SECTION II

RESULTS

Chaque fleur est une âme à la Nature éclose.

Every flower is a soul blossoming in Nature.

Gérard de Nerval

CHAPTER 2

A DATABASE OF FLOWERING-TIME GENES

2.1 • CONTEXT

As extensively reviewed in Chapter 1, the control of flowering time is a very complex process that involves hundreds of genes belonging to interconnected pathways. Currently, more than 300 articles are published yearly on the topic, for the sole species *Arabidopsis thaliana* (Figure 2-1). This huge amount of data is difficult to deal with for young scientists starting a Ph.D. It is also an issue for scientists only briefly interested in the control of flowering time for the purpose of their research. Even the experts in the field are progressively specialized in the processes regulating specific pathways. Besides, Arabidopsis is considered by many as the Rosetta stone of plant molecular biology (Simpson and Dean, 2002), and is used as a source in translational research on other species, such as crops. Several aspects of flowering are indeed conserved and may be used in comparative genomic studies (Izawa *et al.*, 2003; Higgins *et al.*, 2010; Jung *et al.*, 2012). However, a curated list of the flowering-time genes is necessary to standardize those comparative analyses, and such a tool was lacking in Arabidopsis.

This issue came to light during the early stages of my thesis, as I wondered if we could perform a comprehensive survey of the expression of flowering-time genes in the roots of *Arabidopsis thaliana*

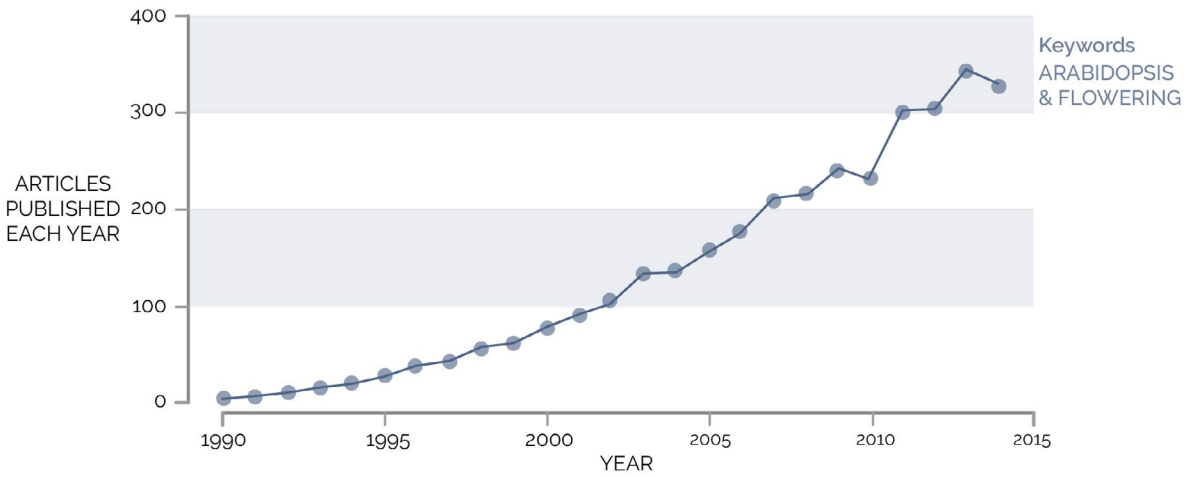

Figure 2-1. Articles published yearly on the flowering-time regulation of *Arabidopsis thaliana*. Results obtained using the research query "Arabidopsis flowering" in PubMed (http://www.ncbi.nlm.nih.gov/pubmed).

(Chapter 4). However, as stated above, such an approach required an accurate and comprehensive list of flowering-time genes. Of course, many reviews were published and some of them provided lists of key genes controlling the floral induction through specific inductive pathways (*i.e.* vernalization, photoperiod, etc.). Unfortunately, none of them included an exhaustive and up-to-date list of flowering-time genes. Therefore, I began to build my own inventory, starting from a document published by the group of Prof. G. Coupland (Fornara *et al.*, 2010). I used different methods - described in the following article - to create an accurate index of characterized flowering-time genes. However, I wanted to be able to sort these data using different parameters, such as the role of the genes in the regulation of flowering, the photoperiodic conditions in which corresponding mutants and transgenic lines were characterized, etc. Hence, I added those pieces of information to my personal list of genes, beginning the construction of a database. I produced a list of more than 300 genes, each of them being associated with relevant publications and complementary information on biological functions and phenotypes.

In parallel, to build a clear picture of the different networks involved in the control of floral induction, I started to draw schemes showing their connections/interactions. Starting from an overview, shown in Figure 2-2, I drew more and more detailed networks depicting the different pathways. To keep track of the references showing the functional links between genes, I also built a table to associate the relevant publications. The next challenge was to display those data in a convenient way. I used an existing tool that allows the mapping of data on custom schemes using R programming. Nevertheless, for several reasons, this workflow was not convenient and did not allow the realization of rapid updates and modifications. At this point, I had neither a structured database - only several very large Excel files - nor a web interface to display the data.

I discussed my project with Guillaume Lobet, a postdoc researcher interested in both the modeling of complex biological processes and the development of computer tools for the analysis of phenotypic traits; in short, a bioinformatician. He found elegant solutions to transform the Excel data into a proper database. Together, we set up a convenient workflow to build and update the interactive schemes. Finally, he designed a website providing online tools and formats that can be either used by "traditional" users or by bioinformaticians (XML format).

About a year later, our FLOWERING INTERACTIVE DATABASE (FLOR-ID) was released. The following article, published in Nucleic Acids Journal, explains the purposes and the methodology used to create the interactive database. To see the actual content of the database, I invite you to visit http://www.flor-id.org.

