A peer-reviewed version of this preprint was published in PeerJ on 19 April 2017.

View the peer-reviewed version (peerj.com/articles/3197), which is the preferred citable publication unless you specifically need to cite this preprint.

van Dijk ADJ, Molenaar J. (2017) Floral pathway integrator gene expression mediates gradual transmission of environmental and endogenous cues to flowering time. PeerJ 5:e3197 https://doi.org/10.7717/peerj.3197
Floral pathway integrator gene expression mediates gradual transmission of environmental and endogenous cues to flowering time

Aalt D.J. van Dijk Corresp., 1, 2, 3, Jaap Molenaar 1

1 Biometris, Department for Mathematical and Statistical Methods, Wageningen University, Wageningen, The Netherlands
2 Laboratory of Bioinformatics, Wageningen University, Wageningen, The Netherlands
3 Bioscience, Wageningen University and Research, Wageningen, The Netherlands

Corresponding Author: Aalt D.J. van Dijk
Email address: aaltjan.vandijk@wur.nl

The appropriate timing of flowering is crucial for the reproductive success of plants. Hence, intricate genetic networks integrate various environmental and endogenous cues such as temperature or hormonal statues. These signals integrate into a network of floral pathway integrator genes. At a quantitative level, it is currently unclear how the impact of genetic variation in signaling pathways on flowering time is mediated by floral pathway integrator genes. Here, using datasets available from literature, we connect *Arabidopsis thaliana* flowering time in genetic backgrounds varying in upstream signalling components with the expression levels of floral pathway integrator genes in these genetic backgrounds. Our modelling results indicate that flowering time depends in a quite linear way on expression levels of floral pathway integrator genes. This gradual, proportional response of flowering time to upstream changes enables a gradual adaptation to changing environmental factors such as temperature and light.
Short title: Linear response of flowering time

Floral Pathway Integrator Gene Expression mediates Gradual Transmission of Environmental and Endogenous Cues to Flowering Time

A.D.J. van Dijk¹,²* and J. Molenaar¹

¹ Biometris, Wageningen UR, Wageningen, The Netherlands
² Bioscience, Wageningen UR, Wageningen, The Netherlands

* Corresponding author: aaltjan.vandijk@wur.nl

Abstract

The appropriate timing of flowering is crucial for the reproductive success of plants. Hence, intricate genetic networks integrate various environmental and endogenous cues such as temperature or hormonal statues. These signals integrate into a network of floral pathway integrator genes. At a quantitative level, it is currently unclear how the impact of genetic variation in signaling pathways on flowering time is mediated by floral pathway integrator genes. Here, using datasets available from literature, we connect Arabidopsis thaliana flowering time in genetic backgrounds varying in upstream signalling components with the expression levels of floral pathway integrator genes in these genetic backgrounds. Our modelling results indicate that flowering time depends in a quite linear way on expression levels of floral pathway integrator genes. This gradual, proportional response of flowering time to upstream changes enables a gradual adaptation to changing environmental factors such as temperature and light.
Introduction

The reproductive success of flowering plants depends on flowering at the right moment. Hence, plants have evolved genetic and molecular networks integrating various environmental cues with endogenous signals in order to flower under optimal conditions [1]. The signal transduction pathways that receive and transmit input signals include the photoperiod pathway, the vernalization pathway, the ambient temperature pathway, and the autonomous pathway [2]. The input from these pathways is integrated by a core set of floral pathway integrator genes [3]. The regulation of flowering time by these various factors has been extensively studied experimentally in the plant model species Arabidopsis thaliana. Substantial qualitative information is available about the factors involved and how these interact genetically, both for the signal transduction pathways and the floral pathway integrator genes [4]. Activation of the photoperiodic flowering pathway leads to transcriptional activation of FLOWERING LOCUS T (FT), an activator of flowering. FT is produced in the leaves and moves to the shoot apical meristem [5], leading to activation of SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) [6] and APETALA1 (AP1) expression [7, 8]. The vernalization (winter cold) pathway inhibits the transcription of FLOWERING LOCUS C (FLC). FLC, together with SHORT VEGETATIVE PHASE (SVP), represses the transcription of SOC1 and FT. Thus FLC acts as a flowering repressor by blocking the photoperiodic flowering pathway. In the ambient temperature pathway, which involves amongst other FLOWERING LOCUS M (FLM) and SVP, small fluctuations in temperature influence flowering time via floral pathway integrators including FT and SOC1 [9, 10]. SOC1 integrates signals from multiple pathways and transmits the outcome to LEAFY (LFY) [11, 12]; SOC1 is supposed to act at least partially via a positive feed-back loop in which AGAMOUS-LIKE 24 (AGL24) is involved upon dimerizing with SOC1 [13]. Autonomous pathway mutants are characterized by delayed flowering irrespective of day length. The proteins
encoded by the genes in the autonomous pathway generally fall into two broad functional
categories: general chromatin remodelling or maintenance factors, and proteins that affect RNA
processing [1]. Gibberellins influence the floral transition through the regulation of \textit{SOC1} and
\textit{LFY} [14]. LFY is a positive regulator of \textit{AP1} [15] and the commitment to flower is ascertained
by a direct positive feedback interaction between \textit{AP1} and \textit{LFY}. Once the expression of \textit{AP1} has
been initiated, this transcription factor orchestrates the floral transition by specifying floral
meristem identity and regulating the expression of genes involved in flower development [16].

