# Rational Design of Minimal Synthetic Promoters for Plants

Yaomin Cai[1], Kalyani Kallam[1], Henry Tidd[1], Giovanni Gendarini[1], Amanda Salzman[1] and Nicola J. Patron[1†]

[1]Engineering Biology, Earlham Institute, Norwich Research Park, Norfolk, NR4 7UZ, UK [†]Corresponding author nicola.patron@earlham.ac.uk

ABSTRACT Promoters serve a critical role in establishing baseline transcriptional capacity through the recruitment of proteins, including transcription factors (TFs). Previously, a paucity of data for *cis*-regulatory elements in plants meant that it was challenging to determine which sequence elements in plant promoter sequences contributed to transcriptional function. In this study, we have identified functional elements in the promoters of plant genes and plant pathogens that utilise plant transcriptional machinery for gene expression. We have established a quantitative experimental system to investigate transcriptional function, investigating how identity, density and position contribute to regulatory function. We then identified permissive architectures for minimal synthetic plant promoters enabling computational design of a suite of synthetic promoters of different strengths. These have been used to regulate the relative expression of output genes in simple genetic devices.
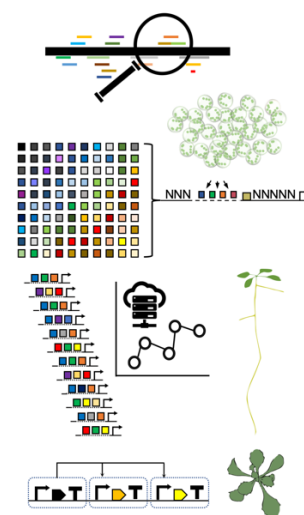
Key words: gene regulation, promoters, transcription factors, cis-regulatory elements, plant biotechnology, synthetic biology

INTRODUCTION

Transgenic techniques are used to investigate the function of plant genes and to develop new products for agriculture and industry. Biotech crops, typically containing at least one transgene, are now planted on over 190 million hectares each year (1), and plants are finding new roles as platforms for biomanufacturing pharmaceuticals (2, 3). For many years, the majority of transgenic events involved only a single gene of interest and a selectable marker gene. However, recent advances in DNA assembly techniques pioneered by the nascent field of synthetic biology have enabled the facile construction of multigene constructs for plants (4). Researchers are now able to apply these tools to design and deploy synthetic genetic circuits and reconstruct heterologous biochemical pathways in plant systems (5, 6). However, the realisation of synthetic genetic circuits that function as expected requires the ability to precisely and predictably regulate gene expression. Spatiotemporal quantities of endogenous gene products are regulated through numerous mechanisms including transcript elongation (7), antisense transcription (8), and several post-transcriptional and translational processes (9, 10). Although these mechanisms could be leveraged to fine-tune the expression of transgenes, information flow from synthetic genetic circuits is initiated by transcription and, therefore, control of transcription is considered the simplest way to balance the expression of transgenes within a synthetic genetic

circuit (11). To achieve this, regulatory elements with predicable characteristics are highly desirable. However, suites of promoters with different levels of expression for plants are not widely available and the promoters that are used are often several kilobases in length and their functional elements have only rarely been identified and characterised. Many plant scientists and biotechnologists still rely on a small set of natural promoters first isolated in the 1980s. In particular, constitutive promoters from plant-infecting DNA viruses or the opine biosynthetic genes found on tumour-inducing (Ti) plasmids of *Agrobacterium tumefaciens* that recruit the host cell's transcriptional machinery (12). These include the 35s promoter from the double-stranded DNA virus, Cauliflower Mosaic Virus (CaMV), which is reported to have at least partial function in numerous plant species as well as in bacteria (13), fungi (14, 15), and vertebrates (16, 17). Eukaryotic promoter sequences are comprised of complex arrangements of motifs and elements. Early deletion and rearrangement studies identified several key functional elements in the promoters of plant-infecting viruses and bacteria, revealing synergistic interactions between *cis*-elements (18–20). Engineered variants were made by swapping domains to achieve promoters of similar strengths with reduced sequence homologies (21). However, progress towards rational-design of synthetic promoters was limited by a lack of comprehensive data for plant transcription factor binding sites (TFBSs) as well as by the technical limitations of rearranging and rewriting

DNA sequences before chemical gene synthesis was widely available. Consequently, although some functional elements were characterised and the possibility of designing synthetic plant promoters was discussed (22), these goals have yet to be fully realised. Progress has been made by engineering variations of natural promoters for tissue-specific expression (23). In addition, a number of synthetic regulatory elements comprised of binding sites for orthogonal TFs fused to a minimal core promoter have been used to enable inducible constitutive expression (24–26). However, to initiate the flow of transcription, these require a non-orthogonal promoter to drive expression of the orthogonal TF, for which natural are typically used (27).

In recent years, significant progress has been made in the design of synthetic regulatory elements for microorganisms, initially with the rational design of ribosome binding sites (28), transcription factors (29) and enhancers (30) and, subsequently, promoters (31–34). Such studies were substantially enabled by comprehensive datasets of TFBSs as well as by the ability to deliver sizable and complex libraries of sequences to populations of cells, sorting, selecting and sequencing cells with desired expression profiles. Equivalent experiments are challenging in plants due to the limitations of DNA-delivery technologies and a paucity of cell-lines. However, genome sequencing technologies have recently shed light on epigenetic states and chromatin accessibility in plant genomes (35) and have identified candidate binding-motifs for many plant transcription factors (36–38). However, genomic datasets alone cannot be used to predict the intrinsic regulatory functions of DNA sequences and assessing the contribution of sequence motifs to regulatory activity is considered essential for characterising function (39). Genome engineering technologies are enabling the functions of specific *cis*-regulatory elements to be dissected (40). Rationally engineered suites of synthetic plant promoters of different strengths have yet to be reported. Here we describe a series of investigations to identify and functionally characterise plant *cis*-regulatory elements, revealing how complexity and relative positions contribute to regulatory functions. We use these data to predict the performance of computationally-designed minimal synthetic constitutive promoters and demonstrate predictable behaviour in dicotyledonous plants in transient expression and when integrated as stable transgenes. Thus, we present suites of minimal synthetic plant promoters of varied strengths, activated by either endogenous or orthogonal transcription factors and demonstrate how these can be used to control the

relative expression of output genes in simple genetic circuits.

MATERIAL AND METHODS

Identification of candidate transcription factor binding sites (TFBSs)

The Position Weight Matrices (PWMs) from the *Arabidopsis thaliana* (Arabidopsis) cistrome dataset (http://neomorph.salk.edu/dev/pages/shhuang/dap_web/pages/browse_table_aj.php) dataset (36) and the Plant Transcription Factor Database (41) were used to create a motif file for the command line version of FIMO (MEME suite) (42). This was used to scan FASTA files of promoter sequences with a threshold p-value of 0.0001. Candidate TFBSs were mapped back to the promoter sequences. Expression data for TF-encoding genes across multiple Arabidopsis tissues was obtained from the Expression Atlas (http://www.ebi.ac.uk/gxa) (43).

Construction of plasmids

All constructs were designed in Benchling (San Francisco, CA, USA), synthesised as double-stranded DNA fragments (Twist Biosciences, San Francisco, CA, USA) and cloned into a universal acceptor plasmid (pUAP1 (44) or pUPD2 (45)), to produce standardised Level 0 Phytobricks, conforming to the plant common syntax standard (44). Expression cassettes and multigene constructs were assembled using the Type IIS DNA assembly protocol described in (46). Synthetic and control promoter parts were assembled with the omega 5' untranslated region from tobacco mosaic virus (5UTR-ΩTMV; pICH41402, Addgene #50285), the coding sequence of firefly luciferase (LucF; pEPAS0CM0008, Addgene T.B.C.), a C-terminal FLAG tag (pICSL50007, Addgene #50308), and a 3' untranslated region and terminator sequence (3UTR) from *Agrobacterium tumefaciens* octopine synthase (*AtuOCS*) (pICH41432, Addgene #50343). A calibrator construct (pEPYC1CB0197, Addgene T.B.C.) for ratiometric quantification was assembled from *A. tumefaciens* nopaline synthase (*AtuNOS*) promoter (pICH42211, Addgene #50255), 5UTR-ΩTMV, the coding sequence of NanoLuc luciferase (LucN, pEPYC0CM0133, Addgene T.B.C.) and *AtuOCS* terminator. For stable plant transformation, synthetic and control promoters were assembled with the 5'UTR from cowpea mosaic virus (CPMV), a chimeric coding sequence consisting of an N'-terminal HiBit (pEPYC0CM0258, Addgene T.B.C.) the *uidA* coding sequence, encoding β-glucuronidase (GUS; pICSL80016, Addgene #50332) and a C'-terminal yellow fluorescent protein (YFP; pICSL50005, Addgene #117536), and *AtuOCS* terminator. This reporter cassette was assembled with

a plant selectable marker cassette conferring resistance to kanamycin (pEPYC1CB0308, Addgene T.B.C.). Synthetic and control promoters were additionally fused to either a transcription activator like effector (TALE) or a synthetic transcription factor comprised of a Gal4 activation domain (GB0900, received from the Orzaez laboratory) and a PhiC3 binding domain (GB_UD_32AB, received from the Orzaez laboratory). A table of all synthetic promoters used in this study is provided in Supplementary Data 1 and tables with the details of all plasmids used and constructed for this study are provided in Supplementary Data 2. All plasmids and their complete sequences have been submitted to the Addgene repository.