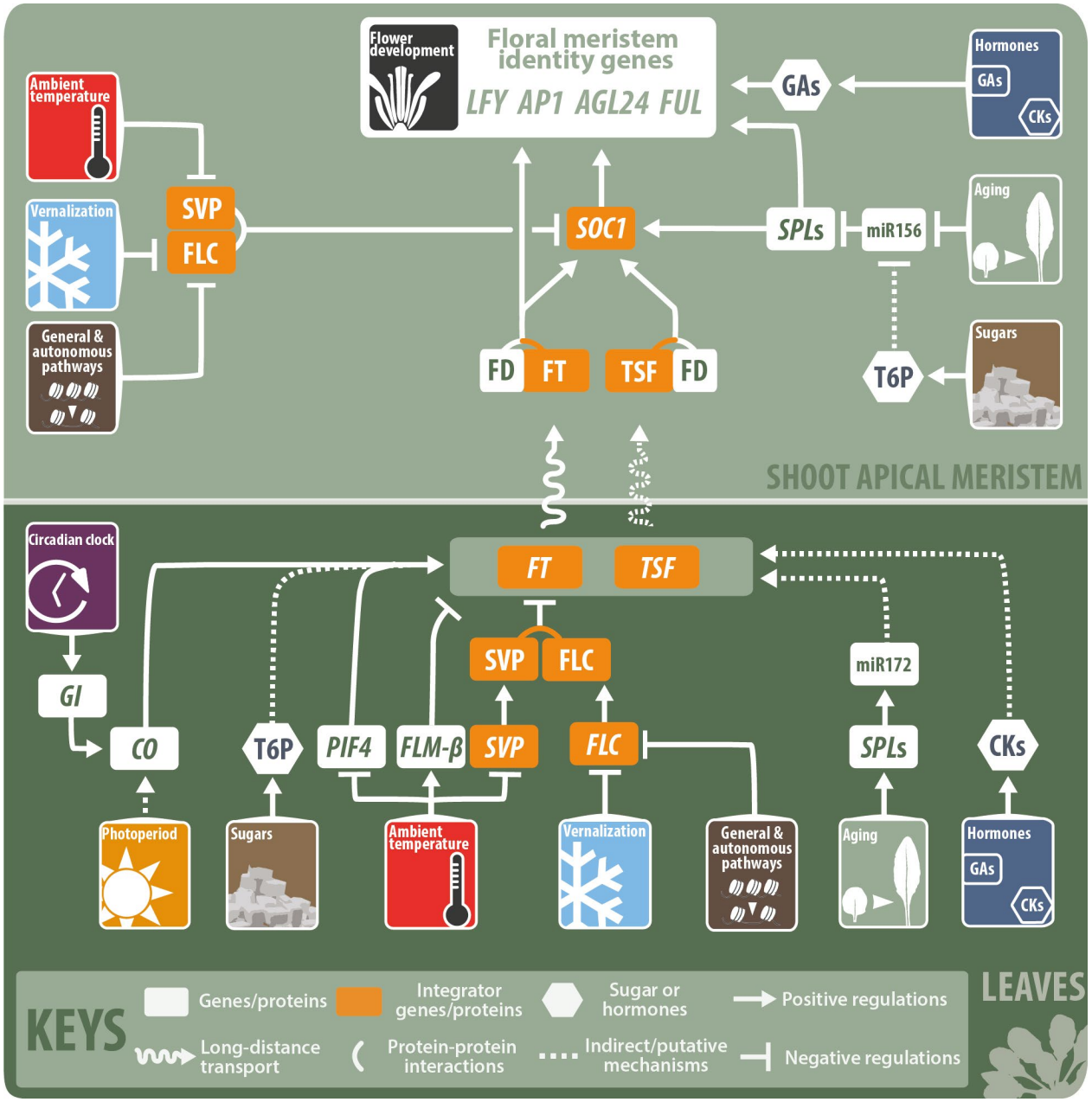

Figure 2-2. Overview of the genetic control of flowering time [adapted from www.flor-id.org].

2.2 • ARTICLE

FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*.

Nucleic Acids Research, 2015. doi: 10.1093/nar/gkv1054

Bouché, F.[1], Lobet, G.[1], Tocquin, P., and Périlleux, C.

PhytoSYSTEMS, Laboratory of Plant Physiology, University of Liège, Quartier Vallée 1 Sart Tilman Campus, 4 Chemin de la Vallée, 4000 Liège, Belgium.

[1]The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors.

FB performed the literature mining, built the data tables and drew the interactive schemes. GL built the MySQL database, the website and the related online tools. CP, PT, GL and FB participated in the amelioration of the database and the writing of the manuscript.

2.2.1 • ABSTRACT

Flowering is a hot topic in Plant Biology and important progress has been made in *Arabidopsis thaliana* toward unravelling the genetic networks involved. The increasing complexity and the explosion of literature however require development of new tools for information management and update. We therefore created an evolutive and interactive database of flowering-time genes, named FLOR-ID (Flowering-Interactive Database), that is available freely at http://www.flor-id.org. The hand-curated database contains information on 306 genes and links to 1595 publications gathering the work of more than 4500 authors. Gene/protein functions and interactions within the flowering pathways were inferred from the analysis of related publications, included in the database and translated into interactive manually drawn snapshots.

Keywords : Flowering, Arabidopsis, Photoperiod, Vernalization.

2.2.2 • INTRODUCTION

The timing of flowering significantly affects plant fitness and crop yield so that the understanding of the underlying mechanisms is a primary source for further improvement of agricultural productivity. Reproductive success requires synchronization with seasonal cycling, thus plants have evolved a complex network of regulatory pathways to sense and integrate external and endogenous signals. Among the environmental cues, daylength and temperature are the main determinants of flowering time (Bernier and Périlleux, 2005). As a result, pioneering forward genetic studies focused on the model plant *Arabidopsis thaliana* led to the categorization of flowering-time mutants according to their altered sensitivity to photoperiod ("photoperiodic pathway") and to winter cold ("vernalization pathway") (Koornneef *et al.*, 1991; Martínez-Zapater and Salinas, 1998). Mutants that were late flowering but remained sensitive to both environmental factors were classified in an "autonomous pathway" whereas the limiting effect of gibberellins, a group of phytohormones, gave its name to the the "gibberellin pathway" (Koornneef *et al.*, 1998; Levy and Dean, 1998; Reeves and Coupland, 2001). Until recently, the convergence of these four flowering pathways towards a few transcription factors, called "integrators", was still the most visible crosstalk reflecting the fine-tuning of flowering time by known stimuli (Putterill *et al.*, 2004).

The view of flowering time control has however exploded over the last years and a picture of great complexity is emerging (Blümel *et al.*, 2015). First, dissection of the genetic networks underlying the pathways revealed multiple links between their components, such as light signalling and circadian timing in the photoperiodic pathway (Song *et al.*, 2015), which explained the pleiotropic phenotypes of some flowering-time mutants (Mishra and Panigrahi, 2015; Shim and Imaizumi, 2015). Second, the identification of RNA processing and epigenetic regulation as major mechanisms of the autonomous and vernalization pathways revealed actors with rather generic roles and even became instrumental in unraveling new layers of gene regulation in plants (Marquardt *et al.*, 2006; Berry and Dean, 2015). Third, additional pathways were uncovered after the investigation of the mechanisms promoting flowering in the absence of photoperiodic cues, as the plant ages or as the surrounding temperature rises (Wang, 2014; Verhage *et al.*, 2014; Capovilla *et al.*, 2015). MicroRNAs were shown to have an important role (Teotia and Tang, 2015), as well as sugar status and signalling (Wahl *et al.*, 2013; Yu *et al.*, 2013; Yang *et al.*, 2013) or other phytohormones (Davis, 2009; D'Aloia *et al.*, 2011).

This fast progress clearly demonstrates that the genetic dissection of flowering time control was greatly facilitated by the focus on Arabidopsis. As the view of the process evolved from a set of discrete pathways towards a complex network of interconnected hubs, the number of genes involved increased from 80 (Putterill, 2001) to 180 (Fornara *et al.*, 2010). The highly intricate nature of these genetic networks, together with the diversity of the experimental evidence supporting their building, create a need to have a consolidated basis on which new players will be added. The question is even more challenging as new -omics high-throughput technologies are being used. We therefore undertook the construction of a core database named FLOR-ID (Flowering-Interactive Database) that could be progressively and regularly updated on the basis of new knowledge. We performed a careful literature survey and created a curated database of 306 flowering-time genes in Arabidopsis. We developed this tool as a website, meeting its dynamic and evolutive aims. A collection of manually drawn snapshots provides an interactive user interface.

2.2.3 • DATABASE CREATION

Identification of flowering-time associated genes

The FLOR-ID database aims at gathering information about genes involved in the regulation of flowering time. We defined "flowering-time genes" as genes whose mutation and/or overexpression alters flowering time in any Arabidopsis accession. We allocated those genes among seven pathways whereby flowering is regulated by: photoperiod, vernalization, aging, ambient temperature, hormones, sugar, or autonomously. Genes under the control of several converging pathways were defined as "flowering-time integrators".

