Dissection of floral induction pathways using global expression analysis

Markus Schmid¹,², N. Henriette Uhlenhaut²,*, François Godard²,†, Monika Demar¹, Ray Bressan³, Detlef Weigel¹,²,‡ and Jan U. Lohmann¹,²

¹Department of Molecular Biology, Max Planck Institute for Developmental Biology, D-72076 Tübingen, Germany
²Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
³Center for Plant Environmental Stress Physiology, Purdue University, West Lafayette, IN 47907, USA

*Present address: German Cancer Research Center, 69120 Heidelberg, Germany
†Deceased
‡Author for correspondence (e-mail: weigel@weigelworld.org)

Accepted 22 August 2003

Development 130, 6001-6012
© 2003 The Company of Biologists Ltd
doi:10.1242/dev.00842

Summary
Flowering of the reference plant Arabidopsis thaliana is controlled by several signaling pathways, which converge on a small set of genes that function as pathway integrators. We have analyzed the genomic response to one type of floral inductive signal, photoperiod, to dissect the function of several genes transducing this stimulus, including CONSTANS, thought to be the major output of the photoperiod pathway. Comparing the effects of CONSTANS with those of FLOWERING LOCUS T, which integrates inputs from CONSTANS and other floral inductive pathways, we find that expression profiles of shoot apices from plants with mutations in either gene are very similar. In contrast, a mutation in LEAFY, which also acts downstream of CONSTANS, has much more limited effects. Another pathway integrator, SUPPRESSOR OF OVEREXPRESSION OF CO 1, is responsive to acute induction by photoperiod even in the presence of the floral repressor encoded by FLOWERING LOCUS C. We have discovered a large group of potential floral repressors that are down-regulated upon photoperiodic induction. These include two AP2 domain-encoding genes that can repress flowering. The two paralogous genes, SCHLAFMÜNTE and SCHRARCHZAPFEN, share a signature with partial complementarity to the miR172 microRNA, whose precursor we show to be induced upon flowering. These and related findings on SPL genes suggest that microRNAs play an important role in the regulation of flowering.

Supplemental data available online

Key words: Arabidopsis, Floral induction, Flower development, Floral homeotic genes, Microarrays

Introduction
In contrast to animals, postembryonic development of many plants is highly plastic. A particularly dramatic example is the timing of the transition from vegetative to reproductive growth. In some species, the formation of the reproductive structures, the flowers, begins within a few days after the seedling has emerged from the seed, while in others it can take years or decades. Even within a species, the onset of flowering can vary tremendously, either because of differences in the environment or because of genetic differences.

Flowering is being studied extensively in the reference plant Arabidopsis thaliana, an ephemeral weed of the crucifer family (Lohmann and Weigel, 2002; Simpson and Dean, 2002). Many wild Arabidopsis strains flower only after several months unless they have experienced an extended period of cold, called vernalization. The vernalization requirement is conferred by a pair of epistatic loci, FRIGIDA (FRI) and FLOWERING LOCUS C (FLC), with FLC acting downstream of FRI. In plants with functional FRI, RNA levels of the floral repressor FLC are high unless the plants have been vernalized. FLC is also upregulated when genes of the so-called autonomous pathway are defective (Michaels and Amasino, 1999; Sheldon et al., 1999).

When FLC is only weakly active, Arabidopsis strains typically flower within a few weeks under long days, but considerably later when days are short. The effects of photoperiod variation are mediated by a signaling cascade that converges on the CONSTANS (CO) transcription factor (Suárez-López et al., 2001; Yanovsky and Kay, 2002), so named because co mutants are much less responsive to changes in day length than wild-type plants are (Redei, 1962). CO acts redundantly with a pathway that requires the phytohormone gibberellin, and gibberellin-deficient co mutants often do not flower at all, even under long days (Reeves and Coupland, 2001).

The different floral induction pathways are integrated by a small set of genes, including FLOWERING LOCUS T (FT), SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1) and LEAFY (LFY) (Blázquez and Weigel, 2000; Borner et al., 2000; Lee et al., 2000; Samach et al., 2000). LFY, together with another transcription factor, APETALA1 (AP1), activates homeotic genes such as APETALA3 (AP3) and AGAMOUS (AG), which specify the identity of the different organ types in newly arising floral primordia (Busch et al., 1999; Lamb et al., 2002; Ng and Yanofsky, 2001).

The critical events of early flower development are confined
Genomic DNA was isolated by a modified CTAB method. 2 g of tissue DNA isolation and labeling (Affymetrix). biotinylated cRNA were obtained. 20 the column twice to increase binding. Second, the eluate was re-

Transcript Labeling kit (Enzo). Biotinylated cRNA was cleaned with synthesis of biotinylated cRNA using the BioArray High Yield oligo(dT)-T7 primer (Genset). The cDNA served as a template for cDNA using the Superscript Choice System (Invitrogen) and an

Nitrogen. Because the expression of many floral regulators is under subjective dawn in about five groups of five from each genotype, and

or with equivalently sized leaf primordia, were frozen in liquid blades under the dissecting microscope at 30

Research article

Materials and methods

Plant material

Plants were grown under a 3:1 or 1:1 mixture of Cool White and Gro-Lux (Wide Spectrum) fluorescent lights, with a fluence rate of 125 to 175 μmol/m²/s and a temperature of about 21°C. Plants were grown initially in short days (9 hours light, 15 hours dark) and then transferred to long days (16 hours light, 8 hours dark).