In addition to qualitative information on pathways involved in flowering time regulation,
recently quantitative information has become available. This includes flowering time
measurements under various conditions and in different genetic backgrounds [17, 18], and time
series of expression for key floral pathway integrator genes [18]. Such quantitative information
has enabled construction of a set of models describing flowering time regulation at the molecular
level [18-21]. Given the above-described complexity, computational models are useful tools to
comprehend flowering time regulation. One example of a quantitative finding from our model
[18] for the network of floral pathway integrator genes is that a disturbance in a particular gene
has not necessarily the largest impact on directly connected genes. For example, the model
predicts that \textit{SOC1} mutation has a larger impact on \textit{AP1}, which is not directly regulated by
\textit{SOC1}, compared to its effect on \textit{LFY} which is under direct control of \textit{SOC1}. This prediction was
confirmed by expression data.

Flowering time regulation facilitates the successful dispersion of flowering plants over the world
[2] by contributing to the adaptation of plants to different environmental conditions. In this
context, it is an important question how genetic variation in the various signaling pathways
influences flowering time regulation. Can we describe the effect of genetic variation in these
signaling pathways by linking the magnitude of flowering time change to the magnitude of expression change of floral pathway integrator genes? If so, what type of relationship exists between expression levels of floral pathway integrator genes and flowering time in genetic backgrounds which differ in signaling components?

The above-mentioned quantitative analyses focus on one specific Arabidopsis genetic background, without genetic difference in signaling pathways being taken into account, leaving these questions so far unanswered. In principle, one could imagine answering these questions by extending these models to include a large number of signaling pathway components. However, construction of such large models would lead to serious complications in terms of e.g. parameter estimation. Here we follow a different route to investigate how the effect of genetic variation in components of upstream signalling pathways on flowering time is mediated by floral pathway integrator genes. We establish a quantitative connection between expression levels of floral pathway integrator genes, and flowering times in various genetic backgrounds differing in upstream signal components. This demonstrates that in many cases, floral pathway integrator genes transmit perturbations to flowering time via gradual, proportional changes in their expression levels. Our current study is complementary to our previous modelling approach which focused on the floral pathway integrator gene network, and not on the input to this network by upstream signalling components. This analysis provides a quantitative understanding of the effect of variation in the various input pathways on flowering time, which will ultimately enable us to better understand plant adaptation.
Methods

Simulations

Predictions from the dynamic flowering time model were obtained using the model as presented in ref. [18]. This consists of a set of Ordinary Differential Equations (ODEs) for the dynamics of AP1, LFY, SOC1, FD, FT and AGL24; SVP and FLC are present as external inputs in the model. In each of the six ODEs, regulation of gene expression is described by one or more terms of the form $\beta*f(x)$, where $f$ is a function of concentrations $x$ of one or more regulators. To simulate the effect of genetic variation in upstream signalling pathways influencing a given gene, the value of each parameter $\beta$ in its equation was modified by multiplying it with a factor $a$ ranging from 0.05 to 10 in steps of 0.05 and subsequently from 10 to 100 in steps of 1. The resulting flowering time after simulating the modified model was obtained, as well as the expression value of the gene itself at day 10 (this timepoint was used because it matches closely with the timepoint used in much of the experimental datasets that we used). Out of the resulting expression values, a range of ten-fold expression change was chosen around the unperturbed expression level at day 10. For SVP and FLC there is no ODE because these genes are present as external inputs in the model. For these, variation in upstream signalling pathways was simulated by simply setting the level of the gene to different fixed levels. For SVP this again involved a range of ten-fold expression change; for FLC this range was arbitrarily made larger because of the small effect of ten-fold expression change.

Experimental data
We use data from a randomly chosen subset of genes for which mutations are described as impacting flowering time [22]. Our dataset has at least several examples per floral pathway integrator gene. Data was extracted from figures or tables in papers describing the effect of mutations of particular genes on flowering time, and presenting the expression level of genes involved in signal integration. Expression measurements in different experiments are made at different days and/or different tissues, but such differences are not taken into account. Also, in particular for $FT$, often values are provided for several timepoints during one day (to capture the circadian rhythm). Although in such case in principle it would be best to record the total area under the curve (sum of expression), for simplicity, the highest observed value was used as approximation in this case.

To analyse the data, a straight line was fitted through each of the datasets: $T=Sensitivity \times x + T_0$, where $T$ is flowering time and $x$ is expression level; $Sensitivity$ and $T_0$ are parameters for which values are obtained in the fit. The R-function `lm` was used for the linear fit, and `cor.test` to test the statistical significance.

One important point in our data analysis is that various datasets were obtained using different ways of normalizing the expression values. Multiplicative normalization should effect $Sensitivity$ in a multiplicative way: if $T=S \times x + T_0$, then for $x'=a \times x$, $T=(S/a) \times x' + T_0$, i.e. $S'=S/a$. Hence, we can compare the value of $Sensitivity$ for different genes only when the same reference gene is used for normalization, and no additional relative normalization is used. The parameter $T_0$ should be independent of the normalization that is used for expression data. It would only depend on the unit of flowering time. This unit was either total leaf number or rosette leaf number; we did not observe a systematic difference for data reported in either unit and hence did not discriminate between these cases in presenting our results.
In addition to separately fitting the various datasets available for a given floral pathway integrator gene, we also obtained one model for each floral pathway integrator gene in which the various datasets were fitted simultaneously. This was performed using the R-function \texttt{nls}. In these models, each dataset obtained its own value of Sensitivity, but only one global value of $T_0$ was used for each floral pathway integrator gene.
Results