To draw up an exhaustive list of flowering-time genes in Arabidopsis, we merged the results of four complementary approaches: [1] we started with a gene list published by Fornara and colleagues (2010); [2] we searched the UniProt database (http://www.uniprot.org) (UniProt Consortium, 2015) for all the Arabidopsis proteins associated with the keyword "flowering", excluding those isolated from pleiotropic mutants; [3] we performed a literature search in NCBI PubMed (http://www.ncbi.nlm.nih.gov/pubmed), limited to the publications of the last 10 years; [4] we analysed recent reviews on the topic. More details are provided as Supplementary Material (**Table S4-1, page S55**), including the full list of the reviews (supplementary references S1-S87).

Gene information retrieval

For each identified flowering-time gene, we retrieved the relevant publications by querying TAIR10 (http://www.arabidopsis.org) (Lamesch *et al.*, 2012), PubMed and Google Scholar (http://scholar.google.com) with the AGI locus identifier or the gene name(s). We narrowed these publication collections to the most representative and informative articles by primarily retaining those describing the cloning of the gene(s) or specifically dedicated to flowering time. When necessary, a further round of selection followed to keep most recent articles and/or those having a high number of citations.

The resulting compilation of more than 3000 full-text PDF files was then analysed to find information about the phenotypes of the mutants, overexpressors, and other transgenic lines. This was performed by contextual searches using FoxTrot Professional Search (CTM Development). Each phenotype was associated with the corresponding publication(s) and, when available, complementary information was added such as the growth conditions used for mutant phenotyping and functional data on the encoded protein. Genes were further sorted according to the flowering pathway in which they are primarily involved.

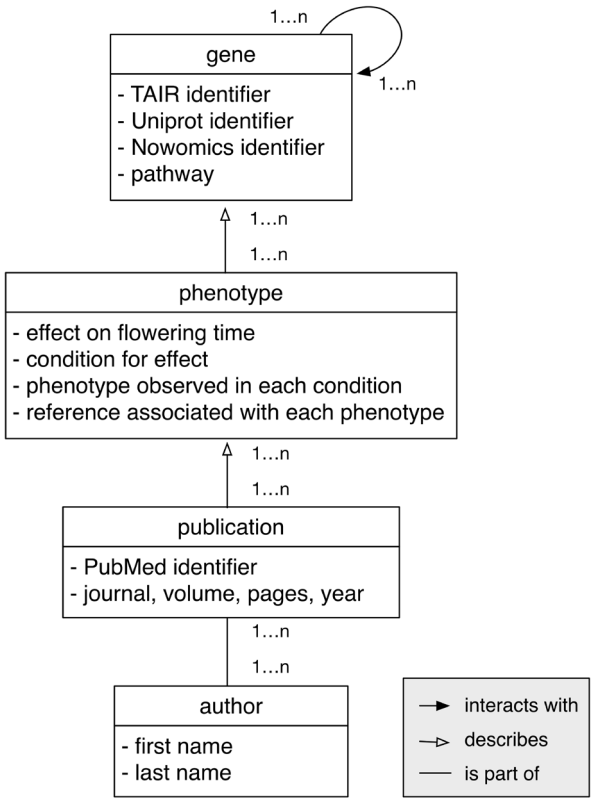

Figure 2-3. Simplified representation of the relational structure of the FLOR-ID database.

Database structure

All collected information about the flowering-time genes was incorporated into a normalized MySQL relational database: FLOR-ID (**Figure 2-3**). In addition to the information gathered from our literature survey, we used the AGI locus identifier and the reference PubMed ID to link FLOR-ID to external databases: TAIR, PubMed, Uniprot and Nowomics (https://nowomics.com/).

2.2.4 • USER INTERFACE

A freely accessible website was designed as a user frontend of FLOR-ID (http://www.flor-id.org). The database can be consulted by two ways: pre-compiled tabulated views and interactive schemes featuring flowering pathways and gene interactions.

Tabulated data

The raw content of the database is available through three thematic tables focused respectively on Genes, Publications and Authors. Data can be further filtered by any textual variable and thereafter exported as text (CSV), Excel or PDF files. On a gene-by-gene basis, the "Gene Details" pages of FLOR-ID give access to phenotypes of mutants and overexpressors, publications, protein function and interactors of the selected gene. "Interactors" here refers to genes that have been shown to act upstream or downstream of the selected gene, whether their interaction is direct or not, or to proteins if protein-protein interactions were demonstrated (**Figure 2-4**). FLOR-ID also supports programmatic extraction of information for meta-analysis purposes. Custom URL addresses are used to access the gene information as eXtended Markup Language (XML) or JavaScript Object Notation (JSON) files. For instance, retrieving information about *FLOWERING LOCUS T,* (AT1G65480) in an XML format is done with the URL http://www.phytosystems.ulg.ac.be/florid/details/?gene=AT1G65480&type=xml.

Interactive schemes

To give a visual structure to the FLOR-ID database, we produced snapshots illustrating different levels of complexity and enabling interactive access to the data. The snapshots are based on a careful analysis of the literature and supported by experimental evidence. One overview level, 7 flowering pathways and 2 complementary pictures on circadian clock and flower development are connected to detailed nested schemes (37 in total). In each scheme, interactions between genes or proteins are shown by different line types and ending styles. From a technical point of view,

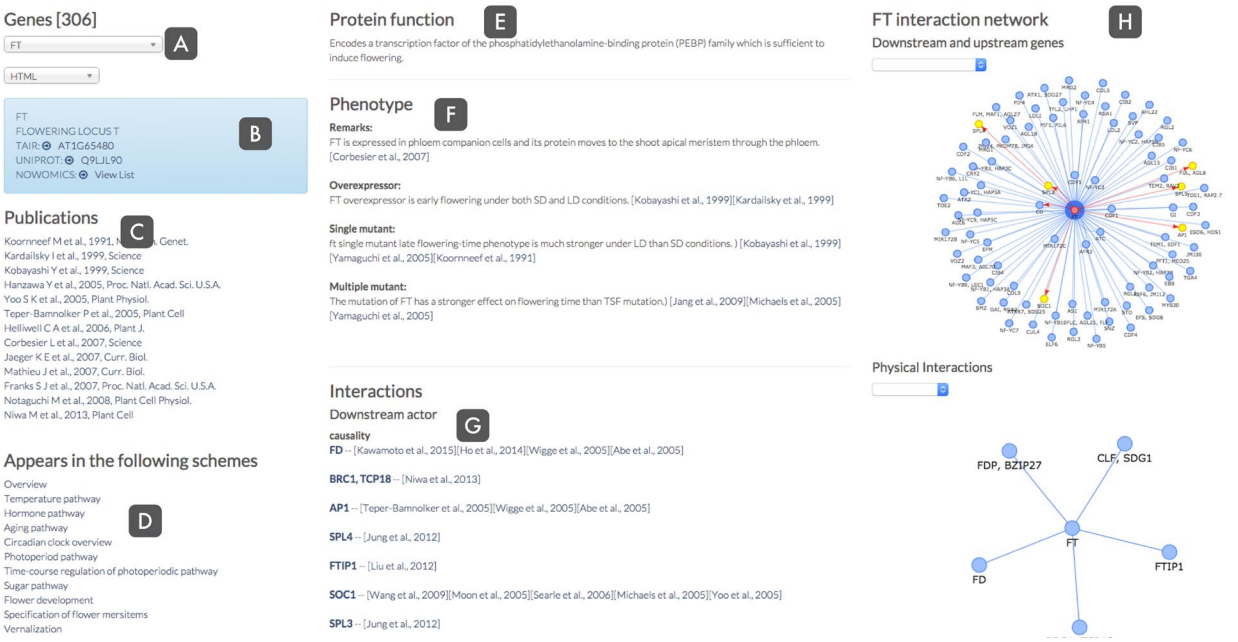

Figure 2-4. Gene-details window displayed in FLOR-ID.