Growth of plant material
Arabidopsis thaliana (Col-0), *Nicotiana benthamiana* and *Brassica rapa* plants were germinated and grown in potting medium (two-parts sieved compost to one-part sand) within controlled environment chambers with a 16 h photoperiod at 22 °C with 120–180 µmol/m^2/s^2 light intensity. For the two days before leaves were harvested for the preparation of protoplasts, the photoperiod was reduced to 8 h.

Protoplast preparation and transfection
Protoplasts were prepared from the leaf tissues of *A. thaliana*, *N. benthamiana* and *B. rapa* as previously described (47) and diluted to 10^4 -10^5/mL for transfection. A total of 4.5 µg purified plasmid DNA, comprising equal molar ratios of the plasmid containing the expression cassette for which expression was measured (*test-p*:ΩTMV:*LucF:AtuNOSt*) and a calibrating plasmid (pEPYC1CB0197; *AtuNOSp*:ΩTMV:*LucN:AtuNOSt*) were added to each designated well of a 2.2 mL 96 deep-well plate containing 200 µL protoplasts (10^4 - 10^5 /mL) and mixed gently by shaking. PEG solution was freshly prepared by mixing 2 g PEG (poly(ethylene glycol), MW 4,000 Da) with 2 mL 500 mM mannitol and 0.5 mL 1M CaCl$_2$ and 220 µL was added to each well. After 5 mins at room temperature, 1.2 mL W5 (154 mM NaCl, 125 mM CaCl$_2$, 5 mM KCl, 2 mM MES pH5.6) was added and protoplasts were collected by centrifugation at 100 g for 2 min and resuspended in 100 µL W5 solution. Resuspended protoplasts were transferred to a round bottom 96-well plate (pre-prepared by incubation with 0.1% bovine serum albumin for 10 mins). Transfected protoplasts were incubated at 22 °C with 100 µmol/m^2/s^2 light intensity for at least 16 hrs. For each batch of protoplasts, a control plasmid, pEPYC1CB0199 (*AtuMASp*: ΩTMV:*LucF:AtuNOSt)* and the calibrator

(pEPYC1CB0197; *AtuNOSp*: ΩTMV:*LucN:AtuNOSt*) were used to transfect three aliquots of protoplasts.

Production of stable transformants
Transgenic Arabidopsis lines were produced by agrobacterium-mediated transformation of floral tissues. Assembled plasmids were transformed into *A. tumefaciens* (GV3101) and liquid cultures were grown from single colonies in growth medium supplemented with 50 µg/mL rifampicin, 25 µg/mL gentamycin and 50 µg/mL kanamycin at 28 °C. *A. tumefaciens* cells were collected by centrifugation and resuspended to OD$_{600}$ 0.8 in 5% sucrose, 0.05% Silvet L-77 and sprayed onto Arabidopsis floral tissues. Plants were sealed in black plastic bags for 24 hours. Seeds were collected from mature siliques and surface sterilised with 70% EtOH for 10 min followed by 3-5% sodium hypochlorite for 10 min. For selection of transgenics, sterilised seeds were germinated and grown on Murashige and Skoog medium supplemented with 75 µg/mL Kanamycin with 16 h light 22 °C.

### Determination of transgene copy number by digital droplet PCR (ddPCR)
Samples of leaf tissue (0.1 g) were ground in liquid nitrogen. DNA was extracted using the cetyltrimethylammonium bromide (CTAB) extraction protocol described in (48) and 2 µg genomic DNA was digested with 20 Units EcoRV for 2 h at 37 °C. 400ng of digested genomic DNA was used in digital droplet PCR (ddPCR) reactions with QX200™ddPCR™EvaGreen®Supermix (Bio-Rad, Hercules, California, USA) and oligonucleotide primers to the *UidA* transgene sequence (5'-CGGCGAAATTCCATACCTGTT and 5'-TCAGCCGATTATCATCACCGA) or a homozygous single-copy reference gene, *AtADH1* (AT1G77120; 5'-ACTTCTCTCTGTCACACCGA and 5'-GGCCGAAGATACGTGGAAAC). Droplets were generated using the QX200™Droplet Generator (Bio-Rad), PCR reactions were run on the C1000 Touch™Thermal Cycler (Bio-Rad) and analysed on the QX200™Droplet Reader (Bio-Rad). Absolute transgene copy number was calculated using the QuantaSoft™software (Bio-rad) to analyse the ratio of droplets in which the target (UidA) was amplified to those in which the reference (*AtADH1)* was amplified.

Quantification of gene expression
Luciferase expression was detected using the Nano-Glo® Dual-Luciferase® reporter assay system (Promega, Madison, WI, USA). Protoplasts were homogenised in 30 µL passive lysis buffer (Promega) containing protease inhibitor (P9599, Sigma-Aldrich, Dorset, UK). Following incubation on ice for 15 min

and centrifugation (100 x g, 2 min, 4 °C), 30 μL supernatant was removed and mixed with 30 μL ONE-Glo™ EX Luciferase Assay Reagent (Promega) and incubated at room temperature for 10 min. LucF luminescence was detected using a Clariostar microplate reader (BMG Labtech, Aylesbury, UK) with a 10 s read time and 1 s settling time. Gain was set at 3,600. LucN luminescence was detected from the same sample by adding 30 μL NanoDLR™ Stop & Glo® Reagent (Promega). After incubation for 10 min at room temperature, luminescence was detected as above. Normalised expression is reported throughout this manuscript as the ratio of luminescence from the test construct (LucF) to the calibrator (LucN; pEPYC1CB0197), normalised to the luminescence of the experiment control (LucF; pEPYC1CB0199/ LucN; pEPYC1CB0197).

Expression from stably-integrated HiBit:GUS:YFP transgenes was quantified using the Nano-Glo® HiBiT Extracellular Detection System (Promega). 10 mg leaf tissue was homogenised in 50 μL passive lysis buffer (Promega) containing protease inhibitor (P9599, Sigma-Aldrich). Homogenised leaf tissues were