We aim to obtain a comprehensive picture of how variation in signalling pathways influences flowering time via affecting floral pathway integrator genes. To do so, we first analysed our recently published mechanistic model for the floral pathway integrator gene network [18]. This model describes regulatory interactions between the various integrator genes and is able to predict the effect of a specific perturbation in one of the genes, on all the other genes in the network. By assessing how this finally influences $AP1$ expression, the model predicts flowering time: flowering is predicted to start when $AP1$ expression passes a certain threshold. This model was developed using expression data and flowering time of wild-type *Arabidopsis thaliana*, as well as mutants of floral pathway integrator genes. In our current work, we focus on genetic variation in upstream signalling pathways, which were not used previously for modelling. To simulate variation in these upstream signalling pathways, parameters describing input to the floral pathway integrator genes were modified in the model (see Methods). This allowed to observe the dependency of predicted flowering time on expression levels of floral pathway integrator genes (Fig. 1). These plots indicate that for each gene, in an expression range of five- to tenfold around its nominal expression, the response of flowering time to expression change is approximately linear. To further analyse the response curves obtained from our model (Fig. 1), a linear model was fitted. The p-value associated with the linear fit is significant ($<10^{-15}$) for all the genes over the full range of expression displayed in Fig. 1. The obtained Pearson $R^2$ values for the linear fits are all above 0.75.

Hence, analysis of our floral pathway integrator gene regulatory network model predicts a gradual and rather linear dependence of flowering time response on changes in input to the floral regulatory network. To assess the validity of this prediction, we chose to analyze large amounts
of datasets available in literature. Numerous studies present measurements of flowering times in various conditions and for various genetic backgrounds. Since one often knows which floral pathway integrator gene is relevant for the specific signalling pathway involved, the expression levels of the specific gene thought to be responsible for mitigating the input from the signal transduction pathway are measured as well. Although one has to extract most of this data manually from tables or figures in relevant publications, it is an advantage that large amounts of data can be analysed in this way. Even though some of the individual datasets are small, in its totality the data consists of over 200 pairs of measurements of expression level and flowering time. This data has so far been scattered throughout literature and we demonstrate that it can be integrated. We use this data as a means to describe in a quantitative way the effect of changes in genetic background in signalling pathway components on flowering time. We start with a specific example regarding the floral pathway integrator gene \textit{SOC1}.

\textit{Introductory example for SOC1}

\textit{SOC1} expression measurements (qPCR) were obtained in different genetic backgrounds (\textit{cry2} and \textit{fri}, affecting the photoperiod pathway and the vernalization pathway, respectively) and different conditions [23]. For the same conditions, flowering time was also measured [23]. It is straightforward to combine these two sets of measurements in a quantitative way, although this has not yet been done so far. As shown in Figure 2A, across the different genetic backgrounds, there is a quite strong linear dependency of flowering time on the expression level of \textit{SOC1} (R$^2$=0.80). It is this dependency that is the focus of investigation of this study, for \textit{SOC1} as well as for floral pathway integrator genes. In our analysis, we focus on the effect of differences in
genetic backgrounds on each particular gene in the floral pathway integrator gene network. For that particular gene, expression level measurements might then be explanatory for flowering time changes. By analysing data as shown in Figure 2A from various publications, we are able to get a comprehensive quantitative picture how floral pathway integrator gene expression mediates transmission of environmental and endogenous cues to flowering time.

When integrating and comparing data for different experiments or different genes, one particular complication is that reported qPCR gene expression levels are normalized in various ways. In order to be able to combine datasets from different publications, one of the two following conditions should hold: (1) The same reference gene was used for normalization, and we assume that the expression level of the reference gene is constant in the different conditions applied in the various publications. In this scenario, expression levels of different genes in various publications can be quantitatively compared. Alternatively, (2) the reported expression level was scaled using wildtype expression levels of the gene of interest. In this case, in order to compare data from different publications, it is essential that the wildtype expression level that is used is the same. This seems less likely than the assumption that a reference gene such as actin or tubulin has a constant gene expression level. In several cases, the two scenarios are combined, in the sense that qPCR data are first normalized to a reference gene but that the reported expression level is subsequently scaled to a wildtype expression level.

For SOC1, the data analysed above were reported after scaling the expression level to wildtype SOC1 expression levels. Two additional examples of data for flowering time and SOC1 expression were obtained in which expression levels were normalized relative to a reference gene [24, 25] (Figure 2B-C). In one of these (Fig. 2B), there was again a clear linear relationships between the observed SOC1 expression levels and flowering time in various
backgrounds, with Pearson $R^2$ value of 0.76. In the third one, there was less evidence for a linear relationship, with Pearson $R^2$ value of 0.46 (p-value 0.3). Remarkably, it can be observed in Figure 2 that one of the two parameters in the linear equation is quite similar for each of the three datasets (78, 98 and 91, respectively). This observation is more generally true, and we will come back to it in the next section. Note that the fits in Fig. 2B and 2C are less robust than the one in Fig. 2A, but we discuss below how we can combine multiple datasets for one gene in a simultaneous fit.

Dependency of flowering time on floral pathway integrator gene expression levels

Datasets reporting gene expression levels for various floral pathway integrator genes in different genetic backgrounds, in combination with flowering time values in these genetic backgrounds, were obtained (Figure 3; Table I). We start by fitting multiple models for each gene (one per dataset). Because in some cases, the number of data points in a dataset is rather small, we subsequently fit one model per floral pathway integrator gene (see below).

