Target Genes of the MADS Transcription Factor SEPALLATA3: Integration of Developmental and Hormonal Pathways in the *Arabidopsis* Flower

Kerstin Kaufmann[1], Jose M. Muiño[2], Ruy Jauregui[3], Chiara A. Airoldi[4], Cezary Smaczniak[1,5], Pawel Krajewski[2], Gerco C. Angenent[1,5*]

1 Business Unit Bioscience, Plant Research International, Wageningen, The Netherlands, **2** Institute of Plant Genetics, Polish Academy of Sciences, Poznań, Poland, **3** BIOBASE GmbH, Wolfenbüttel, Germany, **4** Centre for Plant Sciences, University of Leeds, Leeds, United Kingdom, **5** Centre for BioSystems Genomics (CBSG), Wageningen, The Netherlands

The molecular mechanisms by which floral homeotic genes act as major developmental switches to specify the identity of floral organs are still largely unknown. Floral homeotic genes encode transcription factors of the MADS-box family, which are supposed to assemble in a combinatorial fashion into organ-specific multimeric protein complexes. Major mediators of protein interactions are MADS-domain proteins of the SEPALLATA subfamily, which play a crucial role in the development of all types of floral organs. In order to characterize the roles of the SEPALLATA3 transcription factor complexes at the molecular level, we analyzed genome-wide the direct targets of SEPALLATA3. We used chromatin immunoprecipitation followed by ultrahigh-throughput sequencing or hybridization to whole-genome tiling arrays to obtain genome-wide DNA-binding patterns of SEPALLATA3. The results demonstrate that SEPALLATA3 binds to thousands of sites in the genome. Most potential target sites that were strongly bound in wild-type inflorescences are also bound in the floral homeotic *agamous* mutant, which displays only the perianth organs, sepals, and petals. Characterization of the target genes shows that SEPALLATA3 integrates and modulates different growth-related and hormonal pathways in a combinatorial fashion with other MADS-box proteins and possibly with non-MADS transcription factors. In particular, the results suggest multiple links between SEPALLATA3 and auxin signaling pathways. Our gene expression analyses link the genomic binding site data with the phenotype of plants expressing a dominant repressor version of SEPALLATA3, suggesting that it modulates auxin response to facilitate floral organ outgrowth and morphogenesis. Furthermore, the binding of the SEPALLATA3 protein to *cis*-regulatory elements of other MADS-box genes and expression analyses reveal that this protein is a key component in the regulatory transcriptional network underlying the formation of floral organs.

Citation: Kaufmann K, Muiño JM, Jauregui R, Airoldi CA, Smaczniak C, et al. (2009) Target genes of the MADS transcription factor SEPALLATA3: Integration of developmental and hormonal pathways in the *Arabidopsis* flower. PLoS Biol 7(4): e1000090. doi:10.1371/journal.pbio.1000090

Introduction

In contrast to animals, most developmental processes of plants occur postembryonically and integrate a variety of internal and environmental cues. Modularity of plant development is based on the ability of plants to maintain pools of undifferentiated stem cells throughout the life cycle of the plant. Stem cells near the tip of the growing shoot are located in the shoot apical meristem, where different types of plant organs, such as vegetative leaves or floral organs, can be initiated from the flanks of the meristem. Leaves and floral organs are "variations of a theme": they arise via modifications of a common basic genetic program [1]. Which type of organ is produced by the meristem depends on the developmental phase of the plant. Initially, leaves are produced during the early vegetative phase of the plant, followed by the transition to reproductive phase, which triggers the transformation of vegetative shoot meristems to inflorescence and floral meristems, giving rise to flowers and floral organs, respectively.

Thus, change in the identity of plant organs is initiated by reprogramming within the meristems [2]. Plant developmental biologists have identified a number of key regulatory genes that trigger changes in meristem and organ identity, many of them encoding transcription factors, chromatin remodeling factors, or other signaling molecules like microRNAs (miRNAs). One family of transcription factors that is important in this process is the MADS-box gene family [3]. MADS-box genes play crucial roles in the switches from vegetative to inflorescence and finally to floral meristems. These latter

Academic Editor: Detlef Weigel, Max Planck Institute for Developmental Biology, Germany

Received October 10, 2008; **Accepted** March 9, 2009; **Published** April 21, 2009

Copyright: © 2009 Kaufmann et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: ChIP, chromatin immunoprecipitation; ChIP-CHIP, chromatin immunoprecipitation coupled with DNA microarray hybridization; ChIP-SEQ, chromatin immunoprecipitation coupled with Solexa sequencing; FDR, false discovery rate; GO, gene ontology

* To whom correspondence should be addressed. E-mail: gerco.angenent@wur.nl

Tecan Ex. 2114, Page 1 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Author Summary

Most regulatory genes encode transcription factors, which modulate gene expression by binding to regulatory sequences of their target genes. In plants in particular, which genes are directly controlled by these transcription factors, and the molecular mechanisms of target gene recognition in vivo, are still largely unexplored. One of the best-understood developmental processes in plants is flower development. In different combinations, transcription factors of the MADS-box family control the identities of the different types of floral organs: sepals, petals, stamens, and carpels. Here, we present the first genome-wide analysis of binding sites of a MADS-box transcription factor in plants. We show that the MADS-domain protein SEPALLATA3 (SEP3) binds to the regulatory regions of thousands of potential target genes, many of which are also transcription factors. We provide insight into mechanisms of DNA recognition by SEP3, and suggest roles for other transcription factor families in SEP3 target gene regulation. In addition to effects on genes involved in floral organ identity, our data suggest that SEP3 binds to, and modulates, the transcription of target genes involved in hormonal signaling pathways.

meristems give rise to flowers and floral organs, respectively [4].

Developmental transitions and organ differentiation require global changes in gene expression. The genome of the model flowering plant *Arabidopsis thaliana* is roughly 20-fold smaller than the human genome; still, it encodes about 27,000 protein-coding genes, which is more than found for humans (http://www.arabidopsis.org; [5]). One of the most challenging current questions is how developmental control genes trigger global changes in gene expression during the multiple phase transitions and in organ identity determination, starting from a small pool of undifferentiated cells.

In the present study, we focus on the MADS-box transcription factor SEPALLATA3 (SEP3). SEP3 is a member of the SEP subfamily of MADS-box genes, whose members have nearly redundant functions in the specification of floral meristem identity and in the identity of all types of floral organs: sepals, petals, stamens, and carpels. Triple mutants impaired in *SEP1–3* function have flowers with floral organs converted into sepals and display a loss of determinacy in the center of the flower [6]. This phenotype masks the involvement of the *SEP* genes in processes occurring later in development, e.g., the formation of the ovules as has been shown by Favaro et al. (2003) [7]. The SEP3 protein appears to be the central player, since it is part of at least a dozen different MADS domain dimer complexes [8] and it is expressed throughout flower development, from the floral meristem to fully developed floral organs [9]. This suggests that SEP3 is a multifunctional protein controlling a plethora of developmental processes. According to the current model of flower development, the SEP3 protein is proposed to mediate the higher-order protein complex formation between MADS-domain proteins with more specific floral organ identity functions [10]. Furthermore, it may provide the transcriptional activation potential to the floral homeotic protein complexes [10]. More-recent evidence suggests that the SEP3 protein may also recruit transcriptional corepressors, demonstrating that it can modulate the function of the plant protein complexes in a broader sense, depending on the availability of cofactors [11]. However, evidence for higher-order complex formation between MADS-domain proteins comes mostly from protein interaction studies in heterologous systems and genetic data, and there is no indication for the relevance of these interactions in target-gene recognition *in planta* so far. Another question is how different MADS-domain protein complexes achieve functional specificity, since the in vitro DNA-binding characteristics of MADS-domain proteins appeared rather similar, and the short DNA sequence motifs supposedly bound by MADS-domain proteins are very abundant in the *Arabidopsis* genome [12].

In order to characterize the mode of action and general downstream pathways of floral homeotic genes, we generated genome-wide DNA-binding profiles of SEP3 in its native context. Chromatin immunoprecipitation (ChIP) followed by ultrahigh-throughput Solexa (Illumina) sequencing (ChIP-SEQ) has been shown recently to be a powerful tool to obtain genome-wide DNA-binding patterns of transcription factors [13,14]. The large numbers of short individual sequence reads produced by novel instruments facilitate the digital quantification of DNA sequences that are present in a sample. An alternative method comprises the combination of ChIP and whole-genome microarrays (ChIP-CHIP) to map the genomic DNA regions enriched in the immunoprecipitated sample [15,16]. These genomic tiling arrays are available for *Arabidopsis* and have been used to map binding sites for plant transcription factors [17].

We compared the targets of SEP3 in wild-type and the floral homeotic *agamous* (*ag*) mutant background. In the *ag* mutant, stamens are replaced by petals, and instead of the carpels in the fourth whorl, a new mutant flower is formed [18]. Accordingly, the analysis of this mutant should reveal SEP3 target genes specifying petal development, whereas targets that are specific to stamens and carpels should be absent. We further studied the function of SEP3 in the regulation of downstream pathways by analyzing the effects of a dominant repressor version of SEP3 in plants.

The genome-wide identification of direct target genes of SEP3 provides a framework for a hierarchical transcriptional network underlying the formation of floral organs. SEP3 binds to thousands of genomic regions containing the consensus binding sites for MADS-domain proteins, but it also acts as part of regulatory modules with other transcription factors. These modules link floral homeotic gene functions with organ growth. Our analysis identified multiple links between SEP3 and hormonal pathways, and in particular auxin signaling. Auxin signaling is crucial for the outgrowth and development of lateral organs, and its role in flower development has been suggested previously based on mutant phenotypes [19–21]. Our ChIP-SEQ data and the phenotypes of plants that repress direct SEP3 targets in a dominant-negative fashion suggest cooperation of SEP3 and genes in the auxin pathway in the regulation of floral organ growth and differentiation.

Results

Genome-Wide Mapping of Regions Bound by SEPALLATA3

To identify genomic regions bound in vivo by SEPALLATA3 (SEP3), we used ChIP followed by deep sequencing by ChIP-SEQ [14,15]. In a parallel experiment, the ChIP was

Tecan Ex. 2114, Page 2 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Table 1. Number of Significant Peaks in the SEP3 ChIP-SEQ Datasets

Genetic Background	Total Number of Peaks for Different FDR Thresholds			
	FDR < 0.05	FDR < 0.01	FDR < 0.005	FDR < 0.001
Wild-type	9,197	6,092	5,202	4,282
ag mutant	6,641	4,015	3,983	2,828

doi:10.1371/journal.pbio.1000090.t001

followed by whole-genome tiling array hybridizations to identify enriched regions (ChIP-CHIP). For the ChIP experiments, we used inflorescences including inflorescence meristems and floral buds of stage 1–12 from *Arabidopsis* wild-type and *agamous* (*ag*-1) mutant plants. Protein–DNA complexes were immunoprecipitated using a peptide antibody specific to SEP3. As a negative control, we performed ChIP-SEQ using the same antibody on *sep3*-1 mutant plants. Western blot analysis revealed that the SEP3 antibody reacts exclusively with the SEP3 protein (Figure S1).

From the ChIP-SEQ experiments, we obtained between 3 to 7.5 million approximately 35-bp sequence reads after one to three independent rounds of sequencing for each sample, of which 30%–40% were uniquely mapped on the *Arabidopsis* nuclear genome (Table S1). The uniquely mapped reads were extended to 300 bp in order to recover the average original DNA fragments that were subjected to sequencing in a similar fashion as described by Robertson et al. (2007) [14]. This allows positioning of the maximum of enrichment present in the samples at high resolution. The number of mapped reads was counted for every nucleotide position (defined as "number of hits"), and for each strand independently. For some genomic positions, we observed that the number of hits was only due to reads with identical sequence. Although, it is expected that some reads will have an identical sequence, it is also expected that a true peak of enrichment should be represented by several reads with partly overlapping but different sequences. In order to avoid any artifact due to identical sequence reads, we included the requirement that the number of hits at each genomic position should be supported by reads mapped in both DNA strands. To test for enrichment at each nucleotide position in the sample compared to the control, we used a score based on the Poisson distribution, as it is commonly used for statistical modeling of tag counts [22]. For each genomic region representing a candidate peak, the maximum score value was used to test the significance of the peak (defined as "peak score"). We used the false discovery rate (FDR) to control the error rate of our testing procedure.

The number of significant peaks for the ChIP-SEQ datasets is given in Table 1. Notably, SEP3 binds to thousands of regions in the *Arabidopsis* genome. At FDR <0.001, we found 4,282 significantly enriched regions for the SEP3 in wild-type plants, and 2,828 regions in the *ag* mutant. Thus, at this level of significance and given our data, SEP3 seems to bind a reduced number of regions in the *ag* mutant compared to wild-type plants.

We used biological replicates and comparison of ChIP-SEQ and ChIP-CHIP data in order to evaluate the reproducibility of the generated genome-wide binding profiles of SEP3. To simplify the comparison of ChIP-SEQ replicates, the average score in nonoverlapping 5,000-bp windows for each replicate was calculated. A high correlation between the two different sequencing rounds and biological replicates was found (Figure S2E and S2G). Peaks positions and ranks from independent biological replicates overlap strongly (Figure S2H). Since more sequences were produced for replicate 1, we focused in our further bioinformatic analysis on this replicate.

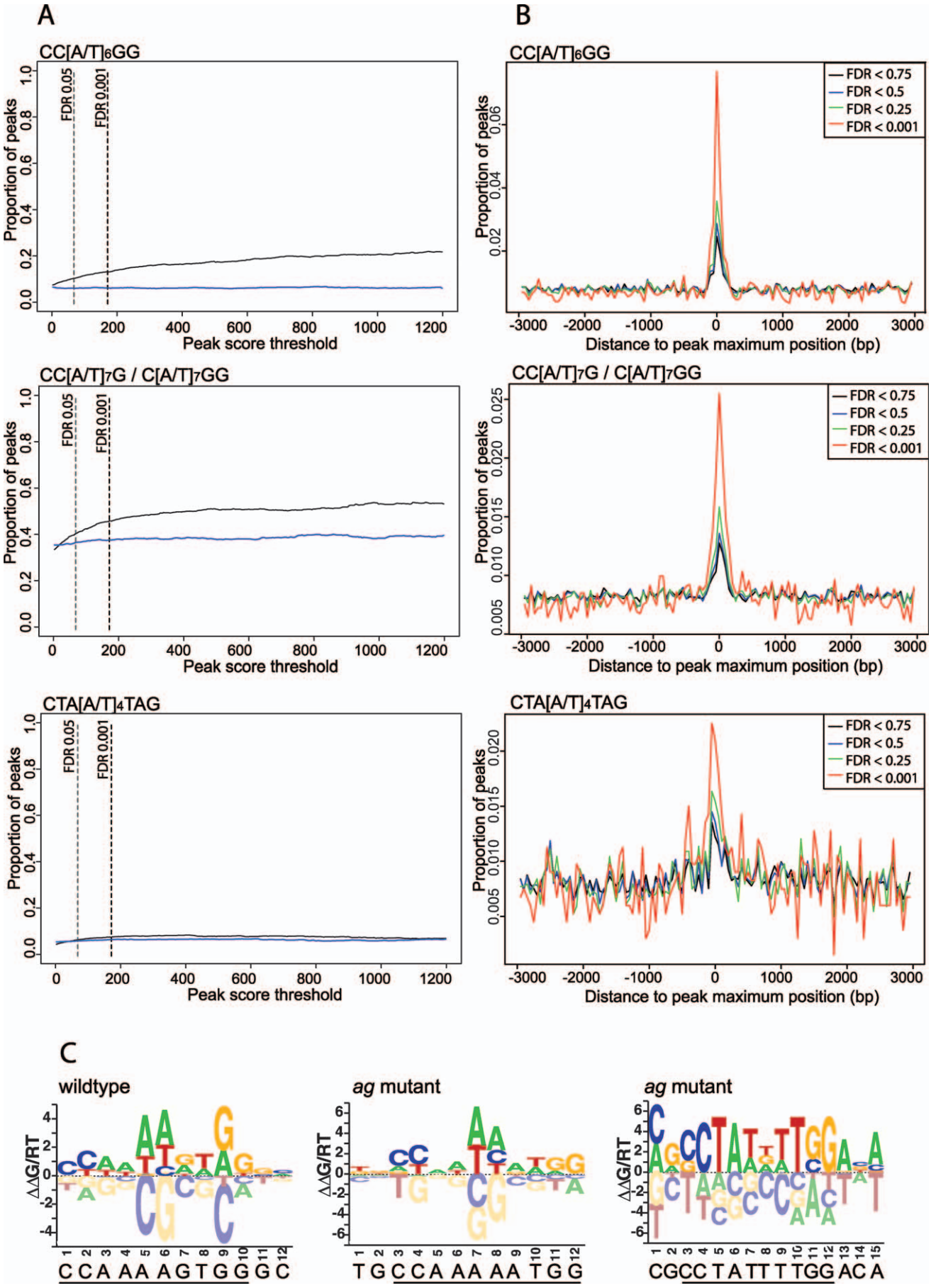
Comparison of the ChIP-CHIP and the ChIP-SEQ experiments reveals a good agreement between the results of the two methods, as reflected in the large overlap in peak positions and similar ranking of the peaks (Figure S2). We were also interested in comparing the positional resolution of the two platforms. The average width of the most strongly significant ChIP-SEQ peaks is around 800 bp. In contrast, the peaks with an equivalent ChIP-CHIP rank have a width of approximately 1,300 bp (Figure S2). This larger window size for the ChIP-CHIP peaks compared to ChIP-SEQ peaks results in a lower positional resolution, which is particularly problematic in regions with multiple binding sites that are close to each other.

