Exploring the role of temperature regulated alternative splicing in flowering time and morphogenesis

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Giovanna Capovilla

aus Padova, Italien

Tübingen

2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	9 Oktober 2017
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Markus Schmid
2. Berichterstatter:	Prof. Dr. Claudia Oecking

ACKNOWLEDGMENTS

I would like to thank,

For the opportunity of this PhD adventure, his constant encouragement, his faith in my abilities and his extraordinary patience,
Markus Schmid.

For his advice over the years, for being part of my PhD committee, for allowing me to work in his lab and escape from my fate in the far north,
Detlef Weigel.

For agreeing to be my second supervisor, for evaluating the thesis, being member of my PhD and Thesis committee,
Claudia Oecking.

For agreeing to be part of my Thesis committee,
Sascha Laubinger and Klaus Harter.

For introducing me to the idea of pursuing a PhD in Germany, encouragement and continuous support,
Silvio Collani.

For invaluable help by replying to my questions and helping me in my projects or in the lab,
Rebecca Schwab, Silvio Collani, Effie Symeonidi, Nicolas Delhomme, Nacho Rubio, Hülya Wicher, Frank Küttner, Christa Lanz, Ilja Bezrukov, David Posé, Patricia Lang, Jathish Ponnu, Emanuele Scacchi, Wangsheng Zhu, Rui Wu, Anette Habring, Lisa Smith, Iryna Shutava, Sonja Kersten and Josip Perkovic.

For the accuracy in translating the Abstract in German that Google Translate doesn't reach,
Patricia Lang.

For being my "guinea pig"- first student I have ever had - and for still wanting to be a scientist, with my relief and peace of mind,
Karin Poersch.

For my fun life in Tübingen, inside and outside the lab, for their friendship and encouragements,

Effie Symeonidi, Kavita Venkataramani, Patricia Lang, Cris Zaidem, Daniela Lazzaretti, Reza Shahidi, Ezgi Dogan, Diep Tran, Emanuele Scacchi, Silvio Collani, Paolo Panza, Leily Rabbani, Fabio Zanini, Wei Yuan, Talia Karasov, Sergio Latorre, Jorge Kageyama, Simon Perathoner, Hernan Burbano, Noemi Skorzinski, Cristina Barragan, Subhashini Muralidharan, Gautam Shirsekar, Nacho Rubio, Rui Wu, Moises Exposito, Beatriz Antunez, Lorenzo Ducci, Dino Jolic, Ulrich Lutz, Fernando Rabanal, The Alice in WonderLAB team and the Climbers' gang.

For their support and friendship from Italy while I was abroad,

Andrea Ghedina, Elena Barbera, Maria Capodiferro, Chiara Broccanello, Enrico Nguyen and Sabrina Giaretta.

For the fun times, for their friendship and encouragements from across the ocean during the thesis writing,

Joe Campanale, Greg Russo, Daisy Muralles, Aurora Rodriguez and Payam Rowghanian.

For always being there for me,

Vale, Effie, Nina, Ale.

For their unconditional support for each and every thing that I do, and for their love,

My parents, my sister, my granny Ada, and my family.

For distributing joy and happiness,

My nieces Emma and Giulia.

For being my sunshine,
Ale.

TABLE OF CONTENTS

ABSTRACT

The regulation of gene expression is a complex, multifaceted molecular process that controls development, growth and environmental adaptation. At its core, gene expression is regulated at the level of transcription, however, post-transcriptional mechanisms such as alternative pre-mRNA splicing contribute to its fine-tuning. While initially believed to be rare in plants, an increasing number of examples have demonstrated the importance of alternative splicing in diverse processes, including the context of flowering time control.

Plant reproductive success depends in part on the correct onset of flowering. Because of its importance, flowering is tightly regulated by several antagonistic floral inductive and repressive pathways. Plants integrate endogenous and exogenous signals to assess the overall environmental conditions and choose accordingly the best time to flower. Among the environmental signals perceived by plants, ambient temperature has received limited attention. One of the processes by which temperature can affect gene expression is through alternative splicing during mRNA maturation. The thermoregulation of alternative splicing and its contribution to flowering time regulation and morphogenesis in A*rabidopsis thaliana* are the focus of my dissertation.

In Chapter 2 I sought to clarify the role of *FLM,* a MADS domain transcription factor involved in flowering time regulation, which is characterized by temperature-dependent alternative splicing. *FLM* gives rise to two major isoforms whose role in the control of flowering

has been discussed controversially. To investigate the specific contribution of these two isoforms I employed CRISPR/Cas9 gene editing to introduce targeted deletions in the *FLM* locus. These novel mutant lines lack exons specific for each main isoform, allowing a systematic analysis of their contribution to flowering time regulation. The results reported here support a central role for FLM-β, the most abundant isoform, in repressing the transition to flowering. Moreover, I show that FLM-δ, which has been proposed as a regulator of flowering induction in recent *in vitro* studies, does not promote flowering *in vivo*, likely because its endogenous expression levels do not reach the required activation threshold.

Plants are highly responsive to temperature fluctuations and are capable of modulating organogenesis and growth rate to adapt to novel conditions. In chapter 3 I describe the function of a novel, *bona fi*de alternative splicing regulator that is essential for correct development and morphogenesis of *Arabidopsis thaliana* at low temperature. Loss of function mutations of this gene, which I named *PORCUPINE* (*PCP*), displayed severe meristem defects when grown at 16°C, whereas at 23°C no obvious phenotype was detectable. At any developmental stage, by solely changing the temperature, plant growth can be arrested or reinitiated. This behavior indicates the presence of a mechanism that allows rapid adaptation of growth and development to abrupt changes in the ambient temperature. The meristem defects detected in the mutants can be largely explained by the misregulation of two genes involved in maintaining the stem cell fate in the shoot apical meristem, *WUSCHEL* and *CLAVATA3*. These findings support the importance of temperature-dependent alternative splicing in plant morphogenesis and establish *PCP* as an important regulator of environmental adaptation.

ZUSAMMENFASSUNG

Die Regulation der Genexpression ist ein komplexer und vielseitiger molekularer Prozess, der die Entwicklung, das Wachstum und die Umweltanpassung kontrolliert. Im Grunde wird Genexpression auf Ebene der Transkription reguliert, jedoch tragen posttranskriptionelle Mechanismen wie alternatives Pre-mRNA-Spleißen zu ihrer Feinabstimmung bei. Während alternatives Spleißen in Pflanzen zunächst als selten galt, bestätigt mittlerweile eine wachsende Zahl an Beispielen die Wichtigkeit alternativen Spleißens in vielfältigen Prozessen, einschließlich im Zusammenhang mit der Blühzeitkontrolle.

Der pflanzliche Fortpflanzungserfolg hängt zum Teil vom korrekten Beginn der Blüte ab. Wegen ihrer Bedeutung ist die Blüte durch mehrere antagonistische florale induktive und repressive Signalwege streng reguliert. Pflanzen integrieren endogene und exogene Signale, um die Gesamtumgebungsbedingungen einzuschätzen und entsprechend den besten Zeitpunkt zum Blühen zu wählen. Unter den von Pflanzen wahrgenommenen Umgebungssignalen wurde der Umgebungstemperatur bisher nur geringe Aufmerksamkeit entgegengebracht. Einer der Prozesse, durch welche die Temperatur die Genexpression beeinflussen kann, ist alternatives Spleißen während der mRNA-Reifung. Die Thermoregulation des alternativen Spleißens und ihr Beitrag zur Blütezeitregulation und Morphogenese in *Arabidopsis thaliana* stehen im Mittelpunkt meiner Doktorarbeit.

Kapitel 2 beschäftigt sich mit der Klärung der Rolle von *FLM*, einem MADS-Domänen-Transkriptionsfaktor, der an der Blütezeitregulierung beteiligt ist und den temperaturabhängiges alternatives Spleißen charakterisiert. *FLM* bildet zwei Haupt-Isoformen, deren Rolle in der Blühzeitkontrolle kontrovers diskutiert wurde. Um den spezifischen Beitrag dieser beiden Isoformen zu untersuchen, verwendete ich CRISPR / Cas9-Geneditierung, um gezielte Deletionen im *FLM*-Locus einzuführen. Diesen neuartigen Mutantenlinien fehlen die für jede Haupt-Isoform spezifischen Exons, so dass eine systematische Analyse ihres jeweiligen Beitrags zur Blütezeitregulierung möglich ist. Die hier berichteten Ergebnisse unterstützen eine zentrale Rolle für FLM-β, die häufigste Isoform, bei der Unterdrückung des Übergangs zur Blüte. Darüber hinaus zeige ich, dass FLM-δ, das in neueren *in vitro* Studien als Regulator der Blüteninduktion vorgeschlagen wurde, die Blüte *in vivo* nicht begünstigt, wahrscheinlich, weil seine endogenen Expressionsniveaus nicht die erforderliche Aktivierungsschwelle erreichen.

Pflanzen reagieren sensibel auf Temperaturschwankungen und sind in der Lage, Organogenese und Wachstumsrate zu modulieren, um sich an neue Bedingungen anzupassen. In Kapitel 3 beschreibe ich die Funktion eines neuartigen, *bona fide* alternativen Spleißreglers, der für die korrekte Entwicklung und Morphogenese von *Arabidopsis thaliana* bei niedriger Temperatur unentbehrlich ist. Nullmutanten dieses Gens, das ich *PORCUPINE* (*PCP*) genannt habe, zeigen bei 16°C schwere Meristemdefekte, während bei 23°C kein offensichtlicher Phänotyp nachweisbar war. In jedem Entwicklungsstadium kann durch bloßes Ändern der Temperatur das Pflanzenwachstum angehalten oder wiederhergestellt werden. Dieses Verhalten zeigt das Vorhandensein eines Mechanismus an, der eine rasche Anpassung von Wachstum und Entwicklung an abrupte

Änderungen der Umgebungstemperatur ermöglicht. Die in den Mutanten nachgewiesenen Meristemdefekte können weitgehend durch die Fehlausrichtung von zwei Genen erklärt werden, die an der Aufrechterhaltung des Stammzellschicksals im apikalen Meristem des Sprosses beteiligt sind, nämlich *WUSCHEL* und *CLAVATA*. Diese Erkenntnisse unterstützen die Bedeutung des temperaturabhängigen alternativen Spleißens in der Pflanzenmorphogenese und etablieren *PCP* als wichtigen Regulator von Umweltanpassung.

CHAPTER 1

The mechanisms regulating expression of flowering time related genes and the importance of temperature

1.1 INTRODUCTION

1.1.1 THE FLOW OF GENETIC INFORMATION

During the last century, the discovery of deoxyribonucleic acid (DNA) as the hereditary material containing genetic information has been an outstanding achievement for the scientific community. Scientists concentrated their subsequent efforts on understanding how genes are able to direct synthesis of proteins.

From the observation that protein synthesis in eukaryotic cells occurs in the cytoplasm (Prescott, 1964), and not in the nucleus where the chromosomal DNA is located, it was clear that the DNA double helix itself could not serve as a template for protein synthesis.

Instead, another molecule chemically very similar to DNA found very abundant in the cytoplasm, the RNA, quickly moved center stage. RNA is a long and unbranched molecule that contains 4 types of nucleotides linked by phosphodiester bonds like in the DNA. RNA differs from DNA for the sugar component (ribose in place of deoxyribose) and the replacement of thymine with the closely related pyrimidine uracil. Unlike DNA, the RNA is typically found in the cells as a single-stranded molecule that often bends on itself to form loops and 3D structures and serves as template for protein syntheses (Watson *et al.* 2015).

The pathway for the flow of genetic information, often seen as the central dogma of molecular genetics and biology, was first

hypothesized in 1956 and then established in 1970 by Francis Crick. The central dogma of molecular biology describes how from a DNA template, in a process called transcription, a ribonucleic acid (RNA) molecule is synthesized and serves as a template for the protein synthesis, or translation. At its core this central dogma is still valid today, at least for protein-coding genes (Watson *et al.* 2015).

The next big question was how the information contained in the genome is expressed, and how its expression is regulated. In simple organisms like bacteria, a very large fraction of the DNA is used to encode proteins. In contrast, in more complex organisms such as many plants and animals, the portion of non-coding regions increases and only a relatively small portion of the genome is eventually translated into proteins (Cooper 2000). It is generally assumed/accepted that the non-coding regions of these more complex genomes harbor different kinds of regulatory elements that are important to regulate gene expression.

In multicellular organisms the regulation of gene expression is extremely important as all the cells contain the same genetic information, However, this information needs to be expressed in different combination at different times or tissues to facilitate growth and development and to allow the organisms to react to changes in their environment (Watson *et al.* 2015).
Since the central dogma has been postulated, we have gained a much deep understanding about the biology of living organisms and the molecular mechanisms used to regulate gene expression.

In this chapter, I will introduce known regulatory molecular mechanisms, particularly at the RNA level, that occur in flowering time pathways in the model plant *Arabidopsis thaliana*. During the

switch from vegetative to flowering, major physiological and developmental changes occur in plants. The induction of flowering is therefore controlled by a large number of genes. These genes form a complex regulatory network that enables plants to respond to changes in both endogenous and exogenous signals to finely regulate the developmental switch from vegetative to flower state (Bouchè *et al.* 2016).

1.1.2 THE MODEL PLANT *ARABIDOPSIS THALIANA*

Arabidopsis thaliana is a member of the mustard family that for several reasons represents a good model to study plant genetics, development, and physiology (Somerville and Koornneef 2002). Both, the plant itself and its genome, are relatively small. Each plant produces in a rapid cycle a large number of seeds, primarily by self-fertilization that can be easily stored. *Arabidopsis thaliana* is native to Eurasia (Sharbel *et al.* 2000) and more than a thousand natural accessions have been collected from natural population in many geographic areas, where they have (presumably) adapted to the local climate. Full genomic sequence data are also available for many of these accessions, facilitating correlative studies between specific phenotypes and sequence polymorphisms (genome wide association studies) or *vice versa*.

Arabidopsis thaliana is also easy to transform by infiltration with disarmed *Agrobacterium tumefaciens* strains using the so-called “flower dip” technique, and crosses between different lines are also possible and relatively easy. For all these reasons, and the rich literature about flowering time genes and pathways, *Arabidopsis thaliana* is the model organism used for all the experiments in this thesis.

1.1.3 REGULATION OF FLOWERING TIME

The timing of flowering is extremely important to ensure plants reproductive success. During the transition from vegetative to reproductive state plants undergo a major physiological change, and flowering plants become more susceptible to environmental changes. For this reason the transition must be tightly regulated, and plants have evolved intricate genetic networks to perceive both endogenous and environmental signals. The integrating signals eventually result in flowering when overall conditions are at the best to allow progeny to develop.

The genetically defined flowering time pathways in *Arabidopsis thaliana* have been well characterized and much of the information available on individual genes and their interactions has been collected in the Flowering Interactive Database, FLOR-ID (Bouchè *et al.* 2016) (Figure 1.1). The regulation of flowering time in *Arabidopsis thaliana* has also been covered in a number of excellent reviews (Srikanth and Schmid 2011, Andrès and Coupland 2012, Song *et al.* 2015). I will therefore only briefly summarize the most important aspects of flowering time regulation and otherwise refer to these reviews for more details.

An environmental factor used by many (but not all) plant species to regulate flowering time are seasonal changes in day length (photoperiod). *Arabidopsis thaliana* is a so called facultative long day (LD) plant, meaning that LD promotes flowering but plants will ultimately flower even under otherwise non-inductive short day (SD) conditions. Interestingly, photoperiod is perceived not at the shoot apical meristems (SAM), where flowers are induced, but in the phloem companion cells in the leaf vasculature. Here, inductive LD results in the stable expression of the CONSTANS (CO) protein,

which promotes transcription of *FLOWERING LOCUS T* (*FT*) (Kobayashi *et al.* 1999). The FT protein acts as a florigen (long distance signal) that travels through the vasculature to the SAM where it interacts with the bZIP transcription factor FD to induce flowering (Liu *et al.* 2013; Giakountis and Coupland 2008) (see 1.2.1 for details on the role of transcriptional regulation in the regulation of flowering).

Temperature is also a major environmental factor regulating flowering time. Two aspects of the temperature-controlled regulation of flowering can be distinguished: 1) the vernalization pathway, which controls the regulation of flowering in response to exposure to a prolonged period of cold, i.e. overwintering, and 2) ambient temperature, which regulates flowering in response to fluctuation of temperatures in the physiological range (Capovilla *et al*. 2014). Interestingly, both vernalization and ambient temperature have been shown to modulate *FT* expression, marking this gene as an important flowering time integrator of environmental cues.

Endogenous cues such as hormones, especially gibberellic acid, carbohydrates, and plant age also play important roles in the regulation of flowering and are perceived in both leaves as well as at the SAM. Several of these signals have been shown to converge on *SUPPRESSOR OF CONSTANS OVEREXPRESSION 1* (*SOC1*), which is another important flowering time integrator in *Arabidopsis thaliana*, particularly at the SAM (Lee and Lee 2010).

Last but not least the so-called autonomous pathway and general processes like chromatin modifications, mircroRNA, protein stability and control of transcription have also been extensively studied and many of its components functions are described in the following paragraphs (Spanudakis and Jackson 2014).

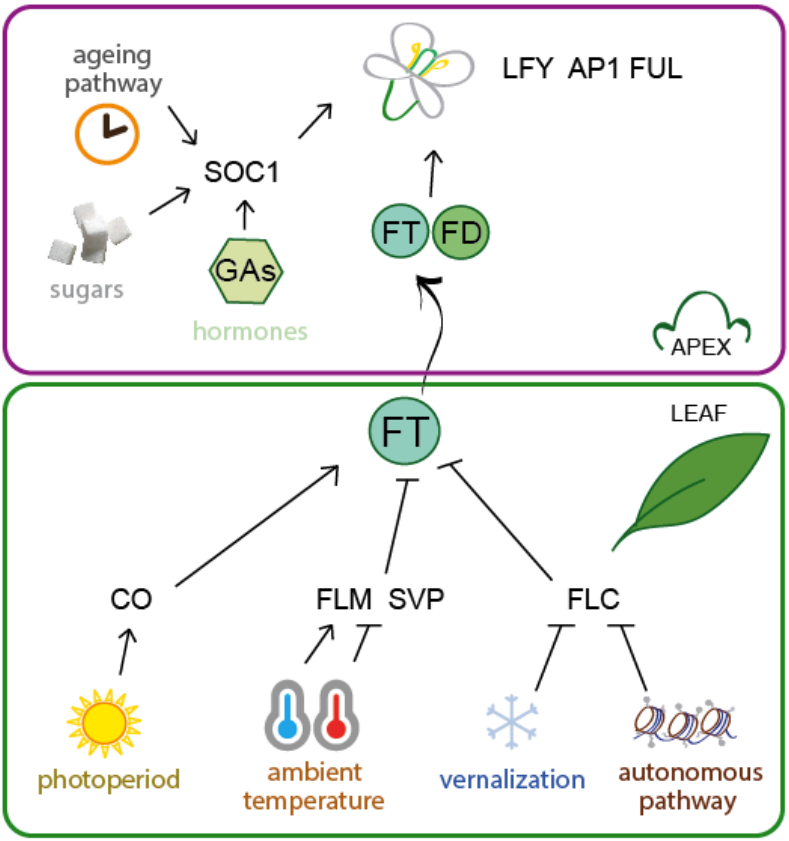

Figure 1.1: Simplified schematic representation of the contribution of the different pathways in regulating flowering time. Each pathway contributes to regulate the flower integrators SOC1, FT and FD, which eventually trigger the transition to flowering.

1.2 MOLECULAR MECHANISMS OF EXPRESSION REGULATION IN FLOWERING TIME PATHWAYS

Because of its importance for plant growth and development, flowering time is regulated at multiple levels. However, genes encode factors regulating chromatin features, transcription factors, transcriptional co-regulators, of RNA processing proteins appear to be overrepresented among the "classic" flowering time genes, suggesting that the regulation of gene expression plays a pivotal role in this process. In the following I will briefly introduce the mechanisms that regulate flowering at different levels, using selected flowering time genes as examples, finishing with the regulation at the RNA level, which is central to this thesis.

1.2.1 REGULATION AT THE CHROMATIN LEVEL

At the most basic level, expression of genes is regulated at the chromatin level, which modulates the access of regulatory proteins such as transcription factors and RNA polymerases to DNA. The basic unit of DNA chromatin packaging is the nucleosome, which consists of DNA wrapped around a histone octamer protein complex. These nucleosomes are the building blocks for more complex DNA structures. Genes located in such highly compact heterochromatic regions are in fact rarely transcribed in comparison with genes located into more open euchromatic regions (Watson *et al.* 2015).
The state of chromatin, heterochromatic vs. euchromatic, is not static but is adjusted to ensure genes are expressed at the right time in a tissue- and cell-specific manner. Central to this dynamic nature of chromatin packaging are modifications at specific amino acid residues of histone proteins (Bannister and Kouzarides 2011). By recruiting histone modifiers the nucleosomes can be modified in ways

that alter the accessibility to one or more genes to the transcriptional machinery and other regulatory proteins and therefore either inhibit or facilitate transcription. Common examples of chromatin modifications are achieved by removing acetyl groups from the histones tails, normally carried out by histone deacetylases, or by adding methyl groups to them in a process called methylation. Methylations mainly occur on the side chains of lysines and arginines, and are carried out by histone lysine methyltransferase (HKMT), or arginine methyltransferase (Bannister and Kouzarides 2011).

In the flowering time field, the regulation of the *FLOWERING LOCUS C* (*FLC*) gene provides a great example of regulatory mechanisms at the chromatin level in the context of vernalization. In addition, *FLC* expression is also regulated by long non-coding RNAs (lncRNA) as discussed later in the chapter.

As briefly mentioned above (1.1.3), vernalization is the process by which prolonged exposure to cold temperatures allows plants to flower the next spring. Certain plant species, including the winter annual accession of *Arabidopsis thaliana,* need to experience a period of several weeks of non-freezing cold to overcome a block to flowering and gain the ability to flower when temperature rises in spring (Chouard 1960, Lang 1965). FLC is the primary determinant of the vernalization requirement in *Arabidopsis thaliana* (Bastow *et al.* 2004, Michaels and Amasino 1999, Sheldon *et al.* 2009, Sung and Amasino 2004).

Before exposure to cold, *FLC* sense mRNA levels are very high, and the flowering repression activity of the FLC protein prevents the plant from undergoing the flower transition prematurely. High levels of *FLC* are ensured in this phase by FRIGIDA (FRI) and by several players involved in the (di- or tri-)methylation on histone 3 (H3) lysine 4 (H3K4) or on H3K36, chromatin modifications commonly associated

with actively transcribed euchromatin (Kouzarides 2007, Wagner and Carpenter 2012). Methylation on H3K4 is carried out by the ATWDR5a protein (Jiang *et al.* 2009), and the Arabidopsis TRITHORAX-RELATED7 (ATXR7), ATX1, and ATX2 proteins (Tamada *et al*. 2009). In contrast, methylation of H3K36 is mainly regulated by SET DOMAIN GROUP 8 (SDG8) (Xu *et al.* 2008).

During the exposure to cold, *FLC* expression is epigenetically silenced and this silencing is stably maintained even when plants are exposed to warmer spring temperatures after winter. Epigenetic silencing of *FLC* is directed by plant homeodomain (PHD) proteins such as VERNALIZATION INSENSITIVE3 (VIN3) (Sung and Amasino 2004, Wood *et al.* 2006), VERNALIZATION5 (VRN5) and VIN3-LIKE 2 (VIL2) (Greb *et al.* 2007, Sung *et al*. 2006, De Lucia *et al.* 2008), that interact with the Polycomb repressive complex 2 (PRC2) (De Lucia *et al.*, 2008). The activity of the PHD–PRC2 complex is responsible for increased trimethylation of histone H3 lysine 27 (H3K27me3) in the nucleation region of the *FLC* locus in a direct correlation with the length of the cold exposure resulting in progressive accumulation of methylation in H3K27 (Angel *et al.* 2011) and eventually in the permanent repression of *FLC* in vernalized vegetative plants.

1.2.2 TRANCRIPTION AND TRANSCRIPTION FACTORS

One outcome of the regulation of DNA accessibility by chromatin modifiers is that in a given cell, only certain regions of the genome are exposed and can be recognized by DNA binding proteins or more likely being transcribed.

Transcription is the process that expresses the genetic material by the production of RNAs, and proceeds through 3 phases called initiation, elongation and termination. The process that catalyzes RNA

synthesis is carried out principally by the RNA polymerases that can initiate the transcription from a DNA template.

During transcription, certain parts of the genome are selected and from one to even thousands of copies are made out of them. The choice of the regions to transcribe, and how extensively, is not random but finely regulated by specific DNA sequences. These include the core promoter needed for the binding to the Pol II machinery, and other promoter proximal elements, enhancers, silencers and other regulatory sequences that can bind regulatory proteins (Watson 2015).

An important group of proteins that directly interacts with short (often palindromic) DNA motives contained within the regulatory regions of their target genes are transcription factors (TFs).

In the *Arabidopsis thaliana* genome over 2000 TFs are encoded (Jin *et al.* 2013), many of which contribute to developmental programs, including the control of flowering time. Depending on their nature TFs can either activate or prevent expression of bound genes and are consequently referred to as activators or repressors. Often DNA binding TFs do not by themselves regulate transcription of nearby genes by rather include transcriptional co-regulators that provide the activating or repressive activity.

One important transcriptional repressor involved in the regulation of flowering is FLC, which has already been introduced as the main target of the vernalization pathway (see 1.2.1). The *FLC* gene encodes a MADS domain transcription factor that actively represses flowering in non-vernalized winter-annual accessions of *Arabidopsis thaliana*. *FLC* is predominantly expressed in the shoot or root apices and in the vasculature, and regulates flowering at least in part through repression of the floral integrator genes *FT* and *SOC1* (Searle *et al.* 2006). Several other MADS domain proteins closely related to FLC,

MADS AFFECTING FLOWERING 1 (MAF1; also known as FLOWERING LOCUS M (FLM)) and MAF2 – MAF5, regulate other aspects of flowering. FLC and several of the MAF proteins have been shown to interact with a more distantly related MADS protein, *SHORT VEGETATIVE PHASE (SVP)*, to form active transcriptional repressor complexes (Ratcliffe *et al.* 2003, Gu *et al.* 2013, Lee *et al.* 2013, Posè *et al.* 2013).
FLM is the main focus of chapter 2, where its role and regulatory mechanism are extensively discussed. *FLM* main isoform is *FLM-*β, which encode for a protein that physically interacts with SVP. The complex SVP-FLM-β is more abundant at low ambient temperature, and prevents flowering by binding to downstream target genes, like *SEP3* and *SOC1* (Posè *et al.* 2013).

Once FLC has been epigenetically silenced, floral promoters such as the zinc finger transcription factor CONSTANS (CO) and the bZIP TF FD can activate flowering. *CO* mRNA is regulated by the circadian clock, a biochemical oscillator entrained by the day-night cycle. Several factors concur in CO regulation, which ultimately enables plants to discriminate between non-inductive short days (SD), under which the CO protein is rapidly degraded, from long day (LD) conditions, that allow CO protein to accumulate towards the end of the day (Suàrez-Lòpez *et al.* 2001, Valverde *et al.* 2004, Laubinger *et al.* 2006). Several lines of evidence indicated that transcriptional activation of the principal florigen *FLOWERING LOCUS T (FT)* under inductive day length conditions in leaves is directly mediated by CO (Samach *et al.* 2000). However the demonstration of direct binding of DNA by CO has only been recently obtained by Gnesutta and colleagues (2017). The authors showed that CO needs to interact with two CCAAT-binding transcription factors, NF-YB2 and NF-YC3, in a trimer called NF-CO. This trimeric complex can efficiently bind

the CORE element of the *FT* promoter and promotes its transcription under LD (Figure 1.2).

The FT protein, functions as a florigen and moves from the leaf to the apex where it interacts with the meristem specific bZIP transcription factor FD and 14-3-3 adapter proteins to form a flower activation complex (FAC) (Abe *et al*. 2005, Wigge *et al.* 2005). This complex triggers flowering through the activation of key floral meristem identity genes like *APETALA 1 (AP1)* and *SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1)* (Abe *et al.* 2005, Wigge *et al.* 2005, Taoka *et al.* 2011).

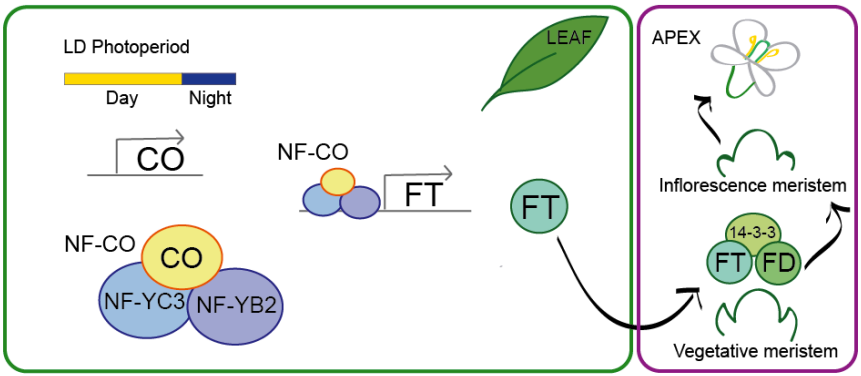

Figure 1.2: Schematic representation of the contribution of CONSTANS (CO) in promoting flowering. In long day conditions (LD) *CO* expression is stimulated by players belonging to the photoperiod pathway. CO physically interact with NF-YC3 and NF-YB2 to form the NF-CO complex which act as a transcription factor and stimulates the expression of the florigen *FT*. FT moves from the leaves to the apex where by binding with its partners FD and 14-3-3 induces the transition to flowering by activating downstream target genes such as *AP1* and *SOC1*. These genes promote the transition of the shoot meristem from vegetative to floral meristem.

1.2.3 REGULATIONS ON RNA LEVEL

1.2.3.1 TYPES OF RNA

There exist several different types of RNA in plants, all of which are the product of transcription. Three of them are required for protein synthesis: the messenger RNA (mRNA), which carries the information from the DNA to site of protein synthesis, the ribosomes; the transfer RNAs (tRNA), that fold in a well defined 3D conformation and serve as adaptors to deliver the amino acids to ribosomes; and finally the most abundant RNA type, the ribosomal RNAs (rRNAs), which are principal components of the ribosomes themselves (Watson *et al*. 2015).
Besides these three extremely well characterized types of RNA, several other types of regulatory RNAs have been discovered more recently.

Of the latter, the so-called CRISPR RNA is of high relevance for this thesis. In prokaryotes CRISPR-Cas forms an adaptive and heritable immune system. DNA from invading phages is incorporated into Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), forming a cluster of characteristic repeating DNA sequences interspaced with phage DNA. This clusters are expressed and processed to give rise to crRNAs that each consist of a one palindromic repeat, plus a short variable sequence obtained from phage DNA. crRNAs hybridize with another non-coding RNA, the so-called tracrRNA, and the RNA hybrid directs an endonuclease, Cas9, to destroy foreign phage DNA that infects the cell. The crRNA and tracrRNA have been engineered into a single molecule, the so called single guide RNA (sgRNA) that (within certain limits) can been reprogrammed to target specific DNA sequences, thereby facilitating

genome engineering in a wide range of eukaryotic species, including *Arabidopsis thaliana* (Mali *et al*., 2013; Belhaj *et al.* 2013). An example of the use of the CRISPR-Cas9 system is reported in chapter 2 of this thesis.

1.2.3.2 RNA SPLICING

The coding sequence of a protein-coding gene is a series of three-nucleotide codons that specifies the linear sequence of amino acids incorporated in the polypeptide product. However, in the vast majority of eukaryotic genes the coding sequences, called exons, are not contiguous but periodically interrupted by stretches of non-coding sequences, or introns (Watson 2015, Sharp 1994). The result is a mosaic of blocks of coding exons and noncoding introns that are transcribed in a primary RNA transcript (pre-mRNA). To create the mature mRNA that can eventually be translated into proteins, the primary transcripts must have their introns removed and the exons joined together. The process of intron removal is called RNA splicing and must occur with great precision to avoid the loss, or addition, of even a single nucleotide to avoid frame shifts in the mRNAs (Sharp 2005). Introns and exons are defined by specific sequences in the pre-mRNA. These sequences are called the 3' and 5' splice sites, denoting their relative locations at each end of the introns, and a third element, the branch site, found within the introns. The most highly conserved sequences found within the regulatory sequences are the GU in the 5' splice site, the AG in the 3' splice site and an A at the branch site, but the rest of the sequences vary. The splicing process is a sequence of two successive transesterification reactions in which phosphodiester bonds within the pre-mRNA are broken and new ones are formed. At first, an A in the branch site attacks a G in the 5' splice site, forming a lariat. In the second reaction, the liberated 3' exon

attacks the 3' splice site. The result of these reactions is the fusion between two exons and the release of the intron folded in a lariat shape structure (reviewed in Kornblihtt *et al.,* 2013). The described splicing process is carried out by a large complex of proteins and RNAs called the spliceosome, made up of 5 RNA-protein complexes called U1, U2, U4, U5, and U6 snRNPs. Each of these comprises an RNA molecule, called the U1 to U6 snRNA, respectively, and a number of proteins (Sharp 2005). The spliceosome is a very dynamic structure, it assembles on the pre-mRNA at different steps recognizing introns and catalyzing their removal. Each subunit joins and leaves the complex, performing a particular function. After catalyzing the splicing reaction, the spliceosome is disassembled, and the spliced mRNA released. By removing from the original RNA all of the introns, intron-containing genes may give rise to a unique mRNA species. But in many cases, splicing can produce several different mRNAs from the same gene by selecting alternative splicing sites and therefore splicing the original pre-mRNA in different patterns. This process is known as alternative splicing (Syed *et al.* 2012).

