

INTEGRATING ROOTS INTO A WHOLE-PLANT MAP OF FLOWERING-TIME GENE NETWORKS IN *ARABIDOPSIS THALIANA*

FRÉDÉRIC BOUCHÉ

Ph.D. Thesis
2015-2016

Thesis Advisor: Prof. Claire Périlleux

Thèse présentée en vue de l'obtention du grade de Docteur en Sciences

LABORATORY OF PLANT PHYSIOLOGY · PHYTOSYSTEMS
DEPARTMENT OF LIFE SCIENCES
FACULTY OF SCIENCES

Université de Liège

INTEGRATING ROOTS INTO A WHOLE-PLANT MAP OF FLOWERING-TIME GENE NETWORKS IN *ARABIDOPSIS THALIANA*

2015 - 2016

Laboratory of Plant Physiology, PhytoSYSTEMS, Department of Life Sciences,
Faculty of Sciences, University of Liège

Thesis Advisor : Prof. Claire Périlleux

Thesis Comitee : Prof. Claire Périlleux (ULg), Prof. Patrick du Jardin (ULg), Prof. Jacques Dommes (ULg), and Dr. Pierre Tocquin (ULg).

Ph.D. Examining Committee : Prof. Maria Albani (University of Köln), Prof. François Chaumont (UCL, Belgium), Prof. Patrick du Jardin (ULg), Prof. Jacques Dommes (ULg), Dr. Marc Hanikenne (ULg), Prof. Claire Périlleux (ULg), and Dr. Pierre Tocquin (ULg).

À mes grands-parents.

ABSTRACT OF THE THESIS

Flowering is a crucial step in plant development that needs to be carefully regulated to occur at the right time of the year, thus ensuring reproductive success. In *Arabidopsis thaliana*, several interconnected molecular networks have been disclosed that mediate flowering response to environmental cues, such as photoperiod and temperature, or to endogenous factors, such as plant age or hormones. Many of these signalling pathways are systemic, *i.e.* involve regulatory mechanisms distant from the shoot apical meristem where floral transition eventually occurs. However, most investigations were focused on the aerial parts of the plant but ignored the roots.

The aim of this Ph.D. thesis was to integrate the roots into a comprehensive overview of the genetic control of flowering in Arabidopsis. A prerequisite was to obtain a full list of known flowering-time genes. This step led to the creation of a database of flowering-time genes, which is accessible online and in which users can navigate through data tables or interactive schemes (www.flor-id.org).

In the second part of the work, we studied the involvement of the roots in the differential developmental rates of plants grown in hydroponics and on soil. This analysis revealed that the growing medium had a striking effect on the juvenile-to-adult phase transition, but had only a moderate impact on flowering time. These differences were associated with significant changes in the root transcriptome, as more than 2,000 genes were differentially expressed. By crossing the list of flowering-time genes with these transcriptomic data, we observed that a large proportion of the genes involved in the initiation of reproductive development was expressed in roots, although only a small subset was differentially expressed in the two growing substrates. Among them, we found the potent repressor of flowering, *FLOWERING LOCUS C* (*FLC*), which was upregulated on soil. However, further experiments with Arabidopsis lines expressing *FLC* to different levels demonstrated that this gene does not mediate the medium effect on plant developmental rate.

In the third part of the work, we used a data mining approach that confirmed the expression of about 200 flowering-time genes in the roots of Arabidopsis. Genes of the photoperiodic pathway were overrepresented but most of the key regulators were not expressed, suggesting that flowering-time genes may be involved in other regulatory processes in the roots. Using a complementary approach, we analysed the root transcriptome to identify early changes occurring during the induction of flowering by a photoperiodic treatment. We found 595 differentially regulated genes, the majority of which were previously qualified as circadian. Interestingly, clock genes such as *GIGANTEA* (*GI*) and *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) showed a delay in expression upon extension of the photoperiod. Enrichment analysis in root-tissue expression, gene ontology terms and promoter sequences converged onto the identification of sugars as putative signals triggering the transcriptomic changes in the roots. Consistently, *TREHALOSE-6-PHOSPHATE SYNTHASE* 1 (*TPS1*) was upregulated during the inductive photoperiod. Moreover, upregulation of cytokinin biosynthesis suggested involvement of the roots in a feed-forward loop promoting flowering.

Collectively, the results presented in this work brought new insights in the regulation of flowering time at the whole-organism scale by integrating the "hidden part" of plants in the current landscape of the molecular processes controlling phase transitions in *Arabidopsis thaliana*.