Genome-Wide Overview of Bound Regions

cis-Regulatory elements controlling gene expression are preferentially found in the promoters of target genes. However, there are also numerous examples of important regulatory sequences in introns, particularly near the 5′ end of genes [23]. A typical example is the second intron in the MADS-box gene *AG*, which is bound by multiple factors [24–26]. AG, on its turn, binds to the downstream region of the *SPOROCYTLESS* gene, demonstrating that *cis*-regulatory elements are not exclusively located in the upstream or intragenic regions of plant genes [27].

We determined the position of the putative binding sites relative to the nearest gene based on our ChIP-SEQ dataset. As evident from our ChIP-SEQ experiments, most in vivo binding sites of SEP3 are close to or within protein-coding genes and only about 6% to 8% of all peaks (FDR <0.001) are not located within 3 kb upstream to 1 kb downstream of any genomic locus (wild-type and *ag* datasets). A total of 3,475 genes are targeted by SEP3 in wild type, whereas 2,424 genes are putative targets in the *ag* mutant at FDR <0.001 (Tables S2 and S3). In agreement with its role as transcriptional regulator, DNA-binding sites of SEP3 are predominantly located in the upstream region of genes (Figure S3; Table S3). Notably, we found the highest enrichment of SEP3 binding sites in a region spanning a few hundred base pairs directly upstream of the annotated transcriptional start of genes (Figure S3). Surprisingly, binding sites are also enriched in

Tecan Ex. 2114, Page 3 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028



Tecan Ex. 2114, Page 4 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Figure 1. Enrichment, Position, and Sequences of CArG Boxes in ChIP-SEQ Peaks

(A) Enrichment of CArG boxes in peak areas with increasing significance threshold. All different types of CArG boxes show an increase in frequency with increasing significance level. Note that the frequency of the MEF2-type CArG box increases only very weakly enriched. Black line: proportion of peaks with at least one binding site. Blue line: proportion of peaks with at least one binding site after permutating the nucleotide positions of the peak areas, keeping the same base composition (i.e., control).

(B) Frequency plots of sequence position of CArG boxes relative to the peak score maximum position ("center of the peak"). Only perfect CArG boxes were considered in the analyses (no mismatches allowed). All three types of CArG boxes show enrichment in the center of the peak, with the canonical SRF motif CC[A/T]$_6$GG displaying the strongest enrichment in the peak maximum score position.

(C) Affinity logos of CArG box–like elements identified by MatrixREDUCE [30]. The logo represents the estimated DNA-binding affinity ($\Delta\Delta G$) for each nucleotide position. One consensus motif resembling a CArG box was identified for the wild-type ChIP-SEQ data, whereas two motifs were found in the *ag* mutant data.

doi:10.1371/journal.pbio.1000090.g001

the downstream region of genes, located just downstream of the 3′ UTR, although this enrichment is clearly less pronounced than the one found near the transcriptional start site of genes. Within genes, peaks are preferentially located in introns and UTR regions (Figure S3 and Table S3).

Position of MADS-Domain Binding Motifs

It is known from in vitro–binding studies that MADS-domain proteins bind to specific DNA elements called CArG boxes (reviewed in [12]). Most well known is the serum response element (SRE or SRF)-type CArG box, which has the consensus CC[A/T]$_6$GG. A related sequence motif is the MEF2-type CArG box, which has the general consensus C[A/T]$_8$G but is usually more strictly defined as CTA[A/T]$_4$TAG. Plant MADS-domain proteins often show relatively broad DNA-binding preferences, recognizing SRF$^-$ and MEF2$^-$, as well as intermediate motifs [28,29]. CArG boxes are frequently found in the *Arabidopsis* genome, so the presence of this motif alone is not sufficient to predict targets of MADS-box transcription factors [12]. Given the large number of MADS-box transcription factor genes present in the *Arabidopsis* genome and the capacity to form an even larger number of heterodimeric transcription factor complexes [8], and considering their divergent functions in plant development, it is important to understand how different MADS-domain proteins (and protein complexes) achieve target-gene specificity. Also, the relevance of the formation of higher-order protein complexes in DNA binding, associated with the binding to more than one CArG box, has not been demonstrated *in planta* yet.

We considered the 1,001 bp surrounding the peak maximum score position, defined as peak area, for the characterization of transcription factor binding sites. The peak areas were searched for the presence of different types of CArG boxes. Our results show that the SRF, MEF2, and intermediate types are enriched in the genomic regions bound by SEP3 in vivo (Figure 1A). In agreement with a true enrichment of CArG boxes in genomic regions bound by SEP3, we found that the "background" frequency of CArG boxes of type CC[A/T]$_6$GG in promoter regions (–1,000 bp upstream) was approximately 7%, whereas in our ChIP-SEQ data at FDR = 0.001, the frequency is approximately 12.5%; and in more strongly bound regions, the frequency increases to more than 20%. SRF and intermediate types of CArG boxes showed considerably more enrichment than MEF2-type CArG boxes, suggesting that they are more frequently bound by SEP3 or SEP3-containing protein complexes *in planta*.

The CArG boxes are usually positioned in the center of the peak (peak maximum score position) (Figure 1B). The strong preference of CArG boxes for the center of the peak demonstrates the high positional resolution of ChIP-SEQ experiments combined with our method of peak detection.

In order to further characterize the features of binding sites recognized by SEP3 and its complex partners *in planta*, we used the ChIP-SEQ peak score as a measure of affinity for binding sites within enriched genomic regions as implemented in the MatrixREDUCE software [30]. We found that the obtained consensus models tend to be relatively flexible in most nucleotide positions (Figure 1). This suggests that the in vivo binding data reflect a mix of affinities of different homo- and heterodimeric SEP3 protein complexes.

Since MADS-domain proteins bind as dimers to a CArG box, CArG boxes can be considered as composite elements, with each half-site contacted by a different monomer with possibly different binding preferences. We estimated the frequencies of all possible half-sites for the CArG boxes of the general consensus CC[A/T]$_6$GG and CC[A/T]$_7$G in the ChIP-SEQ data. Using a binomial test, we found that three out of eight possible half-sites for the consensus CC[A/T]$_6$GG were significantly overrepresented in the SEP3 ChIP-SEQ data, and four out of 24 for the consensus CC[A/T]$_7$G (Table S4). Not all possible combinations of the most frequent half-sites are represented among the most strongly overrepresented core CArG boxes (Table S4), suggesting dependencies between the half-sites. The most frequently represented sequence of type CC[A/T]$_6$GG, which is CCAAAAATGG, is in fact the same or highly similar to the consensus sequences that were identified by MatrixREDUCE (Figure 1C).

Based on the combinations of half-site sequences found in the ChIP-SEQ data, we measured the dependencies between single nucleotides using a chi-square test. Here, we found that strong dependencies exist between nucleotides within the [A/T]-rich core of the CArG box (Table S5). Surprisingly, we also identified dependencies for the nucleotides surrounding the core CArG box, which is in line with experimental evidence suggesting that sites surrounding the core consensus are contacted by MADS-domain proteins and may contribute to DNA-binding specificity as well [31–33]. The dependencies between nucleotide positions in functional CArG boxes could (at least partly) explain why only 7.7% and 5.7% of all CArG boxes perfectly matching the consensus CC[A/T]$_6$GG and C[A/T]$_7$GG, respectively, are bound in vivo by SEP3.

According to the "floral quartet" model [34], higher-order MADS-domain protein complexes bind to two CArG box–like DNA sequences at short distance from each other. Thus, we would expect an enrichment of ChIP-SEQ peaks with regulatory modules consisting of two CArG box elements. In order to identify regulatory modules composed of more than one binding site, we used the Explain software [35]. The module with the highest fitness score (0.748 on a scale of 0 to 1) was composed of a single pair of CArG boxes separated by a DNA stretch varying between 10 to 200 bp in length. This finding supports the idea that MADS-domain proteins act in

Tecan Ex. 2114, Page 5 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

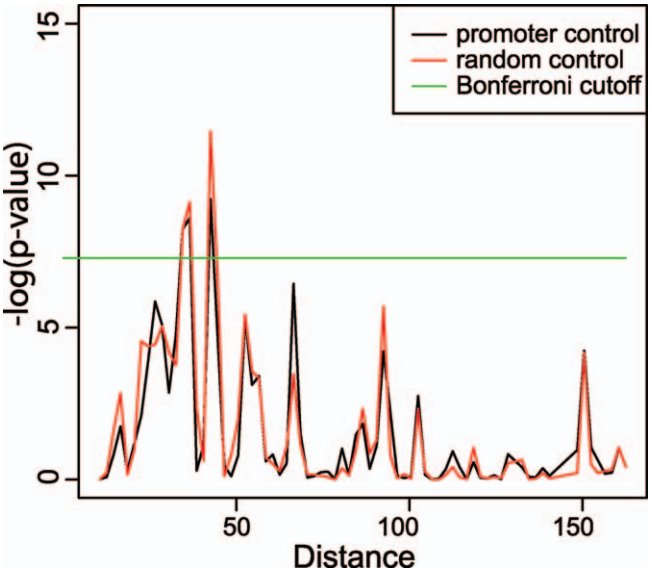


Figure 2. Preferred Distances between CArG Boxes within ChIP-SEQ Peaks

Plotted are the $-\log_e(p\text{-value})$ resulting from a binomial test for enrichment of distances of CArG boxes in wild-type ChIP-SEQ data compared to a random set of promoters (promoter control) and randomized sequences (random control). The Bonferroni correction gives the most conservative cutoff for significantly enriched distances. The most strongly preferred distance is at 42–43 bp.

doi:10.1371/journal.pbio.1000090.g002

complexes composed of two dimers that bind to two adjacent binding sites, as has been predicted by the floral quartet model.

Next, we were interested whether CArG boxes within ChIP-SEQ peaks have preferred distances from each other. To test this, we compared the distribution of distances of CArG boxes within peaks to "background distributions" obtained from random sets of *Arabidopsis* promoters or randomized sequences. As shown in Figure 2, there is a preference for close distances, with the strongest preference around 42–43 bp (which corresponds to four helical turns of the DNA). Aside from these relatively short distances, the frequent occurrence of multiple peaks in the same genomic regions open the possibility that MADS-domain protein complexes can also bridge and bend larger DNA stretches as was suggested in the floral quartet model [34].

Binding Sites for Other Transcription Factors in the ChIP-SEQ Peaks

In addition to our targeted screen for MADS-domain binding sites, we used MatrixREDUCE and MEME [36] in order to identify DNA sequence motifs that are abundant in the regions bound by SEP3. Using these tools, we recovered motifs corresponding to CArG boxes, but interestingly, we also detected sequence motifs potentially bound by non-MADS transcription factors. A motif that was identified using these programs has the consensus sequence CACGTG. This motif has been named "G-box" in the literature and represents a DNA-binding site for bHLH and bZIP transcription factors [37]. We found that this motif is indeed overrepresented in the genomic regions bound by SEP3 (Figure 3A). Similar to CArG boxes, the G-box motif is enriched in the center of the peak (Figure 3B). A second motif that was found using these programs has strong similarity to the DNA-binding consensus of TCP transcription factors with the general consensus motif CCNGGG [38]. We analyzed whether this TCP DNA binding consensus was overrepresented in the regions bound by SEP3. Indeed, it appeared to be enriched with increasing peak score threshold, and it is enriched in the center of the peak (Figure 3A and 3B).

Next, we tested systematically for enrichment of known DNA-binding consensus sequences of transcription factors using information from the Transfac and AGRIS databases. In total, these databases contain information for 105 (Transfac) and 72 (AGRIS) DNA-binding consensus sequences of plant transcription factors. In addition to confirming the enrichment of MADS-, TCP, and bHLH/bZIP binding sites, we found that also ARF, C2H2 (ID1) DNA recognition motifs, and a bHLH (MYC) DNA-binding site similar to the G-box were enriched with increasing peak score threshold, and located in the center of the ChIP-SEQ peaks (Figures 3 and S4).

Overlap of SEP3 Binding Sites in Wild-Type and *agamous* Mutant Backgrounds

The ChIP-SEQ experiments were done with samples from wild-type plants and the *ag* mutant. We were interested in the overlap of DNA-binding sites in these two samples, which could point to target genes involved in the formation of perianth organs. There is clear preference for overlapping genomic positions of SEP3 ChIP-SEQ peak maximum positions in wild type and *ag* mutant (Figure 4A).

The overlap of potential SEP3 targets in wild type and *ag* mutant is also evident from genes that are targeted by SEP3 (Figure 4B). Whereas the number of peaks in *ag* is only approximately 65% of the number of peaks in wild type at FDR = 0.001, the overlap in affected target genes is almost 70% at the same FDR level (Figure 4B). Thus, individual peaks are more likely affected by loss of AG than target genes. With increasing significance level, targets of SEP3 in wild type and *ag* mutant overlap progressively more (Figure 4B), as do individual peaks. Highly enriched target genes are usually common to wild type and *ag* mutant (overlap >90%).

According to these results, only a small fraction of strongly enriched direct target genes is specific to the SEP3 complexes specifying stamen and carpel development (Table S2). These genes may represent candidate genes determining the specific morphologies of stamens and carpels downstream of the floral homeotic genes.

Expression of Candidate SEP3 Targets during Flower Development

In the ChIP-SEQ approach to identify potential SEP3 targets, different floral tissues corresponding to different developmental stages were used. In order to evaluate the relevance of DNA-binding events in the regulation of the genes corresponding to the ChIP-SEQ peaks, we used comprehensive gene expression array data that are publicly available. Mainly, the collection of AtGenExpress experiments provides information about timing of gene expression and changes in different floral homeotic mutants [39].

