# Optimized dCas9 Programmable Transcriptional Activators for Plants

**Matthew H Zinselmeier[1,3,4], J. Armando Casas-Mollano[2,3,4], Jonathan Cors[2,3,4], Adam Sychla[2,3,4], Stephen C Heinsch[3,4,5], Daniel F Voytas[1,3,4,✉], and Michael J Smanski[2,3,4,5✉]**

[1]Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN 55455
[2]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, MN, 55455
[3]Center for Precision Plant Genomics, 1500 Gortner Avenue, Cargill Building, Saint Paul, MN 55108
[4]Biotechnology Institute, University of Minnesota, Saint Paul, MN 55108
[5]Bioinformatics and Computational Biology Graduate Program, University of Minnesota, Saint Paul, MN 55108

**Understanding how gene expression impacts plant development and physiology is important for rationally engineering improved crops. Programmable Transcriptional Activators (PTAs), including CRISPR-Cas activators, have traditionally relied on a limited number of transcriptional activation domains. Usually the VP64 domain, derived from human herpes simplex virus, is fused to a DNA-binding domain to activate target gene expression. We reasoned there was considerable space for PTA improvement by replacing VP64 with a plant-derived activation domain. To address this, we designed, built, and tested a PTA library of 38 putative plant transcriptional activation domains. Domains from *HSFA6b, AvrXa10, DOF1, DREB1*, and *DREB2* genes function as strong activators in *Setaria viridis* and *Arabidopsis thaliana* in protoplast assays and in transgenic plants. Overexpression of multiple endogenous genes (*FT, PAP1, WUS*) reached levels similar to the highly expressed housekeeping gene, *PP2A*, regardless of basal expression level. Further, these domains were effective in different PTA architectures, including the dCas9-SunTag, dCas9-Moontag, and TALE-SunTag systems. Lastly, we demonstrate the ability of these improved PTAs to map enhancer regions that promote gene expression in plants. This work demonstrates the effective and flexible nature of PTAs to activate target genes in plants, providing tools that can be used to improve agronomically relevant traits of interest.**

**dCas9 | Transcriptional activator | Plant engineering | SunTag**
**Correspondence:** ***voytas@umn.edu, smanski@umn.edu***

## Introduction

Controlling the expression of endogenous genes in plants is important for basic research on plant development. Genomes provide the blueprint for plant growth, survival, and reproduction. In order to survive a plant must also respond to surrounding environmental conditions such as temperature, light, or humidity. Gene expression is the process by which the genetic blueprint is put into action; activating genes at different times and places to develop, survive, and reproduce in the face of external stimuli (1, 2). Studies correlating differential gene expression with phenotype allow for predicting which genes are most important for proper development and responding to the environment (3, 4). The ability to control the spatio-temporal expression of these key genes driving phenotypic variation will be critical for trait and yield improvement of crop species in different environments across the world.

To this end, recent advances have led to the development of tools capable of activating gene expression in a user-controlled manner. Programmable Transcriptional Activators (PTAs) are fusion proteins comprised of a transcriptional activation domain (AD) fused to a DNA-binding domain that can be rationally engineered to recognize a DNA sequence of interest. The DNA-binding domain can be a zinc-finger (5), a transcriptional activator-like effector (TALE) DNA-binding domain (6), or a catalytically 'dead' Cas9 protein (dCas9) that is directed to a target sequence via a single guide RNA (sgRNA) (7). PTAs can drive the over- and/or ectopic-expression of endogenous genes when designed to bind to a promoter region (8, 9). The strength of overexpression is inversely correlated to a target gene's basal expression levels; PTAs targeting lowly-expressed genes can achieve higher fold-overexpression values than targeting highly-expressed genes (10). dCas9-based PTAs can also target many promoters in parallel by co-expressing multiple sgRNAs (11). Further, PTAs have been used to explore gene function and engineer barriers to sexual reproduction in insects (12).

The first dCas9-based PTA fused the VP64 AD the dCas9 DNA binding domain. VP64 is a tetrameric repeat of the VP16 protein derived from herpes simplex virus (13, 14). This dCas9-VP64 design has been used in plants to activate target genes (15). An additional improvement to PTAs came with the dCas9 translational fusion to the VPR AD (16). VPR is a translational fusion of three different ADs - VP64 from the Herpes Simplex Virus VP16 protein, RTA from the *Homo sapien RELA* transcription factor, and p65 from the Epstein-Barr virus *BRLF1* gene. While VPR drives strong gene expression in animal species, the same robust effectiveness seen in mammalian cells remains elusive in plants (7).

Several groups have developed improved dCas9-based PTAs for use in plant systems. The most common strategy is to increase the number of ADs recruited to the PTA. dCas9-TV is comprised of 6 copies of the TAL activation domain and two copies of the VP64 activation domain translationally fused to dCas9 (17). The SunTag system uses a non-covalent interaction between single-chain variable fragment antibodies (scFv) and a tandemly-repeated epitope tail (GCN4 motif) to recruit ADs to the dCas9 (18). Using 10 copies of GCN4 fused to dCas9, the SunTag system can activate target genes

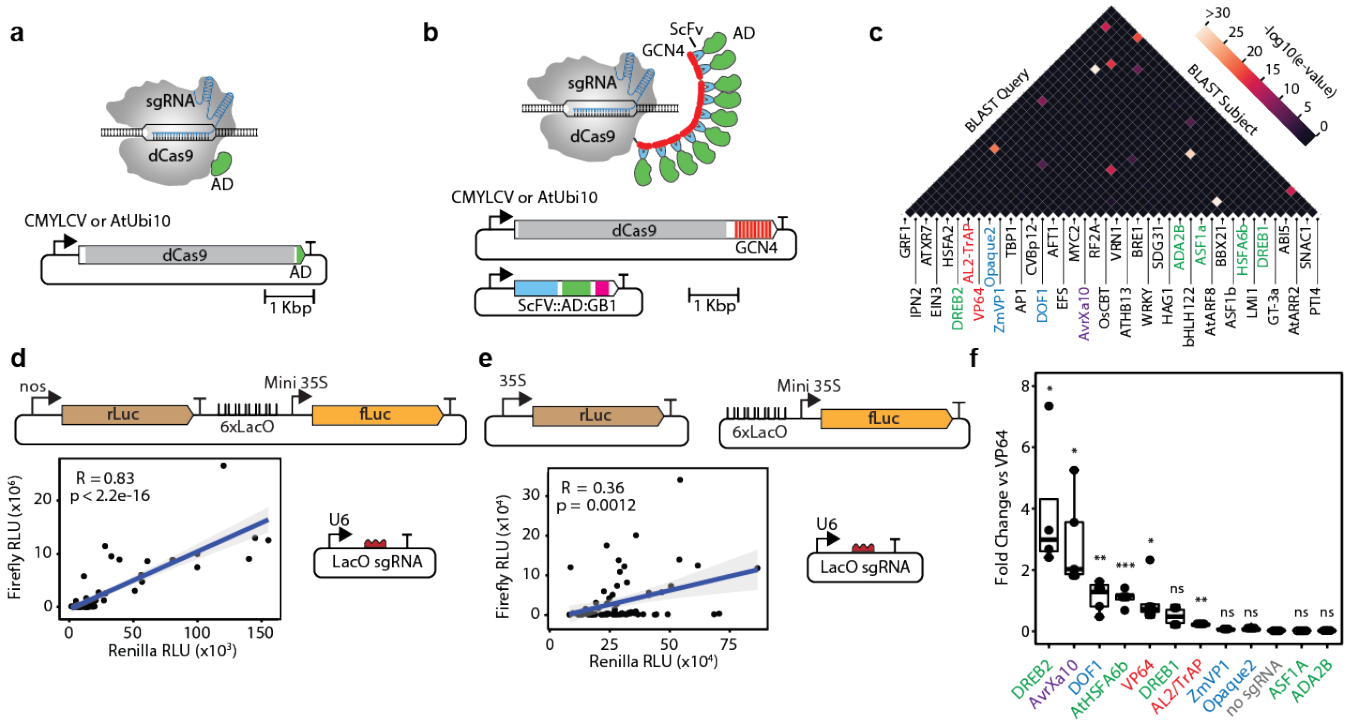



**Fig. 1. Screening a library of plant-derived activation domains with the dual luciferase protoplast assay.** Cartoon images of dCas9-based Programmable Transcriptional Activators (PTAs) including the direct fusion (a) and SunTag scaffolding system (b). AD, activation domain; ScFv, single-chain variable fragment; GB1, solubility tag; GCN4, epitope bound by ScFv. (c) Heatmap showing the similarity of tested activation domains. Black denotes no significant similarity. Color-coded ADs highlight the strongest from the analysis in (f), with the following relationship: green, dicot; blue, monocot; purple, bacterial; red, viral. (d) Genetic construct design for a single-plasmid dual luciferase assay (top) and observed correlation between expression of PTA-controlled firefly luciferase and constitutively expressed Renilla luciferase (bottom). Data is from *S. viridis* protoplasts. (e) Genetic construct design for two-plasmid dual luciferase assay (top) and corresponding low correlation between firefly and Renilla luciferase expression (bottom). Data is from *A. thaliana* protoplasts transfected with the plasmids shown, plus the two plasmids encoding components of the SunTag system in (b). (f) A rank order of the strongest ADs is show for *A. thaliana* protoplasts against the VP64 positive control. Parametric Welch's t-tests were performed comparing each AD with the negative control of scFv-VP64 delivered without a guide (no sgRNA). * corresponds to p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, ns is not significant (p > 0.05).

in *A. thaliana* to produce visible phenotypes (19). Subsequent iterations of the SunTag system have proven highly effective in activating target genes across monocot and dicot species (20). While these PTAs improved performance in plants, many still rely on the VP64 AD. Given the evolutionary distance between mammals and plants, we sought to improve PTA activity by screening ADs that evolved in plant systems. Here we report the performance of several plant-derived ADs for programmable gene expression in monocot and dicot plants.

## Results

**Building a library of putative plant-derived transcriptional activation domains.** We first determined a suitable experimental system for identifying strong ADs. Protoplast isolation and transformation pipelines yield consistent cell numbers and transformation efficiencies, allowing for direct comparison of groups within a given transformation when activating a reporter gene (21, 22).

We compared the ability of the direct-fusion dCas9-VP64 PTA architecture with the TAD-scaffolding PTA architecture SunTag (Figure 1a-b) to activate an endogenous target gene in *A. thaliana*. A tRNA-based multi-guide expression array (23) expressing four sgRNAs targeting a single core promoter for *WUSCHEL (WUS)* was transformed into *A. thaliana* protoplasts along with a dCas9 PTA, followed by RNA extraction and RT-qPCR. The dCas9-PTAs were expressed using the *Ubi10* promoter from *A. thaliana*.

At one day post-transformation, all PTAs increased the expression of *WUS* compared to a no-sgRNA control. While dCas9-TV outperformed dCas9-VP64, the SunTag-VP64 design outperformed both direct-fusion PTAs by two orders of magnitude (Supplemental Figure 1). Given these results, we reasoned that testing a library of putative ADs as fusions to scFv in the SunTag system would result in a dynamic range of activation for AD comparison.

Next a list of ADs was compiled to test in plants. We started with a literature search of plant-derived transcription factors with known DNA binding domains from diverse transcription factor families. We computationally removed native DNA-binding domains, as these regions are highly conserved and are expected to produce off-target effects if retained in PTA fusions. If an activation domain was not empirically determined, we selected motifs enriched in acidic and/or aromatic residues. Acidic and aromatic residue patches are often associated with transcriptional activation domains due to their propensity to form phase separation condensates upon recruitment to a core promoter (24). We also added plant-evolved activation domain sequences from plant pathogen effector proteins such as TALE proteins from *Xanthomonas*, along with plant-derived sequences from tran-

scription preinitiation complexes such as 14-3-3 scaffolding proteins. The terms "plant-derived" and "plant-evolved" are used here to refer to sequences identified in higher plants or plant pathogens, respectively.

The name, sequence, and citations for each domain tested can be found in Supplementary Table 1. We observe low amino acid sequence similarity across coding sequences from which ADs were derived, as illustrated by an all-by-all BLAST analysis (Figure 1c). This highlights the aim of this screen to survey a diverse array of plant-related ADs, versus optimizing on a well-performing family.

Putative AD sequences were codon optimized based on average codon usage tables for *A. thaliana* and *Oryza sativa* for reliable expression across a variety of plant backgrounds and synthesized as dsDNA fragments. We used Type IIS restriction enzyme cloning to assemble putative AD coding sequences into an scFv destination vector. ADs were expressed as C-terminal fusions to scFv, separated by a flexible Glycine-Serine (GS) linker containing Gly-Gly-Gly-Gly-Ser (25). AD coding sequences that failed multiple attempts to assemble in the scFv vector were removed from the study. ADs that failed multiple cloning attempts are denoted by an asterisk in Supplemental Table 1.

**A set of 5 plant-derived ADs show comparable activity to VP64 in a dual luciferase protoplast assay.** A dual luciferase reporter assay was designed to screen the library for domains capable of activating transcription in plant cells (26, 27). On the luciferase reporter plasmid, we designed a synthetic promoter comprising six copies of the *Lac* operator (*Lac*O) upstream of the minimal *35S CaMV* core promoter upstream of firefly luciferase (28). A sgRNA targeting the *LacO* sequence will recruit the dCas9-based transcriptional activators to each of the six *LacO* motifs, such that firefly luciferase expression reports activation domain strength. To provide an internal control of plasmid copy number, a Renilla luciferase controlled by a constitutive promoter was inserted in the same plasmid, upstream of and oriented in the same direction as the firefly luciferase (Figure 1d).

Alternative activation domains were expressed as scFv-AD fusions driven by a *35S* promoter, allowing for strong expression in both monocot and dicot protoplasts. The dCas9-24xGCN plasmid was expressed by an *AtUBI*10 promoter in *A. thaliana* or the viral *CmYLCV* promoter in *S. viridis* (Figure 1b).

The library was first tested in *S. viridis* protoplasts. We isolated and transformed 100,000 cells with 2 micrograms of DNA in 96 well plates. Each transformation received dCas9-24xGCN, the *LacO* sgRNA, the single-plasmid dual luciferase plasmid with *35S* driving Renilla luciferase, and a given scFv-AD. Each scFv-AD was independently transformed twice per protoplast isolation. Following two *S.viridis* protoplast isolations spanning four independent transformations, a correlation was seen between Renilla and Firefly luciferase activity and was particularly pronounced in some batches (Figure 1d, Supplementary Figure 2,3). This could be due to strong PTAs at the 6x Lac operator acting as enhancers for the constitutive nos promoter on the vector. A two-plasmid dual-luciferase construct diminished this correlation (Figure 1e). With this experimental set-up, we compared the ability of alternative ADs to induce expression of Firefly luciferase. Across multiple experiments in different plant species, DREB2, DREB1, DOF1, AvrXa10, and HSFA6b showed good activity in the dual luciferase assay (Figure 1f, Supplementary Figures 4-6). Activity required the full SunTag system, as the ADs alone did not drive substantial luciferase activity (Supplementary Figure 7). Activity of ADs was bench marked against scFv-VP64 delivered both with and without a guide RNA for positive and negative controls respectively ('VP64' and 'no sgRNA', Fig 1f).

**DREB2, DOF1, and AvrXa10 plant-derived activation domains outperform VP64 across endogenous loci.** We next tested the efficiency of HSFA6b, AvrXa10, DOF1, DREB1, DREB2, and VP64 at endogenous loci. Single guide RNAs expressed from the *A. thaliana* U6-26 promoter were constructed to target the core promoters of *FLOWERING LOCUS T* (*FT*), *CLAVATA3* (*Clv3*), *PRODUCTION OF ANTHOCYANIN PIGMENT 1* (*PAP1*), *WUSCHEL* (*WUS*), and *LEAFY COTYLEDON1* (*LEC1*) for activation (Figure 2a-b, Supplemental Figure 8). These target genes were selected either because they are expected to drive a clear phenotype upon overexpression (*FT*), or because they are putative targets for the development of genetic biocontainment via Engineered Genetic Incompatibility, which is a longer term goal of our group (29). All sgRNAs used in this study are previously published (15, 30), except for the *WUS* guide, which was developed for this study. Approximately 500,000 *A. thaliana* protoplasts were isolated and transformed with the SunTag PTA and *A. thaliana* U6-sgRNA plasmids, with either two or three transoformation replicates per activation domain per target gene. After 24 hours, RNA was isolated and RT-qPCR was performed to quantify gene activation. We compared the activity of each AD against a no AD control for every sgRNA tested, in which the dCas9 and sgRNA were transformed without an scFv-AD. For each sample, the target gene and a control housekeeping gene (*PP2A*) were amplified. We report expression relative to the housekeeping gene *PP2A* in Figure 2b for *WUS*, *CLV3*, *FT*, and *PAP1*. By graphing all data points relative to the housekeeping gene, it is easier to observe both the fold-change in gene expression (height of the bar), and the absolute change in gene expression (location relative to the y-axis). This is important, as the fold-change in expression that can be achieved by PTAs is inversely correlated with the basal expression level. Genes already expressed at high levels cannot be overexpressed to the same relative degree as genes with a low basal expression level (10).

The sgRNAs used in these experiments bind the core promoter regions, typically located 0-300 base pairs upstream of an annotated transcription start site (TSS) for a given gene except for the *PAP1* sgRNA used in this study which binds downstream of the annotated TSS. A clear pattern emerged across all sgRNAs tested, with AvrXa10, DOF1, and DREB2 consistently producing greater gene activation than VP64, HSFA6b, and DREB1 (Figure 2b, Supplemental Figure 8).