Wild type was either Landsberg erecta (Ler) or Columbia (Col-0, Col-7). In experiments II and III, Ler and Col-7 contained AG::GUS transgenes (Busch et al., 1999). ify-12 is a strong allele in the Col-0 background (Huala and Sussex, 1992; Weigel et al., 1992) and co-2 and ft-2 are strong alleles in the Ler background (Kardailsky et al., 1999; Kobayashi et al., 1999; Koornneef et al., 1991; Putterill et al., 1995). The FLC FRI-SF2 strain contains the FRI allele of the San Feliu-2 (SF-2) accession introgressed into Col-0 (Lee et al., 1993). flc-3 is a strong loss-of-function allele induced in the FLC FRI-SF2 strain (Michaels and Amasino, 1999).

Scanning electron microscopy (SEM)

After fixation in methanol for 5 minutes, apices were transferred to long days (16 hours light, 8 hours dark). After centrifugation, the pellet was resuspended in 2 ml extraction buffer and carefully mixed with 2 ml lysis buffer (20 mM Tris pH 7.5, 50 mM EDTA, 2 M NaCl, 2% CTAB) and 150 μl N-laurylsarcosine. Incubation at 65°C for 20 minutes was followed by extraction with 8 ml chloroform. After precipitation with isopropanol and sodium acetate, DNA was extracted three times with phenol:chloroform:isoamylalcohol (25:24:1) and once with chloroform, precipitated again with ethanol, and resuspended in 100 μl TE buffer. DNA was fragmented by overnight digestion at 37°C using restriction enzymes Alul and Msel, followed by heat inactivation of the enzymes at 65°C for 20 minutes. DNA was extracted with phenol:chloroform:isoamylalcohol and precipitated with ethanol and sodium acetate. DNA fragments were labeled using the BioPrime System (Invitrogen) according to the manufacturer’s protocol. Labeled DNA was resuspended in 30 μl nuclease-free water and quantified by spectrophotometry. After DNA quality was determined by agarose gel electrophoresis, four individual labeling reactions were pooled to yield at least 30 μg of DNA for hybridization.

Array hybridization

Hybridization of GeneChip arrays was done according the manufacturer’s protocol (Affymetrix). For washing and staining, protocol EuKGe-WS2v4 (Affymetrix) was used. Because there was considerable variation between DNA hybridization experiments, only arrays hybridized with DNA extracted and labeled at the same time were compared (two each for Col and Ler). Using previously described algorithms (Borevitz et al., 2003), all unique features were evaluated for differential hybridization. With 3,806 single feature polymorphisms (SFPs) detected among 92,924 unique features, a false discovery rate of 5.4% was estimated, a number similar to the one reported before (Borevitz et al., 2003).

Analysis of expression data

Expression levels were estimated from Affymetrix hybridization intensity data using the robust multi array analysis (RMA) package implemented in R (Irizarry et al., 2003), or MicroArray Suite 5.0 (Affymetrix, 2001). Expression values were imported into GeneSpring 5.1 (Silicon Genetics) and normalized to the 50th percentile of each array for further analysis.

Analysis of DNA hybridization

Scanned images were saved as .CEL files using default settings of MicroArray Suite 5.0 (Affymetrix). Numeric values representing the signal of each feature were analyzed using scripts and statistical methods developed by Borevitz and colleagues (Borevitz et al., 2003) and implemented in R.

Identification of Col/Ler length polymorphisms

Primers located in the 5′ and 3′ UTRs of candidate polymorphic genes are listed in Table S1 at http://dev.biologists.org/supplemental. Genomic DNA was purified with the DNeasy Plant Mini kit (Qiagen). PCR was carried out using a 1:10 mixture of ExTaq (Takara) and Taq polymerase in ExTaq buffer with 10 pmol of each primer and 50 ng of DNA in 20 μl volume. PCR reactions were cycled for 41 times at 94°C for 20 seconds, 51°C for 30 seconds and 72°C for 5 minutes.

Real time and semi-quantitative RT-PCR

Total RNA was extracted from apices of plants grown in an independent experiment using RNase Mini columns with on-column DNase digestion (Qiagen). Reverse transcription was performed with 1 μg of total RNA, using a Reverse Transcription Kit (Promega). PCR amplification was carried out in the presence of the double-strand DNA-specific dye SYBR Green (Molecular Probes). Amplification was monitored in real time with the Opticon Continuous Fluorescence Detection System (MJR). A list of primers used is shown in Table S2 (http://dev.biologists.org/supplemental).
Identification and analysis of the ALF7 mutant

Arabidopsis plants of the Col strain were transformed with the pSKE1015 activation-tagging vector (Weigel et al., 2000) and several Activation-tagged Late-Flowering (ALF) lines were selected in the T1 generation. Plasmid rescue was used to identify the insertion point in one of these lines, ALF7. The corresponding cDNA and that of its paralog were PCR-amplified from first-strand cDNA generated from shoot apex RNA, and placed behind the CaMV 35S promoter in the pART27 derivative pMLBART (Gleave, 1992). The resulting constructs were introduced into Col wild type by Agrobacterium tumefaciens-mediated transformation (Weigel and Glazebrook, 2002).