We found that about 45% of all genes with significantly enriched peaks in the SEP3 ChIP-SEQ experiment (FDR <0.001) were differentially expressed at very young developmental stages in at least one of the homeotic mutants (*lfy*-12, *ap1*-15, *ap2*-6, *ap3*-6, *ag*-12) (Figure S5). This fraction was higher than the overall genome-wide fraction of differentially

Tecan Ex. 2114, Page 6 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

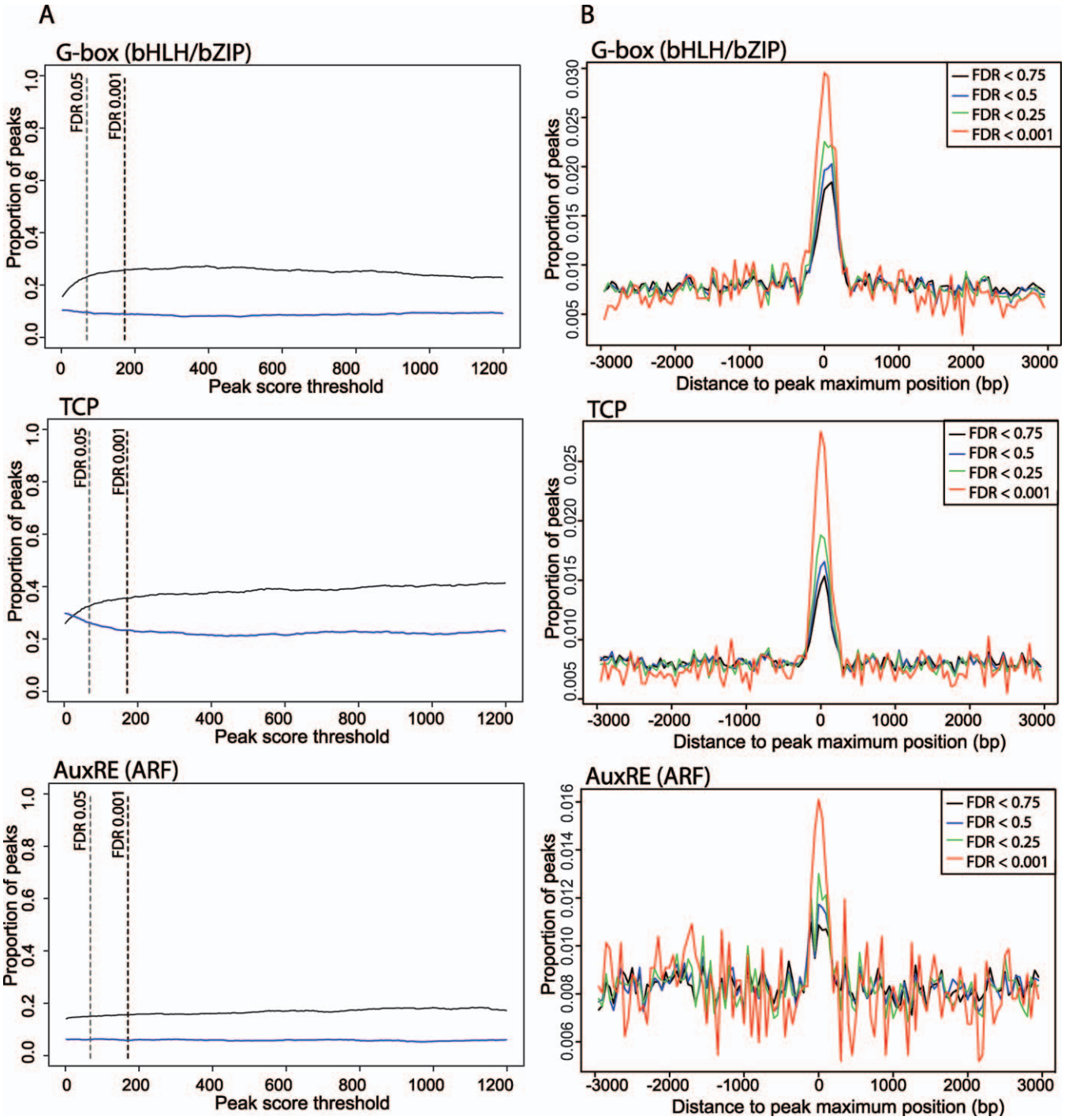


Figure 3. Non–MADS-DNA Binding Motifs Enriched in ChIP-SEQ Peaks

(A) Increase in frequency of the motifs with increasing significance level (peak score threshold). Black line: proportion of peaks with at least one binding site; blue line: proportion of peaks with at least one binding site after permutating the nucleotide positions of the peak areas, keeping the same base composition (i.e., control).

(B) Frequency plots of sequence position of TCP, bHLH/bZIP, and ARF binding sites relative to the peak score maximum position ("center of the peak") (no mismatches allowed). All sequence elements show enrichment in the center of the peak.

doi:10.1371/journal.pbio.1000090.g003

expressed genes in these mutants (29%). Forty-five percent (903/2022) of the genes with ChIP-SEQ peaks in the *ag* mutant are differentially expressed in the *ag*-12 mutant compared to wild type (up to floral stage 12; 28% (5,927/21,039) in the total dataset). Considering SEP3 binding sites in genes that are differentially expressed during development, we found the strongest enrichment for genes that change expression in the meristem during the earliest stages of floral development (FDR <0.001, *p*-value 4.3e–40, binomial test). About 63% of the potential targets of SEP3 are differentially expressed at

Tecan Ex. 2114, Page 7 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

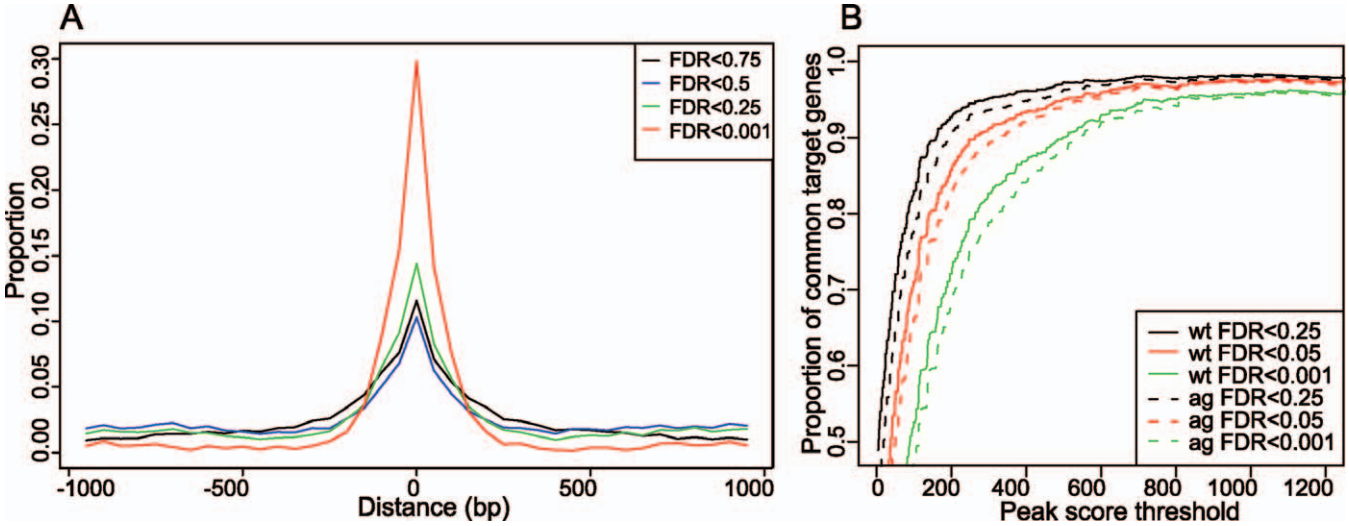


Figure 4. Distance between Peaks in Wild Type and *agamous* Mutant and Overlapping Target Genes

(A) Proportion of distances between significant peaks in wild type and *ag* mutant for different FDR levels. Most peaks cluster within a distance of ±200 bp.

(B) Proportion of common target genes for wild-type and *ag* mutant datasets. The solid line represents the proportion of common targets genes relative to the total number of targets in the *ag* mutant for different FDR levels of the wild-type dataset; the dashed line represents the proportion of common target genes relative to the total number of targets in the wild-type for different FDR levels of the *ag* mutant dataset. Since the total number of significant peaks is lower in *ag* mutant than in the wild-type dataset, the dashed line is below the solid line.

doi:10.1371/journal.pbio.1000090.g004

any stage of reproductive development starting from floral transition to flowers of stage 12 (Figure S5).

In total, 72% of the potential SEP3 targets are differentially expressed during flower development or in any of the homeotic mutants. Although the differential expression can also be due to indirect effects, the data suggest that the majority of potential direct SEP3 targets may also be regulated by SEP3. We also found an enrichment in frequency of genes that are correlated in expression with SEP3 expression with increasing peak score threshold in ChIP-SEQ (Figure S5). The fraction of genes with ChIP-SEQ peaks that are positively coexpressed with SEP3 is clearly higher than that of negatively regulated genes, supporting the idea that SEP3 acts mostly as a transcriptional activator.

Evidence for Direct Regulatory Interactions among MADS-Box Transcription Factors

Genetic and gene expression experiments suggest that the *SEP* genes are required for the up-regulation of floral homeotic genes, and that this up-regulation is crucial for the establishment of the identities of the different floral organs. However, until now, it has not been demonstrated whether this regulation is direct.

We analyzed the binding profiles for the genomic loci corresponding to the floral homeotic genes and found that SEP3 binds to nearly all of these loci (Figure 5). Only *SEEDSTICK* (*STK*) and *CAULIFLOWER* (*CAL*) do not have significantly enriched regions. In most cases, the peaks are located in the promoters of the respective genes. In case of the *APETALA1* (*AP1*), *APETALA3* (*AP3*), *SEP1*, and *SEP2* loci, there are also peaks in the 5′ UTR, whereas a SEP3 binding site is present in the second intron of *AGAMOUS* (*AG*).

Of all homeotic genes, the genomic regulatory sequences controlling the expression of *AG* and *AP3* are best characterized. The spatial expression pattern of *AP3* is driven by regulatory elements within approximately 500 bp upstream of the transcriptional start. CArG boxes in this part of the promoter are important for the positive as well as negative regulation of *AP3* [40]. Our ChIP-SEQ results demonstrate that SEP3 binds to the genomic region comprising positively and negatively acting CArG boxes in the *AP3* promoter, strongly suggesting a direct molecular link between the binding of SEP3 and the regulation of *AP3*.

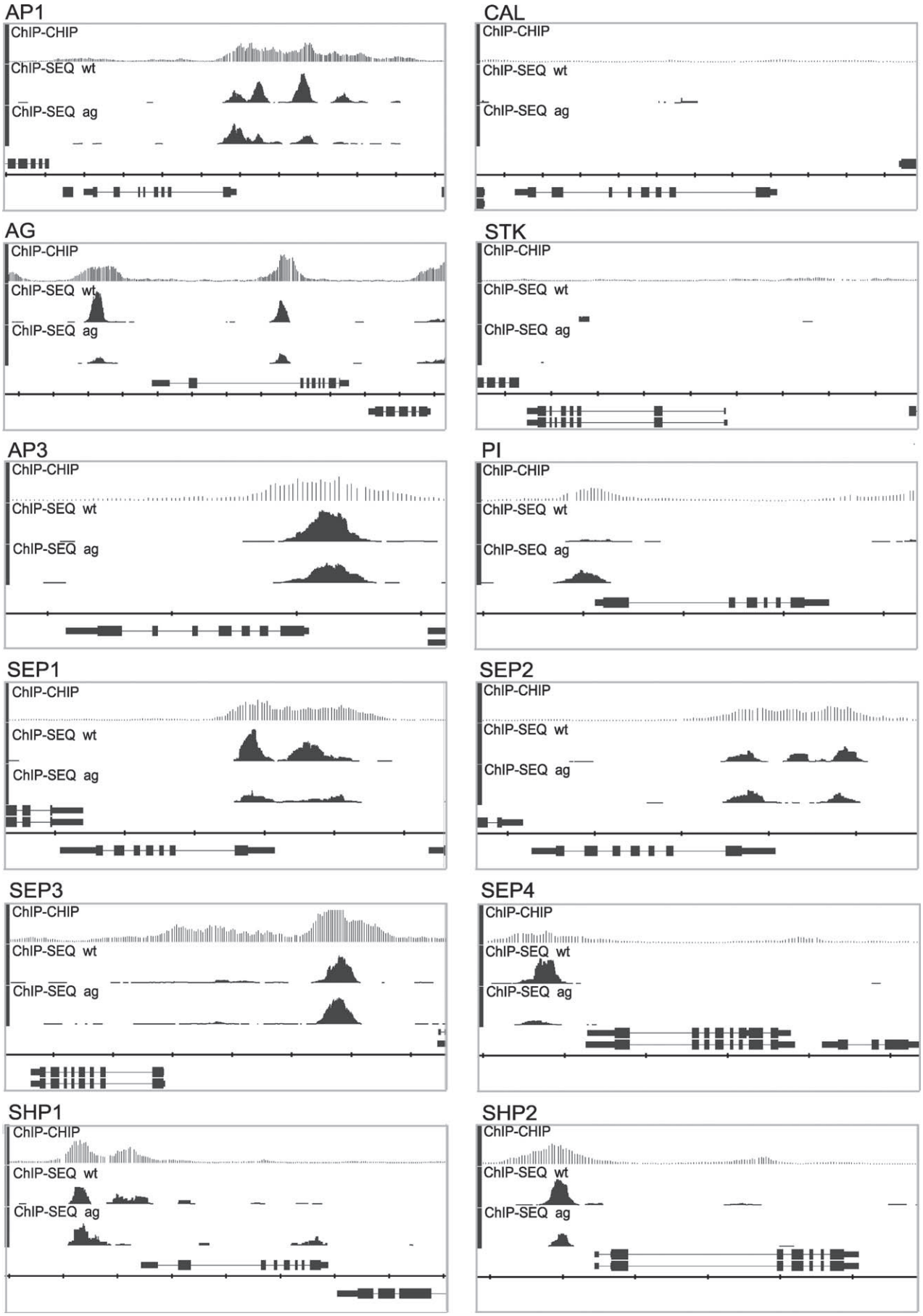
Most regulatory sequences controlling the expression of *AG* are located in its second, 4-kb large intron [24,25]. Consistent with this observation, we identified a peak of enrichment of SEP3 in this intron. More specifically, the peak marks a CArG box in the 3′ activation domain located in this intron. The 3′ activation domain functions in the up-regulation of *AG* in stage 3 floral meristems and is also responsible for maintenance of *AG* expression in developing carpels [24]. This CArG box was also found to be bound by AG itself in previous experiments [26]. Interestingly, we identified a second peak of enrichment in the upstream region of *AG*. Consistent with the idea that an AG/SEP heterodimer is responsible for the positive autoregulation of *AG* [26], the heights of the peaks in the *AG* locus are reduced in the *ag* mutant compared to wild type (Figure 5).

The regulatory sequences controlling the expression of the *SEP1–4* genes are still not well characterized. Our ChIP-SEQ results, however, strongly suggest autoregulation of the redundantly acting *SEP* MADS-box genes.

All floral MADS-box genes that are targeted by SEP3 in wild type, are also targeted in the *ag* mutant, although there is some variation in the heights or presence of individual peaks (e.g., *AP1*, *SEP2*, *SHP1*; see Figure 5). This raises the possibility that different SEP3 complexes may have different affinities to individual binding sites.

