

UC Davis

UC Davis Previously Published Works

Title

Novel synthetic inducible promoters controlling gene expression during water-deficit stress with green tissue specificity in transgenic poplar.

Permalink

<https://escholarship.org/uc/item/4df8q4sc>

Journal

Plant Biotechnology Journal, 22(6)

Authors

Yang, Yongil

Chaffin, Timothy

Shao, Yuanhua

et al.

Publication Date



2024-06-01

DOI

10.1111/pbi.14289

Peer reviewed

Novel synthetic inducible promoters controlling gene expression during water-deficit stress with green tissue specificity in transgenic poplar

Yongil Yang¹, Timothy A. Chaffin¹, Yuanhua Shao^{1,2}, Vimal K. Balasubramanian³, Meng Markillie³, Hugh Mitchell³, Maria M. Rubio-Wilhelmi⁴, Amir H. Ahkami³, Eduardo Blumwald⁴  and C. Neal Stewart Jr^{1,2,*} 

¹Center for Agricultural Synthetic Biology, University of Tennessee Institute of Agriculture, Knoxville, Tennessee, USA

²Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee, USA

³Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA

⁴Department of Plant Sciences, University of California, Davis, California, USA

Received 6 August 2023;

revised 16 November 2023;

accepted 3 January 2024.

*Correspondence (Tel 865 974 6487; fax 865 946 1989; email nealstewart@utk.edu)

Summary

Synthetic promoters may be designed using short *cis*-regulatory elements (CREs) and core promoter sequences for specific purposes. We identified novel conserved DNA motifs from the promoter sequences of leaf palisade and vascular cell type-specific expressed genes in water-deficit stressed poplar (*Populus tremula* × *Populus alba*), collected through low-input RNA-seq analysis using laser capture microdissection. Hexamerized sequences of four conserved 20-base motifs were inserted into each synthetic promoter construct. Two of these synthetic promoters (Syn2 and Syn3) induced GFP in transformed poplar mesophyll protoplasts incubated in 0.5 M mannitol solution. To identify effect of length and sequence from a valuable 20 base motif, 5' and 3' regions from a basic sequence (GTAACTTCAGGGCCTGTGG) of Syn3 were hexamerized to generate two shorter synthetic promoters, Syn3-10b-1 (5': GTAACTTCA) and Syn3-10b-2 (3': GGGCCTGTGG). These promoters' activities were compared with Syn3 in plants. Syn3 and Syn3-10b-1 were specifically induced in transient agroinfiltrated *Nicotiana benthamiana* leaves in water cessation for 3 days. In stable transgenic poplar, Syn3 presented as a constitutive promoter but had the highest activity in leaves. Syn3-10b-1 had stronger induction in green tissues under water-deficit stress conditions than mock control. Therefore, a synthetic promoter containing the 5' sequence of Syn3 endowed both tissue-specificity and water-deficit inducibility in transgenic poplar, whereas the 3' sequence did not. Consequently, we have added two new synthetic promoters to the poplar engineering toolkit: Syn3-10b-1, a green tissue-specific and water-deficit stress-induced promoter, and Syn3, a green tissue-preferential constitutive promoter.

Keywords: synthetic biology, synthetic promoter, green tissue specific promoter, water-deficit stress, *Populus*.

Introduction

Plant synthetic biology has made substantial contribution to sustainable applications for biofuel, pharmaceutical production, and food production using plant cellular systems (Yang *et al.*, 2022a). The current economic success of synthetic biological products calls for more effective techniques and plant platforms to be applied in industry settings (Clarke and Kitney, 2020; Fausther-Bovendo and Kobinger, 2021; Mahmood *et al.*, 2020). To date, various synthetic biology innovations have been applied to biosensors, crop development, and synthetic bioproducts using various plant resources. For these techniques, the design of synthetic promoters is an attractive approach for targeted gene expression in plants (Ali and Kim, 2019; Kumar *et al.*, 2020; Yasmeen *et al.*, 2023).

Native promoters cover many different CREs spanning hundreds of DNA sequences, which cause unnecessary complexity in targeted gene expression by recruiting superfluous transcription factors to overlap narrow DNA fragments in response to stress or developmental cues (Ali and Kim, 2019; Yasmeen *et al.*, 2023). However, in synthetic promoters, the small number of specific

CREs without non-essential elements allows researchers to off-target binding of transcription factors and focus on valuable responses, leading to more practical applications. In a number of recent studies, synthetic promoters have been applied to effectively manage gene editing, transgenic crop improvement, and orthogonal biosensor construction (Moreno-Gimenez *et al.*, 2022; Oliva *et al.*, 2019; Persad *et al.*, 2020; Persad-Russell *et al.*, 2022).

Synthetic promoters are constructed with a transcription factor (TF) that recognizes a sequence containing either a re-formatted CREs or simple repeats of a conserved CRE sequence, as well as a core promoter sequence including the binding site of RNA polymerase II. This format has been employed repeatedly for applications in herbaceous and woody plant species (Cai *et al.*, 2020; Sears *et al.*, 2023; Sultana *et al.*, 2022; Yang *et al.*, 2021, 2022b). Short synthetic promoters composed of head-to-tail repeats of conserved CREs and fused upstream of a – 46 bp 35S promoter can result in a shorter than 100 bp and functional promoter. For example, Sultana *et al.* (2022) demonstrated that the synthetic promoters were induced specifically by soybean cyst nematode application in engineered soybean roots (Liu *et al.*,

2014). In each of these studies, 6–7 copies of short conserved sequences were effective to induce gene expression to specific stimuli.

Synthetic promoters can be used to regulate precise spatio-temporal gene expression at specific developmental stages and tissues/organs. For this purpose, useful CREs must be identified from native promoters in specific tissue/organ at desired plant developmental stages. To date, tissue-specific native promoters have been reported in various tissues of different species. For example, two different vascular-tissue-specific native promoters were identified in switchgrass and *Arabidopsis* (Xu *et al.*, 2018; Zhang *et al.*, 2014). A switchgrass vascular-tissue-specific promoter (PvPfn2) of 1715 bp and its various 5'-end deleted fragment showed vasculature specificity in leaves, sheaths, stems, and flowers of heterologous transgenic rice (Xu *et al.*, 2018). The native promoter of *Arabidopsis thaliana* heat shock protein-related protein (AtHSPR) was also shown to positively induce a GFP reporter gene in vascular tissues of transgenic *Arabidopsis* (Zhang *et al.*, 2014). Additionally, promoters of 37 vascular specific expressed genes, including *AtHSPR*, employed several CRE motifs that responded to phytohormones, light, and other biotic and abiotic stimuli, demonstrating their tissue specificity (Zhang *et al.*, 2014). In addition, three native switchgrass promoters induced GUS marker gene expression in the above-ground organs of transgenic rice (Liu *et al.*, 2018). Those promoters responded to several photosystem-related proteins in the same regions of transgenic switchgrass by fusing with switchgrass MYB4 (Liu *et al.*, 2018). So far, a number of representative tissue- or organ-specific native promoters have been reported: root-specific promoters including tomato SIREO and rice RCc3 (Jones *et al.*, 2008; Xu *et al.*, 1995), the phloem-specific promoter of *Brassica juncea* (Koramutla *et al.*, 2016), seed-specific promoters rice glutelin (Wu *et al.*, 2000), sesame 2S albumin (Bhunja *et al.*, 2014), corn zein (Joshi *et al.*, 2015), and bean arcelin (De Jaeger *et al.*, 2002), and trichome tissue-specific promoters in mint (Ahkami *et al.*, 2015; Vining *et al.*, 2017). A developing xylem (DX) tissue-specific native promoter was cloned from *Pinus densiflora*, which was used to enhance plant biomass and produce biofuel precursors (Cho *et al.*, 2019; Ko *et al.*, 2012).

The majority of synthetic tissue-specific promoters have been designed for use in herbaceous plants. For instance, three synthetic promoters were designed to have root-specific drought inducible activity in transgenic *Arabidopsis* (Jameel *et al.*, 2020). These promoters were developed by combinations of 11 well-known motifs identified from 63 promoters of soybean drought-inducible genes. A minimal –46 to +1 CaMV 35S core promoter was fused with combinations of CREs (Jameel *et al.*, 2020). Additionally, four bidirectional tissue-specific synthetic promoters were shown to be active in green tissues of leaves, sheaths, panicles, and stems of transgenic rice (Bai *et al.*, 2020). Synthetic DNA fragments containing different types of CREs in rice P_{D540} showed various patterns of suppression or activation on gene expression in leaves, roots, young panicles and stems, depending on the selection of CREs, indicating that tissue-specificity may be tuned in rationally designed, tissue-specific promoters (Cai *et al.*, 2007). Recently, additional green-tissue- (Wang *et al.*, 2015), stem- (De Meester *et al.*, 2018), bidirectional vascular- (Lv *et al.*, 2009), and root-specific (Mohan *et al.*, 2017) synthetic promoters have also been developed. In woody plants, native tissue-specific promoters for induction of downstream gene expression in different organs and tissues have also been reported

(Ko *et al.*, 2012; Li *et al.*, 2018). With recent advancements in plant single cell biology and synthetic biology (Plant Cell Atlas *et al.*, 2021; Thibivilliers and Libault, 2021; Yang and Reyna-Llorens, 2023), we can expect the development of robust synthetic tissue and cell type-specific promoters in plants. Poplar stress responsive synthetic promoters have previously been designed from stress inducible promoters using open-source RNA-seq results of whole poplar tissues (Yang *et al.*, 2022b). However, these promoters did not show tissue specificity under water-deficit or high salinity treatment. By combining these methods with cell type-specific RNA-seq techniques, it is possible to generate advanced tissue specific and cell type-specific promoters.

The goal of this study was to expand the toolbox of poplar tissue-specific and water-deficit stress inducible synthetic promoters by leveraging cell type-specific (leaf palisade mesophyll and vascular bundle cells) RNA-seq data. Synthetic promoters containing hexamerized conserved DNA sequences of 20 bases were developed from promoters of differentially expressed genes in two major leaf cell types (palisade-mesophyll and vascular cells) under water-deficit stress conditions. Syn3 (20 bp core sequence) and its two derivatives (10 bp 5' or 3' of Syn3) were screened for tissue/organ-specificity in transgenic poplar under water-deficit stress.

Experimental procedures

Plants growth and water-deficit and salt stress treatment

Populus tremula × *Populus alba* INRA 717 1B4 clones (717-1B4) were propagated in solidified rooting media (RM: 1/2 Murashige and Skoog (MS) medium, 0.5 g/L MES, 30 g/L sucrose, 5 g/L activated charcoal, 3 g/L Phytoagar, 1 g/L Gelzan, and 0.1 mg/L IBA) in Magenta GA7 boxes (Bioworld, Dublin, OH, USA). The clones were grown in a growth chamber (Percival, Perry, IA, USA) under 16/8 h light/dark conditions at 25 °C with 150 μmol/m²s irradiance. Two-month-old root regenerated plants were transferred to ProMix BK25 soil (Premier Tech, Quakertown, PA, USA) in 1 L pots in the trays covered with lid. After acclimation for 2 weeks, lids were opened in the same chamber condition. Trees were watered every 2 days and fertilized with a 14-14-14 nutrient (Seed World, Odessa, FL, USA) solution once a month.

Nicotiana benthamiana seeds were stratified at 4 °C for 4 days in sterile water without light. The seeds were germinated and grown on ProMixBK25 potting mix in 8.9 cm square pots. *N. benthamiana* was grown under 16/8 h light/dark at 25 °C with 150 μmol/m²s light intensity. Plants were watered at three-day intervals.

To generate water-deficit conditions, 3-month established poplar plants in pots or 4 week-old *N. benthamiana* were withheld from watering until leaf-wilting was observed in growth chamber conditions.