centrifuged at 18,000 g 10 min 4 °C and 2 μL supernatant mixed with 48 μL Bradford reagent (ThermoFisher Scientific, Waltham, Massachusetts, USA). Protein concentration was estimated by absorbance at 595 nm and concentrations were normalised. 5 μL normalised extract were diluted to 30 μL in passive lysis buffer and mixed with 30 μL Nano-Glo® HiBiT Extracellular Detection Reagent (Promega) and luminescence was detected as above. GUS expression was visualised by submerging 10-day-old seedlings in 0.5 mM $K_3Fe(CN)_6$, 0.5 mM $K_4Fe(CN)_6$, 100 mM sodium phosphate buffer pH7.0, 10 mM EDTA 0.5 mg/mL X-Gluc) for 24 h at room temperature. To remove chlorophyll, this was replaced with 70% EtOH followed by 100% EtOH for 8 h each. Images were taken using a Leica M205FA stereo microscope (Leica, Wetzlar, Germany). YFP expression was visualised using a SP5 (II) confocal microscope (Leica) with a 20x air objective, excitation 514 nm, emission 530 nm. Final images were prepared using Fiji ImageJ (49) (https://imagej.net/Fiji).
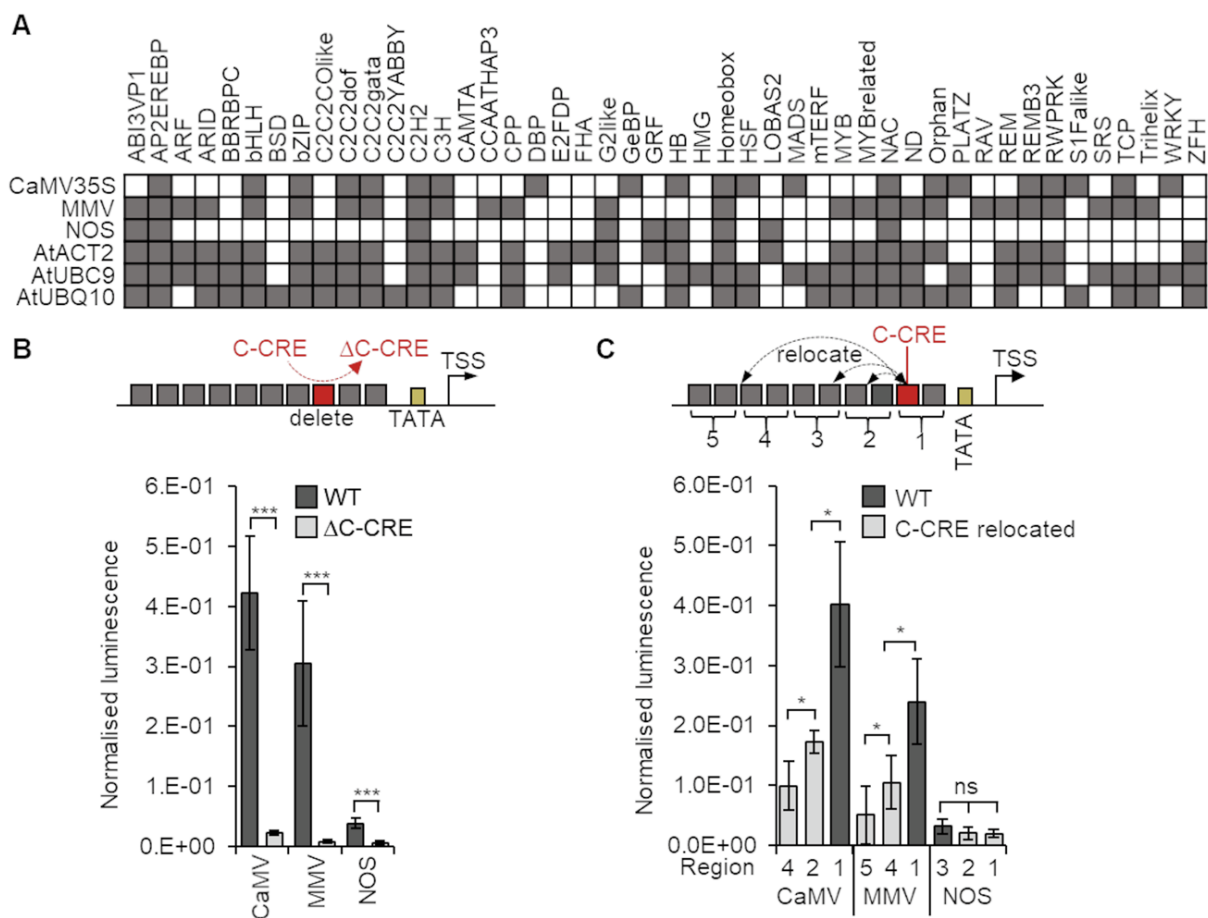
Figure 1. Identification and characterisation of plant *cis*-regulatory elements (CREs). (A) Identification of candidate CREs in constitutive promoters (B) Deletion, or, (C) Relocation of a CRE common to all pathogen promoters (C-CRE) significantly reduces expression. Error bars = 2 x standard error; n=3; P-values were calculated using unpaired two-tailed Student's t-test; *P < 0.05, ***P < 0.001, ns = not significant.

RESULTS

Constitutive promoters are comprised of multiple functional elements with the potential to bind numerous transcription factors

To identify candidate *cis*-regulatory elements (CREs) for use in minimal synthetic constitutive promoters, we analysed promoters widely used for exogenous expression for the presence of candidate TFBSs. These included promoters from vascular plants as well as those from plant-infecting pathogens that recruit the plant's transcriptional machinery, including CaMV35S, *Agrobacterium tumefaciens* nopaline synthase (*AtuNOS*) and Mirabilis Mosaic Virus (MMV). The data indicated that constitutive promoters have, in principle, the ability to bind multiple classes of TFs (Figure 1A). Analysis of Arabidopsis gene expression data indicated that few of the TFs predicted to bind to constitutive promoters show constitutive expression themselves (Supplementary Data 3). We also performed *de novo* motif identification using MEME. This analysis identified the presence of a *cis*-regulatory element common to all 14 pathogen promoters (common-CRE or C-CRE) (Supplementary Data 4). In six of these promoters, the C-CRE contained a predicted binding site for a basic-leucine-zipper (bZIP) TF. These C-CREs can therefore be considered to be equivalent to the previously described Activation Sequence 1 (As-1), shown to directly bind members of the TGACG-motif binding (TGA) family of basic-leucine-zipper (bZIP) TFs (50–52). The other eight pathogen promoters were not predicted to bind bZIP TFs. However, it was previously shown that this region of *AtuNOS* is able to bind TGA4 in the presence of a cofactor, OBP5 (53). Consistent with early studies in which regions of promoters were sequentially deleted, specifically deleting individual C-CREs significantly reduced transcriptional output (Figure 1B). This was in contrast to the majority of candidate CREs, of which deletion did not significantly change expression (Supplementary Data 5). To investigate whether the position of the C-CREs within the promoter was essential, the element was relocated varying its proximity to the TSS. While expression reduced when the C-CREs in CaMV35S and MMV were moved further from the TSS, relocating the C-CRE in *AtuNOS,* which is already located distally to the TSS, had a negligible impact (Figure 1C).

Orthogonal tools with a range of expression levels

To identify a functional basic design for minimal synthetic plant promoters (MinSyn-P), we first built and tested synthetic promoters with a range of expression levels that respond to orthogonal TFs. The initial design was based on previously reported

synthetic Transcriptional Activator Like Effector (TALE)-responsive synthetic elements to which single binding sites for TALES were added (26). These synthetic promoters consist of 19 bps of random sequence, followed by a second region of random sequence of flexible length (to which CREs are added), a TATA box sequence (TATATAA) and a 43 bp minimal core including transcriptional start site (TSS) (Figure 2A). We successfully verified that this general design could be used to build promoters with a range of expression levels by adding different numbers of binding sites for either TALES or recently described synthetic Gal4:ΦC31 TFs (Vazquez-Vilar *et al.,* 2017) (Figure 2B).
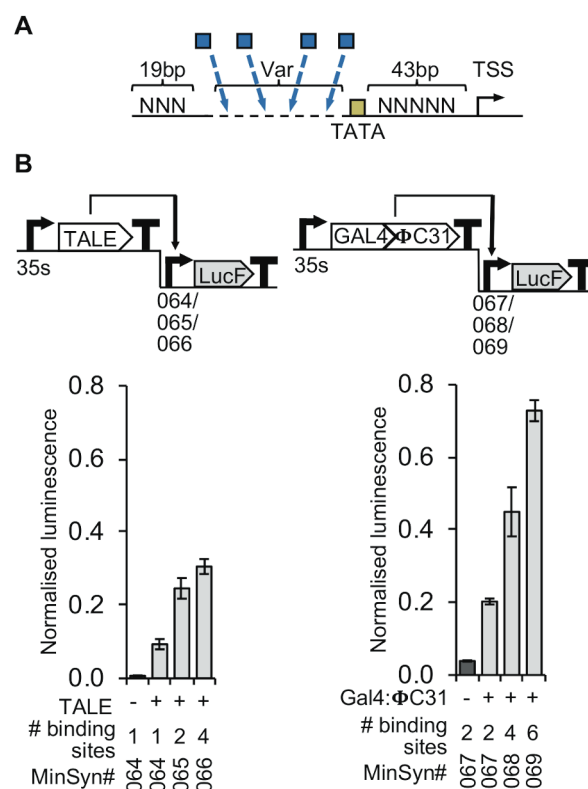
Figure 2. Minimal synthetic promoters (MinSyns) of different strengths regulated by orthogonal transcription factors. (A) MinSyns regulated by Transcriptional Activation Like Effectors (TALES). (B) MinSyns regulated by GAL4:ΦC31 transcription factors; n=3.