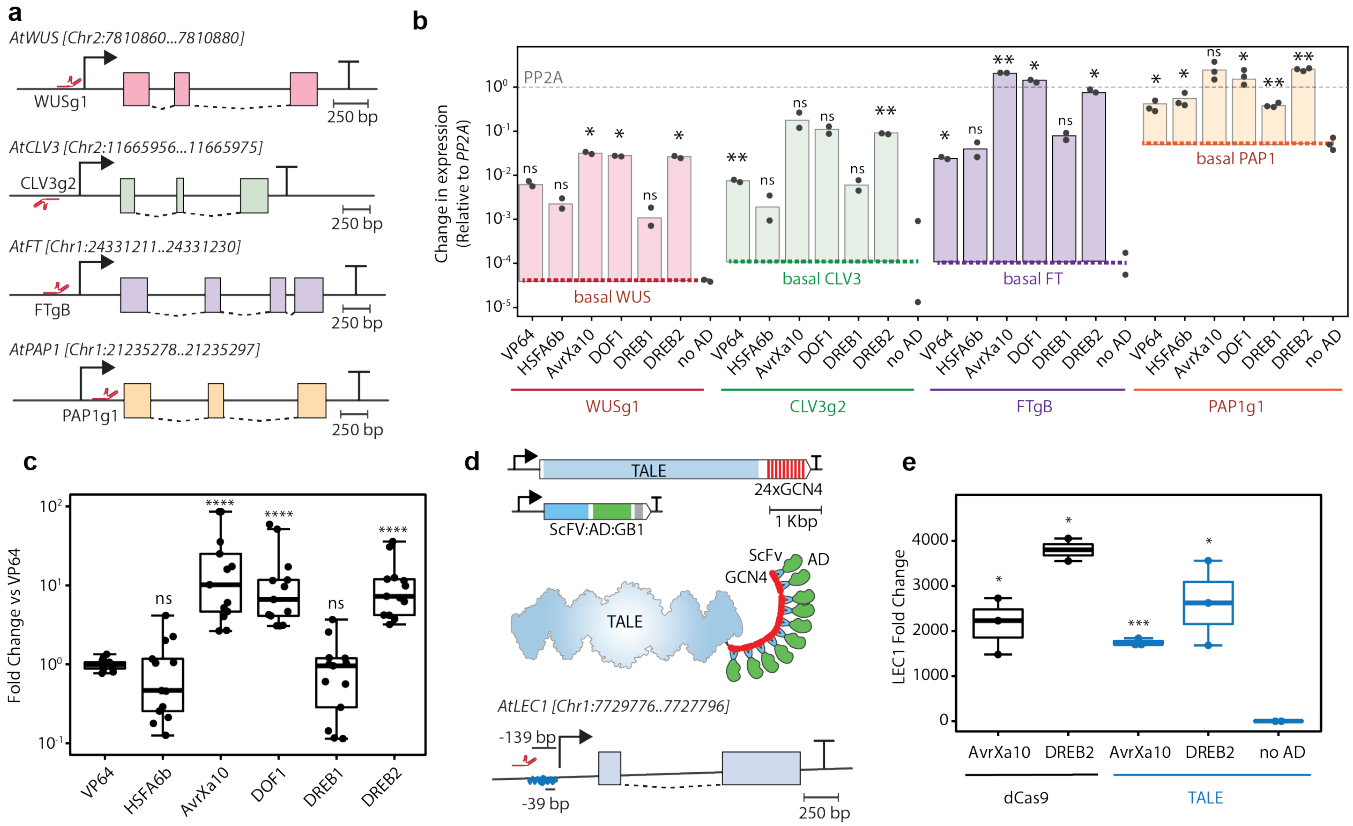



**Fig. 2. Endogenous gene activation in *A. thaliana* protoplasts.** (a) Guide RNA design for a given promoter is shown for *WUS*, *CLV3*, *FT*, and *PAP1* loci. The precise location of sgRNA binding is given in brackets above each gene diagram. (b) Gene activation data for each corresponding gene is shown relative to the housekeeping gene *PP2A*. Basal level of gene expression, corresponding to values measured from no AD controls is unique for each target and is highlighted with colored, dashed line. Statistical significance compared to no AD control was determined conservatively using a Welch's t-test due to low replicate numbers. See symbol key for p-values at the end of this legend. (c) Aggregate gene activation across all genes tested, where each dot represents an independent protoplast transformation. Parametric Welch's t-tests were performed comparing each AD with the negative control NoAD. (d) The TALE DNA binding domain was fused to the GCN4 epitope tail to comprise a TALE-SunTag system (top). The TALE DNA binding domain was engineered to target the LEC1 promoter, compared to dCas9 binding the same promoter (bottom). The abbreviated names in this panel are described in the legend of Figure 1b.(e) The *LEC1* target gene in *A. thaliana* was targeted for gene activation comparing dCas9-SunTag (black) with TALE-SunTag (blue). * corresponds to p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, and **** corresponds to p<= 0.0001

As expected, the maximum fold-overexpression was correlated with basal expression level. For example, *FT* saw a much greater fold-overexpression compared to no-AD control (>10,000) than did *PAP1* ( 50). However this is due to *FT* starting at much lower basal expression levels, and maximum absolute expression attained is comparable. *WUS*, on the other hand, could only be overexpressed to 10% the level of *PP2A*.

We compared each AD's gene activation across the panel of endogenous genes to VP64. After setting VP64 activation values for each gene tested equal to 1, we combined data sets for all sgRNAs across four target genes (n=13). While the rank order changed depending on target gene, AvrXa10, DOF1, and DREB2 produced statistically significant gene activation across all target genes and guide RNAs that were tested when compared to VP64 (Figure 2c).

Because these activation domains described come from plant backgrounds, it is possible that transcriptional networks unrelated to the CRISPR-targeted promoter could be activated. To confirm the specificity of targeting endogenous genes in *A. thaliana* protoplasts, the transcriptional change of downstream target genes for each plant-derived AD were quantified in *A. thaliana* protoplasts after treatment with the dCas9 activation complex targeting the FT locus, and an unrelated target gene. These experiments were performed in *A. thaliana* protoplasts because HSFA6b, DREB1, and DREB2 are sequences taken from endogenous Arabidopsis genes. Downstream target genes for HSFA6b, DREB1, and DREB2 have been identified in the literature (31–36). Primers were designed to quantify the expression of these downstream target genes with RT-qPCR (Supplemental Table 5). The expression of a downstream target gene for the negative control (no AD) was compared to each AD treatment. No statistically significant off-target activation was detected for HSFA6b, DREB1, or DREB2 (Supplemental Figure 9). These results suggest specificity of the plant-derived ADs for the sgRNA-defined target site when fused to dCas9 in the scaffolding SunTag system.

We next tested the portability of these ADs between different programmable DNA binding domains. The 24xGCN4 epitope tail from the SunTag system was cloned as a C-terminal fusion to a TALE programmable DNA binding domain (Figure 2d). A TALE DNA binding domain is engineered by assembling a series of 20 TALE repeats, each of which contain a unique Repeat Variable Diresidue (RVDs) (15). Each RVD within a TALE repeat specifies a single nucleotide for bind-

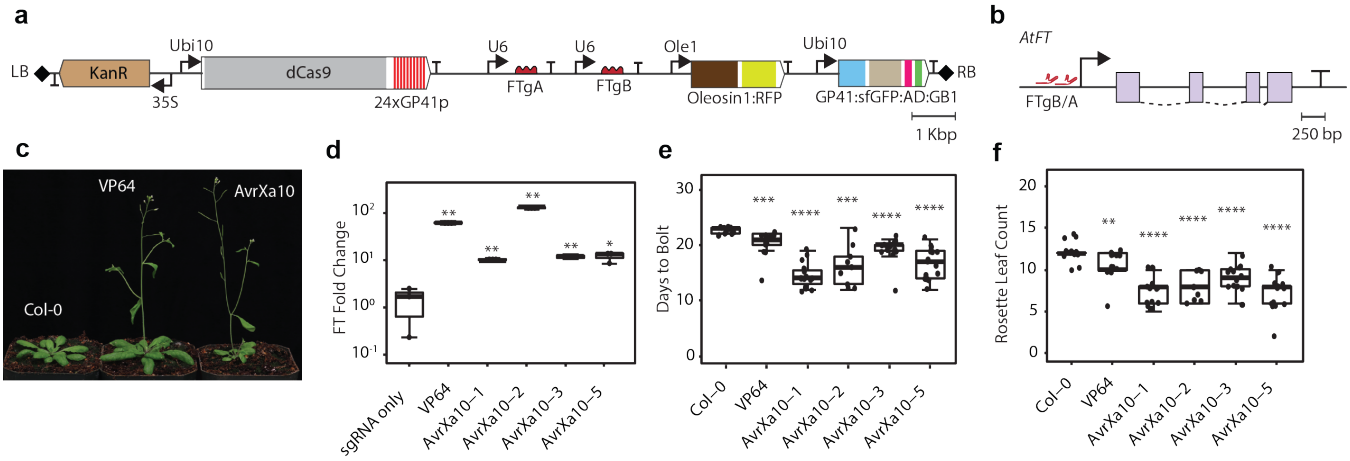



**Fig. 3. FT gene activation in transgenic plants** (a) Genetic construct design for binary vectors floral dipped to generate transgenic *A. thaliana* plants. The abbreviated names in this panel are described in the legend of Figure 1b, in addition to the following: RFP, red fluorescent protein; GP41, camelid nanobody; GP41p, peptide epitope for camelid nanobody; sfGFP, super-folder green fluorescent protein; LB and RB, left and right border, respectively. (b) Illustration of FTgA [Chr1:24331368..24331387] and FTgB [Chr1:24331212..24331230] design relative to the *FT* TSS. (c) Representative images of T2 transgenic plants for Col-0, VP64, and AvrXa10-1 transgenic lines. (d) Molecular quantification (RT-qPCR) of *FT* gene activation from 10d old T2 seedlings. Parametric Welch's t-tests were performed on triplicate technical replicates from six pooled plants from each genotype compared to the sgRNA only negative control. (e) Bolting time quantification in T2 transgenic plants. (f) Rosette leaf quantification one day after bolting in T2 transgenic plants. Parametric Welch's t-tests were performed in panels (e) and (f) comparing each line with the negative control Col-0. * corresponds to p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, and **** corresponds to p<= 0.0001

ing, such that a specific 20 base pair genomic target sequence is specified by 20 RVDs. We built a TALE repeat targeting a 20 bp region 39 bp upstream of the transcription start site for the *A. thaliana* gene *LEC1*. In addition, we cloned a sgRNA targeting a 20 bp region 139 bp upstream of the transcription start site for the same *LEC1* target gene (Figure 2d). We then tested the ability of the AvrXa10 and DREB2 ADs to activate the *LEC1* target gene in conjunction with either dCas9 or TALE DNA binding domains in *A. thaliana*. As expected the dCas9 SunTag system using AvrXa10 and DREB2 domains resulted in strong *LEC1* gene expression with 2100 or 3800 fold greater activation than the No AD negative control (Figure 2e). The TALE SunTag system using AvrXa10 and DREB2 domains also resulted in strong *LEC1* gene expression of 1700 and 2600 fold greater activation than the No AD negative control. These results demonstrate the flexibility of the AvrXa10 and DREB2 ADs to drive strong expression of target genes using different programmable DNA binding domains.

**AvrXa10 activation domain promotes early flowing in transgenic plants.** Next, we tested whether our plant-derived ADs would produce a phenotype upon gene activation in stable transgenic plants. Binary vectors were generated expressing the plant-derived ADs using a new PTA architecture named the MoonTag system (37). The MoonTag system utilizes a similar mechanism as SunTag, in which an epitope-nanobody interaction recruits the AD to the dCas9 molecule (38). In our Moontag system there are 24 copies of the GP41p epitope fused to dCas9, while AvrXa10, DOF1, and DREB2 were fused to the GP41 2H10 nanobody. Two sgRNAs targeting the *FT* core promoter in *A. thaliana* were expressed from *A. thaliana* U6 promoters to drive *FT* gene activation (Figure 3a-b). Plants were floral dipped and T1 transgenic lines were identified using the pFAST *Oleosin1* embryo oil body fluorescent reporter (39, 40). For comparison, we used a validated MoonTag-VP64 transgenic line as the positive control (37) and wild type Columbia-0 (Col-0) or sgRNA-only plants as the negative controls in these experiments (Figure 3c). T1 transgenic seeds were germinated on 1/2 MS agar plates and 1-5 seedlings from AvrXa10, DOF1, DREB2, and Col-0 genotypes were transplanted to soil. Four of 5 AvrXa10 plants displayed an early flowering phenotype, compared to a single early flowering plant of each DOF1 and DREB2 genotypes and none from the Col-0 controls (Supplemental Figure 10). We allowed these plants to set seed and used transgenic plants (without determining hetero- or homozygosity) from the T2 generation for molecular and phenotypic quantification of flowering time.

Sets of 20 seedlings per line were planted on selection media and allowed to germinate. Ten days after germination, seedlings were pooled into groups of six and RNA was extracted from the pooled group. We measured *FT* gene expression in the transgenic lines compared to a control line expressing only sgRNAs against a different target gene. The line used as a negative control expressed only sgRNAs, and lacked a Cas9 altogether. A binary vector map for this negative control line can be found in Supplemental Table S4. We observed strong gene activation in the positive control VP64 line, along with one of the AvrXa10 transgenic lines (Figure 3d).

We also quantified the *FT* overexpression phenotype in the T2 population by selecting 18 RFP-positive seedlings and planting them directly in soil for each transgenic line. Following germination, we waited for the seedlings to produce their first set of true leaves and used this time point as t=0. Two commonly used metrics of flowering time are bolting time and rosette leaf number at time of flowering (41, 42). We noted the day at which bolting was observed and calculated the difference between this day and t=0 to quantify bolt-

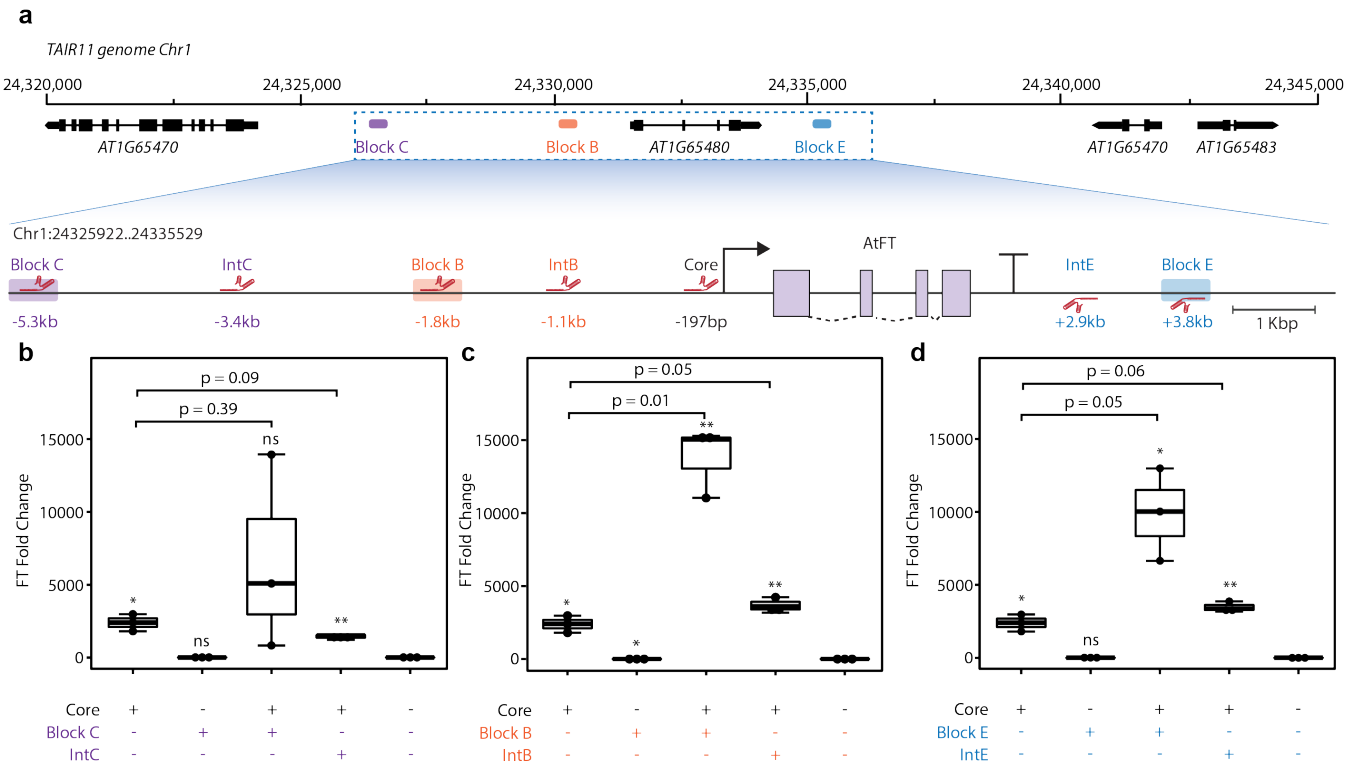



**Fig. 4. PTA-mediated enhancer activation at the *A. thaliana FT* locus using DREB2.** (a) Schematic illustration of sgRNA design in relation to the region of *A. thaliana* chromosome 1 surrounding the *FT* gene. Top depicts entire 25kb locus and bottom shows more detail of enhancer location and sgRNA design. sgRNAs drawn above or below the line denote targeting the sense or antisense strand, respectively. Distances labeled are relative to the *FT* TSS. (b-d) *FT* gene activation, measured by RT-qPCR when targeting Block C (b), Block B (c) or Block E (d) with different combinations of sgRNAs (x-axis). Data plotted in b-d are all from a single experiment but have been organized into three separate subplots (with the same data from the core promoter and negative control shown in each subplot) to facilitate interpretation of the results. Parametric Welch's t-tests were performed comparing each sgRNA combination with the negative control where no sgRNA is delivered. P-values for these analyses are denoted symbolically, where * corresponds to p <= 0.05, ** corresponds to p <= 0.01, and ns means 'not significant' (p > 0.05). Brackets with specific p-values listed are separate parametric Welch's t-tests comparing the groups below the bracket ends.

ing time (Figure 3e). We then counted rosette leaves one day post bolting for each seedling as an additional quantification of the *FT* phenotype (Figure 3f). We observed *FT* activation phenotypes across multiple AvrXa10 transgenic lines, with early bolting and reduced rosette leaf number in comparison to wild type plants. We also noted reduced rosette leaf size in several of the AvrXa10 plants (Figure 3c) compared to Col-0 and VP64 plants. We did not observe strong early flowering phenotypes for DOF1 and DREB2, likely due to a small number of lines recovered. Taken together we observed the plant-derived AvrXa10 AD functions and drives the strongest early flowering phenotype in *A. thaliana*.

**PTA-Mediated Gene Enhancer Activation.** Finally, we demonstrated the ability of our improved plant PTAs to map enhancers. Co-targeting PTAs to both a core promoter and enhancer region has previously been shown to boost observed expression levels in human cells (43). To illustrate this concept in plants, we utilized the *FT* gene activation platform. Of the two sgRNAs tested, FTgB was capable of driving stronger gene activation than FTgA at the core promoter in our protoplast assays (Figure 2b, Supplemental Figure 8). According to previous literature, Block B (-1.8kb), Block C (-5.3kb), and Block E (+3.8kb) are three known enhancers of *FT* located either upstream or downstream of the TSS (44). We designed sgRNAs that target the centers of each 350-400 bp enhancer, along with the *FT* core promoter guide FTgB previously tested. In addition, we designed guides that target sequences located between the annotated enhancer regions and the core promoter (Figure 4a).

