment) and pAbAi-ZmBCH2PFΔ4-5′UTR were introduced into the yeast strain Y1HGold. These strains were retransformed with pGADT7-*ZmGAMYB*. The resulting strains could grow on SD/–Leu medium. Once these yeast strains were transferred to SD/–Leu/AbAi medium (200 ng ml^{-1}), only strain (b3) transformed with pAbAi-ZmBCH2PFΔ4-5′UTR containing an intact AACA motif could grow, whereas the strains transformed with pAbAi (b1) and pAbAi-ZmBCH2PFΔ4-mA-5′UTR (b2) could not.

New Phytologist

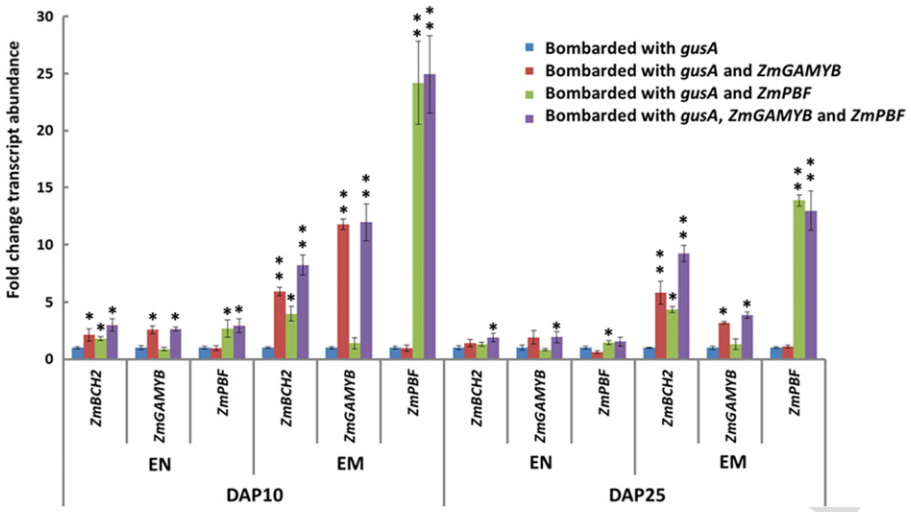


Fig. 5 Expression of *ZmBCH2*, *ZmGAMYB* and *ZmPBF* in maize (*Zea mays*, inbred line B73) endosperm and embryo tissue as measured by transient expression analysis followed by quantitative real-time reverse transcription polymerase chain reaction (qRT-PCR), with mRNA levels normalized against *ZmACTIN* mRNA. Expression of *ZmPBF*, *ZmGAMYB* and *ZmBCH2* is expressed as a percentage relative to the internal control (*gusA* expression). Five biological replicates and three technical replicates were tested. Asterisks indicate statistically significant differences compared with the control (bombarded with pAHC25) according to Student's *t*-test: *, $P < 0.05$; **, $P < 0.01$. Bars indicate ± SD.

14

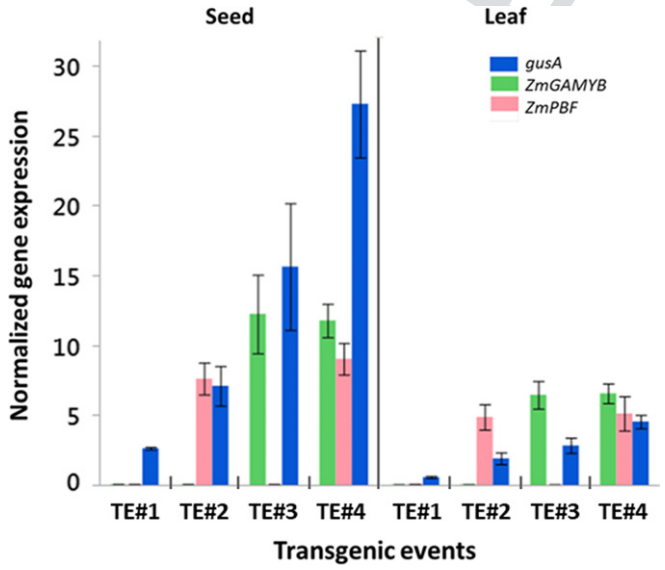
7

interaction was detected for either the presence/absence (ANOVA) or the quantitative (ANCOVA) models, further supporting the independent additive behavior of the transcription factors in both tissues (Fig. 6; Table S6). GUS activity analysis fully supported the *gusA* expression data (Fig. 7; Table S6).