Results and discussion

Experimental design and availability of data

To monitor global changes in gene expression, we used photolithographically produced microarrays in which each gene is represented as a probe set with several oligonucleotide features (Affymetrix GeneChips). Initial experiments were done with the Arabidopsis Genome 1 (AtGenome1) array, which contains 8,297 probe sets; the final experiment was performed with the newer Arabidopsis ATH1 array, which represents 22,810 probe sets (Table 1). For comparison across different arrays, raw data were scaled using the global intensity of all probe sets on each array. Signal intensities for each probe set were estimated from .CEL files using Affymetrix Microarray Suite (MAS) 5.0 (Affymetrix, 2001) or the log-scale robust multi array analysis (RMA) package implemented in R (Irizarry et al., 2003). The analysis presented here is based on RMA, because it produces fewer false positives when using single arrays then MAS or dChip (Li and Wong, 2001). Microarray data discussed here have been deposited with the Gene Expression Omnibus database at the NCBI (http://www.ncbi.nlm.nih.gov/geo/; series accession number GSE576 and GSE577; sample accession numbers GSM8827-8866 and GSM8868-8879). See Table S3 (http://dev.biologists.org/supplemental) for locus identifiers of genes discussed in this work.

Because many floral regulators are expressed exclusively or predominantly at the shoot apex, the site of flower formation, we first compared the sensitivity of the arrays in detecting such genes between whole 30-day-old seedlings and dissected shoot apices. The shoot apex includes the growing point of the plant, the shoot meristem, surrounded by young primordia, which before floral induction develop into leaves, and afterwards into flowers. Several meristem-specific genes, such as CUPSHAPED COTYLEDON 2 (CUC2), SHOOT MERISTEMLESS (STM) and WUSCHEL (WUS) (Aida et al., 1999; Long et al., 1996; Mayer et al., 1998), were not reliably detected in whole seedlings, but easily detectable in shoot apices (Fig. 1A). Moreover, reproducibility of the results was not compromised by the manipulations of dissection, as demonstrated by comparison of replicate arrays (Fig. 1B).

DNA polymorphisms between two wild-type strains

The probes (25mer oligonucleotides) on the Affymetrix arrays were designed using mostly information from the Columbia (Col) reference strain, whose genome has been sequenced (The Arabidopsis Genome Initiative, 2000). Because many flowering time mutants have been induced in Landsberg erecta (Ler), another commonly used laboratory strain, we evaluated the efficiency of detecting Ler sequences by labeling and hybridizing genomic DNA from Col and Ler to AtGenome 1 arrays, using procedures similar to those of Borevitz and colleagues (Borevitz et al., 2003). Less than one percent of loci contained more than six single-feature polymorphisms and these were considered as probably highly polymorphic or deleted in Ler. 26 of 31 loci that were PCR amplified had indeed Ler-specific deletions (see Table S1 at http://dev.biologists.org/supplemental). Among the other five,

### Table 1. Experiments and arrays probed

<table>
<thead>
<tr>
<th>Experiment I (AtGenome 1; 30 SD)</th>
<th>Genotype</th>
<th>+ 0 LD</th>
<th>+ 2 LD</th>
<th>+ 4 LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col (FLC fri-Col)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>FLC FRI-Sf2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ftc-3 fri-Col</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ftc-3FRI-Sf2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Experiment II (AtGenome 1; 30 SD)</td>
<td>Genotype</td>
<td>+ 0 LD</td>
<td>+ 1 LD</td>
<td>+ 2 LD</td>
</tr>
<tr>
<td>Col</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ftc-12 (Col)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ler</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>co-2 (Ler)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ft-2 (Ler)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Experiment III (ATH1; 30 SD)</td>
<td>Genotype</td>
<td>+ 0 LD</td>
<td>+ 3 LD</td>
<td>+ 5 LD</td>
</tr>
<tr>
<td>Col</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ftc-12 (Col)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Ler</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>co-2 (Ler)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ft-2 (Ler)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Numbers indicate arrays probed. In each experiment, plants were grown at the same time, but shoot apices for each array were isolated independently. The array type for each experiment is indicated.

LD, long days; SD, short days; Col, Col; Ler, Ler.

Fig. 1. Characteristics of expression estimates. (A) Expression of marker genes for the shoot apex. (B) Similarity in expression estimates between arrays of duplicate samples. (C,D) Differences in expression estimates for averages from duplicate Col and Ler arrays. Numbers indicate relative expression levels. d, days; rep, replicate number.
the RNA signal of 4 genes (At1g60130, At2g15400, At2g21060, At5g26580) was lower in Ler than in Col, suggesting that these loci are highly polymorphic.