For each of Blocks B, C, and E, targeting the enhancer alone did not result in overexpression above no-sgRNA controls (Figure 4b-d). Targeting both the core promoter and the enhancer led to a significant increase in expression level for Blocks B and E. The Block C enhancer-targeting sgRNA produced greater expression levels in two of three replicates, but a large amount of variance in this experiment prevented it from reaching statistical significance. When targeting enhancers plus core promoters, mean fold overexpression compared to no-sgRNA controls ranged from 6,000-fold (Block C) to 14,000 fold (Block B). This improvement over targeting the core promoter alone (2,400-fold overexpression compared to control) could not be attributed to having a second sgRNA, as the sgRNAs targeting sequences outside of the enhancer regions did not significantly boost expression levels (Figure 4b-d).

Finally, we tested the specificity of PTA-mediated enhancer activation by quantifying the expression of the FAS1 gene (AT1G65470) located upstream of the FT gene (AT1G65480). It is possible that the Block B, C, and E regions of open chromatin serving as enhancers for the FT locus may also act as enhancers for nearby genes when targeted

with dCas9-based activators. To test this, RT-qPCR primers were designed for the FAS1 transcript, and its expression was measured across all enhancer treatments. The enhancer treatments surveyed for FAS1 expression in the study were targeting of the FT core promoter alone, co-targeting of the FT core promoter + Block B, C, and E, and finally the FT core promoter sgRNA + dCas9-24xGCN4 lacking any AD. Little activation was observed at the FAS1 gene when targeting the FT core promoter along with guides targeting the open chromatin enhancer regions of Block B, C, and E when compared to the No AD negative control (Supplemental Figure 11).

## Discussion

We set out to identify plant-derived ADs that function as well or better than VP64 in plant systems. Through a library approach we were able to clone and test putative ADs in protoplast dual luciferase assays to determine which sequences showed activity comparable to VP64. We identified a core set of 5 activation domains comprised of AvrXa10, HSFA6b, DREB1, DREB2, and DOF1 that showed strong activity in both the monocot model *S. viridis* and the dicot model *A. thaliana* dual luciferase protoplast assays. The high throughput nature of protoplast transformation serves as a valuable tool for testing large numbers of genetic parts (22).

While performing these assays, we observed what appeared to be local off-target activation of our dual luciferase reporter. DREB2, DREB1, and HSFA6b consistently produced the largest Renilla luciferase values, even though they were targeted to the firefly luciferase promoter. This resulted in a linear correlation between firefly luciferase and Renilla luciferase values (Figure 1d). Because this correlation was reduced by splitting the reporters into two plasmids (Figure 1e), we conclude expression from the upstream Renilla promoter is the result of trans-activation from the PTA. In this case, the promoter driving the Renilla gene was located approximately 3kb upstream of the sgRNA binding site. Interestingly, targeting a PTA to within 3kb of the endogenous *FT* locus did not cause *FT* overexpression unless at least one sgRNA bound to the core promoter region. It is possible that the proximity effect seen in our single-plasmid dual luciferase assay is dependent on transcriptional activity from the targeted promoter. Nucleosome formation on transiently-transfected, episomal DNA is known to be altered compared to chromosomal DNA(45–47), and this could allow for interactions between reporter construct promoters in our protoplast assay. This phenomena may vary depending on the locus, the number of sgRNAs, or even the AD that is directed to a target site. In genome-wide screens, off-target activation events are rarely seen (48).

We demonstrate endogenous gene activation in which AvrXa10, DOF1, and DREB2 consistently produce the greatest activation across all genes tested. It is interesting to note that the relative strength of ADs did vary both across species and across target loci. For example the AvrXa10 domain shows the largest range of activation among the three strongest ADs across target genes in *A. thaliana* protoplasts. Given that AvrXa10 is derived from a plant pathogen, perhaps there are varying degrees of host response to the expression of this domain across plant species and/or target gene.

We demonstrate the ability of AvrXa10, DOF1, and DREB2 plant-derived activation domains to drive gene activation of the *FT* locus in transgenic plants. T1 plants showed an early flowering phenotype, and this phenotype was maintained in the T2 generation where early flowering was quantified. This work illustrates the feasibility of moving from a protoplast-based discovery platform to validating the tools in stable transgenic lines. Changes in leaf size were also observed, particularly in the AvrXa10 expressing transgenic lines. Decreases in leaf size and number have been observed by others, deemed the "extra-early flowering" phenotype (49).

Finally, we showcase the ability of the DREB2 activation domain to function in a potential gene enhancer discovery pipeline. The *FT* locus contains three known enhancers of gene expression Block B, C, and E. These enhancers are located vast distances from the core promoter, and fail to produce any observable activation when targeted with dCas9 PTAs. While we do observe slight boosts in *FT* activation when co-targeting both the core promoter and an intervening region between enhancers, these boosts were rarely statistically significant and never as dramatic as the overexpression seen when co-targeting an enhancer. Noisy gene activation at distances outside of the core promoter (50-400bp upstream TSS) has been reported before (48), and may be a natural consequence of transcription pre-initiation dynamics.

Determining the location and activity of enhancer sequences can be challenging. Luciferase assays in protoplasts have traditionally been used to confirm enhancer activity by cloning an enhancer upstream of the firefly luciferase reporter and checking for activity (50). PTA-mediated enhancer activation retains the original genomic architecture and chromatin state of the enhancer-promoter pair, while traditional luciferase assays deploy enhancer sequences in non-native genomic architectures on plasmids which lack nucleosomes. Given our understanding of enhancers functioning through a DNA looping mechanism, retaining this genetic architecture may be important.

Once a putative enhancer is identified it can also be difficult to determine the promoter, or promoters, it regulates. Recent studies have utilized methods such as STARR-seq to confirm the transcriptional regulatory capacity of distal accessible chromatin elements in protoplasts (51). We anticipate PTA-mediated enhancer activation can be similarly used for testing distal accessible chromatin activity in plants. One could design a series of sgRNAs to target intergenic regions in tandem with a sgRNA targeting the core promoter for a gene of interest. Conversely, an accessible chromatin region of interest can be targeted along with the core promoters of nearby genes to identify which transcripts are significantly activated by the enhancer. When conducted in a high-throughput transient platform such as protoplasts, this pipeline could reveal novel enhancer-promoter interactions. This type of enhancer mapping can further expand the definition of gene by including enhancer regions in addition to the core promoter, untranslated regions, exons, introns, and

terminator comprising a transcriptional block. Determining the general reliability of this method for mapping enhancers will require follow-up experiments in different genes.

This study has limitations that should be addressed with future work. The PTAs we describe here work great in research and development environments, but they have components derived from human or plant pathogens that would trigger special regulatory approval considerations if used in transgenic plants intended for field release. Replacing the VP64 domain with plant-evolved ADs mitigates that in part, however the dCas9, GB1, and AvrXa10 components are similarly derived from human or plant pathogens (52).

In addition, it has been noted that scFv proteins have difficulty forming disulfide bonds within the reducing environment of the cytosol, leading to poor solubility (53). For these reasons, replacing GB1 with alternate solubility tags could be addressed to further improve the performance of these PTAs in plants. For example, a naturally occurring solubility tag from spider silk shows promise as an alternative to the Streptococcal GB1 sequence (54). Improving the solubility of the PTA complex without compromising its transcriptional activation potency should be of significant interest moving forward.

With this work we showcase potent PTAs in plants capable of driving strong expression of target genes. We present three transcriptional activation domains from plant backgrounds that outperform VP64 in both reporter gene assays and targeting endogenous loci. We further showcase the portability of both DREB2 and DOF1 in diverse plants. They provide stronger expression than VP64 for target genes in *A. thaliana*, *S. viridis*, and *Z. mays*. Finally, we show that the DREB2 domain is equally as potent when fused to an alternative TALE programmable DNA binding domain. Given these results, we anticipate that these domains will be used by plant scientists to effectively activate transcription of target genes within diverse plant species.

**ACKNOWLEDGEMENTS**

MHZ and DV are supported by Department of Energy Grant DE-SC0018277. MHZ, JACM, AS, and MJS are supported by the USDA grant 2018-33522-28747 and are supported by the Advanced Plant Technologies program, DARPA Award HR001118C0146. MHZ and AS are supported by an NIH NIGMS Biotechnology Training grant NIHT32GM008347. SCH was supported by a Doctoral Dissertation Fellowship from the University of Minnesota Graduate School.

**AUTHOR CONTRIBUTIONS**

MHZ, DV, and MJS conceived this study. MHZ designed and built constructs, planned experiments, collected and analyzed data. ACM, AS, and MHZ performed protoplast isolation, transformation, RNA extraction, and RT-qPCR analysis. ACM built constructs and carried out floral dip. ACM provided VP64 transgenic line for *FT* analysis. MHZ, MJS, and DV wrote the manuscript.

**COMPETING FINANCIAL INTERESTS**

MHZ, MJS, and DV are listed as inventors on UMN-owned patents related to this work.

# Bibliography

1. Anna V. Klepikova, Artem S. Kasianov, Evgeny S. Gerasimov, Maria D. Logacheva, and Aleksey A. Penin. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *The Plant Journal*, 88(6):1058–1070, December 2016. ISSN 1365-313X. doi: 10.1111/TPJ.13312. Publisher: John Wiley & Sons, Ltd.
2. Steffen Knauer, Marie Javelle, Lin Li, Xianran Li, Xiaoli Ma, Kokulapalan Wimalanathan, Sunita Kumari, Robyn Johnston, Samuel Leiboff, Robert Meeley, Patrick S. Schnable, Doreen Ware, Carolyn Lawrence-Dill, Jianming Yu, Gary J. Muehlbauer, Michael J. Scanlon, and Marja C.P. Timmermans. A high-resolution gene expression atlas links dedicated meristem genes to key architectural traits. *Genome Research*, 29(12):1962–1973, December 2019. ISSN 1088-9051. doi: 10.1101/GR.250878.119. Publisher: Cold Spring Harbor Laboratory Press.
3. Yunlong Pang, Chunxia Liu, Danfeng Wang, Paul St. Amand, Amy Bernardo, Wenhui Li, Fang He, Linzhi Li, Liming Wang, Xiufang Yuan, Lei Dong, Yu Su, Huirui Zhang, Meng Zhao, Yunlong Liang, Hongze Jia, Xitong Shen, Yue Lu, Hongming Jiang, Yuye Wu, Anfei Li, Honggang Wang, Lingrang Kong, Guihua Bai, and Shubing Liu. High-Resolution Genome-wide Association Study Identifies Genomic Regions and Candidate Genes for Important Agronomic Traits in Wheat. *Molecular Plant*, 13(9):1311–1327, September 2020. ISSN 1674-2052. doi: 10.1016/J.MOLP.2020.07.008. Publisher: Cell Press.
4. Fei He, Wei Wang, William B. Rutter, Katherine W. Jordan, Jie Ren, Ellie Taagen, Noah DeWitt, Deepmala Sehgal, Sivakumar Sukumaran, Susanne Dreisigacker, Matthew Reynolds, Jyotirmoy Halder, Sunish Kumar Sehgal, Shuyu Liu, Jianli Chen, Allan Fritz, Jason Cook, Gina Brown-Guedira, Mike Pumphrey, Arron Carter, Mark Sorrells, Jorge Dubcovsky, Matthew J. Hayden, Alina Akhunova, Peter L. Morrell, Les Szabo, Matthew Rouse, and Eduard Akhunov. Genomic variants affecting homoeologous gene expression dosage contribute to agronomic trait variation in allopolyploid wheat. *Nature Communications 2022 13:1*, 13(1):1–15, February 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-28453-y. Publisher: Nature Publishing Group.
5. Ahmad S. Khalil, Timothy K. Lu, Caleb J. Bashor, Cherie L. Ramirez, Nora C. Pyenson, J. Keith Joung, and James J. Collins. A Synthetic Biology Framework for Programming Eukaryotic Transcription Functions. *Cell*, 150(3):647–658, August 2012. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2012.05.045. Publisher: Elsevier.
6. Tina Lebar and Roman Jerala. Benchmarking of TALE- and CRISPR/dCas9-Based Transcriptional Regulators in Mammalian Cells for the Construction of Synthetic Genetic Circuits. *ACS Synthetic Biology*, 5(10):1050–1058, October 2016. doi: 10.1021/acssynbio.5b00259. Publisher: American Chemical Society.
7. J. Armando Casas-Mollano, Matthew H. Zinselmeier, Samuel E. Erickson, and Michael J. Smanski. CRISPR-Cas Activators for Engineering Gene Expression in Higher Eukaryotes. *CRISPR Journal*, 3(5):350–364, October 2020. ISSN 25731602. doi: 10.1089/CRISPR.2020.0064/ASSET/IMAGES/LARGE/CRISPR.2020.0064_FIGURE2.JPEG. Publisher: Mary Ann Liebert Inc.
8. Pablo Perez-Pinera, D. Dewran Kocak, Christopher M. Vockley, Andrew F. Adler, Ami M. Kabadi, Lauren R. Polstein, Pratiksha I. Thakore, Katherine A. Glass, David G. Ousterout, Kam W. Leong, Farshid Guilak, Gregory E. Crawford, Timothy E. Reddy, and Charles A. Gersbach. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nature Methods 2013 10:10*, 10(10):973–976, July 2013. ISSN 1548-7105. doi: 10.1038/nmeth.2600. Publisher: Nature Publishing Group.
9. Morgan L. Maeder, Samantha J. Linder, Vincent M. Cascio, Yanfang Fu, Quan H. Ho, and J. Keith Joung. CRISPR RNA–guided activation of endogenous human genes. *Nature Methods 2013 10:10*, 10(10):977–979, July 2013. ISSN 1548-7105. doi: 10.1038/nmeth.2598. Publisher: Nature Publishing Group.
10. Anna M. Chiarella, Kyle V. Butler, Berkley E. Gryder, Dongbo Lu, Tiffany A. Wang, Xufen Yu, Silvia Pomella, Javed Khan, Jian Jin, and Nathaniel A. Hathaway. Dose-dependent activation of gene expression is achieved using CRISPR and small molecules that recruit endogenous chromatin machinery. *Nature biotechnology*, 38(1):50, January 2020. ISSN 15461696. doi: 10.1038/S41587-019-0296-7. Publisher: NIH Public Access.
11. Haibo Zhou, Junlai Liu, Changyang Zhou, Ni Gao, Zhiping Rao, He Li, Xinde Hu, Changlin Li, Xuan Yao, Xiaowen Shen, Yidi Sun, Yu Wei, Fei Liu, Wenqin Ying, Junming Zhang, Cheng Tang, Xu Zhang, Huatai Xu, Linyu Shi, Leping Cheng, Pengyu Huang, and Hui Yang. In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR–dCas9-activator transgenic mice. *Nature Neuroscience 2018 21:3*, 21(3):440–446, January 2018. ISSN 1546-1726. doi: 10.1038/s41593-017-0060-6. Publisher: Nature Publishing Group.
12. Ambuj Upadhyay, Nathan R. Feltman, Adam Sychla, Anna Janzen, Siba R. Das, Maciej Maselko, and Michael Smanski. Genetically engineered insects with sex-selection and genetic incompatibility enable population suppression. *eLife*, 11, February 2022. ISSN 2050084X. doi: 10.7554/ELIFE.71230. Publisher: NLM (Medline).
13. Ivan Sadowski, Jun Ma, Steve Triezenberg, and Mark Ptashne. GAL4-VP16 is an unusually potent transcriptional activator. *Nature 1988 335:6190*, 335(6190):563–564, 1988. ISSN 1476-4687. doi: 10.1038/335563a0. Publisher: Nature Publishing Group.
14. Roger R. Beerli, David J. Segal, Birgit Dreier, and Carlos F. Barbas. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proceedings of the National Academy of Sciences of the United States of America*, 95(25):14628–14633, December 1998. ISSN 00278424. doi: 10.1073/PNAS.95.25.14628.
15. Levi G. Lowder, Jianping Zhou, Yingxiao Zhang, Aimee Malzahn, Zhaohui Zhong, Tzung Fu Hsieh, Daniel F. Voytas, Yong Zhang, and Yiping Qi. Robust Transcriptional Activation in Plants Using Multiplexed CRISPR-Act2.0 and mTALE-Act Systems. *Molecular Plant*, 11(2):245–256, February 2018. ISSN 1674-2052. doi: 10.1016/J.MOLP.2017.11.010. Publisher: Cell Press.
16. Alejandro Chavez, Jonathan Scheiman, Suhani Vora, Benjamin W Pruitt, Marcelle Tuttle, Eswar P R Iyer, Shuailiang Lin, Samira Kiani, Christopher D Guzman, Daniel J Wiegand, Dmitry Ter-Ovanesyan, Jonathan L Braff, Noah Davidsohn, Benjamin E Housden, Norbert Perrimon, Ron Weiss, John Aach, James J Collins, and George M Church. Highly efficient Cas9-mediated transcriptional programming. *Nature Methods 2015 12:4*, 12(4):326–328, March 2015. ISSN 1548-7105. doi: 10.1038/nmeth.3312. Publisher: Nature Publishing Group.
17. Zhenxiang Li, Dandan Zhang, Xiangyu Xiong, Bingyu Yan, Wei Xie, Jen Sheen, and Jian Feng Li. A potent Cas9-derived gene activator for plant and mammalian cells. *Nature Plants 2017 3:12*, 3(12):930–936, November 2017. ISSN 2055-0278. doi: 10.1038/s41477-017-0046-0. Publisher: Nature Publishing Group.
18. Marvin E. Tanenbaum, Luke A. Gilbert, Lei S. Qi, Jonathan S. Weissman, and Ronald D. Vale. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*, 159(3):635–646, October 2014. ISSN 0092-8674. doi: 10.1016/J.CELL.2014.09.039. Publisher: Cell Press.