When the experiments were repeated using PFΔ4-mP-5′UTR-*gusA* (the mutated P-box in the truncated *ZmBCH2* promoter, 5′-TGAAAAG-3′ to 5′-TGgAgAG-3′) and pAL76-*ZmPBF*, or using PFΔ4-mA-5′UTR-*gusA* (the mutated AACA motif in the truncated *ZmBCH2* promoter, 5′-AACAAAC-3′ to 5′-gAgAAcC-3′) and pAL76-*ZmGAMYB*, the effect of the two transcription factors was completely abolished in T0 transgenic rice leaf tissue and T1 seeds (Figs 8, S9). These data strongly suggest that the proximal P-box and AACA motif of the *ZmBCH2* promoter are the sites bound *in planta* by ZmPBF and ZmGAMYB, respectively, to activate the *ZmBCH2* promoter.

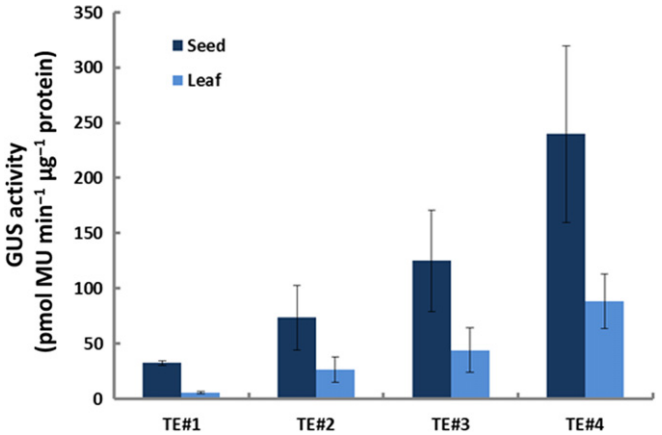
Discussion

The transcriptional regulation of carotenoid biosynthesis genes in monocots is a key intervention point to gain insight into carotenoid metabolism and, in turn, to devise more predictable and robust strategies for metabolic engineering. Very little is known about the regulation of these genes in cereals, or indeed in other plants. A few transcription factors have been shown to influence carotenogenic genes in dicots, including Arabidopsis RAP2.2 and PIF1 (Welsch *et al.*, 2007; Toledo-Ortiz *et al.*, 2010), tomato (*Solanum lycopersicum*) RIN and BBX20 (Martel *et al.*, 2011; Xiong *et al.*, 2018), and citrus CrMYB68 and



15

Fig. 6 Relative average expression (normalized to *OsACTIN*) in rice T3 seeds (25 d after pollination, DAP) and T2 mature flag leaves for two maize transcription factor genes (*ZmPBF* and *ZmGAMYB*) and the reporter gene *gusA*. Bars represent ± SE of the mean for three independent transgenic lines representing each transgenic event (TE): TE#1 expressing *gusA*; TE#2 coexpressing *gusA* and *ZmPBF*; TE#3 coexpressing *gusA* and *ZmGAMYB*; and TE#4 coexpressing *gusA*, *ZmPBF* and *ZmGAMYB*.



16

Fig. 7 Quantitative analysis of β-glucuronidase (GUS) activity in transgenic rice T3 seeds and T2 mature flag leaves. Bars represent ± SE of the mean for three independent transgenic lines representing each transgenic event (TE): TE#1 expressing *gusA*; TE#2 coexpressing *gusA* and *ZmPBF*; TE#3 coexpressing *gusA* and *ZmGAMYB*; and TE#4 coexpressing *gusA*, *ZmPBF* and *ZmGAMYB*.