Salt treatments were applied via watering with 250 mM NaCl every other day. Mock controls of plants were watered normally.

For cell type-specific RNA-seq analysis, water deficit treatments were applied 45 days after rooting, by withholding water until visual stress symptoms (i.e., leaf wilting) appeared (30%–35% relative soil water content) (early water deficit, EWD). Plants were kept for 10 days at 35% soil water content (late water deficit, LWD) and then re-watered with regular fertilization for 3 days

after sampling for the recovery period (REC). Leaf samples from each treatment were collected between 9 am and 10 am (1.30–2.30 h after the lights were turned on).

Cryosectioning and laser capture microdissection (LCM) for leaf cell types isolation

Leaf samples were harvested close to the midvein region comprising both palisade and vascular cell types. The leaf samples were directly embedded in cryomold filled with OCT[®] embedding polymer (VWR, Radnor, PA, USA) and the blocks were frozen using liquid nitrogen. These blocks were used for cryosection process using CryoStar[™] NX70 cryostat that generated 12 μ m cross sections of leaf tissue. The obtained cryosections were placed on PEN membrane glass slides (Thermo Fisher Scientific, Waltham, MA, USA) and washed using ice cold 70% ethanol for 2 min, 85% ethanol for 1 min, and 100% ethanol for 1 min. The slides were dried and used for the LCM procedure described previously (Balasubramanian *et al.*, 2021). Approximately 500–1000 mesophyll palisade cells and about 1000–2000 vascular cells were collected for transcriptome analysis (Figure S1).

RNA-seq and data analysis

RNA isolation and full-length cDNA generation were performed from LCM isolated cell clusters using SMART-Seq[®] v4 plus kit according to the manufacturer's protocol (Takara Bio USA, Mountain View, CA). The cDNA quality was determined by 5200 Agilent Fragment analyzer (Agilent Technologies, Santa Clara, CA). Single-read sequencing with a read length of 150 was performed by NextSeq 500 sequencing system using NextSeq 500/550 high output reagent kit v2.5 (150 cycles) (Illumina, San Diego, CA). With the generated RNA-seq reads, initial Quality control was performed using FastQC (Andrews, 2010). Read trimming was performed using BBduk script integrated in BBmap (Bushnell, 2014). Reads were aligned to the *Populus trichocarpa* v3.0 genome (https://phytozome-next.jgi.doe.gov/info/Ptrichocarpa_v3_0), using Bowtie 2 (Langmead and Salzberg, 2012). Aligned reads mapped to genome features were counted using htseq-count function in HTSeq 2.0 (Putri *et al.*, 2022). Differential expression gene (DEG) analysis was performed using DESeq2 (Love *et al.*, 2014).

DNA motif analysis for selection of water-deficit responsive and tissue-specific synthetic promoter generation

Promoter sequences for DNA motif analysis were profiled from BioMart integrated in Phytozome (v12.1.6; www.phytozome.org) (Goodstein *et al.*, 2012; Smedley *et al.*, 2015). Two kilobase (kb) upstream sequence from each promoter's ATG initial codon was curated from the *P. trichocarpa* v 4.0 genome annotation. The sequences were submitted to the command line application of MEME (v5.5.3) (Bailey *et al.*, 2015) to find conserved DNA motifs of 20 bases in length. We submitted conserved DNA motifs to the PLACE database of plant *cis*-acting regulatory DNA elements (PLACE, <https://www.dna.affrc.go.jp/PLACE/>) with default setting by host web site (Higo *et al.*, 1999).

Single stranded oligonucleotides composed of 6 repeats of a conserved DNA motif were commercially synthesized for synthetic promoter generation (Integrated DNA technologies, Coralville, IA; Table S1). Double stranded DNA fragments were generated by reannealing two complementary single-stranded oligonucleotides as described in our previous study (Yang *et al.*, 2021). These double strand DNA fragments were fused with the –46 to +1

CaMV 35S core promoter (47 bp) together with TMV 5' Ω - leader sequence (63 bp; Table S2). Turbo green fluorescence protein (GFP) or β -glucuronidase (GUS) coding sequences were placed downstream of the synthetic promoter as a reporter coding gene. Turbo red fluorescence protein (RFP) or luciferase (LUC) genes driven by 2 \times CaMV 35S short promoter (Table S2) were included in the T-DNA as a transformation control. Binary vectors for transient and stable transgenic transformation were generated by Golden Gate cloning using MoClo parts (Engler *et al.*, 2014; Werner *et al.*, 2012). Synthetic DNA motifs were inserted into backbone plasmids by Bsa I (New England Biolabs, Ipswich, MA) digestion followed by T7 ligase (New England Biolabs) ligation.

Transient transformation for synthetic promoter activity assay

Poplar mesophyll protoplast preparation, protoplast transformation, and image analysis were performed as described by Yang *et al.* (2021, 2023). For transient protoplast transformation and agroinfiltration assays in *N. benthamiana* leaves, the binary gene constructs incorporating the synthetic promoters fused with GFP were used (Figure 2c). *Agrobacterium* transformation, agroinfiltration of 1-month-old *N. benthamiana* leaves, water cessation, and subsequent fluorescence measurements were performed as previously described (Sparkes *et al.*, 2006; Yang *et al.*, 2021).

Fluorescent images of GFP and RFP expressed in poplar mesophyll protoplast were taken by EVOS M7000 image analyzer (Thermo Fisher Scientific). The fluorescence intensity of leaves was measured using a Fluorolog[®]-3 (HORIBA, Kyoto, Japan) fluorescence spectrophotometer using direct scans on the abaxial side of intact leaves.

Agrobacterium-mediated poplar leaf disk transformation with binary plasmids containing a GUS driven by various synthetic promoters

LUC coding sequence, in place of RFP, was driven by the CaMV 2 \times 35S promoter, while the *uidA* GUS gene replaced GFP and was driven by one of the designed synthetic promoters (Yang *et al.*, 2022b). Synthetic DNA motif digestion and ligation was performed as described above. See Figure 4 for plasmid illustrations.

The binary gene constructs were transfected into *Agrobacterium* (EHA105) via the freeze–thaw method using liquid nitrogen (Hofgen and Willmitzer, 1988). The protocol and reagent chemical composition for *Agrobacterium*-mediated poplar transformation were optimized with leaf disks from previously published methods (Song *et al.*, 2006). Leaf disks were cut from 2-month-old poplar plants that were grown in RM in Magenta GA7 boxes. Leaf disks were pre-cultured for 2 days in the dark on callus-induction media (CIM: 1 \times MS, 30 g/L sucrose, 0.25 g/L MES, 0.1 g/L myo-inositol, 1 \times vitamin solution, 1 μ M NAA, 3 g/L Phytoagar, 1 g/L Gelzan) plates without any antibiotics before dipping in *Agrobacterium* solution that had been cultured overnight in 5 mL of LB (pH 5.4; Fisher Scientific, Hampton, NH, USA) containing 50 μ g/mL of kanamycin (Sigma-Aldrich, St. Louis, MO, USA), 25 μ g/mL rifampin (Sigma-Aldrich), and 20 μ M acetosyringone (Sigma-Aldrich) at 28 $^{\circ}$ C with shaking (200–250 rpm). Leaf segments were then incubated for 3 days in the dark. Co-cultured leaf disks were washed three times in sterile water and then shaken in 40 mL of sterile water including 300 mg/mL of cefotaxime (Sigma-Aldrich) and 300 mg/mL of timentin (Plantmedia, Dublin, OH, USA) for 1 h. Excess liquid surrounding leaf disks was removed on autoclaved Whatman[™]

3MM paper (130 mm; Cytiva, MA) and then leaf disks were put on onto a new CIM plate that included hygromycin, cefotaxime, and timentin for selection. Callus induction, shoot regeneration, shoot elongation, and rooting were performed by following the same process as described in a published protocol (Song *et al.*, 2006).

PCR assays used genomic DNA that was extracted from a leaf of independently regenerated poplar using a DNeasy plant mini kit (Qiagen, Germantown, MD, USA). PCR was performed using DreamTaq DNA polymerase master mix (Thermo Fisher Scientific) with the reaction sequence as follows: 1 cycle of 95 °C for 2 min, 30 cycles repeating 95 °C for 30 s, 57 °C for 30 s, and 72 °C for 30 s, and 1 cycle of 72 °C for 7 min. Two separate reactions were performed using gene specific primers for LUC and GUS (Table S3). PtaUBCc primers were used as an internal control for genomic DNA qualification. Amplification fragments were separated on 1.2% agarose gels running in 1 × TAE.

GUS staining and fluorescent image analysis

The leaves at the 3rd to 5th position from the apical tip, as well as stem and root tissues, were collected from wilted and well-watered trees. Stem samples were collected from 10 to 15 cm below the primary shoot apex, and then sliced for GUS staining. Roots were collected directly below the potting mix surface. Mock transgenic control or wild-type plants were collected in the same manner from well-watered plants.

GUS staining in collected organs were performed by incubation for 6 h at 37 °C in GUS staining buffer (50 mM Na₂HPO₄ (Sigma, St. Louis, MO), 30 mM NaH₂PO₄ (Sigma), 0.4 mM K₃[Fe(CN)₆] (Sigma), 0.4 mM K₄[Fe(CN)₆] (Sigma), 0.1% Triton X-100 (Sigma), and 1 μg/mL 5-bromo-4-chloro-3-indol-glucuronide cyclohexylamine salt (X-Glu) (Gold Biotechnology, Olivette, MO). The stained tissues were bleached by washing overnight with 70% ethanol. Stained tissue images were taken by Stereomaster dissection microscope (Fisher Scientific).

Quantitative polymerase chain reaction (qPCR) analysis

Total RNA was extracted via Plant RNA extraction reagent following the manufacturer's protocol (Invitrogen, Carlsbad, CA). RNA extracts were purified through RNA binding columns using the Spectrum total plant RNA extraction kit (Sigma). Single-stranded cDNA was synthesized from 1 μg of total RNA by reverse transcription reaction with RevertAid first-strand cDNA synthesis kit following manufacturer's manual (Thermo Fisher Scientific). Single stranded cDNAs were two times diluted with nuclease-free water and subjected into 15 μL of total volume of 1 × PowerUp™ SYBR™ green master mix (Applied Biosystems, Waltham, MA) with gene-specific primers for each experiment. Primer sequences are listed in Table S1. The qPCR profile was 50 °C for 2 min, 95 °C for 10 min, and 40 cycles of 95 °C for 15 s and 60 °C for 30 s. The relative gene expression change was determined by the 2^{-ΔΔCt} equation (Livak and Schmittgen, 2001). Non-detected Ct values in qPCR such as GUS and LUC gene detections in wild-type samples' organs were replaced with a Ct value of 36 for calculation of estimated relative fold change compared to samples with detectable Ct values.