In order to characterize the regulatory effects of SEP3 on MADS-box genes that are potential direct targets, we

Tecan Ex. 2114, Page 8 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028



Tecan Ex. 2114, Page 9 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Figure 5. Binding Profiles of Floral Homeotic MADS-Box Gene Loci

For each locus, ChIP-CHIP and ChIP-SEQ profiles are depicted for SEP3 in wild type (wt) and in the *ag* mutant. The TAIR annotation of the genomic loci is shown at the bottom of each panel. If the genomic locus is shown above the scale, it is in forward orientation, and if it is in the bottom of the scale, it is in reverse orientation. The scale division corresponds to 1,000 nt. In most cases, the enrichment is in the upstream regions of the respective genes.
doi:10.1371/journal.pbio.1000090.g005

analyzed the expression of these MADS-box genes upon SEP3 induction using a constitutively expressed translational fusion of SEP3 to the rat glucocorticoid receptor hormone binding domain (GR). For this, seedlings expressing the 35S:SEP3-GR construct were treated with dexamethasone (DEX) for 8 h, 1 d, or 10 d, and the expression of MADS-box genes was determined by real-time reverse transcriptase (RT)-PCR. The relative expression levels in comparison with nontreated plants is shown in Figure 6. Our results reveal that SEP3 is indeed able to activate the expression of other floral homeotic genes as suggested previously [6]. *SEP3* itself is most strongly up-regulated, demonstrating a strong auto-regulatory feedback loop. Although some of the tested genes show an early response to SEP3 induction, others are regulated only after prolonged SEP3 induction, suggesting that SEP3 alone is not sufficient to regulate these genes, but needs to interact with partner proteins that are encoded by the induced MADS-box genes. In particular, *AP3*, *AG*, and *AP1* are strongly activated by SEP3. These three genes correspond to the three major classes of floral homeotic genes according to the classical ABC model: class A (*AP1*), class B (*AP3*, together with PI), and class C (*AG*). Their gene products also represent major protein interaction partners of SEP3, suggesting that later induced targets of SEP3 are targets of the corresponding SEP3-containing protein complexes. Thus, SEP3 can activate the flower developmental program by enhancing the expression of its interaction partners as one of its first steps. Induction of the ABC classes of genes is sufficient to form the flower, which explains the very early flowering and the terminal-flower phenotype that we observed upon SEP3 induction. Whereas flower-specific genes are mostly activated, MADS-box genes that are involved in the floral transition (*AGL24*, *SOC1*, and *SVP*) tend to be down-regulated by SEP3. Together with the fact that SEP3 binds to the promoters of these genes, our results suggest that SEP3 is involved in the down-regulation of these genes during early flower development, possibly as part of protein complexes together with other flower-specific MADS-domain proteins, such as AP1.

Functional Characterization of SEP3 Target Genes

It has been a long-standing question in plant developmental biology whether floral homeotic genes act directly on the structural or metabolic genes that create the final morphology of floral organs, or whether they act via intermediate regulators, i.e., other transcription factors, which in turn regulate subsets of targets conferring the final organ shape and function. To answer this question, we investigated the enrichment of gene ontology (GO) terms [41] among genes that are closest to the peaks as a function of their ChIP-SEQ peak score.

In terms of molecular function, genes encoding transcription factors are clearly the most enriched group of genes (GO:0030528; *p*-value 3.62e−19). When dissecting gene functions according to biological processes, there is a clear enrichment for genes involved in development, in response to hormonal stimuli, and in lipid biosynthesis. Figure 7 presents the top-five most enriched specific GO terms. The SEP3 targets in the GO category "lipid biosynthetic process" include genes involved in hormone biosynthesis (terpenoid and steroid pathways), as well as in sterol and wax synthesis.

Next, we were interested in whether some transcription

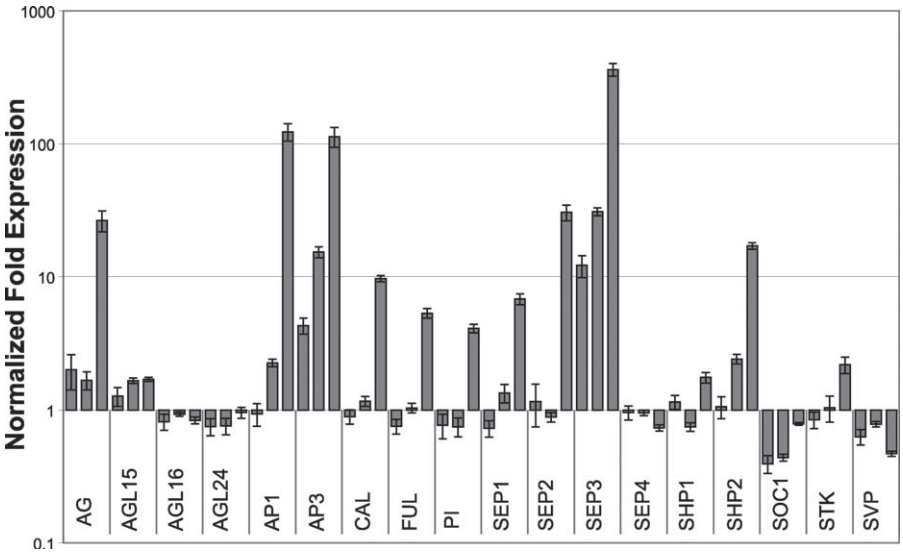


Figure 6. Differential Expression of SEP3 Targets of the MADS-Box Gene Family after SEP3 Induction

SEP3 expression was induced in seedlings for 8 h, 1 d, or 10 d, and the change in expression of selected MADS-box genes relative to noninduced seedlings was determined by quantitative RT-PCR. The standard error represents the variation between two independent biological and technical replicates. All of the depicted genes, with the exception of *CAL* and *STK*, are bound by SEP3 (see Figure 5). In line with being indirect targets, these two genes are only up-regulated 10 d after SEP3 induction.
doi:10.1371/journal.pbio.1000090.g006

Tecan Ex. 2114, Page 10 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

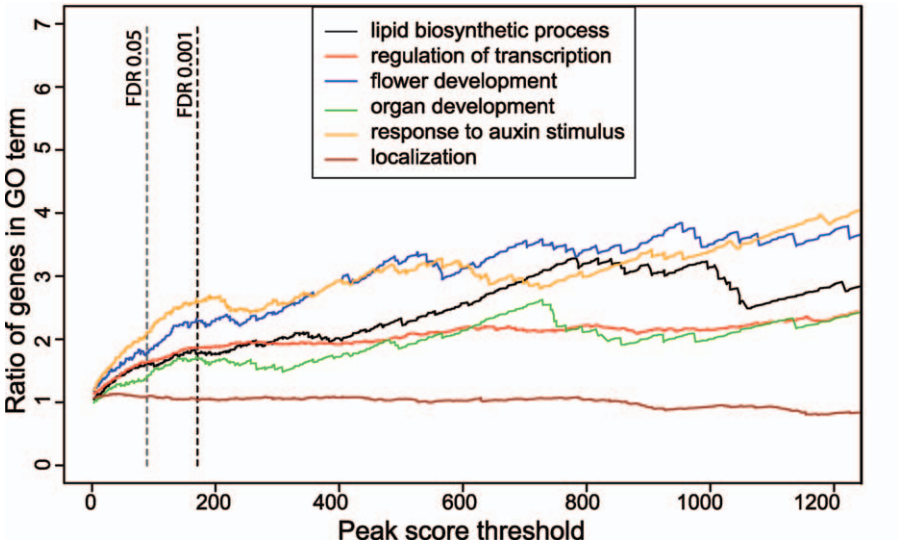


Figure 7. Enrichment for GO Terms in the Category "Biological Process"

The five most strongly enriched GO terms of the lowest possible level (minimum 20 annotated loci) in the GO hierarchy are presented. The GO term "localization" is shown as an example for a nonenriched category.

doi:10.1371/journal.pbio.1000090.g007

Table 2. Significantly Overrepresented Transcription Factor Families

Family	Wild-Type[a]	*ag* Mutant[a]	Common Loci	Total (Genome)	Known Functions[b]
SBP	10 (0.0002)	9 (8.84E−05)	9	16	Transition to flowering [86], leaf development [87]
GRF	6 (0.0010)	4 (0.0121)	3	9	Organ growth [88]
AUX-IAA	11 (0.0288)	12 (0.0007)	10	29	Auxin response [89]
C2C2-Co-like	18 (0.0003)	12 (0.0090)	11	37	Flowering time [90]
C2C2-Dof	14 (0.0142)	11 (0.0186)	9	36	Diverse developmental processes
bHLH	43 (0.0036)	35 (0.0019)	30	127	Diverse processes
TCP	11 (0.0032)	9 (0.0038)	8	23	Organ growth [91]
GRAS	15 (0.0018)	11 (0.0088)	8	33	Meristematic patterning, GA response[92]]
ARF	9 (0.0313)	8 (0.0134)	7	23	Auxin response [89]
C2C2-GATA	8 (0.1508)	9 (0.0106)	7	26	Light regulation [93]
BES1	2 (0.299)	3 (0.039)	2	8	Brassinosteroid & auxin response [94]
HB	32 (0.0080)	19 (0.1344)	16	87	Diverse developmental processes
TAZ	5 (0.0080)	1 (0.4999)	1	9	Stress response [95]
PLATZ	5 (0.0161)	1 (0.5597)	1	10	
Trihelix	8 (0.1508)	8 (0.0311)	7	26	Auxin homeostasis. perianth architecture [96,97]

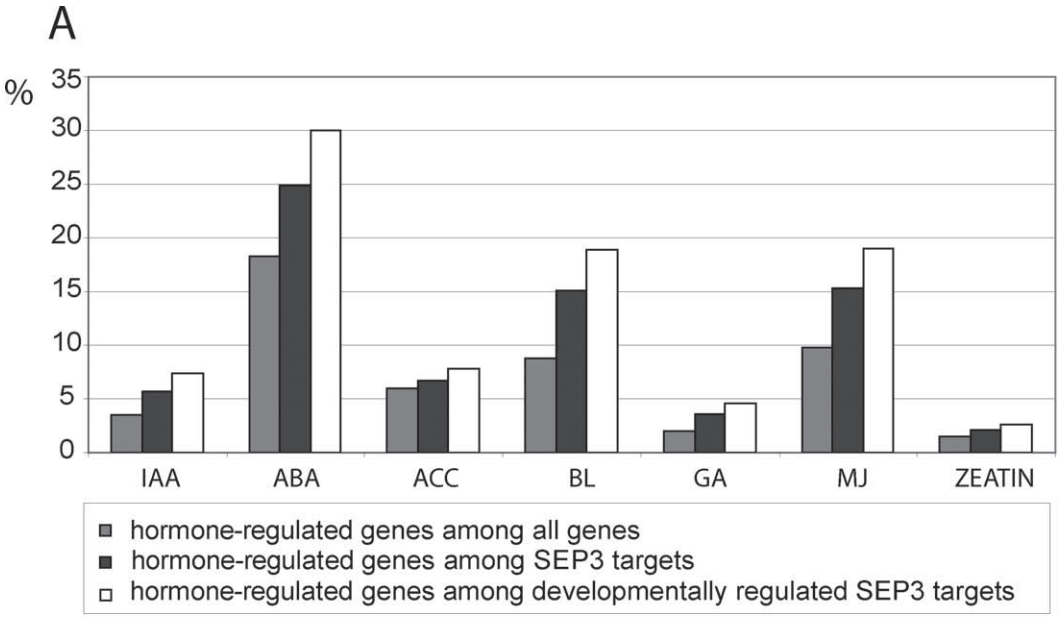
[a]The *p*-values of hypergeometric test for genes near peaks with FDR <0.001 are given in parentheses.
[b]Example original publications or reviews are indicated.

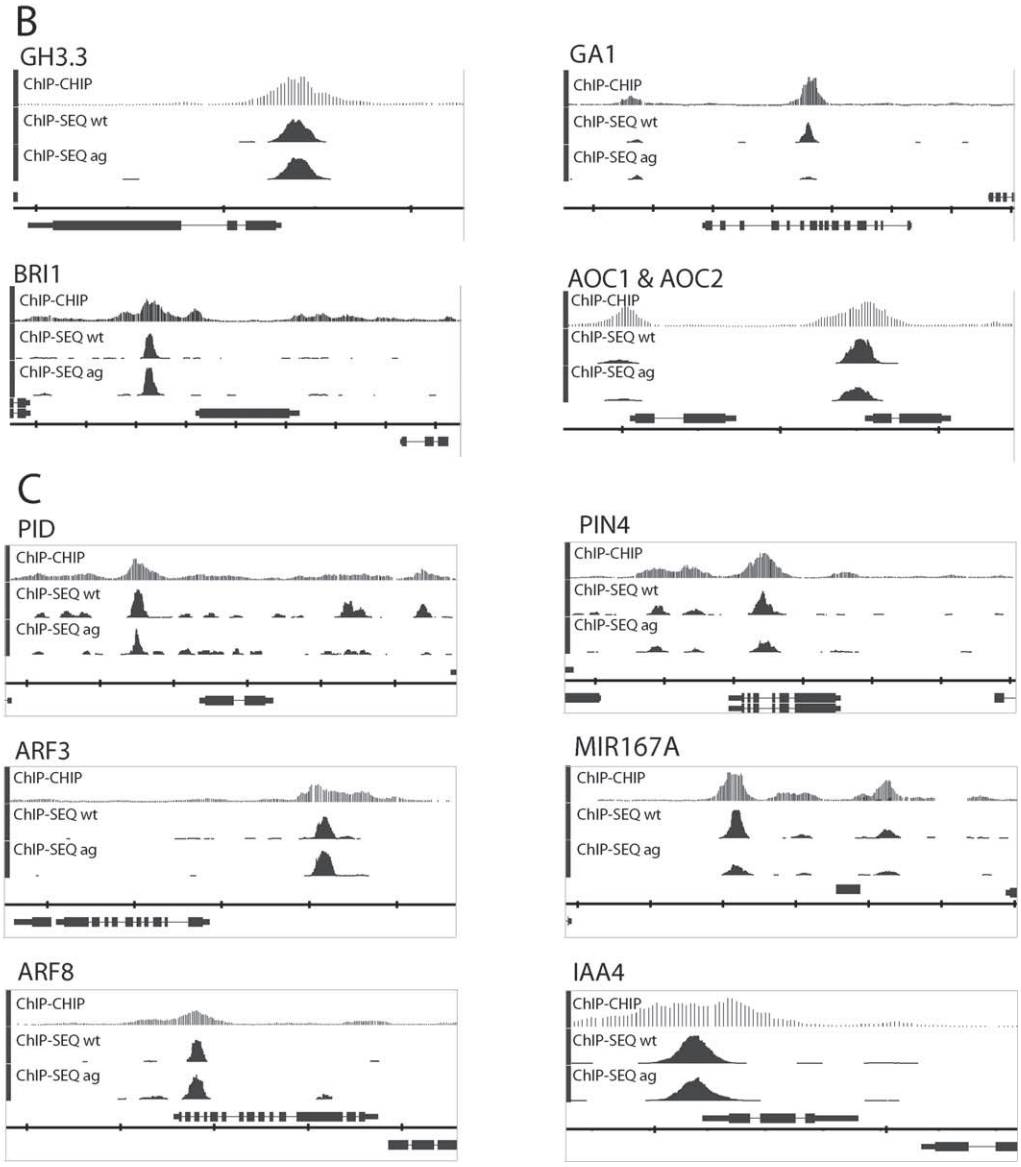
doi:10.1371/journal.pbio.1000090.t002

factor families were more frequently represented among potential direct SEP3 target genes than others. The results shown in Table 2 reveal that 15 transcription factor families were significantly overrepresented among SEP3 targets. Interestingly, we found overrepresentation of families for which we also found enrichment of their DNA-binding sites in the ChIP-SEQ data: bHLH, TCP, and ARF families are overrepresented in both the target and the binding site datasets, which points to the existence of autoregulatory feedback loops (Figure 5, Table 2). In general, we found overrepresentation of transcription factor families with known functions in the control of flowering time (SBP and C2C2-Co-like), organ growth (GRF and TCP), auxin response (AUX-IAA and ARF), brassinosteroid response (BES1), and meristem development (HB and GRAS). Sixty-six percent of the loci belonging to these families identified in wild type, were also found in the *ag* mutant at the same FDR threshold.