Expression from minimal synthetic regulatory elements by passive cooperativity

To define rules for the design of constitutive MinSyns that respond to endogenous TFs, experiments were progressed to test the function of candidate CREs identified from constitutive promoters (Figure 1). To do this, we first inserted three copies of the same CRE into the variable region of the MinSyn. While it was expected that some candidate CREs might be false-positives, would recruit repressors of transcription
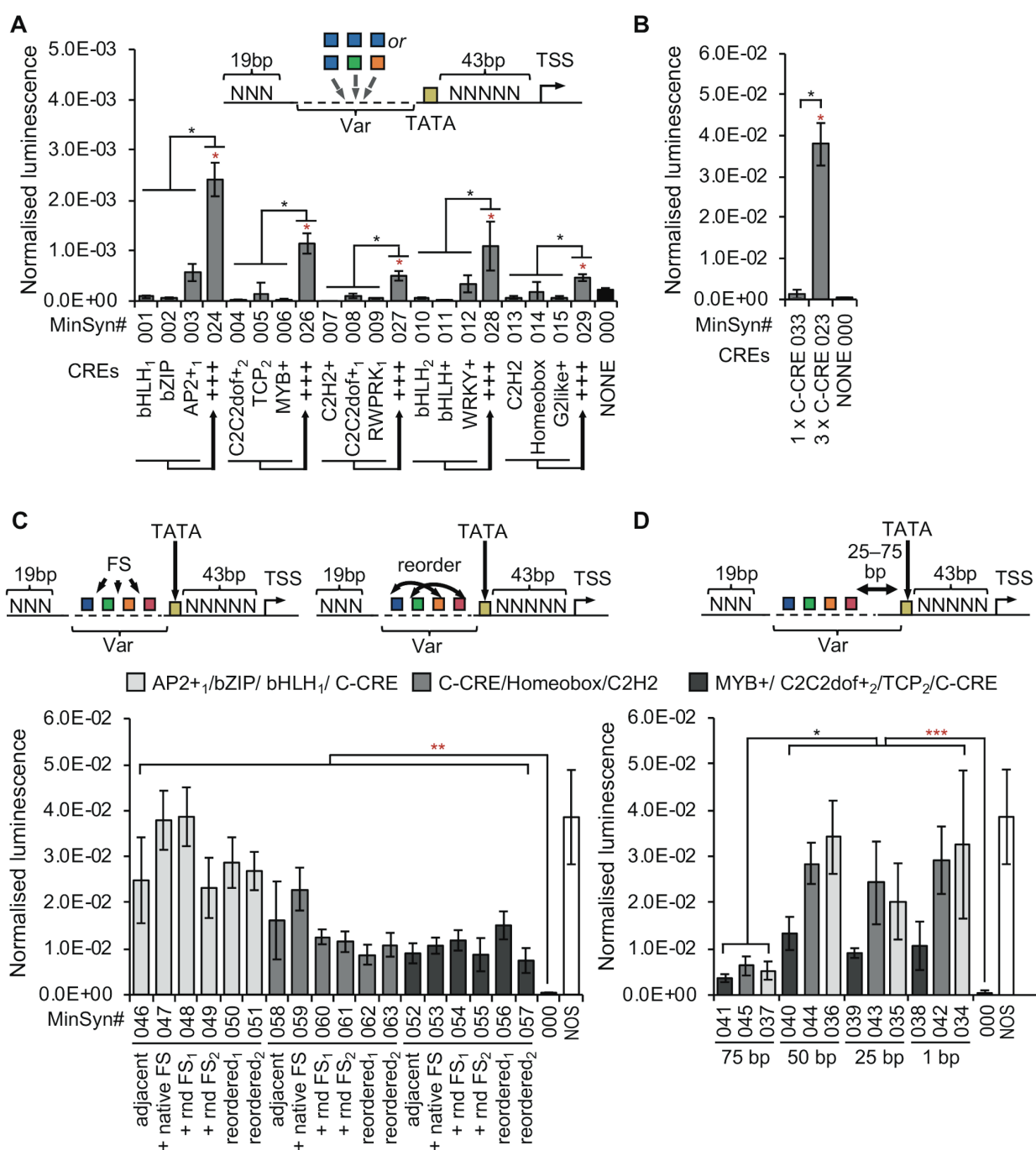
Figure 3. Defining design features of minimal synthetic promoters. (A) Combinations of multiple *cis*-regulatory elements (CREs) enables significant expression compared to repeating copies of the same CRE. (B) The C-CREs enables expression in the absence of other CREs. (C) Rearranging the relative positions of CREs by either inserting native or random flanking sequence (FS) or by reordering does not significantly change expression levels. (D) Relocating CREs greater than 50 base pairs (bp) from the TATA box significantly reduces expression. Error bars = 2 x standard error; n=3; P-values were calculated using unpaired two-tailed Student's t-test; *P<0.05, **P<0.01, ***P<0.001; Red asterisks indicate a significant difference from MinSyn000, which has no CREs. Black asterisks indicate a significant difference from MinSyns indicated by solid black brackets.

(and therefore expression would not be observed), or might need a specific local sequence context, no expression was observed from any MinSyns with only one type of CRE, with the exception of MinSyns containing only multiple copies of C-CREs (Figure 3B). To further investigate the other CREs, we combined random combinations into the variable regions on MinSyns. In the majority of cases, this resulted in significant expression (Figure 3A). In a few cases, combinations of CREs did not result in significant expression (Supplementary Data 6). This is consistent with expectations that some CREs will

mainly recruit transcriptional repressors, and that other CREs need to be correctly co-located to enable TFs to form functional heterocomplexes. To test if expression from MinSyns with multiple CREs was dependent on specific TF-TF interactions, the relative positions and spacing of CREs within the variable region of the MinSyns were altered (Figure 3B). In one set of variants, we added up to 20 bp of additional sequence between the CREs. To control for the effect of local sequence context, we made three variants for each set of CREs, two with random sequence and one with the native flanking sequence from the natural promoter from which the CRE was identified. In a second set of variants, the relative positions of the CREs were permutated. Neither changes to the relative position nor moderate increases in spacing had any significant effect on expression. To determine if the relative location of the CREs to the TATA box was critical and to assess if the minimal length of the MinSyns was limiting function, random sequence was inserted between the variable regions containing the CREs and the TATA box. In chromosomal DNA, DNA looping allows distal enhancer elements to interact with the proximal regions, however, as our design goal was minimal constitutive promoters, further extensions to accommodate such interactions were undesirable. Expression was significantly impacted when more than 50 bps of sequence was inserted between the first CRE and the TATA box (Figure 3C).

Computational design of Minimal Synthetic promoters (MinSyn) with predictable strengths

We applied the knowledge gained from these experiments to develop a script to create a library of 1000 constitutive MinSyn (Supplementary Data 7) for which was predicted. For each MinSyn, the script selects a random number (N) between three and ten that defines the number of CREs in the variable region and creates a random DNA sequence of 5 to 30 bases to comprise the sequence of the variable region. It then selects a single CRE sequences from the pool of previously identified CREs. This pool includes two C-CREs predicted to directly bind TGA TFs and one C-CRE for which direct TGA-binding was not predicted. The first CRE is added to the random DNA sequence and the process repeated N times without replacement. Thus, each MinSyn contains between three and ten different CREs, each added to the variable region in the randomly selected order. From our initial experiments, we observed that the strength of the promoters was most affected by the inclusion of multiple CREs of which the C-CREs had the most significant impact. C-CREs predicted to directly bind TGA TFs have the strongest effect when proximal to the TATA box. In contrast, the relative position is less

important for C-CREs not predicted to directly bind TGA TFs, with strength only decreasing when located more than 150 bps from the TATA box. These observations were used to assign scores to each nucleotide base as follows: bases within CREs were each assigned a specific score and bases within or flanking C-CREs assigned a higher score adjusted by a numerator reflecting the proximity of that base to the TATA box, with different numerators used for TGA-binding and non-TGA binding C-CREs. Finally, scores were summed and the total was divided by the total number of bases. To convert the score into a predicted strength, we applied the prediction to the existing set of tested MinSyns for which strength had been experimentally determined, thus defining a numerator.