As presented above for $SOC1$, linear relationships were observed between flowering time and gene expression levels (SI Figures 1-5; SI Table I). These can be described by the following equation:

$$T = Sensitivity \times ExpressionLevel + T_0$$

Here, $T$ is the observed flowering time, and the coefficients $Sensitivity$ and $T_0$ are specific for each floral pathway integrator gene. This equation describes how the measured flowering time $T$ in a given genetic background can be modelled as a linear function of the expression level of a
floral pathway integrator gene. The parameter Sensitivity describes the slope, in other words, the sensitivity of flowering time to changes in expression of the flowering time integrator network gene. Parameter $T_0$ describes the intercept with the line where ExpressionLevel equals zero. Because, as explained above, expression data can only be directly compared if the same normalization has been applied, we present values of Sensitivity and $T_0$ for each floral pathway integrator gene separately for each possible type of normalization (Table II; SI Table I). SI Figure 6 presents a histogram of the Pearson R$^2$ values obtained with the different models, indicating that in the large majority of cases the value of R$^2$ is higher than 0.75, meaning that more than 75% of the variation is explained by a simple linear model. The majority of the linear models has a significant p-value and this mainly depends on the number of datapoints available; for the cases with more than 5 datapoints, 9 out of 12 have a p-value below 0.05 (SI Table I).

In contrast to Sensitivity, $T_0$ should not depend on normalization applied to the expression data (see Methods for explanation). Hence, $T_0$ values for the same floral pathway integrator gene obtained from different datasets should be quite similar. This was indeed observed for the SOC1 datasets presented above. More generally, although there is some variation, the different values of $T_0$ obtained for a given gene are indeed significantly similar to each other compared to the values for the other genes (SI Text; SI Fig. 7). For the values of Sensitivity this is not the case, in line with our expectation.

One concern with respect to the analysis so far could be that for some of the datasets, the number of data points is rather small. We still chose to analyze such datasets initially separately because the combination of perturbations of various input pathways for the same floral pathway integrator gene allowed to demonstrate the similarity of $T_0$ values. To further deal with the concern that some of the datasets are small, we subsequently fitted one final model per floral
pathway integrator gene. This was done by allowing one $T_0$ value per floral pathway integrator
gene, but a different value of Sensitivity per dataset. In this setup, the number of data points is for
each gene larger than the number of parameters; the number of degrees of freedom ranges from 2
for $SVP$ to 72 for $FT$, and for all genes except $SVP$ and $LFY$ it is at least 30. Comparing the linear
model predictions with the experimental flowering time values indicates in most cases a clear
correspondence (Fig. 4A; SI Fig. 8). Note that $FT$ has the most deviating behaviour in the sense
that the relationship between experimental and predicted flowering time values is less linear.

The values of $T_0$ are ordered as follows: $T_{0,SVP} < T_{0,FLC} \sim T_{0,AGL24} \sim T_{0,LFY} < T_{0,FT} < T_{0,SOC1}$. $T_0$
indicates the flowering time predicted by the linear relation in case of zero gene expression,
which should be later for a flowering activator than for a flowering repressor. Hence, one would
expect activators to have higher values than repressors. This is indeed the case. Given that the
values of $T_0$ indicate the expected flowering time when the level of a specific floral pathway
integrator gene is set to zero, the values of $T_0$ can be used to predict the flowering time for
knock-out mutants of each of the floral pathway integrator genes. To validate these predictions,
we compare them with our recently obtained set of flowering times for knock-out mutants [18]
(Fig. 4B). There is a good correspondence between predictions and experimental data, although
$FT$ deviates from this pattern (Pearson $R^2$ including all cases is 0.38 between $T_0$ and flowering
time of knock-out mutants; excluding $FT$, the value of $R^2$ is 0.96 and the p-value ~0.02). Note
that $LFY$ is not included in this figure because a $lfy$ mutant does not flower properly at all [26].
The discordant behaviour of $LFY$ cannot be predicted by the simple linear analysis presented
here. We provide an alternative analysis of our flowering time ODE model for prediction of $LFY$
mutant flowering time in SI Fig. 9. $LFY$ expression was fixed at given levels and the resulting
flowering time predicted by the ODE model was recorded. For values of LFY below ~ 1nM, the
model predicts that there is no flowering. This behaviour is in accordance with the known
behaviour of the *lfy* null-mutant which was not used for training the model, providing additional
independent validation for the model.

The value of the slope of the fitted line in Figure 4B is much lower than 1. This line relates the
value of $T_0$, our prediction of flowering time, to the observed flowering time in knock-out
mutants. One reason for this small slope could be the fact that knock-out mutants in general will
not have exactly zero expression *in planta*, leading to a smaller effect on flowering time than
predicted. Nevertheless, the clear relationship between predicted and experimental flowering
time provides independent validation of the simple linear model fits from which the value of $T_0$
was obtained. Note that the flowering time and expression data used to obtain these fits are from
genetic backgrounds in which upstream signal components have been mutated. Hence, the input
data are independent from the floral pathway integrator gene knock-out mutants from which
flowering time data is used in Figure 4 for validation.
Discussion

Input from the environment is transduced by signalling pathways and integrated by a small number of floral pathway integrator genes. The complexity of the signalling pathways and their connection with the floral pathway integrator genes is overwhelming. Hence, understanding the effect of genetic variation in signalling pathways on flowering time is a daunting task. Our analysis indicates that in spite of this complexity, the effect of differences in genetic background can be quantitatively understood by focussing on expression level changes of floral pathway integrator genes. Perturbations in upstream signalling pathways effect floral pathway integrator genes mostly in such a way that the effect on flowering time is linear in the change in gene expression level. The fact that a linear response is significant in most cases, and that this response is observed for different floral pathway integrator genes, suggests that it is an import aspect of the way in which plants adapt to their local environment. The measured expression level changes are often up to tenfold or higher (Fig. 2, SI Fig 1-5). Hence, the linearity is observed over a large range of expression values.