The characterization of overrepresented GO categories and transcription factor families suggests that SEP3 is involved in the regulation of hormonal signaling. In order to further understand this link, we analyzed the overlap between the potential direct downstream targets of SEP3 and genes regulated by different hormones [42]. We found that genes regulated by auxin, gibberellic acid, and brassinosteroids were most represented among SEP3 targets, but also genes responding to other hormones were enriched (Figure 8A). Among the top 200 genes targeted by SEP3 are several enzymes involved in hormone biosynthesis (e.g., *GA1* and *AOC2*), signaling (e.g., *BRI1*), or homeostasis (e.g., *GH3.3)* (Figure 8B). *GH3.3* is involved in auxin homeostasis and has

Tecan Ex. 2114, Page 11 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028





Tecan Ex. 2114, Page 12 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Figure 8. Hormonal Signaling Targeted by SEP3

(A) Characterization of hormone-regulated genes among SEP3 targets. The fraction of hormone-regulated genes among all genes represented on the ATH1 microarray is given (grey box), as well as the fraction of hormone-regulated genes among SEP3 targets (black box) and the fraction of hormone-regulated genes among SEP3 targets that are differentially expressed during reproductive development (white box). Hormone-regulated targets are overrepresented among SEP3 targets, with the strongest overrepresentation for IAA-, BL-, and GA-regulated genes (enrichment >1.6-fold among all SEP3 targets and >2.1-fold among developmentally regulated targets). Data on hormone-responsive genes (low stringency) were taken from [42]. Developmentally regulated SEP3 target genes were the ones identified using the AtGenExpress microarray datasets (Figure S5). ABA, abscisic acid; ACC, ethlyene; BL, brassinolide; GA, gibberellic acid; IAA, indole acetic acid; MJ, methyl jasmonate.

(B) Representative key regulatory enzymes and hormonal signaling genes among the top 200 most strongly enriched targets of SEP3. *GH3.3* encodes an IAA-amido synthase that is involved in auxin homeostasis [78]. *GA1* encodes a key enzyme in giberellic acid biosynthesis [79]. *BRI1* encodes a receptor kinase mediating brassinosteroid signal transduction [80]. *AOC1* and *AOC2* gene products catalyze an essential step in jasmonic acid biosynthesis [81].

(C) Different steps in the auxin signaling pathway targeted by SEP3. PID (kinase) and PIN4 (auxin efflux carrier) are important for auxin transport [47,82]. *ARF3* (*ETT*) and *ARF8* are members of the auxin response factor (ARF) family of transcription factors [49,83]. *MIR167A* is one of the genomic loci encoding for a miRNA targeting ARF6 and ARF8 [84]. *IAA4* is an auxin-induced gene [85]; its gene product possibly acts as antagonist of ARF transcription factors. In (B) and (C), the TAIR annotation of the genomic loci is shown at the bottom of each panel. If the genomic locus is shown above the scale, it is in forward orientation, and if it is in the bottom of the scale, it is in reverse orientation. The scale division corresponds to 2,000 nt. wt, wild type.

doi:10.1371/journal.pbio.1000090.g008

been found to be up-regulated as a later response of carpel and stamen induction by AG [26].

Since auxin signaling was consistently found to be over-represented among our potential direct SEP3 targets, we further analyzed SEP3 binding patterns at auxin-related genes (Figure 8C). Most Aux-IAA genes showed similar SEP3 binding patterns, with a peak close to the transcriptional start site (similar to *GH3.3*). Next, to genes involved in auxin transport (e.g., *PIN4* and *PID*) and auxin response factors with known roles in flower and fruit development (e.g., *ARF3*, *ARF6*, and *ARF8*), a miRNA167 locus, which controls *ARF6* and *ARF8* expression, is also found among the targets.

Novel Roles of SEP3 Unraveled by Dominant Repression of Direct Target Genes

SEPALLATA genes have shown to be crucial for the specification of floral meristem and organ identities [6,43]. However, the redundancy among MADS box genes makes the functional characterization of members of this gene family difficult [44]. Indeed, our analysis of potential downstream targets of SEP3 suggests so far unknown links to other developmental and hormonal processes in flower development. Since SEP3 acts mostly as a transcriptional activator, we can alternatively study the functions of this protein by replacing endogenous SEP3 function with that of a chimeric repressor version of SEP3. For this purpose, we fused the genomic coding region of SEP3 to the EAR (ERF-associated amphiphilic repression) domain under the control of a basic SEP3 promoter (−960 bp) [45]. This chimeric repressor blocks the activation of direct target genes of the transcription factor complexes in which SEP3 is present.

Our transgenic approach indicated novel functions of SEP3 in addition to supporting previously proposed roles: transgenic lines with a strong phenotype showed delayed flowering (unpublished data), reduced number and size of floral organs, as well as defects in organ differentiation (Figure 9) and identity. While petals were mostly absent in these plants, the stamen number was strongly reduced, and the stamens were often reduced to filamentous carpelloid structures or fused with the carpels (Figures 9C and 9D). Plants with strong phenotypes were male and female sterile. The carpel of severely affected lines showed severe growth defects: the size of the ovary was greatly reduced or it was even missing, while the gynophore at the bottom and the style at the top of the gynoecium were enlarged. Ovule placentation was abaxialized in a variable manner (Figure 9G). The aberrant carpel morphology closely resembles the phenotype of mutants impaired in auxin biosynthesis or signaling [18]. The carpel phenotypes are similar to those of *pin1* [46], *pinoid* (*pid*) [47,48], or *arf3/ett* mutants [49,50,51] (Figure 9B and 9F, versus 9D and 9E). In *sep1 sep2 sep3* triple-mutant plants, elongated gynophores similar to the *pid* mutant can be observed (Figure 9J and 9K). The receptor kinase PID determines the polar localization of auxin efflux (PIN) proteins, whereas ARF3, along with other ARF transcription factors, mediates auxin response at the gene regulatory level.

Our ChIP-SEQ data suggest multiple links of SEP3 and the auxin signaling pathway: ARF genes and their antagonists, the AUX-IAA factors, as well as ARF-controlling microRNA loci are bound by SEP3 as revealed by the ChIP-SEQ experiments. Also, a genomic region downstream of the *PID* locus is targeted by SEP3, both in wild type and in the *ag* mutant. In agreement with a role of SEP3 regulating these genes, expression microarray data of developmental time series suggest that a majority of ARF transcription factor genes bound by SEP3, as well as *PID* and a smaller number of AUX-IAA genes, are up-regulated in reproductive meristems and young floral stages, in a similar fashion to SEP3 itself (Figure S6).

The enrichment of auxin response elements (ARF binding sites) in the SEP3 ChIP-SEQ peaks suggests that SEP3 cooperates with ARF proteins in target gene regulation, thus the downstream targets of these complexes could be the target of repression by the SEP3-EAR protein. The need for cofactors in auxin response is illustrated by the finding that (positive) auxin response, as measured by the DR5 promoter, is only found in a subset of SEP3-expressing cells. Whereas the expression of SEP3 in undifferentiated meristematic tissues is broader than that of the DR5 marker, the expression domains overlap mostly in growing organs such as the tips of growing sepals (Figure 9H and 9I). DR5 expression has been shown to be also dependent on brassinosteroid signaling, which suggests a complex pattern of upstream regulation of auxin response [52].

Discussion

In order to determine genome-wide binding sites of the MADS-box transcription factor SEP3, we performed ChIP-SEQ and ChIP-CHIP experiments and found a high correlation between the datasets, although the ChIP-SEQ approach provides a higher resolution of the binding pattern. The ChIP

Tecan Ex. 2114, Page 13 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

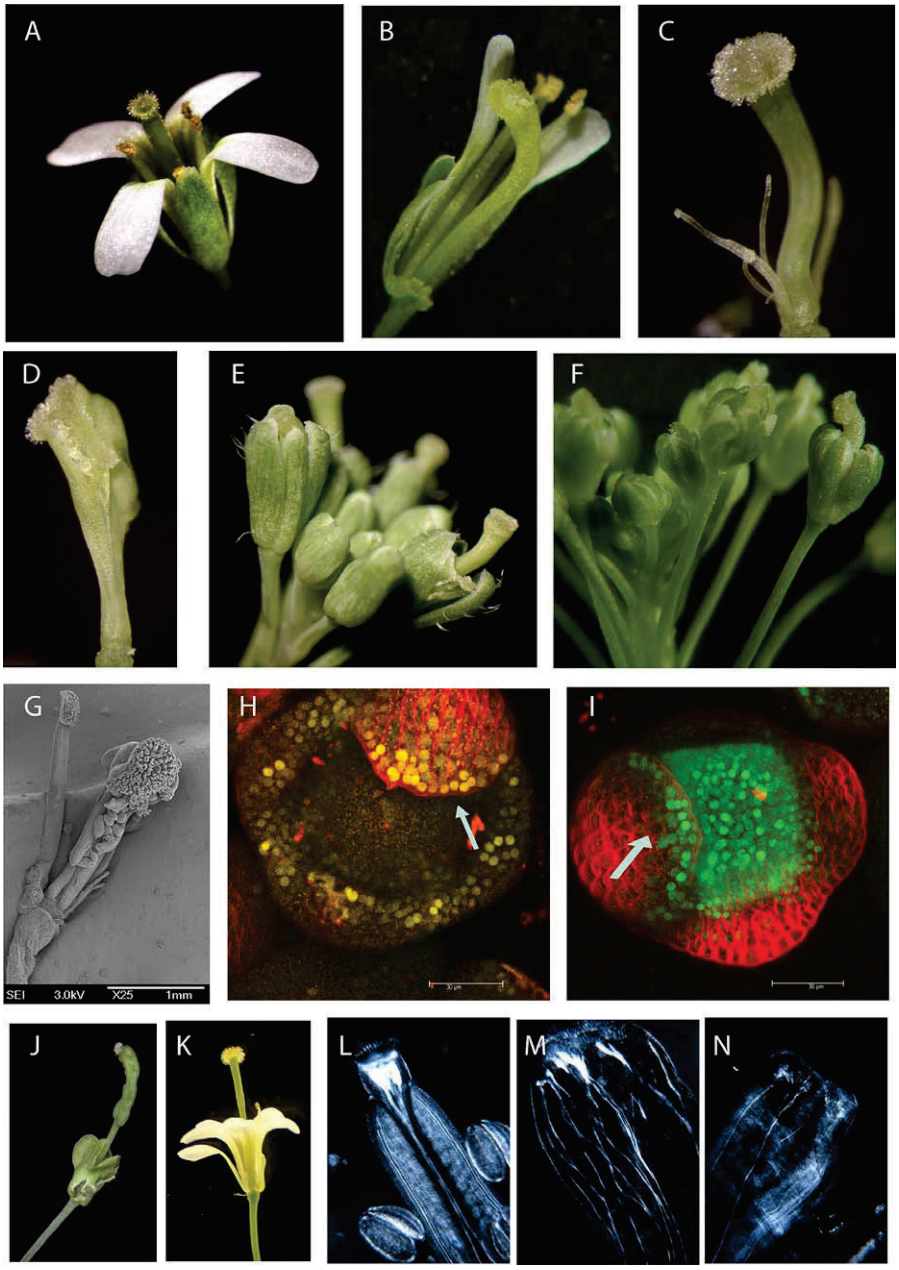


Figure 9. Role of SEP3 in Auxin Signaling

(A) Wild-type *Arabidopsis* flower.
(B) *ett* mutant flower, two sepals and petals removed.
(C and D) Flowers of *SEP3-EAR* plants (*pSEP3::SEP3-EAR* in *sep3–1* mutant) with strong mutant phenotype, sepals (and filamentous organs in [D]) were removed to reveal the inner organs.
(E) Inflorescence of *SEP3-EAR* plant.
(F) Inflorescence of *ett-1* mutant.
(G) SEM picture of a flower from a *SEP3-EAR* plant revealing abnormal ovule placentation and enlarged stigmatic tissue (sepals were removed).
(H and I) Localization of nuclear localized DR5::YFP (yellow) in a floral meristem (H). Localization of SEP3-GFP driven from its own promoter in a floral meristem (I). Arrows indicate similar localization of the DR5 marker and SEP3 in sepal tips. White bar indicates 30 μm.
(J) Flower of a *sep1 sep2/SEP2 sep3* mutant plant with stalked carpel.
(K) *pid* mutant flower with stalked carpel.
(L–N) Venation patterns in wild-type (L) and SEP3-EAR carpels (M) and (N).
doi:10.1371/journal.pbio.1000090.g009

data demonstrate that the floral MADS-box transcription factor SEP3 bind to thousands of regions in the *Arabidopsis* genome. Thus, it apparently acts as global regulator of gene expression during the various stages of floral development from floral meristem initiation to maturity. Also, other MADS-box transcription factors bind to thousands of sites in the *Arabidopsis* genome (AP1; K. Kaufmann and G. C. Angenent, unpublished data), indicating that these transcription factors are key regulators controlling the expression of other regulatory genes and structural genes.

Tecan Ex. 2114, Page 14 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

Functional Relevance of Binding Events

Despite the large number of genomic binding sites, only a small fraction of the potential binding sites that are present in the genome, represented by a CArG box, are indeed bound by SEP3. de Folter and Angenent (2006) [12] calculated that the *Arabidopsis* genome contains far more CArG boxes than the roughly 30,000 genes in the genome, whereas our ChIP-SEQ experiments resulted in enriched binding sites in about 3,400 genes (FDR <0.001). This indicates that the binding of SEP3 to DNA sequences is highly selective. Most binding sites are located in promoters or other regions with potentially regulatory functions, such as introns. These findings suggest that the majority of significant protein–DNA interactions identified in ChIP-SEQ experiments are likely to be relevant. The functional importance of the DNA-binding events detected in our ChIP-SEQ experiments is also supported by the finding that there is enrichment for specific GO annotations among the targets of SEP3.

The question remains whether all of the enriched DNA regions bound by SEP3 are relevant for gene regulation. Since the plant material used for our ChIP-SEQ experiments comprises different developmental stages, a direct correlation of DNA-binding events and stage-specific changes in gene expression is difficult. However, our comparison of the ChIP-SEQ data to comprehensive gene expression microarray data suggest that the majority of the genes that are bound by SEP3 *in planta* are also differentially expressed during flower development in a temporal and/or spatial fashion.

In contrast to the strongly bound regions, the functional relevance of weakly enriched regions detected in ChIP-SEQ is much less clear. They might represent transient interactions of the transcription factors with DNA [53], or DNA binding in only a few cells, resulting in a high dilution with tissues that lack the interaction. In a comparable study to decipher binding sites of transcription factors during *Drosophila* embryo development, Li et al. came to the conclusion that a significant proportion of the poorly bound regions are most likely nonfunctional [15]. These regions corresponded to genes that were poorly modulated in expression, and peaks were often located outside regulatory sequences. It is possible that the weakly enriched regions in our SEP3 ChIP experiment also contain a high proportion of inactive binding sites.

Determinants of In Vivo DNA Binding

Although in vitro studies have revealed consensus binding sites for many transcription factors, including MADS-domain proteins [12,32], it remains unknown what DNA sequence or chromosomal context determines DNA binding site recognition by MADS-box transcription factors in vivo. Our results indicate that the sequence of the cognate binding site, the presence of multiple binding sites, and binding sites for non-MADS cofactors play roles in binding site recognition. The SEP3 ChIP data revealed that not all DNA sequences corresponding to the canonical consensus motifs CC[A/T]$_6$GG and CC[A/T]$_7$G/C[A/T]$_7$GG can serve as functional binding sites for SEP3 and associated MADS-domain proteins *in planta*. Instead, we found significant enrichment for particular types of half-sites in the SEP3 ChIP-SEQ data. In addition, there is clear interdependence between the sequence of individual half-sites, and between half-sites and surrounding bases. Thus, commonly used consensus sequences and position weight matrices, which assume independency of nucleotide positions of a binding site, are highly oversimplified. The enrichment for certain half-sites may reflect the recognition site for SEP3, while the other half of the binding site could match the recognition site of the dimer partner of SEP3. It is known that SEP3 is able to dimerize with many other MADS-domain proteins in yeast assays [8] and *in planta* (K. Kaufmann and G. C. Angenent, unpublished data), which makes it difficult to determine *the* consensus CArG box sequence for SEP3. Combining ChIP-SEQ data obtained from different MADS dimer partners (e.g., SEP3 and AP1) and determining the overlap in binding sites will elucidate the recognition sites for particular MADS-box dimers. We found that there is a clear overlap of target genes in wild type compared to the *ag* mutant, although there are differences in number of binding sites and binding affinity at individual sites. Although at this point we cannot exclude that the wild-type binding sites are enriched for perianth-specific targets, these results support the hypothesis that different MADS-domain protein complexes (e.g., AG-SEP3 and SEP3-AP1) may bind to overlapping sets of target genes, but regulate them in a different way, for instance by recruiting different sets of cofactors leading to differences in activation or repression of genes. In line with the idea that cofactors play a role in differential regulation of downstream targets, it has been shown that SEP3 and AP1 can recruit the corepressor SEUSS, and that this complex acts to repress AG expression in petal development [11]. Our finding that individual peaks are more likely to be affected by loss of AG than target genes makes it also possible that different higher-order complexes may bind to subsets of binding sites present in the promoter. Considering that plant MADS-domain proteins differ in their DNA bending characteristics as determined by gel retardation experiments [28], different protein complexes may have specific effects on the structural properties of the promoter and by that influence target gene expression.