A. Users can select any gene from the database (here *FLOWERING LOCUS T*) and find information displayed either in an HTML or XML format (for automated data retrieval). **B.** Gene information, with links to TAIR, UniProt and Nowomics. **C.** Relevant publications, with direct links to full-length articles **D.** FLOR-ID schemes providing network information for the gene. **E.** Protein function. **F.** Known flowering-time phenotype(s) of the corresponding mutant(s), associated with the original publication(s). **G.** Known interactors, based on stated publications. **H.** Graphical representation of the gene interaction network.

FLOR-ID houses the schemes as Scalable Vector Graphics (SVG). SVG's are natively supported in every modern browser (Chrome 43, Firefox 38, Safari 8, Internet Explorer 11). Their vectorial nature ensures a minimal loading time. It also enables the tagging of each individual element (gene or line) which, combined with modern web technologies (JavaScript), makes the schemes fully interactive (Figure 2-5). As a result, users can access the information stored in the underlying database by clicking on any element: gene names direct to mutant phenotypes, locus information and key publications, while lines and arrows lead to the papers that provided evidence for the interactions shown. This chain of information is displayed in a dynamic lateral panel. This panel also contains links to external resources (TAIR10, UniProt, and direct links to the publications). All snapshots can be downloaded as PDF files.

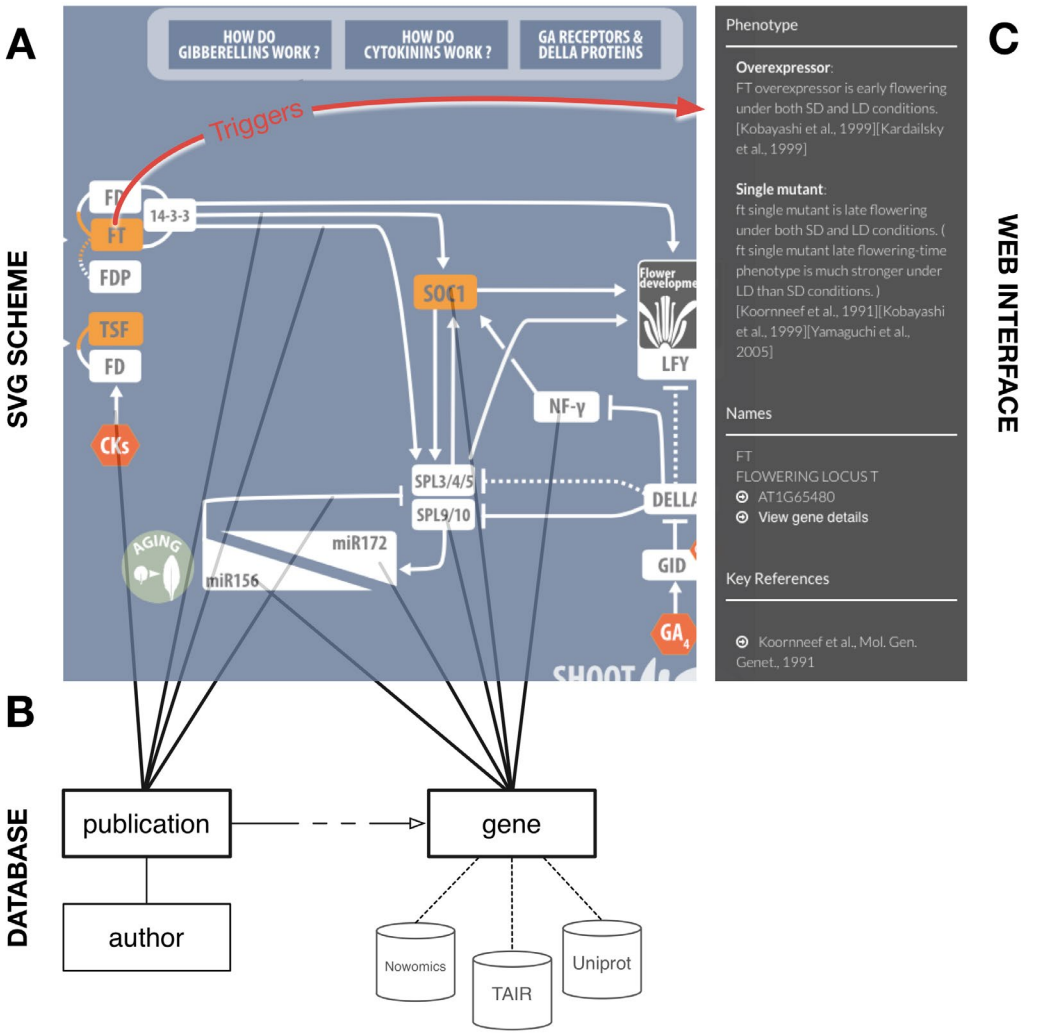

Figure 2-5. Interactive snapshots of FLOR-ID.

A. SVG scheme in which each element (gene, line) is clickable. **B.** Database connected to the scheme. **C.** Web panel opened after clicking on an element of the scheme. Information is retrieved from the database and displayed in a human-reading form.

2.2.5 • DATA CURATION

The curation of the FLOR-ID database is manual by design and was performed by the authors. Manual curation filters the information contained in the database with strong, peer-reviewed experimental data and ensures its quality. Most importantly, all gene and protein interactions shown in the interactive snapshots are supported by experimental data whereas gene interactions inferred from -omics approaches are mainly predictive.

The flowering community is large and active. In order to leverage its assets, we created a user

form allowing to suggest modifications and submit new genes, connections or publications. In a first step, this information will be displayed in a provisional table. In a second step, the database curators will validate the new item. Finally, the update will be incorporated to the public version of the database.

2.2.6 • CONCLUSIONS AND PROSPECTS

Information management is a challenge as new biology approaches disclose increasing levels of complexity and regulatory networks. Evolutive databases provide a mean to cope with the increasing amount of data and to gather information from the scientific community. Tools exist to manage raw experimental data sets, but literature is similarly exploding. The core of FLOR-ID is build around the flowering-time gene networks and contains hand-curated data. It therefore offers a reliable tool to define flowering-time genes in comparative genomic studies or transcriptomic analyses. Complementary plugins can be created to extend its functionality even further, *e.g.* for mapping raw gene expression data onto the flowering pathways.

The content of the FLOR-ID database could be easily incremented with modules expanding beyond flowering time, for example to gametophyte development. The curation of the database could then be shared between different laboratories in order to provide expert updates and ensure content accuracy. We thus hope that FLOR-ID will give the impetus to a collective, shared and structuring approach of literature and knowledge recording. Such initiative will logically contribute to community resources like Araport (Krishnakumar *et al.*, 2015).

Acknowledgements

The authors wish to acknowledge Kevin Mistiaen who contributed to the UniProt search. All members of the Plant Physiology lab are much grateful to Prof. Georges Bernier who transmitted his fascination for flowering. This work was supported by the Interuniversity Attraction Poles Programme, initiated by the Belgian Science Policy Office [P7/29]; and the Fonds de la Recherche Scientifique - FNRS [Ph.D. fellowship FC87200 to F.B., Postdoctoral research grant 1.B.237.15F to G.L.].

