Synthetic TAL effectors for targeted enhancement of transgene expression in plants

Wusheng Liu1, Mary R. Rudis1, Yanhui Peng1, Mitra Mazarei1, Reginald J. Millwood1, Jian-Ping Yang2, Wenzhi Xu1, Jonathan D. Chesnut2 and Charles Neal Stewart Jr1,2,*

1Department of Plant Sciences, The University of Tennessee, Knoxville, TN, USA
2Synthetic Biology Research and Development, Life Technologies, Carlsbad, CA, USA

Received 23 January 2013; revised 19 October 2013; accepted 2 November 2013.
*Correspondence (Tel 865 974 7324; fax 865 946 1989; email nealstewart@utk.edu)

Summary
Transcription activator-like effectors (TALEs), secreted by the pathogenic bacteria Xanthomonas, specifically activate expression of targeted genes in plants. Here, we designed synthetic TALEs that bind to the flanking regions of the TATA-box motif on the CaMV 35S promoter for the purpose of understanding the engineerable ‘hot-spots’ for increasing transgene expression. We demonstrated that transient expression of de novo-engineered TALEs using agroinfiltration could significantly increase reporter gene expression in stable transgenic tobacco expressing the orange fluorescent protein reporter gene pporRFP under the control of synthetic inducible, minimal or full-length 35S promoters. Moreover, the additive effects of a combination of two different synthetic TALEs could significantly enhance the activation effects of TALEs on reporter gene expression more than when each TALE was used individually. We also studied the effects of the C-terminal domain and the activation domain of synthetic TALEs, as well as the best ‘hot-spots’ on the 35S promoter on targeted transgene activation. Furthermore, TALE activation of the Arabidopsis MYB transcription factor AtPAP1 (production of anthocyanin pigment 1) in stable transgenic tobacco gave rise to a dark purple colour on infiltrated leaves when driven by four copies of cis-regulatory elements of pathogenesis-related gene (PR1) with enhancer motifs B and A1 from the 35S promoter. These results provide novel insights into the potential applications of synthetic TALEs for targeted gene activation of transgenes in plants.

Keywords: targeted transgene activation, synthetic TALEs, additive effects, pporRFP, AtPAP1, transgenic plants.

Introduction
The ability to selectively activate or inhibit gene expression has value in basic plant biology and agricultural biotechnology. Targeted gene activation can be achieved by specifically expressing a gene of interest in a temporally and/or spatially defined manner in transgenic plants, or by directly activating endogenous gene expression in its original genetic context. The defined transgene expression can be driven using inducible or tissue-specific promoters, while the endogenous gene activation can be achieved using endogenous or synthetic transcription factors that specifically bind to the promoters of genes of interest and activate gene expression.

The ability to modify genomes and alter gene expression profiles using new synthetic biology tools is one of the most promising advances in agricultural biotechnology (Liu et al., 2013a). Promoter-specific activation of gene expression can be achieved by fusing engineered DNA-binding domains with sequence specificities to activation domains (Bogdanove and Voytas, 2011). The efficiency and precision of targeting mainly depends on the specificity of the DNA-binding domains. Even though helix-turn-helix and leucine zipper motifs are well-known DNA-binding protein folds, zinc finger (ZF) domains have been predominantly used for transcriptional activation (or repression) and recognize targets of various lengths in gene promoters. ZF domains bind to DNA targets as a monomer that consists of a tandem array of 3–6 (even 9) C2H2 fingers (Urnov et al., 2010). A tandem array of 3–6 ZFs corresponds to target sites of 9–18 bp long, permitting it to target unique sites in complex genomes. Nevertheless, the binding specificity of a ZF array is sometimes dependent on the specificities of neighbouring ZFs, requiring highly laborious screening of ZF libraries (Ramirez et al., 2008).

Early experiments suggested that transcription activator-like effectors (TALEs) secreted by the pathogenic bacteria Xanthomonas might provide a superior targeted gene activation system (Bogdanove and Voytas, 2011; Mussolino et al., 2011). The Xanthomonas TALE protein contains an N-terminal translösation domain, a central DNA-binding domain and a C-terminal activation domain (Kay et al., 2007; Romer et al., 2007). The DNA-binding domains of TALEs attach to their corresponding binding sites, which are called UPT boxes (UP regulated by TALEs), in the promoters of their target genes and confer TALE-mediated activation with the help of their activation domains (Kay et al., 2007; Romer et al., 2007). There is a novel DNA-binding domain of Xanthomonas TALEs consisting of 1.5–33.5 (mostly 15.5–19.5) tandem, nearly identical, repeats which are 30–42 (typically 34) amino acids long (Boch and Bonas, 2010). The specificity of each individual repeat is encoded in a repeat-variable diresidue (RVD) at positions 12 and 13 (Figure 1a), with the four most common diresidues being N, N, D and D, specifically binding to nucleotides A, T, C and G, respectively (Boch et al., 2009; Bogdanove and Voytas, 2011; Deng et al., 2012; Moscou and Bogdanove, 2009). These DNA-binding domains are predictable and completely neighbour-independent (Boch et al., 2009; Moscou and Bogdanove, 2009). Design and rearrangement of the DNA-binding repeats allows for the generation of novel synthetic TALE proteins corresponding with DNA-recognition specificities. It was demonstrated that different UPT boxes could
be engineered into the promoter region of a single plant resistance (R) gene in tobacco to recognize multiple TALEs from various pathogens (Romer et al., 2009). After agroinfiltration, de novo-engineered TALEs were used to activate transcription of targeted genes in Arabidopsis, tomato and pepper with known UPT boxes being embedded 46- to 108-bp upstream of transcription initiation sites (Morbitzer et al., 2010).