RÉSUMÉ DE LA THÈSE

Dans le cycle de développement des plantes, la floraison est un processus crucial qui doit être initié alors que les conditions environnementales sont favorables, de manière à assurer le succès du processus reproductif. Dans ce but, la floraison d'*Arabidopsis thaliana* est contrôlée par de nombreux réseaux génétiques interconnectés qui sont influencés par plusieurs signaux environnementaux, tels que la photopériode ou la vernalisation, ainsi que par des facteurs endogènes, tels que l'âge de la plante et son statut hormonal. Les processus moléculaires régulant la floraison impliquent des signaux systémiques qui migrent au travers de la plante, notamment vers le méristème apical de tige, où ils induisent la floraison. Ces processus ont été étudiés de manière assidue au niveau de la partie aérienne, tandis que les racines ont longtemps été ignorées.

Le but de cette thèse de doctorat est de caractériser le contrôle génétique de la floraison à l'échelle de l'organisme entier, en utilisant pour cela une approche holistique visant à intégrer les racines dans les processus moléculaires contrôlant l'initiation du développement reproductif. Un prérequis à cette analyse consistait à dresser une liste exhaustive des gènes de floraison afin d'établir lesquels d'entre eux étaient exprimés dans les racines. Cette étape nous a mené à la création d'une base de données interactive accessible en ligne : FLOR-ID (www.flor-id.org).

Dans la seconde partie de ce travail, nous avons déterminé l'importance de l'environnement racinaire sur le développement des plantes en comparant les transitions de phase ainsi que le transcriptome d'individus cultivés sur terreau et en hydroponie. Cette analyse a révélé un effet conséquent du milieu de culture sur la transition depuis la phase juvénile vers la phase adulte, accélérée en hydroponie, tandis que le moment de floraison n'est que peu affecté. Ces modifications sont accompagnées par des changements significatifs du transcriptome racinaire : plus de 2000 gènes montrent une forte variation d'expression en fonction du milieu de culture. En croisant la liste des gènes de floraison avec cette analyse transcriptomique, nous avons constaté que, dans une importante proportion, les gènes impliqués dans la transition vers le développement reproductif sont exprimés dans les racines. Parmi eux, le répresseur de floraison *FLOWERING LOCUS C* (*FLC*) était exprimé à un niveau plus élevé dans les racines de plantes cultivées sur terreau. Cependant, plusieurs expériences physiologiques menées sur des lignées exprimant *FLC* à des niveaux variables ont révélé que ce gène n'était pas impliqué dans les différences de transitions de phase causées par le milieu de culture.

Dans la dernière partie de ce travail, nous avons employé une méthode d'exploration de données (data mining) de manière à identifier les gènes de floraison couramment détectés dans les racines : cette approche nous a permis de confirmer que nombre d'entre eux y sont exprimés, parmi lesquels de multiples gènes impliqués dans la voie photopériodique. Cependant, la plupart des gènes essentiels à la transition vers le développement reproductif, les « intégrateurs » de la floraison, ne sont pas exprimés dans les racines. Nous avons également employé une méthode complémentaire visant à déterminer les changements transcriptomiques précoces se déroulant dans les racines lors de l'induction photopériodique de la floraison. Cette analyse a révélé l'existence de changements conséquents, puisque 595

gènes étaient différentiellement exprimés dans les racines au cours de l'induction florale. Parmi les processus altérés, nous avons pu déceler l'existence d'un décalage de l'expression de certains gènes de l'horloge circadienne, tels que *GIGANTEA* (*GI*) et *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*). Plusieurs méthodes bio-informatiques ont montré le probable rôle des sucres dans les changements transcriptomiques observés. Cette hypothèse est corroborée par l'augmentation du niveau d'expression du gène *TREHALOSE-6-PHOSPHATE SYNTHASE 1* (*TPS1*), qui est impliqué dans les réactions métaboliques liées aux sucres. Qui plus est, l'induction de l'expression de gènes impliqués dans la biosynthèse de cytokinines suggère l'implication des racines dans une boucle de rétroactivation participant à l'induction de la floraison.

Les résultats récoltés au cours de cette thèse de doctorat ouvrent de nouvelles perspectives concernant la régulation des transitions de phase à l'échelle de l'organisme entier, en y intégrant les racines.

TABLE OF CONTENTS

TABLE OF CONTENTS

SECTION IV – Supplementary Materials

PREFACE

An investment in knowledge pays the best interest.