Supplementary Material are available in **Table S4-1** (**page S55**), including Supplementary References (S1-S87).

2.3 • GENERAL DISCUSSION

2.3.1 • ON THE IMPORTANCE OF DATABASES

The number of publications on popular topics, such as the flowering-time regulation in Arabidopsis, is constantly rising (*e.g.* **Figure 2-1, page 77**), making the management of literature laborious. The database presented above has the aim to be accessible to either flowering time experts, plant genetic students or any scientist interested for any reason in the control of flowering time. While navigating throughout the schemes provided by the database, the user can notice that the information provided expand well beyond the flowering-time regulation, as we also present figures showing the regulation of the circadian clock, the biogenesis of microRNAs, etc. The purpose was to create a self-standing website where all the information necessary to interpret the scheme are available in a graphical form. To the best of our knowledge, such curated database of literature associated with interactive schemes is an innovative way of gathering and summarizing research. As far as we know, the "enhanced snapshots" published by Cell, which consist of PDF snapshots associated with multimedia resources, are the closest publication format to our database. However, these resources are restricted to the limited interactivity allowed in PDF files. Besides, other databases provide literature related to specific processes (*e.g.* microRNAs involved in human diseases; Jiang *et al.*, 2009), but the existence of databases associating interactive snapshots with corresponding literature is uncommon. Although the "wikipathway" website (Pico *et al.*, 2008; Kelder *et al.*, 2012; http://www.wikipathways.org) provides such tool, it does not take full advantage of the modern interactivity features and relies on erratic updates performed by anonymous users. For instance, the last update of the scheme showing the flowering time control in *Arabidopsis thaliana* was performed in January 2012, and some of the provided gene interactions are not documented. We believe that, due to the exploding amount of big data analyses, the necessity for such literature-mining databases will significantly increase during the following years.

Accordingly, we think that databases must be, as much as possible, free and in open access. The journals specialized in the publication of this type of resources (*e.g.* Database (Oxford); the Database issue of "Nucleic Acids Research"; Giga Science) mostly follow this trend. We believe that in the current context of research, in which most labs work on limited funds, making these databases freely accessible to everyone is essential, although this is not always the case. For small-sized databases like FLOR-ID, this question does not arise since the quantity of curated data is relatively small. However, when considering much bigger manually curated databases,

such as The Arabidopsis Information Resource (TAIR; http://www.arabidopsis.org; Lamesch *et al.*, 2012), the question of a paid subscription is more pertinent: the curation of data represents a tremendous amount of work, and data storage requires the use of costly servers. The yearly budget to run, maintain and update TAIR is 1.6 million $ and its principal source of fundings, the US National Science Foundation cut its financing in 2013. To maintain the database, a core group of the TAIR staff created a non-profit organization with the hope provide sustainability to the database. Since TAIR is responsible for the integration of a vast collection of information, its role is essential in the landscape of Arabidopsis research. Therefore, preventing users from the access of such a gathering of knowledge by adding a pay wall is ethically questionable. TAIR curators have chosen an intermediate solution, as they charge users for the early access to new data: new annotations and phenotypes are subject to a 1-year embargo after the initial release on the website. In our opinion, this economical solution is a good compromise, which ensures both the sustainability of the database and the access to curated data, albeit with a 12-month delay, to the laboratories that cannot afford a paid subscription.

The problematic behind that cut in funding had several origins: (i) the TAIR project was funded exclusively by a US foundation but used worldwide by more than 40,000 unique users a month; (ii) TAIR is a gene-centric database that covers only a part of Arabidopsis research as large-scale molecular data are not implemented in this resource. This limitation led to a growing decentralization of the data and, therefore, impaired its attractiveness for funding agencies that wanted to invest in a comprehensive tool gathering all the knowledge about Arabidopsis.

Facing this issue, several Arabidopsis committees from different continents decided to create a new initiative, called the IAIC (International Arabidopsis Informatics Consortium, 2010; 2012), which has the aim to gather the current Arabidopsis informatics resources to rationalize their funding. As the amount of "big-data" is exponentially increasing, centralizing the access to this information is essential. Besides, the development of tools to visualize and quickly sort these data is critical to make it accessible to every scientist. In addition, the IAIC wanted to create a web resource that was also able to integrate smaller specialized databases, such as FLOR-ID. These reflections led to the creation of a portal, called Araport (http://www.araport.org; Krishnakumar *et al.*, 2015). Its aim is to bring together multiple databases and provide the tools required for their exploration and analysis. As the funding is key to the sustainability of such initiative, the IAIC gathered grants from the National Science Foundation (US) and the Biotechnology and Biological Sciences Research Council (UK), and their target is to further increase the worldwide nature of Araport financial support. It is noteworthy that, if Araport is a portal that centralizes the information, it does not replace TAIR, as its role is not to curate

and manually annotate data. Therefore, the Arabidopsis gene annotations still rely on TAIR, and one has to hope that its current funding strategy will be sustainable, albeit not optimal (Chandras *et al.*, 2009).

Overall, as the number of databases and big-data analyses grows continuously, open access portals such as Araport are essential. We are currently in discussion with Araport support to implement our dataset on their portal. Albeit at a modest scale, we hope to integrate FLOR-ID in the current growing set of online Arabidopsis research resources.

2.3.2 • FURTHER DEVELOPMENTS

Although FLOR-ID already provides a valuable overview of the flowering-time regulation in *Arabidopsis thaliana*, additional information and tools are already in preparation. For instance, in addition to the gene categories mentioned in the paper, we added a data table indexing the genes "pending" to be considered as flowering-time genes. This list contains the genes that were either (i) displaying pleiotropic defects, including an altered flowering time, but without any established link to known flowering-time genes or (ii) genes that were described as flowering-related but for which additional information is required to determine their role in the control of flowering time. Those genes may be later added to the main data tables if further experiments confirm their involvement in the control of flowering time in Arabidopsis.

The interest of such a database relies on regular updates. The new key publications, the interaction schemes, the list of flowering-time genes as well as phenotypes described are bound to evolve over time. We intend to perform those updates – including scheme changes and the addition of new genes to the database - on a regular basis (*i.e.* every six months). However, smaller updates requested by users can be performed in a matter of hours, thanks to the robust workflow we use to generate the database. Besides, we would like to implement new tools and functions to the current database structure:

(i) The first major change will be the addition of the type of functional relationship connecting two genes. Currently, the causality link does not indicate the nature of the interaction (protein-protein, protein-DNA, chromatin modifications, etc.), and we intend to add those data in future versions of FLOR-ID.

(ii) The interactions database will be updated through the addition of the techniques that were used to show the functional relationship between genes and/or proteins. This information is valuable, as some methods do not prove that the interaction oc-

curs *in vivo*. For instance, many yeast two-hybrid screenings result in several false-positive or false-negative results. Therefore, adding those pieces of information may add clues to determine the level of confidence we may give to specific genetic links.

(iii) For each technique, a descriptive scheme will summarize the corresponding methods while links will redirect the users to relevant publications (*e.g.* Nature protocols). This new feature intends to increase further the interest of the database as a teaching tool.