In this study, we generated synthetic TALEs by fusing their N-terminal translocation domains and de novo-engineered central DNA-binding domains with DNA-recognition specificities to the activation domain of the virion protein VP16 from human herpes simplex virus. We demonstrated that the engineered TALEs, when introduced using agroinfiltration, were capable of targeting desired promoter sequences of transgenes in stable transgenic tobacco. In addition, we investigated the effects of the C-terminal domain and activation domains of synthetic TALEs, the best ‘hot spots’ on the 35S promoter as well as the additive effects of different TALEs on targeted transgene activation. Finally, we applied the system to a visible reporter gene, AtPAP1, which results in high production of anthocyanin. Activation of AtPAP1 expression by synthetic TALEs could change transgenic tobacco leaf colour from green to deep purple.

Results

Effects of the C-terminal domain and activation domains of synthetic TALEs

T2 homozygous stable transgenic tobacco were generated expressing a single copy of the pporRFP reporter from the hard coral Porites porites (Alieva et al., 2008; Mann et al., 2012) driven by synthetic inducible, minimal or full-length CaMV 35S promoters. Using agroinfiltration, we tested the activation abilities of synthetic TALEs on these stable transgenic plants to determine ‘hot spots’ on the 35S promoter for transgene activation in addition to other TALE features and conditions important for activation.

The synthetic TALEs containing the N-terminal domain, combinations of different DNA-binding domains, the C-terminal domain and activation domains were cloned into destination vector pMDC32 where they were under twice the control of the full-length 35S promoter (i.e. 2 × 35S). Our initially designed DNA-binding sites of synthetic TALEs (i.e. T118, T218, T318) were 19 bp long and located in the flanking regions of the TATA-box motif of the minimal 35S promoter (i.e. −46 35S; Figure 1b) in a forward orientation. The binding site of T318 was located between the TATA-box motif and the transcription initiation site (Figure 1b; Table 1). The other two synthetic TALEs, T118 and T218, were bound to the regions covering A1 and −46 35S domains of the 35S promoter (Figure 1b; Table 1). After fusing to the translocation domain and DNA-binding domain of each synthetic TALE, 3 types of C-terminal domains (including activation domains), that is, truncated or full-length C-terminal domain together with VP16 activation domain or four head-to-tail copies of VP16 activation domains (Beerli et al., 1998; Geibler et al., 2011; Yang et al., 2000), were used in this study and designated as VP16 TR, VP16 FL and VP64, respectively. VP16 TR contains the truncated (i.e. the first 130 aa) C-terminal domain and VP16 activation domain, VP16 FL contains the full-length wild-type C-terminal domain and VP16 activation domain, whereas VP64 contains the full-length wild-type C-terminal domain and four head-to-tail copies of VP16 activation domains.

We first tested the effect of Agrobacterium titre on synthetic T218-VP64-activated transgenic expression in stable transgenic tobacco harbouring 35S::pporRFP or B4 × PR1A::pporRFP that

![Figure 1](image_url)
was driven by the inducible promoter 4 × PR1 containing the enhancer domains B and A1 from the 35S promoter (Benfey et al., 1990; Liu et al., 2011, 2013b). It was demonstrated that OD600 = 0.6 of Agrobacterium containing T218-VP64 resulted in the optimal induction of pporRFP fluorescence in transgenic tobacco harbouring 35S::pporRFP or B4 × PR1::pporRFP, when compared with OD600 = 0.3, 0.45, 0.75 or 1.0 of Agrobacterium containing the same TALE (Figure S1). At OD600 = 0.6, there was a gradual increase in pporRFP expression, with the first significant induction of pporRFP expression observed at 48 or 72 h post-inoculation (hpi; Figure S1). Thereafter, the concentration of Agrobacterium containing each TALE individually or in combination was adjusted to be OD600 = 0.6 for all the experiments.