A

B
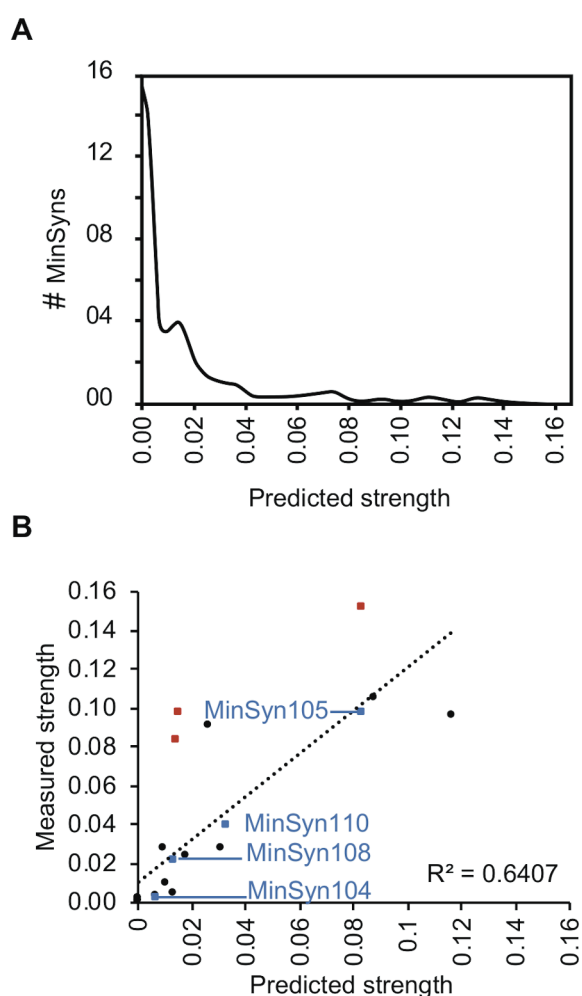
Figure 4. **Computational design of Minimal Synthetic Promoters (MinSyns).** (A) Of a population of 1000 MinSyns, the majority were predicted to have relatively weak expression (B) Comparison of predicted and measured strengths of 24 computational-designed MinSyns. Red squares indicate MinSyns with unintended *cis*-regulatory elements formed at sequence junction that may explain deviance from predicted strength. Blue squares indicate MinSyns selected for further characterisation (Figure 5).

We were therefore able to formulate a predicted expression level for each promoter in the library. As expected for the profile of CREs in the pool, the majority of computational-designed promoters were predicted to have relatively weak expression. Twenty-four MinSyn sequences were selected from the library for synthesis and testing and the predicted and actual levels of expression were compared (Figure 4). For the whole population of 24, the predicted and actual values showed reasonable correlation (Figure 4, red dashed line, R^2=0.6407), however, there were some outliers. We reanalysed the sequences for the presence of known TFBSs that were not present in the pool used to create the library of MinSyns (e.g. those created unintentionally at sequence junctions of CREs). In 17 cases, additional known TFBS were identified but in most cases, there was insufficient data to determine how the TFs predicted to bind might affect expression (if they were activators or repressors). In three cases, the new motifs were predicted to bind additional TGA, NAC or cytokinin-response factor (CRF) or transcriptional activators that would explain the deviation from the predicted activity (Figure 4, red data points). Four MinSyns were selected for further analysis (Figure 4, blue data points).
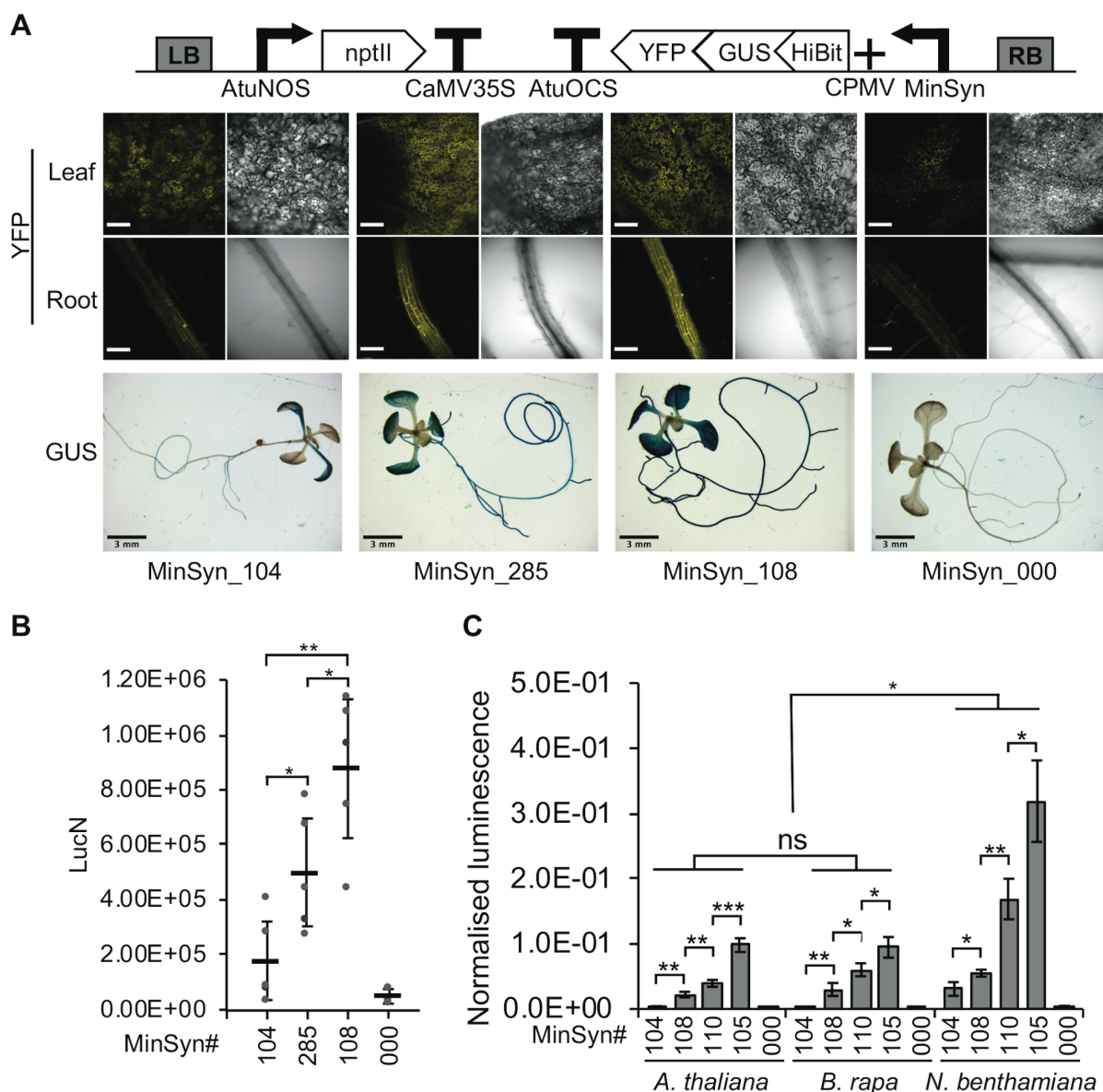
Figure 5. Characterisation of Minimal Synthetic Promoters (MinSyns). (A) Expression from stably integrated transgenes detected by florescence microscopy (yellow fluorescent protein; YFP), scale bar = 150 μm and histochemical staining (β-glucuronidase; GUS), scale bar = 3 mm. (B) Expression levels from stably integrated transgenes quantified by detection of luminescence via Hi-Bit tag and normalised to transgene copy number. (C) Transient expression levels in three plant species. Error bars = 2 x standard error; n=5; P-values were calculated using unpaired two-tailed Student's t-test; *P ≤0.05, **P≤0.01, ***P≤0.001; ns= not significant.