(iv) Given the vectorial nature of the schemes (SVG format), it is relatively easy to add any custom information. One of the major updates of the database will be to take advantage of this flexibility to map custom data on the pathway schemes. The purpose is to allow the user to either browse a database of existing experiments or map his own transcriptomic data on the schemes to see whether specific pathways are affected in their analysis. This mapping would be accompanied by several statistics, such as the number of flowering-time genes detected/absent from the study, the proportion of differentially expressed genes with a positive/negative effect on flowering, etc.

Of course, those updates may be accompanied by other add-ons suggested by users. Hence, we hope that the Arabidopsis flowering time community will participate in the amelioration of the database by providing constructive feedbacks and remarks.

In the following chapters, the list of flowering-time genes indexed the FLOR-ID database are used to filter transcriptomic analyses, thus providing concrete examples of the exploitation of the FLOR-ID list of genes.

2.4 • REFERENCES

Bernier, G. and Périlleux, C. (2005). A physiological overview of the genetics of flowering time control. Plant Biotechnol J **3**: 3–16.

Berry, S. and Dean, C. (2015). Environmental perception and epigenetic memory: mechanistic insight through *FLC*. Plant J **83**: 133–148.

Blümel, M., Dally, N., and Jung, C. (2015). Flowering-time regulation in crops—what did we learn from Arabidopsis? Curr Opin Biotechnol **32**: 121–129.

Capovilla, G., Schmid, M., and Posé, D. (2015). Control of flowering by ambient temperature. J Exp Bot **66**: 59–69.

Chandras, C., Weaver, T., Zouberakis, M., Smedley, D., Schughart, K., Rosenthal, N., Hancock, J.M., Kollias, G., Schofield, P.N., and Aidinis, V. (2009). Models for financial sustainability of biological databases and resources. Database (Oxford) **2009**: bap017–bap017.

Davis, S.J. (2009). Integrating hormones into the floral-transition pathway of *Arabidopsis thaliana*. Plant Cell Environ **32**: 1201–1210.

D'Aloia, M., Bonhomme, D., Bouché, F., Tamseddak, K., Ormenese, S., Torti, S., Coupland, G., and Périlleux, C. (2011). Cytokinin promotes flowering of Arabidopsis via transcriptional activation of the *FT* paralogue *TSF*. Plant J **65**: 972–979.

Fornara, F., de Montaigu, A., and Coupland, G. (2010). SnapShot: Control of flowering in Arabidopsis. Cell **141**: 550–550.e1–2.

Higgins, J.A., Bailey, P.C., and Laurie, D.A. (2010). Comparative Genomics of Flowering Time Pathways Using *Brachypodium distachyon* as a Model for the Temperate Grasses. PloS one **5**: e10065.

International Arabidopsis Informatics Consortium (2010). An international bioinformatics infrastructure to underpin the Arabidopsis community. Plant Cell **22**: 2530–2536.

International Arabidopsis Informatics Consortium (2012). Taking the next step: building an Arabidopsis information portal. Plant Cell **24**: 2248–2256.

Izawa, T., Takahashi, Y., and Yano, M. (2003). Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. Curr Opin Plant Biol **6**: 113–120.

Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., Li, M., Wang, G., and Liu, Y. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res **37**: D98–104.

Jung, C.-H., Wong, C.E., Singh, M.B., and Bhalla, P.L. (2012). Comparative genomic analysis of soybean flowering genes. PloS one **7**: e38250.

Kelder, T., van Iersel, M.P., Hanspers, K., Kutmon, M., Conklin, B.R., Evelo, C.T., and Pico, A.R. (2012). WikiPathways: building research communities on biological pathways. Nucleic Acids Res **40**: D1301–7.

Koornneef, M., Alonso-Blanco, C., Peeters, A.J.M., and Soppe, W. (1998). Genetic control of flowering time in *Arabidopsis*. Annu Rev Plant Physiol Plant Mol Biol **49**: 345–370.

Koornneef, M., Hanhart, C.J., and van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol Gen Genet **229**: 57–66.

Krishnakumar, V. *et al.* (2015). Araport: the Arabidopsis information portal. Nucleic Acids Res **43**: D1003–9.

Lamesch, P. *et al.* (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res **40**: D1202–10.

Levy, Y. and Dean, C. (1998). The transition to flowering. Plant Cell **10**: 1973–1990.

Marquardt, S., Boss, P.K., Hadfield, J., and Dean, C. (2006). Additional targets of the Arabidopsis autonomous pathway members, *FCA* and *FY*. J Exp Bot **57**: 3379–3386.

Martinez-Zapater, J.M., Coupland, G., Dean, C. and Koornneef, M. (1994) The transition to flowering in *Arabidopsis*. In: Meyerowitz,EM and Somerville,CR (eds). Arabidopsis, Cold Spring Harbor Laboratory Press, NY, pp. 403–433.

Mishra, P. and Panigrahi, K.C. (2015). *GIGANTEA* - an emerging story. Front Plant Sci **6**: 8.

Pico, A.R., Kelder, T., van Iersel, M.P., Hanspers, K., Conklin, B.R., and Evelo, C. (2008). WikiPathways: Pathway Editing for the People. PLoS Biology **6**: e184.

Putterill, J. (2001). Flowering in time: genes controlling photoperiodic flowering in *Arabidopsis*. Philos Trans R Soc Lond B Biol Sci **356**: 1761–1767.

Putterill, J., Laurie, R., and Macknight, R. (2004). It's time to flower: the genetic control of flowering time. Bioessays **26**: 363–373.

Reeves, P.H. and Coupland, G. (2001). Analysis of flowe-

ring time control in *Arabidopsis* by comparison of double and triple mutants. Plant Physiol **126**: 1085–1091.

Shim, J.S. and Imaizumi, T. (2015). Circadian clock and photoperiodic response in *Arabidopsis*: from seasonal flowering to redox homeostasis. Biochemistry **54**: 157–170.

Simpson, G.G. and Dean, C. (2002). *Arabidopsis*, the Rosetta stone of flowering time? Science **296**: 285–289.

Song, Y.H., Shim, J.S., Kinmonth-Schultz, H.A., and Imaizumi, T. (2015). Photoperiodic flowering: time measurement mechanisms in leaves. Annu Rev Plant Biol **66**: 441–464.

Teotia, S. and Tang, G. (2015). To bloom or not to bloom: role of microRNAs in plant flowering. Mol Plant **8**: 359–377.

UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res **43**: D204–12.

Verhage, L., Angenent, G.C., and Immink, R.G.H. (2014). Research on floral timing by ambient temperature comes into blossom. Trends Plant Sci **19**: 583–591.

Wahl, V., Ponnu, J., Schlereth, A., Arrivault, S., Langenecker, T., Franke, A., Feil, R., Lunn, J.E., Stitt, M., and Schmid, M. (2013). Regulation of flowering by trehalose-6-phosphate signaling in *Arabidopsis thaliana*. Science **339**: 704–707.

Wang, J.-W. (2014). Regulation of flowering time by the miR156-mediated age pathway. J Exp Bot **65**: 4723–4730.

Yang, L., Xu, M., Koo, Y., He, J., and Poethig, R.S. (2013). Sugar promotes vegetative phase change in *Arabidopsis thaliana* by repressing the expression of *MIR156A* and *MIR156C*. Elife **2**: e00260.

Yu, S., Cao, L., Zhou, C.-M., Zhang, T.-Q., Lian, H., Sun, Y., Wu, J., Huang, J., Wang, G., and Wang, J.-W. (2013). Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. Elife **2**: e00269.

