www.sciencemag.org/cgi/content/full/328/5974/85/DC1

Supporting Online Material for

Orchestration of Floral Initiation by APETALA1

Kerstin Kaufmann, Frank Wellmer, Jose M. Muiño, Thilia Ferrier, Samuel E. Wuest, Vijaya Kumar, Antonio Serrano-Mislata, Francisco Madueño, Pawel Krajewski, Elliot M. Meyerowitz, Gerco C. Angenent, José Luis Riechmann*

*To whom correspondence should be addressed. E-mail: jriechma@caltech.edu

Published 2 April 2010, *Science* **328**, 85 (2010)
DOI: 10.1126/science.1185244

This PDF file includes:

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/328/5974/85/DC1)

Tables S1 to S3, S5, S6, and S9 as zipped Excel files

SUPPORTING ONLINE MATERIAL

Materials and Methods

Plant growth. Plants were grown on a soil:vermiculite:perlite (2:1:1) mixture at 20ºC under constant illumination with cool white fluorescent light, unless indicated otherwise.

Tissue collection for gene expression profiling experiments. For all experiments, we used approximately 4 week-old 35S:AP1-GR *ap1*-1 *cal*-1 plants (*S1*). For each sample, inflorescence tissue from ~25 plants was collected using jeweler's forceps as previously described (*S1*). At least four biologically independent sets of samples were generated for each experiment. For the time-course experiments, we treated inflorescences with a solution containing 10 μM dexamethasone (Sigma-Aldrich), 0.01% (v/v) ethanol, and 0.015% (v/v) Silwett L-77 (De Sangosse), or with an identical mock solution that lacked dexamethasone. Using plastic pipettes, the solutions were directly applied onto the inflorescences so that the cauliflower-like structures were completely drenched. We then collected inflorescence tissue after 2, 4, 8, and 12 hours, as well as from untreated plants (0 hour time point). For cycloheximide experiments, we treated inflorescences of 35S:AP1-GR *ap1*-1 *cal*-1 plants with a dexamethasone-containing or a mock solution (as described above), or with identical solutions that contained in addition 10 μM cycloheximide (Sigma-Aldrich). Tissue was collected 3 hours after the treatment. For qRT-PCR experiments, tissue was collected from untreated control plants, and from plants 12 hours after treatment with either a dexamethasone-containing or a mock solution (as described above).

RNA preparation. Total RNA was isolated from tissue samples using the Plant Total RNA kit (Sigma-Aldrich) according to the manufacturer's instructions. Quality of RNA samples was evaluated using a Bioanalyzer and a RNA Nano 6000 kit (Agilent). RNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific). Total RNA preparations used for qRT-PCR analysis were treated with DNase I (Ambion) to reduce contaminations with genomic DNA.

Microarray setup and experiments. Two different microarray platforms were used in the study: microarrays that were based on the *Arabidopsis* Genome Oligo Set Version 1.0 and on the *Arabidopsis* Genome Oligo Set Version 1.0 Upgrade (Operon, Huntsville, AL), and hereafter referred to as "Operon microarrays"; and custom microarrays manufactured by Agilent (Agilent Technologies, Santa Clara, CA), and hereafter referred to as "Agilent microarrays". The Operon microarrays were prepared as previously described (*S2*), and consist of a total of 30,194 70-mer oligonucleotide probes that correspond to 26,215 annotated genes according to The *Arabidopsis* Information Resource (TAIR) genome annotation version 8. Agilent microarrays (version 1) were designed using the *e*Array software pipeline (https://earray.chem.agilent.com/earray/) and TAIR genome annotation v7, and contain 26,566 60-mer probes corresponding to the same number of annotated genes (pseudogenes, transposons, retrotransposons, and pre-tRNAs were excluded from the design). A second-generation *Arabidopsis* custom *e*Array was designed using TAIR genome annotation v8, and contains probes corresponding to 28,327 annotated genes. The same sets of RNA samples were analyzed using both the Operon and Agilent microarrays.

Operon microarrays were used as previously described (*S1*), except that the Amino Allyl MessageAmpII aRNA Amplification Kit (Applied Biosystems/Ambion, Austin, TX)

was used for labeling of RNA samples with Cy3 and Cy5 fluorescent dyes, following manufacturer's instructions. Microarrays were scanned with both Axon GenePix 4200A and Axon 4000B scanners, using the *Gene Pix* 5.0 analysis software (Molecular Devices, Sunnyvale, CA).

Agilent microarrays were used following manufacturer's instructions. RNA samples were labeled with fluorescent dyes using the Quick Amp Labeling Kit (Agilent). Microarray hybridizations (65ºC, 16h) and washes were performed with Agilent reagents and following standard protocols. Microarrays were scanned using an Agilent DNA Microarray Scanner G2565CA, and data were acquired using Agilent's Feature Extraction Software version 9.5.3.1. Agilent version 1 custom microarrays were used for the time series experiments, whereas version 2 microarrays were used for the cycloheximide experiment (see below).

Four independent sets of biological samples were used for the experiments, and sample setup and pairing for hybridizations was the same for Operon and Agilent microarrays. In each of the biological replicates of the time course experiments, all the samples derived from dexamethasone (Dex)-treated plants were labeled with one dye (i.e., Cy3), and all the samples derived from the corresponding Mock-treated plants were labeled with the alternative dye (i.e., Cy5). The dyes used for labeling RNA from a given treatment type (Dex and Mock) were switched for two of the replicate experiments, to reduce dye-related artifacts. Dex- and Mock-derived samples for each timepoint and biological replicate were co-hybridized. This experimental setup resulted in a total of 5 hybridizations per set (0h, 2h, 4, 8h, and 12h; Dex vs. Mock at each timepoint), and two biological replicate sets labeled with each dye polarity (Mock-Cy3/Dex-Cy5, and vice versa). In the cycloheximide (Chx) experiments, samples from each biological replicate resulted in a set of four

hybridization pairs: Mock vs. Dex, Mock vs. Chx, Mock vs. Dex+Chx, and Chx vs. Dex+Chx. Dye polarities were switched between biological replicates. An additional microarray experiment, a comparison between *ap1-1* mutant and wild-type inflorescences, was also carried out using Agilent version 1 microarrays. The samples used in this experiment corresponded to the inflorescence meristem and floral buds, and have been described previously (*S2*). Three sets of biological replicates were used, and dye polarities were switched between them.

Microarray data analysis. Raw data from the Operon microarrays, and feature extraction software pre-processed data from the Agilent Microarrays, were imported into the Resolver gene expression data analysis system version 7.1 (Rosetta Biosoftware, Seattle, WA) and processed as described (*S1*). Resolver uses a platform-specific error model-based approach to stabilize the variance estimation to improve the specificity and sensitivity in differential gene expression detection (*S3*). The data from the four biological replicates of each condition were combined, resulting in an error-model weighted average of the four. The P-values for differential expression calculated by Resolver were adjusted for multi-hypothesis testing using the Benjamini & Hochberg procedure, as implemented in the Bioconductor *multtest* package in R (http://www.bioconductor.org/packages/bioc/stable/src/contrib/html/multtest.html). Genes for which the Benjamini & Hochberg-adjusted P-value was <0.05 in at least one of the comparisons (i.e., timepoints, for the time-course experiment) and microarray platforms used (Agilent and/or Operon) were initially considered differentially expressed in response to AP1 activation (Table S1). In the time course experiment, genes that were differentially expressed at timepoint 0h were excluded from the analysis. A further selection was

implemented by applying an absolute fold-change (FC) cutoff of 1.8, and resulted in the identification of a total of 1,370 genes that showed robust expression changes during the time-course experiment (Table S1). Genes that were detected as differentially expressed in the Agilent time-course experiment (1017 genes) were subjected to cluster analysis using the k-means algorithm implemented in Resolver (Fig. 1B; Table S9). Data from the cycloheximide experiment, in which Agilent version 2 microarrays were used, were processed in Resolver in the same manner, using a Benjamini & Hochberg-corrected P-value<0.05 and an absolute FC>1.8 as cut-offs for differential gene expression. The analysis of the data from the Mock vs. Dex, Mock vs. Chx, and Mock vs. Dex+Chx hybridizations revealed that the treatment with Chx caused considerable changes in gene expression (a total of 10,270 probes reported differential expression upon the Chx treatment, 9,257 probes upon the Chx+Dex treatment, and 681 probes upon Dex treatment). The data also indicated that the effects of Dex and Chx were largely additive rather than synergistic (data not shown). As a result, additional Chx vs. Dex+Chx hybridizations were performed. With these datasets, genes that responded to AP1 activation in the presence of Chx (and are therefore potential direct AP1-targets), were identified using three different methods. First, genes were selected if they were detected as differentially expressed upon both the Dex and Dex+Chx treatments but not upon the Chx treatment. Second, the Ratio-Split and Re-Ratio functions of Resolver 7.1 were used to computationally generate Chx vs. Dex+Chx comparisons, from which differentially expressed genes were selected with the same cutoffs indicated above. Last, differentially expressed genes were also identified from the direct Chx vs. Dex+Chx hybridizations. Combining these three strategies, a total of 1,578 genes were identified as potential AP1 targets (Table S5). Data from the comparison between *ap1*

mutant and wild-type inflorescences were also processed in Resolver.

qRT-PCR experiments. cDNA synthesis was performed using DNase I-treated total RNA preparations (see above), oligo-dT primers and the RevertAid H Minus M-MuLV reverse transcriptase (Fermentas). Relative transcript abundance of selected genes (see Table S11 for a list of genes and the primers used) was determined using the Roche LightCycler 480 system and the LC480 SYBR Green I Master kit (Roche Applied Sciences). Measurements were taken for three biologically independent sets of samples. In addition, all samples were analyzed twice. LightCycler melting curves were obtained for the reactions, revealing single peak melting curves for all amplification products. The amplification data were analyzed using the second derivative maximum method, and resulting Cp values were converted into relative expression values using the comparative Ct method (*S4*). Two reference genes ('REF1' and 'REF2'; see Table S11) (*S5*) were used to normalize the data. Their Cp values were averaged for each sample.