Using Agrobacterium-mediated infiltration (OD600 = 0.6), the designed TALEs containing fusions of one of the three DNA-binding domains (i.e. T118, T218, T318) and one of the three types of the C-terminal domains and activation domain combinations (i.e. VP16 TR, VP16 FL and VP64) were transferred into T2 homozygous transgenic tobacco plants harbouring 35S::pporRFP or −46 35S::pporRFP (Figure 2). As shown in Figure 2, the basal expression level of the reporter gene pporRFP in transgenic tobacco containing 35S::pporRFP was 34 times higher than the transgenic tobacco harbouring −46 35S::pporRFP. In stable transgenic tobacco expressing 35S::pporRFP, time-course analysis of pporRFP reporter gene expression revealed a gradual increase in pporRFP expression during the time points 24, 48, 72, 96 hpi for Agrobacterium infiltration (Figure 3a–c) and up to 10 days post-inoculation (dpi; Figure S2). The first significant increase in pporRFP fluorescence was always detected at time point 48 hpi, irrespective of the C-terminal domains and binding sites used (Figure 3a–c). Moreover, reporter gene expression increased by up to twofold in stable transgenic tobacco plants harbouring these synthetic TALEs (Figure 3). Among the three C-terminal domains tested, VP64 caused slightly higher induction of pporRFP expression than truncated (VP16 TR) and full-length C-terminal domains (VP16 FL; Figure 3). Among the three binding sites tested, T218 induced the greatest pporRFP fluorescence in transgenic tobacco harbouring 35S::pporRFP (Figure 3b). However, T218 caused relatively lower levels of induction of pporRFP fluorescence in transgenic tobacco harbouring −46 35S::pporRFP (Figure 3d) than in transgenic tobacco harbouring 35S::pporRFP (Figure 3c), and significant induction of pporRFP expression was mainly observed at time point 72 or 96 hpi (Figure 3d), which was much later than on transgenic tobacco harbouring 35S::pporRFP (Figure 3c).

### Table 1 Nucleotide sequences of the DNA-binding sites of synthetic TALEs on the CaMV 35S promoter

<table>
<thead>
<tr>
<th>Tale</th>
<th>Orientation</th>
<th>Length</th>
<th>Domain</th>
<th>Position</th>
<th>Nucleotide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>118</td>
<td>Forward</td>
<td>19</td>
<td>A1, −46</td>
<td>−52 − −34</td>
<td>ttcttgcaagacctttcc</td>
</tr>
<tr>
<td>218</td>
<td>Forward</td>
<td>19</td>
<td>A1, −46</td>
<td>−60 − −42</td>
<td>tccacatctcgcag</td>
</tr>
<tr>
<td>318</td>
<td>Forward</td>
<td>19</td>
<td>−46</td>
<td>−21 − −4</td>
<td>tatttggagaaagc</td>
</tr>
<tr>
<td>418</td>
<td>Forward</td>
<td>25</td>
<td>A1</td>
<td>−77 − −53</td>
<td>baagggttcgcaatcc</td>
</tr>
<tr>
<td>518</td>
<td>Forward</td>
<td>25</td>
<td>A1</td>
<td>−82 − −58</td>
<td>bgttaagggttcgcaatcc</td>
</tr>
<tr>
<td>618</td>
<td>Forward</td>
<td>25</td>
<td>B5, A1</td>
<td>−104 − −80</td>
<td>baagggttcgcaatcc</td>
</tr>
<tr>
<td>718</td>
<td>Forward</td>
<td>25</td>
<td>B4, B5</td>
<td>−128 − −104</td>
<td>tccacacgtcttcaagag</td>
</tr>
<tr>
<td>819</td>
<td>Forward</td>
<td>25</td>
<td>B4</td>
<td>−146 − −122</td>
<td>tgccttggcaag</td>
</tr>
<tr>
<td>919</td>
<td>Forward</td>
<td>25</td>
<td>B3, B4</td>
<td>−169 − −145</td>
<td>tggccaccacagcagcag</td>
</tr>
</tbody>
</table>

The first nucleotide T (i.e. position 0 in Figure 1) of each synthetic TALE binding site is in bold and underlined. The CAAT motif is underlined. The TATA-box motif is located on position −31 − −25 on the 35S promoter.

### ‘Hotspots’ on the 35S promoter for synthetic TALEs activation

Because the binding site of T218 is located in the border region of A1 and −46 35S domains of the 35S promoter (Benfey et al., 1990) and is closer to the 5′-end of the 35S promoter than the other two TALEs (i.e. T118 and T318; Table 1; Figure 1), we designed additional TALEs whose binding sites were located upstream to T218 and tested the activation effects of these additional TALEs for targeted transgene activation. The newly synthesized TALEs contained a translocation domain, full-length C-terminal domain and VP64, and synthetic DNA-binding domains that bound to 25-bp-long DNA-binding sites in the forward orientation. These newly designed TALEs were designated as T418-, T518-, T618-, T718-, T818- and T919-VP64 (for targeted locations and nucleotide sequences; see Figure 1b and Table 1). Among these, the DNA-binding sites of T418-VP64 and T518-VP64 were located in the A1 domain of the 35S promoter and contained a CAAT motif (Figure 1b, Table 1). The binding site of T618-VP64 was located in the border of B5 and A1 domains of the 35S promoter and contained a reverse CAAT motif (Figure 1b, Table 1).