When we used the AT1H array to compare RNA signals of vegetative shoot apices from Col and Ler, we found 961 transcripts with at least a two-fold difference in signal intensity between the two accessions. 553 of these had a lower signal in Ler, again raising the possibility that some of these differences are true polymorphic. However, 408 transcripts produced a higher signal in Ler, suggesting that these differences are true expression changes (see Fig. S1 at http://dev.biologists.org/supplemental). In conclusion, sequence polymorphisms or deletions in the Ler sequence should not be a major concern when using Affymetrix arrays for analysis of Ler-derived samples. However, there appear to be many genuine expression differences between Col and Ler (Fig. 1C,D; Fig. S1 at http://dev.biologists.org/supplemental), and it is important to consider this fact when comparing non-isogenic strains.

**Effect of day length change on two different wild-type strains**

To monitor changes in gene expression during floral induction and early flower development, we grew plants under short photoperiods (which delays flowering) for 30 days, and then transferred them to long days. In a pilot experiment, we had found that many flower-specific markers such as homeotic genes were not detected on day 0, but were robustly induced around day 6. Scanning electron microscopy confirmed that the shoot apex was vegetative at the beginning of the experiment (Fig. 2A,C). After wild-type plants had been grown in long photoperiods for 7 days, the oldest floral primordia at the end of our experiments were around stage 7 (Smyth et al., 1990). Importantly, in addition to floral primordia, release of lateral shoot primordia was evident (Hempel and Feldman, 1994). Thus, we can expect to identify in our experiments at least three classes of genes in addition to genes that are expressed in young flowers: genes that characterize young leaf primordia (which should be repressed); genes that mark the formation of side shoots (which should be induced), and genes that distinguish the shoot apical meristem before and after floral induction.

We assayed gene expression changes in the Col and Ler wild-type strains at multiple time points in two experiments (Table 1). Both experiments were performed in duplicate, using separately prepared shoot apices from plants grown at the same time. To ensure consistency in the dissection of shoot apices, each investigator performing a specific experiment participated in dissecting plants from all genotypes. Here, we focus on experiment III, in which we used AT1H arrays.

Several genes are known to be induced in the shoot meristem proper upon floral induction, including the MADS box genes SOC1 and FRUITFULL (FUL), the SQUAMOSA PROMOTER BINDING PROTEIN LIKE 3 (SPL3), SPL4 and SPL5 genes and the REM1 gene (Borner et al., 2000; Cardon et al., 1999; Cardon et al., 1997; Franco-Zorrilla et al., 2002; Hempel et al., 1997; Lee et al., 2000; Samach et al., 2000). Other genes, such as FLOWERING PROMOTING FACTORI (FPF1), are induced at the periphery (Kania et al., 1997). For the floral primordia proper (Smyth et al., 1990), several stage-specific markers are known. During stage 1, the floral meristem identity genes LFY, API and the API paralog CAULIFLOWER (CAL) are induced (Gustafson-Brown et al., 1994; Kempin et al., 1995; Weigel et al., 1992). During stage 2, SEPALLATAI (SEP1), SEP2 and SEP3 are activated, and shortly thereafter the homeotic genes AP3, PI and AG, which act in combination with the SEP genes (Drews et al., 1991; Flanagan and Ma, 1994; Goto and Meyerowitz, 1994; Jack et al., 1992; Savidge et al., 1995). Upregulation of all genes discussed above was easily detected in both Col and Ler samples (Fig. 3A-F). For the later time points, the MAS software identified almost all of them ‘present’, which is an indication of the ease with which these genes are detected.

A sequence of induction of the homeotic genes could be partially resolved in our experiments, with the C function gene AG, which is expressed in the center of the flower, being activated last (Fig. 3F). Two other well-studied genes that were robustly detected were CRABS CLAW (CRC) and WUS (Fig. S2, http://dev.biologists.org/supplemental). CRC has been reported by in situ hybridization to be activated during stage 6 of flower development, which would be toward the end of our time series (Bowman and Smyth, 1999). This is inconsistent with the profile we observed, suggesting that there is also non-localized induction of CRC. WUS marks a small group of cells in shoot and floral meristems (Mayer et al., 1998), and its
upregulation probably reflects the increase in meristem number after floral induction. Expression of STM, which is expressed more widely in the shoot apical meristem, was less markedly changed (Fig. S2, http://dev.biologists.org/supplemental).

All floral markers were induced more quickly in Ler than in Col. Both strains are relatively early flowering compared to many wild accessions, partially because they have null alleles at the FRIGIDA (FRI) locus, which is required for high expression of the floral repressor FLC (Johanson et al., 2000; Michaels and Amasino, 1999; Sheldon et al., 1999). However, Col flowers several leaves later than Ler under long days (e.g. Kardailsky et al., 1999). One genetic difference between the two strains is that the Ler allele of the floral repressor FLC is only very weakly active (Koornneef et al., 1994; Lee et al., 1994). Accordingly, we detected lower FLC levels in Ler than in Col (Fig. S3, http://dev.biologists.org/supplemental).