Antibodies. The coding region of AP1 was cloned as *Nde*I fragment into a pET-19b vector (Novagen, Madison, WI) and expressed in *E. coli* BL21 as a poly-histidine-tagged fusion protein. The recombinant proteins were purified under denaturing conditions using a guanidinium HCl-containing buffer for cell lysis on Ni-NTA agarose (Qiagen) according to the manufacturer's instructions, and then resolved by preparative PAGE. A band corresponding to full-length AP1 was eluted and used to immunize rabbits. The resulting AP1 antiserum was affinity-purified using a C-terminal fragment of AP1 (which lacked the MADS domain) fused to glutathione S-transferase, which had been purified under native conditions. For ChIP-qPCR validation experiments, a commercial, affinity-purified peptide

antibody against GR (Affinity Bioreagents, PA1-516) was used for independent immunoprecipitations.

Chromatin immunoprecipitation. For ChIP experiments, inflorescences of approximately 4 week-old 35S:AP1-GR *ap1*-1 *cal*-1 plants were treated with a dexamethasone-containing solution (as described above). Inflorescence material from treated and untreated plants (0.8 g) was collected 2 hours after the treatment and fixed for 30 min in 1% (v/v) formaldehyde and then quick-frozen in liquid nitrogen. The chromatin immunoprecipitation experiments and ChIP-Seq sample preparation were performed as described in (*S6*). One microgram of affinity-purified AP1 antibody (see above) was used for each ChIP experiment. Two independent sets of biological samples were used in two replicate experiments.

Illumina ultra high-throughput sequencing and data analysis. ChIP-Seq libraries were sequenced using an Illumina Genome Analyzer II (Fasteris SA, Plan-les-Quates, Switzerland). Thirty-six cycles of sequencing were performed using the 36 Cycle Sequencing Kit v3. One or two flow-cell channels were run per sample, generating a total of more than 7 million sequence reads per sample. The raw sequence data were processed using the Illumina sequence data analysis pipeline GAPipeline1.3.2. Sequences in FASTQ format were mapped to the unmasked *Arabidopsis* genome (ATH1.1con.01222004; ftp://ftp.arabidopsis.org/) using the SOAP program (*S7*) (Table S10). A maximum of two mismatches and no gaps were allowed, and reads were iteratively trimmed from the 5' end until mapped or their length fell below 31 nt. Only uniquely mapped reads were retained. Sequences mapping to the plastid and mitochondrial genomes were eliminated. Uniquely mapped reads were extended directionally to 300 bp, and enrichment of sample compared to

control was calculated using a score based on the Poisson distribution as described previously (*S8*) for each biological replicate independently. False discovery rate (FDR) thresholds were estimated by permutation of reads between sample and control. Afterwards, both biological replicates were combined conservatively by taking their minimum score at each nucleotide position. The most conservative threshold of both replicates was used to control the FDR.

To analyze the AP1 DNA binding regions for CArG box enrichment, perfect matches with CArG box ($CC[W]_6GG$) and CArG box-like ($C[W]_7GG$ and $CC[W]_7G$) consensus sequences were located in the 1,000-bp region around the maximum peak score position. The sequences of the regions were extracted from the *Arabidopsis* reference genome (ATH1.1con.01222004; ftp://ftp.arabidopsis.org/). In order to show the dependency of the enrichment on CArG boxes on the peak score value, only a region of 250-bp around the maximum peak score position was considered. For each score threshold considered, the proportion of significant peaks containing at least one CArG box was estimated. A control-set of sequences was obtained by randomly permuting the nucleotide positions for each original 501-bp sequence, and CArG-box motif consensus sequences were located on the permuted sequences. This procedure was repeated 100 times to obtain the mean and the standard deviation of the proportion of significant peaks containing at least one CArG-box motif when the nucleotide sequences are permuted.

All ChIP-Seq peaks were characterized by the relative position of the peak center (maximum score position) with respect to the 5' and 3' end of all annotated genes (as described in the TAIR8_GFF3_genes.gff file, ftp://ftp.arabidopsis.org/). Genes with at least one AP1 ChIP-Seq peak located within the region from 3 kb upstream of the 5'-end to 1kb

downstream of the 3'-end of the annotated transcript were considered as putative AP1 direct target genes (a total of 2,298 genes). In the majority of cases, the 5'-end of the annotated transcript marked the transcriptional start site (TSS). Out of 28,244 genes encoded in the *Arabidopsis* nuclear genome (TAIR8), 18,937 (67%) were annotated with a 5' UTR, 19,893 (70%) with a 3' UTR, and 18,331 (65%) with both 5' and 3' UTRs. For the 2,298 genes identified as putative direct AP1 targets, the percentages were 79.7 % (5' UTR), 82.5 % (3' UTR), and 78.2 % (5' and 3' UTR).

In order to estimate the level of enrichment of AP1 binding events in a region close to the start of genes, the method of Storey was applied (*S9*). In brief, given a set of values coming from two distributions, a uniform distribution representing the background level, and an unknown non-uniform distribution, representing the enriched positions, the proportion of values coming from each of these distribution is calculated separating the two densities.

To determine the overlap between the AP1 binding sites identified in this study and those determined for SEP3 (*S8*), AP1 and SEP3 sites in each common putative direct target gene were paired as follows. For each gene, the relative peak position regarding the 5' end of the gene (if the peak is upstream of the gene), the 5’ end weighted by the length of the gene (if the peak is located inside the gene), or the 3’ end (if the peak is downstream) was used to unequivocally associate significant (FDR<0.001) AP1 peaks with their closest significant (FDR<0.001) SEP3 peaks. When two or more AP1 peaks had the same closest SEP3 peak, the SEP3 peak was assigned to the closest AP1 peak, and the other AP1 peaks were assigned to the next closest SEP3 peak (if any). In addition, the overlap between the

AP1 and SEP3 binding sites was also determined considering all possible AP1-SEP3 peak pairs per target gene, i.e., in a manner that was not peak-position directed (Fig. S10).

ChIP-qPCR experiments. qPCR experiments were performed on immunoprecipitated DNA that was obtained using the GR antibody (see above) and tissue collected from untreated control plants, and from plants 2 and 12 hours after treatment with a dexamethasone-containing solution (as described above). qPCR experiments were done as two or more technical replicates using a Bio-Rad MyIQ Cycler and Bio-Rad iQ SYBR Supermix. Primers used for PCR amplification were designed surrounding the position of maximum height in the corresponding peak (i.e., the peak center) identified in the ChIP-Seq experiments, using the Primer3 program (http://frodo.wi.mit.edu/primer3/) (Table S12). The qPCR data were analyzed using the Biogazelle qBasePlus version 1.3 program. Two reference genes were used for normalization (see Table S12).

Analysis of nuclear AP1-GR accumulation. Inflorescence tissue was collected from *ap1*-1 *cal*-1 double mutants and from 35S:AP1-GR *ap1*-1 *cal*-1 plants at different time points after treatment with a dexamethasone-containing or a mock solution (as described above). The tissue was immediately frozen in liquid nitrogen. For each sample, nuclei were extracted by grinding 500 mL tissue in liquid nitrogen, followed by addition of 250 mL M1 buffer (10 mM Na phosphate pH7.0, 100 mM NaCl, 1 M hexylene glycol (Sigma), 10 mM 2-mercaptoethanol (Sigma), 1x Complete Protease Inhibitor Cocktail (CPIC; Roche)). The tissue was then ground again on ice and 650 mL M1 buffer were added. Subsequently, the cell suspension was filtered twice through Miracloth (Calbiochem) and the extract was centrifuged at 7,600x*g*, 5ºC for 1 min to pellet the nuclei. The supernatant was removed and the pellet washed in 900 mL M2 buffer (10 mM Na phosphate pH7.0, 100 mM NaCl, 10

mM $MgCl_2$, 0.5% (v/v) Triton X-100 (Sigma), 1 M hexylene glycol, 10 mM 2-mercaptoethanol, 1x CPIC), followed by centrifugation at 7,600x*g*, 5ºC for 1 min. Washes with M2 buffer were repeated twice. Subsequently, the pellet was resuspended in 900 mL M3 buffer (10 mM Na phosphate pH7.0, 100 mM NaCl, 10 mM 2-mercaptoethanol, 1x CPIC) and centrifuged at 2,000x*g*, 5ºC for 5 min. The pellet was then resuspended in 20 mL ChIP dilution buffer (16.7 mM Tris-HCl pH8.1, 167 mM NaCl, 1.2 mM EDTA, 1.1% (v/v) Triton X-100, 0.01% (w/v) SDS, 1x CPIC) and sonicated at high intensity for 5 cycles of 30 sec using a Bioruptor (Diagenode). The resulting nuclear extract was centrifuged at 20,000x*g*, 5ºC for 10 min to pellet membranes and insoluble proteins. After centrifugation, equal amounts of protein from the supernatants of the different samples were separated on 12% SDS-PAGE gels, followed by transfer to a PVDF membrane and staining with Ponceau S. AP1-GR was detected using a polyclonal antiserum raised in rabbit against the rat glucocorticoid receptor (Affinity Bioreagents #PA1-516) at a final concentration of 2 mg/mL in PBS-T (1xPBS with 0.05% (v/v) Tween 20) supplemented with 5% (w/v) milk powder. Histone 2A/B was detected using a polyclonal antiserum raised in rabbit (a gift from Dr. Thomas Kretsch, University of Freiburg), which was diluted 10,000-fold in 5% (w/v) milk PBS-T. The immunoblots were developed using horseradish peroxidase-coupled anti-rabbit IgG antibodies (Sigma; diluted 50,000-fold in PBS-T) and the ECL+ detection kit (GE) or the ECL kit (Pierce) for AP1-GR and histone 2A/B, respectively, according to the manufacturers' instructions.