1.2.3.3. ALTERNATIVE RNA SPLICING

Alternative splicing (AS) of RNA is the primary source for eukaryotes to expand the functional proteome (Nilsen *et al.* 2010). Splice variants can be divided into the following categories: retained introns (RIs), cassette exons (or exon skipping), mutually exclusive exons, competing 5' splice sites, and competing 3' splice sites (Syed *et al.*, 2012) (Figure 1.3). In plants, many RIs were shown to introduce premature termination codons or frame shifts. More interestingly are maybe cassette exons, which in principle can increase the functional

diversity of proteins, especially if the protein domains encoded display distinct functional or structural features.

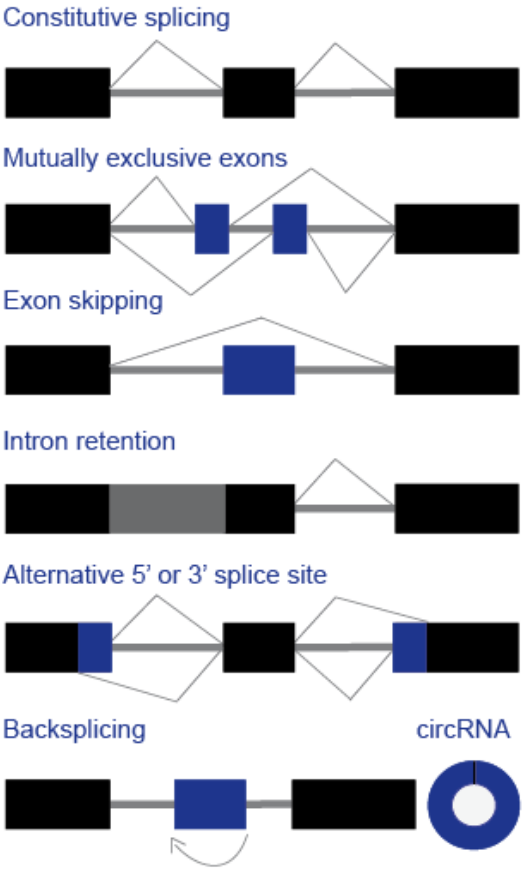

Figure 1.3: Schematic representation of the possible types of alternative splicing and the backsplicing, whose products are the circRNAs discussed in chapter 1.2.3.5.3

As described above, only very short sequences flanking the exons are highly conserved. The challenge for the splicing machinery to recognize and splice at correct sites is facilitated by other proteins, mainly belonging to the SR family or the hnRNPs, which are able to recognize specific sequences in the pre-mRNA and target the spliceosome accordingly. More details about these players, alternative splicing factors, and the occurrence of AS in response to environmental stimuli are presented in chapter 3 of this thesis.

Alternative splicing occurs essentially in all eukaryotes and can have severe developmental consequences. For example, in *Drosophila* sex-determination is controlled by a cascade of splicing events, which determines whether the fly develops as a male or female. A key exon skipping occurs in *Sex-lethal (Sxl)* in response to X chromosome number. In females the inclusion of an alternative 'poison exon' containing a stop codon is repressed, allowing the expression of full-length proteins. On the contrary, in males the poison exon is retained in the mRNA and a truncated Sxl protein is produced (Bell *et al*. 1991). Remarkably, another gene downstream of *Sxl, fruitless (fru)* is also differentially spliced in males and females in order to express a specific isoform of Fru only in males. This gene is in fact necessary for male courtship behavior, and a mis-splicing is sufficient to generate male behavior in females (Demir *et al*. 2005).

In plants alternative splicing has been implicated in the regulation and/or fine-tuning of the circadian rhythm (Seo *et al.* 2012, Wang *et al.* 2012, Jones *et al.* 2012, Marshall *et al.* 2016), seed maturation (Sugliani *et al.* 2010), abiotic stress response (Liu *et al*. 2013, Wang et al. 2015*a*), biotic stress response (Xu *et al*. 2012) or splicing itself (Zhang *et al*. 2009). Several flowering time genes in *Arabidopsis* are also regulated by AS. Two examples are the closely related MADS box transcription factors *MAF2* and *FLM*, which are involved in the temperature pathway. MAF2 is a flowering repressor that can interact

with SVP to bind target genes. By increasing the ambient temperature *MAF2* preferentially retains an intron, that introduces a premature termination codon and results in an inactive isoform (Rosloski *et al.* 2012, Airoldi *et al*. 2015). A second example is given by *FLM*, whose role in flowering time regulation and splicing patterns are extensively discussed in chapter 2.

1.2.3.4 NMD

In a process known as translation the information contained within the nucleotide sequence of the mRNA is decoded into the linear sequence of amino acids of the polypeptide chain, resulting in the synthesis of proteins. Eukaryotic cells rely on a mechanism called Nonsense-Mediated Decay (NMD) that during translation proofreads the mRNAs being translated and monitors their integrity. Transcripts containing a premature stop codon (or nonsense codon) are recognized and eventually degraded by NMD. The recognition of mRNAs with premature stop codons relies on protein complexes that are assembled on the mRNA just upstream of each exon–exon boundary as a consequence of splicing. When the ribosome translates the mRNA, these complexes are displaced as the mRNA enters the decoding center of the ribosome. However, if a premature stop codon is present in the mRNA, for example due to intron retention, the ribosome is released before the displacement of all the exon–junction complexes, and a set of RNA degrading proteins is recruited. These proteins cleave the mRNA, remove the cap at its 5' end and also its poly-A tail, leaving the mRNA unprotected and target of rapid degradation by exonucleases (Reviewed in Lykke-Andersen and Jensen 2015). In *Arabidopsis thaliana* AS coupled by NMD is a key regulatory process that regulates *FLM* expression in different

ambient temperatures (Sureshkumar *et al.* 2016), as discussed in chapter 2.

1.2.3.5 NON CANONICAL RNAs

1.2.3.5.1 LONG NONCODING RNAs

A class of non-coding RNAs longer than 200 nucleotides is known as long non-coding RNAs (lncRNAs). They can serve many roles in processes from translation and splicing to transcriptional regulation (Zhu and Wang, 2012) Even though research in this area in plants is far behind that in animals (Zhang *et al.* 2013, Bai *et al.* 2015), they have been characterized in the flowering time pathway and, more specifically in the regulation of *FLC*. As already described above *FLC* undergoes a progressive trimethylation on H3K27 that results in the epigenetic silencing of the gene when temperatures rises after vernalization. However during the exposure to cold three lncRNAs play a role in down regulating *FLC: COOLAIR, COLDAIR* and *COLDWRAP* (Figure 1.4).

Transcription of *FLC* decreases relatively rapidly in response to exposure to cold and the non-coding antisense transcripts called *COOLAIR* has been implicated in this process (Swiezewski *et al.* 2009). Expression of *COOLAIR* is upregulated very quickly even after a short exposure to cold from a transcription start site immediately downstream of the poly-A site of the *FLC* sense transcript (Marquardt *et al.* 2014). A second cold-induced non-coding transcript from the sense strand of *FLC, COLDAIR*, accumulates later than *COOLAIR*, reaching its maximum levels after 3 weeks of exposure to cold (Heo and Sung 2011). *COLDAIR* is non-polyadenylated, non-spliced and is produced from intron 1. It has been found to interact with the histone methyltransferase subunit (CLF) (Heo and Sung 2011). Together with

the polycomb-binding lncRNA *COLDWRAP*, which is derived from the repressed promoter of *FLC*, *COLDAIR* has been suggested to coordinate the vernalization-mediated silencing of *FLC* (Kim and Sung 2017).

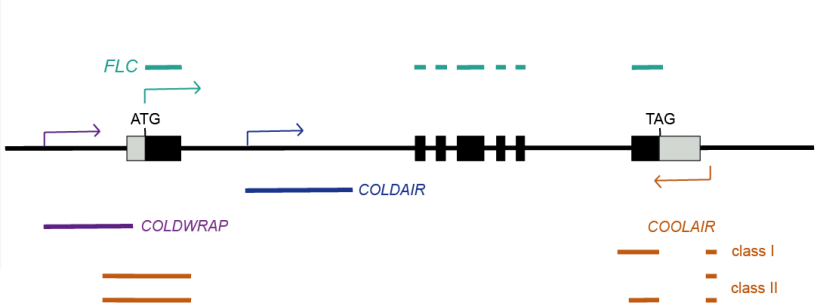

Figure 1.4: Schematic representation of the *FLC* intron-exon structure and transcript details. The full length *FLC* sense transcript is represented in green. The group of *COOLAIR* lncRNA expressed in antisense is represented in dark orange, and the sense lncRNAs *COLDAIR* and *COLDWRAP* are marked in blue and purple, respectively. The transcription starting sites represented as arrows whose colors match the respective transcripts.

1.2.3.5.2 miRNAs

microRNAs (miRNA) are a group of small regulatory RNAs. miRNAs are short RNA molecules ranging from 18 to 24 nucleotides in length. They have been shown to regulate gene expression in a number of plant developmental processes including flowering time, serving either to inhibit or to promote it.

Plant miRNAs are encoded by MIR genes that are transcribed by RNA polymerase II. The primary transcript (pri-miRNA) is processed by DICER-LIKE 1 (DCL1) into a precursor miRNA (pre-miRNA) and

then subsequently into a miRNA/miRNA* duplex. The duplex consists of the guide strand miRNA and a passenger strand (miRNA*) (reviewed in Naqvi *et al*. 2012). The duplex is methylated and transported from the nucleus to the cytoplasm where the ARGONAUTE 1 (AGO1) protein or its paralogue AGO10 bind to the dissociated guide strand and together with accessories proteins form the miRNA-induced silencing complex (RISC) (Naqvi *et al.* 2012). The RISC can eventually inhibit expression of target genes in three ways: by destruction of the mRNA encoded by the target gene, by inhibiting translation of the mRNA, or by inducing chromatin modifications within the target gene and thereby silence its transcription.

Examples of miRNAs involved in flowering time regulation the miR156 and miR172 families and their respective targets, the *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* and AP2-like transcription factors, that contribute to the regulation of both the juvenile and adult phase transition in *Arabidopsis thaliana* (Wu *et al.* 2009, Huijser and Schmid 2011, Yamaguchi and Abe 2012). Several genes of the *miR156* family are highly expressed in the embryo and in young seedlings, but the abundance of mature miR156 declines with increasing plant age. *miR156* targets 10 of the 16 known *SPL* transcription factors, down-regulating their expression by transcript cleavage (Park *et al.* 2005, Franco-Zorrilla *et al.* 2007, Huijser and Schmid 2011, Yamaguchi and Abe 2012). The age-dependent decrease in *miR156* levels is in fact accompanied by a concomitant increase in *SPL* expression (Wu and Poethig 2006; Yamaguchi and Abe 2012). *SPL3, SPL4* and *SPL5*, all of which are targeted by miR156, have been shown to induce expressions of floral promoters like *LEAFY (LFY),* and *AP1*, and indirectly, the transition to flowering (Wu *et al.* 2009). However, more recent findings performed in loss of function lines suggest a role for *SEP3, SEP4* and *SEP5* in promoting floral meristem identity rather than flowering induction (Xu *et al.*

2016). A second group of target *SPLs (SPL9* and *SPL15)* instead, promotes the transition to flowering by inducing the expression of *miR172* (Wu *et al.* 2009, Zhu and Helliwell 2011) known to have the opposite effect of *miR156. SPL9* shows an effect only in LD whereas *SPL15* loss of function delays flowering regardless the photoperiod (Hyun *et al.* 2016). Target of *miR172* are transcripts of six *APETALA 2-like* floral repressors (*AP2, TARGET OF EAT 1 (TOE1), TOE2, TOE3, SCHLAFMÜTZE (SMZ)*, and *SCHNARCHZAPFEN (SNZ)* (Aukerman and Sakai, 2003; Chen, 2004; Yamaguchi and Abe, 2012) that and are repressed by *miR172* by translational inhibition or transcript cleavage (Aukerman and Sakai 2003, Schwab *et al.* 2006, Fang and Spector 2007). This temporally opposite expression of miR156 and miR172 has been suggested to regulate flowering time in regards to the plant age.

1.2.3.5.3 circRNAs

Circular RNAs (circRNAs) are a class of hyper-stable single-stranded and covalently-closed RNA molecules that arise through a form of alternative splicing called backsplicing. Typically, circRNAs are formed by circularization of exonic sequences present in the pre-messenger RNA that undergo exon skipping and are therefore not included in the mRNA (Ye *et al.* 2015).

They have been identified in all eukaryotic kingdoms of life (Wang and Wang 2015), but only recently their regulatory role in *Arabidopsis thaliana* has started to be uncovered. Conn and colleagues (2017) reported that circRNAs can be used as *bona fide* biomarkers of functional, exon-skipped AS variants in *Arabidopsis thaliana*, finding evidence for circRNA formation in MADS-box TF superfamily genes like *SEPALLATA3 (SEP3), SEEDSTICK (STK),* and the flowering

time genes *FLM* and *FLC*. Furthermore, the authors showed a mechanism of action for a circRNA derived by the circularization of the 6th exon of *SEP3*. This circRNA can strongly bind to its cognate DNA locus, forming an RNA:DNA hybrid, which may physically slow transcriptional elongation and thus favor the biogenesis of the exon-skipped AS variant.

1.3 THE IMPORTANCE OF AMBIENT TEMPERATURE

Changes in ambient temperatures have been found to affect living organisms of all kingdoms. For example, among bacteria, the food-borne pathogen *Listeria monocytogenes* offers a strikingly example of the importance of ambient temperature-dependent regulation. *L. monocytogenes* can survive and replicate in a wide range of temperatures and causes severe infections that primarily affect immunocompromised individuals and pregnant women (reviewed in Cossart and Lecuit, 1998). The large temperature-shifts that exist between the outside environment and the host correlates with the synthesis of its virulence proteins. *Listeria* virulence genes are in fact maximally expressed at 37°C and almost silent at 30°C (Johannson *et al.* 2002) . Their expression is controlled by PrfA, a thermoregulated transcriptional activator. Prfa untranslated mRNA folds in a secondary structure masking the ribosome binding region at lower temperature. At high temperature the folding conformation changes and the ribosome binding site is recognized (Johannson *et al.* 2002).

Ambient temperature differences also play a significant role in development and morphology of plants and animals. For example, small changes in temperature play a key role in determining the sex of a number of reptiles. In turtles and alligators there is no genetic predisposition for the embryos to develop as either male or female.

The sex-determining mechanisms are temperature-dependent and therefore at the mercy of the ambient conditions affecting egg clutches in nests. The exposure to environmental temperature during a critical period of the embryonic development determines whether an egg develops as male or female (Reviewed in Crain and Guillette 1998). For example, in many turtle species, such as the European pond turtle, *Emys obicularis*, eggs from cooler nests, below 25°C, hatch as all males. On the contrary eggs at temperatures above 30°C produces all females broods. The threshold temperature at which the sex ratio is even was calculated as 28.5°C (Pieau *et al.* 1994).

In plants, growth rate, leaf morphology and also flowering time are regulated by ambient temperature. Plants are sessile organisms, therefore when temperatures are high, but not so high as to induce a heat stress response, a number of morphological changes occur that are likely to contribute to adaptive growth acclimation. In *Arabidopsis thaliana*, seedlings elongate the hypocotyl, rosette leaves are fewer, smaller and thinner, carry fewer stomata and show hyponastic growth (reviewed in Quint *et al.* 2015). At high ambient temperature plants also exhibit a greater transpiration rates, suggesting that thermomorphogenic adaptations may contribute to high temperature mitigation by enhancing leaf evaporative cooling (Quint *et al.* 2015).

All these phenotypes are thought to assist cooling. The bHLH transcription factor PHYTOCHROME INTERACTING FACTOR 4 (PIF4) has been identified as a key regulator of thermomorphogenic phenotypes, including hyponasty, hypocotyl and petiole elongation (Proveniers and Zanten 2013). PIF4 integrates various environmental signals and induces transcriptional changes on its target genes that subsequently trigger primarily phytohormone-induced elongation responses (reviewed in Quint *et al.* 2015).

The responsiveness of flowering to thermal induction varies between plant species and also within accessions. In *Arabidopsis thaliana*,

importance of ambient temperature in flowering time is highlighted by the fact that a moderate temperature increase from 23°C to 27°C is sufficient to induce flowering under an otherwise non-inductive short-day (SD) photoperiod (Balasubramanian *et al.*, 2006).

The optimal temperature for *Arabidopsis thaliana* growth is 22°C to 23°C (The Arabidopsis Biological Resource Center 2015) and its development speed decreases at lower temperatures. However, the mechanisms of temperature perception in plants are only partially understood but a number of thermoresponsive factors have been described (reviewed in Capovilla *et al.* 2014).

Thermosensitive chromatin remodeling was the first to be implicated in temperature-dependent regulation of flowering time in *Arabidopsis thaliana*. The histone variant H2A.Z is incorporated into nucleosomes by a chromatin-remodeling complex that includes ACTIN-RELATED PROTEIN6 (ARP6) and PHOTOPERIOD-INSENSITIVE EARLY FLOWERING1 (PIE1). This makes DNA less accessible for transcription factors and consequently limits gene expression (Talbert and Henikoff 2014). In response to increasing ambient temperatures, however, H2A.Z nucleosomes are evicted from the DNA, thereby enabling binding of PIF4 to the *FT* promoter, which was proposed to contribute to accelerate flowering (Kumar and Wigge 2010, Kumar *et al.* 2012), even though genetic analyses support only a minor role for *PIF* genes in adjusting flowering time, and that GA biosynthesis might play a more central role (Galvao *et al.* 2015).

Other transcription factors besides PIF4 have also been implicated in the temperature-dependent control of flowering time. MADS-box genes related to *FLC* such as *SVP*, *FLM* and other *MAFs* (*MAF2*, *MAF3* and *MAF4*) have been shown to play a central role in thermoresponsive flowering (Ratcliffe *et al*. 2003, Gu *et al*. 2013, Capovilla *et al.* 2014). At the heart of this regulatory module is SVP, which forms repressor complexes with other MADS box transcription

factors that repress *FT* and *SOC1* transcription at low temperatures but decline in abundance at higher temperatures, relieving repression (Lee *et al.* 2013, Posè *et al*. 2013). The effects of ambient temperature in regulating alternative splicing are exhaustively discussed in both chapter 2 and chapter 3.

1.4 CONCLUSIONS AND THESIS PURPOSE

Alternative splicing is a key mechanism that allows rapid changes in transcript expression in fluctuating ambient temperature. In this thesis, I present two case studies that provide important insights into how AS and temperature can affect two important aspects of plant development in *Arabidopsis thaliana*, flowering time and morphogenesis.

In chapter 2 I describe my contribution to better understand the role of AS of *FLOWERING LOCUS M (FLM)* to the regulation of temperature-dependent flowering. *FLM* was already known to produce alternative splicing variants whose expression change with temperature. The main isoforms were known to regulate flowering time in an opposite fashion. I managed to introduce targeted deletions in *gFLM* using CRISPR technology and delete specific exons required for the correct expression of the two major isoforms. The aim of the project was to investigate the contribution of each isoform in a context as close as possible to the WT.

In chapter 3 I present a *bona fide* AS factor, whose function was previously unknown, that has a dramatic temperature dependent effect on plant morphology. I named this gene *PORCUPINE (PCP)*, after the peculiar phenotype that the knock out mutant shows at low temperature.

CHAPTER 2

Role of *FLM* in temperature-dependent flowering

2.1 INTRODUCTION

2.1.1 THE CONTRIBUTION OF TEMPERATURE IN FLOWERING TIME REGULATION

The correct timing of the transition from vegetative growth to flowering is a crucial decision in the life cycle of plants, and the survival of the species in part depends on it. It is therefore not surprising that plants constantly monitor endogenous and environmental signals such as photoperiod and temperature by an intricate genetic network (reviewed in Srikanth and Schmid 2011) to adjust their flowering time.
Two aspects can be distinguished regarding the contribution of temperature to the regulation of flowering time: the response to prolonged periods of cold (overwintering, vernalization) and the effects of ambient temperature. In *Arabidopsis thaliana* the vernalization response has been extensively explored and the epigenetic mechanism silencing the effect of the flowering repressor *FLC* have been described in chapter 1 (and are reviewed in Sheldon *et al.*, 2009 and in Song *et al.*, 2012).
The importance of ambient temperature is highlighted by the finding that, in short day photoperiod (non-inductive) a moderate temperature increase (from 23°C to 27°C) triggers flowering in *A. thaliana* (Balasubramanian *et al.*, 2006).
Several MADS-domain transcription factors have been implicated in the thermosensory pathway, including *SHORT VEGETATIVE PHASE (SVP)*, *FLC* and other member of its clade such as *FLOWERING LOCUS M (FLM or MADS AFFECTING FLOWERING 1 (MAF1)),*

MAF2, MAF3 and *MAF4.* In rapid-cycling accessions like Col-0 the contribution of *FLC* is quite limited (Blasquez 2003, Lee 2013), whereas *SVP* and *FLM* assume a more substantial regulatory role in this pathway. SVP is a potent floral repressor, its loss of function mutant is in fact early flowering and partially temperature-insensitive (Lee et al. 2007). *SVP* is negatively regulated by FCA and FVE (Lee et al. 2007) and the SVP protein can physically interact with FLM, with which it shares many downstream target genes (Posè et al. 2013).

In this Chapter I first review briefly the current knowledge about the role of FLM in temperature-dependent flowering time and its mechanism of action. Then I discuss how my work contributes to a deeper understanding of the role of specific isoform of *FLM* in this process. Finally, based on an analysis of different natural accessions performed as part of this work and also taking into account previously published data, I discuss the role of *FLM* and its relevance in natural strains of *A. thaliana*.

2.1.2 *FLOWERING LOCUS M (FLM)*

The focus of this Chapter is *FLOWERING LOCUS M (FLM,* also called *MADS AFFECTING FLOWERING 1, MAF1)*, which encodes a MADS-box transcription factor related to FLC and is one of the genes involved in the thermosensory pathway in *A. thaliana*.
By the analysis of loss of function mutants and over-expressing lines Scortecci and colleagues (2001) attributed to *FLM* a role as floral repressor. *flm* mutants flowered earlier than WT in both inductive and non-inductive photoperiods, whereas constitutive expression of *FLM* caused delayed flowering.

The FLM transcript is regulated by alternative splicing, a common mechanism in eukaryotic cells described in chapter 1. Through the selection of alternative splice sites different mature mRNAs are produced from the *FLM* primary transcript. Initially 4 alternative splicing (AS) variants, *FLM-α, FLM-β, FLM-γ* and *FLM-δ*, with alternative $2^{nd}/3^{rd}$ and $8^{th}/9^{th}$ cassette exons were described in the Wassilewskjia (Ws) accession (Scortecci *et al*., 2001) as potentially protein coding transcripts (Figure 2.1). In Col-0 however, only two of these 4 isoforms (*FLM-β* and *FLM-δ),* which differ in the inclusion of the mutually exclusive 2^{nd} and 3^{rd} exon were found (Posé *et al*., 2013).

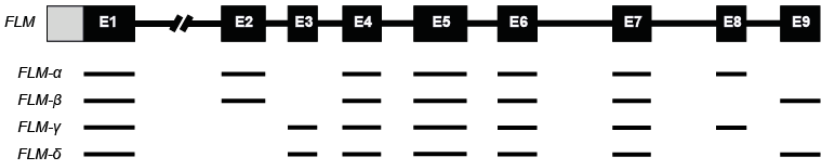

Figure 2.1: Schematic representation of the 4 potentially protein coding AS variants described in Scortecci *et al.* 2001 in Ws background.

FLM has been shown to be alternatively spliced in response to ambient temperature and the main isoform appeared reduced after thermal shift (Balasubramanian *et al.* 2006). The identity of the main active isoform, as well at the role of a larger splice form that appeared after thermal induction described by Balasubramanian and colleagues was not known at that time (2006). However, the authors proposed a mechanism in which temperature-dependent AS overcame the repressive effect of *FLM*.

In 2013 this hypothesis was proven correct by the combined work of Posè and colleagues and Lee and colleagues (2013) and *FLM-β* was identified as the main isoform. *FLM-β* is expressed higher at low ambient temperatures and functions as a flower repressor. In fact it physically interacts with SVP and controls downstream targets genes like *SEP3* and *SOC1* (Posè *et al*., 2013; Lee *et al*., 2013). A second protein coding isoform, *FLM-δ* has been shown to either not respond to changes in temperature (Sureshkumar *et al.* 2016; Lutz *et al.* 2015; Lutz *et al*. 2017) or to be induced by elevated temperatures (Posé *et al*., 2013). Furthermore, in agreement with the original results (Balasubramanian *et al*. 2006), additional splice variants with combinations of intron retention and/or exon skipping, have been identified in Col-0, particularly at elevated ambient temperature (Sureshkumar *et al*., 2016 and this study).

In the past 4 years two hypotheses have been proposed to explain how the AS of *FLM* might control flowering time in *A. thaliana*. The first model is based on evidences provided by Lee *et al*., 2013 and Posé *et al*., 2013. The authors showed that overexpression of *FLM-δ* in both the Col-0 and *flm-3* backgrounds leads to premature flowering. Furthermore, by an elegant electrophoretic mobility shift assay (EMSA) experiment the authors show that FLM-β needs SVP for binding DNA, that FLM-δ can compete with FLM-β for binding with partners like SVP, and that the SVP-FLM-δ complex is not able to bind to and repress target genes. They therefore suggested that *FLM-δ* might act as a dominant negative version of FLM that indirectly promotes the transition to flowering. The ratio between the two *FLM* isoforms would allow to fine-tune flowering time.
Alternatively, Sureshkumar and colleagues (2016) suggested that increasing ambient temperature induces flowering only by reducing

the expression of the repressor *FLM-β*. The mechanism proposed is based on the detection of temperature sensitive splicing sites in *FLM* and an increase of non-canonical isoforms that mostly contain premature stop codons other than *FLM-δ* at high temperature. The authors proved that nonsense-mediated decay (Lykke-Andersen and Jensen 2015) plays an important role in degrading the non-canonical isoforms.

In both models the relative abundance of the floral repressor *FLM-β* decreases at elevated temperatures but the contribution of *FLM-δ* in the regulation of flowering is more controversial. In this Chapter I present how the results obtained during this study contribute to a deeper understanding of the role of the *FLM* isoforms in flowering time.

2.1.3 CRISPR/Cas9

The CRISPR/Cas system has recently emerged as a powerful new approach for genome editing and an efficient alternative to Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) for inducing targeted genetic alterations in plants (Gaj 2013, Ma *et al.* 2016).

CRISPR stands for clustered regulatory interspaced short palindromic repeats and in bacteria and Achaea, CRISPR-Cas systems provide immunity against invading foreign DNA. Short fragments of invading phages are integrated into the CRISPR genomic locus, transcribed as pre-crRNA, and subsequently processed into short CRISPR RNA (crRNA). These crRNAs anneal to trans-activating crRNAs (tracrRNAs) and direct sequence-specific cleavage and silencing of pathogenic DNA by Cas proteins.

This mechanism has been conveniently engineered and represents now a tool to precisely edit DNA sequences in many organisms including *A. thaliana* (Mali *et al.*, 2013; Belhaj *et al.* 2013). Genome editing by CRISPR/Cas9 has three requirements: expression of the Cas9 protein, production of guide RNA (gRNA) that complements the DNA sequences of the target gene, and the presence of a conserved trinucleotide-containing protospacer adjacent motif (PAM) sequence upstream of the crRNA-binding region (Figure 2.2).

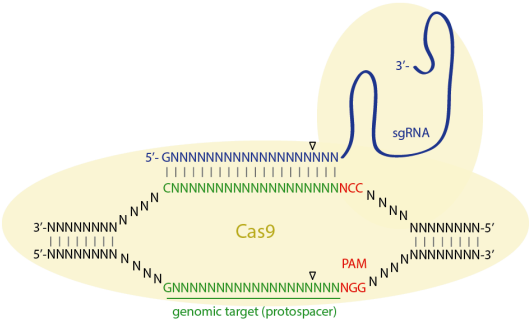

Figure 2.2: Schematic representation of the Cas9-sgRNA system cutting the recognized target sequence adjacent to the PAM sequence motif. (Figure adapted from Belhaj *et al*. 2013)

In this work by using simultaneously two sgRNA I introduced targeted mutations in the *FLM* locus with the goal to delete the whole sequence of exon 2 or of exon 3 in two separate lines.

2.2 A CRISPR/Cas9 APPROACH TO UNDERSTAND THE CONTRIBUTION TO FLOWERING OF THE MAIN *FLM* ISOFORMS

2.2.1 AIM

The aim of this chapter was to clarify the role of the two main isoforms of *FLM* in the Col-0 accession *in vivo* by generating mutant

lines with modified endogenous *FLM* genomic sequence (Figure 2.3). For this purpose CRISPR/Cas9 will be used to introduce specific mutations directed to delete the 2nd or the 3rd exon of *FLM* in two lines called FLM-ΔE2 and FLM-ΔE3, respectively. Because of the engineered *FLM* genomic sequence the two lines are expected to express only one of the two major splicing variants, *FLM-β* in FLM-ΔE3 and *FLM-δ* in FLM-ΔE2. By analyzing the flowering time phenotype of the CRISPRs lines in comparison with the WT, the loss of function mutant *flm-3* and the over-expressing lines it will be possible to draw conclusions regarding the contribution of specific *FLM* isoforms to the regulation of flowering time, especially the role of *FLM-δ*, whose role in this process is discussed controversially.

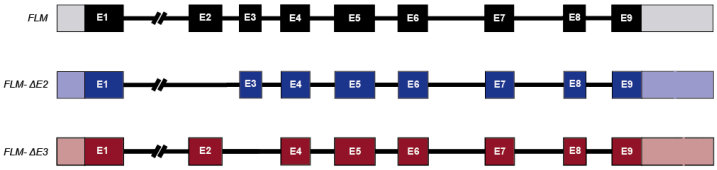

Figure 2.3: Schematic representation of the *FLM* locus in the WT (black) and the planned edited locus in the two engineered lines.

2.2.2 RESULTS

2.2.2.1 DELETION OF ISOFORM-SPECIFIC *FLM* CASSETTE EXONS BY CRISPR-CAS9

With the aim to understand the contribution of each of the two main isoforms, I generated two mutant lines by introducing specific mutation directed to delete exon 2 and exon 3, the two exons responsible for each major isoform.

To generate targeted mutations in *FLM* genomic sequence in Col-0 background using CRISPR-Cas9 system I combined the strategy reported in (Wang et al. 2015*b*) of employing *EGG CELL1* (*EC1.1*) promoter to express Cas9 in the egg cell (germ line) with the convenient *mCherry* based selection as reported in (Gao et al. 2016) to identify Cas9 containing seeds. The sgRNAs were designed in regions flanking the 2nd or 3rd exon, respectively, and expressed under the control of the *U6* promoter.

Using this approach lines lacking the 2nd or 3rd FLM exon were obtained (Figure 2.4)

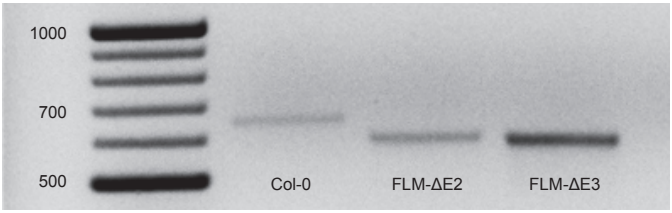

Figure 2.4: PCR amplification of the *FLM* genomic region spanning intron 1 to exon 4 shows the 57 bp and 64 bp deletions in *FLM*-Δ*E2* and *FLM*-Δ*E3*, respectively.

Sanger sequencing confirmed the presence of a 57 bp deletion that covers most of exon 2, normally incorporated in the repressive *FLM-*β splice variant, with the exception of 2 bp at the 5' end and the 2 first bp of intron 3 (Figure 2.3). This line is in the following referred to as FLM-ΔE2. In contrast, FLM-ΔE3 carries a 64 bp deletion that completely covers exon 3, normally found in *FLM-*δ*,* and a small portion on the flanking introns (Figure 2.5).