These second round TALEs as well as T218-VP64 were used to determine whether there were any ‘hot spots’ for transgene activation on the 35S promoter. These were agroinfiltrated into T2 homozygous transgenic tobacco plants expressing 35S::pporRFP or B4 × PR1A::pporRFP. As shown in Figure 4, T218-, T618- and T818-VP64 conferred the highest induction of pporRFP expression, with T218-VP64 resulting in slightly higher induction of pporRFP expression than T618- and T818-VP64, even though T818-VP64 only resulted in a slightly higher induction of pporRFP expression driven by 35S promoter at 96 hpi. The first significant induction by the three TALEs was observed at 24 hpi on transgenic tobacco harbouring 35S::pporRFP and 72 hpi on transgenic tobacco harbouring 4 × PR1A::pporRFP (Figure 4). Because the insertion of 4 × PR1 into B and A1 domains of 35S caused a 5-bp-long mismatch at the 3′-end of the binding site of T618, we were surprised to find that T618-VP64 still caused significant induction in pporRFP expression (Figure 4).

### Additive effects of co-expression of two synthetic TALEs

Because T218-, T618- and T818-VP64 caused the highest induction of pporRFP expression in transgenic tobacco plants (Figure 4), we used combinations of any two of these three synthetic TALEs to study the additive effects of co-expression of
two synthetic TALEs on the target promoter sequence in T₂ homozygous transgenic tobacco expressing 35S::pporRFP or B4 × PR1A::pporRFP.

Time-course analysis of pporRFP expression revealed that the combination of any two of the three TALEs conferred a greater induction of pporRFP expression in transgenic tobacco harbouring 35S::pporRFP or B4 × PR1A::pporRFP than did either individual TALE (Figures 5 and 6, S3). Combining TALEs could be a good strategy for increasing the additive effects of combinations of different TALEs. In our hands, the highest increase was 2.8-fold higher than control, with the first significant induction detected at 24–48 hpi in transgenic tobacco harbouring 35S::pporRFP, but at 48–72 hpi in transgenic tobacco harbouring B4 × PR1A::pporRFP (Figure 5). The additive effects of the co-expression of two TALEs on pporRFP expression driven by the 35S or B4 × PR1A promoter in stable transgenic tobacco was also demonstrated by the fluorescence of pporRFP visualized with an epifluorescent microscope (Figure 6) and with a closed Fluorcam system at 96 hpi (Figure S3).

To further confirm the co-delivery efficiency for two TALEs in combination via Agrobacterium, we conducted real-time RT-PCR to measure the relative expression of each TALE on spots that were infiltrated with each TALE individually or in combination in the same transgenic tobacco leaf expressing 35S::pporRFP, which was performed on 3 individual plants. As T218-, T618- and T818-VP64 are almost identical in nucleotide sequence, only the primers specifically designed for T218 and T618 worked whereas the primers specifically designed for T818-VP64 did not. As shown in Figure S4, the relative expression of T218 and T618 spots infiltrated with each construct individually was comparable with that from spots infiltrated with the two constructs in combination. Thus, the co-delivery efficiency for each TALE in combination is close to 1:1.

Effects of activated AtPAP1 expression on leaf colour change in transgenic tobacco plants by TALEs

The Arabidopsis gene PRODUCTION OF ANTHOCYANIN PIGMENT 1 (AtPAP1) is a conserved R2R3 Myb transcription factor (AtMyb75) involved in expression of a number of anthocyanins of phenylpropanoid biosynthesis (Borevitz et al., 2000). Overexpression of AtPAP1 gene results in activated expression of most anthocyanin biosynthetic pathway genes, leading to high anthocyanin accumulation, which results in purple pigmented transgenic Arabidopsis (Borevitz et al., 2000; Tohge et al., 2005) or tobacco (Xie et al., 2006) when under the control of the constitutive CaMV 35S promoter. We generated stable transgenic tobacco with AtPAP1 driven by –46 35S, B_A and 35S promoters as well as inducible synthetic promoters 4 × PR1 and B4 × PR1A. The transgenic tobacco plants harbouring –46 35S::PAP1 and 4 × PR1::PAP1 had green-coloured leaves, whereas the plants containing B_A::PAP1 and 35S::PAP1 had purple shoots similar to the above-mentioned study (Xie et al., 2006).

Because the PR1 motif is responsive to salicylic acid (SA) treatment and bacterial pathogen infection (Liu et al., 2011, 2013b), we applied SA, Pseudomonas syringae pv. tabaci and P. s. pv. tomato to the transgenic tobacco plants expressing 4 × PR1::PAP1 or B4 × PR1A::PAP1 as reported previously (Liu et al., 2011, 2013b). We found that neither inducible promoter was sufficient to change tobacco leaf colour following SA or...
bacterial treatments (Figure S5). However, the Agrobacterium-mediated transfer of the synthetic TALEs (including T618-VP64) into transgenic tobacco harbouring B49 PR1A::PAP1 resulted in high accumulation of anthocyanins in infiltrated leaves, which, in turn, changed leaf colour from green to deep purple (Figure 7). The leaf colour changes from the activation of the AtPAP1 gene were achieved by all the synthetic TALEs individually or in combination, irrespective of the DNA-binding sites. The purple pigmentation started to develop at 72 hpi as light purple reaching to deep purple at 96–120 hpi and persisted until senescence of the infiltrated leaf. The induced leaf colour change was correlated with the area of agroinfiltration site on the leaves: a larger area of agroinfiltration with synthetic TALEs (Figure 7a) in leaves was capable of changing leaf colour from green to purple, whereas smaller areas of agroinfiltration sites (Figure 7b) in leaves rendered no change.