19. Ashot Papikian, Wanlu Liu, Javier Gallego-Bartolomé, and Steven E. Jacobsen. Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. *Nature Communications 2019 10:1*, 10(1):1–11, February 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-08736-7. Publisher: Nature Publishing Group.
20. Changtian Pan, Xincheng Wu, Kasey Markel, Aimee A. Malzahn, Neil Kundagrami, Simon Sretenovic, Yingxiao Zhang, Yanhao Cheng, Patrick M. Shih, and Yiping Qi. CRISPR–Act3.0 for highly efficient multiplexed gene activation in plants. *Nature Plants 2021 7:7*, 7(7):942–953, June 2021. ISSN 2055-0278. doi: 10.1038/s41477-021-00953-7. Publisher: Nature Publishing Group.
21. Sang Dong Yoo, Young Hee Cho, and Jen Sheen. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols 2007 2:7*, 2(7):1565–1572, June 2007. ISSN 1750-2799. doi: 10.1038/nprot.2007.199. Publisher: Nature Publishing Group.
22. Adam Sychla, Juan Armando Casas-Mollano, Matthew H. Zinselmeier, and Michael Smanski. Characterization of Programmable Transcription Activators in the Model Monocot Setaria viridis Via Protoplast Transfection. In Kan Wang and Zhang Feng, editors, *Protoplast Technology, Methods in Molecular Biology*, volume 2464, pages 223–244. Humana, New York, NY, 2022. doi: 10.1007/978-1-0716-2164-6_16.
23. Tomáš Čermák, Shaun J. Curtin, Javier Gil-Humanes, Radim Čegan, Thomas J.Y. Kono, Eva Konečná, Joseph J. Belanto, Colby G. Starker, Jade W. Mathre, Rebecca L. Greenstein, and Daniel F. Voytas. A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants. *The Plant Cell*, 29(6):1196–1217, July 2017. ISSN 1040-4651. doi: 10.1105/TPC.16.00922. Publisher: Oxford Academic.
24. Ann Boija, Isaac A. Klein, Benjamin R. Sabari, Alessandra Dall'Agnese, Eliot L. Coffey, Alicia V. Zamudio, Charles H. Li, Krishna Shrinivas, John C. Manteiga, Nancy M. Hannett, Brian J. Abraham, Lena K. Afeyan, Yang E. Guo, Jenna K. Rimel, Charli B. Fant, Jurian Schuijers, Tong Ihn Lee, Dylan J. Taatjes, and Richard A. Young. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*, 175(7):1842–1855.e16, December 2018. ISSN 0092-8674. doi: 10.1016/J.CELL.2018.10.042. Publisher: Cell Press.
25. Xiaoying Chen, Jennica L. Zaro, and Wei-Chiang Shen. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews*, 65(10):1357–1369, October 2013. ISSN 0169-409X. doi: 10.1016/j.addr.2012.09.039.
26. Bruce A Sherf, Shauna L Navarro, Rita R Hannah, and Keith V Wood. Dual-Luciferase TM Reporter Assay: An Advanced Co-Reporter Technology Integrating Firefly and Renilla Luciferase Assays. *Promega Notes Magazine Number*, 57:2, 1996.
27. Kenneth R. Luehrsen, Jeffrey R. de Wet, and Virginia Walbot. [35] Transient expression analysis in plants using firefly luciferase reporter gene. *Methods in Enzymology*, 216(C): 397–414, January 1992. ISSN 0076-6879. doi: 10.1016/0076-6879(92)16037-K. Publisher: Academic Press.
28. Philip N. Benfey and Nam Hai Chua. The Cauliflower Mosaic Virus 35S Promoter: Combinatorial Regulation of Transcription in Plants. *Science*, 250(4983):959–966, November 1990. ISSN 00368075. doi: 10.1126/SCIENCE.250.4983.959. Publisher: American Association for the Advancement of Science.
29. Maciej Maselko, Nathan Feltman, Ambuj Upadhyay, Amanda Hayward, Siba Das, Nathan Myslicki, Aidan J. Peterson, Michael B. O'Connor, and Michael J. Smanski. Engineering multiple species-like genetic incompatibilities in insects. *Nature Communications*, 11(1), December 2020. ISSN 20411723. doi: 10.1038/S41467-020-18348-1. Publisher: Nature Research.
30. Joanne E. Lee, Manuela Neumann, Daniel Iglesias Duro, and Markus Schmid. CRISPR-based tools for targeted transcriptional and epigenetic regulation in plants. *PLOS ONE*, 14 (9):e0222778, September 2019. ISSN 1932-6203. doi: 10.1371/JOURNAL.PONE.0222778. Publisher: Public Library of Science ISBN: 1111111111.
31. Mie Kasuga, Qiang Liu, Setsuko Miura, Kazuko Yamaguchi-Shinozaki, and Kazuo Shinozaki. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology*, 17(3):287–291, March 1999. ISSN 1546-1696. doi: 10.1038/7036. Number: 3 Publisher: Nature Publishing Group.
32. Motoaki Seki, Mari Narusaka, Hiroshi Abe, Mie Kasuga, Kazuko Yamaguchi-Shinozaki, Piero Carninci, Yoshihide Hayashizaki, and Kazuo Shinozaki. Monitoring the Expression Pattern of 1300 Arabidopsis Genes under Drought and Cold Stresses by Using a Full-Length cDNA Microarray. *The Plant Cell*, 13(1):61–72, January 2001. ISSN 1040-4651. doi: 10.1105/tpc.13.1.61.
33. Yoh Sakuma, Kyonoshin Maruyama, Yuriko Osakabe, Feng Qin, Motoaki Seki, Kazuo Shinozaki, and Kazuko Yamaguchi-Shinozaki. Functional Analysis of an Arabidopsis Transcription Factor, DREB2A, Involved in Drought-Responsive Gene Expression. *The Plant Cell*, 18 (5):1292–1309, May 2006. ISSN 1040-4651. doi: 10.1105/tpc.105.035881.
34. Yoh Sakuma, Kyonoshin Maruyama, Feng Qin, Yuriko Osakabe, Kazuo Shinozaki, and Kazuko Yamaguchi-Shinozaki. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proceedings of the National Academy of Sciences*, 103(49):18822–18827, December 2006. doi: 10.1073/pnas.0605639103. Publisher: Proceedings of the National Academy of Sciences.
35. Ai Zhang, Keita Matsuoka, Abdul Kareem, Madalen Robert, Pawel Roszak, Bernhard Blob, Anchal Bisht, Lieven De Veylder, Cătălin Voiniciuc, Masashi Asahina, and Charles W. Melnyk. Cell-wall damage activates DOF transcription factors to promote wound healing and tissue regeneration in Arabidopsis thaliana. *Current Biology*, 32(9):1883–1894.e7, May 2022. ISSN 0960-9822. doi: 10.1016/j.cub.2022.02.069.
36. Ya-Chen Huang, Chung-Yen Niu, Chen-Ru Yang, and Tsung-Luo Jinn. The Heat Stress Factor HSFA6b Connects ABA Signaling and ABA-Mediated Heat Responses. *Plant Physiology*, 172(2):1182–1199, October 2016. ISSN 0032-0889. doi: 10.1104/pp.16.00860.
37. J Armando Casas-Mollano, Matthew H Zinselmeier, Adam Sychla, and Michael J Smanski. Efficient gene activation in plants by the MoonTag programmable transcriptional activator. *Nucleic Acids Research*, 51(13):7083–7093, July 2023. ISSN 0305-1048. doi: 10.1093/nar/gkad458.
38. S. Boersma, Deepak Khuperkar, Bram M.P. Verhagen, S. Sonneveld, Jonathan B. Grimm, Luke D. Lavis, and Marvin E. Tanenbaum. Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. *Cell*, 178(2):458–472.e19, July 2019. ISSN 0092-8674. doi: 10.1016/J.CELL.2019.05.001. Publisher: Cell Press.
39. Steven J. Clough and Andrew F. Bent. Floral dip: a simplified method for Agrobacterium -mediated transformation of Arabidopsis thaliana. *The Plant Journal*, 16(6):735–743, December 1998. ISSN 1365-313X. doi: 10.1046/J.1365-313X.1998.00343.X. Publisher: John Wiley & Sons, Ltd.
40. Takashi L. Shimada, Tomoo Shimada, and Ikuko Hara-Nishimura. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *The Plant Journal*, 61(3):519–528, February 2010. ISSN 1365-313X. doi: 10.1111/J.1365-313X.2009.04060.X. Publisher: John Wiley & Sons, Ltd.
41. Shinobu Takada and Koji Goto. TERMINAL FLOWER2, an Arabidopsis Homolog of HETEROCHROMATIN PROTEIN1, Counteracts the Activation of FLOWERING LOCUS T by CONSTANS in the Vascular Tissues of Leaves to Regulate Flowering Time. *The Plant Cell*, 15(12):2856–2865, December 2003. ISSN 1040-4651. doi: 10.1105/TPC.016345. Publisher: Oxford Academic.
42. Martín Krzymuski, Fernando Andrés, Juan I. Cagnola, Seonghoe Jang, Marcelo J. Yanovsky, George Coupland, and Jorge J. Casal. The dynamics of FLOWERING LOCUS T expression encodes long-day information. *The Plant Journal*, 83(6):952–961, September 2015. ISSN 1365-313X. doi: 10.1111/TPJ.12938. Publisher: John Wiley & Sons, Ltd.
43. Y. Esther Tak, Joy E. Horng, Nicholas T. Perry, Hayley T. Schultz, Sowmya Iyer, Qiuming Yao, Luli S. Zou, Martin J. Aryee, Luca Pinello, and J. Keith Joung. Augmenting and directing long-range CRISPR-mediated activation in human cells. *Nature Methods 2021 18:9*, 18(9): 1075–1081, August 2021. ISSN 1548-7105. doi: 10.1038/s41592-021-01224-1. Publisher: Nature Publishing Group.
44. Johan Zicola, Liangyu Liu, Petra Tänzler, and Franziska Turck. Targeted DNA methylation represses two enhancers of FLOWERING LOCUS T in Arabidopsis thaliana. *Nature Plants 2019 5:3*, 5(3):300–307, March 2019. ISSN 2055-0278. doi: 10.1038/s41477-019-0375-2. Publisher: Nature Publishing Group ISBN: 4147701903752.
45. V Mladenova, E Mladenov, and G Russev. Organization of plasmid DNA into nucleosome-like structures after transfection in eukaryotic cells. *Biotechnology and Biotechnological Equipment*, 23(1):1044–1047, 2009. ISSN 1310-2818.
46. Pratibha B Hebbar and Trevor K Archer. Altered histone H1 stoichiometry and an absence of nucleosome positioning on transfected DNA. *Journal of Biological Chemistry*, 283(8): 4595–4601, 2008. ISSN 0021-9258.
47. Fumitaka Inoue, Martin Kircher, Beth Martin, Gregory M Cooper, Daniela M Witten, Michael T McManus, Nadav Ahituv, and Jay Shendure. A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome research*, 27(1):38–52, 2017. ISSN 1088-9051.
48. Luke A. Gilbert, Max A. Horlbeck, Britt Adamson, Jacqueline E. Villalta, Yuwen Chen, Evan H. Whitehead, Carla Guimaraes, Barbara Panning, Hidde L. Ploegh, Michael C. Bassik, Lei S. Qi, Martin Kampmann, and Jonathan S. Weissman. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, 159(3):647–661, October 2014. ISSN 0092-8674. doi: 10.1016/J.CELL.2014.09.029. Publisher: Cell Press.
49. Changtian Pan, Gen Li, Aimee A. Malzahn, Yanhao Cheng, Benjamin Leyson, Simon Sretenovic, Filiz Gurel, Gary D. Coleman, and Yiping Qi. Boosting plant genome editing with a versatile CRISPR-Combo system. *Nature Plants*, 8(5):513–525, May 2022. ISSN 2055-0278. doi: 10.1038/s41477-022-01151-9. Number: 5 Publisher: Nature Publishing Group.
50. Anthony Studer, Qiong Zhao, Jeffrey Ross-Ibarra, and John Doebley. Identification of a functional transposon insertion in the maize domestication gene tb1. *Nature Genetics 2011 43:11*, 43(11):1160–1163, September 2011. ISSN 1546-1718. doi: 10.1038/ng.942. Publisher: Nature Publishing Group.
51. William A. Ricci, Zefu Lu, Lexiang Ji, Alexandre P. Marand, Christina L. Ethridge, Nathalie G. Murphy, Jaclyn M. Noshay, Mary Galli, María Katherine Mejía-Guerra, Maria Colomé-Tatché, Frank Johannes, M. Jordan Rowley, Victor G. Corces, Jixian Zhai, Michael J. Scanlon, Edward S. Buckler, Andrea Gallavotti, Nathan M. Springer, Robert J. Schmitz, and Xiaoyu Zhang. Widespread long-range cis-regulatory elements in the maize genome. *Nature Plants 2019 5:12*, 5(12):1237–1249, November 2019. ISSN 2055-0278. doi: 10.1038/s41477-019-0547-0. Publisher: Nature Publishing Group.
52. Wen-Jing Bao, Yong-Guang Gao, Yong-Gang Chang, Tie-Ying Zhang, Xiao-Jing Lin, Xian-Zhong Yan, and Hong-Yu Hu. Highly efficient expression and purification system of small-size protein domains in Escherichia coli for biochemical characterization. *Protein Expression and Purification*, 47(2):599–606, June 2006. ISSN 1046-5928. doi: 10.1016/j.pep.2005.11.021.
53. Adam C. Fisher and Matthew P. DeLisa. Efficient isolation of soluble intracellular single-chain antibodies using the twin-arginine translocation machinery. *Journal of molecular biology*, 385(1):299, January 2009. ISSN 00222836. doi: 10.1016/J.JMB.2008.10.051. Publisher: NIH Public Access.
54. Nina Kronqvist, Medoune Sarr, Anton Lindqvist, Kerstin Nordling, Martins Otikovs, Luca Venturi, Barbara Pioselli, Pasi Purhonen, Michael Landreh, Henrik Biverstål, Zigmantas Toleikis, Lisa Sjöberg, Carol V. Robinson, Nicola Pelizzi, Hans Jörnvall, Hans Hebert, Kristaps Jaudzems, Tore Curstedt, Anna Rising, and Jan Johansson. Efficient protein production inspired by how spiders make silk. *Nature Communications 2017 8:1*, 8(1):1–15, May 2017. ISSN 2041-1723. doi: 10.1038/ncomms15504. Publisher: Nature Publishing Group.
55. Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, 1990. ISSN 00222836. doi: 10.1016/S0022-2836(05)80360-2.
56. Jianmei Cao, Dongmei Yao, Fan Lin, and Mingyi Jiang. PEG-mediated transient gene expression and silencing system in maize mesophyll protoplasts: A valuable tool for signal transduction study in maize. *Acta Physiologiae Plantarum*, 36(5):1271–1281, March 2014. ISSN 01375881. doi: 10.1007/S11738-014-1508-X/FIGURES/6. Publisher: Polish Academy of Sciences.

## Supplementary Note 1: Materials and Methods

**Plasmid construction.** Putative activation domains were selected from literature analysis for the potential to recruit transcription initiation machinery in plant cells. DNA binding domains from transcription factors were identified and removed on the basis of sequence conservation to experimentally determined DNA binding domain motifs. Patches of acidic/aromatic residues were identified as core motifs for a prospective activation domain, and a given sequence was selected to encompass as many of these patches as possible without including the DNA binding domain. Synthetic DNA fragments were then designed and codon optimized based on average codon usage in *A. thaliana* and *Oryza sativa*. Fragments were synthesized by Genscript with BsmBI cloning sites flanking the prospective activation domain. Fragments were cloned into an scFv destination vector with BsaI and BsmBI cloning sites generating compatible overhangs. The assembled plasmids generate an in-frame C terminal fusion of the activation domain to the scFv. The scFv is driven by the *CaMV 35S* promoter, and also contains a C terminal fusion of the GB1 solubility tag. The scFv-AD is recruited to targets by the dCas9-GCN4. This vector was generated by fusing 24 copies of the GCN4 epitope tag to the C-terminus of *A. thaliana* codon optimized dCas9. Each GCN4 repeat was separated by short linker GS linker sequences. The dCas9 is expressed under the AtUbi10 promoter in dicots, the *CMYLCV* promoter in *S. viridis*, or the *ZmUbi10* promoter in *Z. mays*. Guide RNAs were designed and expressed from *U6* promoters, either the *A. thaliana* U6 promoter for dicots or the *O. sativa U6* promoter for monocots. A complete list of plasmids used in this study can be found in Supplemental Table 1.

**PTA protein sequence similarity.** BLAST (ver. 2.8.1+)(55) was used to perform an all-by-all comparison of the proteins from which the activators were derived. Full length amino acid sequences were used as both query and subject, and no e-value threshold was set. Expectation values were negative log transformed for plotting.

**Protoplast isolation and transformation.** *S. viridis* and *A. thaliana* protoplasts were isolated and transformed as stated in previous protocols(21, 22). ME034 *S. viridis* plants are grown in a growth chamber set to 31/22 degrees C with 12h diurnal cycle. Mesophyll cells can be isolated from leaf tissue 14-21 days post germination. Col-0 *A. thaliana* plants are grown in a growth chamber set to 22C with 16h light cycle. Mesophyll cells can be isolated from leaf tissue 14-21 days post germination. *Z. mays* protoplasts were isolated and transformed as stated in previous protocols. (56) *Z. mays* plants were grown in a growth chamber at 25 degrees C under 16h light cycle. Prior to isolation, plants were 'greened' by first placing the seedlings in the dark for 5 days post-germination followed by 2 days of exposure to light prior to isolation. Plasmids to be transformed in all systems are prepared according to the manufacturer's protocol (Qiagen 12945) to ensure transformation grade endotoxin free DNA.

**Dual luciferase assay.** A dual luciferase assay was designed with the firefly and Renilla luciferase reporter genes from Promega (26). A Renilla luciferase was cloned upstream of a firefly luciferase on a single plasmid. The Renilla CDS was under the control of a constitutive promoter, either the *CaMV 35S* or weaker *Nos* promoter. Upstream of the firefly luciferase CDS was a minimal *CaMV 35S* promoter, containing 60 bp upstream of the TSS. Six copies of the *lac* operator were cloned upstream of the minimal promoter driving firefly luciferase. A sgRNA targeting the *LacO* region results in targeting of the dCas9 activation to the firefly luciferase promoter. We drove dCas9-24xGCN4 with a constitutive promoter (*CMYLCV, AtUbi10, or ZmUbi1*), the *LacO* sgRNA driven by *O. sativa U6* promoter, and the scFv with a constitutive *35S* promoter. We co-transformed *S. viridis*, *A. thaliana*, or *Z. mays* protoplasts with 20 micrograms each of the dCas9-24xGCN4, *Lac*O sgRNA, scFv-AD, and dual luciferase reporter plasmids. Following 24h incubation at 25C, we lysed the cells in 1X Passive Lysis Buffer from Promega. We then quantified firefly and Renilla luciferase luminescence using the GloMax Explorer plate reader (Promega GM3500) equipped with dual injectors with the Dual Luciferase Assay kit (Promega E1960). Firefly luciferase substrate was injected first and luminescence was quantified, followed by injection of the Renilla luciferase substrate and subsequent quantification. We normalized for delivery first by dividing firefly luciferase activity by Renilla luciferase activity unless otherwise described, followed by Fold Change calculation by dividing a given AD by the negative control or VP64. A complete list of plasmids used in this study can be found in Supplemental Table 2.