Analysis of binding sites in *TFL1*. The lines SAIL351H08 and SAIL566G09 (accession: Columbia) were obtained from the European Arabidopsis Stock Centre (http://arabidopsis.info/) and contain T-DNA insertions 418 and 1615 bp, respectively,

downstream of the stop codon of *TFL1* (in the intergenic region; the 3' untranslated region of *TFL1* is 118 nucleotides long). The genotype of these lines was confirmed by PCR. The *tfl1-1* mutant allele (accession: Columbia) has been previously described (*S10*). All plants were grown under standard long day conditions (16h light, 8h dark) at 22ºC.

For complementation assays, we used different *TFL1* constructs. pASM4 has been previously described (*S11*) and contains 1035 bp of *TFL1* coding sequence, 2195 bp of genomic sequence upstream of the start codon and 4613 bp downstream of the stop codon of *TFL1*. It was made by ligating a 7843 bp *Xho*I/*Eco*RI fragment of Landsberg *erecta* genomic DNA into the *Sal*I/*Eco*RI sites of the binary vector pBIN19. The pASM6 and pASAM8 constructs contained fragments including the *TFL1* coding sequence and the genomic sequence upstream of the start codon, as in pASM4, but the sequences downstream of the stop codon were reduced to 432 bp in case of pASM6 and 1625 bp in case of pASM8. Both constructs were derived from pASM4, by replacing the 3' sequence downstream of the stop codon with shorter 3' fragments. These fragments were generated by PCR amplification and were subcloned using naturally occurring restriction sites in the *TFL1* genomic sequence. Full details of the cloning strategy are available upon request (madueno@ibmcp.upv.es). All constructs were transformed into homozygous *tfl1*-1 plants using the floral dip method (*S12*). For each genomic construct, several independent transgenic lines were obtained and analyzed (7 for pASM4, 31 for pASM6 and 65 for pASM8).

Figures and legends

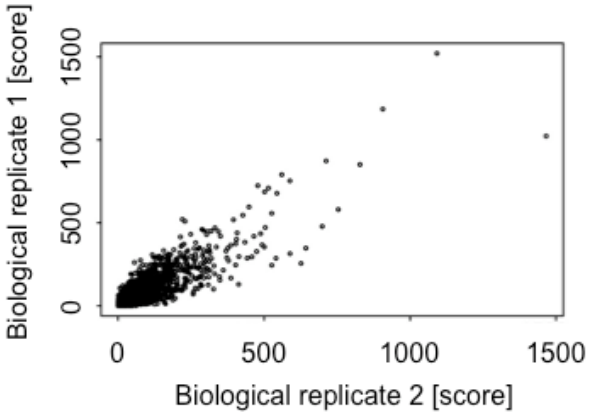

Figure S1: Reproducibility of ChIP-Seq experiments. Average read-enrichment single-nucleotide position scores for non-overlapping 1,000 bp regions across the *Arabidopsis* genome were plotted for the replicate ChIP-Seq experiments. The Pearson correlation coefficient for the data is 0.9 (and 0.84 for non-overlapping 100 bp regions), indicating high reproducibility of the experiments.

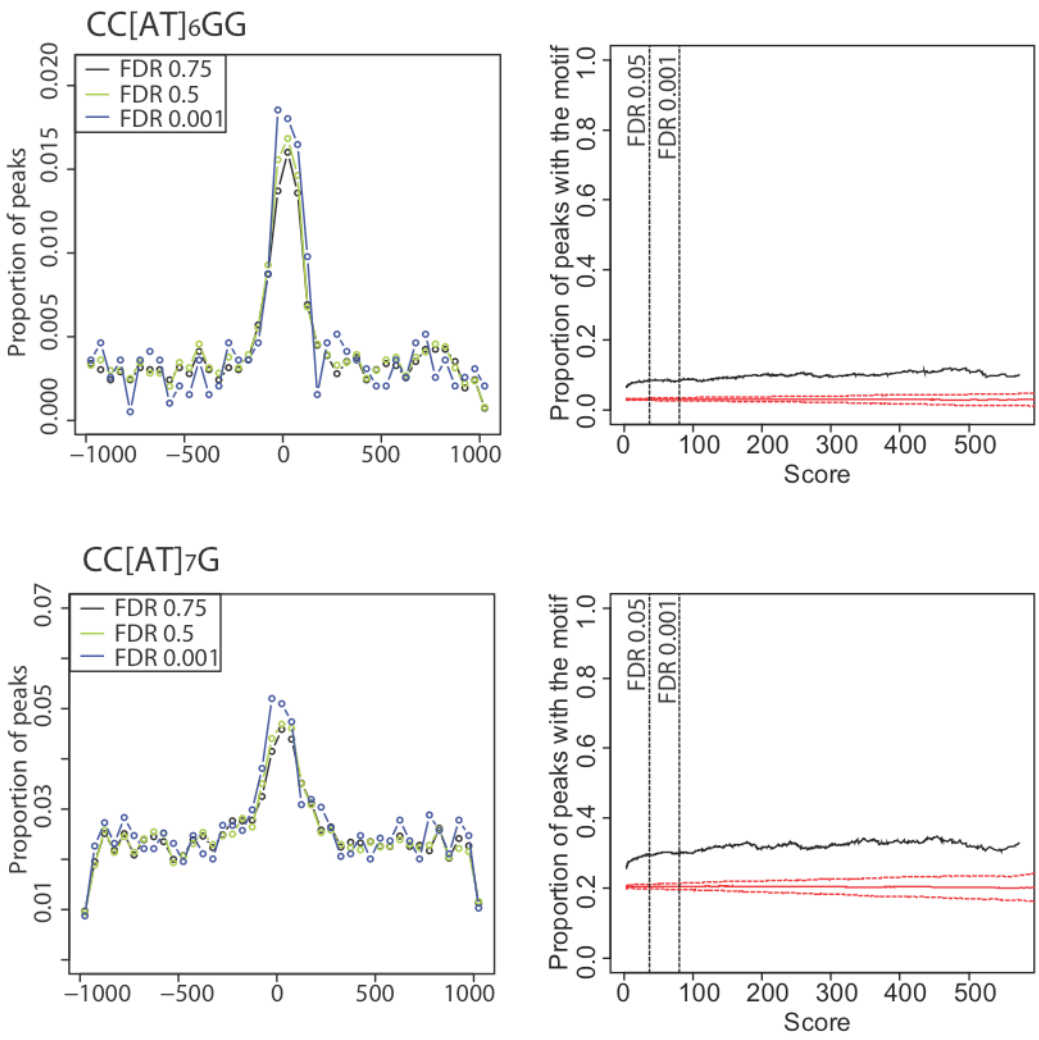

Figure S2: CArG box enrichment in AP1-binding genomic regions. Left panels: location of perfect CArG boxes ($CC[AT]_6GG$) (upper panel) and of CArG box-like sequences of type $CC[AT]_7G$ (lower panel) in genomic fragments identified by ChIP-Seq, using different FDR cut-offs (1942 regions were identified with FDR<0.001). CArG box and CArG box-like sequences were predominantly found in the center of the genomic regions identified by ChIP-Seq (at the location of the peak maximum, or 0 position), and this positional enrichment increased with decreasing FDR cut-off values. Right panels:

distribution of CArG box-like sequences in the genomic regions bound by AP1. Black line: the proportion of genomic regions (defined as a 501 bp region centered around the location of the peak maximum) with CArG box (upper panel) and CArG box-like sequences (lower panel) at different peak score cut-off values (as determined in the ChIP-Seq analysis). Red solid line: permutation-derived background distribution of CArG box and CArG box-like sequences. Dashed red lines indicate one standard deviation from the mean.

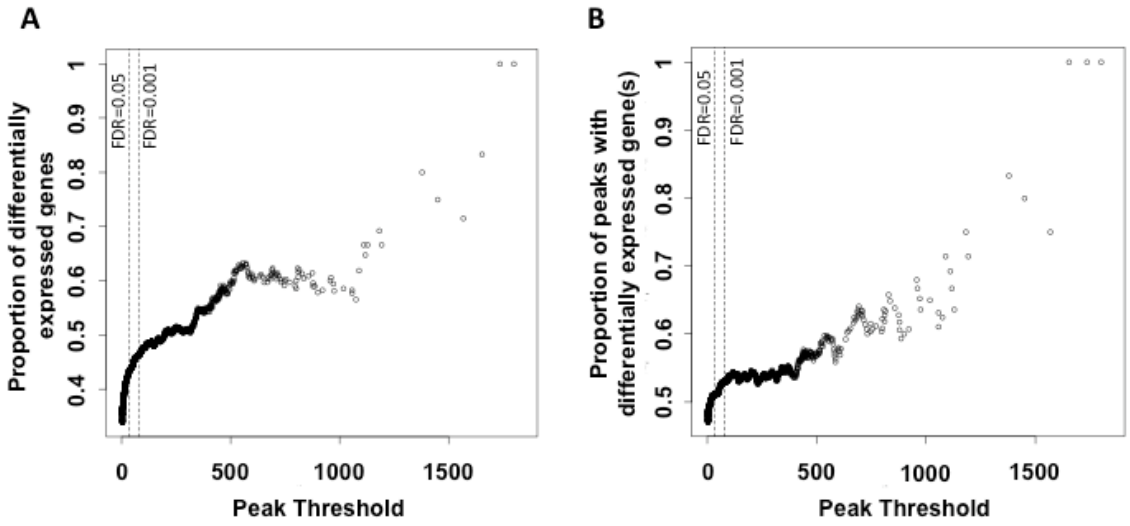

Figure S3: Relationship between AP1 binding events and target gene regulation. (A) Approximately half of the genes identified through ChIP-Seq showed changes in gene expression after AP1 activation. The graph displays the proportion of differentially expressed genes among genes with a significant AP1 ChIP-Seq peak, at different threshold values (indicated on the x-axis). Genes identified through ChIP-Seq were those that showed at least one AP1 binding peak within 3 kb upstream of the 5' end and 1 kb downstream of the 3' end of the annotated gene. With an FDR<0.001, used in this study to define the set of potential AP1 direct targets, a total of 2,298 genes were identified. Of those, 2,183 genes

were represented in the arrays used in this study, and 46.3% of them (1011 genes) were detected as differentially expressed (Agilent and Operon datasets, Benjamini & Hochberg-adjusted P-value <0.05; table S1). **(B)** Approximately half of the AP1 binding sites identified through ChIP-Seq are near genes that showed changes in gene expression after AP1 activation. The graph displays the proportion of AP1 binding sites identified through ChIP-Seq, at different threshold values, that were near a gene that was detected as differentially expressed in the microarray experiments. With an FDR<0.001, used in this study, 1942 AP1 binding peaks were identified, and 1030 (53%) of them were located in the proximity of at least one differentially expressed gene (as defined in A). If only ChIP-Seq peaks that are in the proximity of annotated genes are considered (1,678 of the 1,942), approximately 62% of them are nearby differentially expressed genes.