Discussion
A long-pursued goal in plant molecular biology is transgene expression at a high level and/or in a spatial- and temporal-specific manner. In this study, we demonstrated that synthetic TALEs, when fused to the VP16 or VP64 activation domains, were capable of activating transgene expression by up to a 2.8-fold increase in tobacco (Figure 3–5). This is of special importance in targeted activation of transgenic plants where higher transgene expression is desired. The virion protein VP16 of herpes simplex virus (Triezenberg et al., 1988) is one of the most potent transcriptional activators [such as GAL4 (Fischer et al., 1988), p53 (Liu et al., 1993)] that strongly activate transcription in various expression systems when its activation domain is attached to the DNA-binding domain of a heterologous protein. The VP16 activation domain interacts with basal transcription factors, facilitates the assembly of the pre-initiation complex, and also recruits histone acetyltransferases (Hirai et al., 2010). However, it was reported that the VP16 activation domain could only bring strong activation from a promoter-proximal position (Hagmann et al., 1997). Therefore, the DNA-binding sites for the synthetic TALEs in this study were chosen to be in the flanking regions of the TATA-box motif of the 3SS promoter (within the region from –18 to –169 bp; Table 1; Figure 1).

We observed that T218-, T618- and T818-VP64 conferred the highest induction level of reporter gene expression in transgenic tobacco (Figure 4). This may imply that these three binding sites
Transgenic tobacco leaves were infiltrated with 35S::pporRFP plants harbouring activation by synthetic TALEs containing VP64 in stable transgenic tobacco. Because we used the transcriptional (co)factor binding than the other TALE binding contain fewer or more dispensable motifs for endogenous transcriptional (co)factor binding than the other TALE binding sites. For example, the binding sites of T415 and T515 contain a CAAT motif used to initiate transcription. Because we used the 2 × 35S promoter to drive expression of each TALE, it is conceivable that a positive feedback mechanism was set up for the TALE-mediated transgene activation. Moreover, we found that the additive effects of different combinations of synthetic TALEs could further enhance the transgene expression (Figure 5). The binding of multiple TALEs to one promoter region is in agreement with Romer et al. (2009), in which different TALE binding sites were engineered into the promoter region of a single tobacco resistance (R) gene and allowed the recognition of multiple TALE effectors from various pathogens. The co-delivery efficiency of each of T218 and T618 in combination via multiple TAL effectors from various pathogens. The co-delivery of B4PR1A::pporRFP expression in three independent biological experiments ± standard errors (vertical bars). Significant changes in pporRFP expression (indicated by asterisks) were calculated statistically using ANOVA PROC GLM (LSD, P < 0.05).

contain fewer or more dispensable motifs for endogenous transcriptional (co)factor binding than the other TALE binding sites. For example, the binding sites of T415 and T515 contain a CAAT motif used to initiate transcription. Because we used the 2 × 35S promoter to drive expression of each TALE, it is conceivable that a positive feedback mechanism was set up for the TALE-mediated transgene activation. Moreover, we found that the additive effects of different combinations of synthetic TALEs could further enhance the transgene expression (Figure 5). The binding of multiple TALEs to one promoter region is in agreement with Romer et al. (2009), in which different TALE binding sites were engineered into the promoter region of a single tobacco resistance (R) gene and allowed the recognition of multiple TALE effectors from various pathogens. The co-delivery efficiency of each of T218 and T618 in combination via Agrobacterium was confirmed to be similar (i.e. ~1:1 ratio) using real-time RT-PCR. Interestingly, we found that T318 whose binding site was located between the TATA-box and the start of transcription initiation (Figure 1) could activate the transgene expression (Figure 3c). As many UPT boxes contain TATA-box as well as non-TATA-box, it is possible that TALE itself is capable to recruit Pol II polymerase complex for transcription initiation. The VP16 activation domain does interact with basal transcription factors, facilitates the assembly of the pre-initiation complex and also recruits histone acetyltransferases (Hirai et al., 2010). In addition, another interesting application would be targeted endogenous gene activation in its original genomic context. It is expected that synthetic TALEs could be engineered to activate expression of any endogenous genes of interest in plants, except the epigenetically modified genes (Zhang et al., 2011). These epigenetically modified genes (e.g. the epigenetically silenced oct4 gene in murine neural stem cells) could be dramatically activated by synthetic TALEs when used together with a histone deacetylase inhibitor and a DNA methyltransferase inhibitor (Bultmann et al., 2012). It is also worth investigating whether the reverse-oriented synthetic TALEs have a similar function in gene activation as do the forward oriented TALEs.