**CO- and FT-dependent targets of floral induction**

Activity of the CO gene is essential for perception of photoperiod differences (Koornneef et al., 1991). CO acts through at least two other genes with major effects on flowering time, FT and SOC1 (Samach et al., 2000; Suárez-López et al., 2001). Loss-of-function mutations in all three genes delay flowering under long days, with co mutations having the strongest and soc1 the weakest effects (Koornneef et al., 1998; Onouchi et al., 2000). Because FT and SOC1 integrate other cues in addition to photoperiod, mutations in both genes also delay flowering under short days, where co mutants are normal (Borner et al., 2000; Koornneef et al., 1991; Lee et al., 2000; Onouchi et al., 2000). To assess whether all effects of day length on gene expression in the shoot apex are transduced by the CO pathway, and how much of the CO effect is mediated by FT, we compared the expression profiles of Ler wild-type plants to those of co-2 and ft-2 mutants. By the end of our experiments, floral primordia were just beginning to form in co-2 and ft-2 mutants (Fig. 2I,J).

An examination of known floral marker genes revealed that co and ft had very similar effects (Fig. 3). Overall, the effects of co and ft reflected the sequence of induction in wild type. That is, early response genes, such as FUL, SOC1 and SPL3-5, were attenuated, with FUL showing the smallest change compared to wild type (Fig. 3A,B). Induction of LFY was only attenuated (Fig. 3D). Interestingly, FPF1, which is expressed in a similar temporal pattern as LFY in wild type, is affected more strongly than LFY by co and ft (Fig. 3C). There were several other genes whose expression profile across all data sets was highly correlated with that of FUL (>90%), including that of the SOC1 paralog AGL42 (Fig. 3A). The other floral markers, including AP1, CAL, the SEP genes and the homeotic genes AP3, PI and AG, were not induced in co or ft during the time course of the experiment (Fig. 3D-F). Finally, induction of CRC was only moderately attenuated in co and ft mutants (Fig. S2, http://dev.biologists.org/supplemental). This observation confirms that the CRC expression detected here must be different from the highly localized expression in carpels (Bowman and Smyth, 1999), since neither co nor ft mutants had produced stage 6 flowers by the end of the experiment.

It is notable that SOC1 was affected not only by co, but to a similar extent by ft, indicating cross-regulation between the two CO targets, FT and SOC1. LFY, which is expressed weakly during the vegetative phase (Blázquez et al., 1997; Hempel et al., 1997), was identified as 'present' by the MAS software prior to floral induction. The induction of LFY is attenuated in co mutants, but also in ft mutants, even though genetic analyses clearly show FT and LFY to act in parallel (Kardailsky et al., 1999; Kobayashi et al., 1999; Nilsson et al., 1998; Ruiz-García et al., 1997).
We used reverse transcription followed by quantitative (real-time) PCR to confirm the expression changes of several of these genes in wild-type plants and mutants, using RNAs prepared in a separate experiment from plants at 0 and 7 days after transfer to long days (Fig. S4, http://dev.biologists.org/supplemental). All genes tested were induced more strongly in wild-type plants than in the corresponding mutants, confirming the effects. RT-qPCR resulted in higher estimates for induction of these genes in wild type and mutants, using RNAs prepared in a separate experiment from plants at 0 and 7 days after transfer to long days (Fig. S4, http://dev.biologists.org/supplemental). A comparison of results for this list of genes from two replicate arrays for individual genotype-time point combinations demonstrated that the signals for most of these genes are readily reproducible (Fig. 4A). The effects of the filter are obvious in a scatter plot comparing Ler (day 0) with Ler (day 7) (Fig. 4B).

Comparison of ft (day 0) with Ler (day 0) showed that ft does not have obvious defects in the expression of floral marker genes before transfer to long days (Fig. 4C). Similarly, ft (day 0) and co (day 0) were very much alike (Fig. 4D). The effects of FT and CO on global gene expression were apparent when comparing ft (day 7) with Ler (day 7) (Fig. 4E). Importantly, ft (day 7) and co (day 7) were also very similar (Fig. 4F), confirming the results seen with a smaller selection of genes. Consistent with the morphological changes (Fig. 2I), we found that ft (day 7) was distinct from ft (day 0) (Fig. 4G).

In addition to FT and SOC1, two other CO targets, ACS10 and P5CS2, have been found using an inducible form of CO (Samach et al., 2000). Both genes were detected at high levels in all genotypes that were analyzed, but their levels did not change during the course of the experiment (Fig. S5, http://dev.biologists.org/supplemental). A possible explanation for the discrepancy is that we analyzed only material from the shoot apex, whereas Samach and colleagues (Samach et al., 2000) analyzed whole seedlings.