Our findings on the role of gene expression variation in transducing the effect of genetic background variation to flowering time can be compared with more general analyses focusing on understanding the effect of variation in genetic background on phenotypes. For example, it was found in C. elegans that the effect of genetic background on the severity of RNAi and mutant phenotypes could be predicted from variation in the expression level of the affected gene [27]. Also, it has been observed that genetic variation associated with trait variation is likely to influence expression variation as well [28], suggesting that this expression variation is intermediate in establishing the link between change in genotype and change in phenotype. A recent method estimated genetically regulated gene expression and correlated these estimates
with phenotype values to identify genes involved in causing the phenotype [29]. In a broad perspective, our analysis demonstrates the possibility of analysing the dependence of quantitative traits on expression of key genes involved, which could be applied to a variety of plant traits.

Our findings are based on literature data that differ in experimental conditions such as the day or the timepoint during the day used for measurement. Although this obviously puts limit on the level of comparability between these data sets, our analysis shows that it is possible to integrate such data. One particular complicating factor is the fact that qPCR data are reported in a variety of ways. For one parameter in our model we overcome this problem by comparing data normalized in the same way. For the other parameter, this is not needed because it is independent of normalization. Nevertheless, the use of multiple qPCR reference genes would be of great value, both for better comparability between studies and also to ensure accuracy of measurements [30].

In addition to different ways of reporting expression, also different ways of reporting flowering time are used. The data we used either reported the total number of leaves, or the number of rosette leaves. Days to flowering is not often reported but would be a useful addition, in particular since leaf number and days to flowering are not always congruent [31]. A more systematic storage of qPCR data and of phenotypic measurements [32] such as flowering time would clearly also be helpful to enable large scale comparative analyses such as we present.

The linear model appeared to be successful, but less so for FT than for other genes: the value of $T_0$ obtained for FT did not correlate well with the experimental flowering time of an ft mutant (Fig. 4), and when fitting the various datasets simultaneously for each gene, there was a less clear linear relationship between predicted and observed expression for FT compared to the other
genes (SI Fig. 8). This might relate to the fact that in particular for FT, the mRNA levels measured by qPCR are only a weak proxy for the real amount of active component. This is because FT protein is transported from leaves to meristem before it may exert its effect on SOC1 and FT. Molecular aspects of this transport are not known in much detail yet, but one could imagine that there would be some kind of threshold above which not all FT is transported. If this would be the case, the predicted value of $T_o$ in our analysis would be too low, as is indeed observed when the predicted values are compared with experimental flowering times for mutants (Fig. 4). A similar threshold behaviour seems to be present in SI Figure 8 for FT. A more general scenario in which the response of flowering time to expression level of a particular floral pathway integrator gene would not necessarily be expected to be linear is if multiple floral pathway integrator genes are simultaneously effected by upstream changes. Yet another complicating factor is the fact that various floral pathway integrator genes regulate each other. This could lead to correlations in expression levels of various floral pathway integrator genes, which in turn might influence our analysis. If a gene which is directly influenced by an upstream pathway regulates another floral pathway integrator gene, both might in principle display a clear correlation between flowering time response and expression level.

In the literature, the quantitative, continuous nature of flowering time and its gradual response to changing input is often neglected when analysing the effect of variation on flowering time. In many cases, the measured response of flowering time to perturbations is reported just as leading to early or late flowering. Only a few studies analyse quantitative relationships between gene expression levels and flowering time. This includes a study in which AGL24 is shown to be a dosage-dependent mediator of flowering signals [33]. FLC levels in Arabidopsis accessions are correlated to flowering times of these accessions [34]. For rice, there is one example of analysis
of quantitative relationship between expression of an FT ortholog and flowering time [35]. Our comprehensive quantitative analysis neatly fits with these previous findings and quantifies the dosage dependence of flowering time for various floral pathway integrator genes. It indicates that the effect size of genetic variation in input pathways on flowering time can be understood via expression changes of floral pathway integrator genes. This proportional response of flowering time to upstream changes enables a gradual adaptation to changing environmental factors such as temperature and light. The continuous nature of flowering time is therefore an essential aspect of the potential of plants to adapt to various environments.
Figure 1. Dynamic model predicts linear dependency of flowering time in different genetic backgrounds on floral pathway integrator gene expression levels. The dynamic Ordinary Differential Equation (ODE) model for flowering time regulation in ref. [18] was used to simulate how flowering time (FLT) depends on gene expression level measured at day 10 for (A) AGL24 (B) SOC1 (C) LFY (D) FT (E) SVP (F) FLC. To mimic genetic variation in upstream signalling pathways, parameter values in the ODE model were modified as explained in Methods. Red points indicate the expression level of the gene at day 9-11 in the unperturbed model. Vertical dotted grey lines indicate five-fold expression range around the expression level at day 10. For FLC, the five-fold range is small compared to the displayed range and the two vertical lines fall on top of each other.
Figure 2. Dependency of flowering time (vertical axis) on SOC1 expression levels (horizontal axis) in various genetic backgrounds and various conditions, obtained in three different studies (A-C). Flowering time is reported in number of leaves; expression is normalized by scaling to wildtype expression level (A), normalized to actin (B) or normalized to tubulin (C).
Figure 3. Overview of data and analysis. Available flowering time measurements and expression levels of floral pathway integrator genes (right hand side) were obtained from literature for various genetic backgrounds. The genes from different upstream signalling pathways which were mutated in these genetic backgrounds are indicated at the left hand side. We analyse the data by modelling how expression level changes in floral pathway integrator genes (caused by genetic variation in the upstream signalling pathways) lead to quantitative changes in flowering time. In a first step, several models were obtained for each of the floral pathway integrator genes. Subsequently, one final model was obtained for each of these genes.
Figure 4. Comparison between predictions and experimental data. (A) Comparison between predicted and experimental flowering time for single linear model fitted to various SOC1 datasets. These datasets are the same as the ones used in Fig. 2, but here they are all fitted simultaneously using different values of Sensitivity but one single value of $T_0$. The number of degrees of freedom in this fit is 30. (B) Comparison between $T_0$ and flowering time of knock-out mutants. Based on fits of quantitative relationships between expression levels and flowering time, $T_0$ predicts flowering time in knock-out mutants for different floral pathway integrator genes. These predictions show a good relationship with experimentally observed flowering time for these knock-outs. Each point in this plot represents one particular floral pathway integrator gene; red outlier point indicates $ft$. 
Table I. Datasets obtained from literature\textsuperscript{a}