Results

Identification of leaf tissue specific water deficit stress-induced genes from cell type-specific RNA-seq analysis

To develop tissue-specific and water-deficit stress-induced synthetic promoters, it was necessary to identify the native promoter

sequences of cell type-specific and water-deficit stress upregulated genes. For this process, stress-induced cell type-specific genes (palisade mesophyll and vascular cell type) were identified through low-input transcriptomic analysis of the cell clusters collected by LCM from the leaves of 717-1B4 hybrid poplar under water-deficit stress. Herein, a total of 117 genes were detected to have higher relative expression (defined as a significant fold change of gene expression compared to the same tissue of mock control >4 and adjustive *P* value <0.05) either in leaf palisade or in vascular cell types of water-deficit-stressed leaves of *P. tremula* × *P. alba* (717-1B4) hybrid poplar under EWD, LWD, and REC conditions (Table S4). Ninety-eight of these genes were induced significantly in leaf palisade cells, while 16 genes were significantly induced in vascular cells (Data S1 and S2). To establish coinciding gene expression in other organs and confirm leaf-specificity, we selected three genes that were highly expressed in leaf palisade cells (*PtXaAlbH.02G059400*, *PtXaAlbH.13G089600*, and *PtXaAlbH.03G050900*) and another three genes which were highly expressed in leaf vascular cells (*PtXaAlbH.01G318500*, *PtXaAlbH.05G059900*, and *PtXaAlbH.09G114000*). Gene identity (ID) of *P. trichocarpa* homologues and their functional annotation are listed in Figure 1a. 717-1B4 gene-specific primers were used for qPCR analysis in leaf, stem, and root of water-deficit-stressed poplar (Figure S2). *PtXaAlbH.10G054500* (PtaExtensin) was used as a marker gene to determine water-deficit conditions. When expression of this gene was increased, a plant was assessed to be sufficiently drought-stressed (Figure 1b) (Dash *et al.*, 2018). In the selected palisade-expressed genes, *PtXaAlbH.02G059400* and *PtXaAlbH.13G089600* had higher expression in treated leaves and stems compared with those in mock plants (Figure 1b). Interestingly, the expression neither of these genes remained unchanged in roots under water cessation treatments.

Among the three selected genes displaying high expression in leaf vascular cells, *PtXaAlbH.01G318500* showed high expression in both roots and stems. However, *PtXaAlbH.05G059900* expression was increased during water deficit stress in leaves, stems, and roots (Figure 1b). Most of all, the relative fold changes (FCs) of the expression of both genes were increased by water-deficit conditions compared to the selected genes expressed in palisade cells.

Taken together, four of the selected genes of *PtXaAlbH.02G059400*, *PtXaAlbH.13G089600*, *PtXaAlbH.01G318500*, and *PtXaAlbH.05G059900* were significantly expressed in either leaves or stems during water-deficit stress. However, the gene expression of *PtXaAlbH.02G059400* and *PtXaAlbH.13G089600* was not affected in root tissue by water-deficit stress. These data confirmed that most of the genes selected based on our transcriptomic analysis were expressed in the green tissues under water-deficit stress. Therefore, we hypothesized that the promoters of these genes comprise CREs that were specific to green tissues and were induced by water-deficit stress.

Predicting and screening basic DNA sequences to design synthetic promoters by transient protoplast transformation

We showed previously that the synthetic promoter format using head-to-tail repeats of single CRE motifs is effective to induce marker gene expressions in both transient assays and in stable transgenic plants upon induction by stimuli such as high-salt or water-deficit treatments (Yang *et al.*, 2021, 2022b). To generate rationally designed synthetic promoters with tissue/organ

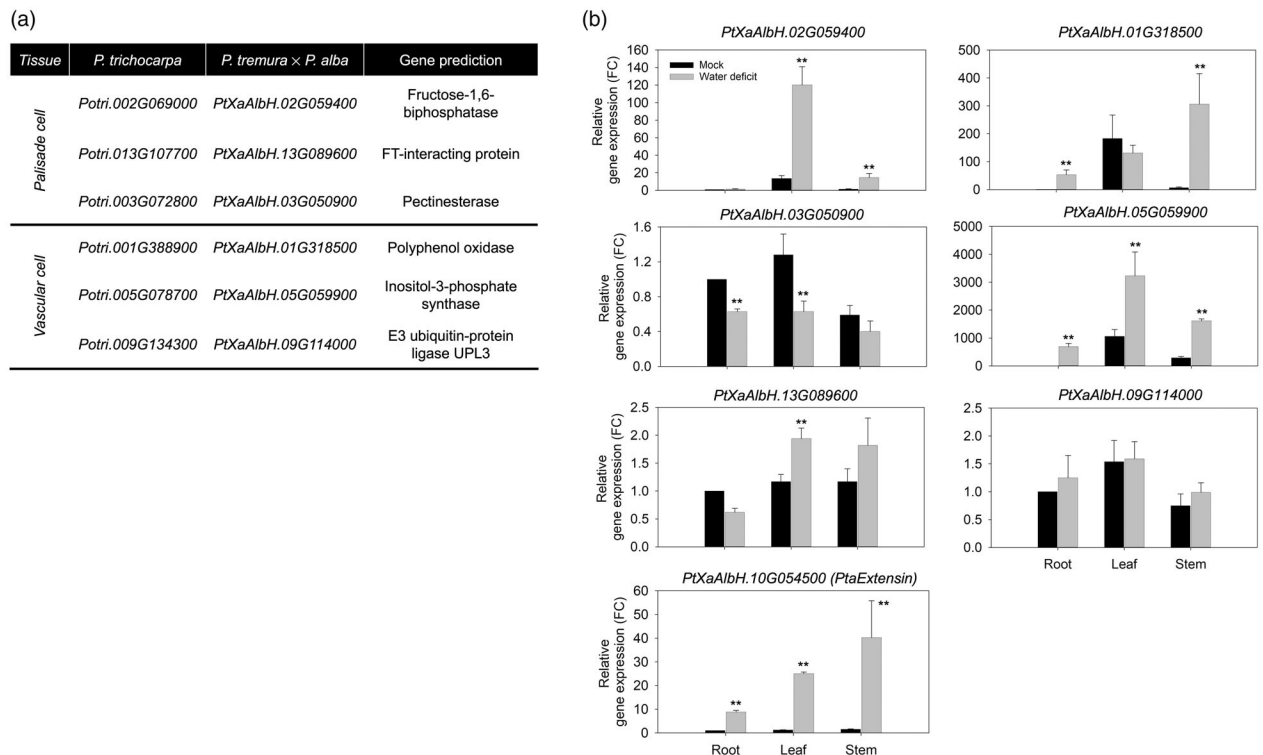


Figure 1 Endogenous gene expression of selected genes based on leaf-specific expression in single-cell RNA-seq analysis. (a) List of potential leaf-specific genes based on leaf single-cell RNA-seq analysis after water-deficit treatment. *P. trichocarpa* genome was used for their analysis. The homologue gene names and predictive functions in *P. tremula* × *P. alba* genome of three *P. trichocarpa* genes are listed. (b) Relative fold change of gene expression of 717-1B4 homologues (*PtXaAlbH.02G059400*, *PtXaAlbH.03G050900*, *PtXaAlbH.13G089600*, *PtXaAlbH.01G318500*, *PtXaAlbH.05G059900*, and *PtXaAlbH.09G114000*) in water-deficit stressed 717-1B4 root, leaf, and stem organs. Data display $2^{-\Delta\Delta Ct}$ value, determined relative to the Ct value of each target gene in mock root sample, which was set to 1. *PtaUBCc* expression was used for normalization of input cDNA. *PtaExtensin* was used as a marker gene to monitor water-deficit condition. Vertical bars represent means ± standard deviation (SD; $n = 3$). Statistical significance gene expression between mock and water-cessation condition was determined by *t*-tests (** $P < 0.01$).

specificity together with water-stress induction, we utilized the same process as our previous studies to identify DNA motifs of interest as sequence to be repeated in synthetic promoters for poplar. Two kilobase upstream sequences from the ATG initiation codon of co-expressed genes in leaf palisade (98 genes) or vascular tissues (16 genes) were submitted separately to predict well-conserved DNA motifs (Table S4). A total of eight conserved sequences of 20 bases in length were predicted with high significance after removal of TATA boxes or single sequence repeats from both analyses ($E < 1 \times 10^{-3}$; Figure 2a). The potential CREs for each was predicted by PLACE and listed in Table 1.

Among eight conserved 20 base sequences, we selected the four sequences with the highest *E*-values to generate synthetic promoters by 6 repetitions of these sequences (total of 120 bp in length) without any spacing or other factors inserted (Figure 2b). When the plasmid constructs (Figure 2c) were transformed into poplar mesophyll protoplasts in 0.5 M mannitol incubation buffer, all transformed protoplasts displayed RFP fluorescence expression, indicating that the protoplast transformation was successful (Figure 2d). However, GFP expression in transformed protoplasts was observed only with Syn2 and Syn3 promoters. Since Syn3 induction was observed in relatively more protoplasts, Syn3 was selected as a template sequence to build a series of synthetic promoters for subsequent analyses.

Agroinfiltration assay with two different hexamerized short sequences (60 bp) and a long sequence (120 bp) on *N. Benthamiana* leaves

New shorter promoter versions were synthesized from the Syn3 sequence since it had the highest activity among the tested 120 bp promoters. Six repeats of the 5' Syn3 sequence (6 × 10 bp; Syn3-10b-1) as well as the Syn3 3' sequence (6 × 10 bp; Syn3-10b-2) were used to replace the full-length sequences of Syn3 in the binary vector used in protoplast transformation (Figure 3a,b). Syn3-10b-1 had no functional CREs reported while Syn3-10b-2 had GGGCC, which has a predicted function for light-induction (Table 1; Figure 3a red box). Since Syn3-10b-1 and -2 were derived from Syn3, the 120 bp of Syn3 was tested along with the two short promoters to compare activities in agroinfiltrated *N. benthamiana* leaves. All of the infiltrated leaves including a vector without any synthetic promoter insertion had lower GFP intensities in mock conditions (Figure 3d). RFP fluorescence was detected in a similar range of fluorescence in all transient transformed leaves, showing that agroinfiltration was successful. The wilted infiltrated leaves resulting from 5-day water cessation treatment, that begun 2 days after infiltration, displayed 2- to 3-fold higher GFP fluorescence in Syn3-10b-1 and Syn3 than in mock conditions. However, Syn3-10b-2 had