Integration of photoperiod and FLC activity

FLC is an important repressor of flowering that acts in parallel with the photoperiod pathway (Borner et al., 2000; Lee et al., 2000; Samach et al., 2000). Because of a deletion in the FLC activator FRI (Johanson et al., 2000), FLC levels are much reduced in Col compared to an isogenic strain with the functional FRI-Sf2 allele (Fig. S3, http://dev.biologists.org/supplemental) (Lee et al., 1993). To determine the effects of FLC on the acute response to photoperiod induction, we compared the expression of floral markers in the congeneric strains FLC FRI-Sf2, FLC fri-Col (Col wild type), flc-3 FRI-Sf2 and flc-3 fri-Col (Michaels and Amasino, 2001). Integration of photoperiod and autonomous pathways appears to be downstream of CO, since CO displays a similar induction profile in all four genotypes. We found that early induction of CAL was only moderately attenuated by FLC activity, whereas SOC1 induction was severely affected, but still detectable. In contrast, FUL induction was abolished in FLC FRI-Sf2 (Fig. 5). Thus, FLC appears to act additively with some regulators of the photoperiod pathway and epistatically with others, consistent with the notion that FLC and CO activities are integrated by the same promoters (Hepworth et al., 2002).

LFY-dependent targets of floral induction

During floral induction, several events can be

Fig. 4. Correlation analysis of the list of ‘top 500 genes in Col and Ler’ (see text for details). d, days; rep, replicate number.
distinguished at the shoot apex. First, expression of genes such as FUL or SOC1, which act upstream of floral identity genes, changes in the shoot meristem itself. After primordia on the flanks of the shoot meristem have acquired floral identity through the activity of proteins such as LFY, genes required for specification of floral organ identity are induced. While mutations in LFY affect the formation of individual flowers, ify mutants have only a small effect on other events associated with floral induction, such as stem elongation. To determine which of the genes that are affected by the transfer from short to long days are likely to be flower specific, we compared expression profiles in Col wild type and the Col-derived ify-12 mutant. In strong ify mutants, the first few flowers are replaced by leaf primordia, in the axes of which lateral shoots arise, while flowers that develop later lack petals and stamens and have some shoot characteristics (Huala and Sussex, 1992; Schultz and Haughn, 1991; Weigel et al., 1992). Scanning electron microscopy showed that, after transfer to long days, ify mutants behaved differently from co or ft mutants, as many more leaves or bracts with incipient axillary meristems in their axils were apparent (Fig. 2G).

We found that only a minority of substantial expression changes caused by transfer from short to long days was LFY dependent. In addition to known LFY targets, which are the homeotic genes AP1, AP3, PI and AG (Fig. 3F) (Busch et al., 1999; Lamb et al., 2002; Liljegren et al., 1999; Wagner et al., 1999; Weigel and Meyerowitz, 1993), the group of LFY-dependent genes includes the homeotic cofactors SEP1-3 (Fig. 3E); all 7 genes are also found in the ‘Col and Ler top 500’ list. A less dramatic effect was seen for the AP1 paralog CAL (Fig. 3D).

Next, we mined the expression profiles for genes that behaved similarly to the homeotic or the SEP genes across all data sets. This procedure resulted in 10 additional genes, of which 6 were again in the ‘Col and Ler top 500’ list (Fig. 6, Table 2). None of them was as strongly induced as the most obvious LFY targets, such as AP1, AP3 or PI. As expected, additional analyses did not identify any genes that were dependent on LFY, but not on CO or FT.

Correlation analysis using the same genes selected for global analysis of co and ft showed that Col (day 0) and ify (day 0) were very similar, indicating that LFY did not affect floral marker gene expression before induction by photoperiod. On day 0, Col and ify are more similar to each other than are Col and Ler (not shown). In contrast to CO and FT, which have dramatic effects on expression of floral marker genes (Fig. 4E, F), a comparison of Col (day 7) and ify (day 7) showed that there were few changes in ify mutants (Fig. 4I), consistent with the finding that only a small number of genes behaved similarly to known LFY targets.

Zik and Irish (Zik and Irish, 2003) have recently reported an analysis of the response of about 6,000 genes to changes in activity of the LFY targets AP3 and PI. The authors identified 47 potential AP3/PI targets, of which 42 are represented on the ATH1 array. Among these, we found only one gene, At5g22430, that is obviously affected in ify mutants (Fig. 6B).

Table 2. Newly identified genes that are activated in a LFY-dependent manner

<table>
<thead>
<tr>
<th>Locus ID</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>At2g01520</td>
<td>Major latex protein (MLP)-related</td>
</tr>
<tr>
<td>At3g04960</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>At4g21590</td>
<td>Putative bifunctional nuclease</td>
</tr>
<tr>
<td>At4g27460</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>At4g31910</td>
<td>Putative protein</td>
</tr>
<tr>
<td>At4g33790</td>
<td>Male sterility 2-like protein</td>
</tr>
<tr>
<td>At5g15150</td>
<td>Homeobox-leucine zipper protein, HAT7</td>
</tr>
<tr>
<td>At5g23900</td>
<td>Photosensitilize-responsive protein PAR-like protein</td>
</tr>
<tr>
<td>At5g37720</td>
<td>Putative protein</td>
</tr>
<tr>
<td>At5g24910</td>
<td>Cytochrome p450, putative</td>
</tr>
</tbody>
</table>

Genes were selected based on 97% correlation in the Col and ify sets with AP1, SEP1-3, PI, AP3 or AG.
Genes repressed upon floral induction