MinSyns function in multiple species and as stable transgenes

Transient expression is used both for rapid experimentation and for production-scale protein expression in plants (3, 54). Therefore, minimal promoters that perform as expected in transient expression are useful for several applications. However, for other applications, the ability to maintain expected levels of expression when stably integrated into the genome is desirable. Other studies have reported a strong correlation between the performance of transiently-expressed and stably-integrated transgenes (24). To investigate the performance of MinSyns in stably-integrated transgenes, MinSyns of varying strengths were fused to multifunctional synthetic reporter protein-fusion enabling qualitative and quantitative detection of expression by luminescence, fluorescence and histochemical staining. Patterns of expression were assessed in five independent transgenic lines by GUS-staining and fluorescence microscopy (Figure 5A). Additionally, protein was extracted and expression quantified by detection of luminescence with data normalised to transgene copy number as determined by digital droplet PCR (Figure 5B). As expected, expression levels varied somewhat between independent lines (most likely the effect of local genomic context), however, the MinSyns expressed in most leaf and root tissues and the trends of expression levels observed in transient assays were maintained in stable lines (Figure 5B). We then compared the performance of MinSyns in two additional plant species, *Brassica rapa* and *Nicotiana benthamiana,* in transient protoplast assays. The overall expression trend observed in these species was maintained, with expression levels in *B. oleracea* being comparable to Arabidopsis, but expression levels in *N. benthamiana* being slightly higher (Figure 5b).

Minimal synthetic elements for plants enable relative control of gene expression in synthetic genetic circuits.

To demonstrate the utility of MinSyns in synthetic genetic circuits, we constructed simple multigene constructs in which all promoter elements were synthetic. Initially, we simply used a MinSyn to initiate transcriptional flow by controlling expression of an orthogonal TF, which activated expression of reporter (Figure 6A). Similar levels of expression were detected to circuits in which the TF was controlled by CaMV35s, which is widely used to initiate transcription in transgenic plants. We then demonstrated the ability to control the relative ratio of expression of two genes using two MinSyns with different numbers of cognate binding sites for an orthogonal TF to control expression of two reporters and a third MinSyn to control expression of the TF (Figure 6B).
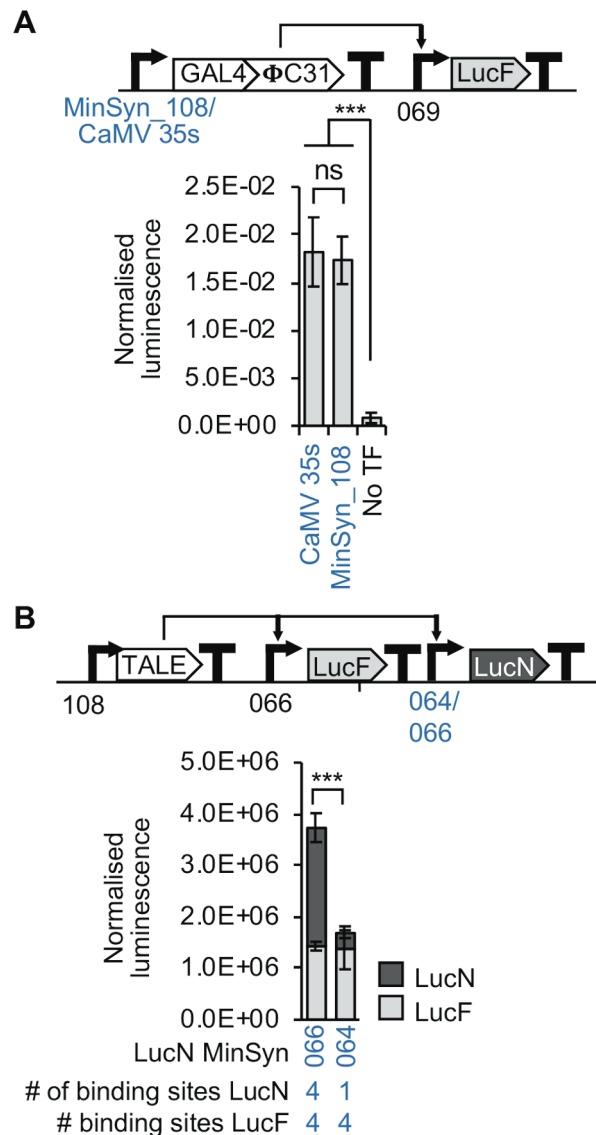
Figure 6. Initiation transcription from simple genetic devices with Minimal Synthetic Promoters (MinSyns). (A) Constitutively expressed MinSyns drive expression of the orthogonal transcriptional factor GAL4:ΦC31, which regulates expression of a reporter. (B) The relative expression of two reporters is regulated using MinSyns with different numbers of binding sites for Transcriptional Activation Like Effectors (TALES). Error bars = 2 x standard error; n=3; P-values were calculated using unpaired two-tailed Student's t-test; ***P≤0.001; ns = not significant.

DISCUSSION

Despite their dominance in plant research and biotechnology, comprehensive sequence analyses of even the most widely used constitutive promoters have not previously been reported. Analysis of expression levels of TF predicted to bind to these

promoters, indicate that constitutive expression is unlikely to depend on steady-state presence of specific TFs across multiple cell types, but rather on the ability to utilise a wider range of TFs present in different cell types (Figure 1A and Supplementary Data 1). This is consistent with data obtained from early experiments in which the use of specific subdomains of CaMV35S resulted in tissue-specific expression (55). Promoters from numerous plant pathogens that have evolved to utilise the plants transcriptional machinery contain a common regulatory (C-CRE), likely to either directly or indirectly bind the TGA sub-class of bZIP TFs (50–53). This C-CRE has significant effect on the expression levels of both natural and synthetic promoters (Figure 1B and C, Figure 3B) and was the only CRE able to promote detectable levels of expression without the presence of additional functional elements (Figure 3A and B). Several bZIP TFs are known to have a role in different disease and stress response pathways (56–58), which could therefore explain their dominance in pathogen regulatory elements. However, these promoters are known to confer broadly constitutive expression of stably-integrated transgenes, including in healthy, non-stressed plants. Several bZIP TFs have been shown to function as pioneer TFs, able to displace nucleosomes in chromatin inaccessible to other TFs, thus enabling the assembly of other TFs (59, 60). It has recently been hypothesised that some bZIP proteins inhibit chromatin compaction, initiating the formation of enhanceosomes (higher-order multicomponent transcription factor– enhancer complexes) (61). Optimal positioning of pioneer TFs, in particular, has been suggested to be necessary for gene expression (62), which could explain the significant impact of relocation (Figure 1C). However, such roles have yet to be determined for plant TGA TFs. Further, although Transfer- DNAs (T-DNAs) integrated into plant nuclear genomes via *Agrobacterium*-mediated transformation might be packaged into chromatin, thus supporting a role for pioneer TFs in the promoters of the *NOS*, *MAS* and *OCS* opine biosynthetic genes from *A. tumefaciens,* the structures of Caulimovirus DNA in plant nuclei are unknown.

Other than bZIP-binding C-CREs, multiple CREs needed to be combined to obtain significant expression from MinSyns (Figure 3A). These data indicate that, in the absence of the bZIP binding motif, multiple TFs are required for the proper recruitment of the transcriptional machinery. Previous studies have presented evidence that TF-complexes enable transcription either through direct protein-protein interactions or through the formation of

enhanceosome complexes, but also without direct protein-protein interactions via a synergistic or collaborative binding process sometimes called passive cooperativity (63, 64). Varying the relative positions and combinations of CREs within the MinSyns variable region revealed that direct protein-protein-interactions were unlikely (Figure 3C), therefore passive cooperativity is a reasonable hypothesis. This is consistent with experiments demonstrating that TFs can be substituted within enhancer complexes, enabling enhancer re-engineering by exchanging TF motifs (65). However, passive cooperativity is proposed to enable the displacement of nucleosomes. While all core histones and the linker histone, H1, have been shown to associate with transiently delivered exogenous DNA in mammalian cells (albeit with aberrant stoichiometry) (66, 67), this has not been investigated in plant cells. Several synthetic promoters and cognate orthogonal TFs for plants, including those that can be induced by chemical signals, have been engineered for plant systems (24–26). In this study we aimed to expand on those efforts, creating regulatory elements of different strengths for use in the construction of larger genetic circuits, particularly biosynthetic pathways, in which it is desirable to control the relative expression levels of different proteins. We provide two options for such constructs: MinSyns of different strengths regulated by endogenous TFs (Figure 5C) or MinSyns of different strengths regulated by synthetic orthogonal TFs (Figure 2 and Figure 6). While the strength of MinSyns that bind orthogonal TFs correlates directly with the number of TF binding sites (Figure 2A, Figure 6A), predicting the strength of constitutive MinSyns that utilise endogenous plant TFs was more challenging. The strength of the computationally designed MinSyns were broadly predictable (Figure 4) but predictability was undermined by the inadvertent introduction of additional TFBSs at sequence junctions. Similar issues were encountered during the creation of synthetic promoters for yeast (68, 69), however the availability of complete datasets of yeast TFBSs allowed programming scripts to be modified to exclude these sequence motifs (68). We considered modifying the script for plant MinSyns to discard sequences in which additional elements were formed, however, as the dataset is incomplete the results would be unpredictable. A second option would be to include any newly created TFBSs in the prediction of strength. This also proved challenging, as relatively few plant TF-DNA interactions have been functionally characterised. In addition, the local context of binding sites has been shown to alter the activity of some TFs from repressors to activators (65, 70), making it difficult to predict the impact on overall expression

levels. Indeed, this phenomenon could also contribute to the difference between predicted and observed strengths of some MinSyns.