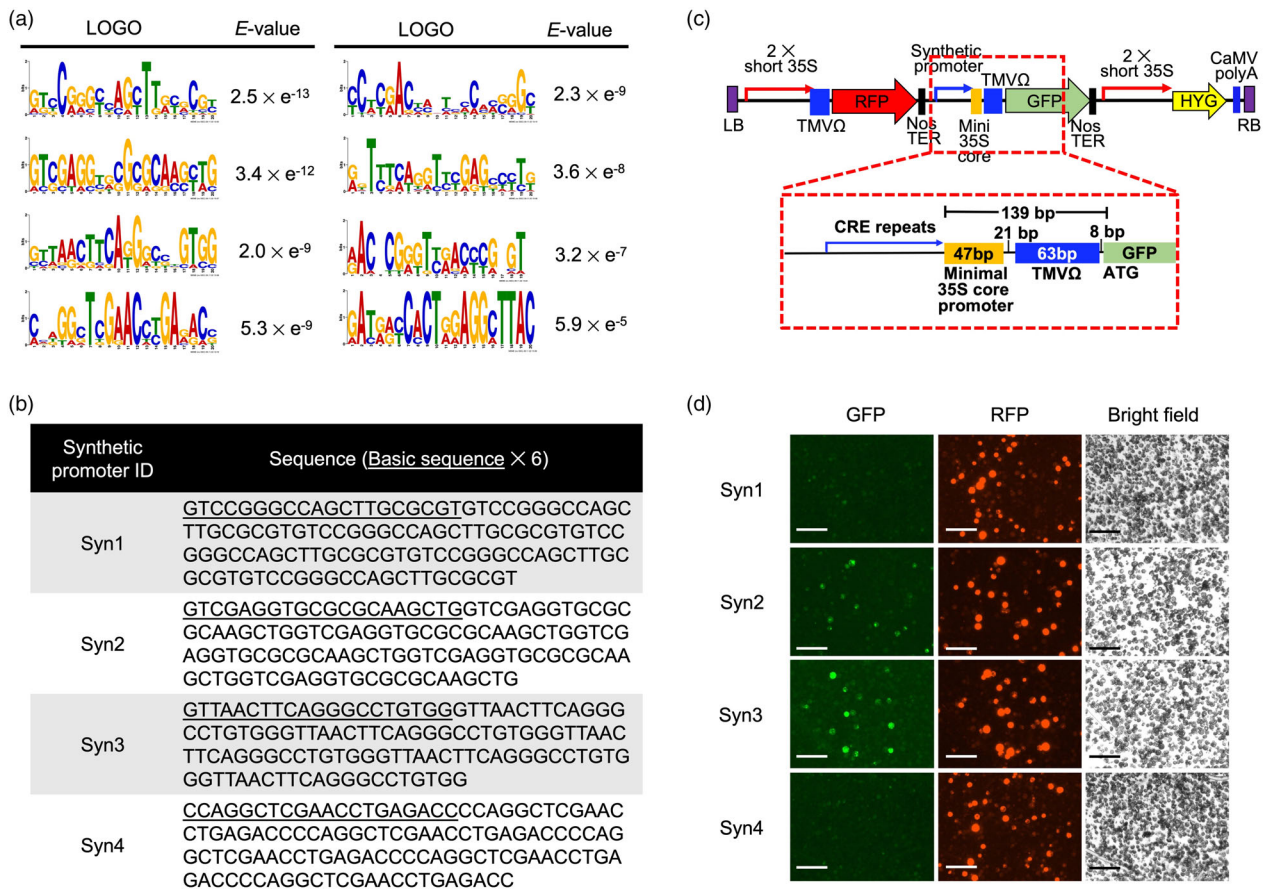


Figure 2 Basic sequence selection for repeats in synthetic promoter gene construct and subsequent screening using poplar mesophyll protoplast transformation. (a) List of conserved DNA sequences. The conserved regions were determined by MEME from 114 co-expressed genes in leaf organs based on single-cell RNA-seq analysis (Table S4). (b) List of synthetic promoters composed of six repeats of a 20-base sequence inserted into a GFP-fused vector as shown in panel c. The plasmids including the synthetic promoter were transformed in poplar mesophyll protoplast and then screened for GFP and RFP expression. (c) Plasmid construction including synthetic promoters. Synthetic promoter was flanked with 139 bases of minimal CaMV 35S core promoter together with a TMV 5'Ω – leader sequence to initiate GFP expression. RFP was constitutively induced by the CaMV35S promoter as a transformation efficiency indicator. (d) GFP and RFP expression images in transformed mesophyll protoplast in 0.5 M mannitol induction solution. Syn2 and Syn3 promoters induced GFP in higher osmotic conditions (bar = 100 μm).

Table 1 CRE prediction of conserved DNA motif by PLACE

Conserved DNA motif	Predicted CRE in PLACE	Sequence
GTCCGGGCCAGCTTGC GCGT	Sequences over-represented in light-induced promoters	GGGCC
	CGCG box	CGCG
GTCGAGGTGCGCGCAAGCTG	CGCG box	CGCG
GTTAACTTCAGGGCCTGTGG	Sequences over-represented in light-induced promoters	GGGCC
CCAGGCTCGAACCTGAGACC	Sulphur-responsive element	GAGAC
CCTGACTAATCCACGGGC	<i>Arabidopsis</i> ARR1 response	AGATT
GGTTTCAGGTTCCAGCCCTG	No prediction	No prediction
AACNCGGGTTGACCCGNGT	No prediction	No prediction

very low GFP intensity similar to vector transformed plants under water-deficit stress (Figure 3d). Interestingly, salt-treatments (250 mM NaCl) did not alter GFP expression (Figure 3d). Thus, we concluded that Syn3 and Syn3-10b-1 promoters were significantly and specifically induced in response to water-deficit stress.

Although Syn3-10b-1 had only half the length of the entire Syn3 sequence, GFP expression was similar between the two promoter constructs (Figure 3d). Given that the 3' sequence (Syn3-10b-2) did not appear to be water stress-induced, we concluded that the critical CRE for water-stress induction was in the Syn3-10b-1 sequence. Noteworthy, although Syn3-10b-2 contained the predicted CRE, GGGCC, for green tissue expression (Table 1, Figure 3a red box), this promoter had apparently the lowest expression in mock and water-deficit treated agroinfiltrated leaves (Figure 3d).

GUS induction in the stable transgenic poplar containing Syn3-10b-1 (*Syn3-10b-1*) in water-deficit conditions

Based on the transient transformation assays, we observed that Syn3 and Syn-10b-1 have water-deficit-specific inducibility in leaves (Figures 2d and 3d). To assess whether the synthetic promoters respond to water-deficit conditions with tissue/organ specificity *in planta*, we engineered poplar with synthetic constructs in which GUS was the induced reporter gene and LUC was the internal control reporter gene under the control of the 35S promoter (Figure 4). PCR analyses showed that most

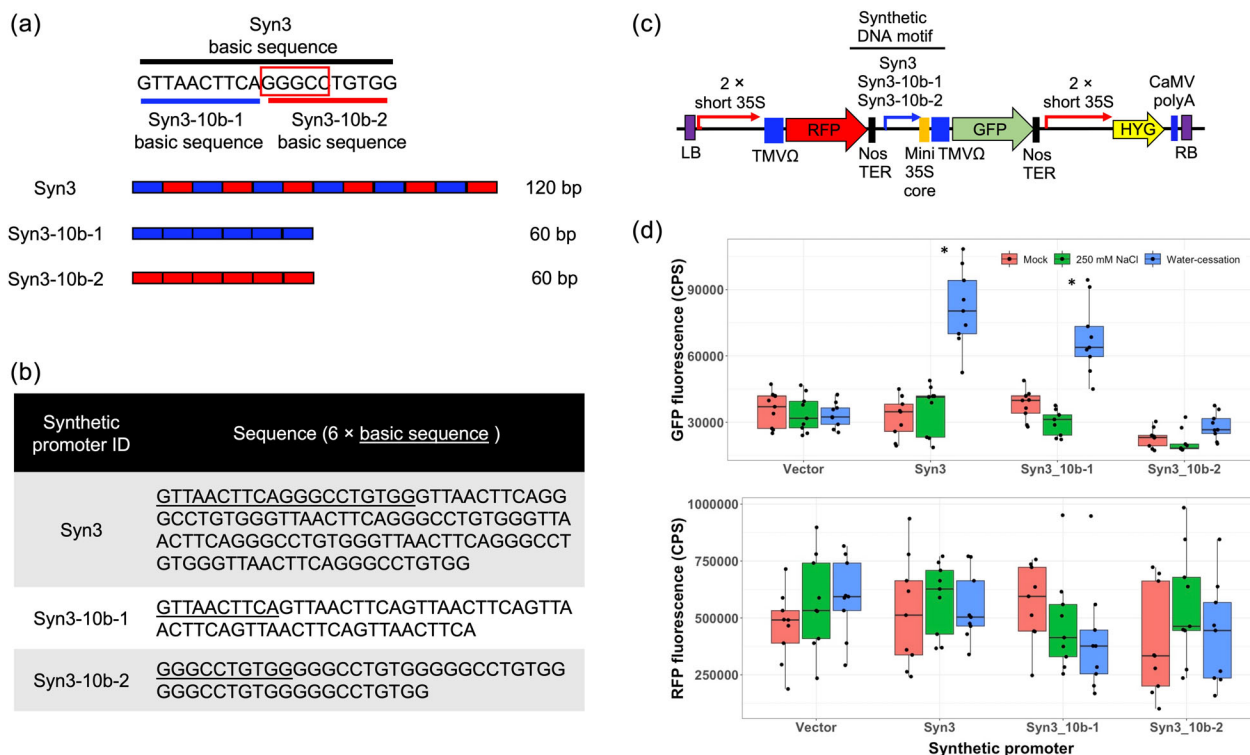


Figure 3 Transient transformation assay of Syn3 and its two derivative synthetic promoters using agroinfiltration of *N. benthamiana* leaves under high salt and water-deficit conditions. (a) Diagram of basic sequences of six-repeat DNA fragments for Syn3 and two derivatives from Syn3. The 5' 10-base sequence (Syn3-10b-1; blue line) and 3' 10 (Syn3-10b-2; red line) from 20 initial bases of Syn3 were hexamerized in synthetic promoters. The synthetic DNA fragments were cloned into the same vector diagrammed in Figure 2b. Red box displays a predicted CRE sequence region by PLACE. (b) The sequence list of hexamerized DNA motifs. (c) Plasmid gene construct including Syn3, Syn3-10b-1, and Syn3-10b-2 for Agroinfiltration of *N. benthamiana* leaves. (d) GFP and RFP intensity on agroinfiltrated leaves of *N. benthamiana* stressed by high salt treatment or water cessation. Note that all transformations were performed successfully based on RFP expression, and GFP was induced only in water-deficit leaves of Syn3 and Syn3-10b-1. None of the tested synthetic promoters responded to salt stress. Statistical significance of gene expression between mock and water-cessation conditions was determined by *t*-tests (**P* < 0.05).

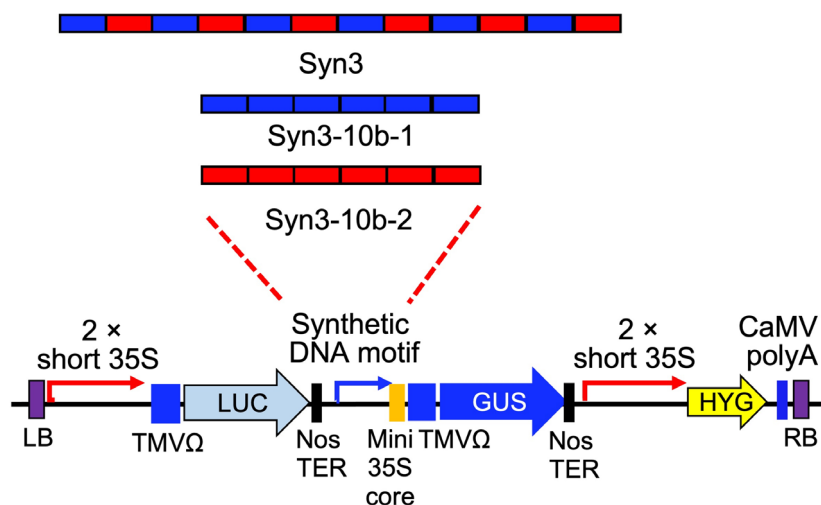


Figure 4 Binary vector construct containing GUS gene driven by synthetic promoter introduced in transgenic poplar. Binary vectors contained a series of synthetic DNA motifs (Syn3, Syn3-10b-1, and Syn3-10b-2), sequences of which were listed in Figure 3b. Synthetic DNA motifs fused with minimal CaMV 35S core promoter and TMV 5'Ω – leader sequence for assistance of GUS expression. Vectors also contained 2 × short 35S driven LUC cassette as an internal control for monitoring transformation efficiency in transgenic trees.

transgenic events contained both GUS and LUC (Figure S3). These events were used in subsequent experiments. No phenotypical differences were seen in regenerated transgenic poplar plants during the 3 months from transplanting into potting mix before water deficit treatment (Figure S2).

We measured Syn3-10b-1's inducibility to water-deficit stress by GUS staining when leaves were wilted after water cessation (~5 days without watering) using well-watered plants as controls. GUS expression resulted in a faint blue colouring on the leaf surface even in mock samples (Figure 5a). Stem slices had slightly darker blue GUS staining at the primary phloem region including the vascular cambium layer on the outside of stem sections. However, GUS-stained cells were not visibly detected in any root samples of all three different transgenic lines, indicating that GUS expression driven by Syn3-10b-1 was slightly activated under non-stress conditions in green tissues but not in root tissues. After water cessation treatment, leaves and stems of wilted transgenic poplars had stronger and more extensive blue staining than mock (Figure 5a). Interestingly, the pith including primary xylem and vascular phloem had the darkest staining as compared to other organs. Based on these observations, the Syn3-10b-1 synthetic

promoter was deemed to be highly activated in leaf and stem tissues under water-deficit conditions.