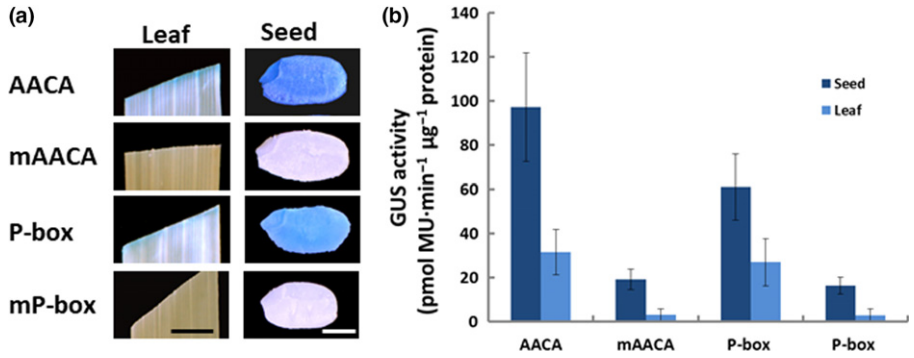


Fig. 8 (a) Histochemical β-glucuronidase (GUS) assays of representative transgenic rice T0 mature flag leaves and T1 seeds (at 25 d after pollination, DAP); (b) quantitative analysis of GUS activity in transgenic rice T0 mature flag leaves and T1 seeds (at 25 DAP). Three biological replicates were tested for GUS activity. Bars represent ± SE of the mean for three independent transgenic lines representing each transgenic event. AACA, transgenic rice events coexpressing *ZmGAMYB* and *gusA* driven by the minimal promoter (PFΔ4-5′UTR-*gusA*); mAACA, transgenic rice events coexpressing *ZmPBF* and *gusA* driven by the mutated AACA motif minimal promoter (PFΔ4-mA-5′UTR-*gusA*); P-box, transgenic rice events coexpressing *ZmPBF* and *gusA* driven by the minimal promoter (PFΔ4-5′UTR-*gusA*); mP-box, transgenic rice events coexpressing *ZmPBF* and *gusA* driven by the mutated P-box minimal promoter (PFΔ4-mP-5′UTR-*gusA*); MU, 4 methylumbelliferone. Bar, 2 mm.

CsMADS6 (Zhu *et al.*, 2017b; Lu *et al.*, 2018). RAP2.2, a member of the *AP2* gene family, binds to the Arabidopsis *PSY* promoter and represses *PSY* and *PDS* (encoding phytoene desaturase) in root-derived callus, thus reducing the carotenoid content (Welsch *et al.*, 2007). Arabidopsis phytochrome-interacting factor 1 (PIF1) directly binds to the *PSY* promoter and thus represses *PSY* expression to inhibit carotenoid accumulation (Toledo-Ortiz *et al.*, 2010). The MADS-box protein RIPENING INHIBITOR (RIN) in tomato directly induces the expression of *SlPSY1* (Martel *et al.*, 2011). SlBBX20, a B-box (BBX) zinc-finger transcription factor, can activate the expression of *SlPSY1* by directly binding to a G-box motif in its promoter, which results in the elevated levels of carotenoids in *SlBBX20* overexpression lines (Xiong *et al.*, 2018). CrMYB68, an R2R3-MYB transcription factor isolated from Green Ougan (MT), a stay-green mutant of *Citrus reticulate* cv Suavissima, can bind to the promoters of both *CrBCH2* and *CrNCED5* (*9-cis-epoxycarotenoid dioxygenase 5*) and significantly inhibits their expression (Zhu *et al.*, 2017b). The MADS-box transcription factor gene *CsMADS6* was recently identified by using the promoter sequence of *CsLCYB1* as bait in a yeast one-hybrid screen for promoter-binding proteins from sweet orange (Lu *et al.*, 2018). Overexpression of CsMADS6 in citrus callus increased the carotenoid content and induced the expression of *CsLCYB1*, *CsPSY*, *CsPDS* and *CsCCD1* (*carotenoid cleavage dioxygenase 1*) by directly binding to their promoters, highlighting the multitarget regulation of carotenoid metabolism by CsMADS6 (Lu *et al.*, 2018). The regulation of carotenogenic genes in monocots is not well understood, even though cereals constitute a key target for metabolic engineering of the carotenoid pathway (Paine *et al.*, 2005; Zhu *et al.*, 2008; Bai *et al.*, 2016).