Supplementary references

S1. Pajoro, A., Biewers, S., Dougali, E., Leal Valentim, F., Mendes, M.A., Porri, A., Coupland, G., Van de Peer, Y., van Dijk, A.D.J., Colombo, L., *et al.* (2014) The (r)evolution of gene regulatory networks controlling *Arabidopsis* plant reproduction: a two-decade history. J Exp Bot **65**: 4731–4745.

S2. Khan, M.R.G., Ai,X.-Y. and Zhang, J.-Z. (2013) Genetic regulation of flowering time in annual and perennial plants. WIREs RNA **5**: 347–359.

S3. Andrés, F. and Coupland, G. (2012) The genetic basis of flowering responses to seasonal cues. Nat Rev Genet **13**: 627–639.

S4. Matsoukas, I.G., Massiah, A.J. and Thomas, B. (2012) Florigenic and antflorigenic signaling in plants. Plant Cell Physiol **53**: 1827–1842.

S5. Posé, D., Yant, L. and Schmid, M. (2012) The end of innocence: flowering networks explode in complexity. Curr Opin Plant Biol **15**: 45–50.

S6. Turnbull, C. (2011) Long-distance regulation of flowering time. J Exp Bot **62**: 4399–4413.

S7. Srikanth, A. and Schmid, M. (2011) Regulation of flowering time: all roads lead to Rome. Cell Mol Life Sci **68**: 2013–2037.

S8. Michaels, S.D. (2009) Flowering-time regulation produces much fruit. Curr Opin Plant Biol **12**: 75–80.

S9. Galvão, V.C. and Fankhauser, C. (2015) Sensing the light environment in plants: photoreceptors and early signaling steps. Curr Opin Neurobiol **34C**: 46–53.

S10. Johansson, M. and Staiger, D. (2014) Time to flower: interplay between photoperiod and the circadian clock J Exp Bot **66**: 719–730.

S11. Romera-Branchat, M., Andrés, F. and Coupland, G. (2014) Flowering responses to seasonal cues: what's new? Curr Opin Plant Biol **21**: 120–127.

S12. Piñeiro, M. and Jarillo, J.A. (2013) Ubiquitination in the control of photoperiodic flowering. Plant Sci **198**: 98–109.

S13. Wigge, P.A. (2011) FT, a mobile developmental signal in plants. Curr Biol **21**: R374–378.

S14. Helliwell, C.A., Anderssen, R.S., Robertson, M. and Finnegan, E.J. (2015) How is *FLC* repression initiated by cold? Trends Plant Sci **20**: 76–82.

S15. Zhu, D., Rosa, S. and Dean, C. (2015) Nuclear organization changes and the epigenetic silencing of *FLC* during vernalization. J Mol Biol **427**: 659–669.

S16. Baulcombe, D.C. and Dean, C. (2014) Epigenetic regulation in plant responses to the environment. Cold Spring Harb Perspect Biol **6**: a019471.

S17. Kim, D.-H. and Sung, S. (2014) Genetic and epigenetic mechanisms underlying vernalization. Arabidopsis Book **12**: e0171.

S18. Rataj, K. and Simpson, G.G. (2014) Message ends: RNA 3' processing and flowering time control. J Exp Bot **65**: 353–363.

S19. Song, J., Irwin, J. and Dean, C. (2013) Remembering the prolonged cold of winter. Curr Biol **23**: R807–R811.

S20. Song, J., Angel, A., Howard, M. and Dean, C. (2012) Vernalization - a cold-induced epigenetic switch. J Cell Sci **125**: 3723–3731.

S21. Zografos, B.R. and Sung, S. (2012) Vernalization-mediated chromatin changes. J Exp Bot **63**: 4343–4348.

S22. Kim, D.-H. and Sung, S. (2012) Environmentally coordinated epigenetic silencing of *FLC* by protein and long noncoding RNA components. Curr Opin Plant Biol **15**: 51–56.

S23. Amasino, R. (2010) Seasonal and developmental timing of flowering. Plant J **61**: 1001–1013.

S24. Spanudakis, E. and Jackson, S. (2014) The role of microRNAs in the control of flowering time. J Exp Bot **65**: 365–380.

S25. Preston, J.C. and Hileman, L.C. (2013) Functional evolution in the plant *SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE* (*SPL*) gene family. Front Plant Sci **4**: 80.

S26. Poethig, R.S. (2013) Vegetative phase change and shoot maturation in plants. Curr Top Dev Biol **105**: 125–152.

S27. Huijser, P. and Schmid, M. (2011) The control of developmental phase transitions in plants. Development **138**: 4117–4129.

S28. Chen, X., Zhang, Z., Liu, D., Zhang, K., Li, A. and Mao, L. (2010) *SQUAMOSA* promoter-binding protein-like transcription factors: star players for plant growth and development. J Integr Plant Biol **52**: 946–951.

S29. Fornara, F. and Coupland, G. (2009) Plant phase transitions make a SPLash. Cell **138**: 625–627.

S30. Poethig, R.S. (2009) Small RNAs and developmental timing in plants. Curr Opin Genet Dev **19**: 374–378.

S31. Xu, H., Liu, Q., Yao, T. and Fu, X. (2014) Shedding light on integrative GA signaling. Curr Opin Plant Biol **21**: 89–95.

S32. Han, Y., Yang, H. and Jiao, Y. (2014) Regulation of inflorescence architecture by cytokinins. Front Plant Sci **5**: 669.

S33. Davière, J.-M. and Achard, P. (2013) Gibberellin signaling in plants. Development **140**: 1147–1151.

S34. Hedden, P. and Thomas, S.G. (2012) Gibberellin biosynthesis and its regulation. Biochem J **444**: 11–25.

S35. Frebort, I., Kowalska, M., Hluska, T., Frebortova, J. and Galuszka, P. (2011) Evolution of cytokinin biosynthesis and degradation. J Exp Bot **62**: 2431–2452.

S36. Ueguchi-Tanaka, M. and Matsuoka, M. (2010) The perception of gibberellins: clues from receptor structure. Curr Opin Plant Biol **13**: 503–508.

S37. Argueso, C.T., Ferreira, F.J. and Kieber, J.J. (2009) Environmental perception avenues: the interaction of cytokinin and environmental response pathways. Plant Cell Environ **32**: 1147–1160.

S38. Hirose, N., Takei, K., Kuroha, T., Kamada-Nobusada, T., Hayashi, H. and Sakakibara, H. (2008) Regulation of cytokinin biosynthesis, compartmentalization and translocation J Exp Bot **59**: 75–83.

S39. To, J.P.C. and Kieber, J.J. (2008) Cytokinin signaling: two-components and more. Trends Plant Sci **13**: 85–92.

S40. Liu, J., Feng, L., Li, J. and He, Z. (2015) Genetic and epigenetic control of plant heat responses. Front Plant Sci **6**: 267.

S41. Proveniers, M.C.G. and van Zanten, M. (2013) High temperature acclimation through PIF4 signaling. Trends Plant Sci **18**: 59–64.

S42. Wigge, P.A. (2013) Ambient temperature signalling in plants. Curr Opin Plant Biol **16**: 661–666.

S43. Song, Y.H., Ito, S. and Imaizumi, T. (2013) Flowering-time regulation: photoperiod- and temperature-sensing in leaves. Trends Plant Sci **18**: 575–583.

S44. Yu, S., Lian, H. and Wang, J.-W. (2015) Plant developmental transitions: the role of microRNAs and sugars. Curr Opin Plant Biol **27**: 1–7.