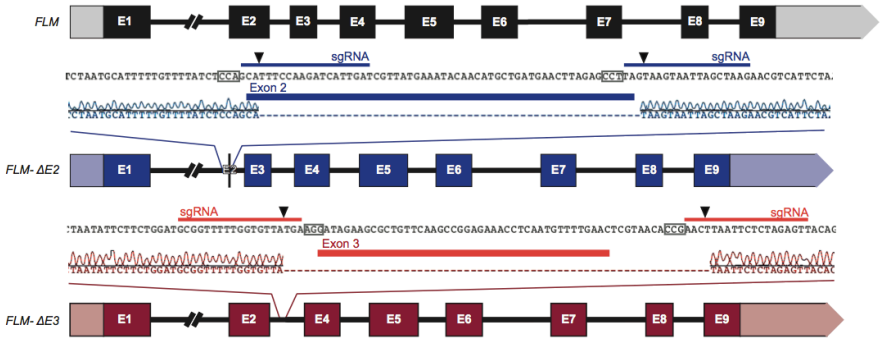

Figure 2.5: Schematic representation of the *FLM* locus in the two *FLM* CRISPR lines. The edited FLM-ΔE2 and FLM-ΔE3 lines are represented in blue and red, respectively. Close-ups provide detailed information on the position of the deletions as determined by Sanger sequencing of the CRISPR lines aligned with wild type *FLM* genomic sequences. The sgRNAs are represented by thin lines and the PAM sites are marked by grey boxes on the WT sequences.

Transgene-free homozygous mutant lines that had lost the Cas9-sgRNA T-DNA by segregation were obtained in T3 generation.

In both cases the Cas9 nuclease cut exactly where predicted, three bases distant from the PAM regions of the two sgRNAs with exception of an additional thymine deleted at the 3' sgRNA (transition between exon 3 and intron 3) in FLM-ΔE3.

To minimize the risk of artifacts in subsequent analyses due to off-target mutations I confirmed by Sanger sequencing that the sequences of the genes most closely related to FLM, *MAF2 to* 5, were fully intact and did not contain any deletions and/or point mutations.

2.2.2.2 DIVERSITY OF *FLM* SPLICE VARIANTS IN COL-0 AND CRISPR LINES

In order to investigate the effect of the CRISPR-induced deletions on *FLM* pre-mRNA splicing I amplified the *FLM* coding sequence by RT-PCR from Col-0 (control), FLM-ΔE3, and FLM-ΔE2 plants grown at 16°C, 23°C, and 27°C.

PCR products were cloned into plasmids and between 34 and 58 colonies for each line and temperature were analyzed by Sanger sequencing (Figure 2.6).

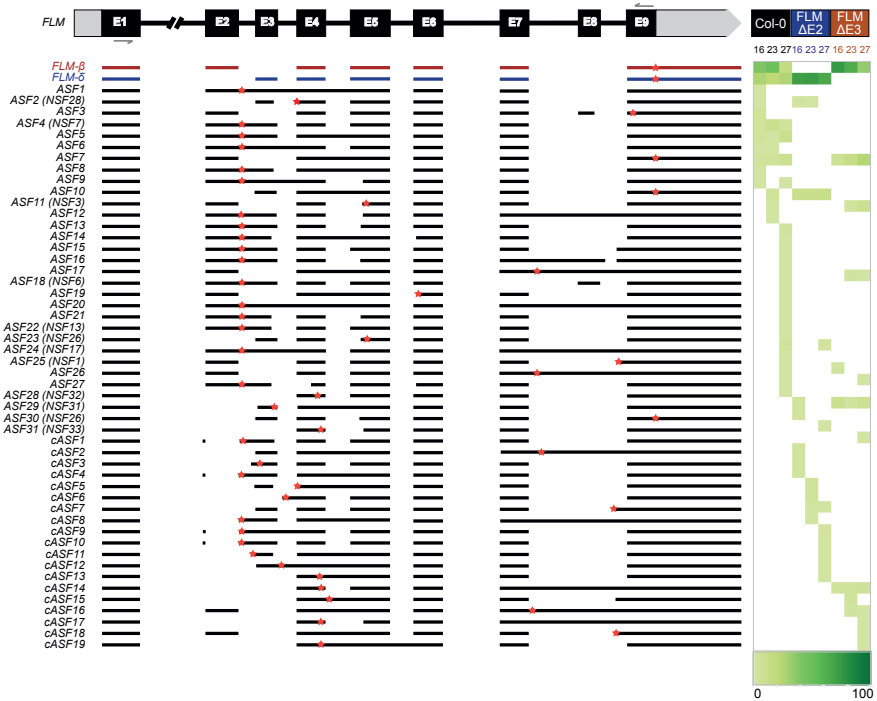

Figure 2.6: Alternative *FLM* splice variants detected by Sanger sequencing. The sequences present in the isoforms are aligned to the annotated *FLM* gene and grey arrows show the positions of the primers used to amplify the cDNA. Stop codons are represented as red stars. Isoforms identified in both Col-0 and at least in one of the CRISPR lines

are listed as alternative splice forms (ASF) 1 to 31. The identifiers of isoforms previously described by Sureshkumar and colleagues (Sureshkumar et al. 2016) are given in brackets. 19 new splice variants detected only in the CRISPR lines are listed as cASF. The heat map shows the frequency of each isoform in Col-0, FLM-ΔE2 and FLM-ΔE3 at 16°C, 23°C, and 27°C. The heat map legend shows a gradient of white to green where 0% of the sequences analyzed is white and 100% is dark green.

As expected from the literature (Posé et al. 2013), depending on the temperature *FLM-β* or *FLM-δ* were preferentially identified in Col-0, with *FLM-β* dominating at low temperature and *FLM-δ* becoming more abundant at 27°C (Figure 2.6 and Table 2.1). In addition, as previously reported by Sureshkumar and colleagues (2006) the frequency of non-canonical isoforms also increased in Col-0 from 30% at 16°C to over 50% at 27°C (Table 2.1).

In FLM-ΔE2 *FLM-δ* was the most abundant isoform (66-70%) at all three temperatures (Figure 2.6 and Table 2.1). Non-canonical splicing variants increased from 29% to 34% in total, however, if considered individually, each splice form accounted for only 2% to 12% of all cloned transcripts (Figure 2.6 and Table 2.1). As expected, *FLM-β* was never detected in this CRISPR line. In contrast, *FLM-β* was the predominant isoform in FLM-ΔE3 at each of the three temperatures, whereas *FLM-δ* could not be detected. By increasing the temperature, the frequency of *FLM-β* decreased from 74% at 16°C to a 35% at 27°C, while other isoforms increased from 25% to 64%.

In total, 31 alternative splice forms (ASF) of *FLM* were found in Col-0 and at least one of the mutated lines, while 19 isoforms (cASF1-19) were detected exclusively in the CRISPR lines (Figure 2.6). These results confirm that *FLM-β* and *FLM-δ* are the most abundant single *FLM* isoforms in Col-0. In addition, our findings demonstrate that plants with targeted deletion of either the 2nd or 3rd exon of *FLM* by CRIPSR/Cas9 predominantly produce either *FLM-β* (FLM-ΔE3) or *FLM-δ* (FLM-ΔE2).

Table 2.1. Percentage of *FLM* isoforms analysed in Col-0 and CRISPR lines at 16°C, 23°C and 27°C.

Isoform	Col-0 16°C	Col-0 23°C	Col-0 27°C	FLM-ΔE2 16°C	FLM-ΔE2 23°C	FLM-ΔE2 27°C	FLM-ΔE3 16°C	FLM-ΔE3 23°C	FLM-ΔE3 27°C
FLM-β	41.5	48.8	17.2	-	-	-	74.4	64.1	35.4
FLM-δ	26.8	22.0	25.9	70.6	72.7	66.0	-	-	-
ASF1	2.4	-	-	-	-	-	-	-	-
ASF2	2.4	-	-	2.9	2.3	-	-	-	-
ASF3	4.9	-	-	-	-	-	-	-	-
ASF4	4.9	9.8	5.2	-	-	-	-	-	-
ASF5	4.9	2.4	10.3	-	-	-	-	-	-
ASF6	2.4	2.4	-	-	-	-	-	-	-
ASF7	4.9	4.9	8.6	-	-	-	7.7	15.4	25.0
ASF8	2.4	-	-	-	-	-	-	-	-
ASF9	2.4	-	1.7	-	-	-	-	-	-
ASF10	-	4.9	-	11.8	11.4	12.8	-	-	-
ASF11	-	2.4	-	-	-	-	-	2.6	10.4
ASF12	-	2.4	-	-	-	-	-	-	-
ASF13	-	-	3.4	-	-	-	-	-	-
ASF14	-	-	1.7	-	-	-	-	-	-
ASF15	-	-	1.7	-	-	-	-	-	-
ASF16	-	-	1.7	-	-	-	-	-	-
ASF17	-	-	1.7	-	-	-	-	2.6	2.1
ASF18	-	-	1.7	-	-	-	-	-	-
ASF19	-	-	1.7	-	-	-	-	-	-
ASF20	-	-	1.7	-	-	-	-	-	-
ASF21	-	-	1.7	-	-	-	-	-	-
ASF22	-	-	1.7	-	-	-	-	-	-
ASF23	-	-	5.2	-	-	4.3	-	-	-
ASF24	-	-	1.7	-	-	-	-	-	-
ASF25	-	-	1.7	-	-	-	2.6	-	-
ASF26	-	-	1.7	-	-	-	-	-	2.1
ASF27	-	-	1.7	-	-	-	-	-	-
ASF28	-	-	-	2.9	-	-	10.3	2.6	8.3
ASF29	-	-	-	2.9	-	-	-	-	-
ASF30	-	-	-	-	-	2.1	-	-	-
ASF31	-	-	-	-	-	-	-	-	2.1
cASF1	-	-	-	2.9	-	-	-	-	-
cASF2	-	-	-	2.9	-	-	-	-	-
cASF3	-	-	-	2.9	-	-	-	-	-
cASF4	-	-	-	-	4.5	-	-	-	-
cASF5	-	-	-	-	4.5	-	-	-	-
cASF6	-	-	-	-	2.3	2.1	-	-	-
cASF7	-	-	-	-	2.3	-	-	-	-
cASF8	-	-	-	-	-	4.3	-	-	-
cASF9	-	-	-	-	-	2.1	-	-	-
cASF10	-	-	-	-	-	2.1	-	-	-
cASF11	-	-	-	-	-	2.1	-	-	-
cASF12	-	-	-	-	-	2.1	-	-	-
cASF13	-	-	-	-	-	-	5.1	5.1	6.3
cASF14	-	-	-	-	-	-	-	5.1	-
cASF15	-	-	-	-	-	-	-	2.6	2.1
cASF16	-	-	-	-	-	-	-	-	2.1
cASF17	-	-	-	-	-	-	-	-	2.1
cASF18	-	-	-	-	-	-	-	-	2.1
% Tot non canonical isoforms	31.7	29.3	56.9	29.4	27.3	34.0	25.6	35.9	64.6
Number of colonies	41	41	58	34	44	47	39	39	48

2.2.2.3 QUANTIFICATION OF MAJOR *FLM* ISOFORMS

To measure the relative expression of transcripts containing either the 2nd or 3rd exon (such as *FLM-β* and *FLM-δ*) and reliably quantify them I established a multicolor TaqMan assay.

Primers were designed in conserved regions in exon 1 and exon 4 of *FLM* to be able to use the same primer pair for the amplification of all splice variants of interest, thereby minimizing the risk of estimating certain isoform expression incorrectly because of differences in primer efficiency. The PCR products were detected in the same reaction by probes labeled with fluorophores with non-overlapping detection range, 6-FAM™ and HEX™2, placed on exons 2 or exon 3, respectively (Figure 2.7, Table 2.2). An extra set of primers and a probe labeled with CY5®, whose emission spectra is not overlapping with the other two fluorophores used in this assay, was used to detect expression of *UBC21*, which was used for normalization purposes (Figure 2.7, Table 2.2)

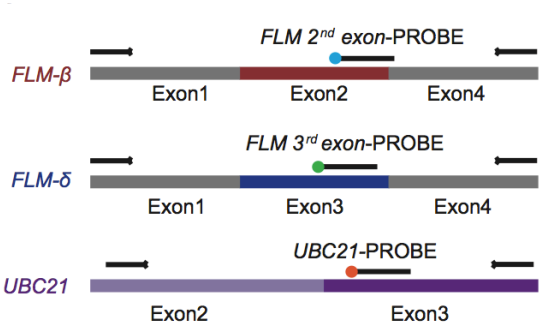

Figure 2.7: Schematic representation of the TaqMan assay designed to detect splice variants with the 2nd exon and 3rd exon of FLM (e.g. *FLM-β* and *FLM-δ*) and the normalization control *UBC21*. Primers are shown as black arrows, probes as black segments. Fluorophores are marked as blue (6-FAM™), green (HEX™2), and orange (CY5®).

Table 2.2. Oligonucleotides used for TaqMan assay.

Gene	Primer	Sequence (5' to 3')	Fluorophore	Detection (nm)	Quencher
FLM-β PROBE	G-36620	CCAAGATCATTGATCGTTATGAAATACAA	6-FAM™	515-530	BHQ-1
FLM-δ PROBE	G-36229	CGGAGAAACCTCAATGTTTTGAACTC	HEX™2	560-580	BHQ-1
FLM Forward	G-36231	CGCTGTTGTCGTCGTATCTG			
FLM Reverse	G-36743	CTAGTAACTCCTTGTGTGGAAG			
UBC21 PROBE	G-38938	GGAGTCCTGCTTGGACGCTTCAGTCTG	CY5®	675-690	BHQ-2
UBC21 Forward	G-38780	CTCCTCAAGTTCGATTCTTG			
UBC21 Reverse	G-38783	CCTGAGTCGCAGTTAAGAGG			

The multicolor TaqMan assay showed that the abundance of transcripts containing the 2nd exon (such as *FLM-β*) decreased in response to increasing temperature in Col-0 (Figure 2.8).

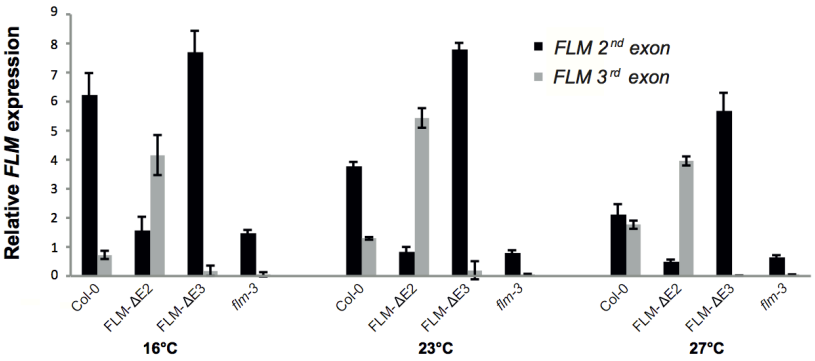

Figure 2.8: Relative expression of transcripts containing the 2nd (e.g. *FLM-β*) or the 3rd exon (e.g *FLM-δ*) measured using the TaqMan assay in Col-0, the CRISPR lines, and the *flm-3* mutant at 16°C, 23°C, and 27°C.

In contrast, the 3rd exon (such as *FLM-δ*) showed a moderate increase (Figure 2.8). A temperature-depended AS of *FLM* was observed also in a transgenic line that expresses the genomic region of *FLM* under the control of the constitutive 35S promoter (Figure 2.9). These findings confirm previous results (Posé et al. 2013) and demonstrate the functionality of the TaqMan assay.

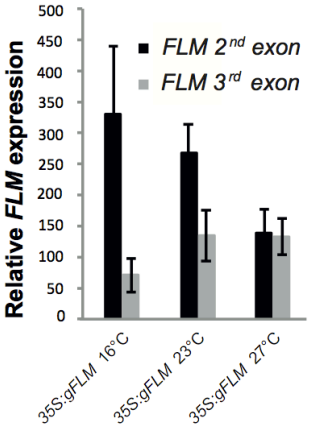

Figure 2.9: Quantification of splice variants with the 2nd exon and 3rd exon of *FLM* by TaqMan assay in a genomic *35S:gFLM* overexpression line at 16°C, 23°C, and 27°C.

A weak signal was detected in the *flm-3* loss of function mutant (Figure 2.8) using the probe against the 2nd exon of *FLM*. Sanger sequencing of the PCR product revealed that this signal originates from the amplification of a fragment of *MAF2*, a gene highly similar in sequence to *FLM*, and indicates that the 2nd exon probe was not absolutely gene specific.

Interestingly, I detected higher levels of exon 3-containing transcripts in FLM-ΔE2 than Col-0 at all temperatures (Figure 2.8). This trend was even more pronounced in the FLM-ΔE3 plants, which showed elevated exon 2 levels (Figure 2.8). Together these findings suggest that the basic expression of *FLM* remains constants regardless of the deletion of either exon 2 or exon 3 but affects the relative ratio of the splice variants.

2.2.2.4 QUANTIFICATION OF THE NON-CANONICAL *FLM* ISOFORMS *ASF7* AND *ASF10*

ASF7 and *ASF10* differ from *FLM-β* and *FLM-δ* respectively only for the retention of the 4th intron, which adds additional amino acids to the protein does not cause any shift in the reading frame downstream (Figure 2.6). To investigate the possible contribution of these two potentially protein-coding non-canonical isoforms to the regulation of flowering, I quantified their relative expression using the most specific primers possible (Fig. 2.10) and performed real Time PCR using SYBR Green technology. Similar to the TaqMan assay described above, *UBC21* was used for normalization purposes. In agreement with previous results, *FLM-β* was the most abundant isoform and its expression was decreasing with increasing temperature (Figure 2.10). Furthermore, *FLM-δ*, was slightly up regulated in warmer temperature and could be confirmed as second most abundant isoform. In contrast, the relative expression of *ASF7* and *ASF10* was very low when compared to *FLM-β* (Figure 2.10) suggesting a minor role of these two non-canonical isoforms to regulation of flowering in Col-0. By analyzing the expression levels, however, I cannot assign a role to the two proteins derived from the non-canonical isoforms, which in principle could encode for hyperactive proteins with a very strong effect. This scenario, albeit unlikely, cannot be excluded.

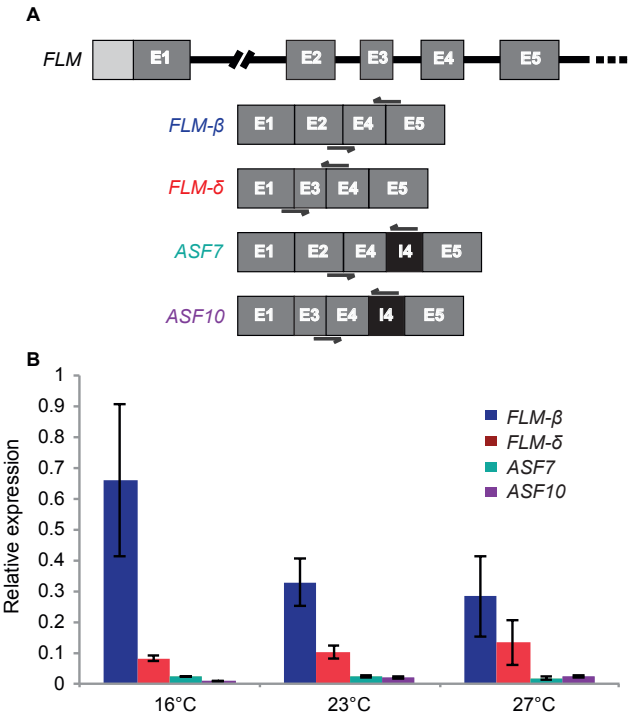

Figure 2.10: Quantification of FLM isoforms by RT-PCR. A. Representation of the relevant part of the *FLM* gene (from Exon 1 to Exon 6) and of the introns or exons included in the isoforms *FLM-β, FLM-δ, ASF7*, and *ASF10*. The primers used for the quantification of each isoform are represented as black arrows. B. Relative expression of each isoform at 16°C, 23°C, and 27°C. Error bars designate the standard deviation of 3 biological replicates.

2.2.2.5 FLOWERING TIME OF *FLM* DELETION MUTANTS

To investigate the contribution of the different isoform to the regulation of flowering I determined the flowering time of the newly generated deletion lines FLM-ΔE2 and FLM-ΔE3 in comparison with Col-0, the *flm-3* loss-of-function mutant and the *35S::gFLM* line at 16°C, 23°C, and 27°C (Figure 2.11).

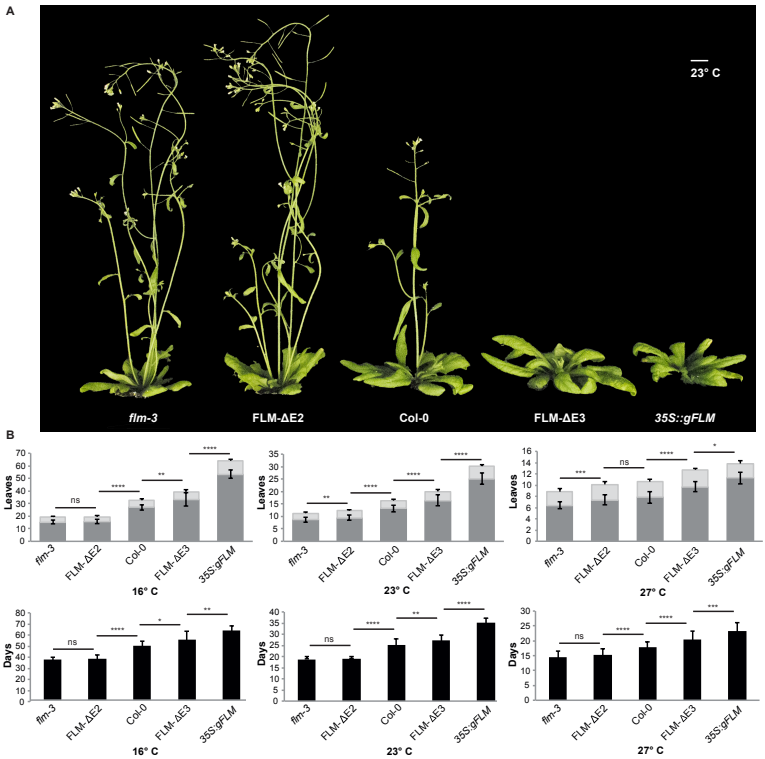

Figure 2.11: Flowering time of *flm* mutants. A. Representative pictures of *flm-3*, FLM-ΔE2, Col-0, FLM-ΔE3, and *35S::gFLM* lines grown at 23°C long days. Scale bar: 1 cm. B. Flowering time given as number of leaves (dark grey: rosette leaves; light grey: cauline leaves) and days to flowering (black) of plants grown at 16°C, 23°C, and 27°C. Error bars indicate standard deviation. *P < 0.05; **P < 0.01;***P < 0.001; ****P < 0.0001; ns, not significant, using Welch's t-test.

Table 2.3. Flowering time of CRISPR lines.

16°C	**Leaves**	**Cauline leaves**	**Days**	**n**
flm-3	15.1 ± 1.5	4.2 ± 0.4	36.3 ± 2.3	15
FLM-ΔE2	15.8 ± 1.7	3.8 ± 1.1	37.5 ± 3.4	14
Col-0	26.9 ± 2.1	5.4 ± 1.2	49.3 ± 3.7	14
FLM-ΔE3	33.2 ± 5.2	6 ± 1.6	54.6± 7.3	13
35S:gFLM	53.4 ± 3.3	10.4 ± 1.7	62.5 ± 4.6	8
23°C	**Leaves**	**Cauline leaves**	**Days**	**n**
flm-3	8.6 ± 1	2.7 ± 0.5	18.8 ± 0.6	16
FLM-ΔE2	9.5 ± 1	2.9 ± 0.3	18.9 ± 0.5	16
Col-0	13.1 ± 1.3	3.2 ± 0.7	25.1 ± 2.2	15
FLM-ΔE3	16.4 ± 2.2	3.6 ± 0.6	27.3 ± 1.6	16
35S:gFLM	25.2 ± 2.3	4.9 ± 0.7	35.2 ± 1.5	11
27°C	**Leaves**	**Cauline leaves**	**Days**	**n**
flm-3	6.4 ± 0.6	2.4 ± 0.5	14.6 ± 1.4	16
FLM-ΔE2	7.4 ± 1	2.8 ± 0.4	15.3 ± 1.5	16
Col-0	7.8 ± 1	2.8 ± 0.4	17.8 ± 1.3	16
FLM-ΔE3	9.7 ± 0.9	3.1 ± 0.3	20.3 ± 2.6	16
35S:gFLM	11.3 ± 1.1	2.6 ± 0.5	23.3 ± 2.1	12

n = number of plants analysed per genotype.

Similar to previous results (Scortecci *et al*. 2001), constitutive expression of *gFLM* delayed flowering in all conditions tested, but the effect was more pronounced at 16°C than at 23°C or 27°C (Figure 2.11 and Table 2.3). This is possibly due to the in reduction in exon 2-containing transcripts relative to exon 3-containing transcripts at 27°C (Figure 2.9).

FLM-ΔE3 line, which expresses higher level of exon 2-containing transcripts (Figure 2.8), was late flowering when compared to Col-0, but not as late as the *35S:gFLM* plants (Figure 2.11), confirming the importance of exon 2 for the repressive function of FLM.
FLM-ΔE2 plants, instead, which produce approx. from 2 to 4 folds higher levels of *FLM-δ* in comparison with the WT, flowered earlier than Col-0, particularly at 16°C and 23°C, but never earlier than the *flm-3* mutant. However, plants expressing the *FLM-δ* open reading frame under the constitutive *35S* promoter in the *flm-3* background flowered significantly earlier than *flm-3* control plants at 16°C (Figures 2.12 and 2.13), confirming previously published results (Posé *et al.* 2013).

Taken together, these data indicate that in Col-0 the reduction in *FLM-β* rather than a dominant negative effect of *FLM-δ* causes plants to flower earlier at elevated ambient temperatures. Nevertheless *FLM-δ* can potentially exert a dominant negative effect on flowering time when expressed at high levels. It remains to be determined whether *FLM-δ* can contribute to the regulation of flowering in more natural settings, in other accessions, or under different growth conditions.

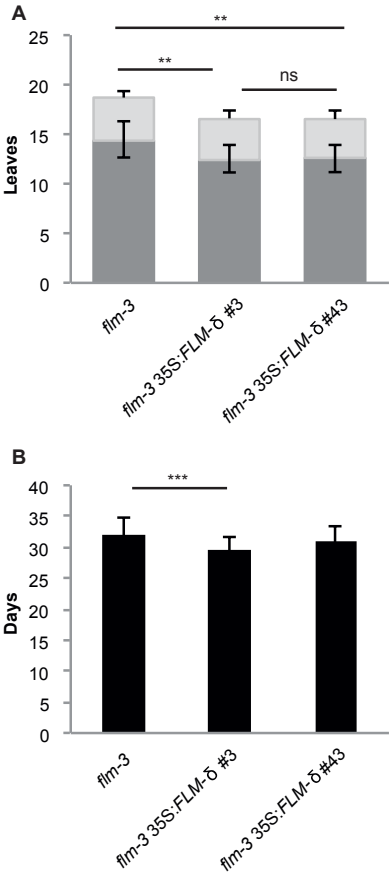

Figure 2.12: Flowering time of *35S::FLM-δ* lines. A. Flowering time given as number of leaves (dark grey: rosette leaves; light grey: cauline leaves) B. Days to flowering (black) of plants grown at 16°C. Error bars indicate standard deviation. **P < 0.01; ***P < 0.001; ns, not significant, using Welch's t-test.

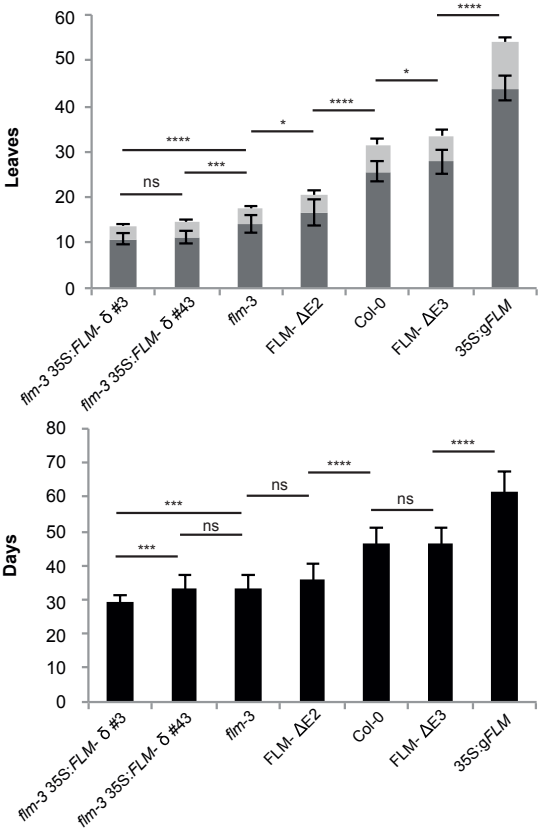

Figure 2.13: Independent replicate of flowering time of all the *FLM* transgenic lines and mutants considered at 16°C long days. In the upper panel the flowering time is given in number of leaves and in the lower panel in days to flowering. Error bars indicate standard deviation. *P < 0.05; **P < 0.01;***P < 0.001; ****P < 0.0001; ns, not significant, using Welch's t-test.

2.2.3 DISCUSSION

In the past years, the contribution of the MADS-domain transcription factor FLM in controlling flowering in response to changes in ambient temperature has been investigated by several groups.

At first Scortecci and colleagues identified FLM as floral repressor (2001), and Balasubramanian and colleagues (2006) associated FLM with the thermal induction of flowering time.

FLM was already known to be subject to alternative splicing (Scortecci *et al.* 2001), and that temperature could regulate the abundance of an active isoform (Balasubramanian *et al.* 2006). But only in 2013 *FLM-β* was identified as the isoform that functions as a flowering repressor and the mechanism of action was revealed (Posé *et al.* 2013; Lee *et al.* 2013). In 2013, Posè and colleagues proposed a model explaining how temperature-dependent alternative splicing contributes to the induction of flowering. This model is based on the finding that the expression of a second protein coding isoform, *FLM-δ*, is promoted at high temperature. The FLM-δ protein is thought to compete with FLM-β for interaction with partners such as SVP, resulting in the formation of non-functional complexes and thereby indirectly inducing flowering at elevated ambient temperature.

Evidences for the potential dominant-negative effect of FLM-δ on flowering time are largely based on the phenotypes of lines expressing this isoform under the constitutive *35S* promoter (Posé *et al.* 2013), which is a highly artificial context. More recently, the importance of FLM-δ in regulating flowering has been challenged and an alternative mechanism has been proposed. Sureshkumar and colleagues (2016) show that at elevated temperature the production of non-canonical *FLM* isoforms targeted for degradation by NMD is induced and this would explain the consequently reduced level of the

functional isoform, *FLM-β* and therefore the early flowering phenotype at high temperatures.

In this thesis, I established to what extent the two most abundant FLM variants, FLM-β and FLM-δ, contribute to the thermosensitive regulation of flowering in a system as close as possible to wild type.

I employed CRISPR/Cas9 to obtain two lines, FLM-ΔE2 and FLM-ΔE3 carrying a targeted deletion of the 2nd or 3rd exon, respectively. The edited lines do not express either *FLM-β* or *FLM-δ* (Figure 2.6) but the expression of other isoforms and the total expression of *FLM* were not compromised (Figure 2.6).

Probably as a consequence of the rearrangements on the *FLM* genomic sequence, I detected 19 isoforms in the CRISPR lines that were not detectable in Col-0 (Figure 2.6). Alternatively, the depth of the Sanger sequencing analysis was simply too low to detect all possible splice variants. Irrespectively of the reason, the new isoforms were detectable only at low frequency (Table 2.1) and contained premature termination codons (PTC) similarly to the majority of the isoforms shared with the wild type. For these reasons, the CRISPRs lines represent the best compromise to study the specific role of isoforms in a WT background. The expression of *FLM-β* in the FLM-ΔE3 line was higher than WT but still down regulated in response to increasing temperature (Figure 2.9, Table 2.1) just like in Col-0. Similarly, *FLM-δ* expression was higher than in Col-0 in the FLM-ΔE2 line at all temperatures.

In agreement with previous findings (Sureshkumar *et al.* 2016), I detected an increase in the frequency of non-canonical isoforms at elevated temperature (Table 2.1) in both wild type and the edited lines. Taken together these results strongly suggest that the edited lines behave mostly like Col-0 and that even if the relative abundance of individual *FLM* isoforms was changed because of the edited

genomic sequence, the overall expression of the *FLM* gene remained constant (Figure 2.8).

As expected due to the high expression of transcripts containing the 2nd exon (*FLM-β*), the FLM-ΔE3 line displayed a significant delay in flowering when compared to the wild type but still flowered earlier than as the transgenic lines overexpressing *gFLM*, especially at low ambient temperature (Figure 2.11). Constitutive expression of *gFLM* delayed flowering in all conditions tested, which is in agreement with a previous report (Scortecci *et al.*, 2001). However, the effect was less pronounced at 27°C, possibly due to the concomitant increase of exon 3-containing transcripts (Figure 2.9) that might buffer the over-expression of the functional repressor.