| Gene / reference | Mutant genotypes                                                                 | Wildtype genotype | Conditions\textsuperscript{b} | Flowering time \textsuperscript{c} |
|------------------|----------------------------------------------------------------------------------|-------------------|-------------------------------|----------------------------------|
| SOC1             |                                                                                   |                   |                               |                                  |
| [23]             | cry2, FLC-Sf2, FRI-Sf2, cry2;FLC-Sf2, cry2;FRI-Sf2, cry2;FRI-Sf2;FLC-Sf2            | Ler, Cvi          | LD, SD; 25C; day 21           |                                  |
| [25]             | gi, 35S::GI, 35S::GI; gi                                                        | Col               | LD; 22C; day 15                |                                  |
| [24]             | 35S::AGL24, 35S::SOC1, agl24                                                     | Col               | SD; 22C; day 21                |                                  |
| FT               |                                                                                   |                   |                               |                                  |
| [36]             | gi; 35S::GI, lhy, lhy; cca1, 35S::GI; lhy, gi; lhy; cca1                          | Ler               | LD; 22C; day 10                |                                  |
| [37]             | agl24, 35S::SVP, svp, soc1                                                          | Col, Ler, C24     | LD, SD; GA; 22C; day 11        |                                  |
| [5]              | cry2, cyr2, spa1                                                                    | Col, RLD          | LD; day 14                     |                                  |
| [38]             | Cdf1, cdf2, cdf3, cdf5                                                            | Col               | LD, SD; day 10                 |                                  |
| [39]             | 35S::JM18, jm18, tissue specific JM18                                              | Col               | LD; 22/18C; day 11             |                                  |
| [25]             | gi, 35S::GI, 35S::GI; gi                                                            | Col               | LD; 22C; day 15                |                                  |
| [40]             | cry2, tissue specific CRY2                                                          | Col               | LD; day 9                      |                                  |
| [41]             | elf3; elf3 enhancer and suppressor lines                                           | Col, Ler          | LL; day 14                     |                                  |
| [42]             | gi; 35S::gi; tissue specific GI                                                    | Col               | LD, SD; 23/16C; day 10         |                                  |
| [43]             | gi, spy                                                                           | Col, Ler          | LD; 22C; day 14                |                                  |
| [44]             | lwd1; lwd2, lwd1; lwd2/LWD1                                                        | Col               | LD, SD; day 18                 |                                  |
| FLC              |                                                                                   |                   |                               |                                  |
| [39]             | 35S::JM18, jm18, tissue specific JM18                                              | Col               | LD; 22/18C;                   |                                  |
Flowering time and expression data for specific floral pathway integrator genes were obtained from literature. Table includes data for each floral pathway integrator gene in which genetic background and expression data was measured. Values obtained from each dataset are presented in Fig. 2 and SI Figures 1-5. Results of fitting these data using a linear model are shown in Table II and SI Table I.

| Gene                          | Genetic Background | Expression Time | Temperature |
|-------------------------------|--------------------|-----------------|-------------|
| [23] cry2, FLC-Sf2, FRI-Sf2, cry2;FLC-Sf2, cry2;FRI-Sf2, cry2;FRI-Sf2;FLC-Sf2 | Ler, Cvi           | day 21          | TL          |
| [45] nox1, nos1, NO-donor treatment | Col                | LD; 22C; day 10 | RL          |
| [46] prmt10, prmt5            | Col                | LD; day 11      | TL          |
| [47] idl1, idl2, idl1/idl2    | Col                | LD; day 10      | TL          |
| [48] ugt87a2                  | Col                | LD; 22C; day 21 | RL          |

**SVP**

| Gene                          | Genetic Background | Expression Time | Temperature |
|-------------------------------|--------------------|-----------------|-------------|
| [41] elf3; elf3 enhancer and suppressor lines | Col, Ler           | LL; day 14      | RL          |
| [37] ft                        | Col                | LD; 22C; day 11 | TL          |