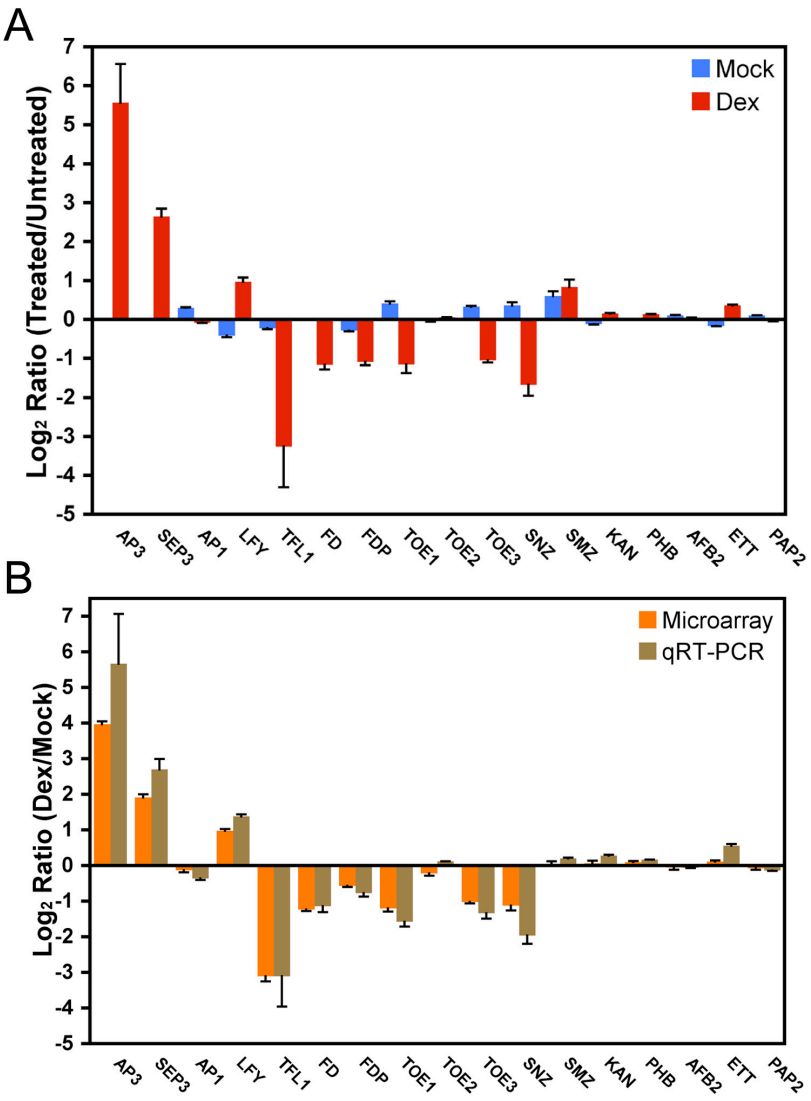

Figure S4: Validation of microarray results for selected genes by quantitative reverse transcriptase (qRT)-PCR. (A) Gene expression levels were determined in inflorescence tissue of 35S:AP1-GR *ap1 cal* plants treated for 12 hours with a dexamethasone-containing (red bars) or a mock solution (blue bars) and compared to those in untreated control plants. **(B)** Comparison of data from Agilent microarray experiments (12 h time point; orange bars) and qRT-PCR assays (brown bars). Log_2 ratios are shown. Error bars indicate standard deviations of the experiments (qRT-PCR) and Resolver-calculated errors (microarrays). Gene aliases and corresponding gene identifiers are listed in Table S11. The putative

promoter regions of all genes tested are bound by AP1 according to the results of the ChIP-Seq experiments, except in the case of *TOE2*, for which the closest AP1 ChIP-Seq peaks were located ~52 kb upstream and ~25 kb downstream from the start and end of the gene, respectively. With the exception of *AP1*, *TOE2*, *SMZ*, *KAN*, *PHB*, *AFB2*, *ETT* and *PAP2*, all genes were judged as differentially expressed in the microarray experiments upon AP1 activation. For *AP1*, primers were used that are specific for the endogenous gene and thus, do not amplify transcripts stemming from the AP1-GR transgene. The agreement between the qRT-PCR results and those obtained in the microarray experiments indicates that the limited overlap between the ChIP-Seq and microarray-selected gene sets was not a result of low sensitivity and/or a reduced dynamic range of the microarrays used for our study. Five of the 7 genes (*KAN*, *AFB2*, *ETT*, *PAP2*, and *PHB*) that were not detected as differentially expressed by either method, had been identified by ChIP-Seq peak(s) specific to the gene (i.e., a peak that was not within 3 kb upstream or 1 kb downstream of any other gene); the *SMZ* peak had another associated gene, *AT3G55000*, which was not detected as differentially expressed in the microarray experiments; and the *AP1* peaks were also associated to *AT1G69130*, which codes for a pre-tRNA (not represented in the microarrays used). These observations suggest that the limited overlap between the ChIP-Seq and microarray-selected gene sets is not primarily a result of the intrinsic difficulties in assigning peaks to genes on the basis of distance.

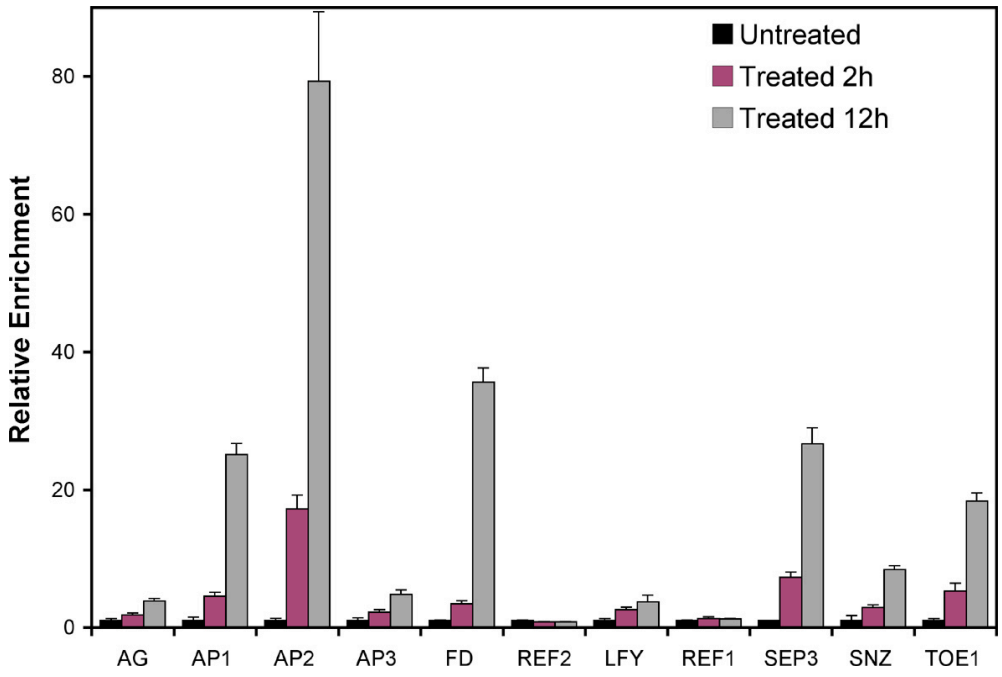

Figure S5: Validation of ChIP-Seq results for selected AP1 binding sites. Quantitative PCR was done using genomic DNA as a template, which was obtained from inflorescence tissue of untreated 35S:AP1-GR *ap1 cal* plants, or from plants 2 or 12h after dexamethasone treatment (as indicated), and then subjected to ChIP with antibodies against the GR portion of the AP1-GR fusion protein. Normalized relative quantities of enrichment were calculated over the uninduced control ChIP sample and normalized using the two reference genomic regions (REF1 and REF2; see Table S12). Error bars indicate standard deviations. In most cases, the genomic regions that were amplified corresponded to the highest ChIP-Seq peaks within the 3 kb upstream to 1 kb downstream region of the respective genomic loci. In case of *AP2* and *TOE1*, we selected peaks that were positioned approximately 4 kb (*TOE1*) or 3 kb (*AP2*) upstream of the transcriptional start of the respective genes. Gene aliases and corresponding gene identifiers are listed in Table S12.