These findings demonstrate that the contribution of the *FLM-β* splice variant to the regulation of flowering can be investigated in the genome-edited line. On a more general level, my results demonstrate that CRISPR/Cas9-generated deletions lines can be employed to address specific questions regarding the role of specific splice variants in plant development.

Thus, the FLM-ΔE2 CRISPR line should be suited to evaluate the contribution of *FLM-δ*, which is expressed at higher levels in than in wild type in all conditions tested, to the regulation of flowering (Figure 2.8). As expected, the FLM-ΔE2 plants always flowered earlier than wild type, which can easily be explained by their inability to produce the floral repressor *FLM-β*, but, surprisingly, never earlier than *flm-3* (Figure 2.11). However, when grown in parallel with the newly generated CRISPR lines, two *FLM-δ* overexpression lines described by Posé and colleagues flowered significantly earlier than the loss of function mutant (Figure 2.12), confirming the published results (Posé *et al.* 2013). Overall these results indicated that *FLM-δ* in principle has the potential to act as a dominant-negative regulator of flowering

time when expressed at non-physiologically high levels. However, in Col-0 under the conditions tested, *FLM-δ* apparently never reaches the expression levels required to realize this potential.

Given that reducing expression of *FLM-β* is sufficient to promote flowering, it remains unclear why the plant invests energy in producing a plethora of alternatively spliced transcripts rather than just shutting down *FLM* transcription. A possible explanation would be that some of the alternative transcripts produced might play an active role in the regulation of flowering. In some conditions the FLM-ΔE2 lines actually flowered moderately late when compared to *flm-3* (Figure. 2.11) especially at elevated temperature, even though no *FLM-β* expression could be detected (Figure. 2.6). This suggests that some *FLM* isoforms lacking exon 2 might contribute to the repression of flowering.

Within all the non-canonical isoforms detected, *ASF7* and *ASF10,* clearly stand out. They have been detected relatively frequently at various temperatures in the CRISPR lines and their structures differ from *FLM-β* and *FLM-δ,* respectively, only by the retention of intron 4. The intro retention does not result in a frame shift and *ASF7* and *ASF10* could encode potentially functional proteins (Figure. 2.6). In Col-0, however, it seems unlikely that *ASF7* and/or *ASF10* play a major role in the regulation of flowering time since the expression levels of these two isoforms were extremely low (Figure. 2.10). Together these findings suggest a potential role for *FLM* in contributing to the plasticity of flowering, which could be of relevance from an evolutionary perspective. If reducing levels of *FLM-β* were sufficient to promote flowering, the reason why plants invest in producing a plethora of alternatively spliced transcripts remains elusive. It could be that because of the structure of the *FLM* transcript, evolving temperature-dependent alternative splicing was the easiest solution to the problem. Alternatively, maintaining the

ability to produce a variety of *FLM* isoforms could provide flowering time plasticity, in fact it seems possible that some of the alternative transcripts produced might play an active role in the regulation of flowering and be relevant in different accessions.

2.3 FLM IN NATURAL ACCESSIONS OF *A. thaliana*

2.3.1 INTRODUCTION

2.3.1.1 NATURAL VARIATION

Phenotypic variations within species are very obvious in domestic animals like in breeds of cats or dogs. In sessile organism like plants instead the environment can greatly influence the final phenotype and mask differences due to genetic traits.
A large collection of *Arabidopsis* germplasms from different geographic areas is available and many of them have been completely sequenced (1001 Genomes Consortium 2016). The genetic differences underlying phenotypic variation can be uncovered by growing these natural accessions under controlled conditions in the lab. The differences found between accessions under the same conditions are thought to reflect particular adaptations to different natural environments and differences in the genetic sequences.

2.3.1.2 *FLM* IN NATURAL VARIATIONS

The role of FLM in the thermosensory pathway has been previously established using two natural strains, Nd-1 (Werner et al. 2004) and Ei-6 (Balasubramanian et al. 2006). These two accessions are largely temperature-insensitive and display an early flowering phenotype.

Using recombinant inbred line populations derived from crosses to Col-0 and subsequent QTL mapping the authors identified *FLM* as the causal gene. Interestingly, both accessions contain a conserved large deletion in *FLM* that results in a non-functional protein.
In 2015, Lutz and colleagues reported that an intronless *FLM* coding sequence under the control of the *FLM* promoter was unable to restore the early flowering phenotype of *flm-3*. Furthermore, the authors identified a site in intron 1 that is required for the correct regulation of *FLM* expression by using a QTL mapping approach. They analyzed F2 segregating populations from a cross between Col-0 and an early flowering Scottish accession, Kil-0, that carries a long insertion in the first intron. Subsequent experiments demonstrated that the position, rather than the sequence of the insertion is responsible for the phenotype, and leads to a down regulation of *FLM* expression in the Kil-0 allele. More recently, other regulatory regions in the *FLM* promoter and non-coding sequences (intron 8) have been identified by testing natural alleles data from the 1001 genomes project (Lutz *et al*. 2017).

In this work, I investigated the contribution of splicing isoforms of *FLM* to the regulation of flowering time in different natural accessions of *A. thaliana*.

2.3.2 RESULTS

To address the role of *FLM* AS in more detail, we investigated to what extent the temperature-dependent splicing of *FLM* pre-mRNA observed in Col-0 is conserved within natural accessions of *A. thaliana*. We performed the TaqMan assay previously described in this chapter to analyze the expression of transcripts containing either the 2nd or 3rd exon (such as *FLM-*β and *FLM-*δ) in 53 non-

vernalization-requiring *A. thaliana* accessions grown at 16°C, 23°C, and 27°C. In addition, we also included 11 non-vernalized winter annual accessions in this analysis (Table 2.4).

Table 2.4. List of natural *A. thaliana* accessions analysed.

Col-0 like	Temperature insensitive	Temperature insensitive at high temperature	Veranlization requiring
Col-0	C24	NFA 8	ice63
Oy-0	Koch-1	Tu V12	ice119
Can-0	ice112	Kil-0	ice60
Aitba-1	ice50	ice92	ice91
Ra-0	ice61	Wal-Has B4	N13
KZ-1	ice127	Fei-0	Tu Sha9
Lag2-2	Cvi-0	HKT2-4	ice134
Van-0	Ts-1	TuWa1-2	Yeg-1
ice73	An-1	Jl-3	ice181
ice130			ice153
ice120			Lerik1-3
Vie-0			
Sei-0			
ice70			
ice75			
ice150			
Tsu-0			
Bor-4			
ice152			
Pu2-7			
Qui-0			
Ey1.5-2			
Sha			
STAR8			
ice71			
Tsu-1			
Rubezhnoe1			
Ru3.1-27			
TuSB30-2			
ice228			
ice1			
Ct-1			
Mt-0			
Nie1-2			

Expression of *FLM* splice variants containing the 2nd or 3rd exon revealed that the majority of accessions, including the 11 vernalization-requiring accessions, showed a trend similar to that observed in Col-0 in (Figure 2.14).

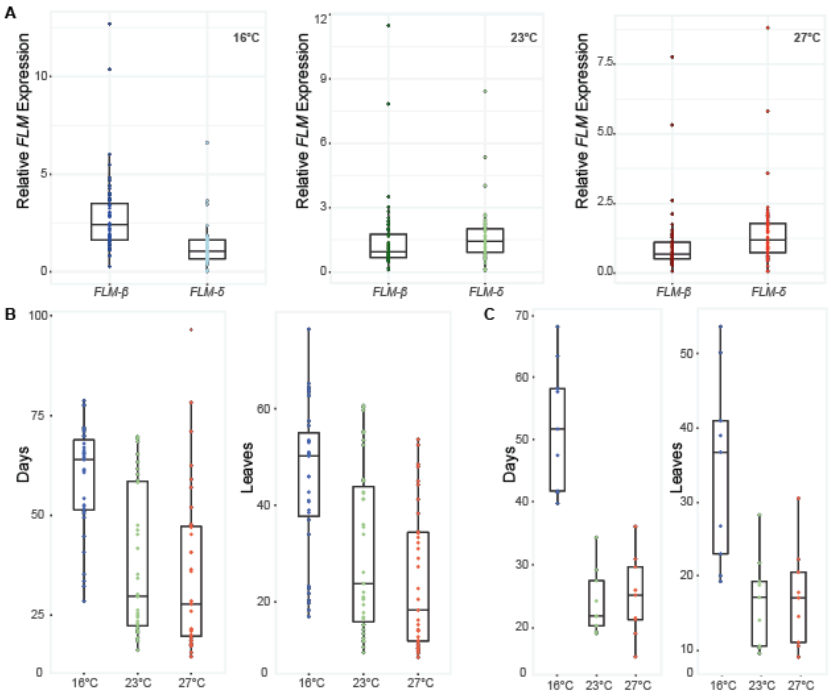

Figure 2.14: Expression of FLM isoforms in natural accessions of *A. thaliana*. A. Overall relative expression of transcripts containing the 2nd (e.g. *FLM-β*) or the 3rd (e.g *FLM-δ*) exon measured using the TaqMan assay in 64 natural *A. thaliana* accessions. B. Mean flowering time in days and leaf number of 34 accessions showing Col-0-like temperature-dependent FLM splicing. C. Mean flowering time in days to flowering and leaf number of 9 accessions insensitive to temperature changes above 23°C.

The mean expression of *FLM* isoforms containing the 2nd exon (such as *FLM-β*) decreased with higher temperatures whereas the 3rd exon-containing variants (such as *FLM-δ*) showed the lowest abundance at 16°C (Figure 2.14A). Despite this largely consistent trend in *FLM* splicing, four distinct groups of accessions can be discriminated based on their flowering time phenotypes among the 53 non-vernalization-requiring accessions (Table 2.4).

Table 2.5. Flowering time of natural *A. thaliana* accessions.

Accession	TL16	D16	TL23	D23	TL27	D27	n
Col-0	23.2 ± 2.7	35.3 ± 3.2	13.4 ± 1.5	19.2 ± 1.6	10.4 ± 1.2	15.7 ± 1.8	86
Can-0	45.9 ± 3.7	52.6 ± 7.4	26.1 ± 6.7	29.6 ± 2.6	14.1 ± 1.8	18.7 ± 3.6	43
Aitba-1	18.3 ± 2.6	28.5 ± 4.3	19.4 ± 4.1	29.8 ± 4.1	18.3 ± 2.6	28.5 ± 4.3	41
Ra-0	40.8 ± 5.3	50.5 ± 6.6	12.5 ± 2.1	19.3 ± 4	10.3 ± 1	17.8 ± 2.8	12
KZ-1	22.8 ± 3.8	64.5 ± 9.3	15.3 ± 2.6	22 ± 2.4	12.8 ± 2.8	26 ± 8.8	12
Lag2-2	41 ± 14	72 ± 13	14.3 ± 1.3	22.8 ± 2.9	10.3 ± 0.5	17.3 ± 2.2	12
Van-0	20.3 ± 1.7	33.6 ± 2.5	16.8 ± 3	20.7 ± 1.5	9.7 ± 1.4	17.8 ± 3.7	81
ice73	50.3 ± 3.5	67 ± 2.6	35.5 ± 9.7	41.8 ± 8.4	29 ± 8	40.8 ± 8.6	11
ice130	50.3 ± 5.4	71 ± 3.7	41.3 ± 3.1	63.3 ± 13.6	34.5 ± 4.9	47.5 ± 7.9	11
ice120	64.5 ± 2.4	69.8 ± 1.9	60.3 ± 9.3	69.5 ± 6.9	53.7 ± 8.1	57 ± 15.6	11
Vie-0	76.5 ± 1.3	77.8 ± 8.1	59.7 ± 4.6	65.3 ± 12.9	52.5 ± 6.2	78.3 ± 6.5	11
Oy-0	53.3 ± 7.3	51.7 ± 5.2	20.5 ± 3.2	21.3 ± 1.7	14 ± 1.1	18.6 ± 0.7	85
Sei-0	16.9 ± 2.6	32.3 ± 2.2	10.6 ± 1.7	16.3 ± 1.4	9.4 ± 1.1	14.6 ± 3.8	85
ice70	56.5 ± 15.4	78.8 ± 11.3	55.3 ± 7.2	60.3 ± 4.2	44.3 ± 5.5	52 ± 8.5	11
ice75	53 ± 4.4	70 ± 6.1	52.5 ± 8.7	58.8 ± 6.9	31 ± 6.4	40.8 ± 8.5	11
ice150	51 ± 3.6	65.8 ± 2.5	45 ± 10.9	58.3 ± 9.4	41.3 ± 4.9	59 ± 6.7	12
Tsu-0	53.5 ± 7.8	65.5 ± 7.6	23.8 ± 4.1	26.3 ± 2.2	16.3 ± 1.5	21.5 ± 1.3	12
Bor-4	63.5 ± 4.7	65 ± 5.4	36 ± 2.2	34.3 ± 1.7	20.5 ± 2.4	21.5 ± 1.3	12
ice152	50.3 ± 4.9	65.5 ± 4	41.3 ± 3.5	69.3 ± 14.2	33.3 ± 4.5	47.7 ± 6	11
Pu2-7	57.5 ± 9.2	77.5 ± 7.8	42.5 ± 1.7	45.3 ± 6.8	38.3 ± 3.2	47 ± 4	9
Qui-0	62.8 ± 10.1	61.5 ± 9.8	45.3 ± 17.8	46.3 ± 16.6	34.3 ± 24.2	36 ± 23.6	11
Ey1.5-2	50.7 ± 3.1	71.8 ± 10.7	16.5 ± 1.7	23.5 ± 1.9	14 ± 0.8	20 ± 2.2	12
Sha	19.8 ± 3	44.8 ± 3.2	9.5 ± 0.6	24.5 ± 1.3	8.5 ± 1	19.5 ± 1.7	12
STAR8	38.5 ± 8.4	49.5 ± 2.6	22 ± 4.7	30.3 ± 5.9	27.3 ± 7.9	36.5 ± 9	12
ice71	46 ± 10.7	68 ± 12.2	42.8 ± 5.3	47.5 ± 5.7	32.3 ± 3.1	45.3 ± 4.3	12
Tsu-1	53.3 ± 5.4	66.3 ± 11.6	20.8 ± 2.1	25.5 ± 1.9	14.3 ± 1.3	21.3 ± 2.5	12
Rubezhnoe-1	51 ± 3.2	54.3 ± 3.3	34 ± 12	35.3 ± 8.5	22.8 ± 3.2	28.5 ± 8	12
Ru3.1-27	42.8 ± 4.1	51 ± 4.8	21 ± 1.8	22 ± 2	15.3 ± 3.4	21 ± 4.1	12
TuSB30-2	39 ± 2	52.3 ± 2.1	17.8 ± 2.6	25 ± 1.4	12 ± 1.8	21.5 ± 2.4	11
ice228	62.8 ± 6.5	60.8 ± 4.8	55 ± 2	61.7 ± 7.6	48.5 ± 4.9	96.5 ± 13.4	9
ice1	64 ± 5.3	64 ± 2	60.7 ± 1.5	69.7 ± 0.6	48 ± 5.7	71 ± 2.8	10
Ct-1	37 ± 3.4	64 ± 4.8	11.8 ± 1.5	21 ± 2.7	10.8 ± 1	20 ± 0.8	12
Mt-0	21.8 ± 2.9	40.8 ± 3.6	12 ± 1.4	18.5 ± 1.3	11.3 ± 1.3	17.5 ± 3	12
Nie1-2	34 ± 8.8	61.3 ± 9.1	15.3 ± 4	22.8 ± 6.3	11.8 ± 2.9	27.8 ± 6.7	12
C24	23.5 ± 2.9	58.8 ± 4.3	26.25 ± 8.1	27.3 ± 3.2	29.25 ± 3.6	28.3 ± 1.7	12
Koch-1	52.75 ± 3.8	55 ± 4.4	56.5 ± 2.1	53.5 ± 2.1	50.75 ± 7.4	60.3 ± 5.9	12
ice112	51.75 ± 6.2	62.3 ± 5.7	73.3 ± 16.8	69 ± 14.4	58.7 ± 1.5	89.3 ± 12.9	10

ice50	35.75 ± 5.9	61.8 ± 1.7	31.5 ± 5.3	28.3 ± 1.9	30.5 ± 4.2	40.3 ± 14.1	12
ice61	45 ± 3.6	58.3 ± 1	52.5 ± 1.9	62.3 ± 3.6	41.75 ± 3.3	53.8 ± 13.8	12
ice127	54 ± 14.1	92 ± 17	56 ± 10.6	78.7 ± 2.5	51.25 ± 8.6	83.8 ± 6.9	9
Cvi-0	13.7 ± 2.9	36 ± 3.7	13.7 ± 3	25.5 ± 4.2	13.5 ± 2.4	18 ± 3.3	81
Ts-1	33.3 ± 5.5	41.7 ± 3.2	31.5 ± 7.8	34.4 ± 12.1	34.5 ± 17.6	32.7 ± 7.2	83
An-1	15.25 ± 1.3	52.5 ± 2.5	9.75 ± 1	18.5 ± 1.7	11.5 ± 0.6	19 ± 3.2	12
NFA 8	23.3 ± 5.4	41.5 ± 4.4	10.8 ± 2.5	18.75 ± 1.9	10.8 ± 1.7	18.5 ± 4.2	12
Tu V12	39.3 ± 8.2	47.25 ± 5.3	22 ± 6.8	28.75 ± 9	22.5 ± 6.2	30.5 ± 7.8	12
Kil-0	27 ± 3.3	58 ± 4.5	17.3 ± 1.2	21.3 ± 0.6	17.3 ± 2.2	25.5 ± 4.4	11
ice92	50.5 ± 2.9	57.5 ± 3.9	28.5 ± 15.5	34 ± 9.6	30.8 ± 6.9	35.75 ± 5.7	12
Wal-Has B4	54 ± 2.8	63.25 ± 3	19.5 ± 1.3	27 ± 3.9	20.8 ± 6.5	29.25 ± 5.4	12
Fei-0	37 ± 4.2	68 ± 10.9	19 ± 3.4	23.75 ± 3.3	18 ± 2.6	24.7 ± 0.6	11
HKT2-4	41.3 ± 4.8	51.5 ± 3.1	14.3 ± 0.5	21.25 ± 2.5	14.8 ± 1.5	20.75 ± 1.7	12
TuWa1-2	20.3 ± 4.5	41.25 ± 3.9	10.5 ± 1	19.75 ± 0.5	11.3 ± 0.5	21 ± 1.2	12
Jl-3	19.5 ± 2.6	39.5 ± 7.7	9.8 ± 1.5	18.5 ± 2.1	9.3 ± 1	14.75 ± 1.7	12

TL16, TL23 and TL27, total leaf number at 16°C, 23°C and 27°C; D16, D23 and D27 days to flowering at 16°C, 23°C and 27°C; n, total number of plants analysed per accession.

Thirty-five accessions behaved similarly to Col-0, flowering later at low temperature and earlier at high temperature, when both the total leaf number and days to flowering are considered (Figure 2.14B). To this group belong Rubezhnoe1 and Tsu-0, which have previously been described as temperature insensitive (Lee *et al.*, 2013). In contrast, a group of 9 accessions, including Jl-3 as previously described (Lee *et al.*, 2013), were relatively late flowering at 16°C but flowered early at both 23°C and 27°C (Figure 2.14C, and Figure 2.15).

Furthermore, 10 accessions were categorized as temperature insensitive, since they flowered at approximately the same time in days to flowering and/or leaf number irrespective of the ambient temperature (Figure 2.16).

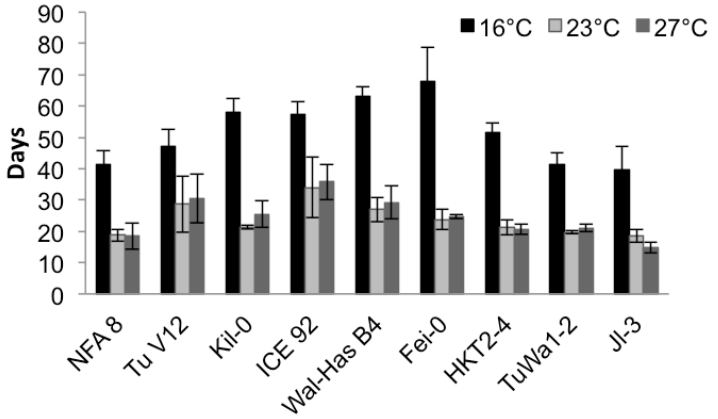

Figure 2.15: Flowering time phenotypes at 16°C, 23°C and 27°C in days (upper panel) and leaves (lower panel) of 9 accessions insensitive to temperature changes above 23°C.

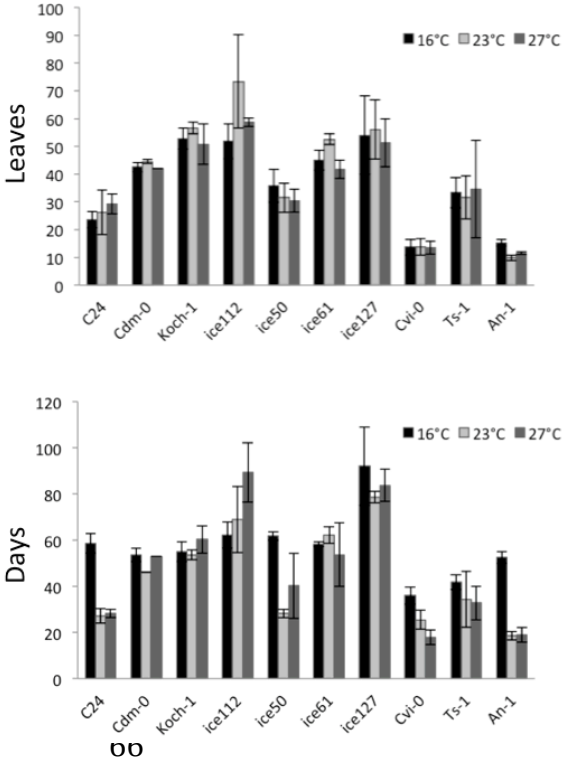

Figure 2.16: Detailed flowering time phenotypes at 16°C, 23°C and 27°C in leaves (left panel) and in days (right panel) of 10 temperature insensitive accessions.

This group includes An-1 and Cvi-0, which have previously been described as temperature insensitive accessions (Balasubramanian *et al.*, 2006; Lee *et al.*, 2013) and Ts-1 that carries an 8 bp deletion in exon 2 of *FLM*, resulting in a frame shift and leading to a premature stop codon (Figure 2.17).

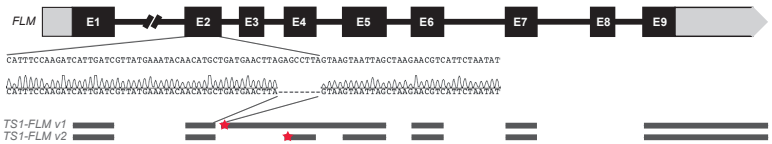

Figure 2.17: *FLM* in Ts-1 carries an 8bp deletion in exon 2. In black the annotated Col-0 *FLM* gene, exons are marked as squares and introns as straight lines. The schematic representation of the two splice isoforms isolated in Ts-1 carrying the 8bp deletion in exon2 are drawn in grey and the close-up show detailed information on the position of the deletion obtained by Sanger sequencing. Stop codons are represented as red stars.

2.3.3 DISCUSSION

Natural accessions are an incredibly valuable source of information that makes it possible to explore the ecological relevance of mechanisms identified in Col-0. In this thesis, I determined the expression of the two *FLM* isoforms most abundant in Col-0, *FLM-*β and *FLM-*δ, with the aim to test how conserved the AS pattern is in natural accessions.

The temperature-dependent AS of *FLM*, including the increased relative abundance of transcripts containing the 3rd exon (such as *FLM-*δ) at elevated temperatures, seems to be well conserved in most

natural accessions tested (Figure 2.14A). However, in this analysis, AS of *FLM* was poorly correlated with flowering of natural accessions, which was not entirely unexpected as flowering time is a complex trait that is regulated by diverse environmental and endogenous cues. In most plants cold temperatures delay flowering, and in fact the majority of the rapid cycling accessions with non-functional *FRIGIDA (FRI)* and/or *FLC* alleles analyzed behaved like Col-0 and showed accelerated flowering in response to elevated ambient temperature (Figure 2.14B). Nevertheless, even accessions that require vernalization to induce flowering (Table 2.4) and that were excluded from the phenotyping because of their extremely late flowering, displayed the typical AS pattern of *FLM* when grown at different ambient temperatures.

Several accessions displayed a partial insensitivity to elevated temperatures (Figure 2.14C and Figure 2.15) or flowered essentially at the same time at all three conditions tested (Figure 2.16). The latter include Ts-1, which carries an 8 bp deletion in exon 2 of *FLM* that results in a frame shift and PTC (Figure 2.17) and flowers later than Col-0 at the three temperatures considered (Table 2.5). The same deletion has been previously described in a different accession, Sf-2 (Sureshkumar *et al.* 2016). However, the temperature-insensitivity of Ts-1 cannot be attributed directly to this mutation alone. As suggested by Sanchez-Bermejo and colleagues (2015), thermal response is associated with many genomic regions including *SVP*, the *MAF* clade genes, *FVE, FCA*, or potentially other elements still uncovered. In addition, temperature-dependent regulation of flowering is intimately linked to light signaling as highlighted by the recent finding that phytochrome B might act as a temperature sensor in plants (Legris *et al.* 2016; Jung *et al.* 2016).

In summary, it can be concluded that in natural accessions the temperature-dependent AS splicing of *FLM* is well conserved. The

active isoform, *FLM-β* is down regulated in increasing the temperature in all the accessions considered. However, in most cases the temperature-dependent abundance of the active *FLM* isoform does not correlate with (or even predict) the phenotype. Clearly, other players than *FLM-β* contribute to the flowering time phenotype, and additional QTL analyses might be useful to identify new thermoregulatory elements.

2.4 METHODS

2.4.1 PLANT MATERIAL AND GROWTH CONDITIONS

Seeds were surface sterilized with 20 mL of thin bleach and 1 mL of 37.5% HCl for 4 h followed by 1.5 h in a laminar flow to evaporate chlorine gas and stratified in 0.1% agar at 4°C in the dark for 72 h before being planted directly on soil. Seeds from natural accessions listed in Table 2.4 belong to the 1001 genomes project, and were obtained from colleagues at the MPI for Developmental Biology, Tuebingen, Germany. *flm-3, flm-3 35S:FLM-δ* #3, and *flm-3 35S:FLM-δ* #43 lines have been previously published (Posé *et al.* 2013). Plants were grown in soil in long day conditions (LD, 16 h light/8 h dark) at a specified temperature either in Percival chambers or in growth rooms. To analyze variation in splicing patterns in response to a change in temperature, plants were grown for 9 days at 23°C LD and then shifted to 16°C, 27°C or kept at 23°C LD for 3 days. Three pools of 10 seedlings for each line in each temperature were collected after the shift at zeitgeber 6 and snap frozen in liquid nitrogen. To analyze flowering time, plants were grown at 16°C, 23°C,

or 27°C and the days to flower as well as the rosette and cauline leaf number were recorded.

2.4.2 RNA-EXTRACTION AND cDNA SYNTHESIS

Total RNA was extracted with TRIzol® reagent or 5:1 acidic phenol:chloroform as described (Box *et al.* 2011). RNA quality was determined by using a Nanodrop ND-2000 spectrophotometer (Nanodrop Technologies) and only high-quality RNA samples (A260/A230>2.0 and A260/A280>1.8) were used for subsequent experiments. To remove possible DNA contamination, RNA samples were treated with DNaseI (Thermo Scientific) for 30 min at 37°C, subsequently inactivated at 65°C with 1ul of 50mM EDTA. 3 µg of RNA was used for complementary DNA (cDNA) synthesis using the RevertAid First Strand cDNA Synthesis kit in accordance with the manufacturer's instructions (Thermo Scientific).

2.4.3 TAQMAN ASSAY

The multicolor TaqMan analysis was carried out in 384-well plates to measure the relative expression of *FLM* splice variants that contained either the 2nd or the 3rd exon. TaqMan technology exploits the 5' nuclease activity of Taq DNA polymerase (Holland *et al.* 1991) and probes designed on the transcript of interest were labeled with fluorescent reporters with non-overlapping detection spectra at the 5' and a quencher dye at the 3' end. Primers and probes are listed in Table 2.2, which also includes details about quenchers and detection channels. iQ^{TM} Multiplex Powermix (BioRad) was used for the PCR following the manufacturer's instructions and PCR was carried out using an annealing temperature of 58°C. The exponential amplification of the fluorescence intensity was measured and

quantified (Lee *et al.* 1993) using a BioRad C1000 Touch Thermal Cycler. The relative expressions were calculated using the delta Ct method and performed in technical triplicates for each of the 3 biological replicats.

2.4.4 PLASMID CONSTRUCTION AND PLANT TRANSFORMATION

The two CRISPR/Cas9 vectors used in this study, pGC001 and pGC002 were assembled using the GreenGate system (Lampropoulos *et al.* 2013). The final constructs contain the p*EC1.1*::AthCas9:trbcs (assembled from GreenGate modules A: *A. thaliana pEC1.1*; B: A. *thaliana* codon-optimized Cas9 (Fauser *et al.* 2014); and C: rbcs terminator), the sgRNAs listed in Table 2.6 under the control of the *A. thaliana U6* promoter (GreenGate modules D and E), and a *pAt2S3::mCherry:tMAS* cassette (GreenGate module F) for seed selection as described by Gao and colleagues (2016). The *p35S*::g*FLM* vector (GC003) was also assembled using the GreenGate system. For this, the genomic region of *FLM*, including UTRs was amplified from Col-0 seedlings and cloned into the GreenGate module C entry vector (Lampropoulos *et al*. 2013). The final GreenGate reaction was performed using modules A: *p35S*, B: empty (pGGB003), the C module carrying the full genomic *FLM* region, D: empty (pGGD002), E: rbcs terminator, and BASTA resistance as selection marker (module F; pGGF001). pGGZ001 was used as destination vector for all the GreenGate reactions described. Plants grown at 23°C were transformed by floral dipping using *Agrobacterium tumefaciens*-mediated gene transfer according to standard protocols (Clough and Bent 1998). Transformants were selected by fluorescence microscopic identification of mCherry-

positive seeds as previously reported in Gao *et al.* (2016) or by BASTA selection.

Table 2.6. Oligonucleotides used in this work.

Oligonucleotides used for cloning sgRNA constructs

sgRNA	Oligo	Sequence (5' -> 3')
FLM-E2-5'	G-40172	gTCAATGATCTTGGAAATGCgttttagagctatgctg
	G-40173	GCATTTCCAAGATCATTGAcaatcactacttcgactc
FLM-E2-3'	G-40174	gCTTAGCTAATTACTTACTAgttttagagctatgctg
	G-40175	TAGTAAGTAATTAGCTAAGcaatcactacttcgactc
FLM-E3-5'	G-40176	gGCGGTTTTTGGTGTTATGAgttttagagctatgctg
	G-40177	TCATAACACCAAAAACCGCcaatcactacttcgactc
FLM-E3-3'	G-40178	gAACTCTAGAGAATTAAGTTgttttagagctatgctg
	G-40179	AACTTAATTCTCTAGAGTTcaatcactacttcgactc

Oligonucleotides for PCR to detect FLM deletions

Position	Oligo	Sequence (5' -> 3')
Intron 1	G-39775	GCACCAGATGATCAGAGTTTCA
Exon 4	G-28145	GATAATTCTGAATTTTTTCTTCAAGATC

Line	PCR product size
Col-0	652
FLM-ΔE2	595
FLM-ΔE3	588

Oligonucleotides for PCR to screen for FLM isoforms

Position	Oligo	Sequence (5' -> 3')
Exon 1	G-36231	CGCTGTTGTCGTCGTATCTG
Exon 9	G-28156	CAGCAACGTATTCTTTCCCAT

Oligonucleotides for qPCR to quantify specific FLM isoforms

Position	Oligo	Sequence (5' -> 3')
Exon 1 to Exon 4	G-43257	CTTAGAGCCTTAGATCTTGAAG
Exon 5 to Exon 4	G-43258	CTTCAAGCTTGCTTTGGACTG
Exon 1 to Exon 3	G-43259	CCTCCGGTGACGAGATAGAAG
Exon 4 to Exon 3	G-43260	GAATTTTTTCTTCAAGATCGAG
Exon 3 to Exon 4	G-43261	GTTTTGAACTCGATCTTGAAG
Intron 4	G-2641	GAGGGGAGAAAAATGTGTCG

2.4.5 SANGER SEQUENCING OF *FLM* CLONES

The *FLM* open reading frame was amplified using primers 5' CGCTGTTGTCGTCGTATCTG 3' and 5' CAGCAACGTATTCTTTCCCAT 3' from the same cDNA used for the TaqMan assay described above. The PCR products were subsequently cloned into pGem®-T Vector System I according to the manufacturer's instructions (Promega) and individual clones were sent for Sanger sequencing.