An unanticipated finding was that there are considerably more genes that are repressed upon transfer from short to long days than are induced; from our ‘Col and Ler top 500’ list, 101 genes were activated and 231 genes repressed (Fig. 7; see Table S4 at http://dev.biologists.org/supplemental, for a list of genes and their expression values). We do not think this is an artifact, because we see a similar ratio if we include a wider range, e.g., top 1000 genes. Previous molecular screens have focused on genes that are activated upon floral induction (e.g. Franco-Zorrilla et al., 1999; Melzer et al., 1990; Samach et al., 2000). Similarly, although forward genetic screens have identified several floral repressors (Mouradov et al., 2002; Simpson and Dean, 2002), only one of them, FLC, is known to be down-regulated by vernalization, a treatment that promotes flowering (Michaels and Amasino, 1999; Sheldon et al., 1999), and none has been identified that is repressed by photoperiod. A more detailed inspection of a subset of repressed genes showed that their behavior was opposite to that of the induced genes across all genotypes, i.e., the Ler response was faster than that of Col, and repression was not completely absent in co and ft mutants (Fig. 8).

Many of the genes that are known to be induced during floral induction belong to two classes of transcription factor genes, the MADS box genes and the SBP box (SPL) genes, with 69 and 15 members, respectively, represented on the ATH1 array. Among the top 101 induced genes, there were 11 MADS box genes and 5 SPL genes. In contrast, among the top 231 repressed genes, there was only one MADS box gene (AGL14) and no SPL gene. Both the enrichment of MADS box and SPL genes in the induced class and the difference between the induced and repressed classes are significant (Fisher’s exact test, P<0.001 and P<0.003, respectively).

A pair of paralogous AP2-domain genes that can repress flowering

An important question is, of course, whether any of the repressed genes play an instructive role in flowering. Coincidentally, we isolated a dominant, activation-tagged late-flowering line, ALF7. Plasmid rescue showed that the activation-tagging vector (Weigel et al., 2000) was inserted next to gene At3g54990, which encodes an AP2-domain protein that we named SCHLAFMÜTZE (SMZ) (Fig. 9A). Analysis of our expression data showed that this gene was repressed upon photoperiod change in Col and Ler wild type as well as lfy mutants, but not in co or ft mutants (Fig. 9B). At3g54990 has a close homolog, At2g39250, which is expressed at lower levels and which was named SCHNARCHZAPFEN (SNZ). The expression profiles of SMZ and SNZ were similar when analyzed by MAS, but down-regulation of SNZ was less apparent when analyzed by RMA (Fig. 9B). For both genes, we generated several transformants in which the coding sequences were placed behind the constitutive 35S promoter from cauliflower mosaic virus. Several lines in which SMZ or SNZ were under the control of the 35S promoter flowered much later than wild type (Fig. 9C), confirming that SMZ and SNZ can repress flowering. Consistent with redundant function of the two genes, SNZ knockouts flower normally. Although SMZ insertions are available, these do not interfere with RNA expression (data not shown).

Control of microRNA precursor expression by floral induction

Because their gene structure had been wrongly annotated, the phylogenetic affinity of SMZ and SNZ with AP2 and its close homologs had not been previously recognized. AP2 (At4g36920) and RAP2.7 (At2g28550) form, together with At5g67180 and At5g60120, a clade of proteins that have two AP2 domains. This clade has been identified as having potential target sites for a group of microRNAs (miRNAs) derived from a family of four precursor genes, MIR172a-1, MIR172a-2, MIR172b and MIR172c (Park et al., 2002). Although SMZ and SNZ have only a single AP2 domain, phylogenetic analysis shows that SMZ and SNZ fall within the clade defined by the other four AP2 domain proteins (not shown).

MiRNA-guided degradation of specific mRNAs has recently
been demonstrated to be important for plant morphogenesis (Palatnik et al., 2003). For the four SMZ and SNZ-related genes, Kasschau and colleagues (Kasschau et al., 2003) have shown that at least a fraction of their mRNAs is cleaved in wild-type inflorescences in the middle of the region that is complementary to the miR172 miRNAs. Experiments with dcl1 mutants and RNA blots indicate that mRNA cleavage is frequent in RAP2.7 and At5g60120, and rarer for AP2 and At5g67180 (Kasschau et al., 2003). SMZ and SNZ share the miR172 complementary motif, but with 3 or 4 mismatches (Fig. S6, http://dev.biologists.org/supplemental). Among the other four, only At5g67180 has also at least 3 mismatches, while the remaining three have 1 or 2 mismatches with at least one miR172 isoform. When we examined the expression profiles of this clade of AP2 domain encoding genes, we found that AP2, RAP2.7 and At5g60120 are down-regulated similarly to SMZ, and that their down-regulation is CO and FT dependent. The expression levels of At5g67180 also responded to floral induction, but in an opposite manner (Fig. 10A).

