Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species

Julia L. Mateos, Vicky Tilmes, Pedro Madrigal, Edouard Severing, René Richter, Colin W. M. Rijkenberg, Pawel Krajewski, and George Coupland

*Department of Plant Developmental Biology, Max Planck Institute for Plant Breeding Research, D-50829 Cologne, Germany; and †Department of Biometry and Bioinformatics, Institute of Plant Genetics, Polish Academy of Sciences, 60-479 Poznań, Poland

Contributed by George Coupland, November 6, 2017 (sent for review November 7, 2016; reviewed by Kerstin Kaufmann and Hao Yu)

Genome-wide landscapes of transcription factor (TF) binding sites (BSs) diverge during evolution, conferring species-specific transcriptional patterns. The rate of divergence varies in different metazoan lineages but has not been widely studied in plants. We identified the BSs and assessed the effects on transcription of FLOWERING LOCUS C (FLC) and PERPETUAL FLOWERING 1 (PEP1), two orthologous MADS-box TFs that repress flowering and confer vernalization requirement in the Brassicaceae species Arabidopsis thaliana and Arabis alpina, respectively. We found that only 14% of their BSs were conserved in both species and that these contained a CARG-box that is recognized by MADS-box TFs. The CARG-box consensus at conserved BSs was extended compared with the core motif. By contrast, species-specific BSs usually lacked the CARG-box in the other species. Flowering-time genes were highly overrepresented among conserved targets, and their CARG-boxes were widely conserved among Brassicaceae species. Cold-regulated (COR) genes were also overrepresented among targets, but the cognate BSs and the identity of the regulated genes were usually different in each species. In cold, COR gene transcript levels were increased in flc and pep1-1 mutants compared with WT, and this correlated with reduced growth in pep1-1. Therefore, FLC orthologs regulate a set of conserved target genes mainly involved in reproductive development and were later independently recruited to modulate stress responses in different Brassicaceae lineages. Analysis of TF BSs in these lineages thus distinguishes widely conserved targets representing the core function of the TF from those that were recruited later in evolution.

Significance

Variation in gene transcription is a major source of phenotypic diversity contributing to adaptation and speciation (1). Understanding how transcriptional patterns arise and evolve is an important question in biology. Transcription is regulated by transcription factors (TFs) that bind, often in combinations, to specific DNA sequences within genes to modulate the activity of RNA polymerase. Variation in transcription of a gene can arise through cis-acting differences that alter TF binding or through trans-acting differences in TF activity. Recently, these issues have been addressed at the genome-wide level by utilizing ChIP followed by next-generation sequencing (ChIP-seq), allowing determination of binding sites (BSs) for individual TFs across the whole genome and comparison of the repertoire of BSs for orthologous TFs in different species. This approach has been widely applied in yeast and metazoans to study evolution of TF BSs (2–4). However, diversification of TF BSs has seldom been studied at the genome-wide level in plant lineages (5). We determined the BSs of the Arabidopsis thaliana TF FLOWERING LOCUS C (FLC) and its ortholog PERPETUAL FLOWERING 1 (PEP1) of Arabis alpina, which are critical repressors of flowering (6–8). Variation in TF binding between species has been extensively studied in yeast and metazoans. Surprisingly, in vertebrates, the BSs of TFs that contribute to developmental processes showed low conservation even among closely related species, and the extent of conservation decreased exponentially with increasing evolutionary distance (9). However, in Drosophila species, TF BSs appear to be more conserved (3, 10), and the extent of conservation decreased at only a linear rate with evolutionary distance (9). In plants, only one comparative study of TF binding has been performed in two sister species (5). BSs of the MADS-box TF SEPALLATA3 (SEP3) were compared between Arabidopsis thaliana and Arabidopsis lyrata, and approximately 21% of BSs were conserved. This rate of divergence resembles that described in vertebrates rather than in insects. However, the extreme variation in genome size observed in plant lineages (11) and the differences in transposon content found between members of the same family suggested that constraints on genome-wide patterns of TF BSs might vary significantly in different parts of the phylogeny (9). Therefore, we focused on PEP1 and FLC, orthologous MADS-box TFs in Arabidopsis and Arabis, respectively. In both species, these TFs confer a flowering response to low winter vernalization | flowering | Arabidopsis | Arabis alpina | binding-site evolution

Author contributions: J.L.M., V.T., and G.C. designed research; J.L.M., V.T., R.R., and C.W.M.R. performed research; J.L.M., V.T., P.M., E.S., P.K., and G.C. analyzed data; and J.L.M., V.T., and G.C. wrote the paper.

Reviewers: K.K., Humboldt University Berlin; and H.Y., National University of Singapore. The authors declare no conflict of interest.

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE89889).

1) J.L.M. and V.T. contributed equally to this work.
2) Present address: Fundación Instituto Leloir, Instituto de Investigaciones Bioquímicas de Buenos Aires-National Scientific and Technical Research Council (CONICET), C1405BWE Buenos Aires, Argentina.
3)Present address: Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SA, United Kingdom.
4)Author contributions: J.L.M., V.T., and G.C. designed research; J.L.M., V.T., R.R., and C.W.M.R. performed research; J.L.M., V.T., P.M., E.S., P.K., and G.C. analyzed data; and J.L.M., V.T., and G.C. wrote the paper.

Reviewers: K.K., Humboldt University Berlin; and H.Y., National University of Singapore. The authors declare no conflict of interest.

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE89889).

1)J.L.M. and V.T. contributed equally to this work.
2)Present address: Fundación Instituto Leloir, Instituto de Investigaciones Bioquímicas de Buenos Aires-National Scientific and Technical Research Council (CONICET), C1405BWE Buenos Aires, Argentina.
3)Present address: Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SA, United Kingdom.
4)Author whom correspondence should be addressed. Email: coupland@mpiz-mainz.mpg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1618075114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1618075114 PNAS | Published online December 4, 2017 | E11037–E11046
temperatures (i.e., vernalization), and their inactivation by muta-
tion causes early flowering (6–8). Transcript levels of FLC and
PEP1 decrease progressively in prolonged cold, so that, after
seven weeks, they reach trough levels. In the annual species
A. thaliana, FLC mRNA level remains low after vernalization,
allowing plants to flower continuously until senescence, and FLC
mRNA level is reset in the progeny. By contrast, in perennial
A. alpina, PEP1 transcript level is reset after vernalization, allowing
the same individual to respond again to cold the following year (8).
Allelic variation at PEP1 and FLC is also important in natural
populations contributing extensively to variation in flowering be-
havior (12, 13). Similar to other MADS-box TFs, FLC binds to
CArG-box motifs with the consensus sequence CC(A/T)6GG (14).
Hundreds of BSs for FLC have been described in the genome of
A. thaliana (15, 16), and these suggest that, in addition to
vernalization response, FLC is associated with vegetative
phase transition, flower development, gibberellin (GA) syn-
thesis and signaling, as well as other environmental responses
(15–20).