2.5 CONTRIBUTIONS

Giovanna Capovilla (GC) and Markus Schmid (MS) conceived and designed the experiments. GC performed the experiments using some plasmids generated by Efthymia Symeonidi (ES), and Rui Wu (RW). GC analyzed the data.

Part of this chapter, with modifications, has been accepted for publication in

Capovilla G, Symeonidi E, Wu R, Schmid M. 2017. Contribution of major *FLM* isoforms to temperature-dependent mediated flowering in *Arabidopsis thaliana*. Journal Experimental Botany.

CHAPTER 3

PORCUPINE (*PCP*) regulates development in response to temperature variations through alternative splicing in *Arabidopsis thaliana*

3.1 INTRODUCTION

3.1.1 ALTERNATIVE SPLICING FACTORS

As already described in chapter 1, the majority of protein-coding genes in eukaryotes contain introns. An essential step in the regulation of gene expression is the removal of these introns through the splicing of their pre-mRNAs, which needs to occur before mRNAs can be transported out of the nucleus (Sharp 2005).

Only very short sequences at the exon-intron boundaries are highly conserved. The challenge for the splicing machinery to recognize the correct sites and initiate splicing is facilitated by specialized proteins, mainly belonging to the serine/arginine (SR)-rich protein family or the heterogeneous nuclear ribonucleoprotein (hnRNP) family (Kornblihtt *et al.* 2013). Some SR proteins work as alternative splicing factors by binding near legitimate splice sites, or *cis*-regulatory sequences, and help recruit the splicing machinery to those sites. In contrast, members of the hnRNP family bind the newly-transcribed pre-mRNA but are unable to recruit the splicing machinery, and consequently repress the use of specific splice sites (Watson 2015). *Cis*-regulatory sequences include exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs), depending on their locations and on how they affect the usage of a splice site (Kornblihtt *et al.* 2013) (Figure 3.1).

The mechanism of RNA splicing and its key players have been extensively studied in human and yeast systems, whereas in plants it remains comparatively poorly understood. Although the splicing mechanisms appear to be overall well conserved between plants and animals, incorrect splicing of plant pre-mRNAs in mammalian systems (and *vice versa*) suggests the presence of plant-specific characteristics (Lorkovik *et al.* 2000).

Genome projects are accelerating the research on splicing-related genes in plants. In the *Arabidopsis thaliana* genome 74 small nuclear RNA (snRNA) and 395 genes encoding splicing-related proteins have been computationally identified, as potential homologs to animal splicing factors (Wang and Brendel 2004). However, for the majority of these factors an effective role in AS has not yet been proven in plants.

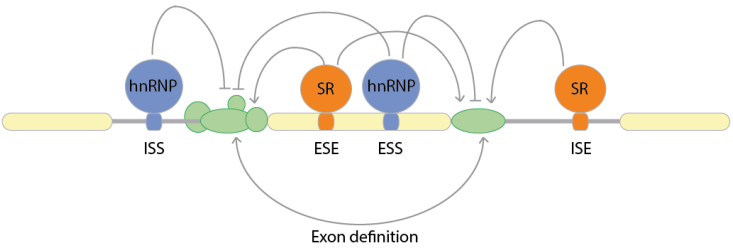

Figure 3.1: The alternative splicing regulatory sequences and factors are schematically represented. In the pre-mRNA *cis*-regulatory sequences, both in introns (grey lines) and in exons (yellow) can be recognized by regulatory proteins. Intronic or exonic splicing silencers (ISS or ESS) (blue) target regulatory components that inhibit the recognition and usage of the site by the spliceosome. In contrast, exonic or intronic splicing enhancers (ESE or ISE)(orange) associate with components of the spliceosome (green) and can mediate exon recognition (Figure adapted from Kornblihtt *et al*. 2013).

A characteristic property of AS factors is that they contain a RNA-recognition motif. Another group of proteins involved in RNA metabolism are the Sm proteins, members of a family of small proteins. These proteins function in multiple aspects of RNA metabolism, included pre-mRNA splicing (Okamoto *et al*. 2016, Cao *et al*. 2011). In *Arabidopsis thaliana* 42 *Sm* genes have been identified (Cao *et al.* 2011). The role of Sm proteins is still largely unknown except for a few cases. The most well studied Sm protein is SAD1, which controls AS in response to heat stress, drought or abscisic acid (Okamoto *et al*. 2016, Xiong et al. 2001).

3.1.2 TEMPERATURE DEPENDENT ALTERNATIVE SPLICING

One way by which AS can modulate transcript levels is by producing new unstable mRNA isoforms that are degraded by nonsense-mediated decay (NMD) (McGlincy and Smith 2008, Kalyna *et al.* 2011). Alternatively, AS can increase the diversity of an organism's proteome by changing or removing functional domains, thereby producing protein isoforms with differences in subcellular localization, stability, or function (Syed *et al*. 2012).

AS is an important mechanism involved in ambient temperature responses, and allows plants to adjust their development to fluctuations in temperatures (Verhage *et al.* 2017). Temperature changes are reflected by the AS of several plant genes, many of which are components of the splicing machinery, whose isoforms in turn can affect splicing of downstream genes (Verhage *et al.* 2017).

Temperature fluctuations have been shown to trigger differential AS both in transcriptome-wide analyses (Leviatan *et al.* 2013, Verhage *et al.* 2017) as well as in several case studies that include *FLM* and *MAF2* (discussed above) and several components of the circadian clock (Capovilla *et al*. 2015).

The circadian clock is a central mechanism that enables organisms to synchronize biological processes with the day–night-cycle by creating oscillations of approx. 24 hours (Salomè and McClung 2004). In plants, this mechanism involves several transcriptional repressors that form the core of a complex negative feedback machinery. The morning-expressed components repress the activity of the evening complex, which in turn represses PSEUDO-RESPONSE REGULATOR (PRR) genes. These close the three-way circle by repressing the morning components (Chow and Kay 2013, Nohales and Kay 2016).

The identity of spliceosome regulators that affect plant clock function is only beginning to be elucidated. PROTEIN ARGININE METHYLTRANSFERASE 5 (PRMT5) has been shown to regulate the AS of *PRR9* (Hong *et al.* 2010, Sanchez *et al.* 2010). Also some SM-like (LSM) genes, which encode core components of the spliceosomal U6 small nuclear ribonucleoprotein complex, were also found to be involved in regulating circadian rhythms (Perez-Santàngelo *et al*. 2014). Furthermore, two AS factors have been linked to temperature compensation in the clock: GEMIN2, (Schlaen *et al*. 2015), and SNW/SKI-INTERACTING PROTEIN (SKIP) (Wang *et al.* 2012).

In this chapter I present the role of a Sm protein, which I called PORCUPINE (PCP), in regulating AS and its importance in ensuring the correct development and morphogenesis at low temperatures.

3.1.3 MERISTEM AND LATERAL ORGAN DEFECTS

The entire aerial part of the plant is derived from a group of undifferentiated cells in the shoot apical meristem (SAM), and a correct formation of the SAM is crucial for a normal development. In *Arabidopsis thaliana* the cells of the shoot meristem are organized in

three clonally distinct populations, or layers: the outermost epidermal L1 layer, the subepidermal L2 layer and the internal layer L3 that contains the organizing center (Steeves and Sussex 1989). The SAM needs to balance the maintenance of undifferentiated stem cells at its center with the repositioning of differentiated cells that will form organ primordia. In fact, at the shoot meristem many cell signaling key events occur, including the commitment to differentiation, the initiation of organ primordia and the establishment of polarities within each organ primordium. The breakdown of either of these processes would cause a morphological disaster for the plant. Genetic analyses of mutants have identified a plethora of genes whose function is necessary for SAM establishment or maintenance.

Key genes for maintaining the stem-cell character of the SAM and promote cell differentiation are *WUSCHEL (WUS)* and the *CLAVATA* genes *CLAVATA1 (CLV1), CLV2,* and *CLV3* (Doerner 2003, Clark 2001, Sparks *et al.* 2013, Heidstra and Sabatini 2014). These genes form a feedback loop in which the homeodomain transcription factor *WUS* is expressed in the organizing center, and the *CLAVATA* genes (*CLV1* and *CLV2*) in the central zone of the SAM. *CLV3* in contrast encodes a small ligand whose expression is limited to cells in the L1 and L2 layer at the tip of the SAM. The CLV3 peptide moves into the underlying L3 cells where it binds to CLV1, a LRR receptor kinase forming a complex that represses *WUS* activity (Clark 2001). In the L3 cells WUS protein, after being synthesized migrates upwards to the stem cells and activates the expression of *CLV3* there by binding to its promoter (Yadav *et al.* 2011).

Of particular relevance in later organ formations is the establishment of the adaxial-abaxial asymmetry that will define the upper to the lower surface of the mature leaves. Many genes have been shown to

be involved in this process, including members of the *YABBY* gene family (Siegfried *et al.* 1999, Eshed *et al.* 2004), the *KANADI* family *(Eshed et al. 2001)*, as well as *PHABULOSA* (*PHB*), *PHAVOLUTA* (*PHV*), and *REVOLUTA* (*REV*) which encode the class III homeodomain-leucine zipper (HD-ZIPIII) proteins (McConnell *et al.* 2001), to name just a few.

3.2 AIM

Recent findings suggest that temperature-regulated AS plays a critical role in controlling the plant response to variations in temperature at the molecular level, allowing quick adaptation to changes in environmental conditions (Verhage *et al.* 2017). How variations in temperature regulate AS and the key players in this process are, however, still largely unknown.
Using RNA-seq in combination with phenotyping candidate mutant lines I have identified a *bona fide* AS regulator that is essential for the correct development of *Arabidopsis thaliana* at low temperature. The gene, now called *PORCUPINE (PCP)*, encodes a Small nuclear ribonucleoprotein family gene (Sm). Here I describe the genetic and molecular characterization of the *pcp* knock out mutant.

3.3 RESULTS

3.3.1 RNA-SEQ ANALYSIS OF TEMPERATURE-DEPENDENT ALTERNATIVE SPLICING

Most eukaryotic genes contain introns in the precursor mRNA (pre-mRNA) that are spliced in order to obtain the mature mRNAs. A powerful post-transcriptional mechanism, the alternative splicing,

allows the control of gene expressions through the production of different mRNAs from the same pre-mRNA (Chapter 1).

The general mechanism of splicing is well studied in human and yeast model organisms. In plants, genome projects are accelerating the research on splicing-related genes. 74 small nuclear RNA (snRNA) and 395 genes encoding splicing-related proteins have been computationally identified in the *Arabidopsis* genome (Wang and Brendel 2004). This list of candidate genes has been created by searching the *Arabidopsis* genome for RNA-binding motifs based on homologies with AS factors from other species. However, the majority of these genes and their potential role in AS has still not been verified experimentally yet.

To identify factors involved in temperature dependent AS I performed a strand-specific RNA-seq in Col-0 seedlings grown in different ambient growth conditions. To avoid differences in developmental stages, plants were grown at 23°C for 9 days and then shifted to 16°C, 27°C, or kept at 23°C for 3 additional days. RNA-seq was performed in a total of 3 independent biological replicates for each temperature, from a pool of 10 seedlings each. RNA-seq data analysis (performed by SC; see Material & Methods for details) identified among the AS factors reported by Wang and Brendel 2004 a total of 25 differentially expressed temperature sensitive genes.

These 25 temperature-responsive AS factors fall into two general expression patterns. 10 genes were significantly downregulated in response to increasing ambient temperature. The expression levels of these genes were highest at 16°C and lowest at 27°C, with an intermediate value at 23°C. (Figure 3.2). The remaining 15 candidates displayed the opposite trend in that they were significantly upregulated in increasing temperature and showed the highest expression at 27°C and the lowest at 16°C (Figure 3.2).

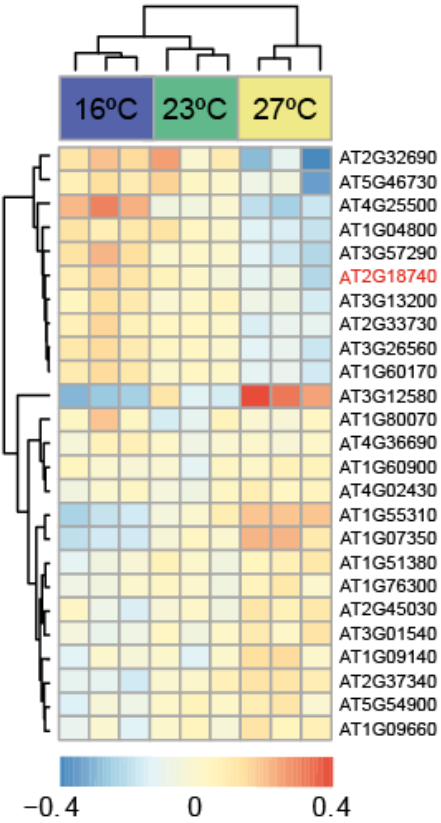

Figure 3.2 Representation of the 25 AS candidates genes significantly differentially expressed in Col-0 seedlings. 3 independent biological replicates are reported in the Heat map for each of the three temperatures considered. Blue colors indicate down regulated genes, and shades of red represent up regulated expression. *PCP* is marked in red.

To test if any of the AS genes that are differentially expressed in response to changes in temperature contribute to *Arabidopsis* development, I obtained T-DNA insertion lines from the stock center. Presences of the T-DNA was confirmed in 10 lines (Table 3.1) by

genotyping (Material & Methods) and verified insertion lines were tested in different environmental conditions.

Table 3.1. T-DNA lines tested

Name	Gene	T-DNA line
GRP23	At2g32690	SALK_004764
rsp35	At4g25500	SALK_118875
/	At2g18740	SALK_089521
/	At2g33730	SALK_100059
SR33	At1g55310	SALK_040864
sr45a	At1g07350	SALK_054457
SMD3	At1g76300	SALK_025193
SR30	At1g09140	SALK_055239
RSZ33	At2g37340	SALK_051523
RBP45A	At5g54900	SALK_041457

3.3.2 TEMPERATURE-DEPENDENT REGULATION OF DEVELOPMENT BY *PORCUPINE* (*PCP*)

3.3.2.1. PCP IS REQUIRED FOR DEVELOPMENT AT LOW AMBIENT TEMPERATURE

Among the candidates tested, the knock out line for *At2g18740* showed a very strong temperature-dependent developmental defect when plants were grown at 16°C but looked essentially like wildtype at 23°C (Figure 3.3). Closer phenotypic analysis of *pcp-1* revealed strong defects of the shoot apical meristem (SAM) and failure in developing properly formed lateral organs (Figure 3.4).

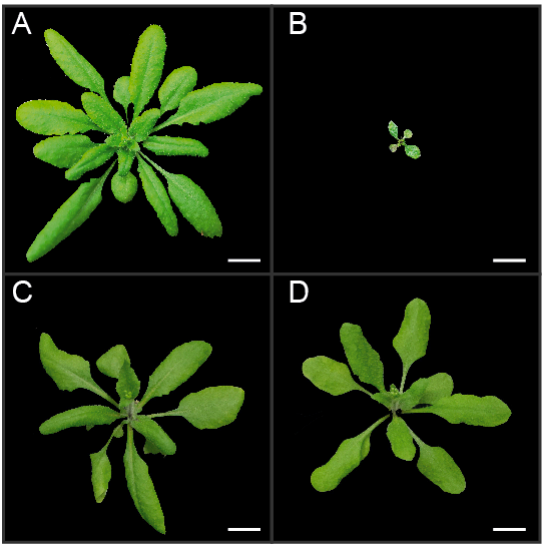

Figure 3.3 Representative pictures of Col-0 and *pcp-1* (SALK_089521) grown at different temperatures. A and B show 35 days old Col-0 (A) and *pcp* (B) plants grown under long day at 16°C. The lower panels (C and D) show 25 days old Col-0 (C) and *pcp* (D) plants grown at 23°C. Bars = 1 cm.

Because of the striking "spiky" phenotype of the mutant the gene has been named *PORCUPINE (PCP).*

Importantly, the *pcp-1* allele is recessive, indicating that the phenotype is caused by loss-of PCP function.

qPCR confirmed the temperature-dependent expression of *At2g18740* (Figure 3.5).

As previously observed by RNA-seq, expression of PCP was highest at 16°C and declined with increasing ambient temperature. The fact that PCP is most strongly expressed at low ambient temperature and that the pcp-1 mutant displays phenotypic defects only at this temperature indicates that PCP function is required for proper plant development particularly at this temperature.

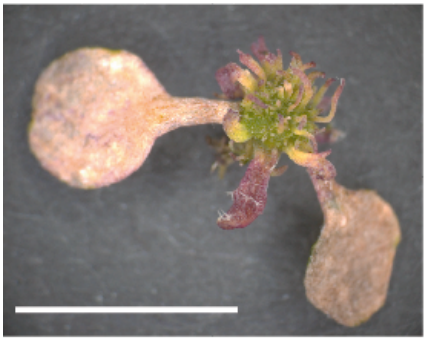

Figure 3.4 Representative pictures of a 70 days old *pcp-1* (SALK_089521) plant grown under long day conditions at 16°C. Bar = 0.5 cm.

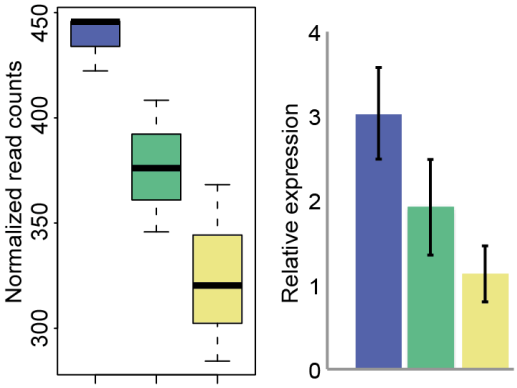

Figure 3.5 Expression analysis of *At2g18740* in Col-0 at three ambient temperatures. On the left a boxplot show the relative expression obtained from the RNA-seq analysis at 16°C (blue), 23°C (green) and 27°C (yellow). Black lines indicate median value for each group and whiskers represent the minimum and maximum values. The barplot show the results obtained from a qPCR analysis on the same biological material. Error bars designate the standard deviation of 3 biological replicates.

3.3.2.2 IDENTIFICATION OF ADDITIONAL *pcp* MUTANT ALLELES

The findings that *PCP* expression is regulated by ambient temperature and that the *pcp-1* mutant displays a strong temperature-dependent developmental phenotype are very interesting. However, these findings do not prove that the T-DNA insertion in the *PCP* gene is causing the phenotype. I therefore obtained and phenotyped three additional potential *pcp* alleles, *pcp-2* to *pcp-4*, from publicly available collections of *Arabidopsis* T-DNA insertion lines (Table 3.2 and Figure 3.6).

Table 3.2. T-DNA lines tested

Name	Gene	T-DNA line	Position	Knock out
pcp-1	*At2g18740*	SALK_089521	exon 2	YES
pcp-2	*At2g18740*	SALK_119088	intron 5	partial
pcp-3	*At2g18740*	SALK_017458	5' UTR	no
pcp-4	*At2g18740*	SALK_022142	promoter	no

Figure 3.6 Schematic representation of *PCP* (*At2g18740*). Exons are drawn as black boxes, introns as dark grey lines and UTRs as grey boxes. The positions of each of the 4 T-DNA lines analyzed are marked with triangles.

Importantly, *pcp-2* (SALK_119088) mutant plants, which carry a T-DNA insertion in the 5th intron, displayed meristem defects when grown at 16°C, even though the phenotype was less severe than those observed in the original *pcp-1* (Figure 3.7). In contrast, SALK_022142 (*pcp-3*) and SALK_017458 (*pcp-4*), in which the T-DNA is inserted in the promoter and 5'UTR of *PCP*, respectively, did not show a phenotype at any of the temperatures tested.

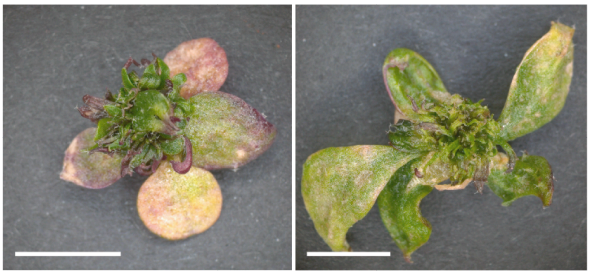

Figure 3.7 Representative pictures of two 70 days old *pcp-2* (SALK_119088) plants grown under long days at 16°C. Bar = 0.5 cm

The identification of a second independent *pcp* allele with a similar phenotype is a good indication that loss-of PCP is causing the temperature-dependent developmental defects that I had observed in the original *pcp-1* mutant. However, to finally prove that *PCP* is the causal gene I performed a complementation test.

3.3.2.3 COMPLEMENTATION OF *pcp-1*

The standard procedure to test whether a mutation is causal for a phenotype is to complement the recessive mutant with a wild type allele. To this purpose, I introduced a transgene carrying a genomic rescue fragment of *PCP* including 232 bp sequence upstream of the ATG and 423 bp downstream of its stop codon into the *pcp-1* mutant,

which is fertile and can be directly transformed when grown at 23°C. Transformants (T1) were identified by resistance to the BASTA herbicide. Importantly, transformants (T1) showed a complete rescue of the *pcp-1* phenotype when grown at 16°C (Figure 3.8).

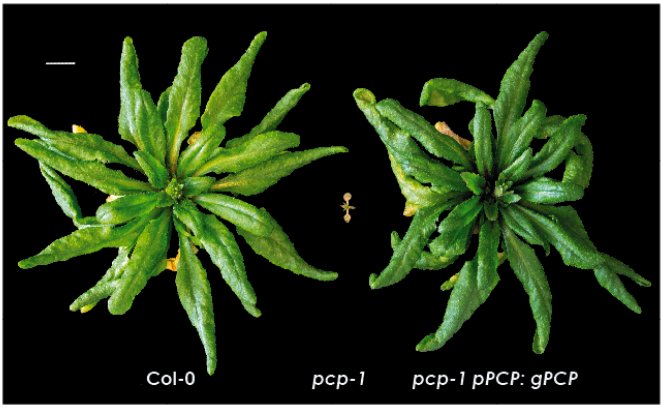

Figure 3.8 Representative picture of the Col-0 wild type, the *pcp-1* defective line and the rescued line in *pcp-1* background. Plants were grown under long day at 16°C for 40 days. Bar = 1 cm

I next tested the expression of *PCP* in Col-0, *pcp-1*, the *pcp-1 pPCP:gPCP* rescue line, and *pcp-2* (SALK_119088) grown at either 16°C or 23°C by qPCR using the expression of *UBC21* for normalization purpose. As expected, *PCP* expression was almost undetectable in the strong *pcp-1* mutant. *PCP* was expressed at similar levels in Col-0 and the *pcp-1 pPCP:gPCP* rescue line (Figure 3.9). Expression of *PCP* was only reduced to approx. 15-20% of WT levels in *pcp-2*, indicating that this allele is not a RNA-null mutant. The remaining expression of *PCP* detected could explain the weaker phenotype of *pcp-2* in comparison with *pcp-1* at 16°C.

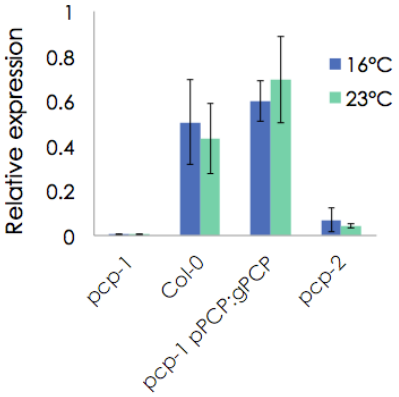

Figure 3.9 Expression analysis of *PCP* (*At2g18740)* in *pcp-1*, Col-0, the *pcp-1 pPCP:gPCP* rescued line, and *pcp-2*. For each line qRT-PCR has been performed in plants grown at 16°C (blue) and 23°C (green). Error bars indicate standard deviation between 3 biological replicates.

3.3.3 *PCP* OVEREXPRESSION LINES

Even though PCP has not been characterized in detail, 2 splice isoforms have been previously reported for *PCP* (Ito *et al*. 2010), to which I will refer in this chapter as *PCP-*α and *PCP-*β (Figure 3.10).

Figure 3.10 Schematic representation of the *PCP* (*At2g18740*) locus. Exons are shown as black boxes, introns as dark grey lines and UTRs as grey boxes. The two known isoforms, *PCP-*α and *PCP-*β are aligned to the gene structure.

To test whether both isoform encode functional variants of the PCP protein capable of rescuing the *pcp-1* loss of function phenotype, I introduced the fully spliced CDSs under the ubiquitous 35S promoter into the *pcp-1* mutant. In addition, the two splice variants were also introduced into Col-0 to evaluate the consequences of ubiquitous overexpression on plant development. Importantly, I could not detect any effect in the overexpressing lines in Col-0 background, irrespectively of whether plants were grown at 23°C or at 16°C (Figure 3.11).

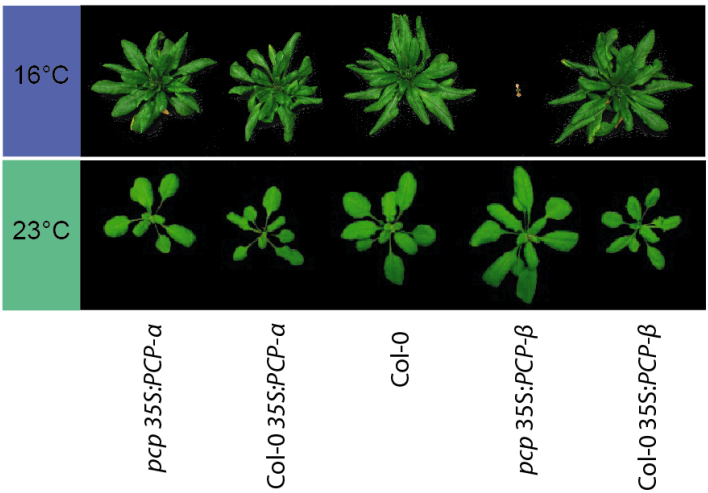

Figure 3.11 Picture of representative individuals of the overexpressing lines and Col-0 wild type, grown under long day at 16°C for 52 days (upper panel), and grown at 23°C for 22 days (lower panel).

Interestingly, the two isoforms had very different effects in *pcp-1* background. *PCP-α* could almost completely rescue the *pcp-1* phenotype at low temperature. The transformants in fact grew normally at the vegetative stage, produced rosette leaves, and even

made the switch to reproductive development, producing a stem with cauline leaves and flowers. However, *pcp-1 35S::PCP-α* plants were infertile and failed to set seeds (Figure 3.11 and 3.12). In contrast, overexpression of *PCP-β* was not able to suppress any of the developmental defects observed in *pcp-1* at low temperature (Figure 3.11) and looked like *pcp-1*. At 23°C degrees, all the overexpressing lines were almost indistinguishable from *pcp-1* (Figure 3.11), indicating that also in the *pcp-1* mutant background, overexpression of the *PCP* isoforms did not negatively affect plant growth and development.

Figure 3.12 Representative picture of an overexpressing *35S::PCP-α* line in *pcp-1* background grown at 16°C for over 2 months. A *pcp-1* plant of the same age and grown at the same conditions is shown for comparison (marked with a red arrowhead). The defective phenotype of *pcp-1* is mostly restored, in *pcp-1 35S::PCP-α* line, but plants fail to set seeds.

Transgenes are randomly inserted into the genome during transformation. The position in the genome can influence their

accessibility and consequently affect their expression. To ensure that *PCP-α* or *PCP-β* were effectively expressed in the transgenic lines, I quantified their expression by qPCR, using *UBC21* to normalize the results. Importantly, all lines, including *pcp-1 35S::PCP-β*, which did not rescue the *pcp-1* phenotype at 16°C, had 15 to 70 times higher levels of *PCP* than Col-0 (Figure 3.13). Interestingly, lines overexpressing *PCP-α* displayed higher *PCP* levels than those overexpressing *PCP-β*. Whether these differences in overall mRNA levels reflect differences in transcription (i.e. due to differences in accessibility of the transgenes) or RNA stability (i.e. faster degradation of *PCP-β* mRNA) remains currently unknown. In general, however, these findings confirm that *PCP-α* is the main *PCP* isoform.

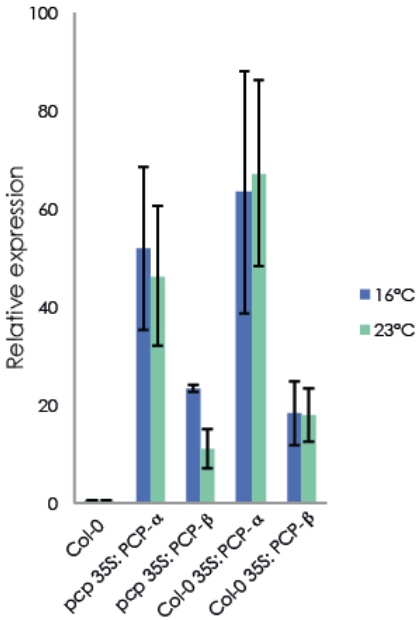

Figure 3.13 Expression analysis of *PCP* (*At2g18740*) in Col-0, and the overexpressing lines *35S::PCP-α* and *35S::PCP-β* in *pcp-1* and Col-0 background. For each line the test has been performed in plants grown at 16°C (blue) and 23°C (green). Error bars indicate standard deviation between 3 biological replicates.

To further investigate if *PCP-β* might modulate or compete with *PCP-α*, I crossed the *pcp-1 35S::PCP-α* with the *pcp-1 35S::PCP-β* transgenic lines and examined the F1 individuals that carried both constructs. All analyzed F1 plants looked similar to the *pcp-1 35S: PCP-α* parents, and did not set seeds at 16°C either. Therefore, I could only assign a function to *PCP-α* isoform.

PCP-α and PCP-β proteins differ only in their C-terminal amino acid sequence (Figure 3.14), suggesting that this part of the protein could be essential for its function. Alternatively it is also possible that the *PCP-β* transcript is not translated efficiently. To discriminate between these two possibilities the relative abundance of the two protein isoforms would need to be determined, for example using isoform-specific antibodies (not available) or reporter lines.

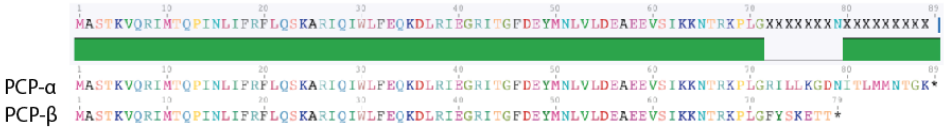

Figure 3.14 Alignment of the protein sequences of the two *PCP* isoforms.

3.3.4 *pcp-1* PHENOTYPES AT DIFFERENT AMBIENT TEMPERATURES

As already described above, *pcp-1* shows a very strong developmental defect only when grown at low ambient temperatures. However, *pcp-1* plants display some phenotypes also when grown at higher temperatures, and in this sub-chapter I will illustrate all the traits that displayed significant differences compared to WT.

3.3.4.1 *pcp-1* PHENOTYPE AT 16°C

At low ambient temperature most of the *pcp-1* plants die at the stage illustrated in Figure 3.3. The plants fail in generating properly developed leaves, and scanning electron microscopy (SEM) pictures reveal interesting details about the leaf-like organs that *pcp-1* produces at low temperature. The most affected organs show radialized symmetry. The stomata appear to be mostly properly formed, but trichomes frequently have more than 3 branches. (Figure 3.15)

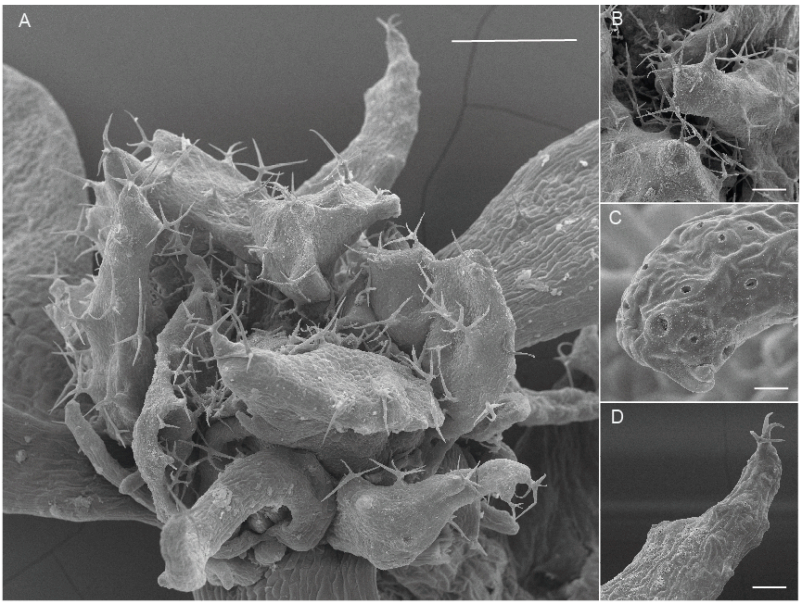

Figure 3.15 Scanning electron microscopy pictures of the *pcp-1* phenotype at 16°C. A. Overview of the meristem defects and the leaf-like organs. B,C,D close-up pictures of the radial organs. Stomata are usually properly formed (C), but many of the trichomes carry more than 3 branches (A,B,D). Bars: 600 μm (A), 200 μm (B), 30 μm (C), 100 μm (D).