**LFY**

| Gene                          | Genetic Background | Expression Time | Temperature |
|-------------------------------|--------------------|-----------------|-------------|
| [45] nox1, nos1              | Col                | LD; 22C; day 10 | RL          |
| [48] ugt87a2                  | Col                | LD; 22C; day 21 | RL          |

**AGL24**

| Gene                          | Genetic Background | Expression Time | Temperature |
|-------------------------------|--------------------|-----------------|-------------|
| [33] AGL24-RNAi, 35S-AGL24    | Col, Ler           | LD; 23C; day 5  | RL          |
| [37] agl24-1, 35S::SVP, svp-41, soc1-2 | Col, Ler, C24     | LD, SD; GA; 22C; day 11 | TL |

a Flowering time and expression data for specific floral pathway integrator genes were obtained from literature.

b Experimental conditions: LD indicates long day, SD indicates short day, LL indicates continuous light, GA indicates gibberelin. Day indicates age of plant for which measurements were taken. If reported, temperature is indicated as well.
Flowering time measurement: RL indicates number of rosette leaves, TL indicates total number of leaves.
Table II. Linear dependencies of flowering time on expression levels

| Gene   | Normalization (number of datasets) | Sensitivity | $T_0$ |
|--------|-----------------------------------|-------------|-------|
|        | scaled (1x)                       | -0.74       | 78.3  |
|        | actin (1x)                        | -72         | 97.5  |
|        | tubulin (1x)                      | -478.9      | 90.8  |
| $SOC1$ | scaled (3x)                       | -0.30 (0.06)| 38.5 (6.0) |
|        | actin (2x)                        | -19.6 (9.95)| 45.4 (11.2) |
|        | tubulin (1x)                      | -11.5       | 29.9  |
|        | IPP2 (4x)                         | -4.0 (1.1)  | 53.4 (15) |
|        | UBQ10 (3x)                        | -363 (451)  | 45.8 (24.0) |
| $FT$   | scaled (7x)                       | 5.8 (7.1)   | 12.7 (5.1) |
|        | actin (1x)                        | 81.0        | 8.1   |
| $FLC$  | scaled (1x)                       | 0.29        | 4     |
|        | tubulin (1x)                      | 37.2        | -12.5 |
| $SVP$  | scaled (3x)                       | -5.0 (1.5)  | 14.6 (2.7) |
| $LFY$  | scaled (3x)                       | -1.7 (1.8)  | 19.6 (2.1) |
| $AGL24$| scaled (3x)                       | -4.0 (1.1)  | 53.4 (15) |

* Values for parameters in linear fit $T = \text{Sensitivity} \times \text{ExpressionLevel} + T_0$ for data shown in Fig. 2 and SI Figures 1-5. Normalization method used in the different datasets is indicated (scaled means normalization by scaling with wildtype or maximum expression value). Different normalization renders values of Sensitivity incomparable, but should not affect comparisons between values of $T_0$. Reported values are average (standard deviation) in case multiple datasets are available for the same normalization. Characteristics of individual datasets are reported in Table I. Values for Sensitivity and $T_0$ in individual datasets are reported in SI Table I.
References

1. Srikanth, A. and M. Schmid, Regulation of flowering time: all roads lead to Rome. Cell Mol Life Sci, 2011. 68(12): p. 2013-37.

2. Andres, F. and G. Coupland, The genetic basis of flowering responses to seasonal cues. Nat Rev Genet, 2012. 13(9): p. 627-39.

3. Simpson, G.G. and C. Dean, Arabidopsis, the Rosetta stone of flowering time? Science, 2002. 296(5566): p. 285-9.

4. Bouche, F., et al., FLOR-ID: an interactive database of flowering-time gene networks in Arabidopsis thaliana. Nucleic Acids Res, 2016. 44(D1): p. D1167-71.

5. Zu, Z., et al., Blue light-dependent interaction of CRY2 with SPA1 regulates COP1 activity and floral initiation in Arabidopsis. Curr Biol, 2011. 21(10): p. 841-7.

6. Yoo, S.K., et al., CONSTANS activates SUPPRESSOR OF OVEREXPRESSSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. Plant Physiol, 2005. 139(2): p. 770-8.

7. Abe, M., et al., FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science, 2005. 309(5737): p. 1052-6.

8. Wigge, P.A., et al., Integration of spatial and temporal information during floral induction in Arabidopsis. Science, 2005. 309(5737): p. 1056-9.

9. Verhage, L., G.C. Angenent, and R.G. Immink, Research on floral timing by ambient temperature comes into blossom. Trends Plant Sci, 2014. 19(9): p. 583-91.

10. Capovilla, G., M. Schmid, and D. Pose, Control of flowering by ambient temperature. J Exp Bot, 2015.

11. Immink, R.G., et al., Characterization of SOC1's central role in flowering by the identification of its upstream and downstream regulators. Plant Physiol, 2012. 160(1): p. 433-49.

12. Michaels, S.D., et al., AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. Plant J, 2003. 33(5): p. 867-74.

13. Lee, J., et al., SOC1 translocated to the nucleus by interaction with AGL24 directly regulates leafy. Plant J, 2008. 55(5): p. 832-43.

14. Eriksson, S., et al., GA4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. Plant Cell, 2006. 18(9): p. 2172-81.

15. Wagner, D., R.W. Sablowski, and E.M. Meyerowitz, Transcriptional activation of APETALA1 by LEAFY. Science, 1999. 285(5427): p. 582-4.

16. Kaufmann, K., et al., Orchestration of floral initiation by APETALA1. Science, 2010. 328(5974): p. 85-9.

17. Jung, C. and A.E. Muller, Flowering time control and applications in plant breeding. Trends Plant Sci, 2009. 14(10): p. 563-73.