**RNA isolation and quantification.** We isolate RNA from either protoplasts or leaf tissue according the the TRIzol manufacturer's protocol (Thermo 15596026). For protoplasts we spun the cells down at 500g for 2 minutes, followed by removal of W5 buffer before re-suspending cells in 1mL of TRIzol. RNA was collected 24h post-transformation for protoplast samples. For plant tissue samples, we first snap froze the sample in liquid nitrogen followed by shaking in a paint shaker apparatus with metal beads to homogenize frozen tissue. We then add 1mL of TRIzol to the homogenized tissue. We then follow the TRIzol protocol according to manufacturer specifications. We then treat the samples with Turbo DNA Free Kit (Invitrogen AM1907) to remove any plasmid and/or genomic DNA from the RNA sample. To quantify a transcript we then perform RT-qPCR using gene-specific primer pairs. A list of primers used for quantification can be found in Supplemental Table 2. We follow a RT-qPCR cycling protocol as defined by the NEB Luna One-Step RT-qPCR kit manufacturer's protocol (NEB E3005L). Primers are designed to have Tm values of 60C and yield amplicon lengths between 75-175bp. The primers typically span an intron,

such that the shorter PCR product corresponds to spliced mRNA. We quantify gene expression for relative comparison between treatments using the delta delta Ct method. All primers can be found in Supplemental Table 2.

**Plant transformation and genotyping.** To generate transgenic *A. thaliana* plants we performed floral dip protocols as previously described (39). Binary vectors were constructed containing an antibiotic resistance gene along with the pFAST Oleosin-RFP transgene (40). *A. thaliana* ecotype Columbia-0 was grown to maturity and floral dipped with Agrobacterium strain GV3101 transformed with the described binary vectors. The antibiotic selection used in this study was either Kanamycin or Bialaphos resistance. T0 seeds were identified by the pFAST system in which RFP-positive seedlings contain the transgene. We placed RFP+ seedlings directly onto soil, or onto selective media, to grow T1 plants. T1 plants were grown to maturity and allowed to set seed. All molecular characterization was performed on T2 lines, along with phenotyping. To quantify *FT* overexpression phenotypes in T2 plants, we planted 18 RFP positive T2 plants from each T1 parent. We chose time point 0 to be the date at which the first set of true leaves emerge from a seedling. We then count the number of days elapsed until bolting is observed. We also count the number of rosette leaves one day after bolting was observed. RT-qPCR was performed to quantify target gene expression, using primers indicated in Supplemental Table 2.

## Supplementary Note 2: Supplementary Table S1. Activation domains

| Name | Domain Source | AA range | AD Seq | DOI | Uniprot |
|---|---|---|---|---|---|
| DOF1 | *Z. mays* | 175-238 | QPGTEDAEAV ALGLGLSDFP SAGKAVLDDE DSFVWPAASF DMGACWAGAG FADPDPACIF LNLP | 10.1093/pcp/pce105 | http://www.uniprot.org/uniprot/P38564 |
| DREB1 | *A. thaliana* | 108-216 | GRSACLNFAD SAWRLRIPES TCAKDIQKAA AEAALAFQDE MCDATTDHGF DMEETLVEAI YTAEQSE-NAF YMHDEAMFEM PSLLANMAEG MLLPLPSVQW NHNHEVDGDD DDVSLWSY | 10.1111/pbi.12057 | http://www.uniprot.org/uniprot/Q9M0L0 |
| DREB2 | *A. thaliana* | 254-335 | SSDMFDVDEL LRDLNGDDVF AGLNQDRYPG NSVANGSYRP ESQQSGFDPL QSLNYGIPPF QLEGKDGNGF FDDLSYLDLEN | 10.1111/pbi.12057 | http://www.uniprot.org/uniprot/O82132 |
| IPN2 | *L. japonicus* | 179-358 | AATNLKGIGP QTIPDMGIMK EFGSPLGFSF QDLDLYGGGG GDQLELQQNM EKPPLDGFMP MNHENL-CLGK KRPNPYSGNN GKSPLMWSDD LRLQDLGSCL QDDPFKGDHH HQIQIAPPSL DRGTEMDPMS EIYDSKPEEK KFDASMKLER PSPRRAPLGE RMSPMITTGT MAQGRSSPFG | 10.1111/nph.12593 | https://www.uniprot.org/uniprot/E5L8F7 |
| PTI4 | *S. lycopersicum* | 170-234 | EPEPVRVTAK RRASPEPASS SGNGSMKRRR KAVQKCDGEM ASRSSVMQVG CQIEQLTGVH QLLVI | 10.1111/pbi.12057 | http://www.uniprot.org/uniprot/O04680 |
| AtARF8 | *A. thaliana* | 350-702 | VSLWEIEPLT TFPMYPSLFP LRLKRPWHAG TSSLPDGRGD LGSGLTWLRG GGGEQQGLLP LNYPSVGLFP WMQQRLDLSQ MGTDNNQQYQ AMLAAGLQNI GGGDPLRQQF VQLQEPHHQY LQQSASHNSD LMLQQQQQQQ ASRHLMHAQT QIMSENLPQQ NMRQEVSNQP AGQQQQLQQP DQNAYLNAFK MQNGHLQQWQ QQSEMPSPSF MKSDFTDSSN KFATTASPAS GDGNLLNFSI TGQSVLPEQL TTEGWSPKAS NTFSEPLSLP QAYPGKSLAL EPGNPQNPSL FGVDPDSGLF LPSTVPRFAS SSGDAEASPM SLTDSGFQNS LYSCMQDTTH ELLHGAGQIN S | 10.1105/tpc.008417 | https://www.uniprot.org/uniprot/Q9FGV1 |
| SNAC1 | *O. sativa* | 172-314 | WEKMQQGKEV KEEASDMVTS QSHSHTHSWG ETRTPESEIV DNDPFPELDS FPAFQPAPPP ATAMMVP-KKE SMDDATAAAA AAATIPRNNS SLFVDLSYDD IQGMYSGLDM LPPGDDFYSS LFASPRVKGT TPRA-GAGMGM VPF | 10.1073/pnas.0604882103 | http://www.uniprot.org/uniprot/A2XNB9 |
| Opaque-2 | *Z. mays* | 41-227 | IDVAAAGHGD GDMMDQQHAT EWTFERLLEE EALTTSTPPP VVVVPNSCCS GALNADRPPV MEEAVTMAPA AVSSAVVGDP MEYNAILRRK LEEDLEAFKM WRADSSVVTS DQRSQGSNNH TGGSSIRNNP VQNKLMNGED PINNNHAQTA GLGVRLATSS SSRDPSPSDE DMDGEVEILG FKM | PMC146487 | http://www.uniprot.org/uniprot/P12959 |
| AvrXa10 | *X. oryzae* | | TVMWEQDAAP FAGAADDFPA FNEEELAWLM ELLPQSGSVG GTI | 10.1094/MPMI.1998.11.8.824 | https://www.uniprot.org/uniprot/Q56830 |
| AL2/C2 /TrAP | Tomato golden mosaic virus | 83-129 | PLHQHQDIPL TNQVQPQPEE SIGSPQGISQ LPSMDDIDDS FWENLFK | 10.1006/viro.1999.9925 | http://www.uniprot.org/uniprot/P03562 |
| AtARR2 | *A. thaliana* | 279-664 | QHQGNMNHSF MTGQDQSFGP LSSLNGFDLQ SLAVTGQLPP QSLAQLQAAG LGRPTLAKPG MSVS-PLVDQR SIFNFENPKI RFGDGHGQTM NNGNLLHGVP TGSHMRLRPG QNVQSSGMML PVADQLPRGG PSMLPSLGQQ PILSSSVSRR SDLTGALAVR NSIPETNSRV LPTTHSVFNN FPADLPRSSF PLASAPGISV PVSVSYQEEV NSSDAKGGSS AATAGFGNPS YDIFNDFPQH QQHNKNISNK LNDWDLRNMG LVFSSNQ-DAA TATATAAFST SEAYSSSSTQ RKRRETDATV VGEHGQNLQS PSRNLYHLNH VFMDGGSVRV KSER-VAETVT CPPANTLFHE QYNQEDLMSA FLKQEGIPSV DNEFEFDGYS IDNIQV | 10.1111/j.1365-313X.2000.00909.x | https://www.uniprot.org/uniprot/Q9ZWJ9 |
| RF2a | *O. sativa* | 56-108 | EILSLPEDLD LCAAGGGDGP SLSDENDEEL FSMFLDVEKL NSTCGASSEA EAE | 10.1074/jbc.M304862200 | http://www.uniprot.org/uniprot/Q69IL4 |
| OsCBT | *O. sativa* | 542-755 | PVTESLLELV LRNRLQEWLV EMVMEGHKST GRDDLGQGAI HLCSFLGYTW AIRLFSLSGF SLD-FRDSSGW TALHWAAYHG RERMVATLLS AGANPSLVTD PTPESPAGLT AADLAARQGY DGLAAY-LAEK GLTAHFEAMS LSKDTEQSPS KTRLTKLQSE KFEHLSEQEL CLKESLAAYR NAADAASNIQ AAL-RERTLKL QTKAIQLANP EIEASEIV | 10.1074/jbc.M504616200 | https://www.uniprot.org/uniprot/Q7XHR2 |
| AP1* | *A. thaliana* | 193-256 | HNMPPPLPPQ QHQIQHPYML SHQPSPFLNM GGLYQEDDPM AMRRNDLELT LEPVYNCNLG CFAA | 10.1023/A:1006273127067 | https://www.uniprot.org/uniprot/P35631 |
| MYC2 | *A. thaliana* | 148-188 | GGVAPSDDAV DEEVTDTEWF FLVSMTQSFA CGAGLAGKAFA | 10.1016/j.celrep.2017.04.057 | https://www.uniprot.org/uniprot/Q39204 |
| ZmVP1 | *Z. mays* | 1-120 | MEASSGSSPP HSQENPPEHG GDMGGAPAEE IGGEAADDFM FAEDTFPSLP DFPCLSSPSS STFSSNSSSN SSSAYTNTAG RAGGEPSEPA SAGEGFDALD DIDQLLDFAS LSMPWDSEPF | 10.1016/0092-86749190436-3 | https://www.uniprot.org/uniprot/P26307 |

## Supplementary Note 3: Supplementary Table S1. Activation domains continued

| Name | Domain Source | AA range | AD Seq | DOI | Uniprot |
|---|---|---|---|---|---|
| OsGRF1 | *O.sativa* | 221-396 | PPPSYYSMDH KEYAYGHATK EVHGEHAFFS DGTEREHHHA AAGHGQWQFK QLGMEPKQST TPLFP-GAGYG HTAASPYAID LSKEDDDEKE RRQQQQQQQQ HCFLLGADLR LEKPAGHDHA AAAQKPLRHF FDEWPHEKNS KGSWMGLEGE TQLSMSIPMA ANDLPITTTS RYHNDE | 10.1093/pcp/pch098 | https://www.uniprot.org/uniprot/A2XA73 |
| AtHSFA2 | *A. thaliana* | 230-341 | KEKKSLFGLD VGRKRRLTST PSLGTMEENL LHDQEFDRMK DDMEMLFAAA IDDEANNSMP TKEEQ-CLEAM NVMMRDGNLE AALDVKVEDL VGSPLDWDSQ DLHD | 10.1111/j.1365-313X.2004.02111.x | https://www.uniprot.org/uniprot/O80982 |
| AtHSFA6b | *A. thaliana* | 252-391 | KEKRKEIEEA ISKKRQRPID QGKRNVEDYG DESGYGNDVA ASSSALIGMS QEYTYGNMSE FEMSELD-KLA MHIQGLGDNS SAREEVLNVE KGNDEEEVED QQQGYHKENN EIYGEGFWED LLNEGQNFDF EGDQENVDVG SSSHTN | 10.1111/j.1365-313X.2004.02111.x | https://www.uniprot.org/uniprot/Q9LUH8 |
| EFS | *A. thaliana* | 820-1230 | KTTGALLDAD IGKTSATYGT ISSDVTHGEM VVDVTIEDSY STESAWVRCD DCFKWRRIPA SVVG-SIDESS RWICMNNSDK RFADCSKSQE MSNEEINEEL GIGQDEADAY DCDAAKRGKE KEQKSKRLTG KQKACFKAIK TNQFLHRNRK SQTIDEIMVC HCKPSPDGRL GCGEECLNRM LNIECLQGTC PAGDLC-SNQQ FQKRKYVKFE RFQSGKKGYG LRLLEDVREG QFLIEYVGEV LDMQSYETRQ KEYAFKGQKH FYFMTLNGNE VIDAGAKGNL GRFINHSCEP NCRTEKWMVN GEICVGIFSM QDLKKGQELT FDYNYVRVFG AAAKKCYCGS SHCRGYIGGD PLNGDVIIQS DSDEEYPELV ILDDDESGEG ILGATSRTFT DDADEQMPQS FEKVNGYKDL APDNTQTQ | 10.1105/tpc.109.070060 | https://www.uniprot.org/uniprot/F4I6Z9 |
| AFT1* | *A. thaliana* | Full length | MAATLGRDQY VYMAKLAEQA ERYEEMVQFM EQLVTGATPAE ELTVEERNLL SVAYKNVIGS LRAAWRIVSS IEQKEESRKN DEHVSLVKDY RSKVESELSS VCSGILKLLD SHLIPSAGAS ESKV-FYLKMK GDYHRYMAEF KSGDERKTAA EDTMLAYKAA QDIAAADMAP THPIRLGLAL NFSVFYYEIL NSSDKACNMA KQAFEEAIAE LDTLGEESYK DSTLIMQLLR DNLTLWTSDM QEQMDEA | 10.1016/S0014-57939801739-6 | https://www.uniprot.org/uniprot/P48349 |
| EIN3 | *A. thaliana* | 310-620 | SLSGGSCSLL MNDCSQYDVEGF EKESHYEVEE LKPEKVMNSSN FGMVAKMHDF PVKEEVPAGNS EFMRKRKPNR DLNTIMDRTV FTCENLGCAH SEISRGFLDR NSRDNHQLAC PHRDSRLPYG AAP-SRFHVNE VKPVVGFPQP RPVNSVAQPI DLTGIVPEDG QKMISELMSM YDRNVQSNQT SMVMENQSVS LLQPTVHNHQ EHLQFPGNMV EGSFFEDLNI PNRANNNNSS NNQTFFQGNN NNNNVFKFDT ADHNN-FEAAH NNNNNSSGNR FQLVFDSTPF DMASFDYRDD MSMPGVVGTM DGMQQKQQDV SIWF | 10.1016/j.jmb.2005.02.065 | https://www.uniprot.org/uniprot/O24606 |
| WRKY50 | *A. thaliana* | 1-106 | MNDADTNLGS SFSDDTHSVF EFPELDLSDE WMDDDLVSAVS GMNQSYGYQT SDVAGALFSG SSSCF-SHPES PSTKTYVAAT ATASADNQNK KEKKKIKGRV AFKTR | 10.3389/fpls.2018.00930 | https://www.uniprot.org/uniprot/Q8VWQ5 |
| ATXR7 | *A. thaliana* | 1250-1423 | DGADVLKMSQ LKARKKHLRF QQSKIHDWGL VALEPIEAEDF VIEYVGELIR SSISEIRERQ YEK-MGIGSSY LFRLDDGYVL DATKRGGIAR FINHSCEPNC YTKIISVEGK KKIFIYAKRH IDAGEEISYN YKFPLEDDKI PCNCGAPNVY CFCEQVPWIA KLKRRTWFSR RN | 10.1105/tpc.109.070060 | https://www.uniprot.org/uniprot/F4K1J4 |
| ASF1a | *A. thaliana* | Full length | MSAIKITNVA VLHNPAPFVS PFQFEISYECL NSLKDDLEWKL IYVGSAEDETY DQLLESVLVGP VN-VGNYRFVFQ ADPPDPSKIQE EDIIGVTVLL LTCSYMGQEF LRVGYYVNNDY EDEQLKEEPPT KVLID-KVQRN ILSDKPRVTK FPIDFHPEEE QTAATAAPPE QSDEQQPNVN GEAQVLPDQS VEPKPEES | 10.1111/pce.12299 | https://www.uniprot.org/uniprot/Q9C9M6 |
| HAG1 | *A. thaliana* | 150-550 | LESSDGGKDG GSSVVGTGV SGTVGGSSIS GLVPKDESVK VLAENFQTSGA YIAREEALKR EEQA-GRLKFVC YSNDSIDEHM MCLIGLKNIF ARQLPNMPKE YIVRLLMDRK HKSVMVLRGN LVVGGI-TYRP YHSQKFGEIA FCAITADEQV KGYGTRLMNH LKQHARDVDG LTHFLTYADN NAVGYFVKQG FTKEIYLEKD VWHGFIKDYD GGLLMECKID PKLPYTDLSS MIRQQRKAID ERIRELSNCQ NVYP-KIEFLK NEAGIPRKII KVEEIRGLRE AGWTPDQWGH TRFKLFNGSA DMVTNQKQLN ALMRALLKTM QDHADAWPFK EPVDSRDVPD YYDIIKDPIDL KVIAKRVESEQ YYVTLDMFVAD ARRMFNNCRT YN-SPDTIYYKC ATRLETHFHS KVQAGLQSGA KSQ | 10.1093/nar/29.7.1524 | https://www.uniprot.org/uniprot/Q9AR19 |
| GT-3A | *A. thaliana* | 110-323 | RYKACETTEP DAIRQQFPFY NEIQSIFEAR MQRMLWSEAT EPSTSSKRKH HQFSSDDEEE EVDEP-NQDIN EELLSLVETQ KRETEVITTS TSTNPRKRAK KGKGVASGTK AETAGNTLKD ILEEFMRQTV KMEKEWRDAW EMKEIEREKR EKEWRRRMAE LEEERAATER RWMEREEERR LREEARAQKR DSLI-DALLNR LNRDHNDDHH NQGF | 10.1016/S0014-57930400222-4 | https://www.uniprot.org/uniprot/Q9SDW0 |
| CVBp12 | Chrym. vir. B | Full | LCVDIYKRAF PRSVNKGRSS YAGGSGGVPG ISYNSKVRDY ILWGVTEVIP | 10.1105/tpc.112.106476 | https://www.uniprot.org/uniprot/P37992 |
| ASF1b | *A. thaliana* | Full length | MSSINITNVT VLDNPAPFVN PFQFEISYEC LTSLKDDLEW KLIYVGSAED ETYDQVLESV LVGPVN-VGNY RFVLQADSPD PLKIREEDII GVTVLLLTCSY MDQEFIRVGY YVNNDYDDEQ LREEPPTKVLI DKVQRNILTD KPRVTKFPINFH PENEQTLGDGP APTEPFADSV VNGEAPVFLE QPQKLQEIEQ FDDSD-VNGEA IALLDQPQNL QET | 10.1111/pce.12299 | https://www.uniprot.org/uniprot/Q9LS09 |
| TBP1* | *A. thaliana* | Full | MTDQGLEGSN PVDLSKHPSG IVPTLQNIVS TVNLDCKLDL KAIALQARNA EYNPKRFAAVI MRIREP-KTTA LIFASGKMVC TGAKSEDFSKM AARKYARIVQ KLGFPAKFKD FKIQNIVGSC DVKFPIRLEG LAYSHAAFSS YEPELFPGLI YRMKVPKIVL LIFVSGKIVI TGAKMRDETY KAFENIYPVL SEFRKIQQ | na | https://www.uniprot.org/uniprot/P28147 |