We focused on the regulation of *ZmBCH2* because this gene controls the accumulation of β-carotene in maize endosperm (Vallabhaneni *et al.*, 2009; Yan *et al.*, 2010), where the gene is preferentially expressed (Li *et al.*, 2010). The *cis*-regulatory elements and corresponding transcription factors that control the endosperm-preferred expression of *ZmBCH2* have not yet been identified. We hypothesized that important regulatory elements would be located not only in the upstream promoter region, but also potentially in the 5′-UTR containing the first intron, as a result of the common phenomenon of intron-mediated transcriptional enhancement in plants (Hernandez-Garcia & Finer, 2014; Gallegos & Rose, 2015). We therefore tested a full-length 5′-flanking region spanning from position −1773 in the *ZmBCH2* upstream promoter to position +263, which includes the entire 5′-UTR (and first intron). When linked to the reporter gene *gusA*, this construct achieved strong GUS expression in the endosperm and embryo of transgenic rice seeds (Fig. 2b), but when the 263 bp 5′-UTR was removed, the GUS activity was reduced 8.4-fold. We therefore concluded that transcriptional enhancement mediated by the 5′-UTR is an important component of the regulation of *ZmBCH2*.

We constructed a series of 5′ deletions, progressively removing parts of the upstream promoter, but found that even the shortest construct (PFΔ4-5′UTR), spanning from position −511 to +263, contained all the necessary *cis*-regulatory elements responsible for high-level *gusA* expression and strong GUS activity in transgenic rice seeds. The mechanism of intron-mediated enhancement is not well understood (Hernandez-Garcia & Finer, 2014; Gallegos & Rose, 2015) so it will be interesting to investigate how the *ZmBCH2* 5′-UTR (including the first intron) boosts the expression of *gusA* in transgenic rice by conducting loss-of-function experiments (site-directed mutagenesis) using the above minimum promoter (PFΔ4-5′UTR).

Among the four *cis*-regulatory elements found in the minimum promoter, a single P-box and a single AACA motif were located within the first 300 bp upstream of the transcription start site, often the key regulatory region in endosperm-specific genes. The P-box is a highly conserved 7 bp sequence (5′-TG(T/A/C)AAA(G/A)-3′) found in the promoters of many cereal seed storage protein genes (Colot *et al.*, 1987; Vicente-Carbajosa *et al.*, 1997). As previously demonstrated using electrophoretic mobility shift assays (EMSAs), maize PBF (Vicente-Carbajosa *et al.*, 1997) and barley and wheat PBFs (Mena *et al.*, 1998) can bind

to storage protein gene promoters containing the P-box. Similarly, EMSAs have also revealed that the AACA motif is recognized by MYB-class transcription factors such as OsMYB5 (Suzuki *et al.*, 1998) and HvGAMYB (Diaz *et al.*, 2002). We confirmed that ZmPBF and ZmGAMYB can interact with the P-box and AACA motifs, respectively, in the proximal region of the truncated *ZmBCH2* promoter by using yeast one-hybrid assays. The minimal promoter fragment (774 bp) also contains two Skn-1 motifs spanning positions −393 to −389 and −434 to −430, which are involved in the endosperm-specific expression of storage protein genes in cereal grains (Washida *et al.*, 1999). The Skn-1 motif in the promoter of the rice storage protein glutelin gene *GluB-1* was found to be necessary for high expression levels in the endosperm in cooperation with other motifs (AACA, GCN4 and ACGT) by site-directed mutagenesis (Washida *et al.*, 1999). It will be interesting to investigate whether the two Skn-1 motifs are important for the regulation of *ZmBCH2* expression.