To validate Syn3-10b-1's positive induction of GUS gene expression as a response to water-deficit conditions, GUS transcript abundances were compared in three organs of leaf-wilted transgenic poplar versus mock controlled plants. As a transformation control, LUC genes were detected in a range from 30- to 600-fold relative changes in all transgenic lines whereas wild-type poplar did not show any LUC gene expression (Figure 5b). Water-deficit stress did not induce more LUC expression in any of the assessed tissues. Meanwhile, GUS transcript abundance was increased in leaf and stem tissues under water-deficit conditions. In leaf samples, GUS was expressed from 4- to 8-fold higher than mock in all transgenic poplar. The stems of three water-deficit transgenic plants had about 3-fold higher GUS transcript than mock controls. However, GUS gene expressions were not increased in water deficit roots. GUS gene expression was minimally detected in all organs of mock samples, suggesting that Syn3-10b-1 may not solely respond in leaf and stem tissues. However, GUS expression was more significantly induced in green tissues/organs than roots under

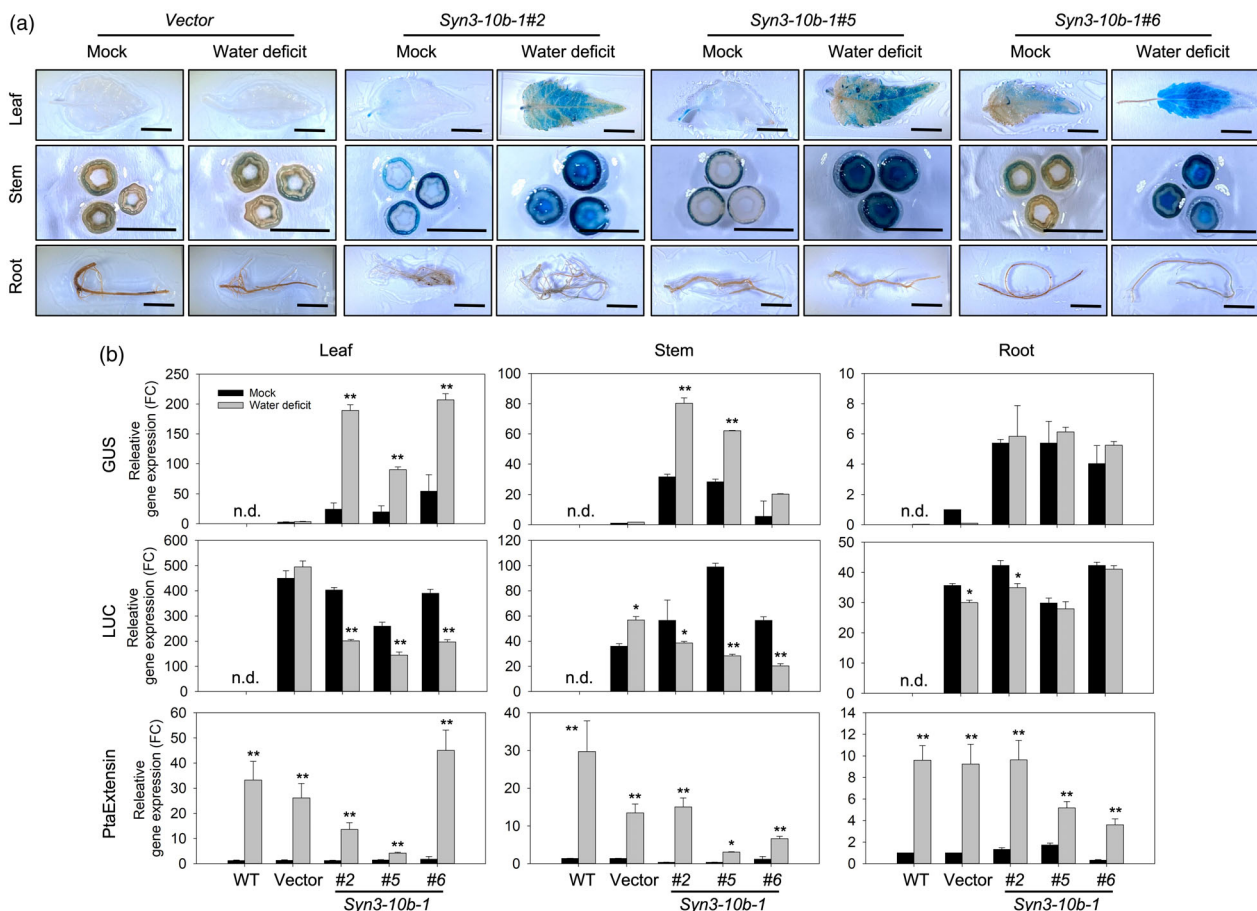


Figure 5 GUS induction in leaf, stem, and root tissues of regenerated Syn3-10b-1 transgenic poplar growing in soil under water-deficit stress. (a) X-glu staining of leaf, stem, and root tissue of transgenic plants with or without water cessation. When the leaves wilted in water-deficit stress plants, samples were collected from leaf, stem, and root tissues and incubated in X-glu solution for 6 h. Chlorophyll was bleached overnight with 70% ethanol (bar = 1 cm). (b) GUS gene abundance in water-deficit stressed Syn3-10b-1 transgenic poplars. Total RNA was extracted from root, leaf, and stem of water-deficit or well-watered (mock) transgenic poplar growing in soil. PtaUBCc was used as internal control for normalization. The relative gene expression was determined by comparison with the expression level of GUS, LUC, and PtaExtensin gene expression in root organs of wild type in mock control. No detection in qPCR denoted as n.d. Bar displays means \pm SD ($n = 3$). Significant differences of fold change of gene expression between mock and water-deficit stressed tissues were determined by t -test (* $P < 0.05$; ** $P < 0.01$).

water-deficit stress (Figure 5b). These observations were in agreement with GUS staining results (Figure 5a).

GUS gene induction by Syn3-10b-2 in stable transgenic poplar (*Syn3-10b-2*) under water-deficit conditions

Two lines containing the Syn3-10b-2 promoter sequence had no visible GUS expression, but Syn3-10b-2#3 showed weak GUS staining. Nonetheless, there was no apparent water-stress induction in any events (Figure 6a).

GUS gene transcripts were quantified in leaf, stem, and root samples, but the levels were about 25-fold less than in Syn3-10b-1 (Figure 6b). Additionally, water cessation did not affect GUS gene expression in any tissues. The consistent LUC gene expression showed that the transgenes were introduced and functioning properly in transgenic poplar. However, the lack of increase in GUS expression confirmed that the lack of water-deficit stress inducibility of Syn3-10b-2.

GUS gene expression in stable transgenic poplar containing Syn3 (*Syn3*) under water deficit conditions

The two short synthetic promoters, Syn3-10b-1 and Syn3-10b-2, responded differently to water-deficit stress in transgenic poplar,

even though they were both derived from the basic Syn3 sequence. To further explore this phenomenon, we tested the hexamerized 20-base Syn3 sequence in a synthetic promoter in transgenic poplar. In the mock treatment, three randomly selected transgenic poplar events had GUS expression in most parts of the leaf unlike Syn3-10b-1 and Syn3-10b-2 (Figure 7a). GUS staining was also observed in vascular cambium and primary phloem regions as well as in primary xylem and pith of Syn3 stems. Unlike Syn3-10b-1 and Syn3-10b-2, the root tips of two transgenic lines also had GUS expression. However, GUS stain detection in green tissues appeared to be less than in mock. Consistent GUS staining in different organs was confirmed by comparing GUS transcript abundance between mock and water-deficit stress conditions (Figure 7b). The GUS transcripts were detected in all Syn3 transgenic lines; however, no significant of GUS gene increase was observed in water-deficit stressed tissues compared to mock.

Thus, Syn3 appears to be a green-tissue active promoter with approximately a 10-to-12 fold less activity in roots compared with leaves of transgenic poplar. The Syn3 was not induced by water-deficit stress treatment. It is a moderately strong promoter in leaves compared to Syn3-10b-1.

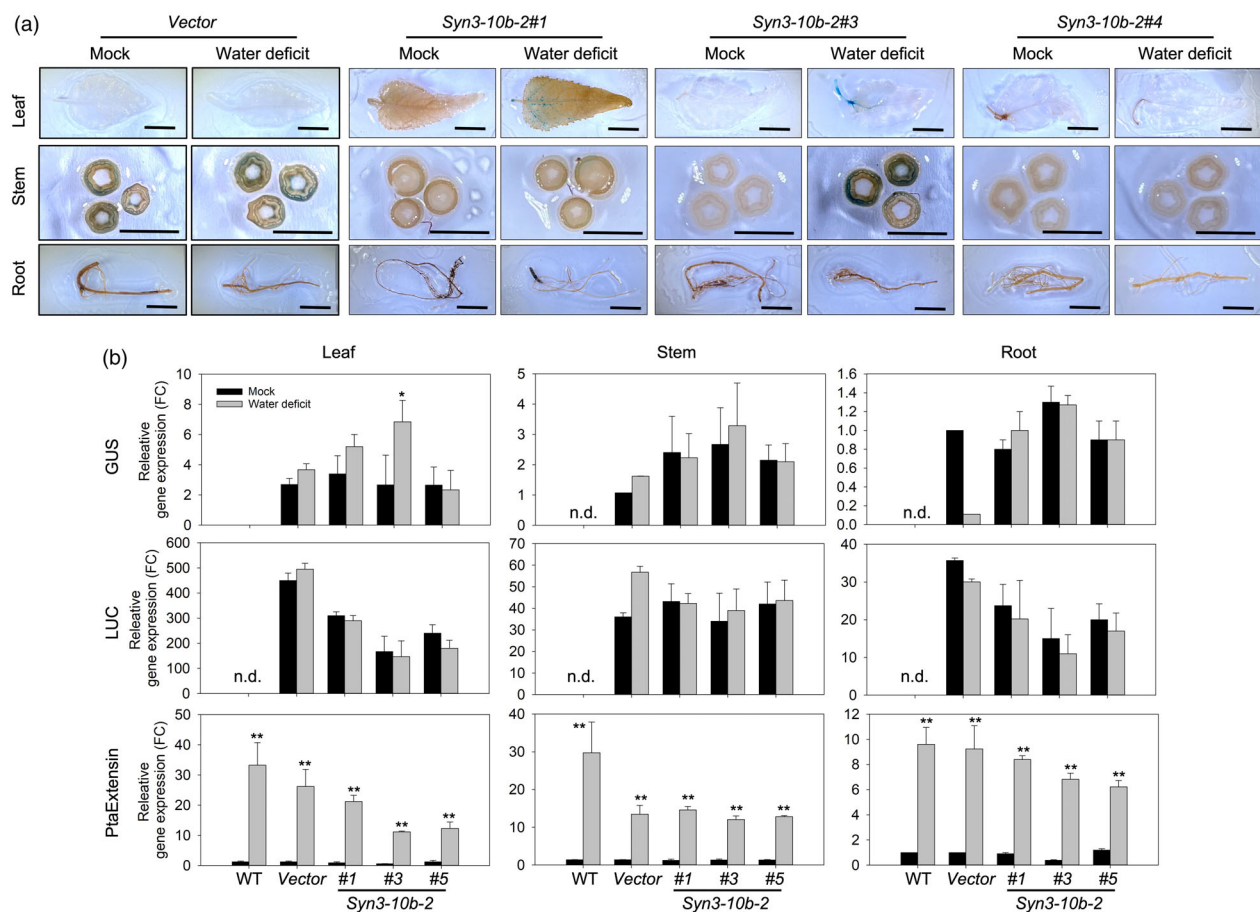


Figure 6 GUS expression in leaf, stem, and root of regenerated Syn3-10b-2 transgenic poplar growing in soil under water-deficit conditions. (a) X-gluc staining of leaf, stem, and root tissues of transgenic plants with or without water-deficit treatment. When leaves wilted under water cessation, leaf, stem, and root tissues were collected from plants, and incubated in X-gluc solution for 6 h. Chlorophyll was bleached overnight with 70% ethanol (bar = 1 cm). (b) GUS gene induction in water-deficit stressed Syn3-10b-2 transgenic poplars. Total RNA was extracted from root, leaf, and stem tissues of water-deficit transgenic poplar growing in soil. PtaUBCc was used as internal control for normalization. Relative gene expression was determined by comparison of GUS, LUC, and PtaExtensin gene expression in well-watered roots of wild type. No detection in qPCR denoted as n.d. Bar displays mean \pm SD ($n = 3$). Significant differences of fold change of gene expression between mock and water-deficit stressed tissues were determined by *t*-test (* $P < 0.05$; ** $P < 0.01$).