18. Leal Valentim, F., et al., A quantitative and dynamic model of the Arabidopsis flowering time gene regulatory network. PLoS One, 2015. 10(2): p. e0116973.

19. Dong, Z., et al., A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. PLoS One, 2012. 7(8): p. e43450.

20. Jaeger, K.E., et al., Interlocking feedback loops govern the dynamic behavior of the floral transition in Arabidopsis. Plant Cell, 2013. 25(3): p. 820-33.

21. Salazar, J.D., et al., Prediction of photoperiodic regulators from quantitative gene circuit models. Cell, 2009. 139(6): p. 1170-9.

22. Lloyd, J. and D. Meinke, A comprehensive dataset of genes with a loss-of-function mutant phenotype in Arabidopsis. Plant Physiol, 2012. 158(3): p. 1115-29.

23. El-Din El-Assal, S., et al., The role of cryptochrome 2 in flowering in Arabidopsis. Plant Physiol, 2003. 133(4): p. 1504-16.

24. Liu, C., et al., Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. Development, 2008. 135(8): p. 1481-91.

25. Gunl, M., et al., Analysis of a post-translational steroid induction system for GIGANTEA in Arabidopsis. BMC Plant Biol, 2009. 9: p. 141.

26. Blazquez, M.A., et al., LEAFY expression and flower initiation in Arabidopsis. Development, 1997. 124(19): p. 3835-44.

27. Vu, V., et al., Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. Cell, 2015. 162(2): p. 391-402.

28. Nicolaie, D.I., et al., Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet, 2010. 6(4): p. e1000888.

29. Gamazon, E.R., et al., A gene-based association method for mapping traits using reference transcriptome data. Nat Genet, 2015.

30. Remans, T., et al., Reliable gene expression analysis by reverse transcription-quantitative PCR: reporting and minimizing the uncertainty in data accuracy. Plant Cell, 2014. 26(10): p. 3829-37.

31. Takahashi, M. and H. Morikawa, Nitrogen dioxide accelerates flowering without changing the number of leaves at flowering in Arabidopsis thaliana. Plant Signal Behav, 2014. 9(10): p. e970433.
32. Krajewski, P., et al., Towards recommendations for metadata and data handling in plant phenotyping. J Exp Bot, 2015. 66(18): p. 5417-27.
33. Yu, H., et al., AGAMOUS-LIKE 24, a dosage-dependent mediator of the flowering signals. Proc Natl Acad Sci U S A, 2002. 99(25): p. 16336-41.
34. Lempe, J., et al., Diversity of flowering responses in wild Arabidopsis thaliana strains. PLoS Genet, 2005. 1(1): p. 109-18.
35. Takahashi, Y., et al., Variations in Hd1 proteins, Hd3a promoters, and Ehd1 expression levels contribute to diversity of flowering time in cultivated rice. Proc Natl Acad Sci U S A, 2009. 106(11): p. 4555-60.
36. Mizoguchi, T., et al., Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis. Plant Cell, 2005. 17(8): p. 2255-70.
37. Li, D., et al., A repressor complex governs the integration of flowering signals in Arabidopsis. Dev Cell, 2008. 15(1): p. 110-20.
38. Fornara, F., et al., Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. Dev Cell, 2009. 17(1): p. 75-86.
39. Yang, H., et al., A companion cell-dominant and developmentally regulated H3K4 demethylase controls flowering time in Arabidopsis via the repression of FLC expression. PLoS Genet, 2012. 8(4): p. e1002664.
40. Endo, M., et al., CRYPTOCHROME2 in vascular bundles regulates flowering in Arabidopsis. Plant Cell, 2007. 19(1): p. 84-93.
41. Nefissi, R., et al., Double loss-of-function mutation in EARLY FLOWERING 3 and CRYPTOCHROME 2 genes delays flowering under continuous light but accelerates it under long days and short days: an important role for Arabidopsis CRY2 to accelerate flowering time in continuous light. J Exp Bot, 2011. 62(8): p. 2731-44.
42. Sawa, M. and S.A. Kay, GIGANTEA directly activates Flowering Locus T in Arabidopsis thaliana. Proc Natl Acad Sci U S A, 2011. 108(28): p. 11698-703.
43. Tseng, T.S., et al., SPINDLY and GIGANTEA interact and act in Arabidopsis thaliana pathways involved in light responses, flowering, and rhythms in cotyledon movements. Plant Cell, 2004. 16(6): p. 1550-63.
44. Wu, J.F., Y. Wang, and S.H. Wu, Two new clock proteins, LWD1 and LWD2, regulate Arabidopsis photoperiodic flowering. Plant Physiol, 2008. 148(2): p. 948-59.
45. He, Y., et al., Nitric oxide represses the Arabidopsis floral transition. Science, 2004. 305(5692): p. 1968-71.
46. Niu, L., et al., Regulation of flowering time by the protein arginine methyltransferase AtPRMT10. EMBO Rep, 2007. 8(12): p. 1190-5.
47. Jiang, D., et al., Arabidopsis relatives of the human lysine-specific Demethylase1 repress the expression of FWA and FLOWERING LOCUS C and thus promote the floral transition. Plant Cell, 2007. 19(10): p. 2975-87.
48. Wang, B., et al., UGT87A2, an Arabidopsis glycosyltransferase, regulates flowering time via FLOWERING LOCUS C. New Phytol, 2012. 194(3): p. 666-75.