These investigations have enabled us to design a suite of minimal synthetic plant promoters of varied strengths, activated by either endogenous or orthogonal transcription factors, that provide numerous options for the construction of large and complex genetic circuits for dicotyledonous plants. We have characterised their performance as stable transgenes, finding that transient assays were broadly predictive of behaviour. In previous work, we have observed that permutations of other components such untranslated regions and terminator sequences also impacts the final expression levels of a synthetic transcriptional unit (71). In this work we have controlled for variance by maintaining the same sequences, allowing us to measure the intrinsic properties of the promoters. Further work will be required to determine if and how the properties of MinSyns are modulated when used in combination with different sequence elements.

FUNDING

CONFLICT OF INTEREST

The authors have no conflicts of interests to declare.

ACKNOWLEDGEMENTS

REFERENCES

1. International Service for the Acquisition of Agribiotech Applications (ISAAA) (2018) Brief 54: Global Status of Commercialized Biotech/GM Crops.

2. Fox,J.L. (2012) First plant-made biologic approved. *Nat. Biotechnol.*, 30, 472.

3. Sainsbury,F. (2020) Innovation in plant-based transient protein expression for infectious disease prevention and preparedness. *Curr. Opin. Biotechnol.*, 61, 110–115.

4. Vazquez-Vilar,M., Orzaez,D. and Patron,N. (2018) DNA assembly standards: Setting the low-level programming code for plant biotechnology. *Plant Sci.*, 273, 33–41.

5. de Lange,O., Klavins,E. and Nemhauser,J. (2018) Synthetic genetic circuits in crop plants. *Curr. Opin. Biotechnol.*, 49, 16–22.

6. Andres,J., Blomeier,T. and Zurbriggen,M.D. (2019) Synthetic switches and regulatory circuits in plants. *Plant Physiol.*, 179, 862–884.

7. Van Lijsebettens,M. and Grasser,K.D. (2014) Transcript elongation factors: Shaping transcriptomes after transcript initiation. *Trends Plant Sci.*, 19, 717–726.

8. Swiezewski,S., Liu,F., Magusin,A. and Dean,C. (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature*, 462, 799–802.

9. Martínez de Alba,A.E., Elvira-Matelot,E. and Vaucheret,H. (2013) Gene silencing in plants: A diversity of pathways. *Biochim. Biophys. Acta - Gene Regul. Mech.*, 1829, 1300–1308.

10. Merchante,C., Stepanova,A.N. and Alonso,J.M. (2017) Translation regulation in plants: an interesting past, an exciting present and a promising future. *Plant J.*, 90, 628–653.

11. Zhang,H. and Jiang,T. (2010) Synthetic circuits, devices and modules. *Protein Cell*, 1, 974–978.

12. Koncz,C., De Greve,H., André,D., Deboeck,F., Van Montagu,M. and Schell,J. (1983) The opine synthase genes carried by Ti plasmids contain all signals necessary for expression in plants. *EMBO J.*, 2, 1597–1603.

13. Assaad,F.F. and Signer,E.R. (1990) Cauliflower mosaic virus P35S promoter activity in Escherichia coli. *Mol. Gen. Genet.*, 223, 517–520.

14. Rüth,J., Hirt,H. and Schweyen,R.J. (1992) The cauliflower mosaic virus 35S promoter is regulated by cAMP in Saccharomyces cerevisiae. *MGG Mol. Gen. Genet.*, 235, 365–372.

15. Sun,L., Cai,H., Xu,W., Hu,Y. and Lin,Z. (2002) CaMV 35S promoter directs β-glucuronidase expression in Ganoderma lucidum and Pleurotus citrinopileatus. *Mol. Biotechnol.*, 20, 239.

16. Vlasák,J., Šmahel,M., Pavlík,A., Pavingerová,D. and Bříza,J. (2003) Comparison of hCMV immediate early and CaMV 35S promoters in both plant and human cells. *J. Biotechnol.*, 103, 197–202.

17. Seternes,T., Tonheim,T.C., Myhr,A.I. and Dalmo,R.A. (2016) A plant 35S CaMV promoter induces long-term expression of luciferase in Atlantic salmon. *Sci. Rep.*, 6, 25096.

18. Ebert,P.R., Ha,S.B. and An,G. (1987) Identification of an essential upstream element in the nopaline synthase promoter by stable and transient assays. *Proc. Natl. Acad. Sci.*, 84, 5745–5749.

19. Benfey,P.N. and Chua,N.H. (1990) The cauliflower mosaic virus 35S promoter: Combinatorial regulation of transcription in plants. *Science*, 250, 959–966.

20. Fang,R.X., Nagy,F., Sivasubramaniam,S. and Chua,N.H. (1989) Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell*, 1, 141–50.

21. Bhullar,S., Chakravarthy,S., Advani,S., Datta,S., Pental,D. and Burma,P.K. (2003) Strategies for development of functionally equivalent promoters with minimum sequence homology for transgene expression in plants: cis-Elements in a novel DNA context versus domain swapping. *Plant Physiol.*, 132, 988–998.

22. Venter,M. (2007) Synthetic promoters: genetic control through cis engineering. *Trends Plant Sci.*, 12, 118–124.

23. Ali,S. and Kim,W.C. (2019) A Fruitful Decade Using Synthetic Promoters in the Improvement of Transgenic Plants. *Front. Plant Sci.*, 10, 1433.

24. Schaumberg,K.A., Antunes,M.S., Kassaw,T.K., Xu,W., Zalewski,C.S., Medford,J.I. and Prasad,A. (2015) Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat. Methods*, 13, 94–100.

25. Zuo,J. and Chua,N.H. (2000) Chemical-inducible systems for regulated expression of plant genes. *Curr. Opin. Biotechnol.*, 11, 146–51.

26. Brückner,K., Schäfer,P., Weber,E., Grützner,R., Marillonnet,S. and Tissier,A. (2015) A library of synthetic transcription activator-like effector-activated promoters for coordinated orthogonal gene expression in plants. *Plant J.*, 82, 707–16.

27. Liu,W. and Stewart,C.N. (2016) Plant synthetic promoters and transcription factors. *Curr. Opin. Biotechnol.*, 37, 36–44.

28. Salis,H.M., Mirsky,E.A. and Voigt,C.A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.*, 27, 946–50.

29. Khalil,A.S., Lu,T.K., Bashor,C.J., Ramirez,C.L., Pyenson,N.C., Joung,J.K. and Collins,J.J. (2012) A synthetic biology framework for programming eukaryotic transcription functions. *Cell*, 150, 647–658.

30. Amit,R., Garcia,H.G., Phillips,R. and Fraser,S.E. (2011) Building enhancers from the ground up: A synthetic biology approach. *Cell*, 146, 105–111.

31. Schlabach,M.R., Hu,J.K., Li,M. and Elledge,S.J. (2010) Synthetic design of strong promoters. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 2538–43.

32. Redden,H. and Alper,H.S. (2015) The development and characterization of synthetic minimal yeast promoters. *Nat. Commun.*, 6, 7810.