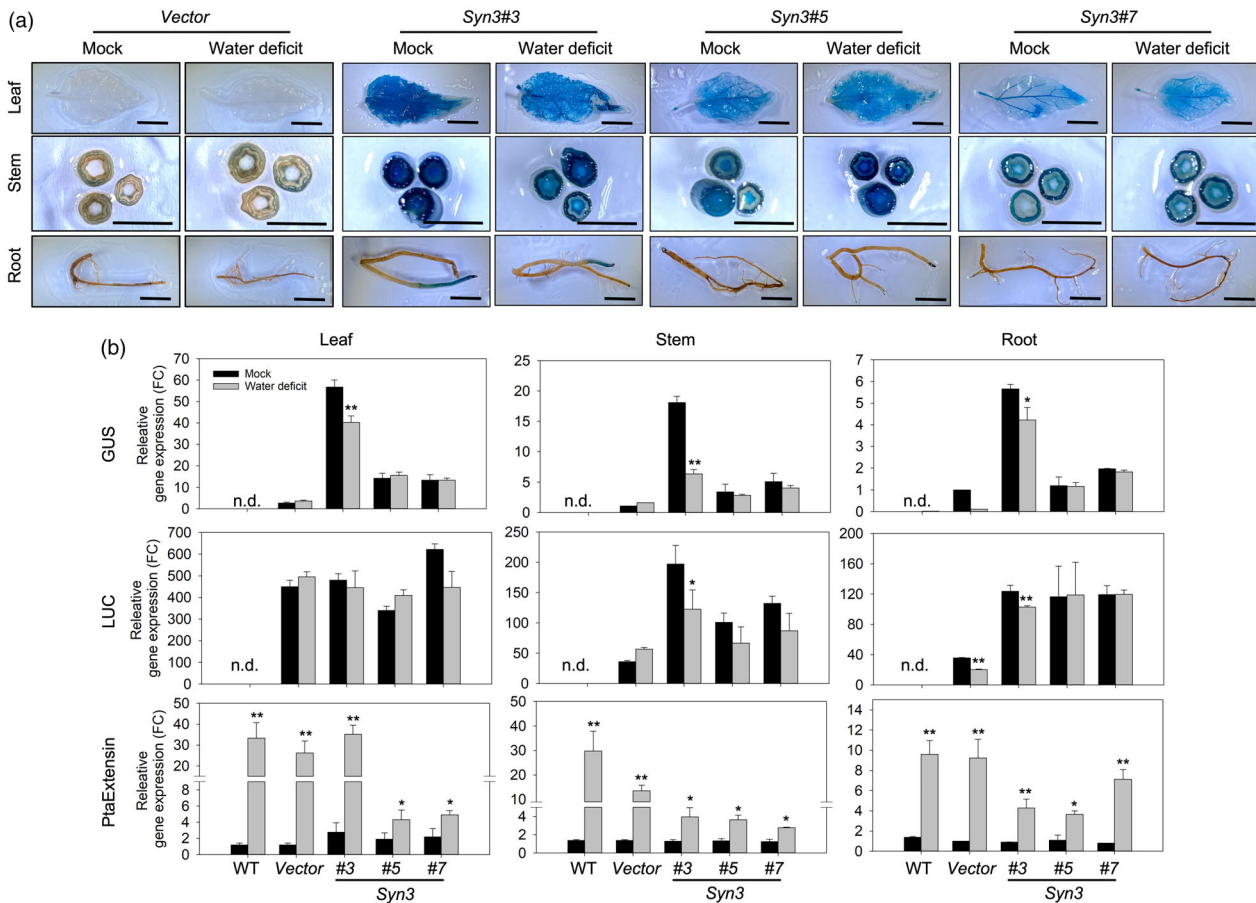


Figure 7 GUS expression in leaf, stem, and root of regenerated Syn3 transgenic poplar growing in soil under water-deficit conditions. (a) X-gluc staining of leaf, stem, and root tissues of transgenic plants with or without water-deficit stress. When the leaf was wilted during water cessation, all samples were collected from leaf, stem, and root tissues and stained in X-gluc solution for 6 h. Chlorophyll was bleached by overnight shaking in 70% ethanol (bar = 1 cm). (b) GUS gene induction in water-deficit stressed Syn3-10b-1 transgenic poplars. Total RNA was extracted from root, leaf, and stem tissues of water deficit transgenic poplar growing in soil. PtaUBCc was used as internal control for normalization. Relative gene expression was determined by comparison of GUS, LUC, and PtaExtensin gene expression in well-watered root organ of wild-type poplar. No detection in qPCR denoted as n.d. Bar displays mean \pm SD ($n = 3$). Significant differences of fold change of gene expression between mock and water-deficit stressed tissues were determined by t-test (* $P < 0.05$; ** $P < 0.01$).

Discussion

Synthetic promoters are useful to spatiotemporally tune transgene expression in plants. In this study, we designed a synthetic promoter that has green-tissue specific activity with water-deficit stress inducibility as well as another quasi-constitutive synthetic promoter. These synthetic promoters are apt additions to the transgenic poplar toolbox. These promoters were constructed by six repeats of a head-to-tail array of conserved sequences found in the collection of 2 kb native promoters whose downstream genes were highly induced in leaf tissues under water-deficit treatment, based on LCM-based cell type-specific transcriptome analysis (Figures 2 and 3). Two green-tissue-specific inducible promoters, Syn3 and Syn3-10b-1, induced the downstream GUS gene in stable transgenic hybrid poplars (Figures 5 and 7). Interestingly, Syn3-10b-1 responded to water-deficit stress in stable transgenic poplar while Syn3 did not, despite containing the sequence of Syn3-10b-1 within its larger sequence. Furthermore, the Syn3-10b-2 synthetic promoter based on the 3' sequence of Syn3 did not induce reliable tissue-specificity or

stress responsiveness (Figure 6). Additionally, the sequence of Syn3-10b-2 did not improve or repress Syn3-10b-1 positive activity in Syn3 (Figure 7). To summarize, the 5' sequence in Syn3 seems to function predominantly for tissue-specificity and water-deficit stress in transgenic poplar, whereas the 3' sequence has no notable response and may even inhibit water deficit response when included in the synthetic promoter sequence (Figures 3 and 5–7). It is evident that synthetic promoter design is still reliant on empirical testing since CREs cannot be validated solely by *in silico* analysis. However, more precise experiments with synthetic promoter constructs containing mutant sequence of present synthetic promoters should also be required to clarify tissue specificity and water-deficit stress responsibility of present synthetic promoters in future.

The traditional architecture of synthetic promoters is consisted of the three parts: (1) the proximal sequence, which is recognized by TFs, (2) a core promoter which binds RNA polymerase, and (3) distal enhancer or repressor sequences to modify transcriptional initiation (Dey *et al.*, 2015). Among these components, the proximal sequences-composed of CREs or a DNA fragment based

on a reliable native promoter-can play a pivotal part in deciding the synthetic promoter's properties and activity. Recent reviews have noted current achievements using well-known CREs in traditional synthetic promoter constructs in various plant species (Ali and Kim, 2019; Yasmeen et al., 2023). The synthetic promoter structure utilizing repeats of a known CRE fused with a minimal core promoter is a rational design that usually functions as designed. To date, our colleagues have facilitated in tandem repeats of well-known CREs fused with a short -46 CaMV 35S core promoter to produce biotic and abiotic stress-responsive phyto-sensors in *Arabidopsis*, soybean, potato, and poplar (Liu et al., 2014; Persad-Russell et al., 2022; Sears et al., 2023; Sultana et al., 2022; Yang et al., 2021, 2022b). Among these studies, pathogen- and radiation-sensing synthetic promoters were developed from copies of the well-known CREs of the S-box pathogen sensing DNA motif and the DNA recognition motif in the *Arabidopsis* RAD51 promoter, following the traditional process for synthetic promoter generation (Persad-Russell et al., 2022; Sears et al., 2023).

To identify new CREs, we and our colleagues have been attempting to find unknown DNA motifs by computational analysis for novel synthetic promoter construction. To do so, DNA motifs are derived from conserved DNA sequences in native promoters of co-expressed genes of interest in species of interest: typically, non-model plants. Recently, several tandem repeats with unknown DNA motifs have been successfully introduced in transgenic soybean and poplar, and clearly induced downstream genes relating to nematode infection in soybean (Liu et al., 2014; Sultana et al., 2022) and abiotic stresses in poplar (Yang et al., 2021, 2022b) from transcriptome analysis including naive microarray analysis. Although these synthetic promoters were activated in transgenic plants by abiotic or biotic stimuli, the responses were not specific to cell types or tissue. The present study more precisely predicted novel DNA motifs from cell type-specific RNA-seq data. Promoter responsiveness was dependent on sequence selection and formation for water-stress response together with tissue/organ specificity, which allows more specific regulation via synthetic promoters than previous achievements.

Notably, Syn3-10b-2 had a CRE motif for light inducibility (GGGCC) in its basic sequence, but it did not endow discernible expression in transiently transformed *N. benthamiana* and stable transformed poplar under water-deficit stress (Figures 2 and 6). Since GGGCC is detected commonly in the most promoter of genes, no response of Syn3-10b-2 in transgenic poplar should be predicted. On the other hand, although no CREs were predicted in the basic sequence of Syn3-10b-1, this synthetic promoter had a water-deficit stress response with specificity for green tissues (Figures 2 and 5). These results indicate that genome-wide databases of CREs are still lacking complete CRE information, and further studies will need to functionally test well-known CREs as well as unknown DNA motifs to develop the most reliable synthetic promoters for non-model plants. Elucidating conserved DNA sequences based on longer promoter sequences may be a good alternative to identify optimal CRE information, as performed in this study. A number of recent successes in high-throughput techniques have given critical information about CRE sequences that bind TFs in diverse plant developmental stages and environmental responses (Marand et al., 2023; Schmitz et al., 2022; Yasmeen et al., 2023). Rapid identification of genome-wide TF binding DNA motifs could be performed generally by chromatin immunoprecipitation sequencing (ChIP-seq) (Lihu and Holban, 2016). Recently, more high-

throughput techniques along with various TF capturing methods such as high-throughput sequencing-fluorescent ligand interaction profiling (HiTS-FLIP), ATAC-seq with nuclei sorting, DNA affinity purification sequencing (DAP-seq), and crosslinking immunoprecipitation sequencing (CLIP-seq) have been used to get the information of genome-wide DNA binding motifs for single TFs (Bartlett et al., 2017; Chen et al., 2008; Liu et al., 2002; Lu et al., 2017; Nutiu et al., 2011). With increasing information of CREs via high throughput techniques, more feasible and reliable designs of synthetic promoters can be made possible.