We compared the BSs and effects on gene expression of FLC
and PEP1 in A. thaliana and A. alpina, respectively, species with
markedly different genome sizes and life histories that are in
different lineages of the Brassicaceae (21, 22). Our study sug-
gests that determining evolutionarily conserved TF BSs in dif-
f erent Brassicaceae lineages can be a powerful approach to
identify target genes contributing to the core function of a TF, as
for FLC/PEP1 in the control of flowering. In addition, it allows
investigation of how the BSs of orthologous TFs have in-
dependently evolved to regulate other processes.

Results

Definition of the Genome-Wide Occupancy and Transcriptional
Network Controlled by PEP1 in A. alpina. To identify the BSs of
PEP1 in A. alpina Pajares, ChIP-seq was performed on long-day
(LD) grown seedlings of A. alpina Pajares and pep1-1 mutant,
which produces no WT PEP1 protein (SI Appendix, Fig. S1). The
assay was performed in three biological replicates for each ge-
notype by using antiserum raised against PEP1 (12), and high
reproducibility was found between biological replicates (SI Ap-
pendix, Fig. S2). This approach identified 204 high-confidence
peaks (false discovery rate < 0.01, present in at least two of
three replicates) at unique genomic regions that represent PEP1
BSs (Fig. 1A and Dataset S1). Among these BSs, 180 were
assigned to 324 annotated neighboring genes, where the center
of a peak resided within 5 kb upstream of the transcriptional
start site (TSS) and no further than 3 kb downstream of the
transcriptional end site (TES) of these genes, which were
therefore referred to as PEP1 direct target genes (Fig. 1A). In
A. thaliana, target genes are usually defined as those in which the
center of the ChIP-seq peak resides within 3 kb upstream of the
TSS and 1 kb downstream of the TES (23, 24), but these dis-
tances were extended for A. alpina because of the longer inter-
genic regions in this species (22). Nevertheless, the PEP1 targets
obtained by using the parameters employed in A. thaliana are
also listed (Dataset S1). MADS-box TFs bind mostly in the 5′
 promoter region or 5′UTR of target genes (15, 16, 23–27).
Consistent with these previous reports, 70% of the detected
PEP1 peaks were in promoter regions (Fig. 1B). Also, several
BSs were validated independently by ChIP-PCR (Fig. 1C).

DNA sequences of PEP1 BSs were examined to identify po-
tisional weight matrices that can be assigned as consensus se-
quences recognized by the TF. De novo motif discovery
identified the canonical CArG-box (CC(A/T)6GG), the recog-
nition sequence of MADS-box TFs, in 79% of the BSs, and most
of these were located close to the center of the peak (Fig. 1D).
In addition, TCP-binding motifs and G-boxes, which are bound
by bHLH and bZIP TFs, were significantly enriched, although the
positions of these elements within the peak were more variable
(Fig. 1D). All three motifs were previously identified in the
binding regions of FLC in A. thaliana (15, 16).

![Image](https://www.pnas.org/cgi/doi/10.1073/pnas.1618075114)

Fig. 1. Characterization of PEP1 targets in A. alpina. (A) Number of significant ChIP-seq peak calls and their associated genes for PEP1 binding across each chromosome. The proportion of peaks associated with genes is shown in black, and others are shown in gray. (B) PEP1 peak distribution over different genomic features in the A. alpina genome. (C) Validation of ChIP-seq for selected PEP1 BSs using ChIP-PCR. For each target, fold enrichment relative to its input is shown. Minus signs indicate primers not flanking predicted BSs used as negative controls. Data are shown as mean ± SEM (n = 3). Asterisks indicate significant enrichment in WT compared with pep1-1 (n.s., not significant; *P < 0.05, **P < 0.01, and ***P < 0.001, Student’s t test). (D) Density plot of distance of CArG-box, G-box, or TCP-binding motif consenus sequences to the center of PEP1 peaks. (Inset) Logo of the motifs found by MEME motif analysis. CArG-box, 161 sites (E-value, 7e–106); G-box, 84 sites (E-value, 4e–3); TCP-binding motif, 128 sites (E-value, e–4). (E) Proportion of PEP1 direct targets among genes that are up- or down-regulated in pep1-1 in leaves or apices. For RNA-seq experiment, plants were grown for 2 wk in LD conditions, and leaves and apices were collected at ZT8.
To analyze the influence of PEP1 on gene expression, RNA sequencing (RNA-seq) was performed on *A. alpina* Pajares and the *pep1-1* mutant. In these experiments, leaves and apices of 2-wk-old seedlings were used. In leaves and apices, 96 and 325 transcripts, respectively, were differentially expressed [log (fold change) > 1.5 and adjusted P < 0.05] in *pep1-1* compared with Pajares (Dataset S2). The gene set regulated by PEP1 differed between the two tissues (SI Appendix, Fig. S3), but approximately 54% of genes differentially expressed in *pep1-1* leaves were also differentially expressed in apices. Overall, more genes were differentially expressed in apices of *pep1-1* mutants than in leaves, which contrasts with what was previously observed in comparisons of FLC and *flc-3 A. thaliana* plants (16). PEP1 therefore appears to have a broader role in *A. alpina* apices than FLC does in *A. thaliana* apices, in which the related protein SHORT VEGETATIVE PHASE (SVP) plays a greater role (16).

To assess the effect of PEP1 on transcription of its direct target genes, differentially expressed genes (DEGs) in *pep1-1* mutants were compared with those identified by PEP1 ChIP-seq. A total of 27 genes representing ~8% of PEP1 direct targets were detected as DEG in either tissue analyzed (Fig. 1E).

In addition, PEP1 acted almost exclusively as a repressor of transcription (Fig. 1E), because nearly all differentially expressed direct target genes were increased in expression in the *pep1-1* mutant compared with Pajares. A similar result was previously obtained for FLC in *A. thaliana* (15, 16). PEP1 direct target genes that are differentially expressed in the *pep1-1* mutant include the orthologs of genes that are associated with different aspects of flowering in *A. thaliana*, including FLOWERING LOCUS T (FT), SUPPRESSOR OF OVEREXPRESSSION OF CO 1 (SOC1) (28, 29), CONSTANS LIKE1 (COL1) (30), SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 15 (SPL15) (31, 32), SEPALATA 3 (SEP3) (33), and PHYTOCHROME INTERACTING FACTOR 3 (PHF3) (34). PEP1 therefore directly controls genes that are likely to be involved in several flowering-related processes in *A. alpina*.