S45. Lunn, J.E., Delorge, I., Figueroa, C.M., Van Dijck, P. and Stitt, M. (2014) Trehalose metabolism in plants. Plant J **79**: 544–567.

S46. Tsai, A.Y.-L. and Gazzarrini, S. (2014) Trehalose-6-phosphate and SnRK1 kinases in plant development and signaling: the emerging picture. Front Plant Sci, **5**: 119.

S47. Bolouri Moghaddam, M.R. and Van den Ende, W. (2013) Sugars, the clock and transition to flowering. Front Plant Sci **4**: 22.

S48. Nolte, C. and Staiger, D. (2015) RNA around the clock - regulation at the RNA level in biological timing. Front Plant Sci **6**: 311.

S49. Cui, Z., Xu, Q. and Wang, X. (2014) Regulation of the circadian clock through pre-mRNA splicing in *Arabidopsis*. J Exp Bot **65**: 1973–1980.

S50. Hsu, P.Y. and Harmer, S.L. (2014) Wheels within wheels: the plant circadian system. Trends Plant Sci **19**: 240–249.

S51. McClung,C.R. (2014) Wheels within wheels: new transcriptional feedback loops in the *Arabidopsis* circadian clock. F1000Prime Rep **6**: 2.

S52. Staiger, D., Shin, J., Johansson, M. and Davis, S.J. (2013) The circadian clock goes genomic. Genome Biol **14**: 208.

S53. Henriques, R. and Mas, P. (2013) Chromatin remodeling and alternative splicing: pre- and post-transcriptional regulation of the *Arabidopsis* circadian clock. Semin Cell Dev Biol **24**: 399–406.

S54. Nagel, D.H. and Kay, S.A. (2012) Complexity in the Wiring and Regulation of Plant Circadian Networks Curr Biol **22**: R648–R657.

S55. Kusakina, J. and Dodd, A.N. (2012) Phosphorylation in the plant circadian system. Trends Plant Sci **17**: 575–583.

S56. Farré, E.M. (2012) The regulation of plant growth by the circadian clock. Plant Biol **14**: 401–410.

S57. McClung, C.R. (2011) The genetics of plant clocks. Adv Genet **74**: 105–139.

S58. Imaizumi, T. (2010) *Arabidopsis* circadian clock and photoperiodism: time to think about location. Curr Opin Plant Biol **13**: 83–89.

S59. de Montaigu, A., Tóth, R. and Coupland, G. (2010) Plant development goes like clockwork. Trends Genet **26**: 296–306.

S60. Song, Y.H., Ito, S. and Imaizumi, T. (2010) Similarities in the circadian clock and photoperiodism in plants. Curr Opin Plant Biol **13**: 594–603.

S61. Yakir, E., Hilman, D., Harir, Y. and Green, R.M. (2007) Regulation of output from the plant circadian clock. FEBS J **274**: 335–345.

S62. Jarillo, J.A. and Piñeiro, M. (2015) H2A.Z mediates different aspects of chromatin function and modulates flowering responses in *Arabidopsis*. Plant J **83**: 96–109.

S63. Prakash, S., Singh, R. and Lodhi, N. (2014) Histone demethylases and control of gene expression in plants. Cell Mol Biol **60**: 97–105.

S64. Holec, S. and Berger, F. (2012) Polycomb group complexes mediate developmental transitions in plants. Plant Physiol **158**: 35–43.

S65. Buzas, D.M., Tamada, Y. and Kurata, T. (2012) *FLC*: a hidden polycomb response element shows up in silence. Plant Cell Physiol **53**: 785–793.

S66. He, Y. (2012) Chromatin regulation of flowering. Trends Plant Sci **17**: 556–562.

S67. Ietswaart, R., Wu, Z. and Dean, C. (2012) Flowering time control: another window to the connection between antisense RNA and chromatin. Trends Genet **28**: 445–453.

S68. Amasino, R.M. and Michaels, S.D. (2010) The timing of flowering. Plant Physiol **154**: 516–520.

S69. Simpson, G.G. (2004) The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of *Arabidopsis* flowering time. Curr Opin Plant Biol **7**: 570–574.

S70. Becker, A. and Ehlers, K. (2015) Arabidopsis flower development -of protein complexes, targets, and transport. Protoplasma, **doi**: 10.1007/s00709-015-0812-7.

S71. Wellmer, F., Graciet, E. and Riechmann, J.L. (2014) Specification of floral organs in *Arabidopsis*. J Exp Bot, **65**, 1–9.

S72. Ó`Maoiléidigh, D.S., Graciet, E. and Wellmer, F. (2014) Gene networks controlling *Arabidopsis thaliana* flower development. New Phytol **201**: 16–30.

S73. Wellmer, F., Bowman, J.L., Davies, B., Ferrándiz, C., Fletcher, J.C., Franks, R.G., Graciet, E., Gregis, V., Ito, T., Jack, T.P., *et al.* (2014) Flower development: open questions and future directions. Methods Mol Biol **1110**: 103–124.

S74. Teo, Z.W.N., Song, S., Wang, Y.-Q., Liu, J. and Yu, H. (2014) New insights into the regulation of inflorescence architecture. Trends Plant Sci **19**: 158–165.

S75. Zhu, Q.-H. and Helliwell, C.A. (2011) Regulation of flowering time and floral patterning by miR172. J Exp Bot **62**: 487–495.

S76. Irish, V.F. (2010) The flowering of *Arabidopsis* flower development. Plant J **61**: 1014–1028.

S77. Kaufmann, K., Pajoro, A. and Angenent, G.C. (2010) Regulation of transcription in plants: mechanisms controlling developmental switches. Nat Rev Genet **11**: 830–842.

S78. Banday, Z.Z. and Nandi, A.K. (2015) Interconnection between flowering time control and activation of systemic acquired resistance. Front Plant Sci **6**: 174.

S79. Navarro, C., Cruz-Oró, E. and Prat, S. (2015) Conserved function of FLOWERING LOCUS T (FT) homologues as signals for storage organ differentiation. Curr Opin Plant Biol **23**: 45–53.

S80. Liu, L., Zhu, Y., Shen, L. and Yu, H. (2013) Emerging insights into florigen transport. Curr Opin Plant Biol **16**: 607–613.

S81. Ream, T.S., Woods, D.P. and Amasino, R.M. (2013) The molecular basis of vernalization in different plant groups. Cold Spring Harb Symp Quant Biol **77**: 105–115.

S82. Pin, P.A. and Nilsson, O. (2012) The multifaceted roles of FLOWERING LOCUS T in plant development. Plant Cell Environ **35**: 1742–1755.

S83. Siriwardana, N.S. and Lamb, R.S. (2012) The poetry of reproduction: the role of *LEAFY* in *Arabidopsis thaliana* flower formation. Int J Dev Biol **56**: 207–221.

S84. Yaish, M.W., Colasanti, J. and Rothstein, S.J. (2011) The role of epigenetic processes in controlling flowering time in plants exposed to stress. J Exp Bot **62**: 3727–3735.

S85. Li, J., Li, Y., Chen, S. and An, L. (2010) Involvement of brassinosteroid signals in the floral-induction network of *Arabidopsis*. J Exp Bot **61**: 4221–4230.

S86. Xing, D. and Li, Q.Q. (2011) Alternative polyadenylation and gene expression regulation in plants. WIREs RNA **2**: 445–458.

S87. Lee, J. and Lee, I. (2010) Regulation and function of SOC1, a flowering pathway integrator. J Exp Bot **61**: 2247–2254.