The spacing between CRE copies and the proximity of CRE copies relative to the core promoter are also known to be important factors affecting synthetic promoter activity (Yasmeen et al., 2023). In a previous study, 6 copies of 7-bp of ABRE were spaced with a 23 bp flanking sequence, which affected synthetic promoter activity (Wu et al., 2018). Additionally, stress-inducible synthetic promoters in poplar, including SD9 (20 bp of a basic sequence) and its shorter derivative (7–8 bp) repeats of SD9-1, SD9-2, and SD9-3 showed different responsiveness in individual transgenic poplar events against environmental stresses (Yang et al., 2022b). In the current study, the Syn3 basic sequence (20 bp) was used to construct 6 repeats of Syn3-10b-1 (10 bp) and Syn3-10b-2 (10 bp), of which the former had positive activity in transgenic poplar and the latter did not (Figures 5 and 6). That is, Syn3 was constructed with 6 repeats of a positively active DNA fragment (Syn3-10b-1) spaced by a non-functional fragment (Syn3-10b-2) (Figure 4). Interestingly, Syn3 was highly active in leaves and to a lesser degree in stems and roots under non-stress conditions while Syn3-10b-1 was not (Figures 5 and 7). In addition, the regulatory activity of Syn3 was not stimulated by water-cessation. Meanwhile, Syn3-10b-1 was shown to highly induce visible GUS expression and GUS gene transcription under water-deficit conditions with tissue-specificity in leaf and stem, but not root tissues (Figure 5). Therefore, it seems to be apparent that the spacing between Syn3-10b-1 fragments by insertion of Syn3-10b-2 led to higher constitutive activity, which is congruent to the observation by Wu et al. (2018) with ABRE repeats. However, in terms of water-deficit inducibility, Syn3-10b-1 had a more precise activity than either Syn3 or Syn3-10b-2. Therefore, spacing requirements of DNA motifs or CRE repeats may be dependent on specific applications. Additionally, the distance between DNA motif repeats with a minimal core promoter is also an important factor for reliable synthetic promoter activity. The studies investigating insertions between the core promoter including the TATA box and CRE array showed that a distance over 50 bp between two factors significantly reduced the promoter activity (Cai et al., 2020). Our current synthetic promoters have 4 bp distance between DNA motif repeats and the core promoter, which may be too proximal, and could potentially prevent optimal TF recruitment for transcription. The present results have shown positive synthetic promoter activity, but we may be able to further optimize the activity and specificity via spacing and CRE repeat number as well optimizing proximity between repeats and the core promoter. Furthermore, the present study did not determine if these synthetic promoters can be tuned accurately to different levels of water stress. In order to create optimal designs for synthetic biology applications, it will be necessary to experimentally quantify the exact correlation between current synthetic promoters and their environmental stress cues, including water deficit and salt stresses in future.

The GUS induction driven by Syn3-10b-1 and Syn3 was enriched in leaf and stem tissues compared to roots (Figures 5

and 7). In application, cell walls in the vascular cambium and xylem are important for bioenergy and bioproduct synthesis. As a practical example of synthetic promoter application for sustainable biomass increase, three copies of a secondary wall NAC binding element (SNBE) of 103 bp in length induced cinnamoyl-coenzyme A reductase 1 (CCR1) specifically in vessel tissue by utilizing a 35S core promoter in transgenic *Arabidopsis* (De Meester *et al.*, 2018). The CCR1-specific induction in stem vascular tissue driven by the SNBE synthetic promoter recovered vascular collapse in a *ccr1* mutant, resulting in a 59% increase of stem biomass compared to wild type. These transgenic *Arabidopsis* had a 4-fold increase in total sugar yield, which could be useful for applications in biofuel feedstocks. Additionally, SNBE regulated spatiotemporal expression of a CRISPR construct showed reduced CCR1 function in secondary cell wall structures (Yu *et al.*, 2021). The reduction of CCR1 function in stem tissue resulted in reduction of recalcitrance, which sequentially increased saccharification in transgenic plants. The stem-tissue-specific properties of Syn3-10b-1 and Syn3 in the present study seem similar to the SNBE synthetic promoter. Therefore, application of these synthetic promoters along with a poplar secondary cell wall biosynthesis related genes could allow manipulation of biomass-related metabolic pathways at the transcriptional level.

Acknowledgement

This work was funded from the Biological and Environmental Research in the U.S. Department of Energy Office of Science (DE-SC0018347).

Author contributions

YY designed and performed the experiments, analyzed the data, and prepared the manuscript. YY, TAC, and YS generated gene constructs and transgenic poplars, and maintained the regenerated transgenic plants. They additionally collected whole fluorescence measurement data, GUS staining images, and qPCR analysis. VB, MM, HM, MMR, and AHA generated transcriptome data from laser captured microdissection isolated leaf cell types. AHA, EB, and CNS conceived of the study and its design and coordination and assisted with interpretation of results and versions of the manuscript. All authors read and approved the final manuscript.

Data availability statement

All pertinent data are included in the submission.

References

Ahkami, A., Johnson, S.R., Srividya, N. and Lange, B.M. (2015) Multiple levels of regulation determine monoterpenoid essential oil compositional variation in the mint family. *Mol. Plant*, **8**, 188–191.

Ali, S. and Kim, W.C. (2019) A fruitful decade using synthetic promoters in the improvement of transgenic plants. *Front. Plant Sci.* **10**, 1433.

Andrews, S. (2010) *FastQC: a quality control tool for high throughput sequence data*. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

Bai, J., Wang, X., Wu, H., Ling, F., Zhao, Y., Lin, Y. and Wang, R. (2020) Comprehensive construction strategy of bidirectional green tissue-specific synthetic promoters. *Plant Biotechnol. J.* **18**, 668–678.

Bailey, T.L., Johnson, J., Grant, C.E. and Noble, W.S. (2015) The MEME Suite. *Nucleic Acids Res.* **43**, W39–W49.

Balasubramanian, V.K., Purvine, S.O., Liang, Y., Kelly, R.T., Pasa-Tolic, L., Chrisler, W.B., Blumwald, E. *et al.* (2021) Cell-type-specific proteomics analysis of a small number of plant cells by integrating laser capture microdissection with a nanodroplet sample processing platform. *Curr. Protoc.* **1**(5), e153.

Bartlett, A., O'Malley, R.C., Huang, S.C., Galli, M., Nery, J.R., Gallavotti, A. and Ecker, J.R. (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* **12**, 1659–1672.

Bhunia, R.K., Chakraborty, A., Kaur, R., Gayatri, T., Bhattacharyya, J., Basu, A., Maiti, M.K. *et al.* (2014) Seed-specific increased expression of 2S albumin promoter of sesame qualifies it as a useful genetic tool for fatty acid metabolic engineering and related transgenic intervention in sesame and other oil seed crops. *Plant Mol. Biol.* **86**, 351–365.

Bushnell, B. (2014) *BBMap: A fast, accurate, splice-aware aligner*. Available online at: <https://www.osti.gov/servlets/purl/1241166>

Cai, M., Wei, J., Li, X., Xu, C. and Wang, S. (2007) A rice promoter containing both novel positive and negative cis-elements for regulation of green tissue-specific gene expression in transgenic plants. *Plant Biotechnol. J.* **5**, 664–674.

Cai, Y.M., Kallam, K., Tidd, H., Gendarini, G., Salzman, A. and Patron, N.J. (2020) Rational design of minimal synthetic promoters for plants. *Nucleic Acids Res.* **48**, 11845–11856.

Chen, X., Guo, L., Fan, Z. and Jiang, T. (2008) W-AlignACE: an improved Gibbs sampling algorithm based on more accurate position weight matrices learned from sequence and gene expression/ChIP-chip data. *Bioinformatics*, **24**, 1121–1128.

Cho, J.S., Jeon, H.W., Kim, M.H., Vo, T.K., Kim, J., Park, E.J., Choi, Y.I. *et al.* (2019) Wood forming tissue-specific bicistronic expression of PdGA20ox1 and PtrMYB221 improves both the quality and quantity of woody biomass production in a hybrid poplar. *Plant Biotechnol. J.* **17**, 1048–1057.

Clarke, L. and Kitney, R. (2020) Developing synthetic biology for industrial biotechnology applications. *Biochem. Soc. Trans.* **48**, 113–122.

Dash, M., Yordanov, Y.S., Georgieva, T., Wei, H. and Busov, V. (2018) Gene network analysis of poplar root transcriptome in response to drought stress identifies a PtaJAZ3PtaRAP2.6-centered hierarchical network. *PLoS One*, **13**, e0208560.

De Jaeger, G., Scheffer, S., Jacobs, A., Zambre, M., Zobell, O., Goossens, A., Depicker, A. *et al.* (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. *Nat. Biotechnol.* **20**, 1265–1268.

De Meester, B., de Vries, L., Ozparpucu, M., Gierlinger, N., Corneillie, S., Pallidis, A., Goeminne, G. *et al.* (2018) Vessel-specific reintroduction of CINNAMOYL-COA REDUCTASE1 (CCR1) in dwarfed *ccr1* mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol.* **176**, 611–633.

Dey, N., Sarkar, S., Acharya, S. and Maiti, I.B. (2015) Synthetic promoters in planta. *Planta* **242**, 1077–1094.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D., Patron, N.J. *et al.* (2014) A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* **3**, 839–843.

Fausther-Bovendo, H. and Kobinger, G. (2021) Plant-made vaccines and therapeutics. *Science* **373**, 740–741.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186.

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

Hofgen, R. and Willmitzer, L. (1988) Storage of competent cells for Agrobacterium transformation. *Nucleic Acids Res.* **16**, 9877.

Jameel, A., Norman, M., Liu, W., Ahmad, N., Wang, F., Li, X. and Li, H. (2020) Tinkering Cis Motifs Jigsaw puzzle led to root-specific drought-inducible novel synthetic promoters. *Int. J. Mol. Sci.* **21**, 1357.

Jones, M.O., Manning, K., Andrews, J., Wright, C., Taylor, I.B. and Thompson, A.J. (2008) The promoter from SIREO, a highly-expressed, root-specific *Solanum lycopersicum* gene, directs expression to cortex of mature roots. *Funct. Plant Biol.* **35**, 1224–1233.