A Small Proportion of PEP1 BSs in *A. alpina* Are Orthologous to *A. thaliana* FLC BSs. PEP1 BSs in *A. alpina* and FLC BSs in *A. thaliana* were compared to determine their conservation and the extent to which they regulate similar biological processes. To provide an FLC dataset directly comparable to that described here earlier for PEP1, *A. thaliana* plants were grown under similar conditions to *A. alpina* and ChIP-seq was performed. The sequences obtained after ChIP-seq of both TFs generated similar genome coverage (Dataset S3). The ChIP-seq of FLC was performed by using an antiserum that was previously described (15) (SI Appendix, Fig. S4, and Dataset S1). Again, high reproducibility was found between biological replicates (SI Appendix, Fig. S2). A total of 297 FLC BSs were identified. The number of FLC BSs in *A. thaliana* appears therefore to be higher than the number of PEP1 BSs in *A. alpina*. Genes were defined as direct targets of FLC if the center of a ChIP-seq peak was within 3 kb upstream of the TSS or 1 kb downstream of the TES. By using these criteria, 487 direct target genes were identified under the LD conditions used here, and these showed 50% overlap with those previously found in plants grown under short-day (SD) conditions (16) (SI Appendix, Fig. S4B). Some of the differences may represent photoperiod-dependent variation in activity of FLC. Furthermore, 61% of the targets detected here overlap with those identified by Deng et al. (15) (SI Appendix, Fig. S4B).

As expected, FLC BSs were mostly located in promoter regions (SI Appendix, Fig. S4C). The conservation of FLC BSs and PEP1 BSs was examined by direct sequence alignment comparisons using BLAST (Fig. 2A). Only 28 PEP1 BSs (representing 14% of all PEP1 BSs) were found to align to FLC BSs and are referred to as conserved BSs (Fig. 2A and B, blue lines). Thus, the binding landscapes of PEP1 and FLC are very different in *A. alpina* and *A. thaliana*, respectively. Target gene conservation was approached independently of BS conservation by comparing directly the target genes identified in each species (Dataset S1). A total of 39 genes were identified as bound by FLC and PEP1, hereafter referred to as common target genes (Fig. 2A and B and Dataset S1). In addition, PEP1 targets were compared with FLC targets that were detected by Deng et al. (15) but not found in this study. Twenty genes were identified (Dataset S1), which represented 4% of the FLC targets unique to the list of Deng et al., further supporting the low proportion of target genes conserved in both species. Genes associated with the 28 conserved BSs were then compared with the 39 common target genes detected in our data. Among the common target genes, 28 contained conserved BSs and were referred to as conserved target genes, whereas the remaining 11 contained BSs that differed between the two species (visualized in SI Appendix, Fig. S5).

Whether PEP1 and FLC BSs are located in conserved syntenic regions was then examined. Each BS in one species was aligned

![Diagram](https://example.com/fig2.png)

**Fig. 2.** Comparison of direct targets of FLC and PEP1 as well as of the DEGs in the respective mutants. (A) Flow diagram illustrating the rationale of the method to assess BS conservation and target gene conservation for FLC and PEP1 in *A. thaliana* and *A. alpina*, respectively. (B) Summary of the analysis described in A shown as a Circos plot. Heat map showing the percentage of identity of sequence alignments of BSs to the orthologous locus in the other genome (percentage identity: <50, gray; 50–60%, very light blue; 60–70%, blue; >70%, dark blue). Connecting lines represent alignments between FLC and PEP1 common BSs. Selected flowering related genes are labeled. (C and D) Venn diagram comparing DEG in leaves and apices of the *A. thaliana flc-3* mutant and the *A. alpina pep1-1* mutant. FLC transcriptional data were obtained from ref. 16. (E) Proportion of DEG in *pep1-1* among common target genes and all *A. alpina* PEP1 target genes. P value indicates a statistically significant difference (hypergeometric test).
to the orthologous locus in the other genome. This approach demonstrated that all conserved BSs that are associated with genes (24 BSs in *A. alpina* and 27 BSs in *A. thaliana*) are located in orthologous genomic positions. Furthermore, the majority of species-specific BSs (72% of PEP1 and 88% of FLC BSs; Fig. 2 *A* and *B*) are present in the other genome, with a sequence identity of 50–80% (*SI Appendix*, Fig. S6, and Dataset S5). Therefore, the sequences containing species-specific BSs are mostly present in the genome of the other species but are not recognized by the orthologous TF.

The effects of PEP1 and FLC on the transcriptome of *A. alpina* and *A. thaliana*, respectively, also differed greatly (Fig. 2 *C* and *D*), consistent with the low conservation of BSs. Among PEP1 direct targets, a higher proportion of conserved target genes showed altered expression in the mutants compared with all target genes (Fig. 2 *E*; validation of differences for a subset of genes is shown in *SI Appendix*, Fig. S7). Thus, under standard growth conditions, conserved targets are more likely to show a change in expression than species-specific targets.

In summary, these results indicate that PEP1 and FLC targets include a small core set of conserved genes, but that, overall, their genome-wide binding landscapes are very different.

**Sequence Variation at CArG-Box Motifs Contributes to Divergence in PEP1 and FLC Binding.** FLC and PEP1 proteins show high sequence identity, which reaches 100% in the DNA-binding MADS-domain (*SI Appendix*, Fig. S8), yet their genome-wide BSs are highly divergent (Fig. 2). Comparison of the overall sequence conservation of *A. alpina* BSs in *A. thaliana* demonstrated that BSs that are conserved across species showed similar levels of sequence identity to those that are specific to *A. alpina* (*SI Appendix*, Fig. S6). This result differs from that described for SEP3 in *A. thaliana* and *A. lyrata*, in which genomic regions that were bound by SEP3 in both species were significantly more conserved than regions that were bound specifically in either species (Fig. 2). SEP3 showed altered expression in the mutants compared with similarly conserved BSs that are conserved across specific species. This motif was significantly enriched in species-specific BSs as well as conserved BSs, and the motifs showed high similarity between species (Fig. 3A), suggesting that both TFs recognize the same cis-elements. However, the consensus sequences of CArG motifs in conserved and species-specific BSs differed slightly (Fig. 3A): only the long core motif of conserved BSs showed significant (Z-score >3) enrichment of a TTT extension at the 5′ end of the core motif (Fig. 3A). This suggests that the 5′ TTT is a functional part of the CArG-box at the conserved BSs, similar to the 3′ AAA extension that was found in all BSs and that was previously found in FLC BSs (15). Interestingly, ChiP-seq peaks present at conserved BSs are also generally more significant than those present at specific BSs (Mann–Whitney *U* test for PEP1, *P* = 0.00611; Mann–Whitney *U* test for FLC, *P* = 0.0352; Fig. 3A), suggesting that in vivo FLC/PEP1 might bind more strongly to those sites containing both the 5′ TTT and 3′ AAA sequences.