Benjamin Franklin

PREFACE

Plants are crucial for humanity as they provide oxygen, food, and are the source of a considerable number of drugs used in modern medicine. Plant genetic material accounts for more than half the DNA content on earth (Landenmark *et al.*, 2015) and approximately 20 % of the calories consumed in the world come from wheat only (Regina *et al.*, 2006). Because of the continuous growth in human population, the global demand for agricultural products is increasing and will continue so for decades (Tilman *et al.*, 2011). Major challenges are thus to meet these higher needs by sustainable ways (Dirzo and Raven, 2003; Vitousek *et al.*, 2008) and to take into account how climate changes will affect productivity (Craufurd and Wheeler, 2009). Plant research is therefore essential. The "golden rice" is a striking example of the applications of plant genetics. Deficiency in Vitamin A kills more than 600,000 children yearly (Black *et al.*, 2008). To answer this major health problem, scientists re-engineered the whole pathway of beta-carotene - the precursor of Vitamin A - to increase its contents in transgenic rice (Ye *et al.*, 2000; Paine *et al.*, 2005), thus helping to prevent shortage in concerned populations.

One of the key research topics closely related to potential improvements in agriculture is the genetic regulation of flowering time. Indeed, a better understanding of the molecular mechanisms controlling this developmental step is necessary to ensure adaptation of the cultivated materials to local climate conditions and to improve yields.

Most economically important plants are large, have relatively long life cycles and possess complex genomes, thus making genetic studies difficult. This complexity is the reason why many scientists rather focused on a small subset of "model plants". Over the last 25 years, the pole position was claimed by a small weed, *Arabidopsis thaliana*. The absence of economic interest and its apparent simplicity both played a crucial role in its selection as a worldwide studied model organism. The Arabidopsis genome was the first one to be sequenced among flowering plants (Arabidopsis Genome Initiative, 2000), leading to the rapid development of molecular and bioinformatics tools that allowed fast progress in genetics studies. Currently, thousands of peer-reviewed articles mentioning Arabidopsis in their title are published yearly: over 2000 according to our estimation (**Figure P1**; more than 3500 according to Provart *et al.*, 2015). However, the rapid improvement of high-throughput sequencing techniques and the development of efficient genome editing methods will likely shed more and more light on eco-

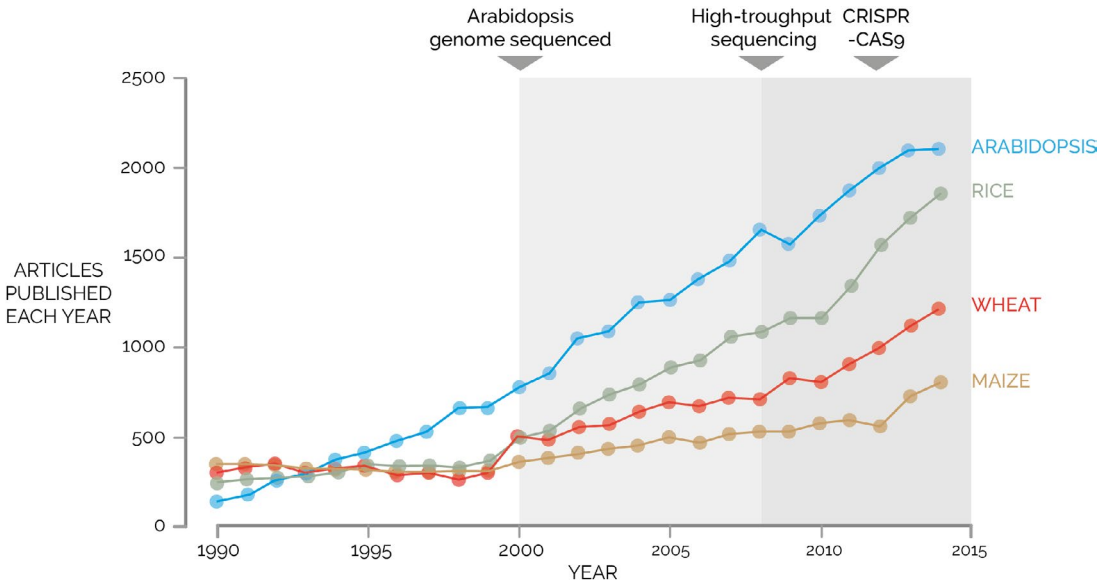


Figure P1. Articles published on model flowering plants over the last 25 years.
Data were retrieved from PubMed (http://www.ncbi.nlm.nih.gov/pubmed) using a query formatted to retrieve papers containing species names in their title.

nomically relevant crops, such as wheat, rice or maize. The rise in crop-related publications is indeed highly significant since 2010 (Figure P1). Nevertheless, Arabidopsis will likely remain a resource in translational research as it allows tackling complex regulatory mechanisms (*e.g.* microRNA biogenesis, chromatin remodeling, characterization of biologically active peptides, etc.). The regulation of flowering time is one of those complex processes where Arabidopsis demonstrated its relevance, as more than 300 genes have been identified in this species, which define several interconnected pathways (see Chapter 2, page 77; Bouché *et al.*, 2015a).