- Joshi, J.B., Geetha, S., Singh, B., Kumar, K.K., Kokiladevi, E., Arul, L., Balasubramanian, P. et al. (2015) A maize alpha-zein promoter drives an endosperm-specific expression of transgene in rice. *Physiol. Mol. Biol. Plants* **21**, 35–42.
- Ko, J.H., Kim, H.T., Hwang, I. and Han, K.H. (2012) Tissue-type-specific transcriptome analysis identifies developing xylem-specific promoters in poplar. *Plant Biotechnol. J.* **10**, 587–596.
- Koramutla, M.K., Bhatt, D., Negi, M., Venkatachalam, P., Jain, P.K. and Bhattacharya, R. (2016) Strength, Stability, and cis-Motifs of In silico Identified Phloem-Specific Promoters in Brassica juncea (L.). *Front Plant Sci* **7**, 457.
- Kumar, K., Gambhir, G., Dass, A., Tripathi, A.K., Singh, A., Jha, A.K., Yadava, P. et al. (2020) Genetically modified crops: current status and future prospects. *Planta*, **251**, 91.
- Langmead, B. and Salzberg, S. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359.
- Li, C., Ma, X., Yu, H., Fu, Y. and Luo, K. (2018) Ectopic Expression of PtoMYB74 in Poplar and Arabidopsis Promotes Secondary Cell Wall Formation. *Front. Plant Sci.* **9**, 1262.
- Lihu, A. and Holban, S. (2016) A review of ensemble methods for de novo motif discovery in ChIP-Seq data. *Brief. Bioinform.* **17**, 731.
- Liu, X.S., Brutlag, D.L. and Liu, J.S. (2002) An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nat. Biotechnol.* **20**, 835–839.
- Liu, W., Mazarei, M., Peng, Y., Fethe, M.H., Rudis, M.R., Lin, J., Millwood, R.J. et al. (2014) Computational discovery of soybean promoter cis-regulatory elements for the construction of soybean cyst nematode-inducible synthetic promoters. *Plant Biotechnol. J.* **12**, 1015–1026.
- Liu, W., Mazarei, M., Ye, R., Peng, Y., Shao, Y., Baxter, H.L., Sykes, R.W. et al. (2018) Switchgrass (*Panicum virgatum* L.) promoters for green tissue-specific expression of the MYB4 transcription factor for reduced-recalcitrance transgenic switchgrass. *Biotechnol. Biofuels*, **11**, 122.
- Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods*, **25**, 402–408.
- Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
- Lu, Z., Hofmeister, B.T., Vollmers, C., DuBois, R.M. and Schmitz, R.J. (2017) Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. *Nucleic Acids Res.* **45**, e41.
- Lv, X., Song, X., Rao, G., Pan, X., Guan, L., Jiang, X. and Lu, H. (2009) Construction vascular-specific expression bi-directional promoters in plants. *J. Biotechnol.* **141**, 104–108.
- Mahmood, N., Nasir, S.B. and Hefferon, K. (2020) Plant-Based Drugs and Vaccines for COVID-19. *Vaccines (Basel)* **9**, 15.
- Marand, A.P., Eveland, A.L., Kaufmann, K. and Springer, N.M. (2023) cis-Regulatory Elements in Plant Development, Adaptation, and Evolution. *Annu. Rev. Plant Biol.* **74**, 111–137.
- Mohan, C., Jayanarayanan, A.N. and Narayanan, S. (2017) Construction of a novel synthetic root-specific promoter and its characterization in transgenic tobacco plants. *3 Biotech*, **7**, 234.
- Moreno-Gimenez, E., Selma, S., Calvache, C. and Orzaez, D. (2022) GB_SynP: A Modular dCas9-Regulated Synthetic Promoter Collection for Fine-Tuned Recombinant Gene Expression in Plants. *ACS Synth. Biol.* **11**, 3037–3048.
- Nutiu, R., Friedman, R.C., Luo, S., Khrebtkova, I., Silva, D., Li, R., Zhang, L. et al. (2011) Direct measurement of DNA affinity landscapes on a high-throughput sequencing instrument. *Nat. Biotechnol.* **29**, 659–664.
- Oliva, R., Ji, C., Atienza-Grande, G., Huguet-Tapia, J.C., Perez-Quintero, A., Li, T., Eom, J.S. et al. (2019) Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* **37**, 1344–1350.
- Persad, R., Reuter, D.N., Dice, L.T., Nguyen, M.A., Rigoulot, S.B., Layton, J.S., Schmid, M.J. et al. (2020) The Q-system as a synthetic transcriptional regulator in plants. *Front. Plant Sci.* **11**, 245.
- Persad-Russell, R., Mazarei, M., Schimel, T.M., Howe, L., Schmid, M.J., Kakeshpour, T., Barnes, C.N. et al. (2022) Specific bacterial pathogen phytosensing is enabled by a synthetic promoter-transcription factor system in potato. *Front. Plant Sci.* **13**, 873480.
- Plant Cell Atlas, C., Jha, S.G., Borowsky, A.T., Cole, B.J., Fahlgren, N., Farmer, A., Huang, S.C. et al. (2021) Vision, challenges and opportunities for a Plant Cell Atlas. *Elife*, **10**, e66877.
- Putri, G.H., Anders, S., Pyl, P.T., Pimanda, J.E. and Zanini, F. (2022) Analyzing high-throughput sequencing data in python with HTSeq 2.0. *Bioinformatics*, **38**, 2943–2945.
- Schmitz, R.J., Grotewold, E. and Stam, M. (2022) Cis-regulatory sequences in plants: Their importance, discovery, and future challenges. *Plant Cell*, **34**, 718–741.
- Sears, R.G., Rigoulot, S.B., Occhialini, A., Morgan, B., Kakeshpour, T., Brabazon, H., Barnes, C.N. et al. (2023) Engineered gamma radiation phytosensors for environmental monitoring. *Plant Biotechnol. J.* **21**, 1745–1756.
- Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., Arnaiz, O. et al. (2015) The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res.* **43**, W589–W598.
- Song, J., Lu, S., Chen, Z.Z., Lourenco, R. and Chiang, V.L. (2006) Genetic transformation of Populus trichocarpa genotype Nisqually-1: a functional genomic tool for woody plants. *Plant Cell Physiol.* **47**, 1582–1589.
- Sparkes, I.A., Runions, J., Kearns, A. and Hawes, C. (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* **1**, 2019–2025.
- Sultana, M.S., Mazarei, M., Millwood, R.J., Liu, W., Hewezi, T. and Stewart, C.N., Jr. (2022) Functional analysis of soybean cyst nematode-inducible synthetic promoters and their regulation by biotic and abiotic stimuli in transgenic soybean (*Glycine max*). *Front. Plant Sci.* **13**, 988048.
- Thibivilliers, S. and Libault, M. (2021) Plant single-cell multiomics: Cracking the molecular profiles of plant cells. *Trends Plant Sci.* **26**, 662–663.
- Vining, K.J., Johnson, S.R., Ahkami, A., Lange, I., Parrish, A.N., Trapp, S.C., Croteau, R.B. et al. (2017) Draft genome sequence of mentha longifolia and development of resources for mint cultivar improvement. *Mol. Plant* **10**, 323–339.
- Wang, R., Zhu, M., Ye, R., Liu, Z., Zhou, F., Chen, H. and Lin, Y. (2015) Novel green tissue-specific synthetic promoters and cis-regulatory elements in rice. *Sci. Rep.* **5**, 18256.
- Werner, S., Engler, C., Weber, E., Gruetzner, R. and Marillonnet, S. (2012) Fast track assembly of multigene constructs using golden gate cloning and the MoClo system. *Bioeng Bugs* **3**, 38–43.
- Wu, C., Washida, H., Onodera, Y., Harada, K. and 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. *Plant J.* **23**, 415–421.
- Wu, R., Duan, L., Prunedo-Paz, J.L., Oh, D.H., Pound, M., Kay, S. and Dinneny, J.R. (2018) The 6xABRE synthetic promoter enables the spatiotemporal analysis of ABA-mediated transcriptional regulation. *Plant Physiol.* **177**, 1650–1665.
- Xu, Y., Buchholz, W.G., DeRose, R.T. and Hall, T.C. (1995) Characterization of a rice gene family encoding root-specific proteins. *Plant Mol. Biol.* **27**, 237–248.
- Xu, W., Liu, W., Ye, R., Mazarei, M., Huang, D., Zhang, X. and Stewart, C.N., Jr. (2018) A profilin gene promoter from switchgrass (*Panicum virgatum* L.) directs strong and specific transgene expression to vascular bundles in rice. *Plant Cell Rep.* **37**, 587–597.
- Yang, J.S. and Reyna-Llorens, I. (2023) Plant synthetic biology: exploring the frontiers of sustainable agriculture and fundamental plant biology. *J. Exp. Bot.* **74**, 3787–3790.
- Yang, Y., Lee, J.H., Poindexter, M.R., Shao, Y., Liu, W., Lenaghan, S.C., Ahkami, A.H. et al. (2021) Rational design and testing of abiotic stress-inducible synthetic promoters from poplar cis-regulatory elements. *Plant Biotechnol. J.* **19**, 1354–1369.
- Yang, Y., Chaffin, T.A., Ahkami, A.H., Blumwald, E. and Stewart, C.N., Jr. (2022a) Plant synthetic biology innovations for biofuels and bioproducts. *Trends Biotechnol.* **40**, 1454–1468.
- Yang, Y., Shao, Y., Chaffin, T.A., Lee, J.H., Poindexter, M.R., Ahkami, A.H., Blumwald, E. et al. (2022b) Performance of abiotic stress-inducible synthetic promoters in genetically engineered hybrid poplar (*Populus tremula* × *Populus alba*). *Front. Plant Sci.* **13**, 1011939.
- Yang, Y., Shao, Y., Chaffin, T.A., Ahkami, A.H., Blumwald, E. and Stewart, C.N., Jr. (2023) Synthetic promoter screening using poplar mesophyll protoplast transformation. *Bio. Protoc.* **13**, e4660.

- Yasmeen, E., Wang, J., Riaz, M., Zhang, L. and Zuo, K. (2023) Designing artificial synthetic promoters for accurate, smart, and versatile gene expression in plants. *Plant. Commun.* **4**, 100558.
- Yu, H., Liu, C. and Dixon, R.A. (2021) A gene-editing/complementation strategy for tissue-specific lignin reduction while preserving biomass yield. *Biotechnol. Biofuels* **14**, 175.
- Zhang, L., Yang, T., Li, X., Hao, H., Xu, S., Cheng, W., Sun, Y. *et al.* (2014) Cloning and characterization of a novel Athspr promoter specifically active in vascular tissue. *Plant Physiol. Biochem.* **78**, 88–96.

Supporting information

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

Figure S1 Laser Capture Microdissection (LCM)-based cell isolation from hybrid poplar leaf tissue. The left panels show leaf cryosections before palisade and vascular cell isolation, and the right panels show the cryosections after cell isolation. The scale bar in palisade panels represents 150 μm , and the scale bar in vascular panels represents 300 μm .

Figure S2 Representative images of stressed plants. Wild-type and transgenic poplar of 3 months after root regeneration were used for all stress treatments and mock controls. (a) Representative image of a wilted leaf wild type plant after water-cessation (~5 days in growth chamber condition). (b) Representative image of mock and wilted transgenic plants after water cessation. Syn10b-1#2 is shown in this panel. We collected leaf, stem, and root based on wilted leaf appearance after water cessation for X-gluc histochemical staining. (c) The phenotype of transgenic poplars containing Syn3, Syn3-10b-1, and Syn3-10b-2 in non-stressed conditions. The bar displays 10 cm.

Figure S3 Genotype of transgenic poplars containing Syn3-10b-1, Syn3-10b-2, Syn3, and vector without a synthetic promoter.

GUS- and LUC-specific primers were used to detect transgene construct transformation in the transgenic plant genome. PtaUBCC was used as an internal control to monitor genomic DNA quality. Asterisks indicate lines selected for X-gluc histochemical staining and GUS gene expression analysis.

Table S1 Oligonucleotide list used in synthetic promoter generation.

Table S2 Nucleotide sequences of accessory components including 2 \times 35S short promoter, minimal 35S core promoter, and TMV 5' Ω – leader sequence in synthetic promoter gene constructs.

Table S3 Primer sequences used for qPCR analysis in this study.

Table S4 Selected upregulated genes in 717-1B4 leaf palisade and vascular tissue-specific expressed genes, collected using a single-cell laser capture following water-deficit stress treatment. Two kilobase (kb) upstream sequences from the ATG initial codon of co-expressed genes in leaf palisade (98 genes) or vascular tissues (16 genes) were collected from BioMart and submitted separately to predict well-conserved DNA motifs.

Data S1 Differentially expressed genes in leaf palisade cells from water stress-treated plants vs controls. Significantly Upregulated ($\text{Log}_2 \text{FC} > 4$) and downregulated ($\text{Log}_2 \text{FC} < -4$) genes were highlighted with green and red colour, respectively (adjusted P ($P_{\text{adj}} < 0.05$, shown in yellow) in each treatment. P_{adj} value was determined by DESeq2 analysis. EWD: Early water deficit; LWD: Late water deficit; REC: Recovery.

Data S2 Differentially expressed genes in leaf vascular cells from water stress-treated plants vs controls. Significantly upregulated ($\text{Log}_2 \text{FC} > 4$) and downregulated ($\text{Log}_2 \text{FC} < -4$) genes were highlighted with green and red colour, respectively (adjusted P ($P_{\text{adj}} < 0.05$, shown in yellow) in each treatment. P_{adj} value was determined by DESeq2 analysis. EWD: Early water deficit; LWD: Late water deficit; REC: Recovery.