These differences between conserved and species-specific CArG-box motif sequences are found in both species, and therefore do not explain the divergence in the BSs of PEP1 and FLC between *A. alpina* and *A. thaliana*, respectively. To analyze these sequences in more detail, position-specific PhastCons scores, which represent the probability that a given nucleotide is part of a conserved region, were calculated for the CArG-box motifs and flanking nucleotide positions. Analysis of position-specific PhastCons scores within CArG-box motifs and 10 flanking nucleotides on each side in the conserved and specific BSs revealed that the CArG motif in the conserved BSs tend to have higher PhastCons scores (Fig. 3B). These results indicate that CArG-boxes in the conserved BSs are located within more conserved regions than CArG-boxes of *A. alpina*-specific BSs. Furthermore, a strong association was found between BS conservation across species and the presence of a CArG-box in these sites in both species (Fig. 3B, *Inset*). Approximately 80% of conserved PEP1 BSs contain a CArG-box in the *A. thaliana* orthologous sequence that is bound by FLC. In contrast, only 45% of the *A. alpina* specific PEP1 BSs contain this motif in their *A. thaliana* counterparts (Fig. 3B, *Inset*). This dramatic reduction in CArG-box conservation at the *A. thaliana* orthologs of specific PEP1 BSs probably largely explains the lack of FLC binding at these sites. Absence of CArG-boxes in these regions of *A. thaliana* was generally not caused by large-scale rearrangements, because overall homology of the BSs sequence between species is mostly maintained (*SI Appendix*, Fig. S6). Examples of modifications that change the CArG-box motif are shown in Fig. 3C. In each of these cases, the CArG-box is absent in *Aethionema arboicum* and in the *Arabidopsis* species examined, suggesting that they were gained in a lineage leading to the *Arabidopsis* genus. Similarly, for *A. thaliana*-specific BSs, CArG-boxes were not present in most sequence counterparts in the *A. alpina* genome (Fig. 3B, *Inset*) and regions surrounding the cis-elements are more conserved in common BSs than in specific ones (Fig. 3B). In contrast to CArG-boxes, the conservation of G-box elements and TCP-BBs at PEP1 BSs did not strongly differ between conserved and specific BSs and thus did not correlate with binding conservation (*SI Appendix*, Fig. S9). In summary, conservation of binding at orthologous sequences across species is strongly correlated with presence of the CArG-box motif in both species (Fig. 3A and B), whereas species-specific PEP1 binding correlated with absence of the CArG-box in *A. thaliana* but not with large scale rearrangements.

Although the absence of a CArG-box motif is strongly correlated with loss of binding, ~40% of specific BSs retained a CArG-box in the species in which no binding was detected (Fig. 3B, *Inset*). Lack of binding at these sites was also not correlated with absence of other motifs, such as the G-box or the TCP-binding motif (*SI Appendix*, Fig. S9). CArG-box motifs are recognized by many MADS-box TFs in addition to FLC, some of which are shown to interact with FLC and COP1 (23, 35), and the genome-wide BSs of several of these have been identified in *A. thaliana* (15, 16, 23–26, 36–38). Conservation of CArG-boxes at species-specific BSs might be necessary to allow binding of other MADS-box TFs to these sites in the species in which FLC/PEP1 does not bind. Binding of different MADS-box TFs could allow repurposing of a cis-element, as was described in human and mouse (39). Comparison of the BSs detected for FLC/PEP1 with those previously described for other MADS-box TFs in *A. thaliana* showed that ~80% of FLC-specific BSs are bound by at least one other MADS-box TF, whereas, for 40% of PEP1-specific BSs, the orthologous *A. thaliana* sequence has been found to bind another MADS-box TF (Fig. 3D). *A. thaliana*-specific BSs containing a conserved CArG-box are slightly more frequently bound by other MADS-box TFs compared with all species-specific BSs [Fig. 3D, row 1 (red) vs. row 2 (black)], suggesting that these CArG-boxes might be selected for because other TFs recognize these regions. Interestingly, BSs that are conserved between *A. thaliana* and *A. alpina* show a higher frequency of overlap with BSs of other MADS-box TFs than the species-specific BSs (Fig. 3D, row 3). The conserved function of PEP1 and FLC might involve interaction with some of these other MADS-box TFs, or these regions might be recognized by multiple MADS-box complexes.

In summary, sequence analysis of BSs showed that PEP1/FLC binding conservation is associated with the presence of a CArG-box at the BSs in both species, but CArG-box conservation is not sufficient for binding to occur. The binding of other MADS-box TFs to the same BS as PEP1/FLC might explain why, in some cases, the CArG-box is maintained even though conserved binding is not observed.

**Target Genes of PEP1 and FLC Show Related Biological Functions.** Gene Ontology (GO) enrichment analysis of direct target genes was performed to address whether differences in the binding landscapes of PEP1 and FLC cause divergence in the
biological processes they regulate. In contrast to the high divergence in the identity of FLC and PEP1 direct target genes (Fig. 2), the biological processes associated with the targets of each TF were highly similar (Fig. 4A, Left). Consistent with the early-flowering phenotypes of flo and pep1 mutants, the most enriched gene set regulated directly by FLC and PEP1 contained genes involved in flowering-time control with representation factors of 8.6 and 5.0 for PEP1 and FLC, respectively (Fig. 4A, Left, and Dataset S1). In both species, direct targets include genes that control flowering in distinct pathways and tissues of the plant, particularly in the photoperiod/light sensing/circadian clock pathway, in GA metabolism or signaling, and in meristem response and development (Dataset S1). Among genes bound by PEP1 whose orthologs are associated with the photoperiodic pathway were CONSTANS LIKE 5 (COL5), FT, NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT B-2 (NYF C-9), SPAIN-RELATED 2 (SPAN2), CIRCADIAN 1 (CIR1), AGAMOUS LIKE-15 (AGL15), and the photoreceptor-binding gene PHOTOTRANSITION B PHOTOB1 (PHOT1).

Fig. 3. PEP1 binding conservation depends on cis-element conservation. (A, Left) CArG-boxes in conserved and species-specific BSs in A. alpina and A. thaliana (consensus sequences are boxed). The number of CArG-boxes identified in each subset is indicated. Z-scores indicate significance of TTT enrichment at positions 1–3. (A, Right) Binding significance varies between groups (test based on hypergeometric distribution). Binding strength is higher among conserved BSs compared with specific BSs. (C) Three examples of A. alpina-specific BSs that contain a CArG-box in A. alpina that is lost in other Brassicaceae. (Left) G-browse capture of PEP1 binding and CArG-box sequence. Changes relative to the consensus motif in A. thaliana are highlighted in red. (Right) Alignment of A. alpina sequence around CArG-box with orthologous sequences in several Brassicaceae species. CArG-box from A. alpina is highlighted in red. (D) Percentages of different subsets of FLC or PEP1 BSs that overlap with BSs described for other MADS-box TFs in A. thaliana. For comparison of A. alpina PEP1 BSs, the orthologous sequence of BSs in A. thaliana was used. Subsets of BSs for each species are specific (red), specific with conserved CArG-box (black), and conserved BSs (black).
Analysis of the genes that were directly bound by PEP1 and differentially expressed in pep1-1 mutants again identified categories related to flowering (Fig. 4A, Right). Thus, conserved binding of PEP1 to flowering-time genes is often associated with repression of their transcription. On the contrary, GO categories related to temperature responses were not enriched among those differentially expressed in pep1-1 (Fig. 4A), indicating that direct binding to temperature related genes is not related with controlling transcription under normal growth temperatures. However, as is shown later, some of these genes were differentially expressed in pep1-1 compared with WT at lower temperatures.