The interplay between different proteins implicates that the dynamics in the floral developmental network is strongly dependent on relative quantities and affinities of individual homeotic proteins competing for common protein interaction partners and key downstream targets. The large number of targets suggests that floral homeotic protein complexes globally control and modify the genetic programs that are active in all plant organs, so that only a limited number of targets would be expected to be unique for the individual floral homeotic protein complexes.

In addition to CArG box elements, we found several other transcription factors consensus DNA-binding sequences to be enriched in the center of the peaks, suggesting that these transcription factors act in a combinatorial fashion with MADS-domain proteins in common regulatory modules. Until now, there has been very limited information about interactions between plant MADS-domain proteins and other transcription factors, and this rare information is based on artificial yeast and/or in vitro protein interaction data. The types of transcription factors whose binding sites are enriched in the peaks link MADS-domain proteins with other cellular, developmental, and hormonal pathways. TCP transcription factors have been shown to be important for cell growth [54–56], and ARF transcription factors are key mediators of auxin response (reviewed in [57]). Whether

Tecan Ex. 2114, Page 15 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

there are direct protein interactions between these classes of transcription factors still needs to be resolved.

In addition to primary sequence characteristics, the accessibility of binding sites, and thus the chromatin structure, can influence the recruitment of MADS-box transcription factors to specific sites in the genome. We observed a strong enrichment of binding sites close to the transcriptional start site of genes, suggesting that the recognition sites are not randomly distributed in the genome. Assuming that any small sequence motif is randomly distributed in the genome, it suggests that not only the primary sequence itself, but also the position of the *cis*-element is relevant for transcription factor binding. Chromatin remodeling, which is active throughout plant development, is likely affecting the accessibility of transcription factors and transcriptional activity. It occurs during post-embryonic developmental transitions (e.g., flowering) and is required for maintenance of meristematic cell identity as well as for the formation of organ primordia, with different chromatin factors acting at different stages ([58] and references therein). Current models of gene regulation suggest that chromatin remodeling and transcription factor binding dynamically alternate due to transient exposure of DNA by displacement of nucleosomes ([59] and references therein). MADS-box transcription factors form large complexes *in planta* (K. Kaufmann and G. C. Angenent, unpublished data), making it even possible that there is a direct interaction between these transcription factors and chromatin remodeling factors. Interesting in this respect is that we identified four out of five members of ATP-dependent chromatin remodeling complexes as targets of SEP3 in the ChIP experiments, suggesting an interdependence relationship.

Auxin Signaling as a Target Pathway of SEP3

Our ChIP-SEQ data indicate that there are multiple direct molecular links between floral homeotic genes and hormonal pathways. Interestingly, ARF genes are targets of SEP3 but could also be coregulators, suggesting that autoregulatory circuits exist involving members of unrelated transcription factor families.

In addition to specifying floral meristem and organ identity, an important role of SEP and AP1-like proteins is to trigger organ growth. The *ap1 cal* double mutant produces multiple undifferentiated meristems, in which organ outgrowth and differentiation are impaired [60]. Similar mutant phenotypes arise by combining *ap1* and *sep* mutant alleles, suggesting partial redundancy between members of these subfamilies of MADS-box genes [43]. This hypothesis of a mutual role of AP1/CAL and SEP proteins in organ outgrowth provides a link to auxin-mediated organ development. Interestingly, the expression of SEP3 fused to the EAR suppression domain led to phenotypes that mimicked developmental aberrations observed in *pin1*, *pid*, or *ettin* (*arf3*) mutants or plants treated with the auxin transport inhibitor NPA [61]. These plants are characterized by defects in lateral organ outgrowth (floral buds and floral organs) and a pistil lacking a functional ovary. Although we observed variable homeotic conversions in SEP3-EAR expressing plants, which would be expected from the down-regulation of SEP3 function, the majority of plants was only affected in outgrowth and differentiation of the floral organs. The mode of action of the SEP3-EAR fusion protein and how it interferes with outgrowth without affecting the homeotic function remain to be studied further. A possible explanation could be that auxin homeostasis is more sensitive to the dominant repression by the SEP3-EAR protein than the floral homeotic functions. Alternatively, SEP3-EAR may suppress ARF function by interacting with ARF proteins in a larger transcriptional complex, a model that is supported by our SEP3-ChIP experiments.

Transcriptional Networks in Flower Development

Developmental transcription networks are composed of positive and negative regulatory loops, which often unite into larger transcription units. Positive feedback loops that are made up of two transcription factors regulating each other result in a robust expression in response to a transient developmental signal in order to establish and maintain a developmental program [62]. However, negative regulation is also very important in developmental processes to rapidly switch from one program to another (e.g., phase transitions).

The expression of MADS-box genes is often regulated by transcription complexes composed of their own gene products. One of the earliest examples for this phenomenon was the finding that floral homeotic genes responsible for the formation of petals and stamens (the B class genes in the ABC model) are up-regulated by dimers consisting of the encoded proteins [63]. More examples of autoregulatory feedback loops in the MADS-box gene family were reported afterwards. Also, *SEP* genes and their protein products are part of these regulatory networks [6]. The ability of the SEP3 protein to form many different dimer combinations [8] and the interaction of SEP3 protein and other floral homeotic proteins in larger protein complexes [10,64] illustrate that SEP3 is a key component in the network. It acts as a hub by linking various developmental programs that occur in the inflorescence and floral meristems and at later stages during organ differentiation. Our ChIP-SEQ data demonstrate that SEP3 is indeed able to bind to the promoters of many floral homeotic genes, supporting the conclusion that SEP3 is a key regulator of flower development (Figure 10).

From Figure 10, it is apparent that most MADS-box genes are targeted by a combination of several other MADS-domain proteins. The finding that most of these MADS-domain proteins also physically interact with each other suggests that combinatorial interactions of floral homeotic MADS-domain proteins and SEP3 protein are required for multiple positive autoregulatory feedback and feedforward loops. These mutual interdependencies may have evolved in order to enable stable organ-specific expression patterns by attenuating stochastic fluctuations in expression levels that may be more frequent if regulation was mediated by single proteins instead of heteromeric protein complexes.

Previously, negative feedback loops were suggested between AP1, which is a protein interaction partner of SEP3, and some of the flowering-time MADS-box genes to prevent the expression of these vegetative factors in the flower [65]. Our data also suggest that SEP3 has a role in repression of genes that control flowering time (e.g., *SOC1* and *AGL24*). In line with this idea, binding sites of SEP3 and AP1 overlap at the *SOC1* promoter ([65] and our ChIP-SEQ data). Interestingly, the binding sites also partially overlap with those of positive regulators at the *SOC1* locus ([66] and our ChIP-SEQ

Tecan Ex. 2114, Page 16 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

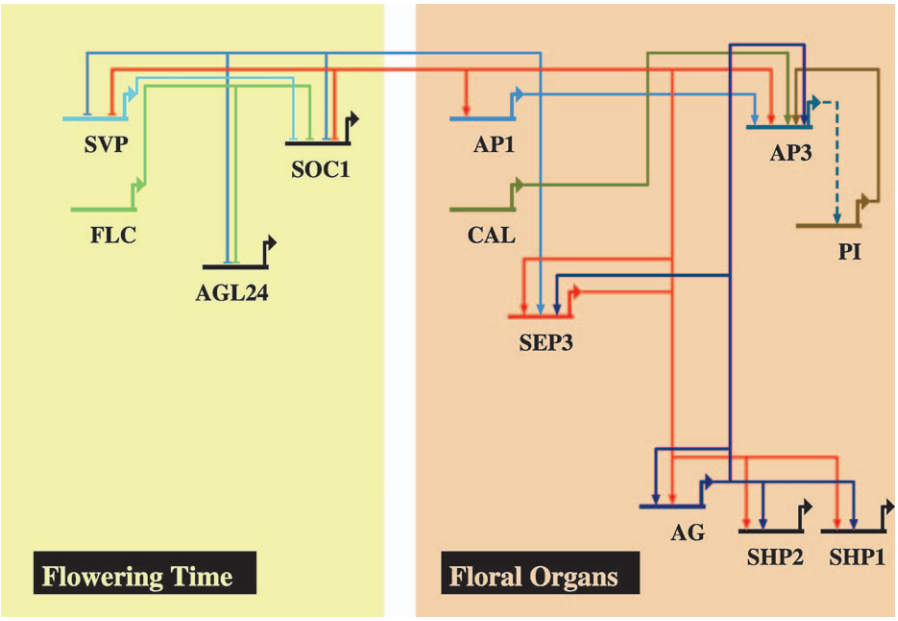


Figure 10. Autoregulatory Network of MADS Box Transcription Factors in *Arabidopsis* Flower Development

The network was visualized using the BioTapestry program [55]. According to our results, SEP3 is involved in the direct repression of flowering time genes, as well as in the activation of floral homeotic genes by binding to their respective promoters. The regulation of *PI* by AP3 is not confirmed by experimental approaches so far and might be indirect (dashed line).

doi:10.1371/journal.pbio.1000090.g010

data). It is possible that positively and negatively acting factors compete for the same binding sites. Since SEP3 is also able to interact with SOC1, it is also possible that a SEP3-SOC1 protein dimer is important for negative autoregulation of the SOC1 locus in floral meristems [8]. In general, direct negative (auto)regulatory feedback loops may enable an efficient and persistent switch between developmental phases, i.e., from inflorescence identity to floral identity.

Transcription factor complexes and their target genes are major components of transcriptional cascades that drive plant developmental processes. Once we have genome-wide datasets describing more of these interactions, we can integrate the data into transcriptional networks and models describing the interactions of the components of the network and predicting the outcome of modulation of the biological system [67,68]. Here, we have shown that ChIP-SEQ is a powerful tool providing these essential datasets to address questions in plant biology. Our findings suggest that many direct regulatory interactions exist in plant developmental networks. Additional information is needed, in particular the dynamics of the interactions in time and space. Furthermore, not only DNA–protein interactions, but also protein–protein interactions of transcription factors need to be resolved for a better understanding of developmental processes in complex organisms.

Materials and Methods

Plant material and growing conditions. *Arabidopsis thaliana*, wild-type (Col-0), *sep3*-1 (Col-0), and *ag*-1 (Ler) mutant plants were grown under standard greenhouse conditions (20 °C, long-day light regime: 16-h light, 8-h dark cycle). Flower material was harvested from primary and secondary inflorescences of 5–7-wk-old plants.

Chromatin immunoprecipitation (ChIP). ChIP experiments were performed essentially as described in [69]. For the IP, we used an antibody raised against a C-terminal peptide of SEP3. The antibody was tested in western analyses on plant extracts of wild-type and *sep3* mutant plants. Total extracts were produced using a standard protocol [70]; nuclei extracts were produced following the protocol that was used for the ChIP experiments, only that the flower material was not fixed. Cross-reaction with other SEPALLATA MADS-domain proteins was tested in western blots using proteins produced by in vitro translation.

Sample preparation for ChIP-SEQ and ChIP-CHIP, tiling array hybridization, Solexa sequencing. Linker annealing, amplification, and gel purification for the Solexa sequencing were essentially performed as instructed by the Illumina protocol with small modifications. The gel purification was done after the amplification step. We used individual, complete ChIP samples for each amplification reaction for the wild-type and *ag*-1 mutant samples. For the *sep3*-1 mutant sample (negative control), we pooled the DNA of three different ChIP experiments to obtain sufficient material for amplification. The amplified material was subjected to Solexa sequencing following Illumina's instructions. The sequence datasets were submitted to Gene Expression Omnibus (GEO) (accession number GSE14600).

For the amplification of the ChIP-DNA for ChIP-CHIP, we used a protocol published by [71] with modifications. The amplified DNA was partially digested with DNAse I (fragments <150 bp) and labeled with biotin. For the control experiment, we used unamplified, sheared chromatin from the same biological sample as the ChIP-DNA ("input DNA"). The DNA was fragmented and labeled in the same way as the ChIP-DNA. The labeled samples were hybridized to GeneChip *Arabidopsis* Tiling 1.0R Arrays (Affymetrix) as biological duplicates. Tiling array data were submitted to GEO (accession number GSE14635).

Processing of the tiling array data. Enrichments in ChIP sample hybridizations relative to input were calculated from raw intensity (CEL) files using a nonparametric statistical method implemented in the Affymetrix Tiling Analysis Software (TAS) [72]. Biological replicates were combined in the analysis. Once the significance (p-value) was obtained, we define as ChIP-CHIP peaks the genomic regions with a p-value lower than 0.05, and not separated by more than 100 bp with the highest p-value not higher than 0.05.

Primary statistical analysis of the ChIP-SEQ data. The 35- or 36-nucleotide (nt) reads were mapped to the unmasked *Arabidopsis* reference genome (ATH1.1con.01222004; ftp://ftp.arabidopsis.org/) using the SOAP software [73], allowing a maximum of two mismatches and no gaps. Iteratively, one base was discarded from

Tecan Ex. 2114, Page 17 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

the end of the nonmapped reads until the reads were uniquely mapped or fell below a minimum read length of 30 nt. Only uniquely mapped reads were retained. In order to recover the average length of the original DNA fragments that were subjected to Solexa sequencing, the reads were extended directionally to 300 nt.

The data analysis was aimed at a comparison of the enrichment found in the wild-type plant and, independently, in the *ag-1* mutant, against the enrichment found in the *sep3*-1 mutant treated as a negative control. For this, data concerning positions of the mapped reads were transformed into numbers characterizing all nucleotide positions in the genome in the following way. Define x_{is}, where i is the nucleotide position and $s = 1,2$ for the examined and control sample respectively, as the minimum of the counts of extended mapped reads that overlap at the position i on the forward and on the reverse strand in the sample s. This value is a conservative estimate of the representation of the nucleotide i in the sequenced samples, supported by both the strands independently. By the transformation

$$y_{is} = \frac{x_{is} \times 300 \times 10^6}{\sum_i x_{is}},$$

observations y_{is} independent of the number of sequenced reads were obtained. The examined sample values y_{i1} were then normalized with respect to the mean and variance of the distribution of control values y_{i2}. For the comparison of the observed examined and control counts at position i, the one-sided test based on the Poisson distribution was made according to the probability (test statistic) formula

$$t_i = 1 - \sum_{k=0}^{y_{i1}-1} \frac{e^{-z_{i2}} z_{i2}^k}{k!},$$

where z_{i2} is the maximum of y_{i2} and the global coverage obtained for the control whose value is the product of the number and the length of the extended mapped reads divided by the mappable genome length. All genomic regions consisting of nucleotides characterized by calculated probability values smaller than 0.05 and not interrupted by a gap of 100 nt or more were identified and assumed to contain a candidate enrichment peak. For presentation purposes, the calculated probabilities were transformed into scores of $-\log_e(t_i)$, and the maximum score value for each candidate peak was used to test the significance.

Permutation tests were used to estimate the FDR for the peaks. For this, each mapped read was considered as having the label "sample" when it belonged to the examined sample and "control" if it belonged to the control sample. To obtain the distribution of the test statistic under the null hypothesis of no differences between the examined and the control samples, the labels of the reads were randomly permuted, and for each permutation, the methodology explained above to test differences in distribution was applied. The permutations were run until at least 65,000 test statistic values for calculation of the null distribution were obtained.