To determine whether the miR172 miRNAs might mediate transcript accumulation of this clade of AP2-related genes in response to floral induction, we monitored expression of four MIR172 precursor RNAs by semi-quantitative RT-PCR. We detected PCR products for four precursors and found that at
least one of them, the MIR172a-2 precursor, was up-regulated after floral induction in a CO- and FT-dependent manner (Fig. 10B). The miR172 miRNA is detected in young flowers, consistent with a role in down-regulating genes that repress flowering. Furthermore, overexpression of miR172 has the opposite effect to SMZ or SNZ overexpression, early flowering (Chen, 2003). Interestingly, miR172 appears to act also through translational repression, as deduced from overexpression experiments with one of the targets, AP2 (Chen, 2003).

The three up-regulated SPL genes discussed earlier, SPL3, SPL4 and SPL5 (Fig. 3B), have also been identified as miRNA targets (Kasschau et al., 2003; Rhoades et al., 2002). When we examined the other SPL genes represented on the Affymetrix array, we found that SPL2, SPL6, SPL9, SPL10, SPL11, SPL13 and SPL15 behave similarly to SPL3, SPL4 and SPL5, but that they reacted less strongly to floral induction (Fig. 10C). We noted that the latter three are distinguished from the rest by the presence of the miR156 miRNA target motifs in the 3’ UTR rather than the coding sequence.

Conclusions

In developmental biology, global expression analysis has been used to date mainly to discover genes or pathways affecting specific processes, but only a few studies (e.g. Hu et al., 2002; Ma et al., 2003; Strand et al., 2003) have exploited this methodology to better understand the effects of mutants with related phenotypes. We have further demonstrated the power of this approach, by analyzing the dynamic behavior of a small organ system, the shoot apex, across multiple time points and multiple genetic backgrounds.

The parallel analysis of many known floral regulatory genes, along with the analysis of a large group of newly identified genes that respond to a change in photoperiod, has allowed us to draw several important conclusions. First, two genes previously identified as CO targets by CO overexpression, ACS10 and PCSS2, do not change at the shoot apex, implying that CO also affects processes outside the region where flowers are formed. Second, consistent with the observation that the two other known CO targets, FT, has more dramatic effects than SOC1 (Onouchi et al., 2000; Samach et al., 2000), the very similar expression profiles of co and ft mutants suggest that, at the shoot apex, FT is the major output of CO. Third, the effects of the floral repressor FLC and photoperiod are additive, resulting in expression profiles of floral marker genes that are similar in plants with and without FLC, but with overall much lower levels in the presence of FLC. This finding also confirms that the similar expression profiles of co and ft are not simply due to the fact that flower formation is delayed in both mutants, since plants with high FLC levels flower even later than co or ft mutants. Fourth, compared to CO and FT, a mutation in LFY has much more subtle effects, indicating that LFY acts further downstream in the floral induction cascade, even though genetically FT and LFY act in parallel downstream of CO.

There are several additional discoveries that we have made by inspecting our data set for genes without a known role in flowering. First, we found that forward genetic analysis has been very successful in identifying many of the genes that are most strongly activated in response to floral induction. However, an equally important response to floral induction may be the repression of regulatory genes. That at least some of these repressed genes indeed have a role in flowering is confirmed by the analysis of the SMZ and SNZ genes. Second, two classes of transcription factor genes, one coding for MADS domain proteins and the other for SBP domain proteins, are highly overrepresented among the genes that are induced in response to photoperiod, both when compared to the overall complement of these families in the genome and when compared to the class of repressed genes. This observation suggests that flower-specific expression is the ancestral state for many genes in these two families. We have also found that there is a large class of genes that produce differential RNA signals between two different wild-type strains, Col and Ler (Fig. S1, http://dev.biologists.org/supplemental), which provides a rich source of candidates controlling phenotypic differences between these two strains.

How floral inductive signals are transmitted from genes such as CO and FT to downstream effectors such as LFY and AP1 is not well understood, and the newly discovered set of genes dependent on CO and FT, but not LFY, constitute a source of potential factors playing important roles in this process. We noticed several paralogous gene pairs with very similar CO and FT responses in this group, which suggests that many of these genes were not identified in forward genetic screens because of redundancy. We have discovered two groups of potential miRNA targets, a clade of AP2-domain-encoding genes and a large group of SPL genes, as being regulated by CO and FT. This observation raises the possibility that miRNAs perform a critical function in mediating the effects of floral induction, which is supported by a recent report on the consequences of miR172 overexpression (Chen, 2003). The analysis of other flowering mutants in a similar experimental design as the one used here should further clarify the regulatory interactions between the many genes already known to play a role in flowering.

This paper is dedicated to the memory of the late François Godard. We thank Justin Borevitz for help with the DNA analysis and the .cdf file translated into the Bioconductor format; Ilha Lee and Rick Amasino for the gift of FRI-Sf2 and ftc-3 strains; Heinz Schwarz and Jürgen Berger for help with SEM; Norman Warthmann for discussion and help with R. This work was supported by fellowships from the Deutsche Forschungsgemeinschaft (M.S.) and Human Frontiers Science Program (J.L.); by grants from NIH (GM62932) and HFSP to D.W., and by the Max Planck Society. D.W. is a Director of the Max Planck Institute.

Note added in proof

Aukerman and Sakai recently showed that At2g28550 (named TOE1) and At5g60120 (named TOE2) are also floral repressors (Aukerman and Sakai, 2003).

References


References