The sequences of three conserved targets with key regulatory roles in flowering or flower development were analyzed (Fig. 4B). In these cases, the predicted CArG-box BSS and surrounding sequences were highly conserved between A. alpina and A. thaliana (Fig. 4B). Moreover, the predicted BSS were highly conserved in other Brassicaceae species, and two were conserved in A. arabicum, the most basal lineage within the Brassicaceae family, and Tarenaya hassleriana, a member of the sister family (Fig. 4C and SI Appendix, Fig. S10). Among the whole set of conserved target genes, the CArG-boxes are widely conserved in the Brassicaceae, but only a small subset is conserved in all species (SI Appendix, Fig. S11). These findings suggest that the regulation of these genes by PEP1/FLC is broadly conserved in the Brassicaceae, although this remains to be directly tested.

In summary, to control flowering, FLC and PEP1 regulate transcription of several orthologous genes through binding to highly conserved CArG motifs, whereas other biological processes regulated by both TFs mainly involve their binding to different target genes, suggesting that they arose by convergent evolution.

**PEP1 and FLC Modulate Responses to Short-Term Cold Exposure by Regulating Different Sets of Genes.** Several direct target genes of FLC and PEP1 were found to be associated with the functional categories “response to cold” or “response to temperature” (Figs. 1C, 4A, and 5A). Moreover, PEP1 bound to orthologs of 33 of the robust list of 1,279 A. thaliana cold-regulated (COR) genes (47), and orthologs of 62 COR genes were differentially regulated in pep1-1 (17% of pep1-1 differentially regulated genes; Fig. 5B and Dataset S6). Comparison of the expression levels of these 62 genes in the pep1-1 mutant with published expression levels of their A. thaliana orthologs after 1 h of cold treatment (47) revealed highly correlated patterns (SI Appendix, Fig. S12). These results suggest that mutation of PEP1 in A. alpina has a similar effect on the transcription of these genes as cold treatment in A. thaliana, and therefore PEP1 might negatively regulate the cold response. Similarly, FLC target genes include those belonging to the response to cold or response to temperature GO categories and 44 COR genes (Fig. 5B and Dataset S6). Furthermore, 70 COR genes are differentially regulated in flc-3 (26% of flc-3 differentially regulated genes) compared with FLC (Dataset S6), and expression patterns of these genes in WT background after 1 h of cold treatment resemble those in the flc-3 mutant (SI Appendix, Fig. S12). Together, these findings suggest that FLC and PEP1 are negative regulators of the cold response. However, comparison of COR genes regulated by FLC and those regulated by PEP1 revealed only a small overlap (Fig. 5B), indicating that most of the cold-network genes controlled by PEP1 and FLC are different.

PEP1 was found to bind to many COR genes (Figs. 4A and 5A and B), but analysis of direct targets that are also differentially regulated in pep1-1 under ambient temperature did not identify cold-related terms (Fig. 4A). Therefore, to test whether PEP1 directly regulates the cold response, the expression of these target genes was analyzed at low temperature in WT and pep1-1 plants. Even though PEP1 expression was slightly increased in response to short-term cold treatment (SI Appendix, Fig. S13), similar to what was described for FLC in A. thaliana (48), many PEP1 direct target genes were expressed at higher levels in the pep1-1 mutant compared with WT (Fig. 5C), suggesting that PEP1 functions in cold temperature to modulate the cold response. Expression of COR genes was also analyzed.
is eventually silenced, and the Arabidopsis thaliana CBF1 was detectable in the first weeks of vernalization (SI Appendix, Fig. S14), correlating with PEP1 expression (8). The expression patterns of orthologs of the COR genes that were direct targets of PEP1 were examined in response to short-term cold treatment in A. thaliana (Fig. 5C). The flc-3 mutation had a similar effect on the amplitude of induction of these cold-responsive genes as pep1-1 did in A. alpina (Fig. 5C). However, only three of the five tested were direct targets of FLC, and, of those, only COL1 is bound at a conserved BS, but this BS does not contain a conserved CARG-box identified in the analysis of conserved target genes (SI Appendix, Figs. S5 and S11, and Dataset S1). This further supports our suggestion that PEP1 and FLC negatively regulate the cold response but do so by directly binding to different genes or BSs.

During cold exposure, plant growth is reduced (49), and, in A. thaliana, CBF1 overexpression caused severe growth reduction (50), suggesting that active repression of growth is part of the response to cold. Stronger expression of cold-responsive genes, including CBF1 (Fig. 5C), in pep1-1 prompted us to test whether the pep1-1 mutation also affects plant growth in response to cold. Compared with WT plants, growth of the pep1-1 mutant was significantly reduced during cold treatment (Fig. 5D). By contrast, before or after cold exposure or under control conditions in which plants were not exposed to cold, this effect was not detected or was even reversed. A stronger growth retardation in response to cold in pep1-1 suggests that, by modulating expression of cold-responsive genes, PEP1 might promote growth in cold, a function not previously assigned to this gene to our awareness.

Taken together, PEP1 and FLC negatively regulate cold induction of cold-responsive genes, but, in contrast to the regulation of reproductive development (Fig. 5E), the COR genes controlled by FLC and PEP1 are highly divergent.

Discussion
Phenotypic differences between closely related species are often caused by changes in gene transcription (51, 52). Variation in the activity of regulatory TFs or in the sequence of their BSs can cause alterations in the transcription of target genes (53). The rate with which genome-wide patterns of TF binding change during evolution seems to vary in different lineages (9), but has not been extensively studied in plants. Here we show that the BSs of the orthologous MADS-box TFs FLC and PEP1 vary greatly between A. thaliana and A. alpina, two species in the Brassicaceae family, but a core set of targets with roles in reproductive development is highly conserved. We argue that the functions of PEP1 and FLC in reproductive development are highly conserved but that they have independently acquired functions in abiotic stress responses.