Around 20-30% of individuals grown at 16°C do not die at the stage illustrated in Figure 3.3 but instead at approx. 60-70 days of age start to develop into small, bushy plants, which do not produce a WT like

rosette but rather develop more than one shoot at the same time (Figure 3.16).

Figure 3.16 ~25% of 65-80 days old *pcp-1* mutants slowly recover into a dwarf and bushy male-sterile plant when grown at 16°C. Bar = 1cm.

In comparison, a WT plant of the same age would be fully grown and starting the senescence process (Figure 3.17).

Figure 3.17 Representative picture of the comparison between a *pcp-1* mutant starting to develop into the bushy plant (left) and a Col-0 plant of the same age grown in the same conditions (right). Plants are 72 days old and were grown at 16°C long day.

The lack of a primary shoot in the *pcp-1* escapees at 16°C can possibly be explained by the fact that the mutant apparently produces multiple potential meristems in the “porcupine” stage as seen in cross sections of 60 days old plants (Figure 3.18). It seems possible that *pcp-1* at the time of recovery initiates multiple shoots from this reservoir of meristems but fails to determine a main shoot apical meristem (SAM).

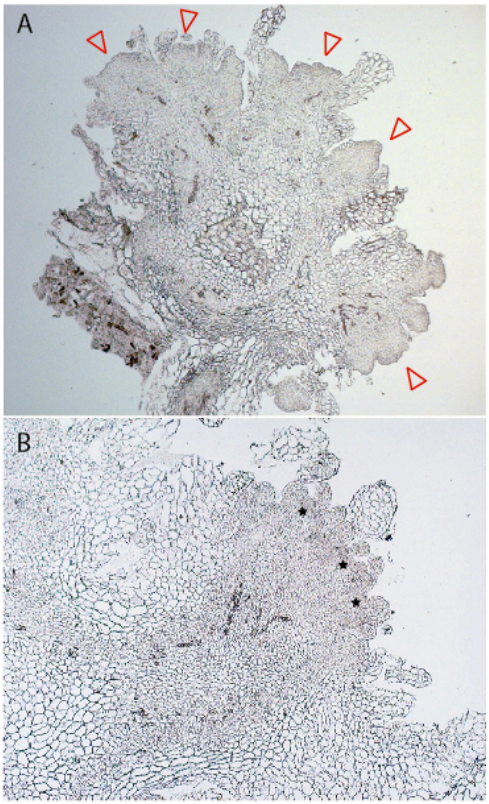

Figure 3.18 Cross section of the meristematic region of *pcp-1*. A. Cross section of a 67 days old *pcp-1* mutant (overview). Red arrowheads mark potential SAMs. B. Cross section of a 67 days old *pcp-1* mutant (magnification) multiple potential meristems are marked with black stars.

However, none of the *pcp-1* escapees at low temperature set seeds even when fully grown (Figure 3.19).

Figure 3.19 Representative picture of a fully grown *pcp-1* grown for over 3 months at 16°C. Plants fail in setting seeds. Bar = 1cm.

This can be easily explained by the failure of *pcp-1* stamens to properly elongate (Figure 3.20 A), making it impossible for the anthers to reach the stigma at the tip of the carpels. Also, it would appear as if these plants do not produce viable pollen grains. Consequently, SEM pictures failed to detect grains on the anthers (Figure 3.20 B). However, when pollinated manually with Col-0 pollen, *pcp-1* grown at low temperatures can produce viable seeds, indicating that the failure *pcp-1* escapees grown at 16°C are female fertile.

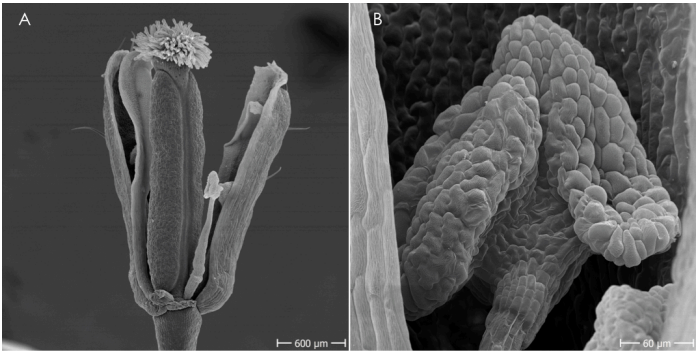

Figure 3.20 Representative SEM picture of a male sterile *pcp-1* flower grown at 16°C (A). Detailed image of the pollenless anther (B). Bar = 600 µm (A) and 60 µm (B).

Finally, the developmental defects of *pcp-1* are apparently not restricted to the shoot. The root apical meristem seems also affected and 14 days old *pcp-1* seedlings grown at 16°C on MS plates clearly show an arrest in root growth when compared with the WT Col-0 (Figure 3.21).

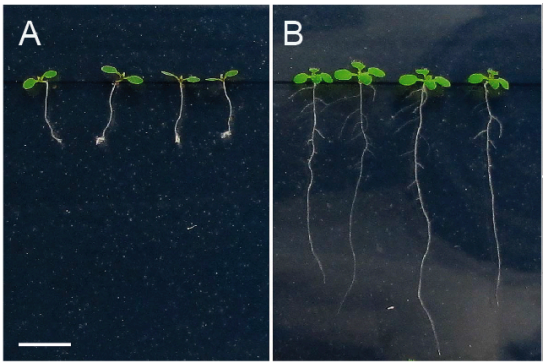

Figure 3.21 Representative picture of 4 individual 14 days old *pcp-1* seedlings grown at 16°C (A) and 4 Col-0 seedlings grown in the same conditions (B). Bar = 1cm

3.3.4.2 *pcp-1* PHENOTYPE AT 23°C

At high temperature, the *pcp-1* mutant grows similarly to WT. It produces leaves in a rosette and in both long days (LD) and short days (SD), the timing of flowering is not significantly different from Col-0 (using Welch's t-test) (Figure 3.22).

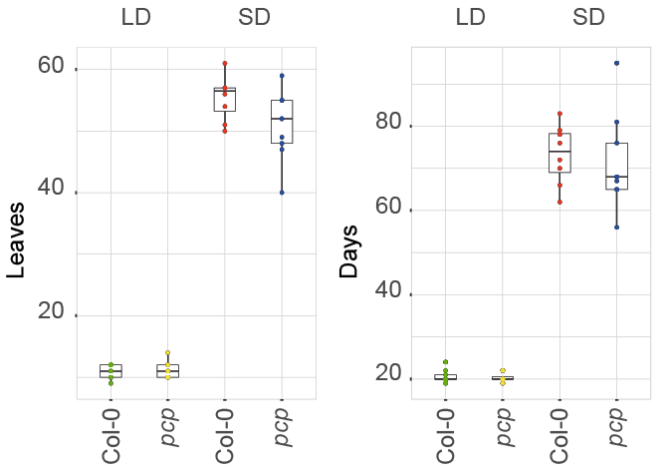

Figure 3.22 Flowering time phenotype measured at 23°C in number of leaves (left panel) or in days from germination (right panel) N. of plants tested ranges between 8 and 9 for each line in each condition.

Even though no differences were observed in flowering time, other rather minor phenotypes were observed at low frequency in *pcp-1* grown at 23°C LD when compared to WT. For example, *pcp-1* displayed a significantly elevated (p value = $1.47*10^{-4}$ using Welch's t-test) number of seedlings with three cotyledons in comparison with the WT (Figure 3.23 A). In the mutant, a significantly increased (p value = 0.03 using Welch's t-test) occurrence of flowers carrying 5 or

rarely 6 petals was observed (Figure 3.23 B) together with a disturbed phyllotaxis and an uneven distribution of siliques along the stem (Figure 3.23 C) (for details on the data collection see the Material & Methods section in this chapter).

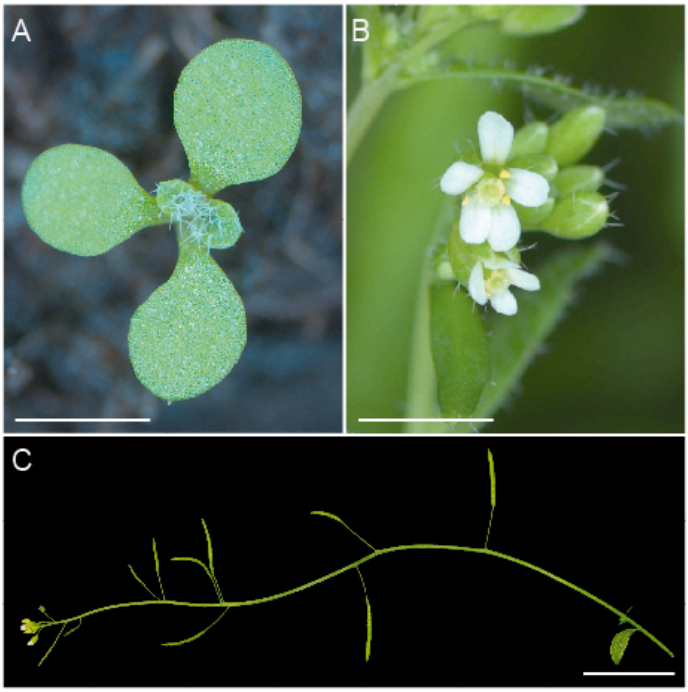

Figure 3.23 Developmental phenotypes of *pcp-1* grown at 23°C. A. Representative picture of a *pcp-1* seedling grown at 23°C carrying 3 cotyledons. B. Representative pictures of *pcp-1* inflorescence meristem and a flower with 5 petals. Representative picture of a *pcp-1* shoot grown at 23°C. Siliques are spaced unevenly along the stem and the angle between subsequent siliques is frequently incorrect. Bars =1cm.

Taken together, these data suggest a minor contribution of *PCP* to plant development even at 23°C. However, the random distribution of these events point to a possible regulatory role for PCP rather than a constitutive one.

3.3.5 *pcp-1* PHENOTYPE IS SENSITIVE TO TEMPERATURE FLUCTUATIONS

As already mentioned, when grown at 23°C *pcp-1* is overall very similar to WT. In order to test in which developmental stages *pcp-1* plants were sensitive to ambient temperature changes I performed shifting experiments from low to high temperature and vice versa and observed the phenotypes.

When the rosette area is considered, *pcp-1* is indistinguishable from WT when continuously grown at 23°C (Figure 3.24 A). However, when 14-day-old plants were shifted from 23°C to 16°C, WT continued to grow normally whereas growth of *pcp-1* slowed down significantly (Figure 3.24 B), indicating that the *pcp-1* mutant is actively responding to environmental changes. This further indicates that PCP is not only required to establish basic patterns during the early seedling stage but is necessary throughout a plant's live to ensure proper growth and development.

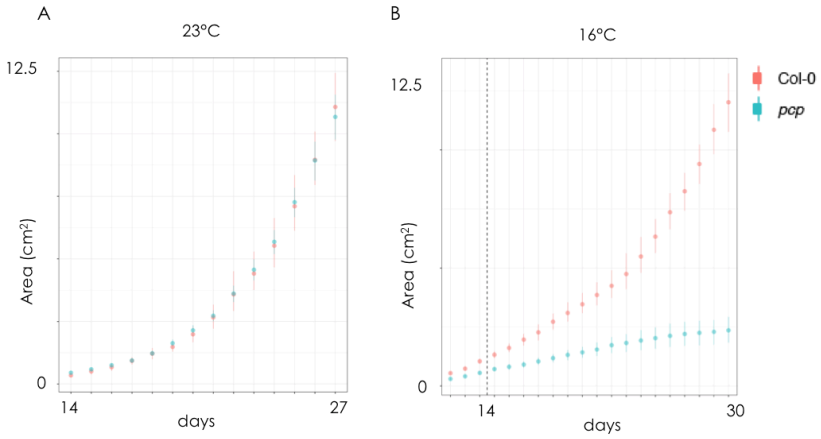

Figure 3.24 Rosette phenotype of *pcp-1*. A. Total average area in cm^2 of Col-0 (red) and *pcp-1* plants (blue) grown at 23°C. The measurement has been performed every day starting from the 14th day after germination for 2 weeks until the 27th day after germination. B. Total average area in cm^2 of Col-0 (red) and *pcp-1* plants (blue) of seedlings grown at 23°C for the first 13 days after germination and then shifted to 16°C for 17 more days, until the 30th day after germination.

Importantly, *pcp-1* is sensitive to temperatures not only during the vegetative stage. For example, male sterility of *pcp-1* grown at 16°C phenotype is rescued when plants are shifted to more permissive temperatures (Figure 3.25 A), and plants grown at 23°C arrest their development or turn male sterile when shifted to lower temperatures (Figure 3.25 B).

These results suggest a role for PCP during the whole life cycle and that plants can sense and adjust rapidly to changes in ambient temperature.

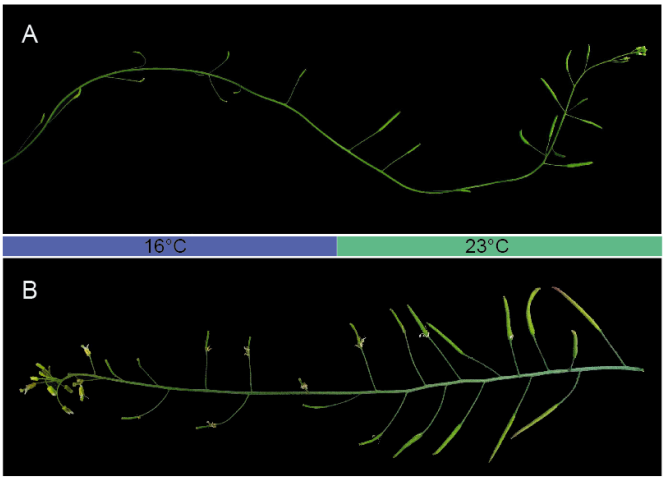

Figure 3.25 Representative picture of a *pcp-1* shoot grown at 16°C and then shifted to 23°C (A), and one of a *pcp-1* shoot grown at 23°C and then shifted to 16°C (B).

3.3.6 MERISTEM DEFECTS IN *pcp-1*: A ROLE FOR WUS AND CLV3

The entire aerial part of the plant derives from cells in the shoot apical meristem (SAM). For this reason the correct formation of the SAM is crucial for a normal plant development. The maintenance of the SAM is controlled by an intricate feedback loop that involves *WUS* and *CLV3*. Briefly, *WUS* is expressed in the organizing center of the SAM and establishes stem-cell fate by promoting *CLV3* expression in the central region of the meristem. *CLV3* encodes a mobile peptide that represses *WUS* expression in the underlying cells and promotes cell differentiation (Reviewed in Doerner 2003, Clark 2001, Sparks *et al.* 2013, Heidstra and Sabatini 2014). To better understand the cause for the meristem defects in *pcp-1*, I investigated the spatial and temporal expression of *WUS* and *CLV3* by (semi-) quantitative RT-PCR and RNA *in situ* hybridization.

qPCR and semi-qPCR performed using RNA extracted from manually dissected apices of plants grown at 16°C clearly showed that the expression of *WUS* was elevated in *pcp-1* in comparison with Col-0 (Figure 3.26). Probably not surprisingly because of the regulatory feedback loop between the two genes, the expression of *CLV3* was reduced in the mutant.

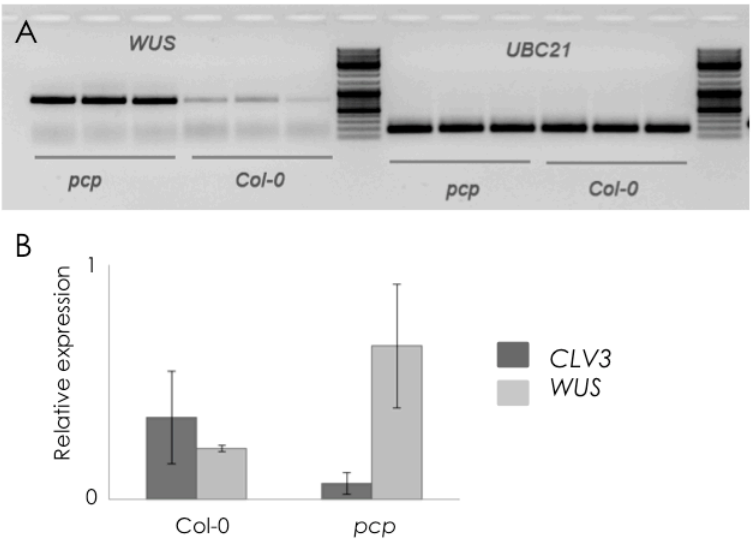

Figure 3.26 Quantification of *WUS* and *CLV3*. A. Semi-quantitative PCR of *WUS* and *CLV3* performed in 24 days old Col-0 and *pcp-1* plants. 3 biological replicates are shown and amplification of *UBC21* is reported for comparison. B. Expression analysis of *CLV3* and *WUS* in 12 days old Col-0 and *pcp-1*. Error bars indicate standard deviation between biological replicates. In both analyses the RNA has been extracted from pools of 15 apices for each biological replicate and 3 biological replicates were considered. Plants were grown at 16°C.

These findings suggest that *WUS* and *CLV3* are misexpressed in *pcp-1*, which could in part explain the *pcp-1* mutant phenotype. Alternatively, the differences observed in *WUS* and *CLV3* expression

could reflect differences in the number of cells that express these genes, rather than true expression differences.

To discriminate between these two possibilities, I next analyzed expression of *WUS* and *CLV3* in *pcp-1* by RNA *in situ* hybridization. The principle behind *in situ* hybridization is the specific annealing of a labeled antisense probe designed on the gene of interest in a fixed tissue, followed by the visualization of the probe. This technique can at best be considered semi-quantitative, because the intensity of the colorimetric reaction used to detect the probe can have biases due to practical challenges. However, the signal obtained provides information on the spatial activation of a specific gene.

Using RNA *in situ* hybridization I examined the spatial expression of *PCP, WUS,* and *CLV3* in longitudinal section of fixed apices of Col-0 or *pcp-1* plants grown at 23°C for 9 days or grown at 23°C for 9 days and then shifted to 16°C for 3 additional days. The *PCP* antisense probe consistently detects the *PCP* transcript in the WT but not in the *pcp-1* mutant, demonstrating the accuracy of the probe and the experimental conditions (Figure 3.27). Furthermore, these results demonstrate that *PCP* is expressed in the SAM and leaf primordia.

As expected, *WUS* was detected exclusively in the organizing center of the SAM, underneath the stem cells in all four samples. In contrast, *CLV3* expression is more dynamic and its expression domain, which in WT can be detected in the apex of the SAM, appears to be shifted downwards in *pcp-1* at low temperature (Figure 3.27). *CLV3* expression is restricted to the stem cells in *pcp-1* at 23°C (Figure 3.27), suggesting that PCP might participate in the temperature-dependent spatial control of *CLV3* expression.

Importantly, the shift in *CLV3* expression is not always complete in all the *pcp-1* individuals. In some cases, I could still detect *CLV3* in a part of the L1 and L2 layers, but the shift downwards can be observed

in all the samples analyzed (Figure 3.28) and is also clear in horizontal consecutive sections (Figure 3.29).

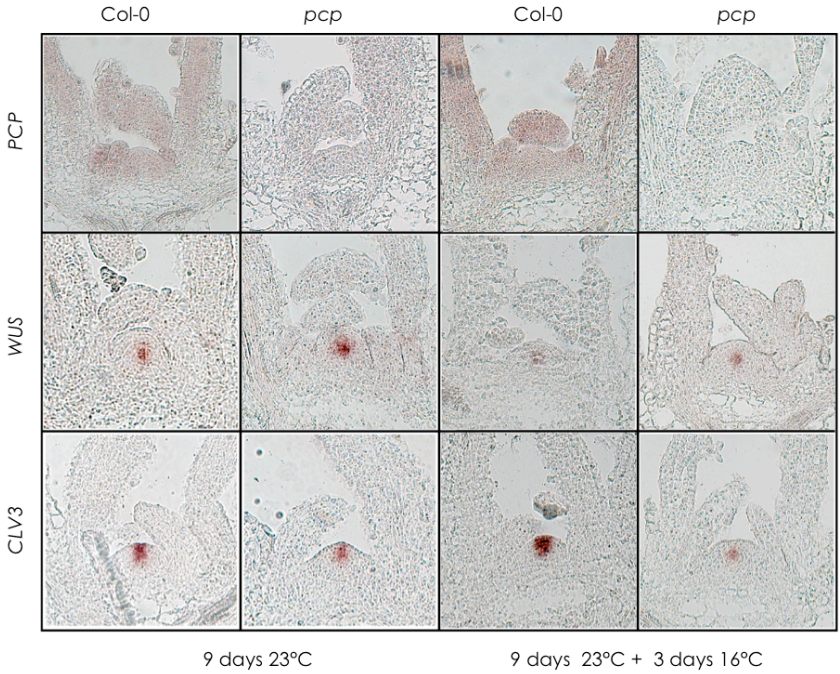

Figure 3.27 Detection of *PCP, WUS,* and *CLV3* expression domains by RNA *in situ* hybridization in 9 or 12 days old Col-0 and *pcp-1* in longitudinal section of the apices.

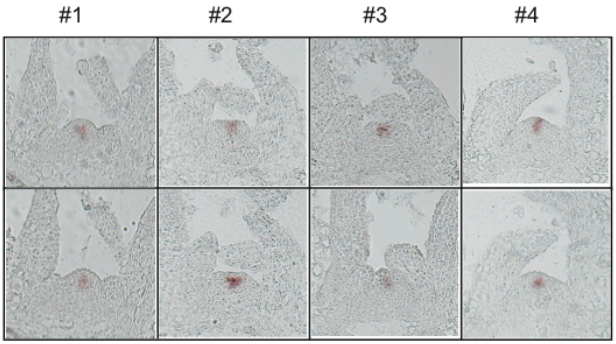

Figure 3.28 Detection of *CLV3* expression domain by *in situ* hybridization in 4 individual 12 days old *pcp-1* plants grown at 23°C for 9 days and shifted to 16°C for 3 more days. For each sample, 2 consecutive longitudinal sections of the apices are reported.

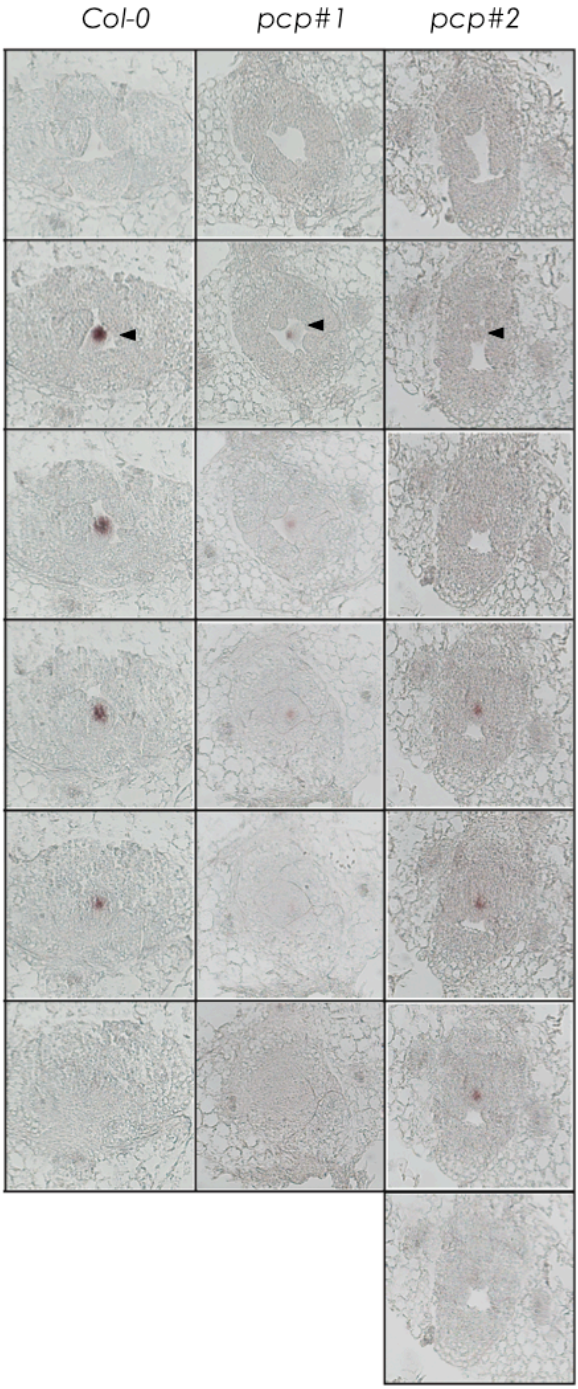

Figure 3.29 Detection of *CLV3* expression domain by *in situ* hybridization in Col-0 and two *pcp-1* individuals. Plants were grown at 23°C for 9 days and shifted to 16°C for 3 days. For each sample, all the consecutive horizontal sections of the apices are reported, including the first slide before the tip of the apex (marked with a black arrowhead) and the first slide without the *CLV3* signal. In the first *pcp-1* individual (#1) *CLV3* signal can only be detected at a small portion of the tip of the apex. In the second individual (#2), the signal cannot be detected before the third consecutive slide from the tip of the apex.

To assess the function of *PCP* in specific areas of the SAM I expressed the fully spliced active isoform *PCP-α* using *pWUS* and *pCLV3*. Interestingly, in T1 transformants neither of the constructs was able to rescue *pcp-1* phenotype at 16°C, suggesting that *PCP* needs to be expressed more broadly and has a more general upstream role in SAM maintenance.

The link between PCP and the misexpression of *CLV3* or the upregulation of *WUS* has not yet been clarified in detail, but the lack of stability in this important loop can certainly explain part of the meristem defects affecting *pcp-1* at low temperature.

3.3.7 A CLOSELY RELATED GENE TO *PCP*

PCP is part of the Small nuclear ribonucleoprotein (Sm) gene family, which in *Arabidopsis* consists of 42 members. Cao and colleagues (2011) have identified a gene*, At4g30330*, that is particularly closely related to *PCP* and whose amino acid sequence differs from PCP only for 2 amino acids (Figure 3.30).

Figure 3.30 Alignment of the protein sequences of PCP-α and At4g30330.

Interestingly, the relative expression of this gene is upregulated in increasing the ambient growth conditions, a trend opposite to that observed in *PCP* (Figure 3.31).

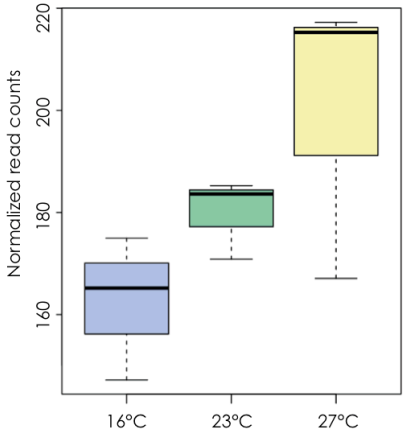

Figure 3.31 Boxplot showing the relative expression obtained from the RNA-seq analysis at 16°C (blue), 23°C (green) and 27°C (yellow) of *At4g30330*.

To determine if this PCP-like gene might also participate in the temperature-dependent regulation of plant development I isolated two T-DNA lines, SALK_139504 and SALK_139760, which both carry insertions in the 3' UTR of At4g30330. Unfortunately, both of these lines still expressed *At4g30330* at levels similar to Col-0 (Figure 3.32), indicating that these two lines are not effective knock out mutants. Consequently, it was not surprising that neither of these two lines displayed a phenotype at any of the temperatures analyzed and no function could be assigned to this PCP-like gene. To address the function of this gene, true loss-of-function alleles would need to be generated, for example by CRISPR/Cas9 genome editing.

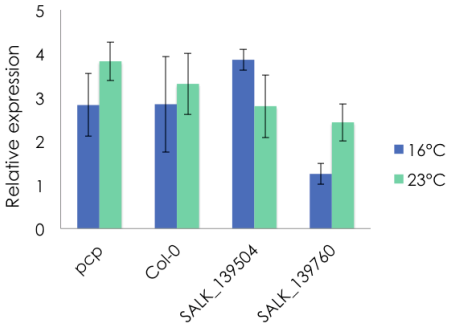

Figure 3.32 Relative expression analysis of *At4g30330* in *pcp-1*, Col-0 and two SALK lines carrying an insertion in *At4g30330* at 16°C (blue) and 23°C (green). Error bars represent the standard deviation between the 3 biological replicates considered.

3.3.8 RNA-SEQ AND CANDIDATES GENES

In collaboration with N.D and I.S who analyzed the data, I performed RNA-seq of *pcp-1* and Col-0 seedlings grown at 16°C and 23°C and plants shifted from one temperature to the other as listed in materials and methods (chapter 3.5). The aim of this experiment was to identify genes regulated by *PCP* to understand both the origin of the developmental defects in the *pcp-1* mutants at low ambient temperature and the rescuing mechanism coming into play at higher temperature. A differential expression (DE) analysis was conducted to characterize (i) the genotype effect, (ii) the temperature effect and (iii) their possible interaction (see Methods in chapter 3.5).

RNA-seq analysis revealed that the effect of genotype contributes to a differential expression of 1440 genes, at a False-Discovery Rate (FDR) cutoff of one percent and with expression log2 fold-changes greater than 0.5; *i.e.* when the two backgrounds, *pcp-1* and Col-0 are compared regardless of in which temperature the plants were grown at (Figure 3.33, green circle). With regards to the temperature-specific effect, when samples are grouped based on the temperature (either

16°C or 23°C) at collection (i.e. after the temperature shift), regardless of the genotype or the temperature in which they have been sown, 1903 genes were found to be differentially expressed (Figure 3.33, orange circle). Finally, when testing the interaction between *pcp-1* and the final temperature (at the time of sample collection), 319 differentially expressed genes were identified (Figure 3.33, purple circle).

Overall, my results indicate that the predominant gene expression differences are due to the combination of *pcp-1* genotype and the effect of cold temperature, which is in agreement with the temperature-dependent nature of the phenotype. Furthermore, expression patterns of *pcp-1* in warm conditions are more similar to the expression patterns occurring in the WT at low temperatures rather than in WT at 23°C.

By the overlap of these three sets (genotype, temperature, and the interaction effects) of genes I obtained a list of 47 genes (Figure 3.33) that react to both the temperature and the genotype in an antagonistic or synergistic way. The goal was in fact to identify genes differentially expressed between *pcp-1* and Col-0 that also show differences between temperatures. Many of the 47 identified genes are involved in basic biological processes like transcription regulation (for instance, *AT1G09250, AT3G02790, AT3G58120*), or in temperature responses, like *LHY, (AT1G01060), ADS1, (AT1G06080)* and *AT5G51440* (Table 3.3).

In addition, the list also contains *CML42 (AT4G20780)* and *NYC1 (AT4G13250),* genes involved in trichome branching and in chlorophyll catabolic process, respectively. The differential expression of these genes could explain the dark green phenotype observed in *pcp-1* young seedlings grown at low temperature and the presence of trichomes with more than 3 branches.

Taken together, the list of 47 genes affected by both variables - genotype, temperature, and their interaction - in the analysis can explain part of the downstream ultimate effects on *pcp-1* mutants, but not their cause. Similar results were obtained when the analysis was performed at the transcripts rather than at the gene level. There was no evidence of a causal isoform.

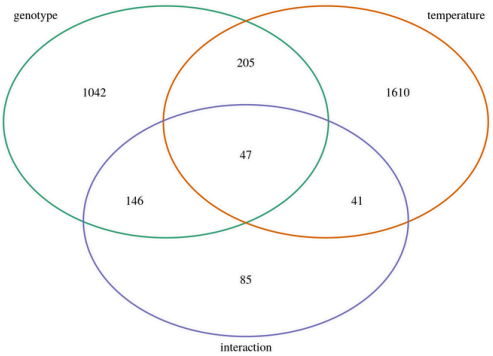

Figure 3.33. Venn diagram showing the number of differentially expressed gene sets for temperature, genotype and their interaction.