Having identified likely candidates for the transactivation of the minimal *ZmBCH2* promoter, we confirmed that the genes for both transcription factors (*ZmPBF* and *ZmGAMYB*) were preferentially transcribed in maize endosperm and regulated developmentally in a similar manner to *ZmBCH2* itself (Fig. 3). We then carried out transient expression experiments in maize, which revealed that the expression of *ZmPBF* and/or *ZmGAMYB* upregulated the endogenous *ZmBCH2* gene in the endosperm and embryo (Fig. 5). Taken together, these results confirmed that *ZmPBF*, *ZmGAMYB* and *ZmBCH2* have similar spatiotemporal expression profiles, so the transcription factors have the opportunity to regulate *ZmBCH2*, and that, when present, they can transactivate the *ZmBCH2* promoter, strongly supporting their role in the regulation of *ZmBCH2 in planta*. The 5′-flanking region of *ZmPSY1* contains one AACA motif (−827 to −821), two P-boxes (−1221 to −1215 and −1341 to −1335) and one reverse and complementary P-box (−1600 to −1594) in the distal region (Fig. S4). However, the transient expression of *ZmPBF* and/or *ZmGAMYB* did not significantly upregulate endogenous *ZmPSY1* expression in maize endosperm and embryo tissues, suggesting that these elements are not essential for the regulation of *ZmPSY1* promoter activity (Fig. S6). Other transcription factors are therefore needed for the control of *ZmPSY1* expression in maize endosperm.

We also investigated the regulatory potential of ZmPBF and ZmGAMYB in transgenic rice plants expressing *gusA* under the control of the minimal *ZmBCH2* promoter (PFΔ4-5′UTR), because this isolates the regulatory controls from any potentially confounding factors such as the maize Opaque2 (O2) transcription factor, which interacts with ZmPBF in maize (Vicente-Carbajosa *et al.*, 1997). In the absence of either transcription factor, GUS activity in the transgenic rice lines expressing the PFΔ4-5′UTR-*gusA* construct alone was low but detectable in seeds, probably reflecting the presence of rice orthologs of these maize transcription factors, but negligible in leaves (Fig. 2b), where even rice orthologs of ZmPBF and ZmGAMYB would be absent. The expression of either transcription factor stimulated the expression of *gusA*, whereas the presence of both showed an additive effect (Fig. 7). The overexpression of both *ZmPBF* and *ZmGAMYB* did not result in additive transactivation of the *ZmBCH2* promoter in the maize endosperm and embryo (Fig. 5), which might be a result of the higher levels of endogenous *ZmBCH2* mRNA in maize endosperm and embryo tissue (Fig. 3). We did not detect *gusA* reporter gene expression or GUS activity in wild-type rice leaves or seeds. Our results indicate that the ZmPBF and ZmGAMYB transcription factors act in a largely independent manner in both maize and rice. Similarly, the combination of rice PBF and GAMYB with the 0.2 kb *GluB-1* promoter linked to *gusA* did not reveal any significant synergistic transactivation, again suggesting little or no interaction between the two transcription factors (Yamamoto *et al.*, 2006). By contrast, barley PBF interacts with the GAMYB transcription factor to activate barley endosperm-specific genes during seed development (Diaz *et al.*, 2002), suggesting that the regulatory mechanism used by these transcription factors is species-dependent. The clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) system is an effective tool for targeted gene mutagenesis in cereals, including maize (Zhu *et al.*, 2017a). Knockouts of *ZmPBF* and/or *ZmGAMYB* in maize using CRISPR/Cas9 could therefore provide further insight into the regulatory mechanisms leading to the biosynthesis and accumulation of β-carotene in maize grains.