Conservation and Diversification of TF BSs. The genome-wide landscapes of BSs for orthologous TFs differ markedly among related species of yeast and mammals. For example, analysis of BSs of the liver-specific TF CEBPα in five different vertebrates and six rodent species showed an exponential decrease in the proportion of BSs conserved with evolutionary distance (4, 54, 55). Analysis of genome-wide BSs of several TFs in different tissues of human and mouse also showed low conservation (39). However, genome-wide BSs for TFs among Drosophila species seem to diverge more slowly than in mammals. For instance, more than 60% of the BSs of the TF Twist, which is involved in embryonic development, were conserved among Drosophila species that diverged more than 30 Mya (3). Similar results were obtained for six TFs that regulate segmentation (10). This apparent lower rate of TF BS divergence among Drosophila species might result from their more compact genome size and higher effective population size than those of mammals (9, 56).
In plants, in which genome size and structure vary greatly even among closely related species (57, 58), little is known of the rate of divergence of genome-wide patterns of TF binding. The one example available so far showed that, between the sister plant species A. lyrata and A. thaliana, only ~20% of BSs of the SEP3 TF were conserved, suggesting a fast rate of divergence more similar to that observed in mammals. However, analysis of more TFs in a wider range of species is required to provide a broader picture of the divergence of TF binding in plants. We analyzed the genome-wide binding patterns of the orthologous MADS-box TFs FLC and PEP1 of A. thaliana and A. alpina,
respectively, two species in different lineages of the Brassicaeae, and found that fewer than 20% were shared between species. Although the CARG-box motifs recognized by different MADS-box TFs can differ subtly in sequence (14, 59–61), this is unlikely to be the cause of the divergent binding patterns of FLC and PEP1 because de novo motif discovery of the ChiP-seq peaks indicated that these TFs recognize the same consensus sequences. Similar observations were made for orthologous metazoan TFs with divergent genome-wide binding patterns (3, 4, 10). Nevertheless, for FLC and PEP1, the conserved BSs appear to differ qualitatively from the species-specific BSs. De novo motif discovery performed specifically on their common BSs identified an extended palindromic CARG-box consensus. MADS-box TFs bind DNA as tetramers, and each dimer binds one CARG-box motif (62). Variation in the CARG-box sequence recognized in vivo by SEP3 was previously proposed to result from it acting in different homo- and heterodimers (23). Flowering-time genes were enriched among those targets containing common BSs for FLC and PEP1 and were overrepresented among genes bound by FLC and its partner MADS-box TF SVP (16, 35). Therefore, common targets might be predominantly recognized by the SVP/FLC complex, whereas species-specific BSs might be recognized more frequently by FLC acting together with other MADS-box partners. Defining such combinatorial activities of TFs can allow their BSs to be predicted more accurately by using computational approaches (63).

Much of the variation in TF binding patterns among species has been explained by species-specific sequence variation at DNA motifs recognized by the TFs (3, 4, 10) or by insertion of transposable elements (4, 5, 64). The divergent landscapes of PEP1 and FLC binding could partially be explained by different distributions of CARG-box motifs. The majority of A. thaliana genomic regions orthologous to A. alpina-specific BSs did not contain a consensus CARG-box. Similarly, in comparisons of SEP3 binding between A. thaliana and A. lyrata, species-specific binding was often associated with the absence of a CARG-box in the species in which no binding was detected (5). Also, in vertebrates, the liver TFs HNF1α, HNF4α, and HNF6 of mouse recognized human-specific BSs in mouse cells containing a human chromosome, emphasizing the importance of cis-acting variation in generating species-specific binding patterns (65). No evidence was obtained for the direct involvement of transposons in the generation of the specific BSs for PEP1 and FLC between A. alpina and A. thaliana, in contrast to what was observed for SEP3 in A. lyrata (5) and despite the high transposon content of A. alpina genome (22). Alternatively, lack of binding often correlated with the absence of a CARG-box, 45% of PEP1-specific BSs retained the CARG-box at the orthologous site in A. thaliana. Thus, at these sites, the presence of a CARG-box is not sufficient for FLC binding. In other systems, loss of TF binding despite the conservation of the binding motif has been explained in different ways. Variation in regions flanking the core motif can impair binding of a TF because it prevents combinatorial interactions with a second TF that binds to an adjacent site (16, 54, 55) or it affects the DNA structure (66) or chromatin accessibility (67). Furthermore, changes in methylation patterns could also account for differences in TF binding (68).

These observations emphasize the importance of testing binding directly and not relying on predictions based on conservation of the binding motif. Other features of these sites might preclude FLC binding, such as the absence of binding of another TF complex at a neighboring site that acts in a combinatorial fashion with FLC or a more general feature of the chromatin structure. However, de novo motif discovery performed on the species-specific sites of FLC/PEP1 binding in the species in which binding occurred did not identify any sequence motif apart from the CARG-box, suggesting that the additional features present at these BSs are not sufficiently conserved in different genes to be identified by this approach or that they are not recognizable in the primary DNA sequence.

The genome of A. alpina is threefold larger than that of A. thaliana, and larger genome sizes have been proposed to correlate with rapid evolutionary turnover of TF BSs, as described here earlier in the comparison of Drosophila species and mammals. Variation in intergenic distances in the genomes of different plant species might also contribute to divergence in the rate at which TF BSs exhibit turnover, although the difference in genome size between A. alpina and A. thaliana is much smaller than that between flies and vertebrates. Despite the larger genome of A. alpina, we found fewer PEP1 BSs than FLC BSs in A. thaliana. However, to determine if this correlates with a lower BS turnover in A. alpina, it will be necessary to consider rates at which BSs were gained and lost in both species. Nevertheless, the rapid rate of turnover of FLC/PEP1 BSs described here, and that of SEP3 BSs in the only other study on genome-wide turnover of plant TF BSs of which we are aware (5), suggest that, in plants, as in vertebrates, BSs of TFs involved in developmental processes evolve rapidly.

**Conservation of the Core Developmental Function of TFs.** In metazoans, despite rather low conservation of BSs of TFs involved in development, their regulatory functions are often maintained. This apparent paradox seems to be explained by high conservation of binding to key genes involved in the core developmental function of the TF (3, 4, 54, 69). Sometimes, even though binding to the same gene occurs in different species, the BSs in the orthologs are at different positions (4, 39, 64, 69). Similarly, FLC and PEP1 each bound to hundreds of genes, but only 39 (Dataset S1) were identified as common target genes. Thus, despite the similar early-flowering phenotypes of fcl and pep1 mutants, the cis-regulatory networks regulated by these orthologous TFs are strongly divergent. Nevertheless, as described here earlier for metazoan BSs, the conservation of core developmental function of flowering-time control were highly overrepresented among the common targets of FLC and PEP1. Early flowering is the most evident phenotype of these fcl and pep1 mutants, suggesting that the major evolutionarily conserved function of these TFs is to repress flowering. Similarly, many SEP3 target genes involved in flower development were conserved in two sister species (5). Thus, the conservation of core functions accompanied by rapid turnover of other BSs appears to be found widely in diverse organisms from vertebrates to plants.