Using synthetic TALEs for targeted gene activation provides a similar (even better) powerful strategy for selective alteration of gene expression in vivo as ZFs (Bogdanove and Voytas, 2011). The DNA-binding domains of both types of proteins can be designed and synthesized to bind to practically any promoter regions in plant cells. The major advantage of this strategy is its versatility. It offers the possibility of targeting any activation domains to any genomic loci when fused to the designed DNA-binding domains (Urnov and Rebar, 2002). The precision and specificity of these DNA-binding domains are of exceptional importance, considering the complexity of plant genomes and off-target binding could cause malfunction or even deleterious effects on plant cells. Both ZF proteins and TALEs provide very stringent DNA sequence specificity for targeted gene activation (Huang et al., 2011; Li et al., 2011; Morbitzer et al., 2010; Tesson et al., 2011; Townsend et al., 2009) and are capable of discriminating two nucleotide differences between the binding sites of two paralogous genes in a single plant genome (Curtin et al., 2011; Townsend et al., 2009; Zhang et al., 2013). Compared with ZF proteins, TALEs provide more predictable sequence specificity than ZFs whose desired specificity requires highly laborious screening of ZF libraries (Kim et al., 2010; Ramirez et al., 2008). Surprisingly, T618–VP64 still caused significant induction of reporter gene expression even though the binding site of T618–VP64, which is 25 bp long, and contained a 5-bp-long mismatch at its 3′-end on the B4 × PR1A promoter. Because synthetic TALEs (such as T118, T218 and T318) whose DNA-binding sites were 19-bp-long were capable of activating transgene expression (Figure 3), it is possible that the addition of a 5-bp-long mismatch at the 3′-end of its binding site does not severely disrupt the binding ability to its target. Zhang et al. (2011) also demonstrated that a mismatch of 4 nucleotides within the DNA-binding site could confer about up to 40–50% of the original reporter activity by TALEs in human cells. However, we cannot exclude the possibility that large amounts of TALEs present in the cells could drive expression of the transgene in a non-specific fashion.

Remarkably, targeted activation by synthetic TALEs changed leaf colour from green to deep purple in transgenic tobacco plants harbouring B4 × PR1A::PAP1 (Figure 7). We found that neither hormone (salicylic acid) treatment nor bacterial pathogen infections (Pseudomonas syringae pv. tabaci and P. syringae pv. tomato) could, alone, produce purple leaves in transgenic tobacco plants harbouring B4 × PR1A::PAP1 (Figure S5). We hypothesized that the inducible signal was too weak and/or transient to synthesize anthocyanin, for two reasons: (i) the B4 × PR1A promoter was inducible by both bacterial pathogen infection and salicylic acid treatment in transgenic tobacco and Arabidopsis plants harbouring B4 × PR1A::pporRFP (Liu et al., 2011, 2013b) and (ii) the

![Figure 4](image-url) 'Hot spots' on the 35S promoter for targeted transgene activation by synthetic TALEs containing VP64 in stable transgenic tobacco plants harbouring 35S::pporRFP (a), or B4 × PR1A::pporRFP (b). Stable transgenic tobacco leaves were infiltrated with Agrobacterium containing negative control vector or each individual TALE (OD₆₀₀ ~ 0.6). Expression of pporRFP was quantified with a Fluorolog®-3 system (Jobin Yvon and Glencore Instruments Spectra, Edison, NJ) at 0, 24, 48, 72, or 96 hpi. Three individual plants were used as independent replicates with two agroinfiltration targets per leaf on each plant. The infiltrated spots were excited at 530 nm, and emission spectra was scanned and recorded from 560 to 610 nm. Intensity was measured at 591 nm in counts per second (cps). Each bar represents the mean of pporRFP expression in three independent biological experiments ± standard errors (vertical bars). Significant changes in pporRFP expression (indicated by asterisks) were calculated statistically using ANOVA PROC GLM (LSD, P < 0.05).
induction of pporRFP expression in transgenic tobacco by 2 × 35S::TALEs remained until 10 dpi with agroinfiltration (Figure S2). It is reasonable to expect that anthocyanin production resulted from the activation of the AtPAP1 gene by the synthetic TALEs (Figure 7). It was also noted that the area (size) of the agroinfiltration site on tobacco leaves affected the colour change (Figure 7). The AtPAP1 gene could be utilized as a potential novel reporter gene when visual monitoring of the phenotype changes requires strong and persistent expression.

The modular nature of the TALE DNA-binding domain provides an attractive solution for sequence-specific DNA recognition and interaction in plant cells. Replacement of its activation domain with other functional modules (such as nucleases, recombinases, methyltransferases) permits precise genome modification in its target sequence. Even though almost all the TALEs could activate transgene expression (Figures 3 and 4), the large range of detected reporter gene activity suggests that other factors might affect the TALE targeting efficiency, such as differences in binding strength of individual repeats, position- or context-dependence of repeat binding strength, the number of mismatches or off-target activities. It is also worthwhile to test the effects of the native activation domains of TALEs because Geibler et al. (2011) showed that the native activation domain is more effective in plants than VP16 or VP64. In addition, the synthetic TALEs used in this research have 1340 amino acids, which renders them much larger than ZFs, and the proper folding of these proteins might be of concern in plant cells in the light of the high synthesis of proteins expressed via agroinfiltration.