Table 3.3 List of the 47 genes reacting to genotype and temperature

ID	Name	Involved in
AT1G01060	LHY	Cold stress
AT1G02800	CEL2	Cell wall organization
AT1G06080	ADS1	Cold stress
AT1G09250	AIF4	Regulation of transcription
AT1G11850		Unknown
AT1G14280	PKS2	Hypocotyl photomorphism
AT1G23020	ATFRO3	Oxidation-reduction process
AT1G31290	AGO3	Regulation of transcription
AT1G43800	FTM1	Fatty acid biosynthetic process, oxidation-reduction process
AT1G53887		Molecular function in guard cells
AT1G73805	SARD1	Salicylic acid synthesis
AT2G02100	LCR64	Defense response
AT2G03760	ATSOT1	Brassinosteroid metabolic process, defense response
AT2G20590		Biological process in guard cells

AT2G21650	RSM1	Early morphogenesis
AT2G27920	SCPL51	Proteolysis
AT2G41730		Anaerobic respiration
AT2G43510	TI1	Defense response
AT2G46790	PRR9	Circadian rhythm
AT2G47000	ABCB4	Auxin efflux and influx
AT2G47440		Protein folding
AT3G02790	MBS1	Transcription regulation
AT3G10200		Methyltransferase activity
AT3G13080	ABCC3	Transmembrane transport
AT3G13520	AGP12	Biological process, transmembrane
AT3G16530		Defense response
AT3G18290	BTS	Zinc ion binding
AT3G19680		Putative protein
AT3G25882	NIMIN-2	Systemic acquired resistance
AT3G46370		Phosphorylation
AT3G47800		Galactose metabolic process
AT3G48090	EDS1	Defense response
AT3G50930	ATBCS1	Cell death
AT3G51290	APSR1	Phosphate starvation response
AT3G58120		Regulation of transcription
AT4G10270		Response to wounding
AT4G13250	NYC1	Chlorophyll catabolic process
AT4G15700	GRXS3	Cell redox homeostasis
AT4G16680		RNA helicase activity
AT4G20780	CML42	Trichomes branching
AT4G21730		Unknown
AT4G22690		Secondary metabolite biosynthetic process
AT5G45820		Intracellular signal transduction
AT5G46240	KAT1	Regulation of ion transmembrane transport
AT5G51440		Response to heat
AT5G53450	ORG1	Gene expression
AT5G66080	APD9	Protein defosphorilation

Next, I generated a hand-curated list of 287 genes involved in meristem identity and maintenance, in leaf primordia initiation, leaf morphology, and genes whose mutants are known for being involved in developmental defects based on a literature survey. From these 287 potential candidates I filtered those whose expression was significantly altered between *pcp-1* and Col-0, or whose expression differs in *pcp-1* at different temperatures, especially when the same trend was not occurring in the WT at the same conditions.

Using this approach I identified three interesting candidate genes for further analyses. The first one is *BREAST CANCER ASSOCIATED RING 1* (*BARD1*), which encodes a protein known to bind to the *WUS* promoter and whose overexpression represses *WUS* expression. Loss of function mutations in *BARD1* cause defects in meristem organization due to failure to repress *WUS* expression (Han *et al*. 2008) that are similar to those observed in *pcp-1*. *BARD1* expression was significantly lower at 16°C temperatures in the *pcp-1* mutant than at 23°C (Figure 3.34). In contrast, no significant differences in *BARD1* expression were detected between temperatures in Col-0.

A second potentially interesting gene is *YABBY3 (YAB3)*, whose expression is higher in Col-0 when compared to *pcp-1*. YAB3 has transcription factor activity and is involved in specifying abaxial cell fate (Siegfried et al. 1999). However, no significant differences in its expression were detected in *pcp-1* grown at different temperatures.

The third potential candidate is *AINTEGUMENTA (ANT)*, which is required for control of cell proliferation. Expression was originally described to be limited to the chalaza and the floral organ primordia (Elliott et al. 1996, Klucher et al. 1996), but *AINTEGUMENTA-LIKE (AIL)* family of transcription factors are expressed in all dividing tissues in the plant, where they have central roles in developmental processes such as meristem maintenance, organ positioning, and growth (Reviewed in Horstman et al. 2014). When the effects of the combined loss of function of *ANT*, *AIL6* and *AIL7* are investigated in the *ant-4 ail6-2 ail7-1* triple mutants, the shoot apical meristem terminated after the production of a few leaves (Mudunkothge and Krizek 2012). Defects in the meristems are associated with reduced cell division, stem cell differentiation and altered expression of meristem regulators such as *WUS*, *CLV3* and *SHOOT MERISTEMLESS* (*STM*) (Mudunkothge and Krizek 2012).

In this study, *ANT* expression is higher in WT than in *pcp-1*, and in the mutant its expression is higher in warmer temperature than at 16°C.

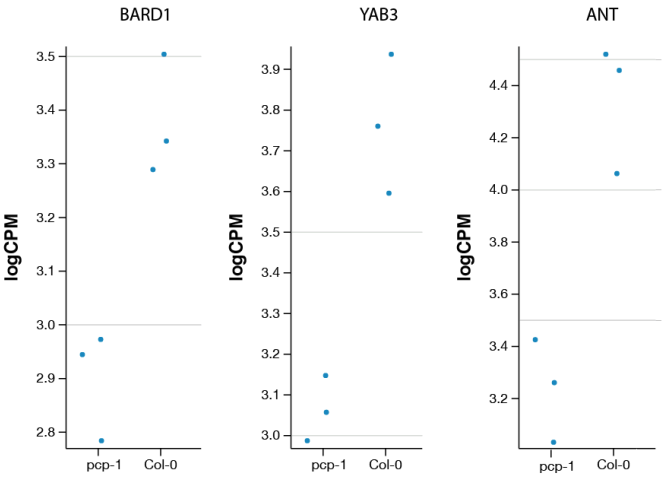

Figure 3.34. Dot plot showing the differential expression of three candidate genes between *pcp-1* and Col-0 at 16°C. Each dot represents a biological replicate.

In addition to the analysis of significant changes in gene expression and transcripts abundance presented above, we also performed a differential exon usage analysis. We identified 53 and six genes with differences in exon usage between mutant and WT at 16°C and 23°C, respectively. Of these, two genes were differentially spliced regardless of the temperature: *PCP* and *U1 SMALL NUCLEAR RIBONUCLEOPROTEIN-70K (U1-70K; At3g50670*). U1-70K is a fundamental component of the spliceosome and is involved in the initial definition of the 5′ splice site in both constitutive and alternative splicing events (Mount *et al*. 1983, Black *et al*. 1985, Rosbash and Seraphin 1991). Exon usage for *U1-70K* differs between *pcp-1* and Col-0 as well as *PCP* at both temperatures, suggesting that PCP

might regulate splicing on the global scale by modulating U1-70K function.

We also tested the differential splicing events that occurred within each genotype between the two temperatures tested, 16°C and 23°C, by considering common exons with differential level of expression. In WT, we detected statistically significant differences in splicing events between 16°C and 23°C in 1867 transcripts, whereas in the *pcp-1* mutant this number was reduced to 22. Only 10 genes were differentially spliced regardless of the temperatures. This data suggest that PCP has a role in the AS and that its absence cause the loss of a great number of splicing events. Interestingly, when the overall significant changes in AS events are considered, *pcp-1* at 23°C shows a more similar pattern to Col-0 at 16°C (Figure 3.35).

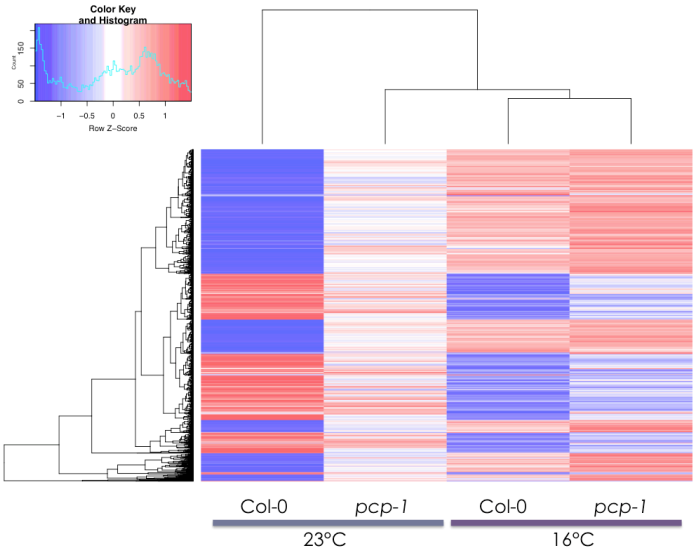

Figure 3.35. Heat map representative of clusters of expression values of significantly differentially used exons between Col-0 and *pcp-1* at the two temperatures considered.

I further performed an initial analysis of two candidates described above, *BARD1* and *U1-70K*, to test if dysregulation of these genes could contribute to the *pcp-1* mutant phenotype. However, my preliminary results show that overexpression of *BARD1* CDS or *U1-70K* CDS under control of the 35S promoter was unable to rescue the *pcp-1* phenotype at low temperature. For this reason the role of these two genes in mediating plant development downstream of PCP is still unknown.

3.4 DISCUSSION

Plant development and growth are influenced by environmental factors that may display high degrees of variability, both in space and in time. The ability to modulate specific molecular programs in response to changes in the environment results in phenotypic plasticity, which is very important to guarantee evolutionary success. Plants are highly responsive to temperature fluctuations and are capable of modulating organogenesis and growth rate to adapt to novel conditions. Recent findings suggest that temperature-regulated alternative splicing (AS) plays a critical role in controlling the plant response to variations in temperature at the molecular level, allowing quick adaptation to the current conditions (Verhage *et al.* 2017).

AS, by producing more than one mRNA from the same gene, allows organisms to increase the diversity of their transcriptomes and work as regulatory mechanism when unstable mRNA are preferentially spliced and subsequently degraded. Genome-wide transcriptome mapping has shown that in plants the extent of AS ranges between 42% to 61% (Klepikova *et al.* 2016, Marquez *et al.* 2012, Reddy *et al.* 2013).

Some examples of genes involved in temperature-dependent AS in the flowering pathway are *FLM* and *MAF2* (Rosloski *et al*. 2013, Posè

et al. 2013, Airoldi *et al.* 2015). Many components of the circadian clock are also affected in their AS by temperature fluctuations. The effects of alternatively spliced isoforms in either flowering time or clock genes are well known. In contrast, the splicing factors involved in the temperature-dependent AS decisions are largely unknown.

In this thesis I describe the function of PORCUPINE (PCP), a *bona fi*de AS regulator, which is essential for correct development of *Arabidopsis thaliana* at low temperature.

In plants, genome and transcriptome projects are accelerating the research on splicing-related genes. In 2004, 74 small nuclear RNA (snRNA) and 395 genes encoding splicing-related proteins were computationally identified in the *Arabidopsis thaliana* genome (Wang and Brendel 2004). With the aim to identify thermoresponsive AS factors, I first generated RNA-seq libraries from *Arabidopsis* seedlings grown at 16°C, 23°C or 27°C. Subsequently I identified those candidates listed in (Wang and Brendel 2004) that were significantly and consistently differentially expressed in response to temperature changes, eliminating all genes whose expression was not changing significantly or changing in complex patterns (for example being highly expressed at both 16°C and 27°C in comparison to 23°C). This analysis revealed 25 candidates whose role could be relevant at either high or low ambient temperature, assuming that transcript expression and protein abundance corresponded. Among the 25 candidates 10 genes were downregulated in increasing the temperature and 15 were upregulated (Figure 3.2). Next, I embraced a reverse genetics approach aimed to find what phenotypes arise as a result of knock out mutations on the candidates and determine their role. I first selected 10 candidates (4 up and 6 down regulated in temperature increases) (Table 3.1) and determined at the phenotype of T-DNA ko mutant lines homozygous for the insertion. All mutants were grown at 16°C, 23°C or 27°C together with Col-0 and I observed

a striking difference in development in SALK_089521, which carries an insertion in *At2g18740*, specifically at 16°C, whereas no obvious phenotype was detectable at 23°C (Figure 3.3). Because of the "spiky" phenotype of the mutant, the gene has been named *PORCUPINE (PCP),* and the SALK line *pcp-1*. To prove *PCP* as the causal gene for the phenotype I tested other available independent T-DNA lines, finding a second potential allele, SALK_119088 or *pcp-2,* which displayed a similar phenotype (Figure 3.7). Developmental defects in *pcp-2* at low temperatures were not as extreme as the ones observed in *pcp-1.* In the second allele the knockout was not complete but reduced mRNA levels to approx. 15 to 20% in comparison to the WT (Figure 3.9). The remaining *PCP* expression in *pcp-2* that can explain the phenotype is probably due to the position of the T-DNA insertion. In fact in *pcp-1* the insertion occurs in the second exon, whereas in *pcp-2* the insertion occurs in the fifth intron, and possibly by splicing a small percentage of fully spliced *PCP* can still be produced. The identification of a second independent allele was already a good indication that the loss of *PCP* was causing the phenotype. However, final proof of this was obtained by a complementation test, in which a transgene carrying a genomic rescue fragment of *PCP* including its own promoter in *pcp-1* completely rescued the phenotype at 16°C in T1 generation (Figure 3.8).

At high temperatures, *pcp-1* is almost indistinguishable from the WT, and transforming *pcp-1* is relatively easy because the mutant grow, flowers and sets seeds. However, minor phenotypic anomalies were observed also at 23°C. The occurrence of seedling with 3 cotyledons, and flowers with 5 (and rarely 6) petals were significantly higher in the mutant than in the WT. This could be due to an increase in meristem size and *clv3* loss of function mutants show in fact similar traits (Clark *et al.* 1995). Furthermore, siliques were unevenly spaced along the

shoot in *pcp-1*, indicating a distorted phyllotaxis (Figure 3.23). Key to the establishment of stable phyllotactic patterns is the distribution of auxin through the efflux carrier PIN1 (Jönsson *et al.* 2006), so PCP could somehow be involved in regulating one or more players in the auxin pathway as well.

At low temperatures, *pcp-1* cotyledons appear darker than in WT, and the meristem fails to produce properly developed leaves. In fact, some of the most strongly affected leaves show signs of radialization (Figure 3.15). The majority of individuals grown at 16°C die in the stage pictured in Figure 3.4, however 20-30% of individuals at an age of approx. 60-70 days, when a WT would be already starting to senest (Figure 3.17), develop into a bushy and dwarf plant that will remain male sterile (Figure 3.16). It seems possible that these *pcp-1* escapees at the time of recovery fail to determine a main shoot apical meristem (SAM). The lack of a primary shoot in the *pcp-1* escapees at 16°C can possibly explain the bushy phenotype, and could be dues to the fact that the mutant apparently produces multiple potential meristems in the "porcupine" stage as seen in cross sections of 60 days old plants (Figure 3.18). The male sterility instead can be explained by the *pcp-1* stamens that had short filaments and small anthers without pollen grains (Figure 3.20).

With the aim to test whether both the two known *PCP* isoforms (Ito *et al.* 2010) encode for functional PCP proteins able to rescue *pcp-1* phenotype, I introduced their fully spliced CDSs under the constitutive 35S promoter. Interestingly only one of the isoforms, PCP-α was able to complement most of the phenotype in *pcp-1* background at 16°C, but the plants were still male sterile, whereas *PCP-*β, whose transcript differ from *PCP-*α only at the 3' end could not rescue the *pcp-1* phenotype at all (Figure 3.11).

One hypothesis was that part of the phenotype, like the male sterility could be due to lack of the second isoform, *PCP-*β. Alternatively the

35S promoter might not be active enough in reproductive tissues. By crossing the two overexpressing lines in *pcp-1* background I was never able to obtain a fully recovered plant, so the role of *PCP-β*, if any, could not be determined. These results also suggest that the C-terminal of PCP might be essential for its function.

Misexpression of a wild-type gene product can also cause mutant phenotypes. To evaluate the consequences of ubiquitous overexpression of either *PCP* isoforms I also introduced the constructs described above in Col-0 background, but no effect in development has been detected irrespectively of the temperature at which the primary transformants were grown (Figure 3.11).

Interestingly, growth can be arrested at any developmental stage by shifting *pcp-1* mutant plants grown at 23°C to lower temperatures (Figure 3.24 and 3.25). Reversely, by exposing the *pcp-1* mutants to 23°C the growth arrest occurred at low temperature can be rescued (Figure 3.24). These finding prove that PCP is essential in all developmental stages at low temperatures. Furthermore, the fast reprogramming suggests the existence of one or more factors that are partially redundant with PCP and are able to take over part of PCP function at high temperature. A candidate for this function could be a gene closely related to *PCP, At4g30330,* whose amino acid sequence differs from that of PCP only in 2 amino acids (Figure 3.30) and whose expression has an opposite trend than *PCP* (Figure 3.31). To address the function of this PCP-like gene in the temperature-dependent regulation of plant development, loss-of-function alleles need to be developed, because unfortunately none of the T-DNA lines available for this gene are complete knock out lines.

RNA-Seq analysis on the differential exon usage supports a role for PCP in AS regulation. In fact, the number of genes differentially spliced across temperatures in *pcp-1* mutant drop drastically to just

22 in comparison with the WT, where over 1800 genes have been identified as temperature-dependent differentially spliced.

At low temperature, the AS events occurring in *pcp-1* and in the WT are more similar than they are at warm temperature (Figure 3.35). On the contrary, the phenotype differs greatly only at low temperature.

When gene expression analysis is considered, expression patters of *pcp-1* at 16°C are very distinct in comparison with the rest of the samples. These results suggest that at low temperature only relatively few AS events differ between the WT and mutant, but that their effects seem to affect a very large number of genes and lead to their differential expression first, and eventually to the differences in phenotype. Conversely, at high temperature regardless the great diversity of AS events between the WT and *pcp-1* the final effect on the plant is minimal.

The literature is rich in mutants showing developmental defects in lateral organ formation regardless of the temperature conditions. Among many others, loss of function of *KANADI* family genes (Emery *et al.* 2003) cause adaxialization of lateral organs, as do gain-of-function alleles of members of the class III HD-ZIP gene family, like *PHAVOLUTA* and *PHABULOSA* (McConnell *et al.* 2001). However, none of the well characterized genes involved in the differentiation of the abaxial and adaxial leaf blades was found differentially expressed (or spliced) specifically in *pcp-1* at low temperature.

The entire aerial part of the plant derives from cells in the shoot apical meristem (SAM), and its correct formation is crucial for normal plant development. The maintenance of the SAM is controlled by an intricate feedback loop that involves *WUS* and *CLV3*, which are expressed in the organizing center of the apex or at its tip, respectively (Figure 3.27). Probably due to the low representation of their expression domains in comparison with the tissues used for the RNA-seq experiments (whole seedling) their expression levels did not

show significant differences. However interestingly, when RNA was extracted from apices and analyzed by qPCR, *WUS* expression was elevated in *pcp-1* when compared to WT (Figure 3.26), but no substantial differences were found in the dimensions or position of its expression domain (Figure 3.27). Conversely, expression of *CLV3* was reduced in the mutant, probably because of the regulatory feedback loop between the two genes, and *CLV3* expression domain appeared shifted downwards in *pcp-1* at low ambient temperatures, (Figure 3.27). The link between PCP and the misexpression of *CLV3* or the upregulation of *WUS* has not yet been clarified in detail, but the lack of stability in this important loop can certainly explain part of the meristem defects affecting *pcp-1* at low temperature.

Taken together, these findings demonstrate the importance of temperature-dependent AS in development and establish *PCP* as an important regulator of this process.

3.5 MATERIALS AND METHODS

PLANT MATERIAL AND GROWTH CONDITION

T-DNA lines listed in Table 3.1 and in Table 3.2 were obtained from the Nottingham *Arabidopsis* Stock *Centre* (*NASC*). Homozygosity of mutants was verified via polymerase chain reaction (PCR) genotyping (primers are listed in Table 3.7). Seeds were sterilized and sown as previously described in chapter 2.4.1.

RNA-seq libraries to analyze changes in expression in AS candidates in Col-0 were obtained from plants grown for 9 days at 23°C LD and then shifted to 16°C, 27°C, or kept at 23°C LD for 3 days. To analyze changes in expression and in exon usage in Col-0 and *pcp-1* at 16°C and 23°C, 9 days old and 12 days old plants were used. Col-0 and *pcp-1* seedlings were grown at 16°C and at 23°C for nine days and

either harvested immediately (9 day samples) or after a plants had been shifted to the other temperature for an additional three days (12 day samples). Three pools (biological replicates, grown at the same time) of 10 seedlings each were collected at zeitgeber time 6 and snap frozen in liquid nitrogen for each sample.

To analyze *WUS* and *CLV3* expression, RNA extraction was performed on manually dissected apices collected at zeitgeber time 6 from 12 days old plants grown at 16°C. Three pools of 15 apices (biological replicates), were considered for each line (Col-0 and *pcp-1*). All plants were all grown in long days (LD) condition in different temperatures as reported, with the exception of the flowering time phenotyping test of *pcp-1* and Col-0 at 23°C that was performed also in short days (SD), as described in the results.

TRANSGENIC PLANTS

The genomic sequence of *PCP*, CDS of *PCP-*α and *PCP-*β were amplified by PCR using Phusion DNA polymerase (New England Biolabs) (Table 3.7), and assembled into a "C" GreenGate module. The Green Gate system was employed as described in chapter 2, to obtain final vectors. GC004 was obtained with a final GreenGate reaction performed using modules A: empty, B: empty, the C module carrying the full genomic *PCP* region including 232 bp sequence upstream of the ATG and 423 bp downstream of its stop codon, D: empty, E: rbcs terminator, and F: BASTA resistance as selection marker. GC005 and GC006 were obtained with a similar reaction using the modules: A: *p35S*, B: empty, the C module carrying the full *PCP CDS (alpha in* GC005 and *beta in* GC006*)*, D: empty, E: rbcs terminator, and F: BASTA resistance as selection marker (module F; pGGF001).

RT-PCR

RNA extraction, cDNA synthesis, qPCR amplification have been performed as described in chapter 2. Expression was normalized against *UBC21* (Table 3.7).

RNA-SEQ LIBRARIES PREPARATION

Total RNA was extracted with Qiagen Rneasy Plant kit, treated with Dnase, purified with a phenol-chloroform extraction and precipitated with ethanol. 1ug of total RNA was used to prepare the libraries with the TruSeq Stranded Total RNA library preparation kit (Illumina) following the manufacturer's instruction. The mRNA selection process was performed by using rRNA Removal Mix-Plants (RiboZero Plant kit, Illumina). As recommended by Illumina, DNA fragments were amplified by 15 PCR cycles, and a temperature fragmentation step at 94 degrees for 2 minutes was used for size selection. Paired-ended sequencing was performed on the Hiseq 2000 or Hiseq 3000 machines operated by the sequencing center at the MPI for Developmental Biology. Two of the biological replicates of *pcp-1* grown at 16°C and shifted to 23°C were sequenced twice to increase coverage.

PRE-PROCESSING OF RNA-SEQ DATA, DIFFERENTIAL EXPRESSION AND DIFFERENTIAL USAGE ANALYSES

The data pre-processing was performed following the guidelines described here: http://www.epigenesys.eu/en/protocols/bio-informatics/1283-guidelines-for-rna-seq-data-analysis. Briefly, the quality of the raw sequence data was assessed using FastQC

(http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), v0.11.4. Residual ribosomal RNA (rRNA) contamination was assessed and filtered using SortMeRNA (v2.1; Kopylova *et al.* 2012; settings --log --paired_in --fastx--sam --num_alignments 1) using the rRNA sequences provided with SortMeRNA (rfam-5s-database-id98.fasta, rfam-5.8s-database-id98.fasta, silva-arc-16s-database-id95.fasta, silva-bac-16s-database-id85.fasta, silva-euk-18s-database-id95.fasta, silva-arc-23s-database-id98.fasta, silva-bac-23s-database-id98.fasta and silva-euk-28s-database-id98.fasta). Data were then filtered to remove adapters and trimmed for quality using Trimmomatic (v0.36; Bolger *et al.* 2014; settings TruSeq3-PE-2.fa:2:30:10 SLIDINGWINDOW:5:20 MINLEN:50). After both filtering steps, FastQC was run again to ensure that no technical artefacts were introduced. Read counts were obtained using kallisto (v0.43.0, Bray *et al.*, 2016) with the parameters quant -b 100 --pseudobam -t 1 --rf-stranded and using the TAIR10 cDNA sequences as a reference (retrieved from the TAIR resource; Berardini *et al.*, 2015). An overview of the data, including raw and post-QC read counts and pseudo-alignment rates is given in (Table 3.4). The kallisto abundance values were imported into R (v3.3.2; R Core Team 2015) using the Bioconductor (v3.3; Gentleman *et al.* 2004) tximport package (v.1.2.0; Soneson *et al.*, 2015). For the data quality assessment (QA) and visualisation, the read counts were normalised using a variance stabilising transformation as implemented in DESeq2. The biological relevance of the data - e.g. biological replicates similarity - was assessed by Principal Component Analysis (PCA) and other visualisations (e.g. heatmaps), using custom R scripts.

Table 3.4

SampleID	Raw reads	rRNA filtered reads	rRNA filtered reads (%)	total rRNA	5S	5.8S	Arc16S	Bac16S	18S	Arc23S	Bac23S	28S
pcp 9 d 16 +3 d 23	53035681	11627937	22.1%	77.9%	0.3%	1.1%	12.6%	1.8%	12.3%	21.0%	16.5%	12.3%
pcp 9 d 16 +3 d 23	45786931	7817723	17.2%	82.8%	0.3%	1.1%	13.4%	1.9%	13.5%	22.0%	17.3%	13.4%
pcp 9 d 16 +3 d 23	50766841	47159758	93.4%	6.6%	0.0%	0.0%	0.1%	0.4%	0.3%	0.3%	5.1%	0.4%
Col-0 9 d 16 +3 d 23	42340804	39449470	93.7%	6.3%	0.0%	0.0%	0.1%	0.4%	0.3%	0.3%	4.8%	0.3%
Col-0 9 d 16 +3 d 23	44691756	41695015	93.8%	6.2%	0.0%	0.0%	0.1%	0.4%	0.3%	0.3%	4.9%	0.3%
Col-0 9 d 16 +3 d 23	47929272	44606754	93.6%	6.4%	0.0%	0.0%	0.1%	0.5%	0.3%	0.3%	4.9%	0.3%
pcp 9 d 16	54491682	39236937	72.9%	27.1%	0.0%	0.1%	3.6%	0.9%	3.7%	6.5%	8.3%	4.0%
pcp 9 d 23	49287105	45842509	93.5%	6.5%	0.0%	0.0%	0.2%	0.5%	0.4%	0.4%	4.7%	0.3%
pcp 9 d 23	51113072	47711449	93.8%	6.2%	0.0%	0.0%	0.1%	0.5%	0.3%	0.3%	4.6%	0.3%
pcp 9 d 23	46468379	43170260	93.4%	6.6%	0.0%	0.0%	0.2%	0.5%	0.4%	0.4%	4.7%	0.5%
Col-0 9 d 23	46920425	43713456	93.6%	6.4%	0.0%	0.0%	0.2%	0.5%	0.3%	0.4%	4.6%	0.4%
Col-0 9 d 23	45107165	42155485	93.9%	6.1%	0.0%	0.0%	0.1%	0.5%	0.3%	0.3%	4.6%	0.3%
pcp 9 d 16	53164725	49170501	93.2%	6.8%	0.0%	0.0%	0.2%	0.6%	0.3%	0.4%	5.0%	0.3%
Col-0 9 d 23	44413781	41376115	93.7%	6.4%	0.0%	0.0%	0.1%	0.5%	0.3%	0.4%	4.7%	0.4%
pcp 9 d 23 +3 d 16	50307767	46997817	93.9%	6.1%	0.0%	0.0%	0.1%	0.5%	0.3%	0.4%	4.5%	0.4%
pcp 9 d 23 +3 d 16	43991403	41048407	93.8%	6.2%	0.0%	0.0%	0.1%	0.5%	0.3%	0.4%	4.5%	0.4%
pcp 9 d 23 +3 d 16	42284016	39607745	94.2%	5.9%	0.0%	0.0%	0.1%	0.4%	0.3%	0.3%	4.4%	0.4%
pcp 9 d 16	50106245	46660048	93.7%	6.3%	0.0%	0.0%	0.1%	0.6%	0.3%	0.2%	4.7%	0.3%
Col-0 9 d23 +3 d 16	72150164	67576679	94.1%	5.9%	0.0%	0.0%	0.1%	0.4%	0.3%	0.3%	4.3%	0.4%
Col-0 9 d23 +3 d 16	45197701	42364820	94.2%	5.8%	0.0%	0.0%	0.1%	0.5%	0.3%	0.3%	4.3%	0.4%
Col-0 9 d23 +3 d 16	36317335	33859594	93.7%	6.3%	0.0%	0.0%	0.2%	0.4%	0.3%	0.5%	4.3%	0.6%
Col-0 9 d 16	55363396	51696361	94.0%	6.0%	0.0%	0.0%	0.1%	0.4%	0.3%	0.3%	4.6%	0.3%
Col-0 9 d 16	49071763	45473752	93.3%	6.7%	0.0%	0.0%	0.2%	0.5%	0.4%	0.4%	4.8%	0.4%
Col-0 9 d 16	51853894	33701711	65.5%	34.5%	0.1%	0.2%	5.3%	1.0%	4.8%	8.8%	9.4%	5.1%
pcp 9 d 16 +3 d 23 resequencing	213853250	91969223	21.5%	78.5%	0.4%	1.1%	13.5%	1.7%	11.6%	21.9%	16.0%	12.2%
pcp 9 d 16 +3 d 23 resequencing	201676655	67509619	16.7%	83.3%	0.4%	1.1%	14.5%	1.7%	12.8%	23.1%	16.6%	13.1%

SampleID	Trimmed reads	Trimmed reads (%)	Forward	F%	Reverse	R%	Dropped	D%	Pseudoaligned	Pseudoaligned %
pcp 9 d 16 +3 d 23	2298671	95.9%	67257	2.8%	13269	0.6%	17803	0.7%	1988379	86.5%
pcp 9 d 16 +3 d 23	7534015	96.4%	196421	2.5%	40531	0.5%	46756	0.6%	6642387	88.2%
pcp 9 d 16 +3 d 23	44237098	93.8%	2273911	4.8%	317313	0.7%	331436	0.7%	38928738	88.0%
Col-0 9 d 16 +3 d 23	37110031	94.1%	1787565	4.5%	294934	0.8%	256940	0.7%	32538525	87.7%
Col-0 9 d 16 +3 d 23	39269703	94.2%	1858105	4.5%	301396	0.7%	265811	0.6%	34499093	87.9%
Col-0 9 d 16 +3 d 23	42039888	94.3%	1965658	4.4%	310786	0.7%	290422	0.7%	36684048	87.3%
pcp 9 d 16	37107797	94.6%	1733394	4.4%	195520	0.5%	200226	0.5%	30422529	82.0%
pcp 9 d 23	43342513	94.6%	1887186	4.1%	337974	0.7%	274836	0.6%	37421436	86.3%
pcp 9 d 23	45079937	94.5%	1984277	4.2%	361164	0.8%	286071	0.6%	38876457	86.2%
pcp 9 d 23	40688088	94.3%	1836479	4.3%	322598	0.8%	323095	0.8%	35010089	86.0%
Col-0 9 d 23	41280949	94.4%	1814611	4.2%	325707	0.8%	292189	0.7%	35968439	87.1%
Col-0 9 d 23	39711012	94.2%	1886627	4.5%	287148	0.7%	270698	0.6%	34598926	87.1%
pcp 9 d 16	46697998	95.0%	1988502	4.0%	260290	0.5%	223711	0.5%	38628508	82.7%
Col-0 9 d 23	39108537	94.5%	1703916	4.1%	312705	0.8%	250957	0.6%	34065655	87.1%
pcp 9 d 23 +3 d 16	44494944	94.7%	1874164	4.0%	352381	0.8%	276328	0.6%	36979947	83.1%
pcp 9 d 23 +3 d 16	38709867	94.3%	1785160	4.4%	287956	0.7%	265424	0.7%	32137462	83.0%
pcp 9 d 23 +3 d 16	37491472	94.7%	1591840	4.0%	293842	0.7%	230591	0.6%	31323772	83.5%
pcp 9 d 16	44585423	95.6%	1643960	3.5%	234231	0.5%	196434	0.4%	37640400	84.4%
Col-0 9 d23 +3 d 16	63936637	94.6%	2737213	4.1%	516638	0.8%	386191	0.6%	53686701	84.0%
Col-0 9 d23 +3 d 16	40013512	94.5%	1734695	4.1%	322396	0.8%	294217	0.7%	33477685	83.7%
Col-0 9 d23 +3 d 16	31923326	94.3%	1456210	4.3%	250801	0.7%	229257	0.7%	27078385	84.8%
Col-0 9 d 16	49286552	95.3%	1926842	3.7%	250585	0.5%	232382	0.5%	42128124	85.5%
Col-0 9 d 16	43452324	95.6%	1608700	3.5%	226308	0.5%	186420	0.4%	36335575	83.6%
Col-0 9 d 16	32144729	95.4%	1233489	3.7%	176973	0.5%	146520	0.4%	27023929	84.1%
pcp 9 d 16 +3 d 23 resequencing	43815834	96.0%	1291160	2.8%	198506	0.4%	346903	0.8%	37903169	86.5%
pcp 9 d 16 +3 d 23 resequencing	32423457	96.5%	835965	2.5%	134669	0.4%	198448	0.6%	28546237	88.0%

Statistical analysis of gene and transcript differential expression (DE) between conditions was performed in R using the Bioconductor DESeq2 package (v1.14.1; Love *et al.* 2014), with the following model: ~genotype * temperature to account for both the genotype (WT or pcp) and the temperature at harvesting (16 or 23 degrees Celsius). FDR adjusted p-values were used to assess significance; a common threshold of 1% was used throughout. The DE results at the

gene and transcript level were compared to identify differential transcript usage (DTU). Statistical analysis of differential exon usage (DEU) was performed using the R (v3.4.0) Bioconductor (v3.4) DEXSeq package (v1.22.0; Anders *et al.*, 2012) with the model: ~ sample + exon + condition:exon for four comparisons, where condition represented: 1) genotype (pcp vs. WT) at 16 degree Celsius, 2) genotype (pcp vs. WT) at 23 degree Celsius, 3) temperature (warm vs. cold) for the WT genotype and 4) temperature (warm vs. cold) for the pcp genotype. All the expression results were generated in R, using custom scripts. The differentially expressed genes (DEGs) obtained at the previous step for the temperature effect, the genotype effect or their interaction were imported into AtGenIE.org (Sundell *et al.*, 2015) and used for Gene Ontology (GO) enrichment analyses. GO tree maps were generated using REVIGO (Supek *et al.*, 2011) with default settings but for the background database that was set to "*Arabidopsis thaliana*", and custom R scripts.