## Supplementary Note 4: Supplementary Table S1. Activation domains continued

| Transcription Factor | Gene ID | AA range | AD sequence | Source | Uniprot |
|---|---|---|---|---|---|
| Ada2b | *A. thaliana* | 91-487 | RVMGNLTFPL ICPDWSADDE MLLLEGLEIY GLGNWAEVAE HVGTKSKEQC LEHYRNIYLN SPFF-PLPDMS HVAGKNRKEL QAMAKGRIDD KKAEQNMKEE YPFSPPKVKV EDTQKESFVD RSFG-GKKPVS TSVNNSLVEL SNYNQKREEF DPEYDNDAEQ LLAEMEFKEN DTPEEHELKL RVLRIYSKRL DERKRRKEFI IERNLLYPNP FEKDLSQEEK VQCRRLDVFM RFHSKEEHDE LLRNVVSEYR MVKR-LKDLKE AQVAGCRSTA EAERYLGRKR KRENEEGMNR GKESGQFGQI AGEMGSRPPV QASSSYVNDL DLIGFTESQL LSESEKRLCS EVKLVPPVYL QMQQVMSHEI FKGNVTKKSD AYSLFKIDPT KVDRVY-DMLV KKGIAQL | na | https://www.uniprot.org/uniprot/Q9ATB4 |
| BRE1 | *A. thaliana* | 700-876 | DIQQGSAYAS FLSKKSSRIE DQLRFCTDQF QKLAEDKYQK SVSLENLQKK RADIGNGLEQ ARSR-LEESHS KVEQSRLDYG ALELELEIER FNRRRIEEEM EIAKKKVSRL RSLIEGSSAI QKLRQELSEF KEILKCKACN DRPKEVVITK CYHLFCNPCV QKLTGTRQKK CPTCSASFGP NDIKPIYI | na | https://www.uniprot.org/uniprot/Q8RXD6 |
| bHLH122 | At1g51140 | 1-300 | MESEFQQHHF LLHDHQHQRP RNSGLIRYQS APSSYFSSFG ESIEEFLDRP TSPETERILS GFLQTTDTSD NVDSFLHHTF NSDGTEKKPP EVKTEDEDAE IPVTATATAM EVVVSGDGEI SVNPEVSIGY VASVS-RNKRPR EKDDRTPVNN LARHNSSPAG LFSSIDVETA YAAVMKSMGG FGGSNVMSTS NTEASSLTPR SKLLPPTSRA MSPISEVDVK PGFSSRLPPR TLSGGFNRSFG NEGSASSKLTA LARTQSGGLD QYKTKD-EDS ASRRPPLAHH MSLPKSLSDI EQLLSDSI | DAP seq dataset | https://www.uniprot.org/uniprot/Q9C690 |
| LMI1 | AT5G03790 | 140-234 | AKQLEQLYDS LRQEYDVVSR EKQMLHDEVK KLRALLRDQG LIKKQISAGT IKVSGEEDTV EISSVVVAHPR TENMNANQIT GGNQVYGQYN NPMLVASSGW PSYP | DAP seq dataset | https://www.uniprot.org/uniprot/Q9LZR0 |
| SDG31 | AT3G04380 | 196-462 | CCANCKGNCLS ADFPCTCARE TSGEYAYTKE GLLKEKFLDT CLKMKKEPDS FPKVYCKDCP LERDHDKGTY GKCDGHLIRK FIKECWRKCG CDMQCGNRVV QRGIRCQLQV YFTQEGKGWG LRTLQDLPKG TFICEYIGE ILTNTELYDR NVRSSSERHT YPVTLDADWG SEKDLKDEEA LCLDAT-ICGN VARFINHRCE DANMIDIPIE IETPDRHYYH IAFFTLRDVK AMDELTWDYM IDFNDKSHPV KAFRCCCGSE SCRDRKI | DAP seq dataset | https://www.uniprot.org/uniprot/Q8W595 |
| VRN1 | At3g18990 | 100-240 | EINYHSTGLM DSAHNHFKRA RLFEDLEDED AEVIFPSSVY PSPLPESTVP ANKGYASSAI QTLFTG-PVKA EEPTPTPKIP KKRGRKKKNA DPEEINSSAP RDDDPENRSK FYESASARKR TVTAEERERA IN-AAKTFEPT | DAP seq dataset | https://www.uniprot.org/uniprot/Q8L3W1 |
| ATHB13* | AT1G69780 | 151-294 | RQFDTLKAEN DLLQTHNQKL QAEIMGLKNR EQTESINLNK ETEGSCSNRS DNSSDNLRLD ISTAPP-SNDS TLTGGHPPPP QTVGRHFFPP SPATATTTTT TMQFFQNSSS GQSMVKEENS ISNMFCAMDD HS-GFWPWLDQ QQYN | DAP seq dataset | https://www.uniprot.org/uniprot/Q8LC03 |
| ABI5 | AT2G36270 | 1-340 | MVTRETKLTS EREVESSMAQ ARHNGGGGGE NHPFTSLGRQ SSIYSLTLDE FQHALCENGK NFGSMN-MDEF LVSIWNAEEN NNNQQQAAAA AGSHSVPANH NGFNNNNNNG GEGGVGVFSG GSRGNEDANN KRGIANESSL PRQGSLTLPA PLCRKTVDEV WSEIHRGGGS GNGGDSNGRS SSSNGQNNAQ NGGE-TAARQP TFGEMTLEDF LVKAGVVREH PTNPKPNPNP NQNQNPSSVI PAAAQQQLYG VFQGTGDPSF PGQAMGVGDP SGYAKRTGGG GYQQAPPVQA GVCYGGGVGF GAGGQQMGMV GPLSPVSSDG LGHGQVDNIG GQYGVDMGGL | DAP seq dataset | https://www.uniprot.org/uniprot/Q9SJN0 |
| LBD2 | AT1G06280 | 45-206 | FPARKTKEFQ AVHKVFGVSN VQKMVRTVRE EDRTKLSDSL TWEALWRQKD PVLGSYGEYR RICEELKLYK SLVHNQPLIG WDNNQRVFNN NSNNKNGLAM TNSSGSGGFS VNNNGVGVNR EIVNGGYASR NVQGGWENLK HDQRQQCYAV INNGFKQHYL PL | DAP seq dataset | https://www.uniprot.org/uniprot/Q9LNB9 |
| BBX21 | AT1G75540 | 110-331 | SSVYKPTSKS SSSSSSNQDF SVPGSSISNP PPLKKPLSAP PQSNKIQPFS KINGGDASVN QWGSTSTISE YLMDTLPGWH VEDFLDSSLP TYGFSKSGDD DGVLPYMEPE DDNNTKRNNN NNNNNNNNTV SLP-SKNLGIW VPQIPQTLPS SYPNQYFSQD NNIQFGMYNK ETSPEVVSFA PIQNMKQQGQ NNKRWYDDGG FTVPQITPPP LSSNKKFRSF W | DAP seq dataset | https://www.uniprot.org/uniprot/Q9LQZ7 |

* These activation domains did not assemble correctly into target plasmids after multiple attempts and were not carried forward.

## Supplementary Note 5: Supplementary Table S2. sgRNA constructs

914

| Name | Link |
|---|---|
| AtU6 LacO sgRNA | https://benchling.com/s/seq-tOShh49fIMXhHFR2n00n?m=slm-VN5W2LGHQAlKLQ9ZvgQH |
| AtWUS 4x sgRNA repeat | https://benchling.com/s/seq-flgpONlt0q2lJnBwHnBc?m=slm-93jZN4A3soTozlEVIM4u |
| OsU6 LacO sgRNA | https://benchling.com/s/seq-vIyHY56Msirwh7nw742r?m=slm-ZR73jW5PrQLFUOOZOn6P |
| AtCLV3g1 | https://benchling.com/s/seq-QWrmfZc1Ec4jfRGqrffm?m=slm-yvBAe0D1ae4A2FLpiaIH |
| AtCLV3g2 | https://benchling.com/s/seq-asNJ6KJvO89swjwsrLNi?m=slm-vk1jze9mtjoP1TzoXRiw |
| FT BlockB sgRNA | https://benchling.com/s/seq-5ziJjbPSXUO7pftspKNn?m=slm-yPNQBDddE3vTFPT7e5wA |
| FT BlockC sgRNA | https://benchling.com/s/seq-PjFkpaHYbWxWtlo9kHf7?m=slm-dhtgd1CTfa1Clvdpywxk |
| FT BlockE sgRNA | https://benchling.com/s/seq-00MeZfV66n40zjyVIC9L?m=slm-vOxkbC9Faka3ax0Zt1Ts |
| FTgA sgRNA | https://benchling.com/s/seq-ZniKsGa2bnk8ZBummIJX?m=slm-bwtcu2gQfWNGLkHDabep |
| FTgB sgRNA | https://benchling.com/s/seq-qPKXU5P5KsZ4M9ekzaPX?m=slm-sjmjHtZt7ZGI5AqzW7qw |
| FT IntB sgRNA | https://benchling.com/s/seq-Lz9NPOLRl0r5fhwCDMUw?m=slm-CWTHhnhIBfr4TKMhU7hJ |
| FT IntC sgRNA | https://benchling.com/s/seq-5CjumRVX0cewolyPVvhQ?m=slm-esyIgOR1PRJkSNgaqKtk |
| FT IntE sgRNA | https://benchling.com/s/seq-UkhXeIHENHQ9mqmUaQLL?m=slm-sTSgfk7BIj0FNqndoEpg |
| *PAP1* sgRNA | https://benchling.com/s/seq-gDuRWc02rOPJXxh6z9n8?m=slm-LlXCm27EI7u6OFjqeXKH |
| WUS sgRNA | https://benchling.com/s/seq-wk3d0aJUfeAMj9Hhfv5J?m=slm-Z1fhErOAFrgFMZzfX0zR |

## Supplementary Note 6: Supplementary Table S3. coding sequences

| Name | Link |
|---|---|
| AtUbi10-dCas9-24xGCN4 | https://benchling.com/s/seq-V82BZGxKbmZPsEALuIjz?m=slm-1VUiIIysQL7H07sPVKiB |
| CmYLCV-dCas9-24xGCN4 | https://benchling.com/s/seq-pi1LDtQbPCT8AkOZA9zI?m=slm-6goMxLqKdLM5Fueln4Sf |
| p6xLacOfLuc | https://benchling.com/s/seq-sdCa6QHGpDqKMOJmuoxQ?m=slm-k8CSiyZg2MnRl3vEoSdX |
| pABI5 | https://benchling.com/s/seq-tJH9toq3q0h5kmb0QFsZ?m=slm-hcIBNlydbsbwtjHqMEAs |
| pAda2b | https://benchling.com/s/seq-L4JonwyGnfLpC9v0p7ec?m=slm-AMxCq3zOQfdyJKo8Vnyt |
| pAFT1 | https://benchling.com/s/seq-LE9lqv9JuA1uHB1QsyBC?m=slm-euBk5W2gIPNXvWZtWGQv |
| pAL2-TrAP | https://benchling.com/s/seq-ToKRHFff0aMbT075SfCl?m=slm-Jhx3STbRueBmY8qDcLQK |
| pAP1 | https://benchling.com/s/seq-GND9iOIG4BvBqFy6RSOd?m=slm-xFfl4xOmu58VBl2NG8DA |
| pAsf1a | https://benchling.com/s/seq-hmnBk1OyBu9hzoNLl9nm?m=slm-QtifPCmgB4qQWpyT0exz |
| pAsf1b | https://benchling.com/s/seq-MYpDPxFoa4gKIS8dNRyR?m=slm-OxPcHnRB8QWTltm7XIez |
| pAtARF8 | https://benchling.com/s/seq-CUyCV0yHwdn39DnX9dXn?m=slm-mnRDYOeryvBQ7rkYb0uZ |
| pAtARR2 | https://benchling.com/s/seq-cNWHH2ierYKqPGig2e0q?m=slm-wErAczXmKCdE9GLiz1H2 |
| pATHB13 | https://benchling.com/s/seq-AASedSoGrBNOWDYNV66e?m=slm-Sl3u4Dg2oj97cboM8695 |
| pAtHSFA2 | https://benchling.com/s/seq-OzY54QANgYCK9k4kpHk4?m=slm-jmnSGI9SNhebcZ6qD8eC |
| pAtHSFA6b | https://benchling.com/s/seq-YKjpgyzzdrTIea4UkFZ7?m=slm-nr84ggjYXlxIa0o4PMns |
| pATXR7 | https://benchling.com/s/seq-e1PrHK1m4z0ioVEim1ZE?m=slm-DJqzrV2QQLkRDuGLGqlF |
| pAvrXa10 | https://benchling.com/s/seq-rJPtr5GAijv4WAH2tR9b?m=slm-203WHqlY6rhsamLtiCq7 |
| pBBX21 | https://benchling.com/s/seq-C4EDu8ltbug9MlnfDtKM?m=slm-79dpnpKf5csnCSLuZzVO |
| pbHLH122 | https://benchling.com/s/seq-buRLTIFkqu2EAPfoDVBh?m=slm-4OgZoyHMmprl7LdDTIJ4 |
| pBRE1 | https://benchling.com/s/seq-ykNFHHaV7m1E8ORBfHLn?m=slm-cXwHxkoVVqvoS7HBlJse |
| pCVBp12 | https://benchling.com/s/seq-Hq82z6HXPJ172OOkI9Lc?m=slm-OJsKRAu4TgTdUmLyqeNd |
| pDOF1 | https://benchling.com/s/seq-N4SpuY16Zs8GttGdwBls?m=slm-b54eDByR6kaiuIyX4rO4 |
| pDREB1 | https://benchling.com/s/seq-7fAmR3xhXziwnjsd25OY?m=slm-cvgl5MtnNy7uSwalys7h |
| pDREB2 | https://benchling.com/s/seq-kf27C99GuU8wOc3EzfAp?m=slm-lXGDP2cRbp2HJckBt8Ew |
| pDREB2-ZmUbi1 | https://benchling.com/s/seq-TD5fXK41gOUJXps1xiA3?m=slm-gZKIeCKHdewH61B8OjFH |
| pDualLuc-*35S* | https://benchling.com/s/seq-elzbv9rmzIOLX0Z7N2Nr?m=slm-yJQeaA6mRVQ4ZDHlAeE3 |
| pDualLuc-Nos | https://benchling.com/s/seq-u5wil7EGkWkB8JNNwudh?m=slm-WhnJihUzchH5egyIUzfy |
| pEFS | https://benchling.com/s/seq-oPwaKkVBQoWcH2bPtGqW?m=slm-TX6UkTfjFxo49ssqf04Q |
| pEIN3 | https://benchling.com/s/seq-swdaXVh0CRDo42dYz9my?m=slm-E9bN7KhHneY5Gk7GW53Z |
| pGT3a | https://benchling.com/s/seq-aAZ4BrhfCyKhdYw49oaZ?m=slm-738xO8DT1a0GrJPDKPzM |
| pHAG1 | https://benchling.com/s/seq-B2EAERI1sTtqEg4amng6?m=slm-ICBODhq1AwINVeX6Urpv |
| pIPN2 | https://benchling.com/s/seq-Q6mY20wLg3hbcOniyd52?m=slm-oWckt6efGF9RzfciVID5 |
| pLBD2 | https://benchling.com/s/seq-ev22Ubl92DVWTDOFsuqR?m=slm-g79wFGa3hb0TNYLsQ2Ps |
| pLMI1 | https://benchling.com/s/seq-KcSZODmVbyAaXL44f7Wl?m=slm-J1jhCIEpEPl2xvvXcX2M |
| pMYC2 | https://benchling.com/s/seq-xs50vrNANfcsle9ZBWFv?m=slm-P7KuN1OKf5mfduEICpiz |
| pOpaque2 | https://benchling.com/s/seq-yp9GBuPLSVrfHnROFZVW?m=slm-0aVbVlI6Pv6JJDQnSL6c |
| pOsCBT | https://benchling.com/s/seq-AamuHHSCGlxaG8HJL2d5?m=slm-YzpITtVPWz1B1isDZIFe |
| pOsGRF1 | https://benchling.com/s/seq-EWHSTwdEJw1RuX0jhHT4?m=slm-sG63DEWVZbNzdwKhpKtK |
| pPTI4 | https://benchling.com/s/seq-7Kd4fwQVcn7ozVbrcmYU?m=slm-7a8NlOTMIUKsoXG9qEyL |
| pRen-*35S* | https://benchling.com/s/seq-n2Xdx9MT7PwAekvj2BaJ?m=slm-OFczZX5aaj45PbIGsR2a |
| pRF2a | https://benchling.com/s/seq-QtvnHqeaayU3efKeWXr9?m=slm-8dsILuwyfaOawm9VKKge |
| pSDG31 | https://benchling.com/s/seq-Q9bcGFZScL6VhTuTdm6M?m=slm-impam7UOTlt0zW2vmUKU |
| pSNAC1 | https://benchling.com/s/seq-x8MispCo89a7Xt63olse?m=slm-x5VdYCjB5oZXKgoC1LMv |
| pTBP1 | https://benchling.com/s/seq-xvVAE0mKixTQabsbmAVd?m=slm-z2y1tZDjNTyk1ERFBaea |
| pVP64 | https://benchling.com/s/seq-un9gakC4TY2o41seLbf1?m=slm-mPRpshMFwuOb5BFCcCGd |
| pVRN1 | https://benchling.com/s/seq-W320IrY6qyddeO89nld6?m=slm-o2cz2Pooiwl9IKy3uOGC |
| pWRKY50 | https://benchling.com/s/seq-c7EdSUEJZL9Npl1g1NE8?m=slm-h4oxln4BGxYSfQveAbEc |
| pZmUbi1-dCas9-24xGCN4 | https://benchling.com/s/seq-FjAPPyUNZMZsLAK7NJ8z?m=slm-UyTR5cy9mRRoIg9T0ALw |
| pZmVP1 | https://benchling.com/s/seq-MsTk4ikHQdAEAgcwNaJO?m=slm-mClYrRWp5u7OVJMZYPWU |