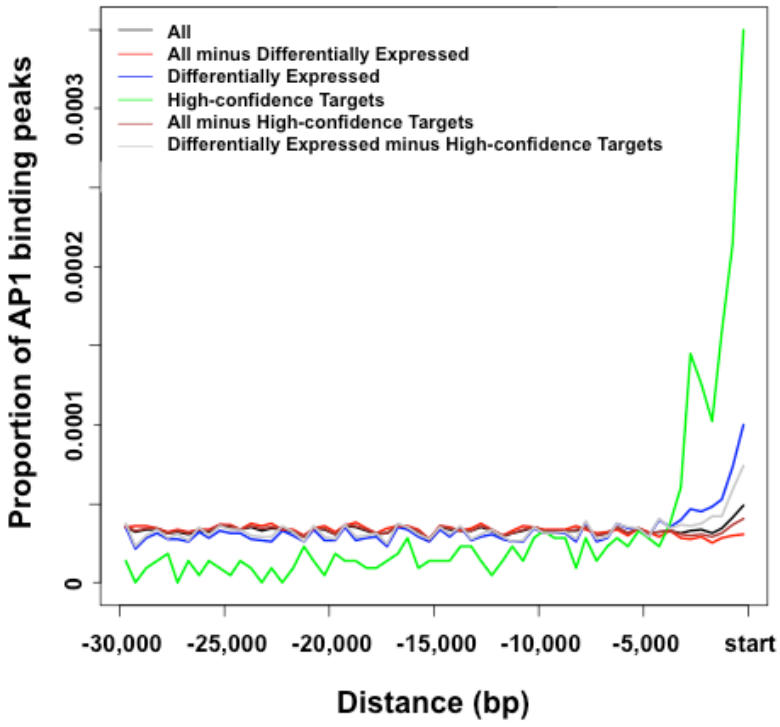

Figure S6: Relative AP1 ChIP-Seq peak position distribution with respect to the 5' end of genes. AP1 binding sites are enriched near the start site (5' end) of genes, and the level of enrichment varies when different groups of genes are considered. All significant AP1 peaks (FDR<0.001) within a distance of 30 kb upstream from the start of each gene in the respective group are included in the analysis. For each group of genes, the proportion of distances coming from the background level is calculated assuming that distances coming from it follow a uniform distribution, and that distances coming from a true enriched signal are non-uniform. To separate these two sets of distances, the method of Storey (*S9*) was used. Gene sets are as follows: all TAIR8 annotated genes (black line; the proportion of distances stemming from the background level is 98.7%); high-confidence AP1 target genes (as defined in the article: adjusted P-value< 0.05 and fold change> 1.8 in Agilent and/or Operon microarray experiments, plus an AP1 ChIP-Seq peak in the -3kb to +1kb region with a FDR< 0.001) (green; 27%); genes that are differentially expressed (adjusted P-value<

0.05, no fold-change cut-off, Agilent dataset) (blue; 83.5%); genes that are differentially expressed minus the high-confidence AP1 target genes (grey; 92.9%); all TAIR8 annotated genes minus the high-confidence AP1 target genes (brown, 99.5%); and all TAIR8 annotated genes minus genes that are detected as differentially expressed (i.e., genes that are not differentially expressed) (red; 100%). When all high-confidence AP1 target genes are excluded from the group of differentially expressed genes, there is still enrichment of AP1 binding sites with respect to the background level.

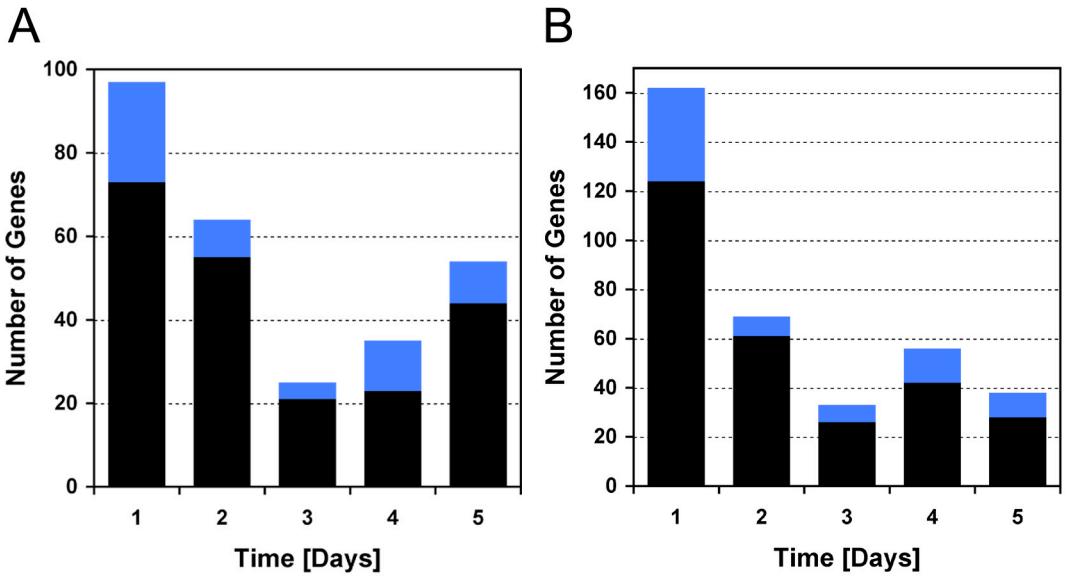

Figure S7: Transcriptional responses of genes bound by AP1 during early flower development. (A) A subset of the genes that did not show a transcriptional response using our statistical selection criteria (adjusted P-value> 0.05), or **(B)** that responded only weakly (adjusted P-value< 0.05 and fold change< 1.8) to AP1 binding within the first 12h after AP1 activation, showed significant changes in expression at more advanced stages of flower development. Both sets of genes were compared with a previously generated dataset stemming from the analysis of differential gene expression during early flower development

(days 1 to 5 after AP1 activation in AP1-GR *ap1 cal* inflorescences; data from (*S1*)). For each of the two sets of genes (**A** and **B**), blue bars indicate the number of AP1-bound genes that showed robust expression changes (adjusted P-value< 0.05 and fold change≥ 1.8) during early flower development, and black bars show the number of genes that responded only weakly (adjusted P-value< 0.05 and fold change< 1.8). The graphs display the number of genes first identified as differentially expressed in a given time point (i.e., genes are counted only once). All together (considering the two sets of genes in A and B), 136 genes (7% of the total) showed a robust transcriptional response, and an additional 497 genes (25%) a weak response.

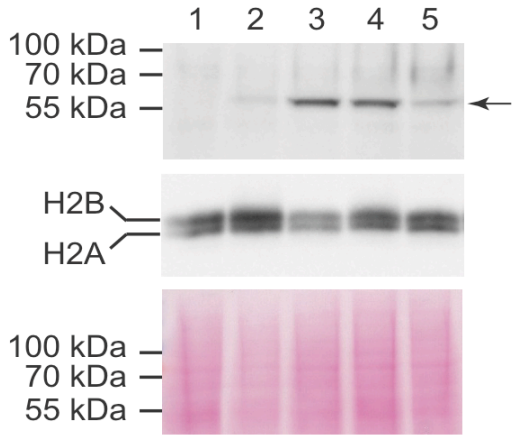

Figure S8: Nuclear accumulation of the AP1-GR fusion protein after dexamethasone treatment. Protein extracts were generated from nuclei isolated from inflorescence tissue of either *ap1 cal* double mutants (lane 1) or of 35S:AP1-GR *ap1 cal* plants (lanes 2-5). The latter plants were either treated with a mock solution (lane 2) or with a dexamethasone-containing solution (lanes 3-5). From the plants treated with dexamethasone, tissue was collected after 1 d (lane 3), 3 d (lane 4) or 5 d (lane 5). The protein extracts were subjected to SDS-gel electrophoresis and then transferred onto a polyvinylidene fluoride (PVDF)

membrane for immunoblotting. Upper panel: results of an immunoblot using an α-GR antiserum. A band corresponding to the AP1-GR fusion protein (~63 kDa) is marked by an arrow. The approximate positions of protein size markers are indicated on the left. The results of the immunoblot show that AP1-GR strongly accumulated in nuclear extracts of dexamethasone-treated plants 1 and 3 d after the treatment. After 5 d, the amount of fusion protein detected in nuclear extracts was considerably lower than in the 1 and 3 d samples, suggesting that most of the AP1-GR protein that entered the nucleus in response to the dexamethasone treatment had been degraded or had undergone nuclear export. Center panel: the membrane shown in the upper panel was stripped and re-probed with an α-histone 2A/B antiserum. Bands corresponding to histone 2A and B are marked. Lower panel: Ponceau S staining of the PVDF membrane after protein transfer showing that approximately equal amounts of protein were present in each lane. The approximate positions of protein size markers are indicated on the left.

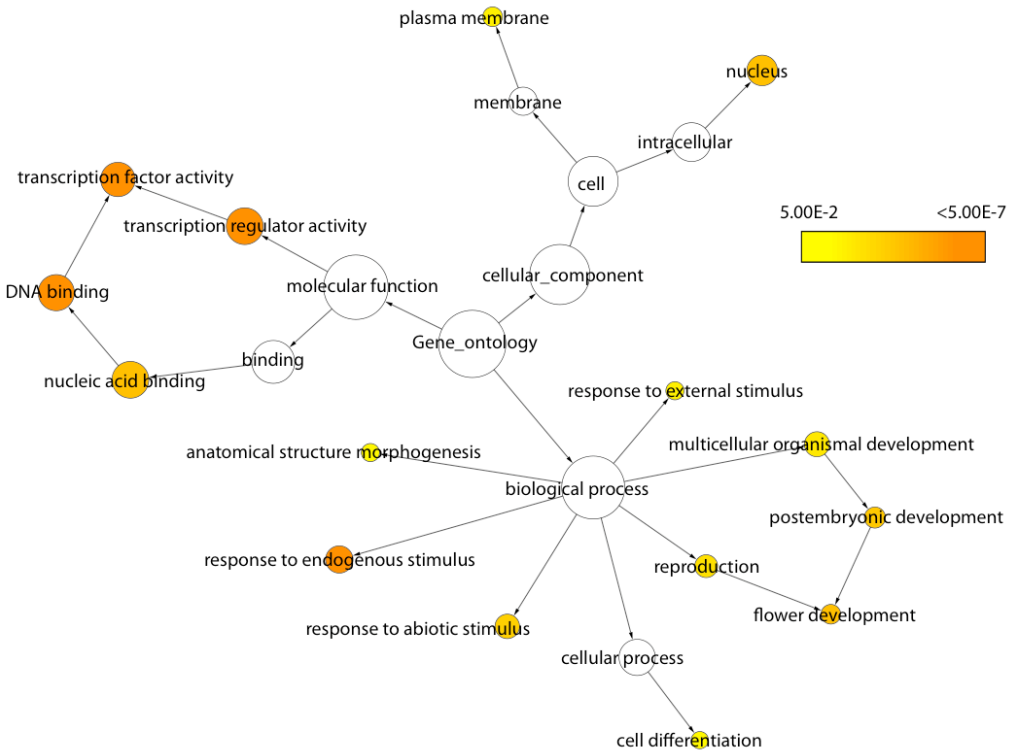

Figure S9: Network visualization of gene ontology (GO) categories enriched in the set of high-confidence AP1 target genes. GOslim categories with significant enrichment (adjusted P-values< 0.05) in the dataset are highlighted in color. Different colors represent different significance levels (as indicated). Diagram was generated using BiNGO version 2.3 (*S13*) and a GO Slim plant database.