Agrobacterium-mediated transient expression has been widely used as an efficient and rapid screening approach before conducting stable transformation experiments in plants, even though it cannot make absolute predictions for the behaviours of the stable transgenics. As synthetic TALEs could not activate the pporRFP expression driven by −46 35S promoter as well as driven by 35S promoter (Figure 3), we expect that TALEs probably will not be useful to reduce the line to line variation in stable transgenic plants (i.e. position effects).

**Experimental procedures**

**Synthesis of de novo-engineered TALEs**

TALE plasmid constructs were created at Life Technologies (www.lifetechnologies.com) by first using target site sequence
information from the 35S promoter to identify appropriate TALE RVD monomers using the code NN = G, NI = A, NG = T, HD = C (Boch et al., 2009). Arrays of 18 or 24 monomers were assembled using a modified Golden Gate/Type IIS assembly (Geibler et al., 2011) from a library of TALE RVD monomers. The synthetic TALEs containing the N-terminal domain, combinations of different DNA-binding domains, the C-terminal domain and activation domains were cloned into destination vector pMDC32 where they were under the control of the 2 × 35S promoter. The first set DNA-binding site targets for synthetic TALEs (i.e. T118, T218, T318) were 19 bp long and located in the flanking regions of the TATA-box motif of the minimal 35S promoter (i.e. −46 35S; Figure 1b) in a forward orientation. The binding site of T318 was located between the TATA-box motif and the transcription initiation site (Figure 1b; Table 1). The other two synthetic TALEs, T118 and T218, were bound to the regions covering A1 and −46 35S domains of the 35S promoter (Figure 1b; Table 1). After fusing to the translocation domain and DNA-binding domain of each synthetic TALE, 3 types of C-terminal domains (including activation domains), that is, truncated (TR) or full-length (FL) C-terminal domain together with VP16 activation domain or four head-to-tail copies of VP16 activation domain (VP64) (Beerli et al., 1998; Geibler et al., 2011; Yang et al., 2000), were used in this study and designated as VP16 TR, VP16 FL and VP64, respectively. VP16 TR contains the truncated (i.e. the first 130 aa) C-terminal domain and VP16 activation domain, VP16 FL contains the full-length wild-type C-terminal domain and VP16 activation domain, whereas VP64 contains the full-length wild-type C-terminal domain and four head-to-tail copies of VP16 activation domains.

**Plasmid construction**

The Gateway Entry vector containing each synthetic TALE was subsequently crossed with a destination vector pMDC32 (Curtis and Grossniklaus, 2003) by LR reactions. Thus, all the synthetic TALEs were under the control of 2 × 35S promoter in the destination vector pMDC32. The activation domain of VP64 was removed from pENTR-T218-VP64 with HindIII and EcoRI as well as the following T4 polynucleotide kinase (Promega, Madison, WI) treatment. The remaining pENTR-T218 was ligated with T4 ligase (Promega) and cloned into the destination plasmid pMDC32. This new construct, without an activation domain was named pMDC-TC and used as the negative control vector.

The details of the construction of B_A−, −46 35S−, B4 × PR1A− and 4 × PR1−pporRFP reporter fusion cassettes in pZP222 binary vector plasmids were reported previously (Liu et al., 2011). pZP B4 × PR1A pporRFP and pZP 4 × PR1 pporRFP contained four head-to-tail copies of cis-acting regulatory ele-

![Figure 6](image6.png) **Figure 6** Visual images of activation of pporRFP expression in stable transgenic tobacco by additive effects of synthetic TALEs. Fluorescence of pporRFP was visualized with an epifluorescent microscope (Olympus stereo microscope model SZX12, Olympus America, Center Valley, PA) using a tdTomato filter set (535/30 nm excitation and 600/50 nm band pass emission) and QCapture 2.56 imaging software, and an Olympus Q-colour 5 camera (Olympus, Center Valley, PA). Agrobacterium containing each TALE individually or in combination (i.e. 1:1 ratio) as well as negative control vector was applied to four different locations on the same leaf. Three individual plants were used as independent replicates. Typical and representative results are shown.

![Figure 7](image7.png) **Figure 7** Leaf colour change in transgenic tobacco harbouring AtPAP1 driven by the inducible promoter B4 × PR1A after activation by synthetic TALEs. Large area of agroinfiltration in the tobacco leaf with synthetic TALEs (a; shown in the blue outline) was capable of changing leaf colour, which was persistent until senescence, while small size of agroinfiltration sites (b; shown in the black outline) were not. The images were taken using a Nikon D90 camera with a 60 mm lens (Nikon Inc, Melville, NY).
ments of pathogenesis-related gene (PR1), whose sequence was reported previously (Liu et al., 2013b; Mazarei et al., 2008). The construct PZ 35S pporRFP was developed by PCR amplification of full-length CaMV 35S and insertion into pSK 4 × PR1 pporRFP (Liu et al., 2011) using XbaI and BamHI, followed by integration into plasmid pZP222 (Hajdukiewicz et al., 1994) with SacI and HindIII. The constructs pZP 4 × PR1 PAP1 and pZP B4 × PR1A PAP1 were constructed by PCR amplification of the Arabidopsis PAP1 gene (also called ATM875; AT1g56650) and insertion into pSK 4 × PR1 pporRFP (Liu et al., 2011) using BamHI and HindIII, followed by integration into pZP 222 with SacI and HindIII. Prior to that, PCR-mediated site-directed mutagenesis was conducted to remove the SpeI restriction site at the position of +573 in AtPAP1 cDNA with nucleotide A replaced by T but without changing the encoded amino acid.