Common targets in flowering-time control indicate that FLC and PEP1 repress the earliest stages of floral induction in the shoot meristem. SOC1 and SPL15 are both direct targets and were recently shown to cooperate to activate target genes in the meristem under noninductive conditions (31). SOTH is also an early-acting gene in photoperiodic response (29). Therefore, by repressing both of these genes, activity of the SPL15-SOC1 complex would be strongly reduced, and floral induction effectively repressed under different environmental conditions. Similarly, SEP3 is a member of many MADS-box complexes involved in reproductive development, and its promoter was bound by both PEP1 and FLC (70). The CARG-boxes recognized by PEP1 and FLC in these three genes are conserved in all Brassicaceae genomes tested except A. arabicum, in which only the CARG-box present in SOC1 was conserved. Thus, these genes were probably regulated by FLC orthologs early during the divergence of the Brassicaceae, before separation of the Arabis and Arabidopsis lineages, and these BSs must be under strong selective pressure; however, binding of FLC orthologs to these sites remains to be tested experimentally in a broader range of species. Previously, FT and FD, which encode photoperiodic flowering pathway components, were identified as targets of FLC by ChIP-PCR (71). Neither here nor in other genome-wide analyses of FLC targets was FD identified as a target of FLC/PEP1 (15, 16). Therefore, despite the additional features present at these BSs are not directly regulated by FLC. By contrast FT is bound by FLC and PEP1 in A. thaliana and A. alpina, respectively (15, 16). Finally, the list of highly conserved targets includes genes that have not previously been shown to have roles in flowering, but are now candidates for testing, including several encoding TFs, a transporter, and a helicase (SI Appendix, Fig. S11).
Nonconserved Targets of FLC and PEP1 Support Convergent Evolution on Stress Responses. Analysis of all BSs of FLC or PEP1 identified many of the same enriched GO categories, particularly hormone stimuli or the response to cold, suggesting that these TFs regulate abiotic responses as well as reproductive development. Similarly, the LEAFY TF, which has developmental roles in conferring floral meristem identity, was shown to contribute to biotic stress responses in Arabidopsis thaliana (72). Among yeast species, cold divergence in TF-binding patterns was proposed to facilitate adaptation to different environments (2). FLC and PEP1 both bind many COR genes, which are regulated by CBF TFs and are implicated in adaptation to cold. Also, overexpression of CBFs was previously shown to increase FLC transcription (48), suggesting a complex interaction between abiotic stress responses to cold and FLC activity. We detected increased transcription of COR genes in pep1-1 and flc-3 compared with WT when plants were grown in cold, suggesting that PEP1 and FLC function specifically in cold to repress COR gene induction. As cold stress tolerance comes along with a retardation of growth (49), PEP1 and FLC might modify this trade-off to maintain growth under cold but nonfreezing temperatures. Interestingly, most of the COR genes bound by FLC or PEP1 are specific for each TF or the COR target genes are orthologous but the BSs are different, as shown for CBF1. This dramatic difference in the set of COR genes regulated by FLC and PEP1 suggests that these TFs were independently recruited after divergence of the lineages containing Arabidopsis, alpina and A. thaliana to modulate responses to cold, and this might represent convergent evolution to adapt to colder climates after the Brassicaceae split from the Cleomaceae (21).

Perspectives. As observed in metazoans, analyzing TF BSs in plant species from different lineages of the Brassicaceae is informative in understanding TF function and evolution. Performing similar studies with more TFs will demonstrate how generally applicable these observations are, and analyzing species in other families will indicate whether conservation of BSs extends beyond the Brassicaceae. FLC orthologs and their regulation by vernalization have been described even in monocotyledonous species (73). Consistent with this, the CArG-boxes to which FLC and PEP1 bind in SOC1 and several other flowering-time genes were present in the orthologs of T. hassleriana, a member of the Cleomaceae, a sister family to the Brassicaceae. However, the Cleomaceae are found in semitropical regions and do not show a vernalization response, suggesting that this CArG-box is bound by other MADS-box TFs in T. hassleriana and has been repurposed in the Brassicaceae to bind FLC orthologs that contribute to vernalization response or that the FLC orthologs present in T. hassleriana (SI Appendix, Fig. S15) bind to this site but do not confer vernalization response. Deepening our knowledge of the function and topology of these networks in different species will help in the understanding of the evolution of reproductive development and of plant transcriptional networks more generally.

Methods

Growth Conditions, Plant Materials, and Phenotypic Analysis. A. alpina plants from Pajares accession and pep1-1 mutant (8) were grown under LD conditions (16 h light/8 h dark at 20 °C). The A. thaliana FRI introgression line (Col FRI) (74) was used as WT. The flc allele used was flc-3 (6). Plants were grown on soil under controlled conditions. For phenotypic analysis of growth under cold conditions, plants were grown for 2 wk under LD conditions at 21 °C and then transferred to SD conditions at 4 °C or 21 °C. Further details of expression analysis are provided in the SI Appendix.

Genome-wide Transcriptome Studies. RNA for RNA-seq experiments was obtained from three biological replicates, and total RNA was isolated by using an RNeasy Plant Mini Kit (Qiagen) and subsequently digested with RNase-free DNase (Ambion) according to the manufacturer’s protocol. A total of 4 μg of RNA was used for library preparation using TruSeq RNA Sample Preparation (Illumina). Libraries were gel-purified from 200 to 350 bp and pair-end sequencing (100 bases read) was performed at the Cologne Center for Genomics, University of Cologne.

ChIP Experiments. For ChIP, plants were grown in LD conditions for 2 wk and above-ground tissue was collected at Zeitgeber time (ZT) 8. Three independent biological replicates were performed for PEP1 ChIP-seq assays, and two were performed for FLC ChIP-seq. For ChIP on PEP1, the Pajares and pep1-1 genotypes were used with 1 μl of PEP1 antiserum raised against the C-terminal domain of PEP1 (12). For ChIP on FLC, the A. thaliana genotypes Col FRI and flc-3 FRI were analyzed and 2 μl of FLC antiserum was used (15). After cross-linking, the ChIP and ChIP-seq were performed as in ref. 16 for both species. Further details are given in the SI Appendix.

ChIP-Seq and RNA-Seq Data Analysis. After quality checking, the illumina sequence reads were mapped to the A. alpina Pajares reference genome sequence (22). The numbers of reads obtained for each replicate are listed in Dataset S3. ChIP-seq peak calling was performed for the replicates of Pajares against those of the flc-3 mutant negative control. Recommended guidelines were followed for the analysis of Chip-seq data for quality control, read mapping, normalization, peak-calling, and assessment of reproducibility among biological replicates (75). The tools “ranger” and “wig” were used in the software PeakRanger (76) to identify read-enriched genomic regions (P = 1 × 10−10, q-value = 0.01, remainder of parameters set to default settings) and to generate variable-step wiggle files of read coverage. Further details of these methods and analysis of the RNA-seq data are given in the SI Appendix. Other bioinformatics analyses were carried out following standard procedures (77).