My Ph.D. thesis aims to bring some new insights into the regulation of flowering time in *Arabidopsis thaliana* by examining a question that has been under-investigated so far: the role of the roots.

The manuscript is structured around four major parts:

(i) The introduction (Chapter 1, page 13) broaches several aspects of the development of Arabidopsis. Each section of the introduction is linked with one of the following chapters, so that the introduction can be read in parts before going to the corresponding experimental chapter and results.

(ii) The first part (Chapter 2,page 77) logically follows the literature survey as it describes the creation of an interactive database indexing the flowering-time ge-

nes described in *Arabidopsis thaliana*. This chapter has been accepted for publication (Bouché *et al.*, 2015a) and the database is available online at http://www.flor-id.org:

> **Bouché, F., Lobet, G., Tocquin, P. and Périlleux, C.** (2015a) FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. Nucleic Acids Res, **doi**: 10.1093/nar/gkv1054.

(iii) The second part (Chapter 3, page 97) aims at evaluating the impact of the growing medium on plant development, through the comparison of plants grown in hydroponics and on soil. This section combines physiological, molecular and genetic analyses.

(iv) The third part (Chapter 4, page 137) focuses on the participation of the roots in the flowering process. First, we carried out a data-mining analysis of the flowering-time genes expressed in roots. Then, we performed a global transcriptomic analysis of the roots during the photoperiodic induction of flowering. Those analyses led to the selection of mutants that were subsequently characterized. This chapter is presented as a research paper, submitted for publication:

> **Bouché, F., D'Aloia, M., Tocquin, P., Lobet, G., Detry N., and Périlleux, C** (Submitted), Rooting the flowering process in *Arabidopsis thaliana*.

Supplemental Chapters S1 (page S3) and S2 (page S33) show some of the side projects of my thesis, dealing with the process of vernalization. These results are published:

> **Périlleux, C., Pieltain, A., Jacquemin, G., Bouché, F., Detry, N., D'Aloia, M., Thiry, L., Aljochim, P., Mathieu, A.-S., Lutts, S., and Tocquin., P.** (2013) A root chicory MADS box sequence and the Arabidopsis flowering repressor *FLC* share common features that suggest conserved function in vernalization and de-vernalization responses. Plant J **75:** 390–402.

> **Bouché, F., Detry, N. and Périlleux, C.** (2015b) Heat can erase epigenetic marks of vernalization in Arabidopsis. Plant Signal Behav **10:** e990799.

Finally, Chapter S3 (page S49) provides additional technical information about the analysis of microRNAs.

While beginning my thesis, I wished to learn a diversified set of skills. I believe that I mostly achieved this goal, as I learned techniques ranging from plant physiology to bioinformatics and genetic engineering. This document thus deals with different aspects of plant biology, such as plant physiology, molecular biology, genetic engineering, transcriptomics, and database creation.

REFERENCES

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature, **408**, 796–815.

Black, R.E., Allen, L.H., Bhutta, Z.A., Caulfield, L.E., de Onis, M., Ezzati, M., Mathers, C. and Rivera, J. (2008) Maternal and child undernutrition: global and regional exposures and health consequences. The Lancet, **371**, 243–260.

Bouché, F., Lobet, G., Tocquin, P. and Périlleux, C. (2015a) FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. Nucleic Acids Res **doi**: 10.1093/nar/gkv1054

Bouché, F., Detry, N. and Périlleux, C. (2015b) Heat can erase epigenetic marks of vernalization in Arabidopsis. Plant Signal Behav **10:** e990799.

Craufurd, P.Q. and Wheeler, T.R. (2009) Climate change and the flowering time of annual crops. J Exp Bot **60**: 529–2539.

Dirzo, R. and Raven, P.H. (2003) Global State of Biodiversity and Loss. Annu Rev Environ Resour **28**: 137–167.

Landenmark, H.K.E., Forgan, D.H. and Cockell, C.S. (2015) An Estimate of the Total DNA in the Biosphere. PLoS Biol **13**: e1002168.

Paine, J.A., *et al.* (2005) Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nature Biotechnol, **23**, 482–487.

Périlleux, C., Pieltain, A., Jacquemin, G., Bouché, F., Detry, N., D'Aloia, M., Thiry, L., Aljochim, P., Mathieu, A.-S., Lutts, S., and Tocquin., P. (2013) A root chicory MADS box sequence and the Arabidopsis flowering repressor *FLC* share common features that suggest conserved function in vernalization and de-vernalization responses. Plant J **75:** 390–402.