Positional characterization of ChIP-SEQ peaks. All peaks were characterized by their location with respect to the annotated genes (as described in the TAIR7_GFF3_genes.gff file, ftp://ftp.arabidopsis.org/). The gene affected by binding at each peak was selected by the following algorithm. First, for all peaks, the affected gene was selected as the one with the peak inside it and the minimum distance to the start. For the peaks outside of the genes, the affected gene was then selected as the one with the peak in its 3,000-bp upstream or 1,000-bp downstream region and the minimum distance to the start or end, respectively. Thus, it was assumed that a DNA-binding event affects the closest neighboring gene, which results in a conservative estimate of the number of genes controlled by each of the transcription factors.

The genes affected by binding inside them were then characterized with respect to the precise annotation of the position of the affecting peak, by considering three categories: 5′ UTR, 3′ UTR, other exon regions (equivalent to CDS for protein coding genes), or "not annotated." This characterization was done on the basis of the first available splicing variant for a gene; in practice, this meant using variant 1 for most of the genes and variant 2 for some.

De novo motif discovery. Default parameters for MatrixREDUCE [30] and MEME [36] algorithms were used in order to identify motifs de novo. Sequences 500 bp upstream and downstream of the maximum score position for significant peaks (FDR <0.001) were obtained from the *Arabidopsis* reference genome (ATH1.1-con.01222005; ftp://ftp.arabidopsis.org/). Repeat and low-complexity regions were eliminated using RepeatMasker (A. F. A. Smit, R. Hubley, and P. Green, RepeatMasker at http://repeatmasker.org). Sequence affinity logo representations were prepared using AffinityLogo [30]. The MatrixReduce algorithm uses genome-wide occupancy/affinity data for a transcription factor and associated nucleotide sequences to discover the sequence-specific binding affinity of the transcription factor. It utilizes a statistical-mechanical model to describe the relationship between the nucleotide sequences and the occupancy/affinity-related score, therefore avoiding the need of selecting any background sequence model. In contrast, MEME only uses the occupancy/affinity-related score to define a group of nucleotide sequences from which a general binding site consensus will be obtained.

Search for known transcription factor binding sites. Perfect match motif consensus sequences were located in the 1,000-bp region around the maximum score position (defined as "peak area") for significant peaks (FDR <0.001) using a perl script and the *Arabidopsis* reference genome. The motif consensus sequences were obtained from AGRIS and Transfac databases [37,74,75]. Each consensus sequence was associated with the score value of the corresponding ChIP-SEQ peak in order to calculate enrichment.

DNA binding site characterization in genomic regions bound by SEP3. For each significant peak, the nucleotide sequence 500 bp around the position of the maximum peak score location were extracted and associated with the peak score value. To obtain the proportion of peaks with a given DNA binding site consensus at a given peak score threshold level, among the nucleotide sequences associated with a peak score value bigger than the threshold level, the proportion of sequences with at least one DNA binding site consensus was calculated for the sample and control set. Two control sets were generated: (1) one control set was generated randomly permutating the nucleotide sequences and the peak score value, and therefore destroying any relationship between them, and (2) another control set was generated permutating the nucleotides within their sequence for each nucleotide sequence independently. Each set was characterized by the number and location (regarding the center of the sequence) of the DNA consensus binding sites studied. We considered binding sites to be overrepresented in the ChIP-SEQ data only when they were enriched relative to both controls. For simplicity, only control 2 is shown in Figures 1 and 3.

The proportion of the distance of the DNA binding site consensus to the peak score location was calculated as the distance from the center position of the DNA consensus to the peak score location for each peak with a score bigger than the corresponding threshold at a given FDR level. Nonoverlapping distance ranges of 50 bp were considered to calculate the proportion of distance values within each range among the total number of distance values considered. All graphs were generated with R software.

Identification of regulatory modules. ExPlain [35] was used to identify transcription factor binding sites in the sequence probes generated by the ChIP-SEQ experiment, and to generate promoter models of regulation modules composed of more than one binding site. All CArG box position weight matrices from Transfac [37] were collected into a set, and thresholds were set to minimize false negatives. The algorithm MATCH [37] was used to identify putative sites matching the CArG box matrices in the profile. Site frequencies were compared to a control sequence set consisting of randomly sampled *Arabidopsis* promoters, assuming a binomial distribution. The algorithm Composite Module Analyst (CMA, [76]) was used to identify transcription modules overrepresented in the query dataset, based on the binding site prediction by MATCH. CMA parameters were set to find modules composed of one or more pairs of sites separated by 10, 11,..., 200 nt. CMA evaluates a fitness score, which includes functions measuring normality, *t*-test, site orientation and distance, sequence match score, and model complexity.

Analysis of the distance distribution between CArG box site pairs. In order to establish whether CArG box pairs presented conserved distances between them, a comparison between the distance distribution between sites found in the real data and those found in two control datasets was performed. The control sets generated consisted of (1) real promoter sequences from genes chosen at random from the *Arabidopsis* genome, and (2) random permutations of the sequences described above. The control or background set using real promoters (1) makes a more stringent test, because real binding sites, including CArG boxes, can still be found in the sequence. This test shows the specific selection of a given distance between CArG box site pairs in the SEP3-dependent gene's promoters, when compared against other real promoter sequences. The second background set, made out of randomly permuted sequences (2), presents sites generated by the random variation of the genome composition, without (biological) selection, so that this

Tecan Ex. 2114, Page 18 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

comparison shows the selection of distances between these sites, when compared with an absolutely random site and distance distribution. The MATCH algorithm was applied to these three sequence sets, using the CArG box profile described above, and the positions of the matching sites collected. The distance between adjacent sites was calculated and the distributions were compared using a binomial test, to identify overrepresented distances.

Gene ontology analysis. Analysis for GO term enrichment was carried out using the AMIGO server [77] for the top 1,000 genes of the ChIP-SEQ dataset (Database version 2008-05-15). Only GO terms with more than 20 annotated loci were taken into account. We considered the most specific categories, ones that were found to be enriched by AMIGO, for further analyses. To study in more detail the enrichment for the top-five most significant, most specific GO terms (GO:0008610 "lipid biosynthetic process," GO:0045449 "regulation of transcription," GO:0009908 "flower development," GO:0048513 "organ development," and GO:0009733 "response to auxin stimulus") and the nonenriched term GO:0051179 "localization," each gene was associated with the maximum score value among the peaks that were affecting it. We mapped each gene to a GO term if the gene belongs to one of these GO terms or its children, using the Arabidopsis Information Resource (TAIR) ATH-GO-GOSLIM 2008-05-10.

Gene expression microarray analysis. We downloaded $\log_2$ gcRMA normalized expression values for experiments of interest from the AtGenExpress developmental atlas [39]. Control probe sets and probe sets matching no or several loci in the *Arabidopsis* genome were ignored in the analysis. After back-transforming the $\log_2$ expression values to original scale, a two-sided Student test statistic was applied to test differential expression between sample and control. These genes with a p-value lower than 0.01 were considered as differentially expressed.

Generation of the SEP3-EAR construct and transgenic plants. To generate the dominant repressor, we used the SRDX domain [45], which is a modified version of the EAR domain of the SUPERMAN protein. The promoter and coding region including introns of SEP3 was amplified from genomic. We added the SRDX domain in two subsequent round of PCR reactions. The primer sequences are available on request. The PCR product recombined into pDONR207, and subsequently into the destination vector pFP101-35SGa. The destination vector was obtained digesting the plasmid pFP101 with BamHI and HindIII, filled in with Klenow enzyme and blunt-end ligated with Gateway cassetteB. The construct was transformed into the *sep3*-1 mutant plants and Col-O wild-type plants. SEP3-EAR expression in wild type and *sep3* mutant showed similar phenotypes; however. we focused our phenotypic analysis on *pSEP3:SEP3-EAR* in the *sep3* mutant.

SEP3-GR induction and quantitative gene expression analysis. Transgenic seedlings were germinated and grown on plates without DEX for 10 d and transferred to medium containing 10 mM DEX for 8 h or 1 d, respectively. Alternatively, they were directly germinated on DEX medium and grown for 10 d. cDNA synthesis was produced using the iScript cDNA synthesis kit, and qPCRs were performed using the SYBR green I–based system from BioRad following the manufacturer's instructions. Two technical and two independent biological replicates were analyzed using the MyIQ program. The data were normalized with two reference genes (*TUB* and *EF*). All primers are available as supplementary information (Table S6).

Coexpression analysis. The expression values for development stages (floral transition to flowers stage 12; AtGenExpress experiments 6, 8, 29, 31, 32, 33, and 39) were obtained as explained in "Gene expression microarray analysis" above. After back-transforming the $\log_2$ expression values to original scale, the Spearman rank correlation test was applied to check the coexpression for each gene on the array with SEP3. For each peak score threshold, the proportion of genes that correlate with SEP3 expression at different control levels (0.05, 0.01, 0.005, and 0.001) among the genes affected by at least one ChIP-SEQ peak was calculated. In a similar way, the proportion of genes that correlate with SEP3 expression among all the genes was calculated.

Supporting Information

Figure S1. Western Blot Analysis

(A) Western blot of in vitro–translated MADS-domain proteins probed with the SEP3 antibody used for the ChIP experiments (upper panel). The same proteins were detected using Streptavidin-Alkaline Phosphatase to test for expression of the biotin-labeled proteins (loading control; lower panel).
(B) Western blot on plant extracts. Left: comparison of SEP3 antibody signal in total and nuclei extracts. Right: detection of SEP3 in wild-type and sep3 mutant lines. Riken line PST1193 has an insertion in the upstream region of SEP3 near the transcriptional start. Riken line PST20678 has an insertion in an exon.

Found at doi:10.1371/journal.pbio.1000090.sg001 (1.57 MB PDF).

Figure S2. Comparison ChIP-SEQ and ChIP-CHIP Data

(A) Histogram of distances between the maximum score positions of significant peaks in ChIP-SEQ and ChIP-CHIP: 90.7% of the peaks in ChIP-CHIP have a ChIP-SEQ peak nearby (distance <1,000 bp).
(B) Distance between the top 4,000 ChIP-SEQ peaks and the top 4,000 ChIP-CHIP peaks as a function of the sum of their ranks. Peaks with low rank sums (high significance in ChIP-SEQ and ChIP-CHIP) are mostly a short distance from each other. Peaks with the higher rank sums are also mostly overlapping or only a short distance apart; however, more variation is observed.
(C) Box-plot of the widths of ChIP-SEQ peaks for different ranks (groups of 1,000). The lower the rank, the more significant is a peak.
(D) Box-plot of the widths of ChIP-CHIP peaks for different ranks. In (C) and (D), only the lowest 10,000 ranks are shown.
(E) Scatter plot of the average number of overlapping extended reads at each nucleotide position based on the raw sequence data for the two sequencing replicates of the SEP3 wild-type ChIP sample. The average was obtained over nonoverlapping windows of length 5,000 bp (Pearson correlation, 0.97).
(F) Scatter plot of rank positions of ChIP-SEQ and ChIP-CHIP peaks. For each ChIP-SEQ peak, the closest ChIP-CHIP peak was identified (at a distance not greater than 500 bp). The scatter plot shows the relationship between the peak rank positions in ChIP-CHIP and ChIP-SEQ. Only the 14,000 most significant peaks are shown. The peaks that are in a similar position in CHIP-SEQ and in ChIP-CHIP (not farther apart than 500 bp) have a similar rank position in both experiments. This relationship is stronger for the more significant peaks. In order to minimize the effects of technical differences related to the different experimental platforms, we did not consider the value of the test statistics for each peak, but rather its rank (the most significant peak has a rank of one).
(G) Scatter plot of the average number of overlapping extended reads at each nucleotide position based on the raw sequence data for the two biological replicates of the SEP3 wild-type ChIP sample. The average was obtained over nonoverlapping windows of length 5,000 bp (Pearson correlation, 0.81).
(H) Distance between the top 4,000 ChIP-SEQ peaks of two biological replicates as a function of the sum of their ranks. Peaks with low rank sums (high significance in ChIP-SEQ and ChIP-CHIP) are mostly a short distance from each other. Peaks with higher rank sums are also mostly overlapping or a short distance apart.

Found at doi:10.1371/journal.pbio.1000090.sg002 (914 KB PDF).

Figure S3. Peak Position within Genomic Features as a Function of SEP3 ChIP-SEQ Peak Score

(A) Enrichment of peaks within promoters (black line, up to 3 kb upstream) and downstream regions (green, up to 1 kb downstream) with increasing ChIP-SEQ score. Peaks within genes (red line) and peaks without any neighboring gene (blue line) are also shown in the graph.
(B) Preferred position of peaks within introns (blue line), 5′ UTRs (black line) and 3′ UTRs (green line). Peaks in coding sequences (red line) are also indicated in the graph.
(C) Position of the ChIP-SEQ peak maxima relative to the transcriptional start of genes (zero position). Enrichment of peaks (as indicated by proportion) depends on the FDR used.
(D) As in (C) but now for the end of genes (zero position is end of the gene).

Found at doi:10.1371/journal.pbio.1000090.sg003 (539 KB PDF).

Figure S4. Enrichment of MYC and C2H2 (ID1) Binding Sites with Increasing Peak Score Threshold

The frequency of ID1 binding sites in the genome and in the ChIP-SEQ data is very low, the enrichment however is supported by two types of controls. Black line: ChIP-SEQ data (wild-type); red line: Permutation of score and sequence; blue line: Randomization of sequences.

Found at doi:10.1371/journal.pbio.1000090.sg004 (537 KB PDF).

Figure S5. Expression Analysis of Potential Direct SEP3 Targets Compared to Genome-Wide Data

(A) Gene expression analysis of different floral homeotic mutants

Tecan Ex. 2114, Page 19 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

(meristematic tissue up to floral stage 7 for *lfy*-12, *ap1*-15, *ap2*-6, *ap3*-6, and *ag*-12) or different developmental stages (meristem before floral transition up to flowers of stage 12).
(B) Enrichment of genes correlated in expression with SEP3. The Spearman rank correlation *p*-value was calculated as described in Material and Methods.

Found at doi:10.1371/journal.pbio.1000090.sg005 (582 KB PDF).

Figure S6. Developmental Time Series Expression Data of Auxin-Related Genes Targeted by SEP3

Expression in vegetative above-ground organs and in flowers and reproductive meristems is shown. The expression of *PIN1*, which is itself not a target of SEP3, is also visualized.

Found at doi:10.1371/journal.pbio.1000090.sg006 (607 KB PDF).

Table S1. Overview of Results from Solexa Sequencing and Mapping of the Reads

(A) Results of Solexa sequencing and mapping of the reads to the genome.
(B) Reads mapped to chromosomes (in percent) in relation to chromosome length.

Found at doi:10.1371/journal.pbio.1000090.st001 (46 KB DOC).

Table S2. List of Genes Targeted by SEP3 in Wild Type and *agamous* Mutant Based on the ChIP-SEQ Experiments at FDR <0.001

The file also contains a table with data on differential expression of potential target genes in the *ag* mutant compared to wild type.

Found at doi:10.1371/journal.pbio.1000090.st002 (921 KB XLS).

Table S3. Genomic Positions of ChIP-SEQ Peaks in Different Categories of Genes

Genomic positions of ChIP-SEQ peaks relative to closest neighboring genomic loci in wild type (A) and *ag* mutant (B), and within genes in wild type (C) and *ag* mutant (D).

Found at doi:10.1371/journal.pbio.1000090.st003 (80 KB DOC).

Table S4. Frequency of Half-Sites and Complete Sequences Corresponding to the Consensus $CC[A/T]_6GG$ and $C[A/T]_7GG$

Overrepresentation was calculated given the observed number of these sequences in wild-type and *agamous* ChIP-SEQ peaks compared to their genome-wide frequencies using a binomial test.

Found at doi:10.1371/journal.pbio.1000090.st004 (41 KB XLS).