In summary, the maize gene encoding β-carotene hydroxylate 2 (ZmBCH2) is regulated by a combination of *cis*-regulatory elements in the upstream promoter and 5′-UTR (including the first intron). The upstream promoter contains sufficient information to restrict *ZmBCH2* expression to the seed, but the 5′-UTR is required for high-level expression, which suggests that *ZmBCH2* expression is partially dependent on 5′-UTR-mediated transcriptional enhancement. Nested deletion analysis revealed that two *cis*-regulatory elements determine the activity of the *ZmBCH2* promoter, namely the proximal AACA motif and P-box in the −300 bp region of the promoter, which is consistent with results from other genes preferentially expressed in seeds. We isolated the candidate transcription factors (ZmPBF and ZmGAMYB) that bind these sequences based on the known regulatory mechanisms in other cereal gene promoters and confirmed that: the corresponding genes (*ZmPBF* and *ZmGAMYB*) are expressed in a spatiotemporally consistent profile shared with *ZmBCH2*; the transcription factors physically interact with the corresponding elements in yeast one-hybrid assays; and the transient expression of either or both transcription factors in maize increases the expression of endogenous *ZmBCH2*. These data show that the two transcription factors act independently. In transgenic rice plants expressing the minimal *ZmBCH2* promoter *gusA* construct (allowing us to isolate the regulatory controls), each transcription factor was able to transactivate the *gusA* expression when present alone, and when both were present the effect was additive. Statistical analysis to confirm the independent activity of the two factors revealed a stronger correlation when the quantity of each transcription factor was built into the model compared with a model based purely on presence/absence, providing a strong confirmation of our original hypothesis that the proximal P-box and AACA motifs in the *ZmBCH2* promoter are the sites bound *in planta* by ZmPBF and ZmGAMYB, respectively. Our results

indicate that transcriptional regulation by two transcription factors provides an elegant strategy for the regulation of carotenogenic gene expression in metabolically engineered plants. This sets the stage for the broader application of our approach (i.e. transcriptional control mediated by independently acting transcription factors) in other metabolic pathways.

Acknowledgements

This work was supported by MINECO (Ministry of Economy and Competitiveness), Spain (BIO2014-54441-P; BIO2014-54426-P) and the National Natural Science Foundation of China (31470394).

Author contributions

CZ conceived and designed the research. XJ, CB, LB, CN and CZ conducted the experiments. XJ, IR and RMT contributed statistical analysis. XJ, CB, LB, CN, IR, RMT, PC and CZ analyzed the data. XJ, CB, RMT, PC and CZ wrote the manuscript. All authors read and approved the manuscript. XJ and CB contributed equally to this work.

ORCID

Chao Bai https://orcid.org/0000-0002-9732-4276
Ludovic Bassie https://orcid.org/0000-0002-4791-0963
Paul Christou https://orcid.org/0000-0001-6358-7396
Xin Jin https://orcid.org/0000-0003-2684-1353
Carmina Nogareda https://orcid.org/0000-0002-6443-3388
Ignacio Romagosa https://orcid.org/0000-0001-6676-2196
Richard M. Twyman https://orcid.org/0000-0002-8764-3003
Changfu Zhu https://orcid.org/0000-0003-2834-2742

References

Bai C, Capell T, Berman J, Medina V, Sandmann G, Christou P, Zhu C. 2016. Bottlenecks in carotenoid biosynthesis and accumulation in rice endosperm are influenced by the precursor-product balance. *Plant Biotechnology Journal* **14**: 195–205.

Berman J, Zorrilla-López U, Sandmann G, Capell T, Christou P, Zhu C. 2017. The silencing of carotenoid β-hydroxylases by RNA interference in different maize genetic backgrounds increases the β-carotene content of the endosperm. *International Journal of Molecular Sciences* **18**: ~~pii: E~~2525.

Christou P, Ford TL, Kofron M. 1991. Production of transgenic rice (*Oryza sativa* L.) plants from agronomically important indica and japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. *Nature Biotechnology* **9**: 957–962.

Colot V, Robert LS, Kavanagh TA, Bevan MW, Thompson RD. 1987. Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO Journal* **6**: 3559–3564.

Diaz I, Martinez M, Isabel-La Moneda I, Rubio-Somoza I, Carbonero P. 2005. The DOF protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development. *The Plant Journal* **42**: 652–662.

Diaz I, Vicente-Carbajosa J, Abraham Z, Martínez M, Isabel-La Moneda I, Carbonero P. 2002. The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. *The Plant Journal* **29**: 453–464.

Farré G, Blancquaert D, Capell T, Van Der Straeten D, Christou P, Zhu C. 2014. Engineering complex metabolic pathways in plants. *Annual Review of Plant Biology* **65**: 187–223.