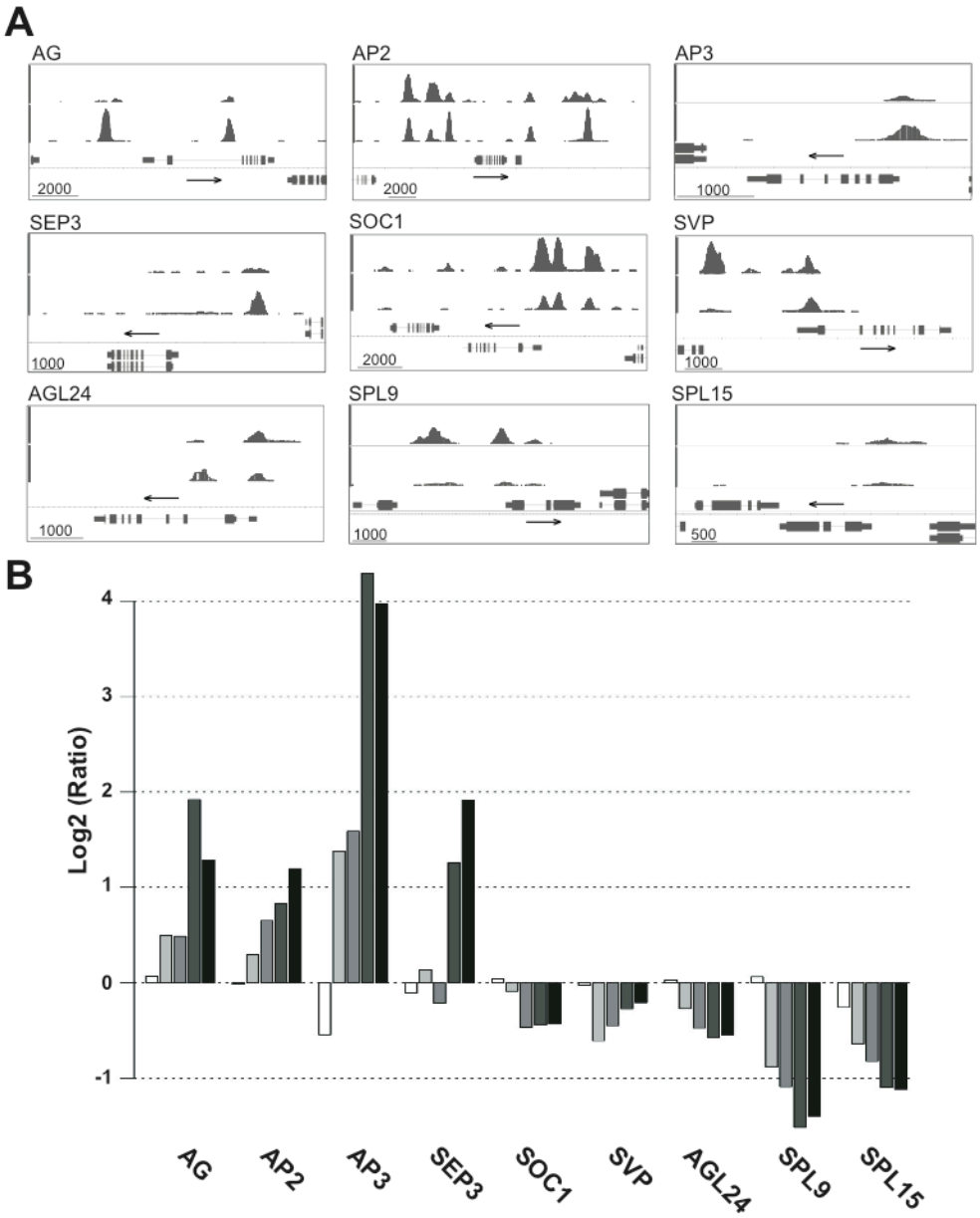

Figure S10: ChIP-Seq and microarray results for selected AP1 target genes. (A) ChIP-Seq results for selected target genes (as indicated). In each panel, the uppermost trace represents AP1 ChIP-Seq data followed by those for SEP3 (*S8*), which are shown for comparison. Genes found in the genomic regions analyzed, as well as their exon-intron structures, are indicated in the lower half of each panel. Scale bars indicate sequence lengths (in bp), and arrows gene orientations. The scale of the Y axis (peak height) is adjusted for

individual traces for visual clarity. **(B)** Transcriptional responses of the genes shown above after AP1 activation. The plot was generated using $\log_2$ ratios derived from (Agilent) microarray experiments. Timepoints are: 0h (white bar), 2h, 4h, 8h (grey scale), and 12h (black).

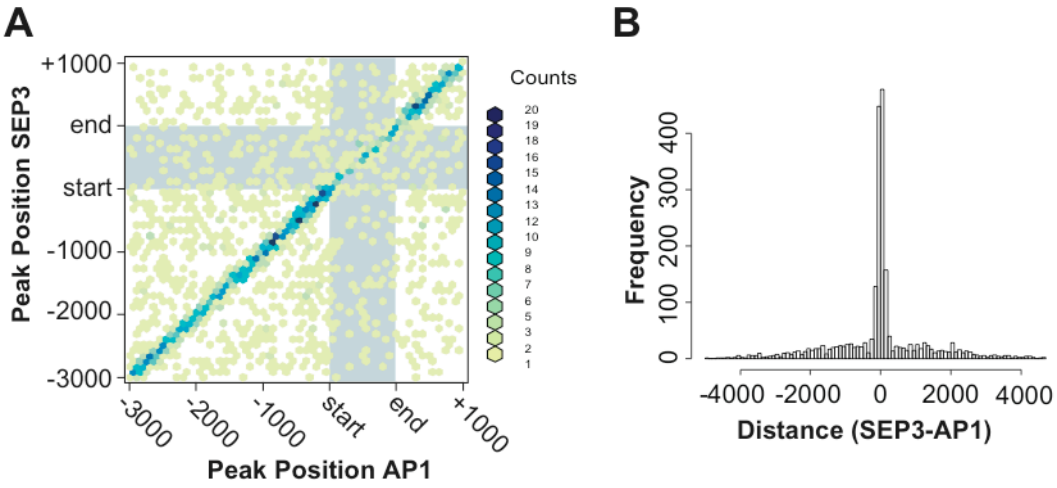

Figure S11. Spatial distribution of AP1 and SEP3 binding sites. (A) All possible pair-wise position combinations of SEP3 and AP1 ChIP-Seq peaks in common target genes are depicted in the diagram. Binding data for SEP3 were taken from (*S8*). Transcribed regions are shaded in grey. Within transcribed regions, peak positions were calculated as a fraction of the gene length (see Fig. 1F). The density of peak pairs in the graph is represented with a color scale (Counts; number of peak pairs per area unit). Because for the AP1 and SEP3 peaks associated with a target gene, no positional information (with respect to the gene and to each other) is considered for pairing the peaks, but rather, all possible combinations are calculated, the results exclude the possibility of positive biases caused by position-directed peak pairing. **(B)** Distance distribution of SEP3 and AP1 binding sites in common target genes. Each bar represents a 100-bp window.

Tables and legends

Table S1: Genes identified as differentially expressed after AP1 activation. Genes for which the Benjamini & Hochberg-adjusted P-value was <0.05 in at least one of the timepoints and microarray platforms used (Agilent and/or Operon) were initially considered as differentially expressed in response to AP1 activation ("Agilent and_or Operon_BH-0h" worksheet). Genes that were differentially expressed at timepoint 0h were excluded from the analysis of the corresponding platform data. Results are summarized per time point (0h, 2h, 4h, 8h, 12h), platform (Agilent or Operon) and, in the case of the Operon oligo microarrays, per scanner used (GenePix 4200A or 4000B), and arranged accordingly in columns (0h Agilent, 0h Operon 4200 SCAN, 0h Operon 4000 SCAN, 2h Agilent, etc.) using values of "0" (no differential expression) and "1" (differential expression). Fold change, P-value, and Benjamini & Hochberg-adjusted P-value (BH) are also shown. A gene may be listed more than once if probes assigned to different gene models or variants were present, either within or across platforms. A total of 6600 different genes (using probes for 6731 different gene/gene variants) were identified as differentially expressed. A further selection for robust differential expression was implemented by applying an absolute fold-change (FC) cutoff of 1.8 ("Agilent and_or Operon BH&FC-0h" worksheet), which led to the identification of 1366 different genes (using probes for 1392 gene/gene variants).