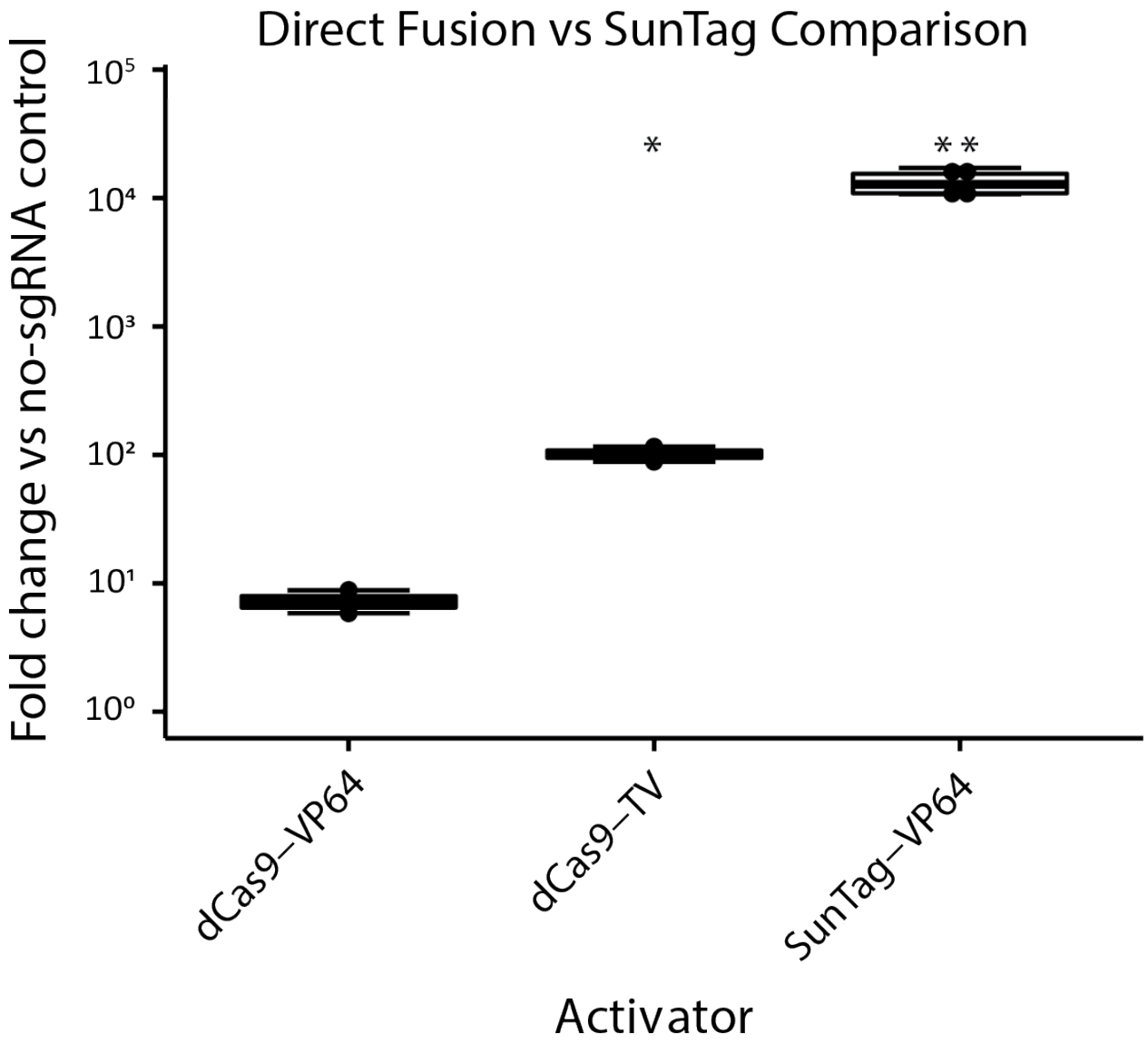
## Supplementary Note 7: Supplementary Table S4. Binary vector maps

| Name | Link |
|---|---|
| MT-AvrXa10-FT | https://benchling.com/s/seq-s8MMHPuv5O8ga4TWVYws?m=slm-zG7Srw29V6PMVhpTotUp |
| MT-DOF1-FT | https://benchling.com/s/seq-nVV28Wj4XlLoZvZzCZs1?m=slm-LttliC4AY8VRoTJ010CR |
| MT-DREB2-FT | https://benchling.com/s/seq-7rK40cVvEoHDMFPqs4LG?m=slm-UYlQe7OqoCpYZXcbzTJl |
| Off-target guide T-DNA | https://benchling.com/s/seq-2wVFIi9WBvRD5n8UfwAO?m=slm-IwbvmdT3J9cS2k9ia6eH |

## Supplementary Note 8: Supplementary Table S5. RT-qPCR primers

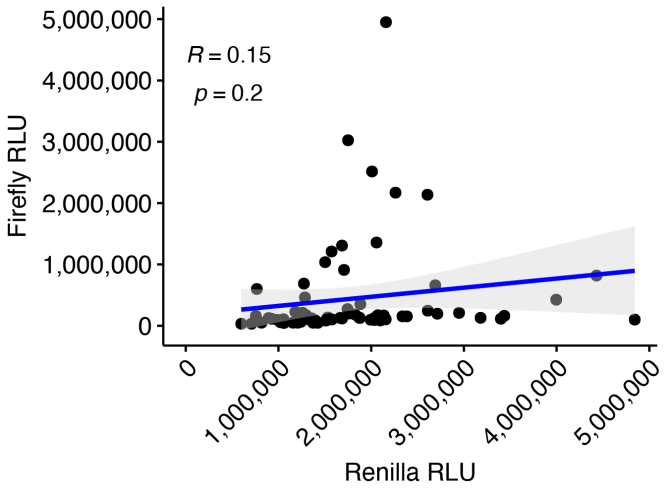
| Name | Sequence (5' to 3') |
|---|---|
| AtPP2A Fwd | AACGTGGCCAAAATGATGC |
| AtPP2A Rev | AACCGCTTGGTCGACTATCG |
| AtWUS Fwd | CAGCTTCAATAACGGGAATTT |
| AtWUS Rev | GTTGTAAGGTGCAGATGAGT |
| AtLEC1 Fwd | CAGAGAAACAATGGAACGTG |
| AtLEC1 Rev | TGAGACGGTAAGGTTTTACG |
| AtClv3 Fwd | GTTCAAGGACTTTCCAACCGCAAGATGAT |
| AtClv3 Rev | CCTTCTCTGCTTCTCCATTTGCTCCAACC |
| AtFT Fwd | CCCTGCTACAACTGGAACAAC |
| AtFT Rev | CACCCTGGTGCATACACTG |
| AtPAP1 Fwd | AGTATGGAGAAGGCAAATGGC |
| AtPAP1 Rev | CACCTATTCCCTAGAAGCCTATG |
| AtRD29a Fwd | GTTGGAGGAAGAGTCGGCTG |
| AtRD29a Rev | TGCTTCTCGTCGACAAGTCTC |
| AtKIN1 Fwd | ACAAGAATGCCTTCCAAGCC |
| AtKIN1 Rev | CGGTCTTGTCCTTCACGAAG |
| AtLEA7 Fwd | TGGTGAAACCAGAGGCAAGG |
| AtLEA7 Rev | CGTTTGTGCAGCTTGAGACG |
| AtDREB2a Fwd | CAAGAAGACTAAACACGAAAGCG |
| AtDREB2a Rev | ACAGTAGTACCGTCACCTCTAC |
| AtCEL3 Fwd | CATGGCTGCTAAGAGCAACG |
| AtCEL3 Rev | AGATGAAATTCTCAGCTGCTTGC |
| AtPME18 Fwd | GCCTGCCAGAGACGATCTC |
| AtPME18 Rev | GTATTGCTGTTCTCCGGTGC |
| AtFAS1 Fwd | ATCTGAAAGGCTTTCACCAGC |
| AtFAS1 Rev | AAGAAGTCCAGGAGGGTCAAG |

## Supplementary Note 9: Supplementary Figure 1



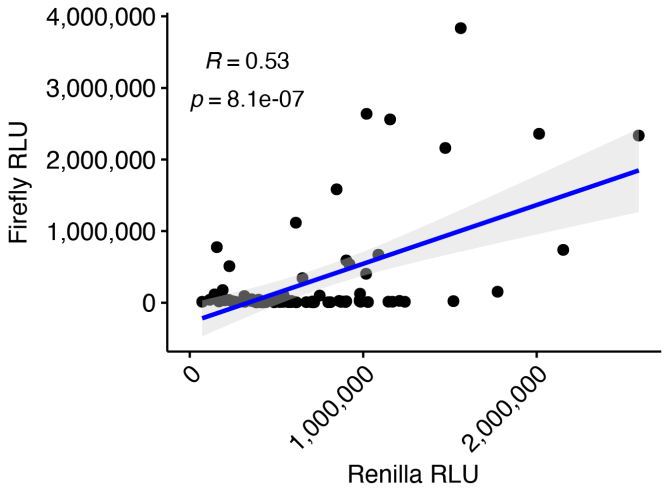

**Supplemental Figure 1. Direct fusion PTA vs SunTag PTA comparison in *A. thaliana* protoplasts**. The direct fusion class of PTAs containing dCas9-VP64 and dCas9-TV, along with the SunTag-VP64 activator, was transformed into protoplasts with a 4x *Wuschel* sgRNA array targeting the core promoter for gene activation. RNA was extracted and activation was quantified by RT-qPCR using the delta delta Ct method to generate the box-and-whisker plots. Parametric Welch's t-tests were performed comparing dCas9-TV or SunTag-VP64 with the dCas9-VP64 direct fusion at the same locus. A link to the sequence file of the 4X sgRNA repeat is given in Supplementary Table S2. All data is normalized against Ct values obtained using a no-sgRNA control (not shown). * corresponds to p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, and **** corresponds to p<= 0.0001.

## Supplementary Note 10: Supplementary Figure 2



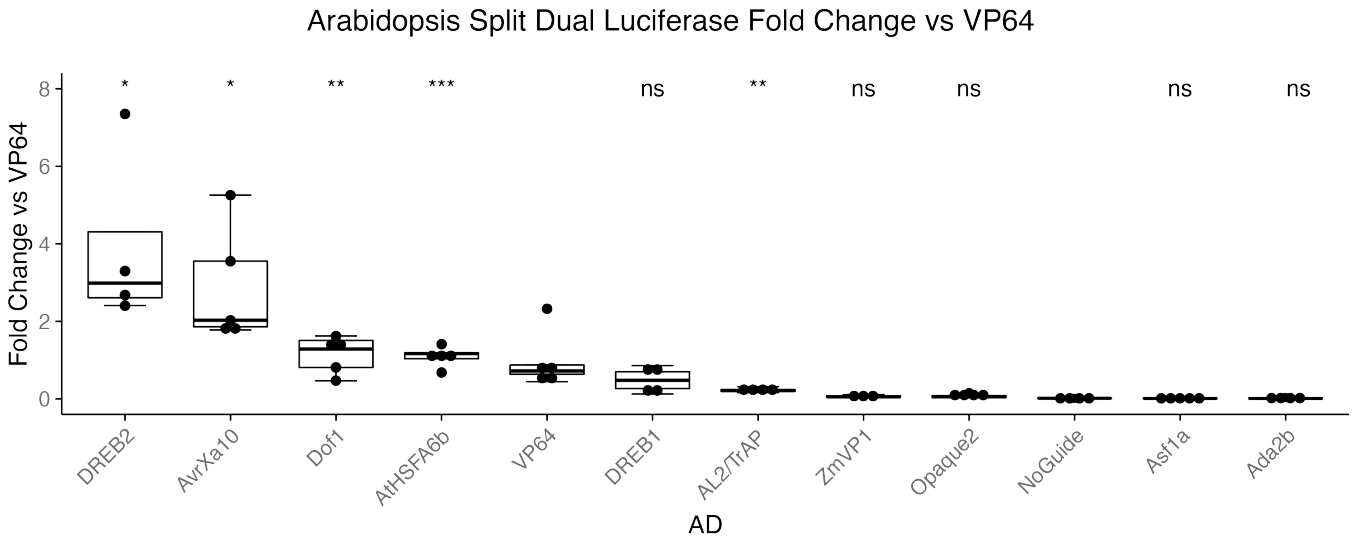

**Supplemental Figure 2. *35S* Batch One *S. viridis* Dual Luciferase Assay** The first protoplast isolation for *S. viridis* transformed with the library of scFv-ADs and the single-plasmid dual luciferase vector. Renilla luciferase is driven by a *35S* promoter. Data was fit to a linear regression model, with R and p values shown above.

## Supplementary Note 11: Supplementary Figure 3



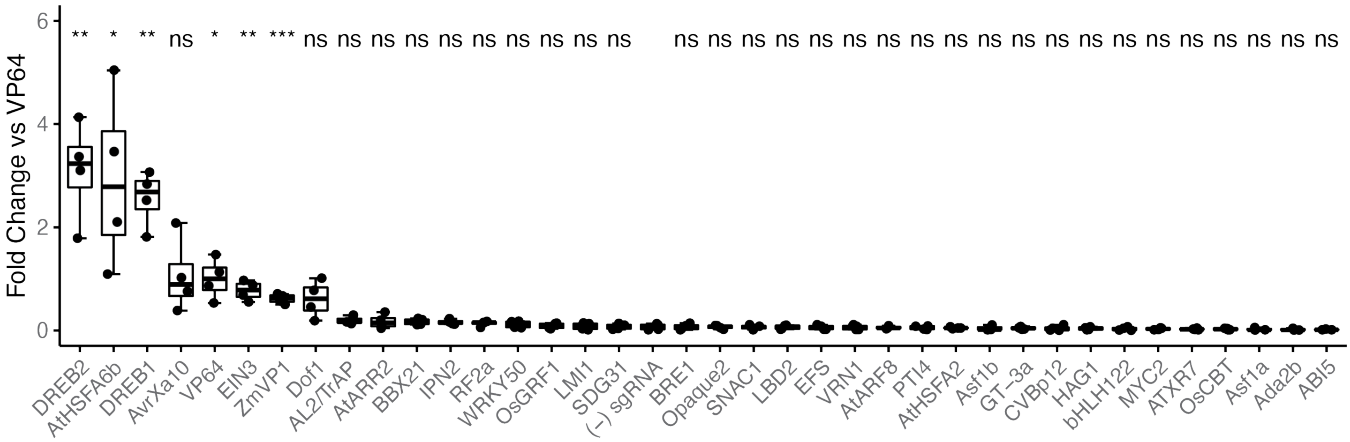

**Supplemental Figure 3. *35S* Batch Two *S. viridis* Dual Luciferase Assay** The second protoplast isolation for *S. viridis* transformed with the library of scFv-ADs and the single-plasmid dual luciferase vector. Renilla luciferase is driven by a *35S* promoter. Data was fit to a linear regression model, with R and p values shown above.

## Supplementary Note 12: Supplementary Figure 4



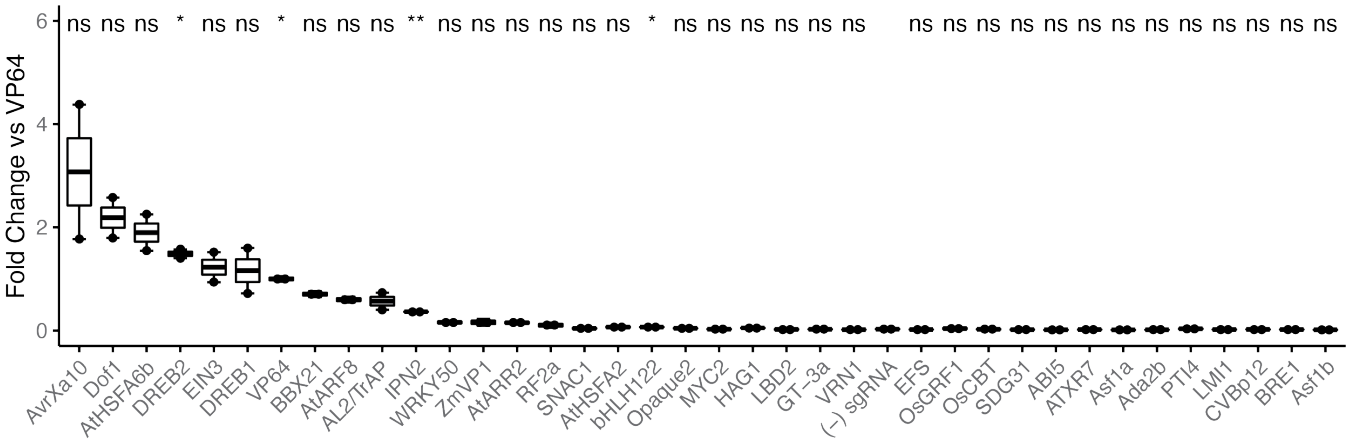

**Supplemental Figure 4. *A. thaliana* Split Dual Luciferase Assay** A smaller library of scFv-ADs was transformed into *A. thaliana* protoplasts along with the split dual luciferase vectors. The firefly luciferase vector contained the *6xLacO-mini35S* promoter, while the Renilla luciferase vector was driven by the constitutive *35S* promoter. Firefly luciferase RLUs were divided by Renilla luciferase RLUs to normalize for transformation. The scFv-AD normalized RLUs were then divided by the scFv-VP64 normalized RLU to generate the plot show above.