**Plant transformation**

The stable transformation of plasmids containing the pporRFP reporter gene driven by B_A, −46 35S, B4 × PR1A and 4 × PR1 into Nicotiana tabacum (cv Xanthi) plants was reported by Liu et al. (2013b). The plasmids pZP 35S pporRFP, pZP B4 × PR1 A PAP1 and pZP 4 × PR1 PAP1 were transformed into Nicotiana tabacum (cv Xanthi) by the leaf disc transformation method (Horsch et al., 1985) using Agrobacterium tumefaciens strain GV3850. Transgenic tobacco plants were selected on Murashige and Skoog (MS) media with 200 mg/L gentamicin.

**Transgenic plants**

Transgenic tobacco plants harbouring a single copy of each construct were grown in a growth chamber at 25 °C under fluorescent white light with a 16:8 h light/dark cycle for three weeks, then in a greenhouse with a 16:8 h light/dark cycle at 27–30 °C with self-fertilization to eventually generate single copy, T₂ homozygous plants. For targeted gene activation experiments, transgenic tobacco plants were grown under the same growth chamber conditions as above for 6–7 weeks.

**Agroinfiltration of synthetic TALEs on the stable transgenic tobacco leaves**

The agroinfiltration of each synthetic TALE construct individually or in combination as well as the negative control vector in stable transgenic tobacco expressing 35S::pporRFP DNAse I was used to remove trace contamination of genomic DNA following to manufacturer’s instructions (Promega). Approximately 2 µg of total RNA was used for reverse transcription to generate cDNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA). Real-time RT-PCR was conducted in triplicate to measure the levels of TALE mRNA abundance from each agroinfiltrated spot using Power SYBR Green PCR master mix (Applied Biosystems) following to the manufacturer’s protocol. A tobacco L25 gene was used as an internal control (Schmidt and Delaney, 2010). Specific primers were designed (Table S1) for specific amplification of a single product from each corresponding gene as confirmed by the melting temperature of the amplicons and gel electrophoresis. Relative quantification was calculated using the standard curve method with R² value being above 0.99 and Efficiency (E) being 99.5–102%, and the quantity of each gene was normalized to the expression of the tobacco L25 gene.

**Data analysis**

Data analysis was conducted as in Liu et al. (2011). The raw data collected at each time point were normalized against data collected from non-transgenic Xanthi, which was infiltrated with infiltration solution at each respective time point. Fold change of pporRFP fluorescence was calculated using the normalized data at different time points 24, 48, 72 or 96 hpi divided by the normalized data at time point 0 h. Analysis of variance (anova) using PROC GLM was performed for statistical analyses (P < 0.05) (SAS 9.2 for Windows; SAS Institute Inc, Cary, NC).

**Acknowledgements**

We gratefully acknowledge support from USDA–NIFA, the University of Tennessee and Life Technologies. Life Technologies produced the synthetic TALEs and has a financial interest in producing TALEs. We appreciate 4 anonymous reviewers for providing critical assessment of the paper, which greatly improved it.

**References**


Synthetic TAL effectors for transgene activation


**Supporting information**

Additional Supporting information may be found in the online version of this article:

- **Figure S1** Effect of *Agrobacterium* titre on T218-VP64-activated expression of the *pporRFP* reporter in stable transgenic tobacco plants harbouring 35S::*pporRFP* (a) or B4 × PR1A::*pporRFP* (b) as measured by orange fluorescence.
- **Figure S2** Effects of the C-terminal domain and activation domain of synthetic T118, T218 or T318 on expression of the *pporRFP* reporter in stable transgenic tobacco plants harbouring 35S::*pporRFP* at time point 10 days post-inoculation (dpi).
- **Figure S3** Visual images (representative) of activation of *pporRFP* expression in stable transgenic tobacco harbouring 35S::*pporRFP* by additive effects of synthetic TALEs.
- **Figure S4** The efficiency of *Agrobacterium*-mediated co-delivery of two TALEs measured by real-time PCR at 96 hpi.
- **Figure S5** Visual images (representative) of AtPAP1 expression in stable transgenic tobacco harbouring B4 × PR1A::PAP1 by either 4 mM salicylic acid treatment (a; shown in the black outline) or bacterial pathogen *Pseudomonas syringae* pv. *tabaci* infection at OD600 = 0.03 (b; shown in the black outline) at 72 hpi.
- **Table S1** Optimized primer sequences and reaction conditions used for real-time PCR.