Supplementary file: Table S1.xls

Table S2: Genomic regions identified as enriched in AP1 ChIP-Seq experiments. A total of 14,174 genomic regions were identified in the AP1 ChIP-Seq experiments ("ChIP-Seq peaks" worksheet), 1,942 of which were classified as statistically significant (FDR<0.001, which corresponded to a peak score >79.6). Of those 1,942 enriched regions, 1,678 were within 3 kb upstream from the start to 1 kb downstream from the end of TAIR8 annotated genes: a total of 2298 different genes were identified, which resulted in 2,653 different peak-gene combination ("ChIP-Seq peak - Gene pairs" worksheet; one peak can be associated with more than one gene, and vice versa). These 2,298 genes are classified as "Putative direct targets". The position of the peak with a score higher than 79.6 (or peaks, if more than one peak above that threshold is associated with a gene) with respect to the corresponding gene is indicated by dividing the gene region into different segments: from 3000 bp to 2001 bp upstream of the start, from 2000 to 1001 bp upstream, from 1000 bp upstream to the 5' end, from 250 bp upstream to the 5' end, within the annotated gene region (5' - 3' end), and from the end of the annotated gene to 1000 bp downstream. The maximum peak score within the indicated regions is shown in the columns. A value of 0 indicates that no peak in that region was above the score threshold of 79.6. Gene symbols and sequence annotation (Sequence Description) are provided for the identified genes. Columns in the "ChIP-Seq peaks" worksheet show: a unique identifier for each peak, which consists of the chromosome number and the coordinate position for the peak maximum (ChIP-Seq peak ID column); the start and end coordinates of the peak region (Start coordinates, End coordinates); the position, in chromosome coordinates, of the nucleotide within the peak with the highest score (ChIP-Seq peak score position); the peak score, which is the maximum score of any of the nucleotides encompassed by the peak (ChIP-Seq peak score); and the distance between the

start and the end of the peak (Peak length). Additional columns in the ChIP-Seq peak - Gene pairs worksheet show: the position of the peak with respect to the gene (Relative position to the start of the gene, and Relative position to the end of the gene), a unique identifier for the gene associated with the peak (Gene ID (At#)), and the distance from the start to the end of the gene according to TAIR8 annotation (Gene length).

Supplementary file: Table S2.xls

Table S3: High-confidence AP1 targets. High-confidence target genes were those identified as differentially expressed (BH< 0.05 and FC> 1.8) and for which a ChIP-Seq AP1 binding site was detected (FDR< 0.001, peak score> 79.6) within the region from 3 kb upstream from the start to 1 kb downstream from the end of the annotated gene (see Fig. 1D). For each gene, both the ChIP-Seq results and the microarray data are presented in the table. The position of the peak with a score higher than 79.6 (or peaks, if more than one peak above that threshold is associated with a gene) with respect to the corresponding gene is indicated by dividing the gene region into different segments: from 3000 bp to 2001 bp upstream of the start, from 2000 to 1001 bp upstream, from 1000 bp upstream to the 5' end, from 250 bp upstream to the 5' end, within the annotated gene region (5' - 3' end), and from the end of the annotated gene to 1000 bp downstream. The maximum peak score within the indicated regions is shown in the columns. A value of 0 indicates that no peak in that region was above the score threshold of 79.6. Gene symbols and sequence annotation (Sequence Description) are provided for the identified genes. Microarray results are summarized per time point (0h, 2h, 4h, 8h, 12h), platform (Agilent or Operon) and, in the case of the Operon oligo microarrays, per scanner used (GenePix 4200A or 4000B), and arranged accordingly in

columns (0h Agilent, 0h Operon 4200 SCAN, 0h Operon 4000 SCAN, 2h Agilent, etc.) using values of "0" (no differential expression) and "1" (differential expression). Fold change, P-value, and Benjamini & Hochberg-adjusted P-value (BH) are also shown. The table is non-redundant, and genes are listed only once.

Supplementary file: Table S3.xls

Table S4: Classification of all genes represented on the Agilent or Operon microarrays with respect to their differential expression and their identification as AP1 target. The proportion of differentially expressed genes showing a strong transcriptional response (BH< 0.05 and FC> 1.8) was significantly larger among AP1 targets (11.41%; 249 out of 2183) (targets selected at FDR< 0.001) than among non-targets (4.31%; 1117 out of 25942) as revealed by a chi-square test for equality of proportions (with 1 degree of freedom, P-value< 0.001).

Gene		Was AP1 target?		Margin
		No	Yes	
Was differentially expressed?	No	24825	1934	26759
	Yes	1117	249	1366
Margin		25942	2183	28125

Table S5: Microarray results of the cycloheximide experiment. Three different strategies were used to identify genes that respond transcriptionally to AP1 activation in the presence of the translational inhibitor cycloheximide (see Materials and Methods): direct Dex vs. Dex+Chx microarray hybridizations ("Direct comparison" worksheet); Dex vs. Dex+Chx comparisons created via a re-ratio analysis ("Re-ratio comparison" worksheet); and comparison of the Mock vs. treatment hybridizations, where genes were selected if they were

detected as differentially expressed upon both the Dex and Dex+Chx treatments but not upon the Chx treatment ("Indirect comparison" worksheet). In all cases, genes were deemed as differentially expressed if Fold change>1.8 and Benjamini & Hochberg-adjusted P-value (BH) <0.05. Fold change, P-value, and adjusted P-value (BH) are shown. The table is redundant, and genes may be listed more than once, because for some genes more than one probe was present on the microarrays used. Agilent version 2 microarrays were used in the cycloheximide experiments, and a total of 87 out of the 197 high-confidence targets (44%) that were detected as such with Agilent version 1 arrays in the time course experiment, responded transcriptionally in the absence of protein biosynthesis. The proportion of differentially expressed genes in the presence of CHX was significantly larger among Agilent-identified high-confidence AP1 targets (87 out of 197, 44.16%) than among non-targets (1491 out of 28130, 5.30%) as revealed by a chi-square test for equality of proportions (with 1 degree of freedom, P-value < 0.001) (worksheet "chi-square").

Supplementary file: Table S5.xls

Table S6: Necessity and sufficiency of AP1 for affecting the expression of high-confidence target genes. To obtain evidence for the necessity of AP1 as an inductive (or repressing) factor of the high-confidence target genes, we performed a gene expression profiling experiment comparing *ap1-1* mutant and wild-type inflorescences using Agilent version 1 microarrays (worksheet "AP1 necessity (ap1 vs wt)"). If AP1 activity were necessary, AP1 target genes that were found to be induced upon AP1-GR activation would be expected to be down-regulated when comparing gene expression in *ap1* versus wild-type inflorescences and, conversely, AP1 target genes that were repressed after AP1-GR

activation would be expected to be up-regulated in the *ap1* versus wild type comparison, as *ap1-1* is a loss-of-function mutant. Sixteen genes (out of the 197 high-confidence targets that were identified through the Agilent microarrays; Table S3) showed expression changes in the *ap1* vs wild type comparison that followed the above logic. That this percentage is low (8.1%) is expected, given that the plant materials and the developmental stages that were assayed in the *AP1-GR ap1 cal* time course and in the *ap1* versus wild type comparison were necessarily different, and because in *ap1* mutant plants the (partially) functionally redundant homolog *CAL* is present (and therefore masks part of the effects of the loss of *AP1* function). For the sixteen genes, the table shows: Fold change and P-value for the *ap1* vs. wild type experiment; and Fold Change, P-value, and adjusted P-value (BH) for the timepoint (indicated) at which strong differential expression (FC>1.8, BH<0.05) was first detected in the time course experiment. The complete *ap1* versus wild type microarray dataset is available in the Gene Expression Omnibus under accession number GSE20138. To obtain evidence for the sufficiency of AP1 as an inductive (or repressing) factor of the high-confidence target genes, we compared the list of such genes with the results of the cycloheximide experiment (worksheet "AP1 sufficiency (CHX)"). Targets that are induced (or repressed) after AP1-GR activation, and that show a similar change in expression when AP1-GR is activated in the presence of cycloheximide (i.e., in the absence of protein synthesis) represent genes for which AP1 is, in a floral context, sufficient to exert the expression change. A total of 81 targets (41%) were identified. For the 81 genes, the table shows Fold Change, P-value, and adjusted P-value (BH) for the timepoint (indicated) at which strong differential expression (FC>1.8, BH<0.05) was first detected in the time course

experiment, and for the three different analysis methods used in the cycloheximide experiment (the cycloheximide results are described in full in table S5).

Supplementary file: Table S6.xls

Table S7. Classification of transcription factor-coding genes represented on Agilent or Operon microarrays with respect to their differential expression and their identification as AP1 target. The fraction of transcription factor-coding genes showing a strong transcriptional response (BH< 0.05 and FC> 1.8) is significantly larger among AP1 targets (18.51%; 62 out of 335) (targets selected at FDR < 0.001) than among non-targets (7.43%; 102 out of 1373) as revealed by a chi-square test for equality of proportions (with 1 degree of freedom, P-value < 0.001).

Gene		Was AP1 target?		Margin
		No	Yes	
Was differentially expressed?	No	1271	273	1544
	Yes	102	62	164
Margin		1373	335	1708

Table S8: Phenotypes observed for the main inflorescences of *tfl1* T-DNA insertion mutants and of transgenic lines derived from the transformation of *tfl1-1* plants with fragments of the *TFL1* genomic region.