Table S5. Dependencies between Nucleotide Positions in CArG Boxes of Type $CC[A/T]_6GG$

The *p*-values of the chi-square test were used to measure dependence. The table shows the dependence in $-\log_e$ scale. Only values bigger than $-\log_e(0.05)$ are shown.

Found at doi:10.1371/journal.pbio.1000090.st005 (17 KB XLS).

Table S6. List of Primers Used in the Real-Time RT-PCR Experiments

Found at doi:10.1371/journal.pbio.1000090.st006 (61 KB XLS).

Acknowledgments

Some parts of the computations were made at the Poznań Supercomputing and Networking Center. We thank Stefan de Folter for initial advice on the ChIP experiments and for the SEP3 antibody, as well as Wolfgang Busch for advice on the TAS software. We also thank Richard Immink and Ruud de Maagd for comments on the manuscript. Mark Aarts and Remko Offringa are acknowledged for providing pictures of *ettin* and *pinoid*, respectively. We would like to thank Renze Heidstra for providing the DR5::YFP plant line. We would like to thank Laurent Farinelli and Magne Osteras (FASTERIS) for support with the ILLUMINA sequencing experiments.

Author contributions. KK and GCA conceived and designed the experiments. KK and CS performed the experiments. KK, JMM, RJ, CS, PK, and GCA analyzed the data. JMM, CAA, and PK contributed reagents/materials/analysis tools. KK and GCA wrote the paper.

Funding. KK, JMM, RJ, and CAA were supported by fellowships of the EU-Marie Curie Research Training program (Transistor-MRTN-CT-2004–512285). This project was cofinanced by the Centre for BioSystems Genomics (CBSG) and Netherlands Proteomics Centre (NPC), which are part of the Netherlands Genomics Initiative. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests. The authors have declared that no competing interests exist.

References

1. Coen E (2001) Goethe and the ABC model of flower development. C R Acad Sci III 324: 523–530.
2. Chuck G, Hake S (2005) Regulation of developmental transitions. Curr Opin Plant Biol 8: 67–70.
3. Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, et al. (2000) A short history of MADS-box genes in plants. Plant Mol Biol 42: 115–149.
4. Ferrario S, Immink RG, Angenent GC (2004) Conservation and diversity in flower land. Curr Opin Plant Biol 7: 84–91.
5. International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. Nature 431: 931–945.
6. Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature 405: 200–203.
7. Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, et al. (2003) MADS-box protein complexes control carpel and ovule development in Arabidopsis. Plant Cell 15: 2603–2611.
8. de Folter S, Immink RG, Kieffer M, Parenicova L, Henz SR, et al. (2005) Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. Plant Cell 17: 1424–1433.
9. Mandel MA, Yanofsky MF (1999) The Arabidopsis AGL 9 MADS box gene is expressed in young flower primordia. Sex Plant Reprod 11: 22–28.
10. Honma T, Goto K (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 409: 525–529.
11. Sridhar VV, Surendrarao A, Liu Z (2006) APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. Development 133: 3159–3166.
12. de Folter S, Angenent GC (2006) trans meets cis in MADS science. Trends Plant Sci 11: 224–231.
13. Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of in vivo protein-DNA interactions. Science 316: 1497–1502.
14. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, et al. (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods 4: 651–657.
15. Li XY, MacArthur S, Bourgon R, Nix D, Pollard DA, et al. (2008) Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. PLoS Biol 6: e27. doi:10.1371/journal.pbio.0060027
16. Borneman AR, Gianoulis TA, Zhang ZD, Yu H, Rozowsky J, et al. (2007) Divergence of transcription factor binding sites across related yeast species. Science 317: 815–819.
17. Thibaud-Nissen F, Wu H, Richmond T, Redman JC, Johnson C, et al. (2006) Development of Arabidopsis whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. Plant J 47: 152–162.
18. Bowman JL, Smyth DR, Meyerowitz EM (1989) Genes directing flower development in Arabidopsis. Plant Cell 1: 37–52.
19. Cheng Y, Dai X, Zhao Y (2006) Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. Genes Dev 20: 1790–1799.
20. Benjamins R, Quint A, Weijers D, Hooykaas P, Offringa R (2001) The PINOID protein kinase regulates organ development in Arabidopsis by enhancing polar auxin transport. Development 128: 4057–4067.
21. Youfa Cheng YZ (2007) A role for auxin in flower development. J Integr Plant Biol 49: 99–104.
22. Robinson MD, Smyth GK (2007) Moderated statistical tests for assessing differences in tag abundance. Bioinformatics 23: 2881–2887.
23. Rose AB, Elfersi T, Parra G, Korf I (2008) Promoter-proximal introns in Arabidopsis thaliana are enriched in dispersed signals that elevate gene expression. Plant Cell 20: 543–551.
24. Deyholos MK, Sieburth LE (2000) Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron. Plant Cell 12: 1799–1810.
25. Sieburth LE, Meyerowitz EM (1997) Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. Plant Cell 9: 355–365.
26. Gomez-Mena C, de Folter S, Costa MM, Angenent GC, Sablowski R (2005) Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. Development 132: 429–438.
27. Ito T, Wellmer F, Yu H, Das P, Ito N, et al. (2004) The homeotic protein AGAMOUS controls microsporogenesis by regulation of SPOROCYTELESS. Nature 430: 356–360.
28. West AG, Causier BE, Davies B, Sharrocks AD (1998) DNA binding and

Tecan Ex. 2114, Page 20 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

dimerisation determinants of Antirrhinum majus MADS-box transcription factors. Nucleic Acids Res 26: 5277–5287.
29. Tang W, Perry SE (2003) Binding site selection for the plant MADS domain protein AGL15: an in vitro and in vivo study. J Biol Chem 278: 28154–28159.
30. Foat BC, Morozov AV, Bussemaker HJ (2006) Statistical mechanical modeling of genome-wide transcription factor occupancy data by MatrixREDUCE. Bioinformatics 22: e141–149.
31. Pellegrini L, Tan S, Richmond TJ (1995) Structure of serum response factor core bound to DNA. Nature 376: 490–498.
32. Huang H, Tudor M, Su T, Zhang Y, Hu Y, et al. (1996) DNA binding properties of two Arabidopsis MADS domain proteins: binding consensus and dimer formation. Plant Cell 8: 81–94.
33. Ciannamea S, Kaufmann K, Frau M, Tonaco IA, Petersen K, et al. (2006) Protein interactions of MADS box transcription factors involved in flowering in Lolium perenne. J Exp Bot 57: 3419–3431.
34. Theissen G, Saedler H (2001) Plant biology. Floral quartets. Nature 409: 469–471.
35. Kel A, Voss N, Jauregui R, Kel-Margoulis O, Wingender E (2006) Beyond microarrays: finding key transcription factors controlling signal transduction pathways. BMC Bioinformatics 7 Suppl 2: S13.
36. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34: W369–373.
37. Wingender E, Dietze P, Karas H, Knuppel R (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res 24: 238–241.
38. Kosugi S, Ohashi Y (2002) DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J 30: 337–348.
39. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, et al. (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501–506.
40. Tilly JJ, Allen DW, Jack T (1998) The CArG boxes in the promoter of the Arabidopsis floral organ identity gene APETALA3 mediate diverse regulatory effects. Development 125: 1647–1657.
41. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25–29.
42. Goda H, Sasaki E, Akiyama K, Maruyama-Nakashita A, Nakabayashi K, et al. (2008) The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. Plant J 55: 526–542.
43. Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF (2004) The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Curr Biol 14: 1935–1940.
44. Martienssen R, Irish V (1999) Copying out our ABCs: the role of gene redundancy in interpreting genetic hierarchies. Trends Genet 15: 435–437.
45. Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. Plant J 34: 733–739.
46. Okada K, Ueda J, Komaki MK, Bell CJ, Shimura Y (1991) Requirement of the auxin polar transport system in early stages of Arabidopsis floral bud formation. Plant Cell 3: 677–684.
47. Bennett SRM, Alvarez J, Bossinger G, Smyth DR (1995) Morphogenesis in pinoid mutants of Arabidopsis thaliana. Plant J 8: 505–520.
48. Christensen SK, Dagenais N, Chory J, Weigel D (2000) Regulation of auxin response by the protein kinase PINOID. Cell 100: 469–478.
49. Nemhauser JL, Feldman LJ, Zambryski PC (2000) Auxin and ETTIN in Arabidopsis gynoecium morphogenesis. Development 127: 3877–3888.
50. Sessions RA, Zambryski PC (1995) Arabidopsis gynoecium structure in the wild and in ettin mutants. Development 121: 1519–1532.
51. Pekker I, Alvarez JP, Eshed Y (2005) Auxin response factors mediate Arabidopsis organ asymmetry via modulation of KANADI activity. Plant Cell 17: 2899–2910.
52. Hardtke CS (2007) Transcriptional auxin-brassinosteroid crosstalk: who's talking? Bioessays 29: 1115–1123.
53. Tanay A (2006) Extensive low-affinity transcriptional interactions in the yeast genome. Genome Res 16: 962–972.
54. Li C, Potuschak T, Colon-Carmona A, Gutierrez RA, Doerner P (2005) Arabidopsis TCP20 links regulation of growth and cell division control pathways. Proc Natl Acad Sci U S A 102: 12978–12983.
55. Cubas P, Lauter N, Doebley J, Coen E (1999) The TCP domain: a motif found in proteins regulating plant growth and development. Plant J 18: 215–222.
56. Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, et al. (2003) Control of leaf morphogenesis by microRNAs. Nature 425: 257–263.
57. Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol Biol 49: 373–385.
58. Guyomarc'h S, Bertrand C, Delarue M, Zhou DX (2005) Regulation of meristem activity by chromatin remodelling. *Trends Plant Sci* 10: 332–338.
59. Workman JL (2006) Nucleosome displacement in transcription. Genes Dev 20: 2009–2017.
60. Kempin SA, Savidge B, Yanofsky MF (1995) Molecular basis of the cauliflower phenotype in Arabidopsis. Science 267: 522–525.
61. Reinhardt D, Mandel T, Kuhlemeier C (2000) Auxin regulates the initiation and radial position of plant lateral organs. Plant Cell 12: 507–518.
62. Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8: 450–461.
63. Schwarz-Sommer Z, Hue I, Huijser P, Flor PJ, Hansen R, et al. (1992) Characterization of the Antirrhinum floral homeotic MADS-box gene deficiens: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. Embo J 11: 251–263.
64. Egea-Cortines M, Saedler H, Sommer H (1999) Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. EMBO J 18: 5370–5379.
65. Liu C, Zhou J, Bracha-Drori K, Yalovsky S, Ito T, et al. (2007) Specification of Arabidopsis floral meristem identity by repression of flowering time genes. Development 134: 1901–1910.
66. Liu C, Chen H, Er HL, Soo HM, Kumar PP, et al. (2008) Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. Development 135: 1481–1491.
67. Belostotsky DA, Rose AB (2005) Plant gene expression in the age of systems biology: integrating transcriptional and post-transcriptional events. Trends Plant Sci 10: 347–353.
68. Alvarez-Buylla ER, Benitez M, Davila EB, Chaos A, Espinosa-Soto C, et al. (2007) Gene regulatory network models for plant development. Curr Opin Plant Biol 10: 83–91.
69. de Folter S, Urbanus SL, van Zuijlen LG, Kaufmann K, Angenent GC (2007) Tagging of MADS domain proteins for chromatin immunoprecipitation. BMC Plant Biol 7: 47.
70. Lamb RS, Irish VF (2003) Functional divergence within the APETALA3/PISTILLATA floral homeotic gene lineages. Proc Natl Acad Sci U S A 100: 6558–6563.
71. Lee TI, Johnstone SE, Young RA (2006) Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat Protoc 1: 729–748.
72. Ghosh S, Hirsch HA, Sekinger E, Struhl K, Gingeras TR (2006) Rank-statistics based enrichment-site prediction algorithm developed for chromatin immunoprecipitation on chip experiments. BMC Bioinformatics 7: 434.
73. Li R, Li Y, Kristiansen K, Wang J (2008) SOAP: short oligonucleotide alignment program. Bioinformatics 24: 713–714.
74. Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, et al. (2003) AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. BMC Bioinformatics 4: 25.
75. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, et al. (2006) TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 34: D108–110.
76. Waleev T, Shtokalo D, Konovalova T, Voss N, Cheremushkin E, et al. (2006) Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm. Nucleic Acids Res 34: W541–545.
77. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, et al. (2009) AmiGO: online access to ontology and annotation data. Bioinformatics 25: 288–289.
78. Staswick PE, Serban B, Rowe M, Tiryaki I, Maldonado MT, et al. (2005) Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. Plant Cell 17: 616–627.
79. Sun TP, Kamiya Y (1994) The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. Plant Cell 6: 1509–1518.
80. Li J, Chory J (1997) A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. Cell 90: 929–938.
81. Stenzel I, Hause B, Miersch O, Kurz T, Maucher H, et al. (2003) Jasmonate biosynthesis and the allene oxide cyclase family of Arabidopsis thaliana. Plant Mol Biol 51: 895–911.
82. Friml J, Benkova E, Blilou I, Wisniewska J, Hamann T, et al. (2002) AtPIN4 mediates sink-driven auxin gradients and root patterning in Arabidopsis. Cell 108: 661–673.
83. Ulmasov T, Hagen G, Guilfoyle TJ (1999) Dimerization and DNA binding of auxin response factors. Plant J 19: 309–319.
84. Wu MF, Tian Q, Reed JW (2006) Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. Development 133: 4211–4218.
85. Conner TW, Goekjian VH, LaFayette PR, Key JL (1990) Structure and expression of two auxin-inducible genes from Arabidopsis. Plant Mol Biol 15: 623–632.
86. Cardon GH, Hohmann S, Nettesheim K, Saedler H, Huijser P (1997) Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a novel gene involved in the floral transition. Plant J 12: 367–377.
87. Lee J, Park JJ, Kim SL, Yim J, An G (2007) Mutations in the rice liguleless gene result in a complete loss of the auricle, ligule, and laminar joint. Plant Mol Biol 65: 487–499.
88. Kim JH, Choi D, Kende H (2003) The AtGRF family of putative

Tecan Ex. 2114, Page 21 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028

transcription factors is involved in leaf and cotyledon growth in Arabidopsis. Plant J 36: 94–104.
89. Teale WD, Paponov IA, Palme K (2006) Auxin in action: signalling, transport and the control of plant growth and development. Nat Rev Mol Cell Biol 7: 847–859.
90. Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80: 847–857.
91. Cubas P, Lauter N, Doebley J, Coen E (1999) The TCP domain: a motif found in proteins regulating plant growth and development. Plant J 18: 215–222.
92. Jiang C, Fu X (2007) GA action: turning on de-DELLA repressing signaling. Curr Opin Plant Biol 10: 461–465.
93. Manfield IW, Devlin PF, Jen CH, Westhead DR, Gilmartin PM (2007) Conservation, convergence, and divergence of light-responsive, circadian-regulated, and tissue-specific expression patterns during evolution of the Arabidopsis GATA gene family. Plant Physiol 143: 941–958.
94. Li L, Deng XW (2005) It runs in the family: regulation of brassinosteroid signaling by the BZR1-BES1 class of transcription factors. Trends Plant Sci 10: 266–268.
95. Du L, Poovaiah BW (2004) A novel family of Ca2+/calmodulin-binding proteins involved in transcriptional regulation: interaction with fsh/Ring3 class transcription activators. Plant Mol Biol 54: 549–569.
96. Brewer PB, Howles PA, Dorian K, Griffith ME, Ishida T, et al. (2004) PETAL LOSS, a trihelix transcription factor gene, regulates perianth architecture in the Arabidopsis flower. Development 131: 4035–4045.
97. Li X, Qin G, Chen Z, Gu H, Qu LJ (2008) A gain-of-function mutation of transcriptional factor PTL results in curly leaves, dwarfism and male sterility by affecting auxin homeostasis. Plant Mol Biol 66: 315–327.

Tecan Ex. 2114, Page 22 of 22
QIAGEN Sciences, LLC v. Tecan Genomics - IPR2025-00028