Bioinformatic Analysis on cis-Elements and Identification of Ortholog Sequences. All bioinformatics approaches to identify enriched cis-elements, distribution of motifs across peaks, and identification of ortholog peak sequences are given in the SI Appendix.

ACKNOWLEDGMENTS. We thank Uciel Chorostecki from the Institute of Molecular and Cell Biology of Rosario, Argentina, for bioinformatics assistance and Wen-Biao Jiao and Korbinian Schneeberger for providing the Arabis montbretiana genome assembly and annotation. This study was supported by the Alexander von Humboldt-Stiftung and Argentinean National Council of Sciences (J.L.M.), Deutsche Forschungsgemeinschaft via the Cluster of Excellence on Plant Science Grant EXC 1028 (to V.T.), EU Marie Curie Innovative Training Network SYSPLO 237909 (to P.M. and P.K.), the ERC through HyLife (G.C.), and a core grant of the Max Planck Society (to G.C.). Some computations were performed at the Poznani Supercomputing and Networking Center.

1. Romero IG, Ruvinsky I, Gilad Y (2012) Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet 13:505–516.
2. Borneman AR, et al. (2007) Divergence of transcription factor binding sites across related yeast species. Science 317:815–819.
3. He Q, et al. (2011) High conservation of transcription factor binding and evidence for combinatorial regulation across six Drosophila species. Nat Genet 43:414–420.
4. Schmidt D, et al. (2010) Five-vertebrate Chip-seq reveals the evolutionary dynamics of transcription factor binding. Science 328:1036–1040.
5. Muñoz JM, et al. (2016) Evolution of DNA-binding sites of a floral master regulatory transcription factor. Mol Biol Evol 33:185–200.
6. Michaels SD, Amasino RM (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11:949–956.
7. Sheldon CC, et al. (1999) The FFL MADS box gene: A repressor of flowering in Arabidopsis by bipolarization and methylation. Plant Cell 11:445–458.
8. Wang R, et al. (2009) PEP1 regulates perennial flowering in Arabidopsis. Nature 459:423–427.
9. Villar D, Fliche P, Odorn DT (2014) Evolution of transcription factor binding in metazoans-Mechanisms and functional implications. Nat Rev Genet 15:221–233.
10. Bradley RK, et al. (2010) Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related Drosophila species. PLoS Biol 8:e1000343.
11. Michael TP (2014) Plant genome size variation: Bloating and purging DNA. Brief Funct Genomics 13:308–317.
12. Allard M-C, et al. (2012) PEP1 of Arabis alpina is encoded by two overlapping genes that contribute to natural genetic variation in perennial flowering. PLoS Genet 8:e1003130.
13. Michaels SD, He Y, Scortecci KC, Amasino RM (2003) Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc Natl Acad Sci USA 100:10102–10107.
14. de Folter S, Angenent GC (2006) Trans acts across in MADS science. Trends Plant Sci 11:224–231.
15. Deng W, et al. (2011) FLOWERING LOCUS C (FLC) regulates development pathways throughout the life cycle of Arabidopsis. Proc Natl Acad Sci USA 108:6680–6685.
16. Mateos JL, et al. (2015) Combinatorial activities of SHORT VEGETATIVE PHASE and FLOWERING LOCUS C define distinct modes of flowering regulation in Arabidopsis. Genome Biol 16:31.
of pathways directly regulated by SHORT Mateos et al.

PLOS Comput Biol

egetative and reproductive development in Arabi-

Nat

229:1

Dev Cell

138:2198

out of cross-species analysis.

www.pnas.org/cgi/doi/10.1073/pnas.1618075114

288:1613

328:

405:200

| Image

18. Edwards KD, et al. (2006) FLOWERING LOCUS C mediates natural variation in the

17. Chiang GC, Barua D, Kramer EM, Amaimon RM, Donohue K (2009) Major flowering

temperature regulatory network. Plant J 82:193–207.

39. Denas O, et al. (2015) Genome-wide comparative analysis reveals human-mouse

38. Wuest SE, et al. (2012) Molecular basis for the specification of floral organs by

26. Immink RG, et al. (2012) Characterization of SOC1's central role in flowering by

25. Gregis V, et al. (2013) Identification of pathways directly regulated by SHORT

24. Schwarz S, Grande AV, Bujdoso N, Saedler H, Huijser P (2008) The microRNA regulated

23. Kaufmann K, et al. (2009) Target genes of the MADS transcription factor SEPALLATA3:

22. Willing E-M, et al. (2015) Genome expansion of Arabis alpina linked with retro-

21. Couvreur TL, et al. (2010) Molecular phylogenetics, temporal diversification, and

20. Willing E-M, et al. (2015) Genome expansion of Arabis alpina linked with retro-

19. Mentzer L, Yee T, Wang T-Y, Himelblau E (2010) FLOWERING LOCUS C influences the

18. Edwards KD, et al. (2006) FLOWERING LOCUS C mediates natural variation in the

17. Chiang GC, Barua D, Kramer EM, Amasino RM, Donohue K (2009) Major flowering

temperature regulatory network. Plant J 82:193–207.

39. Denas O, et al. (2015) Genome-wide comparative analysis reveals human-mouse

38. Wuest SE, et al. (2012) Molecular basis for the specification of floral organs by

26. Immink RG, et al. (2012) Characterization of SOC1's central role in flowering by

25. Gregis V, et al. (2013) Identification of pathways directly regulated by SHORT

24. Schwarz S, Grande AV, Bujdoso N, Saedler H, Huijser P (2008) The microRNA regulated

23. Kaufmann K, et al. (2009) Target genes of the MADS transcription factor SEPALLATA3:

22. Willing E-M, et al. (2015) Genome expansion of Arabis alpina linked with retro-

21. Couvreur TL, et al. (2010) Molecular phylogenetics, temporal diversification, and

20. Willing E-M, et al. (2015) Genome expansion of Arabis alpina linked with retro-

19. Mentzer L, Yee T, Wang T-Y, Himelblau E (2010) FLOWERING LOCUS C influences the

18. Edwards KD, et al. (2006) FLOWERING LOCUS C mediates natural variation in the

17. Chiang GC, Barua D, Kramer EM, Amaimon RM, Donohue K (2009) Major flowering

temperature regulatory network. Plant J 82:193–207.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Mateos, JL; Tilmes, V; Madrigal, P; Severing, E; Richter, R; Rijkenberg, CWM; Krajewski, P; Coupland, G

Title:
Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species

Date:
2017-12-19

Citation:
Mateos, J. L., Tilmes, V., Madrigal, P., Severing, E., Richter, R., Rijkenberg, C. W. M., Krajewski, P. & Coupland, G. (2017). Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 114 (51), pp.E11037-E11046. https://doi.org/10.1073/pnas.1618075114.

Persistent Link:
http://hdl.handle.net/11343/250577

File Description:
published version

License:
CC BY-NC-ND