33. Kotopka,B.J. and Smolke,C.D. (2020) Model-driven generation of artificial yeast promoters. *Nat. Commun.*, 11, 2113.

34. Sharon,E., Kalma,Y., Sharp,A., Raveh-Sadka,T., Levo,M., Zeevi,D., Keren,L., Yakhini,Z., Weinberger,A. and Segal,E. (2012) Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.*, 30, 521–530.

35. Klemm,S.L., Shipony,Z. and Greenleaf,W.J. (2019) Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.*, 20, 207–220.

36. O'Malley,R.C., Huang,S.S.C., Song,L., Lewsey,M.G., Bartlett,A., Nery,J.R., Galli,M., Gallavotti,A. and Ecker,J.R. (2016) Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell*, 165, 1280–1292.

37. Smaczniak,C., Angenent,G.C. and Kaufmann,K. (2017) SELEX-seq: A method to determine DNA binding specificities of plant transcription factors. *Methods Mol. Biol.*, 1629, 67–82.

38. Ricci,W.A., Lu,Z., Ji,L., Marand,A.P., Ethridge,C.L., Murphy,N.G., Noshay,J.M., Galli,M., Mejía-Guerra,M.K., Colomé-Tatché,M., *et al.* (2019) Widespread long-range cis-regulatory elements in the maize genome. *Nat. Plants*, 5, 1237–1249.

39. Hernandez-Garcia,C.M. and Finer,J.J. (2014) Identification and validation of promoters and cis-acting regulatory elements. *Plant Sci.*, 217–218, 109–119.

40. Rodríguez-Leal,D., Lemmon,Z.H., Man,J., Bartlett,M.E. and Lippman,Z.B. (2017) Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell*, 171, 470–480.

41. Jin,J., Tian,F., Yang,D.C., Meng,Y.Q., Kong,L., Luo,J. and Gao,G. (2017) PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.*, 45, D1040–D1045.

42. Grant,C.E., Bailey,T.L. and Noble,W.S. (2011) FIMO: Scanning for occurrences of a given motif. *Bioinformatics*, 27, 1017–1018.

43. Petryszak,R., Keays,M., Tang,Y.A., Fonseca,N.A., Barrera,E., Burdett,T., Füllgrabe,A., Fuentes,A.M.P., Jupp,S., Koskinen,S., *et al.* (2016) Expression Atlas update - An integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.*, 44, D746–D752.

44. Patron,N.N.J., Orzaez,D., Marillonnet,S., Warzecha,W., Matthewman,C., Youles,M., Raitskin,O., Leveau,A., Farré,G., Rogers,C., *et al.* (2015) Standards for Plant Synthetic Biology: A Common Syntax for Exchange of DNA Parts. *New Phytol.*, 208, 13–19.

45. Vazquez-Vilar,M., Quijano-Rubio,A., Fernandez-Del-Carmen,A., Sarrion-Perdigones,A., Ochoa-Fernandez,R., Ziarsolo,P., Blanca,J., Granell,A. and Orzaez,D. (2017) GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data. *Nucleic Acids Res.*, 45, 2196–2209.

46. Patron,N. (2016) DNA Assembly for Plant Biology. *Curr. Protoc. Plant Biol.*, 1, 604–616.

47. Yoo,S.D., Cho,Y.H. and Sheen,J. (2007) Arabidopsis mesophyll protoplasts: A versatile cell system for transient gene expression analysis. *Nat. Protoc.*, 2, 1565–72.

48. Raitskin,O., Schudoma,C., West,A. and Patron,N.J. (2019) Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: An expanded toolkit for precision genome engineering. *PLoS One*, 14, e0211598.

49. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: An open-source platform for biological-image analysis. *Nat. Methods*, 9, 676–682.

50. Lam,E., Benfey,P.N., Gilmartin,P.M., Fang,R.X. and Chua,N.H. (1989) Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. *Proc. Natl. Acad. Sci. U. S. A.*, 86, 7890–7894.

51. Lam,E., Katagiri,F. and Chua,N.H. (1990) Plant nuclear factor ASF-1 binds to an essential region of nopaline synthase promoter. *J. Biol. Chem.*, 265, 9903–9913.

52. Lam,E. and Lam,Y. kam P. (1995) Binding site requirements and differential representation of TGA factors in nuclear ASF-1 activity. *Nucleic Acids Res.*, 23, 3778–37785.

53. Zhang,B., Chen,W., Foley,R.C., Buttner,M. and Singh,K.B. (1995) Interactions between distinct types of DNA binding proteins enhance binding to ocs element promoter sequences. *Plant Cell*, 7, 2241–2252.

54. Sainsbury,F. and Lomonossoff,G.P. (2008) Extremely high-level and rapid transient protein production in plants without the use of viral replication. *Plant Physiol.*, 148, 1212–8.

55. Benfey,P.N., Ren,L. and Chua,N.H. (1990) Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. *EMBO J.*, 9, 1677–1684.

56. Kesarwani,M., Yoo,J. and Dong,X. (2007) Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in Arabidopsis. *Plant Physiol.*, 144, 336–346.

57. Zander,M., La Camera,S., Lamotte,O., Métraux,J.-P. and Gatz,C. (2010) Arabidopsis Thaliana class-II TGA Transcription Factors Are Essential Activators of Jasmonic Acid/Ethylene-Induced Defense Responses. *Plant J.*, 61, 200–210.

58. Zander,M., Thurow,C. and Gatz,C. (2014) TGA transcription factors activate the salicylic acid-suppressible branch of the ethylene-induced defense program by regulating ORA59 expression. *Plant Physiol.*, 165, 1671–1683.

59. Biddie,S.C., John,S., Sabo,P.J., Thurman,R.E., Johnson,T.A., Schiltz,R.L., Miranda,T.B., Sung,M.H., Trump,S., Lightman,S.L., *et al.* (2011) Transcription Factor AP1 Potentiates Chromatin Accessibility and Glucocorticoid Receptor Binding. *Mol. Cell*, 43, 145–155.

60. Collins,C., Wang,J., Miao,H., Bronstein,J., Nawer,H., Xu,T., Figueroa,M., Muntean,A.G. and Hess,J.L. (2014) C/EBPα is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, 111, 9899–9904.

61. Hörberg,J. and Reymer,A. (2019) BZip Transcription Factors Modulate DNA Supercoiling Transitions. *bioRxiv*.

62. Grossman,S.R., Engreitz,J., Ray,J.P., Nguyen,T.H., Hacohen,N. and Lander,E.S. (2018) Positional specificity of different transcription factor classes within enhancers. *Proc. Natl. Acad. Sci. U. S. A.*, 115, E7222–E7230.

63. Reiter,F., Wienerroither,S. and Stark,A. (2017) Combinatorial function of transcription factors and cofactors. *Curr. Opin. Genet. Dev.*, 43, 73–81.

64. Deplancke,B., Alpern,D. and Gardeux,V. (2016) The Genetics of Transcription Factor DNA Binding Variation. *Cell*, 166, 538–554.

65. Stampfel,G., Kazmar,T., Frank,O., Wienerroither,S., Reiter,F. and Stark,A. (2015) Transcriptional regulators form diverse groups with context-dependent regulatory functions. *Nature*, 528, 147–51.

66. Hebbar,P.B. and Archer,T.K. (2008) Altered

histone H1 stoichiometry and an absence of nucleosome positioning on transfected DNA. *J. Biol. Chem.*, 283, 4595–601.

67. Mladenova,V., Mladenov,E. and Russev,G. (2009) Organization of plasmid dna into nucleosome-like structures after transfection in eukaryotic cells. *Biotechnol. Biotechnol. Equip.*, 23, 1044–1047.

68. Curran,K.A., Crook,N.C., Karim,A.S., Gupta,A., Wagman,A.M. and Alper,H.S. (2014) Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nat. Commun.*, 5, 4002.

69. de Boer,C.G., Vaishnav,E.D., Sadeh,R., Abeyta,E.L., Friedman,N. and Regev,A. (2020) Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. *Nat. Biotechnol.*, 38, 56–65.

70. Le,D.D., Shimko,T.C., Aditham,A.K., Keys,A.M., Longwell,S.A., Orenstein,Y. and Fordyce,P.M. (2018) Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc. Natl. Acad. Sci. U. S. A.*, 115, E3702–E3711.

71. Engler,C., Youles,M. and Grüetzner,R. (2014) A Golden Gate modular cloning toolbox for plants. *ACS Synth. Biol.*, 3, 839–843.