Genotype	n[c]	Cauline leaves[d]	Axillary flowers[d,e]	Determinate coflorescences[d,f]	Indeterminate coflorescences[d,g]	No. of flowers[d,h]	Terminal flower (%)[i]
Wild type Columbia	20	3±0.9	0	0	3±0.9	36.3±4.8	0
tfl1-1	47	1.5±0.7	1.4±0.7	0	0	2.2±1.2	100

tfl1-1 pASM4[a]	7	4.4±1.3	0	0	4.4±1.3	43.7±0.5	0
tfl1-1 pASM8[a]	65	2.4±0.9	0.03±0.2	2.3±0.9	0	5.3±2.9	100
tfl1-1 pASM6[a]	31	1.5±08	1.4±08	0.±0.2	0	2.7±1.2	100
∇ 1615[b]	8	1.4±0.5	0.1±0.4	1.3±0.7	0	6.4±1.2	100
∇ 418[b]	11	1.2±0.6	1±0.5	0.2±0.4	0	3.8±1.2	100

[a] Transgenic lines derived from the transformation of *tfl1-1* with fragments of the *TFL1* gene containing the intergenic region upstream the start codon, the *TFL1* coding sequence and 4613 bp (pASM4), 1625 bp (pASM8) or 432 bp (pASM6) of the sequence downstream the *TFL1* stop codon.
[b] Lines with T-DNA insertions at 1615 bp or 418 bp downstream of the stop codon of *TFL1*.
[c] Number of plants analyzed. For *tfl1-1*, ∇1615 and ∇418, homozygous mutant plants were analyzed. For *tfl1-1* plants transformed with genomic constructs, the data correspond to the average of the indicated number of independent T1 plants.
[d] Values are expressed as mean±SD.
[e] Number of solitary flowers formed in the axil of a cauline leave.
[f] Number of coflorescences that formed a terminal flower.
[g] Number of coflorescences that did not form a terminal flower
[h] Number of flowers produced by the main inflorescence stem before cessation of growth.
[i] Percentage of plants whose main inflorescence formed a terminal flower.

Table S9: Clustering analysis of differentially expressed genes. Genes identified as differentially expressed (Benjamini & Hochberg-adjusted P-value <0.05 and FC>1.8) in the Agilent time course microarray experiments and that were subjected to cluster analysis are shown in the table. The cluster in which each gene was grouped (Cluster Number; 1 through 6, as in Fig. 1B) is indicated.

Supplementary file: Table S9.xls

Table S10: Mapping results for the Illumina sequencing runs. Results are shown per flow cell lane. The number of reads (sequences) that were mapped in the *Arabidopsis* chromosomes (chr1-5), and in the chloroplast and mitochondrial genomes, is indicated, as

well as the total number of mapped reads. Column "% chr1-5" shows the percent of the total number of reads that was mapped to chromosomes 1 through 5 (i.e., excluding chloroplast- and mitochondria-mapped reads).

Illumina Sequencing run (lanes)	**Mapping results**					**Total number of reads**
	chrC	chrM	chr1-5	Total number of mapped reads	% chr1-5*	
Replicate 1 uninduced 1	108263	50471	1715147	1873881	49.3	3479131
Replicate 1 uninduced 2	166430	149641	3234895	3550966	43.7	7405077
Replicate 1 induced	99155	42967	3355117	3497239	46.9	7154722
Replicate 2 uninduced 1	92817	14725	2221675	2329217	48.0	4627140
Replicate 2 uninduced 2	183799	39433	3725948	3949180	38.7	9636656
Replicate 2 induced 1	85198	16261	3151437	3252896	64.9	4854426
Replicate 2 induced 2	179239	43248	6321287	6543774	60.9	10371555

* Regarding the total number of reads

Table S11: Primers used for qRT-PCR assays. Gene identifiers, gene aliases and primer sequences are shown.

Gene ID	**Alias**	**5' Primer**	**3' Primer**
AT3G26810	AFB2	5'-ACTTCAGCGGTTATGGATTTTGG-3'	5'-AGATTTCCACACCTCTTTGCTTG-3'
AT1G69120	AP1	5'-ATTAGCAGGCTCCTTTCTTCTTTTG-3'	5'-CCCTGATACCCGATCTGAAGAATG-3'
AT3G54340	AP3	5'-TCCGGACTCAGATCAAGCA-3'	5'-CATTGCTAGCGATAGATCGCTC-3'
AT2G33860	ETT	5'-AACAACTGGTCCCAAGAGAAGC-3'	5'-ACATCTGGGTCTATCATCACCA-3'
AT4G35900	FD	5'-TCTCTCTACTCCCTCTCTGCG-3'	5'-CCGAGGCTAGATGTGTTGTAGAG-3'
AT2G17770	FDP	5'-CAGGCATATACAAACGAACTTGAGC-3'	5'-TGAGAGTCTGGTTCTTGTTTGTAGC-3'
AT5G16560	KAN	5'-GTAACTCTTCAGGA	5'-GCAAGACCCTCTGGAAT

		GAGACATGGC-3’	CTCAATG-3’
AT5G61850	LFY	5’-CGTTTATCGTAACG GAGCCTGG-3’	5’-TTGAGTTGCTTCGGCTA TTTTCAG-3’
AT4G29080	PAP2	5’-CAGTGACTGGATGC TGGTCG-3’	5’-CTAGACCTTCTTCGGTC ACAGC-3’
AT2G34710	PHB	5’-CGAACCTGGTCGTA AAGCTCTTTG-3’	5’-CTTACCGTATAAACCCA TCGTGGAC-3’
AT1G13320	REF1	5’-AAGCGGTTGTGGAG AACATGATACG-3’	5’-AGTTCACGTTCATTCTCT CTGACC-3’
AT4G34270	REF2	5’-ATCTGCGAAAGGGT ATCCAGTTGAC-3’	5’-TGGAGAGCTTGATTTGC GAAATACCG-3’
AT1G24260	SEP3	5’-CTCTAAGACTAAGG TTAGCTGATGGG-3’	5’-CTGAATACTTGGTTCGT CACCTTGG-3’
AT3G54990	SMZ	5’-CAATGAGTTGGGAA AGGGAGAAGG-3’	5’-CTTACCGTATAAACCCA TCGTGGAC-3’
AT2G39250	SNZ	5’-AGTTGGGTGCCCAT AGTAAAGG-3’	5’-CCATGCTTGACGCACAT ACTCC-3’
AT5G03840	TFL1	5’-TCCTTCTTCTGTTTC CTCCAAG-3’	5’-TTCCATTTCATCCTCAAG ACG-3’
AT2G28550	TOE1	5’-CAGCGTGGAGTTAG CTTGAGG-3’	5’-CTTCCCATGCAAACTCG TCAC-3’
AT5G60120	TOE2	5’-GGTCTCTATGTTATC CACCCATCTG-3’	5’-CGTTCCAGTAAAGGCGA TGATCC-3’
AT5G67180	TOE3	5’-GATCTTAGCTCAGA GACGACGAG-3’	5’-GCTGTGTAGAACCTTGA GTTATGCG-3’

Table S12: Primers used for ChIP-qPCR assays. Gene identifiers, gene aliases and primer sequences are shown.

Gene ID	**Alias**	**5’ Primer**	**3’ Primer**
AT4G35900	FD	5’-TATCTTCTTTTGGCC CAATTTCATTT-3’	5’-TTGGTTGCTTAAAAGA ATACACAGCA-3’
AT1G69120	AP1	5’-GGAGATCATGAGCA AAAGTAGTAAGCA-3’	5’-TGGAGATGTAACAAA GGCGAAGATAA-3’
AT2G39250	SNZ	5’-CAATAATGGAGTGA AATGGAACGTGT-3’	5’-TCTACCTCTCTCTCCC CTCCACATAC-3’
AT2G28550	TOE1	5’-CCATCATGGTAAGTG GTAACCAAGTC-3’	5’-GAGACCCATTATTGGG AGTAACCAAA-3’
AT4G36920	AP2	5’-TCATATATGTCGCAT AACGAGGGAGT-3’	5’-TGCTCATATAAAACAA AACCAAAGAAACA-3’
AT5G61850	LFY	5’-TGCATGCATTCACAC ATAGTACACAT-3’	5’-TATTATCCGCCGAGCA ATAGACTGTA-3’
AT4G18960	AG	5’-GAACGTTGTGATGTT	5’-TCAACAACCCATTAAC

		ACTCGGACAAG-3’	ACATTGGGTA-3’
AT1G24260	SEP3	5’-AAACCCAGACGTGA CTTGTTTGACG-3’	5’-TGAGAATCGGACGGC TTTGAGG-3’
AT3G54340	AP3	5’-CAATTGATTTAAGCA GTGTC-3’	5’-GGAAAGTATTGCCTA ATCCATGAAAG-3’
AT4G26930	REF1	5’-TCTCCGACCTTTCTT CACACCCATTCC-3’	5’-GTCTCCGCTTAGGAGC ACGAAAGCTATC-3’
AT4G17740	REF2	5’-GCTATCCACAGGTTA GATAAAGGAG-3’	5’-GAGAAAGATTGTGTG AGAATGAAA-3’

References

S1. F. Wellmer *et al.*, *PLoS Genetics* **2**, e117 (2006).

S2. F. Wellmer, J. L. Riechmann, M. Alves-Ferreira, E. M. Meyerowitz, *Plant Cell* **16**, 1314 (2004).

S3. L. Weng *et al.*, *Bioinformatics* **22**, 1111 (2006).

S4. K. J. Livak, T. D. Schmittgen, *Methods* **25**, 402 (2001).

S5. T. Czechowski, M. Stitt, T. Altmann, M. K. Udvardi, W. R. Scheible, *Plant Physiol* **139**, 5 (2005).

S6. K. Kaufmann *et al.*, *Nature Protocols* **5**, (in press) (2010).

S7. R. Li, Y. Li, K. Kristiansen, J. Wang, *Bioinformatics* **24**, 713 (2008).

S8. K. Kaufmann *et al.*, *PLoS Biol* **7**, e1000090 (2009).

S9. J. D. Storey, *J R STAT SOC B* **64**, 479 (2002).

S10. D. Bradley, O. Ratcliffe, C. Vincent, R. Carpenter, E. Coen, *Science* **275**, 80 (1997).

S11. E. J. Sohn *et al.*, *Proc Natl Acad Sci U S A* **104**, 18801 (2007).

S12. S. J. Clough, A. F. Bent, *Plant J* **16**, 735 (1998).

S13. S. Maere, K. Heymans, M. Kuiper, *Bioinformatics* **21**, 3448 (2005).