## Supplementary Note 13: Supplementary Figure 5



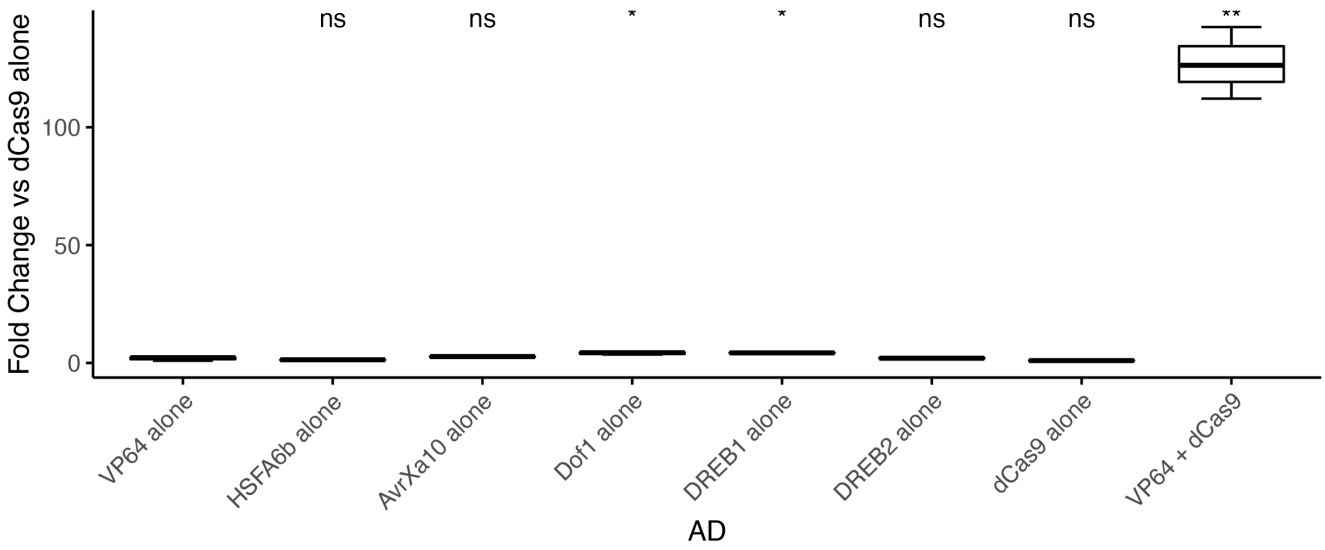

**Supplemental Figure 5. *S. viridis* Single-Plasmid Dual Luciferase Assay** The full library of scFv-ADs was transformed into *S. viridis* protoplasts along with the single-plasmid dual luciferase vectors. The single luciferase reporter contained the firefly luciferase driven by the *6xLacO-mini35S* promoter, while the Renilla luciferase was driven by the constitutive *35S* promoter. Firefly luciferase RLUs were compared to the positive control VP64. A parametric Welch's t-test was performed comparing the negative control (-) sgRNA with each respective AD. p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, and **** corresponds to p<= 0.0001.

## Supplementary Note 14: Supplementary Figure 6



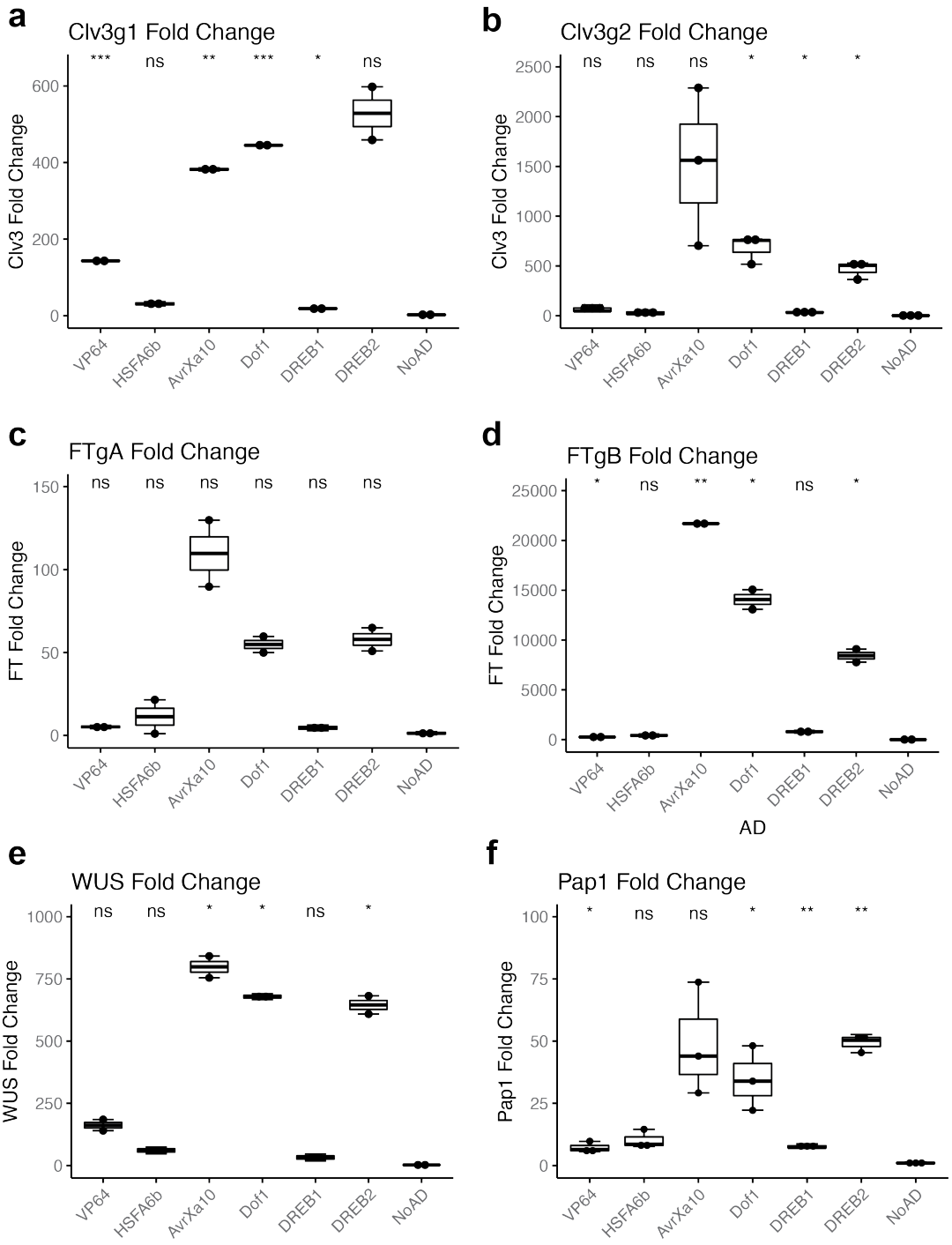

**Supplemental Figure 6.** ***A. thaliana* Single-Plasmid Dual Luciferase Assay** The full library of scFv-ADs was transformed into *A. thaliana* protoplasts along with the single-plasmid dual luciferase vectors. The single luciferase reporter contained the firefly luciferase driven by the *6xLacO-mini35S* promoter, while the Renilla luciferase was driven by the constitutive *35S* promoter. Firefly luciferase RLUs were compared to the positive control VP64. A parametric Welch's t-test was performed comparing the negative control (-) sgRNA with each respective AD. p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, and **** corresponds to p<= 0.0001.

## Supplementary Note 15: Supplementary Figure 7



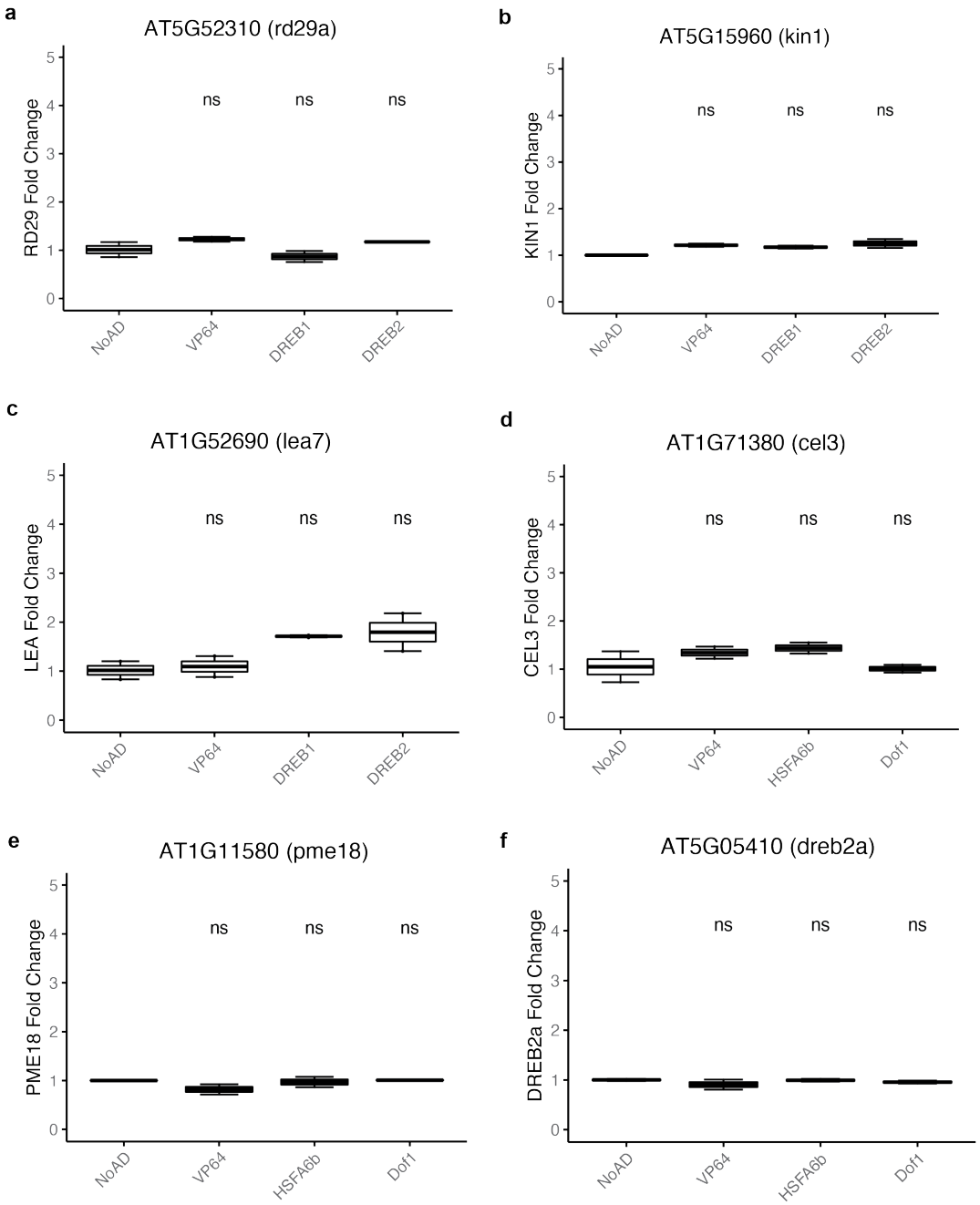

**Supplemental Figure 7. Non-specific Activation Domain strength in *S. viridis* protoplasts.** The non-specific activity of each strong plant-derived AD was measured by transforming *S. viridis* protoplasts with the indicated scFv-AD fusion and dual luciferase reporter plasmid. A parametric Welch's t-test was performed to compare the negative control 'VP64 alone' with each individual scFv-AD treatment. The positive control of 'VP64+ dCas9' serves as the positive control, in which the dual luciferase, sgRNA, dCas9, and scFv-VP64 were co-transformed into *S. viridis* protoplasts. A parametric Welch's t-test was performed comparing the negative control of 'VP64 alone' with each respective AD. * corresponds to p <= 0.05, ** corresponds to p <= 0.01, *** corresponds to p <=0.001, and **** corresponds to p<= 0.0001.

## Supplementary Note 16: Supplementary Figure 8



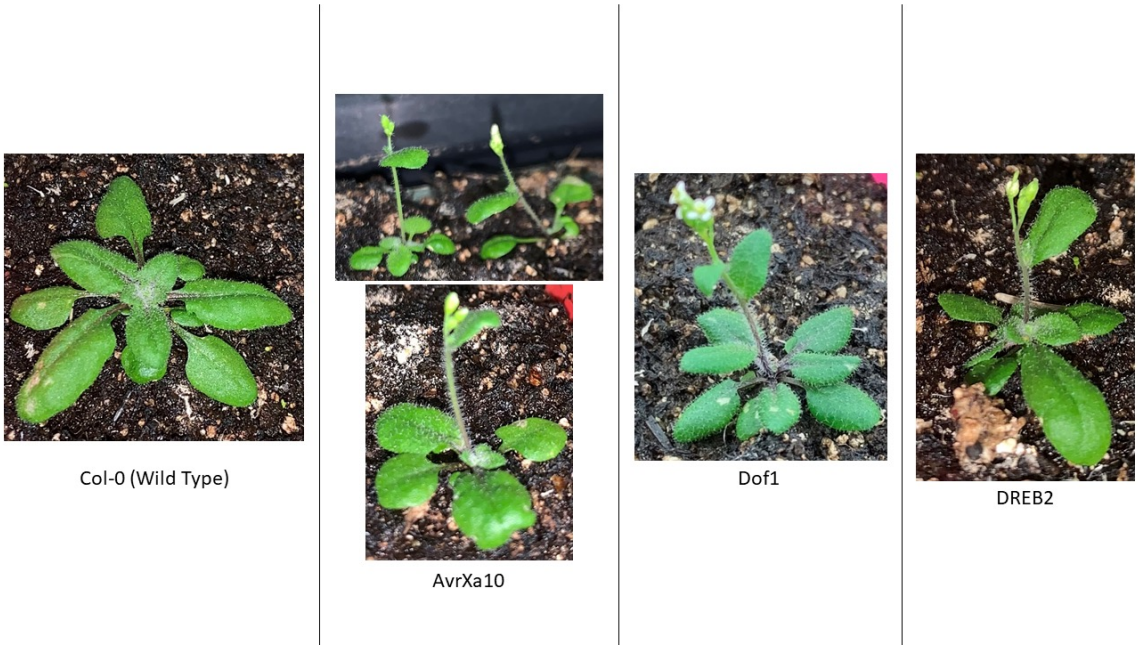

**Supplemental Figure 8. *A. thaliana* protoplast endogenous gene activation across all sgRNAs tested** Protoplasts were isolated from *A. thaliana* and transformed with Ubi10-dCas9-24xGCN, *35S*-scFv-AD, and AtU6-sgRNA constructs for the indicated target gene. RNA was isolated from cells 24h post-transformation, and RT-qPCR was performed to quantify gene expression. The delta delta Ct method was used to calculate Fold Change vs the NoAD negative control. *PP2A* was used as the housekeeping gene. Parametric Welch's t-tests to compare each AD with the NoAD control are denoted by asterisks, where * corresponds to p<= 0.05, ** corresponds to p<=0.01, *** corresponds to p<= 0.001, and **** corresponds to p<=0.0001.

## Supplementary Note 17: Supplementary Figure 9



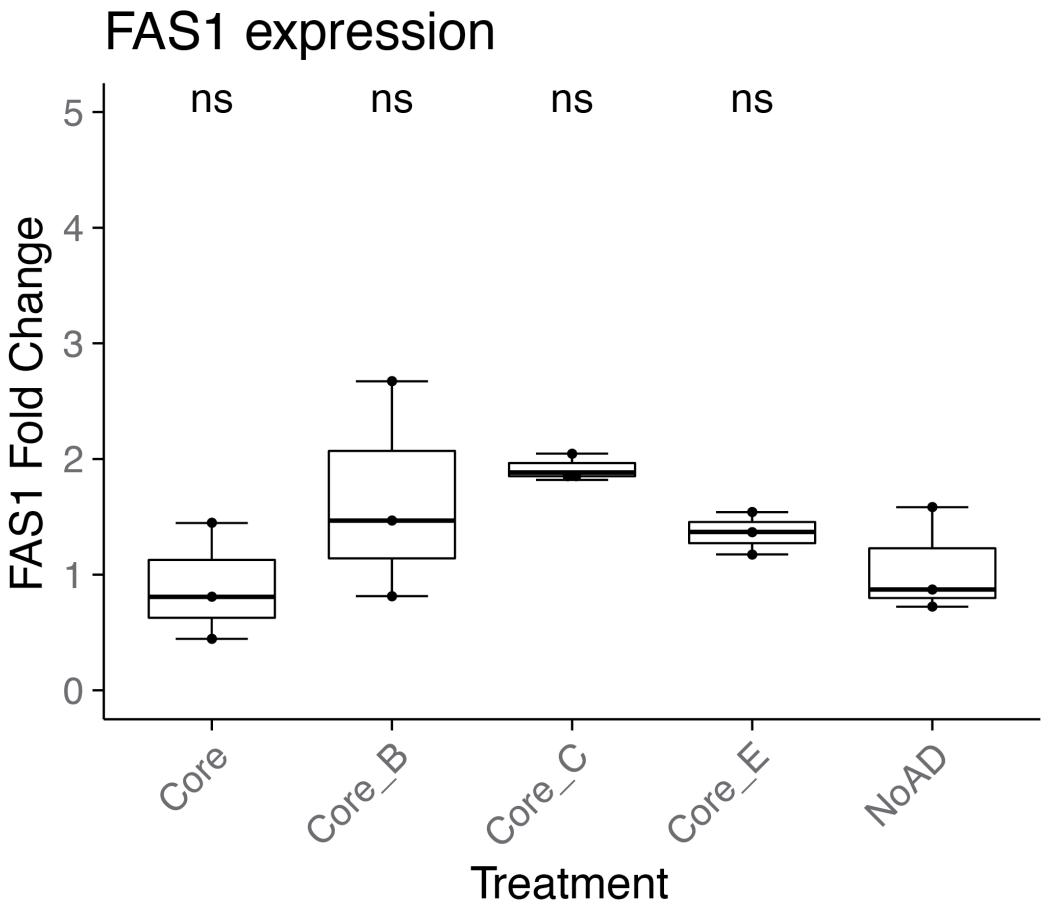

**Supplemental Figure 9. Off-target gene activation in *A. thaliana* protoplasts.** The off-target activity of strong plant-derived activation domains was quantified using RT-qPCR for downstream target genes of the endogenous transcription factors. A parametric Welch's t-test was performed to compare downstream gene expression for a given AD with the negative control "NoAD" in which protoplasts were transformed with dCas9-24xGCN4 and sgRNA but lacking the AD plasmid. Downstream target genes were identified by a literature search. The on-target guide in these experiments was the sgRNA targeting the WUS promoter region.

## Supplementary Note 18: Supplementary Figure 10




**Supplemental Figure 10. T1 *A. thaliana* Parental Lines** Images showing the T1 parental lines and an age-matched wild type Col-0 *A. thaliana* plant. The binary vectors for these plants are identical, except for the AD fused to NbGP41. Labels below each image denote the AD in each parental line. While a Col-0 line was used as a negative control side-by-side with the FT overexpression lines in this experiment, we observed the same phenotype with an age-matched no-sgRNA control(37).

## Supplementary Note 19: Supplementary Figure 11




**Supplemental Figure 11. Off-target activation of nearby FAS1 gene in A. thaliana protoplasts.** The off-target activity of FAS1 (AT1G65470) was quantified by RT-qPCR in *A. thaliana* protoplasts. The on-target guide in these experiments are denoted below, in which the FT locus was targeted with sgRNAs to activate both the core promoter and nearby enhancers Block B, C, and E as described in other figures. Little FAS1 expression change was measured when comparing across groups using a parametric Welch's t-test to compare each treatment with the FT Core + No AD negative control.