Provart, N.J., *et al.* (2015) 50 years of Arabidopsis research: highlights and future directions. New Phytologist, **doi**: 10.1111/nph.13687.

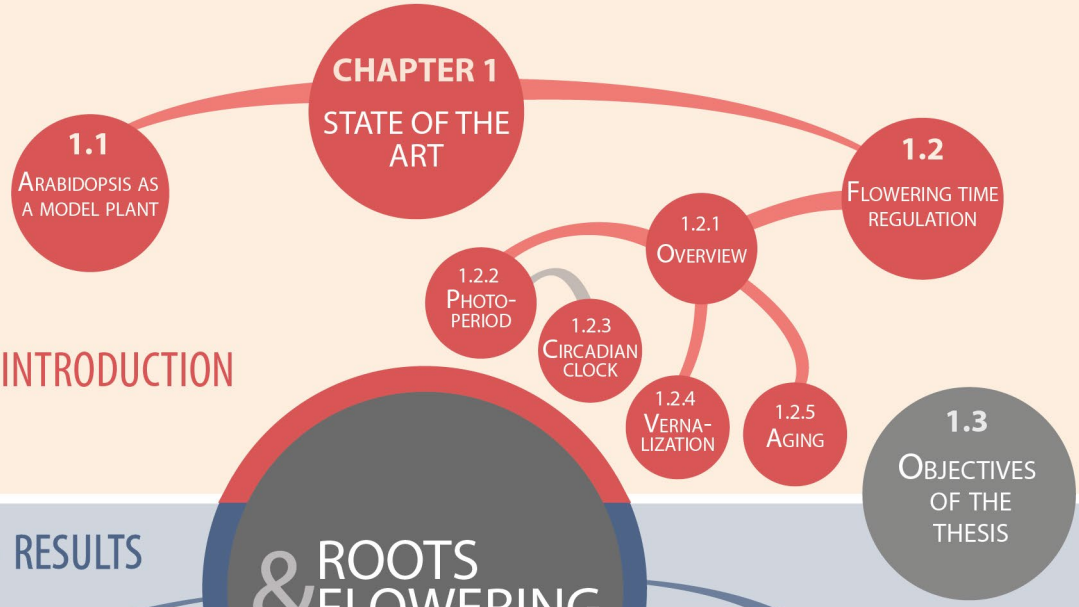
Regina, A., Bird, A., Topping, D., *et al.* (2006) High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. PNAS, **103**, 3546–3551.

Tilman, D., Balzer, C., Hill, J. and Befort, B.L. (2011) Global food demand and the sustainable intensification of agriculture. PNAS, **108**, 20260–20264.

Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H. and Tilman, D.G. (2008) Human Alteration of the Global Nitrogen Cycle: Sources and Consequences. Ecol Appl **7**: 737–750.

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science **287**: 303–305.

THESIS STRUCTURE



SECTION I • INTRODUCTION

SECTION II • RESULTS

ROOTS & FLOWERING

CHAPTER 2
FLOWERING DATABASE

HOW TO BUILD A COMPREHENSIVE LIST OF FLOWERING TIME-GENES?

CREATION OF A DATABASE OF FLOWERING-TIME GENES: **FLOR-ID.**

CHAPTER 3
DEVELOPMENT & GROWING MEDIUM

EFFECT OF THE GROWING MEDIUM ON PLANT DEVELOPMENT?

EFFECT OF THE GROWING MEDIUM ON THE ROOT TRANSCRIPTOME?

A ROLE FOR *FLC* IN THE SUBSTRATE-SPECIFIC CONTROL OF FLOWERING?

CHAPTER 4
ROOTING THE FLOWERING PROCESS

ARE FLOWERING-TIME GENES EXPRESSED IN ROOTS?

DOES THE PHOTOPERIODIC INDUCTION OF FLOWERING MODIFY THE ROOT TRANSCRIPTOME?

DO SELECTED MUTANTS DISPLAY ROOT ARCHITECTURE AND FLOWERING TIME MODIFICATIONS?

SECTION III • GENERAL DISCUSSION

SECTION IV • SUPPLEMENTARY MATERIALS

- CHAPTER S1 — *FLC* ORTHOLOG IN ROOT CHICORY
- CHAPTER S2 — DEVERNALIZATION IN ARABIDOPSIS
- CHAPTER S3 — TECHNICAL NOTE ABOUT MICRORNAS ANALYSIS
- CHAPTER S4 — SUPPLEMENTARY MATERIALS