Fraser PD, Bramley PM. 2004. The biosynthesis and nutritional uses of carotenoids. *Progress in Lipid Research* **43**: 228–265.

Gallegos JE, Rose AB. 2015. The enduring mystery of intron-mediated enhancement. *Plant Science* **237**: 8–15.

Giuliano G. 2017. Provitamin A biofortification of crop plants: a gold rush with many miners. *Current Opinion in Biotechnology* **44**: 169–180.

Giuliano G, Tavazza R, Diretto G, Beyer P, Taylor MA. 2008. Metabolic engineering of carotenoid biosynthesis in plants. *Trends in Biotechnology* **26**: 139–145.

Gubler F, Kalla R, Roberts JK, Jacobsen JV. 1995. Gibberellin-regulated expression of a *myb* gene in barley aleurone cells: evidence for Myb transactivation of a high-pI alpha-amylase gene promoter. *Plant Cell* **7**: 1879–1891.

Gubler F, Raventos D, Keys M, Watts R, Mundy J, Jacobsen JV. 1999. Target genes and regulatory domains of the GAMYB transcriptional activator in cereal aleurone. *The Plant Journal* **17**: 1–9.

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET *et al.* 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **319**: 330–333.

Hernandez-Garcia CM, Finer JJ. 2014. Identification and validation of promoters and *cis*-acting regulatory elements. *Plant Science* **217–218**: 109–119.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Research* **27**: 297–300.

Hwang YS, Ciceri P, Parsons RL, Moose SP, Schmidt RJ, Huang N. 2004. The maize O2 and PBF proteins act additively to promote transcription from storage protein gene promoters in rice endosperm cells. *Plant and Cell Physiology* **45**: 1509–1518.

Jefferson AR, Kavanagh TA, Bevan MW. 1987. GUS fusion: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal* **6**: 3901–3907.

Kosugi S, Ohashi Y, Nakajima K, Arai Y. 1990. An improved assay for β-glucuronidase in transformed cells: methanol almost completely suppresses a putative endogenous β-glucuronidase activity. *Plant Science* **70**: 133–140.

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research* **30**: 325–327.

Li Q, Farré G, Naqvi S, Breitenbach J, Sanahuja G, Bai C, Sandmann G, Capell T, Christou P, Zhu C. 2010. Cloning and functional characterization of the maize carotenoid isomerase and β-carotene hydroxylase genes and their regulation during endosperm maturation. *Transgenic Research* **19**: 1053–1068.

Lu S, Zhang Y, Zhu K, Yang W, Ye J, Chai L, Xu Q, Deng X. 2018. The *Citrus* transcription factor CsMADS6 modulates carotenoid metabolism by directly regulating carotenogenic genes. *Plant Physiology* **176**: 2657–2676.

Martel C, Vrebalov J, Tafelmeyer P, Giovannoni JJ. 2011. The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner. *Plant Physiology* **157**: 1568–1579.

Marzábal P, Gas E, Fontanet P, Vicente-Carbajosa J, Torrent M, Ludevid MD. 2008. The maize Dof protein PBF activates transcription of gamma-zein during maize seed development. *Plant Molecular Biology* **67**: 441–454.

Mena M, Vicente-Carbajosa J, Schmidt RJ, Carbonero P. 1998. An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *The Plant Journal* **16**: 53–62.

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL *et al.* 2005. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nature Biotechnology* **23**: 482–487.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2nd edn. Plainview, NY, USA: Cold Spring Harbor Laboratory Press.

Suzuki A, Wu CY, Washida H, Takaiwa F. 1998. Rice MYB protein OSMYB5 specifically binds to the AACA motif conserved among promoters of genes for storage protein glutelin. *Plant and Cell Physiology* **39**: 555–559.

Toledo-Ortiz G, Huq E, Rodriguez-Concepcion M. 2010. Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proceedings of the National Academy of Sciences, USA* **107**: 11626–11631.

Vallabhaneni R, Gallagher CE, Licciardello N, Cuttriss AJ, Quinlan RF, Wurtzel ET. 2009. Metabolite sorting of a germplasm collection reveals the *hydroxylase3* locus as a new target for maize provitamin A biofortification. *Plant Physiology* **151**: 1635–1645.

Vicente-Carbajosa J, Moose SP, Parsons RL, Schmidt RJ. 1997. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. *Proceedings of the National Academy of Sciences, USA* **94**: 7685–7690.

Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F. 1999. Identification of *cis*-regulatory elements required for endosperm expression of the rice storage protein glutelin gene *GluB-1*. *Plant Molecular Biology* **40**: 1–12.

Welsch R, Maass D, Voegel T, DellaPenna D, Beyer P. 2007. Transcription factor RAP2.2 and its interacting partner SINAT2: stable elements in the carotenogenesis of *Arabidopsis* leaves. *Plant Physiology* **145**: 1073–1085.

Wu C, Washida H, Onodera Y, Harada K, Takaiwa F. 2000. Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene expression. *The Plant Journal* **23**: 415–421.

Xiong C, Luo D, Lin A, Zhang C, Shan L, He P, Li B, Zhang Q, Hua B, Yuan Z *et al.* 2018. A tomato B-box protein SlBBX20 modulates carotenoid biosynthesis by directly activating *PHYTOENE SYNTHASE 1*, and is targeted for 26S proteasome-mediated degradation. *New Phytologist* **221**: 279–294.

Yamamoto MP, Onodera Y, Touno SM, Takaiwa F. 2006. Synergism between RPBF Dof and RISBZ1 bZIP activators in the regulation of rice seed expression genes. *Plant Physiology* **141**: 1694–1707.

Yan J, Kandianis CB, Harjes CE, Bai L, Kim EH, Yang X, Skinner DJ, Fu Z, Mitchell S, Li Q *et al.* 2010. Rare genetic variation at *Zea mays crtRB1* increases β-carotene in maize grain. *Nature Genetics* **42**: 322–327.

Zhu C, Bortesi L, Baysal C, Twyman RM, Fischer R, Capell T, Schillberg S, Christou P. 2017a. Characteristics of genome editing mutations in cereal crops. *Trends in Plant Science* **22**: 38–52.

Zhu F, Luo T, Liu C, Wang Y, Yang H, Yang W, Zheng L, Xiao X, Zhang M, Xu R *et al.* 2017b. An R2R3-MYB transcription factor represses the transformation of α- and β-branch carotenoids by negatively regulating expression of *CrBCH2* and *CrNCED5* in flavedo of *Citrus reticulate*. *New Phytologist* **216**: 178–192.

Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P, Capell T. 2008. Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proceedings of the National Academy of Sciences, USA* **105**: 18232–18237.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Quantitative analysis of β-glucuronidase activity in transgenic rice.

Fig. S2 Alignment of maize and barley GAMYB sequences.

Fig. S3 Schematic representation of expression vectors pAHC25, pAL76-*ZmGAMYB* and pAL76-*ZmPBF*.

Fig. S4 Sequence of the maize phytoene synthase 1 (*ZmPSY1*) promoter and part of the cDNA region.

Fig. S5 RNA gel blot analysis of maize phytoene synthase 1 gene (*ZmPSY1*) expression in different tissues of maize.

Fig. S6 Quantitative RT-PCR analysis of maize phytoene synthase 1 gene (*ZmPSY1*) expression in endosperm and embryo determined by transient expression.

Fig. S7 Expression of maize transcription factor and β-glucuronidase genes in transgenic rice.

Fig. S8 Quantitative analysis of β-glucuronidase activity in transgenic rice.

Fig. S9 Expression analysis of maize *ZmGAMYB* and *ZmPBF* gene in transgenic rice.

Table S1 Primers used for RNA blot analysis.

Table S2 Primers used for β-glucuronidase reporter gene vector construction.

Table S3 Primers used to prepare the constructs pAL76-*ZmPBF* and pAL76-*ZmGAMYB*.

Table S4 Primers used for qRT-PCR analysis.

Table S5 Carotenogenic genes in maize (*Zea mays*).

Table S6 Partition of the observed variance in the relative expression of *gusA* and GUS activity in rice based on standard analysis of variance (ANOVA) and covariance (ANCOVA).

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.