IN SITU RNA HYBRIDIZATION, HISTOLOGY AND MICROSCOPY

WUS, *CLV3*, and *PCP* were amplified by PCR from start to stop codons (using primers listed in Table 3.7) and cloned into pGEM-T easy (Promega). Digoxigenin-labeled antisense probes were synthesized with T7 or SP6 RNA polymerase (Roche). Hydrolysis of probes longer than 450 bp (*WUS*) was performed with a standard carbonate reaction at 60°C for 20 minutes as previously described (Palatnik *et al.* 2003). Shoot apices from 9 or 12 days old long-days-grown plants were manually dissected and fixed in formalin/acetic acid/ethanol (1:1:18). Paraplast-embedded material was sectioned to

8 µm thickness. Hybridization and detection were performed as previously described (Palatnik *et al.* 2003).

Scanning electron microscopy (SEM) was performed as previously described (Dinneny *et al.* 2004). Embedding of plant material in JB-4 media, sectioning and staining with Toluidine Blue was performed as previously described (Roeder *et al.* 2003).

PHENOTYPING

To analyze flowering time, plants were grown at 16°C or 23°C and the days to flower as well as the rosette and cauline leaf number were recorded.

To test the occurrence of seedlings with 3 cotyledons and flowers with 5 petals, *pcp-1* and Col-0 plants were grown at 23°C. 2018 and 792 seedlings were scored for *pcp-1* and Col-0 respectively. The number of seedling with 3 cotyledons detected were 64/2018 for *pcp-1* and 1/792 for Col-0 (Table 3.5). To test the occurrence of flowers with more than 4 petals 10 plants for each line (*pcp-1* and Col-0) were considered, and all the flowers developing from the principal shoot were analyzed daily. In *pcp-1* the total number of flowers considered was 287, and 12 of them carried more than 4 petals (11 flowers with 5 petals and 1 with 6 petals). In Col-0 among a total of 493 flowers tested only 1 carried 5 petals (Table 3.6).

To analyze the growth rate in *pcp-1* seedlings, plants were grown at 23°C or shifted to 16°C after 13 days and Col-0 was included as control. A total of 25 plants (20 *pcp-1* and 5 Col-0) were kept at 23°C, and a total of 29 plants (21 *pcp-1* and 8 Col-0) were used for the shifting experiment. Images were acquired daily at 14:45 in top view using two cameras per tray, which were mounted between the light tubes in the growth room. The cameras were equipped with an OmniVision OV5647 sensor with a resolution of 5 Megapixels and a

focal length of 3.6 mm. Each camera was attached to a Raspberry Pi computer (Revision 1.2, Raspberry Pi Foundation, UK), which performed the image recording and storage. The image pairs for each tray were subsequently merged into one image using transformations that were determined previously using a calibration pattern. Images of individual plants were extracted using a predefined mask for each plant. Segmentation of plant leaves and background was then performed in three steps. First, an initial segmentation was performed by removing the background voxels by applying a threshold on *a* and *b* channels in the *Lab* color space. Second, a GrabCut-based automatic postprocessing was applied (Cheng et al. 2015). In the third step, unsatisfactory segmentations were manually corrected using a graphical user interface where the user could assign background and foreground regions in order to improve the GrabCut segmentation. The leaf area of each plant was then calculated based on the segmented plant images. The image processing pipeline was implemented in Python.

Table 3.5

Tracking the number of seedlings in *pcp-1-1* and Col-0 with more than 2 cotyledons. The percentage of seedlings with more than 2 cotyledons is reported for each line.

Type	Pot	2-cotyledons	3-cotyledons	Tot	Tot seedlings
pcp-1	1	313	16	330	2018
pcp-1	2	202	4	206	
pcp-1	3	285	12	297	3-cot seedlings tot
pcp-1	4	152	6	158	64
pcp-1	5	211	5	216	
pcp-1	6	104	3	107	%
pcp-1	7	52	2	54	3.171456888
pcp-1	8	24	0	24	
pcp-1	9	71	1	72	
pcp-1	10	93	1	94	
pcp-1	11	87	2	89	
pcp-1	12	149	3	152	
pcp-1	13	128	2	130	

pcp-1	14	82	7	89	
Type	Pot	2-cotyledons	3-cotyledons	Tot	Tot seedlings
Col-0	1	14	0	14	792
Col-0	2	7	0	7	
Col-0	3	33	0	33	3-cot seedlings tot
Col-0	4	24	0	24	1
Col-0	5	40	0	40	
Col-0	6	43	0	43	%
Col-0	7	29	0	29	0.126262626
Col-0	8	203	0	203	
Col-0	9	114	0	114	
Col-0	10	56	0	56	
Col-0	11	63	0	63	
Col-0	12	92	1	93	
Col-0	13	73	0	73	

Table 3.6

Tracking the number of petals in *pcp-1* and Col-0 flowers. The percentage of flowers with more than 4 petals is reported for each line.

pcp-1	flowers tested	4 petals	3 petals	5 petals	6 petals	tot >4 p	% > 4 p	Tot flowe
pcp-1 plant1	32	32	0	0	0	0	0	287
pcp-1 plant2	32	32	0	0	0	0	0	Flowers >
pcp-1 plant3	37	36	0	1	0	1	2.702702703	12
pcp-1 plant4	25	24	0	1	0	1	4	% occurre
pcp-1 plant5	27	22	0	5	0	5	18.51851852	4.181184
pcp-1 plant6	21	20	0	1	0	1	4.761904762	
pcp-1 plant7	32	31	0	0	1	1	3.125	
pcp-1 plant8	23	22	0	1	0	1	4.347826087	
pcp-1 plant9	29	28	0	1	0	1	3.448275862	
pcp-1 plant10	29	28	0	1	0	1	3.448275862	
Col-0	flowers tested	4 petals	3 petals	5 petals	6 petals	tot >4 p	% > 4 p	Tot flowe
Col-0 plant1	53	53	0	0	0	0	0	493
Col-0 plant2	45	44	1	0	0	0	0	Flowers >
Col-0 plant3	52	52	0	0	0	0	0	1
Col-0 plant4	48	47	0	1	0	1	2.083333333	% occurre
Col-0 plant5	28	28	0	0	0	0	0	0.202839
Col-0 plant6	57	57	0	0	0	0	0	
Col-0 plant7	54	54	0	0	0	0	0	
Col-0 plant8	52	52	0	0	0	0	0	
Col-0 plant9	51	51	0	0	0	0	0	
Col-0 plant10	53	52	1	0	0	0	0	

Table 3.7

Primers used in this chapter

Oligonucleotides used for CDS amplification

Gene	Primer	Sequence (5' to 3')	Product	Hydrolysis
CLV3 **Forward**	G-42323	ATGGATTCGAAGAGTTTTCTG	291 bp	no
CLV3 **Reverse**	G-42324	TCAAGGGAGCTGAAAGTTGTT		
WUS **Forward**	G-42325	ATGGAGCCGCCACAGCATCAG	879 bp	yes
WUS **Reverse**	G-42326	CTAGTTCAGACGTAGCTCAAG		
PCP-α **Forward**	G-41551	ATGGCGAGCACCAAAGTTCAAAG	267 bp	no
PCP-α **Reverse**	G-41552	TCACTTTCCCGTGTTCATCATC		

Oligonucleotides used for genotyping T-DNA lines

T-DNA line	Primers	Sequence (5' to 3')
SALK_004764	Forward	CTCCTCCTCCTAAACCACCAC
	Reverse	TGGTGAACTGATGATTGCATTC
SALK_118875	Forward	AGAACAATCCACGAAAGCATG
	Reverse	TGTTTTGGGATTCTTCAGTGG
SALK_089521	Forward	CTCCGATTCACCAGACTTGAG
	Reverse	GCCGAAGAGAATGACACAATC
SALK_100059	Forward	CTGAGCTTCATGAGGGTTTTG
	Reverse	TGTCGAAGATAATCGGTTTCG
SALK_040864	Forward	CTGATGCAAAACATCACATGG
	Reverse	AATGCTTTGGTTGATACGACG
SALK_054457	Forward	GACATTAGCGCATTCAAAAGC
	Reverse	CCTCTCAACAAGCTCCTGATG
SALK_025193	Forward	CAACTACCTATGCTGAAGCCG
	Reverse	CATTCGAGGCAGCTTAGTCAG
SALK_055239	Forward	AAAGAAGAAGAGCGTGGAAGC
	Reverse	CATTTAAGCTCCCGGGACTAG
SALK_051523	Forward	ACAGAACCAAAACTATGCCCC
	Reverse	GCAAATGTGATCAGAGTGTGC
SALK_041457	Forward	GCACTGAACAGACGTTTAGGC

	Reverse	ATCAAACGGTCCATGTGAGAC
SALK_119088	Forward	GATTGTGTCATTCTCTTCGGC
	Reverse	CGGTCAGTTCACCAAACAAAC
SALK_017458	Forward	GGCTTACACTAACAGCGTTGG
	Reverse	TGCTCAAATAGCCAAATCTGG
SALK_022142	Forward	TGCTCAAATAGCCAAATCTGG
	Reverse	TCGAAGATTCAATGGTTGGAC
SALK_139504	Forward	CACAAACTCAATGGTGTTCCC
	Reverse	GCTTTAACAGAACGTGTTCGG
SALK_139760	Forward	TCATTGCTATCAATTGGTCCC
	Reverse	AATCGTTAATGATCAATGGCG

Universal SALK T-DNA insertion primer (Forward):
GCGTGGACCGCTTGCTGCAACT

Oligonucleotides used for qPCR amplification

Gene	Primer	Sequence (5' to 3')
***PCP-α* Forward**	G-41551	ATGGCGAGCACCAAAGTTCAAAG
***PCP-α* Reverse**	G-41552	TCACTTTCCCGTGTTCATCATC
***CLV3* Forward**	G-42323	ATGGATTCGAAGAGTTTTCTG
***CLV3* Reverse**	G-42324	TCAAGGGAGCTGAAAGTTGTT
***WUS* Forward**	G-43220	AAGCCATATCCCAGCTTCAA
***WUS* Reverse**	G-43221	CCATCCTCCACCTACGTTGT
***UBC21* Forward**	G-38780	CTCCTCAAGTTCGATTCTTG
***UBC21* Reverse**	G-38783	CCTGAGTCGCAGTTAAGAGG
***At4g30330* Forward**	G-43077	AAGCTCGGATTCAGATTTGG
***At4g30330* Reverse**	G-43078	AAGCAAAATCCGACCAAGTG

3.6 CONTRIBUTIONS

Giovanna Capovilla (GC) and Markus Schmid conceived and designed the experiments. GC performed the experiments. Nicolas Delhomme analyzed most of the RNA-seq data with help from Iryna Shutava and Silvio Collani. Ilja Bezrukov analyzed the growth rate data. GC performed the rest of the analysis. Jürgen Berger and GC took the SEM pictures.

CHAPTER 4

4.1 CONCLUSIONS

Alternative splicing, the choice of one splice site over another during the processing of primary transcripts, is in part affected by temperature fluctuations. This thermoregulated alternative splicing and its contribution in two important aspects of A*rabidopsis thaliana* development, flowering time and morphogenesis, is the focus of this dissertation.

First I clarified the role of *FLM,* a MADS-domain transcription factor known to produce alternative splicing variants in response to changes in temperature, in the regulation of flowering. The contribution of the main isoforms of *FLM* in flowering time was controversial, and two mechanisms for its action had been proposed. The first model proposed that one particular isoform, FLM-δ, which is preferentially expressed at high temperatures, acts as a dominant-negative protein that competes with the main isoform, FLM-β, for binding with interaction partners, thereby indirectly promoting flowering (Posè *et al.* 2013, Lee *et al.* 2013). The second model rejected an active role for FLM-δ in the regulation of flowering. Instead, AS would preferentially produce non-canonical isoforms that are subjected to NMD at high temperatures, resulting in a decrease in expression of the floral repressor FLM-β (Sureshkumar *et al.* 2016). To investigate the role of each main *FLM* isoform, I deleted the cassette exons that distinguish FLM-β and FLM-δ from the genomic *FLM* locus using CRISPR technology.

My findings confirm the role of FLM-β as the main isoform that represses the transition to flowering. Loss of FLM-β leads to early flowering, whereas its over-expression has the opposite effect. In contrast, FLM-δ, which in principle has the potential to act as a dominant-negative regulator of flowering time, never reaches the expression levels required to realize this potential in WT conditions. The temperature sensitive AS in *FLM* could be of relevance from an evolutionary perspective. Maintaining the ability to produce a variety of *FLM* isoforms could provide flowering time plasticity, and some of the many alternative transcripts produced by this gene, particularly at elevated ambient temperature, might play an active role in the regulation of flowering and be relevant in different accessions of *Arabidopsis thaliana* or species. This could be a useful advantage, given climate change.

In the second part of my dissertation I screened mutant lines in *Arabidopsis thaliana* with the aim to identify putative AS factors that controlled plant development in a temperature-dependent manner. Using this approach I have identified a *bona fi*de AS regulator, which is essential for correct development of *Arabidopsis thaliana* at low temperature, and I phenotypically characterized the loss of function mutant. I called this previously uncharacterized gene *PORCUPINE* (*PCP*) because of the peculiar "spiky" phenotype it displays at low temperatures. Mutations of *PCP* show severe meristem defects only when grown at 16°C, while at 23°C no obvious phenotype is detectable. Interestingly, growth can be arrested at any developmental stage by shifting *pcp* mutant plants grown at 23°C to lower temperatures. Reversely, the phenotype can be rescued by exposing *pcp* mutants that had been grown at low temperatures to 23°C. This behavior indicates the presence of a mechanism that allows adaptation of meristem development to changes in the

ambient temperature. RNA-seq analysis supports a role for PCP in AS regulation. The shoot apical meristem (SAM) defects observed in the mutant can be associated with the misregulation of two genes involved in maintaining the stem cell fate in the SAM, *WUS* and *CLV3*, particularly at low ambient temperatures. Whether the misregulation of *WUS* and *CLV3* are causal or rather a consequence of a more general disturbance of SAM development in the *pcp* mutant remains to be determined.

My findings establish *PCP* as an AS factor involved in regulating development in plants. Furthermore, they open up a field of research that links temperature regulated morphogenesis with AS, and a new function for the Sm genes.

REFERENCES

1001 Genomes Consortium. Electronic address: magnus.nordborg@gmi.oeaw.ac.at. 2016. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. Cell **166**, 481–491.

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T. 2005. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science **309**, 1052–1056.

Airoldi CA, McKay M, Davies B. 2015. MAF2 Is Regulated by Temperature-Dependent Splicing and Represses Flowering at Low Temperatures in Parallel with FLM. PloS one **10**, e0126516.

Anders S, Reyes A, Huber W. 2012. Detecting differential usage of exons from RNA-seq data. Genome research **22**, 2008–2017.

Andrés F, Coupland G. 2012. The genetic basis of flowering responses to seasonal cues. Nature reviews. Genetics **13**, 627–639.

Angel A, Song J, Dean C, Howard M. 2011. A Polycomb-based switch underlying quantitative epigenetic memory. Nature **476**, 105–108.

Aukerman MJ, Sakai H. 2003. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. The Plant cell **15**, 2730–2741.

Bai Y, Dai X, Harrison AP, Chen M. 2015. RNA regulatory networks in animals and plants: a long noncoding RNA perspective. Briefings in functional genomics **14**, 91–101.

Balasubramanian S, Sureshkumar S, Lempe J, Weigel D. 2006. Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. PLoS genetics **2**, e106.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. Cell research **21**, 381–395.

Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. Nature **427**, 164–167.

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V. 2013. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant methods **9**, 39.

Bell LR, Horabin JI, Schedl P, Cline TW. 1991. Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. Cell **65**, 229–239.

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E. 2015. The Arabidopsis information resource: Making and mining the 'gold standard' annotated reference plant genome. Genesis **53**, 474–485.

Black DL, Chabot B, Steitz JA. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell **42**, 737–750.

Blázquez MA, Ahn JH, Weigel D. 2003. A thermosensory pathway controlling flowering time in Arabidopsis thaliana. Nature genetics **33**, 168–171.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**, 2114–2120.

Bouché F, Lobet G, Tocquin P, Périlleux C. 2016. FLOR-ID: an interactive database of flowering-time gene networks in Arabidopsis thaliana. Nucleic acids research **44**, D1167–D1171.

Box MS, Coustham V, Dean C, Mylne JS. 2011. Protocol: A simple phenol-based method for 96-well extraction of high quality RNA from Arabidopsis. Plant methods **7**, 7.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nature biotechnology **34**, 525–527.

Cao J, Shi F, Liu X, Jia J, Zeng J, Huang G. 2011. Genome-wide identification and evolutionary analysis of Arabidopsis sm genes family. Journal of biomolecular structure & dynamics **28**, 535–544.

Capovilla G, Pajoro A, Immink RGH, Schmid M. 2015. Role of alternative pre-mRNA splicing in temperature signaling. Current opinion in plant biology **27**, 97–103.

Capovilla G, Schmid M, Posé D. 2014. Control of flowering by ambient temperature. Journal of experimental botany **66**, 59–69.

Chen X. 2004. A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science **303**, 2022–2025.

Cheng M-M, Mitra NJ, Huang X, Torr PHS, Hu S-M. 2015. Global Contrast Based Salient Region Detection. IEEE transactions on pattern analysis and machine intelligence **37**, 569–582.

Chouard P. 1960. Vernalization and its Relations to Dormancy. Annual review of plant physiology **11**, 191–238.

Chow BY, Kay SA. 2013. Global approaches for telling time: omics and the Arabidopsis circadian clock. Seminars in cell & developmental biology **24**, 383–392.

Clark SE. 2001. Cell signalling at the shoot meristem. Nature reviews. Molecular cell biology **2**, 276–284.

Clark SE, Running MP, Meyerowitz EM. 1995. CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. Development **121**, 2057–2067.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal: for cell and molecular biology **16**, 735–743.

Conn VM, Hugouvieux V, Nayak A, *et al.* 2017. A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation. Nature plants **3**, 17053.

Cooper GM. 2000. *The Complexity of Eukaryotic Genomes*. Sinauer Associates.

Cossart P, Lecuit M. 1998. Interactions of Listeria monocytogenes with mammalian cells during entry and actin-based movement: bacterial factors, cellular ligands and signaling. The EMBO journal **17**, 3797–3806.

Crain DA, Guillette LJ Jr. 1998. Reptiles as models of contaminant-induced endocrine disruption. Animal reproduction science **53**, 77–86.

De Lucia F, Crevillen P, Jones AME, Greb T, Dean C. 2008. A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. Proceedings of the National Academy of Sciences of the United States of America **105**, 16831–16836.

Demir E, Dickson BJ. 2005. fruitless splicing specifies male courtship behavior in Drosophila. Cell **121**, 785–794.

Dinneny JR, Yadegari R, Fischer RL, Yanofsky MF, Weigel D. 2004. The role of JAGGED in shaping lateral organs. Development **131**, 1101–1110.

Doerner P. 2003. Plant Meristems: A Merry-Go-Round of Signals Review. Current biology: CB **13**, R368–R374.

Elliott RC, Betzner AS, Huttner E, Oakes MP, Tucker WQ, Gerentes D, Perez P, Smyth DR. 1996. AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. The Plant cell **8**, 155–168.

Emery JF, Floyd SK, Alvarez J, Eshed Y, Hawker NP, Izhaki A, Baum SF, Bowman JL. 2003. Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. Current biology: CB **13**, 1768–1774.

Eshed Y, Baum SF, Perea JV, Bowman JL. 2001. Establishment of polarity in lateral organs of plants. Current biology: CB **11**, 1251–1260.

Eshed Y, Izhaki A, Baum SF, Floyd SK, Bowman JL. 2004. Asymmetric leaf development and blade expansion in Arabidopsis are mediated by KANADI and YABBY activities. Development **131**, 2997–3006.

Fang Y, Spector DL. 2007. Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. Current biology: CB **17**, 818–823.

Fauser F, Schiml S, Puchta H. 2014. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant journal: for cell and molecular biology **79**, 348–359.

Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, García JA, Paz-Ares J. 2007. Target mimicry provides a new mechanism for regulation of microRNA activity. Nature genetics **39**, 1033–1037.

Gaj T, Gersbach CA, Barbas CF 3rd. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology **31**, 397–405.

Galvão VC, Collani S, Horrer D, Schmid M. 2015. Gibberellic acid signaling is required for ambient temperature-mediated induction of flowering in Arabidopsis thaliana. The Plant journal: for cell and molecular biology **84**, 949–962.

Gao X, Chen J, Dai X, Zhang D, Zhao Y. 2016. An Effective Strategy for Reliably Isolating Heritable and Cas9-Free Arabidopsis Mutants Generated by CRISPR/Cas9-Mediated Genome Editing. Plant physiology **171**, 1794–1800.

Gentleman RC, Carey VJ, Bates DM, *et al.* 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome biology **5**, R80.

Giakountis A, Coupland G. 2008. Phloem transport of flowering signals. Current opinion in plant biology **11**, 687–694.

Gnesutta N, Kumimoto RW, Swain S, Chiara M, Siriwardana C, Horner DS, Holt BF, Mantovani R. 2017. CONSTANS imparts DNA sequence-specificity to the histone-fold NF-YB/NF-YC dimer. The Plant Cell Online.

Greb T, Mylne JS, Crevillen P, Geraldo N, An H, Gendall AR, Dean C. 2007. The PHD finger protein VRN5 functions in the epigenetic silencing of Arabidopsis FLC. Current biology: CB **17**, 73–78.

Gu X, Le C, Wang Y, Li Z, Jiang D, Wang Y, He Y. 2013. Arabidopsis FLC clade members form flowering-repressor complexes coordinating responses to endogenous and environmental cues. Nature communications **4**, 1947.

Han P, Li Q, Zhu Y-X. 2008. Mutation of Arabidopsis BARD1 causes meristem defects by failing to confine WUSCHEL expression to the organizing center. The Plant cell **20**, 1482–1493.

Han P, Zhu Y-X. 2009. BARD1 may be renamed ROW1 because it functions mainly as a REPRESSOR OF WUSCHEL1. Plant signaling & behavior **4**, 52–54.

Heidstra R, Sabatini S. 2014. Plant and animal stem cells: similar yet different. Nature reviews. Molecular cell biology **15**, 301–312.

Heo JB, Sung S. 2011. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science **331**, 76–79.

Holland PM, Abramson RD, Watson R, Gelfand DH. 1991. Detection of specific polymerase chain reaction product by utilizing the 5'----3'exonuclease activity of Thermus aquaticus DNA polymerase. Proceedings of the National Academy of Sciences **88**, 7276–7280.

Hong S, Song H-R, Lutz K, Kerstetter RA, Michael TP, McClung CR. 2010. Type II protein arginine methyltransferase 5 (PRMT5) is required for circadian period determination in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America **107**, 21211–21216.

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

Hyun Y, Richter R, Vincent C, Martinez-Gallegos R, Porri A, Coupland G. 2016. Multi-layered Regulation of SPL15 and Cooperation with SOC1 Integrate Endogenous Flowering Pathways at the Arabidopsis Shoot Meristem. Developmental cell **37**, 254–266.

Ito J, Batth TS, Petzold CJ, Redding-Johanson AM, Mukhopadhyay A, Verboom R, Meyer EH, Millar AH, Heazlewood JL. 2011. Analysis of the Arabidopsis cytosolic proteome highlights subcellular partitioning of central plant metabolism. Journal of proteome research **10**, 1571–1582.

Jiang D, Gu X, He Y. 2009. Establishment of the winter-annual growth habit via FRIGIDA-mediated histone methylation at FLOWERING LOCUS C in Arabidopsis. The Plant cell **21**, 1733–1746.

Jin J, Zhang H, Kong L, Gao G, Luo J. 2014. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic acids research **42**, D1182–7.

Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. 2002. An RNA thermosensor controls expression of virulence genes in Listeria monocytogenes. Cell **110**, 551–561.

Jones MA, Williams BA, McNicol J, Simpson CG, Brown JWS, Harmer SL. 2012. Mutation of Arabidopsis spliceosomal timekeeper locus1 causes circadian clock defects. The Plant cell **24**, 4066–4082.

Jönsson H, Heisler MG, Shapiro BE, Meyerowitz EM, Mjolsness E. 2006. An auxin-driven polarized transport model for phyllotaxis. Proceedings of the National Academy of Sciences of the United States of America **103**, 1633–1638.

Jung J-H, Domijan M, Klose C, *et al.* 2016. Phytochromes function as thermosensors in Arabidopsis. Science **354**, 886–889.

Kalyna M, Simpson CG, Syed NH, *et al.* 2012. Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic acids research **40**, 2454–2469.

Kim D-H, Sung S. 2017. Vernalization-Triggered Intragenic Chromatin Loop Formation by Long Noncoding RNAs. Developmental cell **40**, 302–312.e4.

Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA. 2016. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. The Plant journal: for cell and molecular biology **88**, 1058–1070.

Klucher KM, Chow H, Reiser L, Fischer RL. 1996. The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. The Plant cell **8**, 137–153.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. 1999. A pair of related genes with antagonistic roles in mediating flowering signals. Science **286**, 1960–1962.

Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics **28**, 3211–3217.

Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nature reviews. Molecular cell biology **14**, 153–165.

Kouzarides T. 2007. Chromatin modifications and their function. Cell **128**, 693–705.

Kumar SV, Lucyshyn D, Jaeger KE, Alós E, Alvey E, Harberd NP, Wigge PA. 2012. Transcription factor PIF4 controls the thermosensory activation of flowering. Nature **484**, 242–245.

Kumar SV, Wigge PA. 2010. H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell **140**, 136–147.

Lampropoulos A, Sutikovic Z, Wenzl C, Maegele I, Lohmann JU, Forner J. 2013. GreenGate - A Novel, Versatile, and Efficient Cloning System for Plant Transgenesis. PloS one **8**, e83043.

Lang A. 1965. Physiology of flower initiation. Handbuch der Pflanzenphysiologie / Encyclopedia of Plant Physiology. Differenzierung und Entwicklung / Differentiation and Development. Springer, Berlin, Heidelberg, 1380–1536.

Laubinger S, Marchal V, Le Gourrierec J, *et al.* 2006. Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. Development **133**, 3213–3222.

Lee LG, Connell CR, Bloch W. 1993. Allelic discrimination by nick-translation PCR with fluorgenic probes. Nucleic acids research **21**, 3761–3766.

Lee J, Lee I. 2010. Regulation and function of SOC1, a flowering pathway integrator. Journal of experimental botany **61**, 2247–2254.

Lee JH, Ryu H-S, Chung KS, Pose D, Kim S, Schmid M, Ahn JH. 2013. Regulation of Temperature-Responsive Flowering by MADS-Box Transcription Factor Repressors. Science **342**, 628–632.

Lee JH, Yoo SJ, Park SH, Hwang I, Lee JS, Ahn JH. 2007. Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. Genes & development **21**, 397–402.

Legris M, Klose C, Burgie ES, Rojas CCR, Neme M, Hiltbrunner A, Wigge PA, Schäfer E, Vierstra RD, Casal JJ. 2016. Phytochrome B integrates light and temperature signals in Arabidopsis. Science **354**, 897–900.

Leviatan N, Alkan N, Leshkowitz D, Fluhr R. 2013. Genome-wide survey of cold stress regulated alternative splicing in Arabidopsis thaliana with tiling microarray. PloS one **8**, e66511.

Liu J, Sun N, Liu M, Liu J, Du B, Wang X, Qi X. 2013*a*. An autoregulatory loop controlling Arabidopsis HsfA2 expression: role of heat shock-induced alternative splicing. Plant physiology **162**, 512–521.

Liu L, Zhu Y, Shen L, Yu H. 2013*b*. Emerging insights into florigen transport. Current opinion in plant biology **16**, 607–613.

Lorković ZJ, Wieczorek Kirk DA, Lambermon MH, Filipowicz W. 2000. Pre-mRNA splicing in higher plants. Trends in plant science **5**, 160–167.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology **15**, 550.

Lutz U, Nussbaumer T, Spannagl M, Diener J, Mayer KF, Schwechheimer C. 2017. Natural haplotypes of FLM non-coding sequences fine-tune flowering time in ambient spring temperatures in Arabidopsis. eLife **6**.

Lutz U, Posé D, Pfeifer M, Gundlach H, Hagmann J, Wang C, Weigel D, Mayer KFX, Schmid M, Schwechheimer C. 2015. Modulation of Ambient Temperature-Dependent Flowering in Arabidopsis thaliana by Natural Variation of FLOWERING LOCUS M. PLoS genetics **11**, e1005588.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nature reviews. Molecular cell biology **16**, 665–677.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. Science **339**, 823–826.

Marquardt S, Raitskin O, Wu Z, Liu F, Sun Q, Dean C. 2014. Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. Molecular cell **54**, 156–165.

Marquez Y, Brown JWS, Simpson C, Barta A, Kalyna M. 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome research **22**, 1184–1195.

Marshall CM, Tartaglio V, Duarte M, Harmon FG. 2016. The Arabidopsis sickle Mutant Exhibits Altered Circadian Clock Responses to Cool Temperatures and Temperature-Dependent Alternative Splicing. The Plant cell **28**, 2560–2575.

Ma X, Zhu Q, Chen Y, Liu Y-G. 2016. CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. Molecular plant **9**, 961–974.

McConnell JR, Emery J, Eshed Y, Bao N, Bowman J, Barton MK. 2001. Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. Nature **411**, 709–713.

McGlincy NJ, Smith CWJ. 2008. Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends in biochemical sciences **33**, 385–393.

Michaels SD, Amasino RM. 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. The Plant cell **11**, 949–956.

Mount SM, Pettersson I, Hinterberger M, Karmas A, Steitz JA. 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell **33**, 509–518.

Mudunkothge JS, Krizek BA. 2012. Three Arabidopsis AIL/PLT genes act in combination to regulate shoot apical meristem function. The Plant journal: for cell and molecular biology **71**, 108–121.

Naqvi AR, Sarwat M, Hasan S, Roychodhury N. 2012. Biogenesis, functions and fate of plant microRNAs. Journal of cellular physiology **227**, 3163–3168.

Nilsen TW, Graveley BR. 2010. Expansion of the eukaryotic proteome by alternative splicing. Nature **463**, 457–463.

Nohales MA, Kay SA. 2016. Molecular mechanisms at the core of the plant circadian oscillator. Nature structural & molecular biology **23**, 1061–1069.

Okamoto M, Matsui A, Tanaka M, Morosawa T, Ishida J, Iida K, Mochizuki Y, Toyoda T, Seki M. 2016. Sm-Like Protein-Mediated RNA Metabolism Is Required for Heat Stress Tolerance in Arabidopsis. Frontiers in plant science **7**, 1079.

Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D. 2003. Control of leaf morphogenesis by microRNAs. Nature **425**, 257–263.

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. 2005. Nuclear processing and export of microRNAs in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **102**, 3691–3696.

Pederson T. 2015. Molecular Biology of the Gene: by James D. Watson: W. A. Benjamin (1965): New York, New York. FASEB journal: official publication of the Federation of American Societies for Experimental Biology **29**, 4399–4401.

Perez-Santángelo S, Mancini E, Francey LJ, Schlaen RG, Chernomoretz A, Hogenesch JB, Yanovsky MJ. 2014. Role for LSM genes in the regulation of circadian rhythms. Proceedings of the National Academy of Sciences of the United States of America **111**, 15166–15171.

Pieau C, Girondot M, Richard-Mercier N, Desvages G, Dorizzi M, Zaborski P. 1994. Temperature sensitivity of sexual differentiation of gonads in the European pond turtle: Hormonal involvement. The Journal of experimental zoology **270**, 86–94.

Posé D, Verhage L, Ott F, Yant L, Mathieu J, Angenent GC, Immink RGH, Schmid M. 2013. Temperature-dependent regulation of flowering by antagonistic FLM variants. Nature **503**, 414–417.

Prescott DM. 1964. Cellular sites of RNA synthesis. Progress in nucleic acid research and molecular biology **3**, 33–57.

Proveniers MCG, van Zanten M. 2013. High temperature acclimation through PIF4 signaling. Trends in plant science **18**, 59–64.

Quint M, Delker C, Franklin KA, Wigge PA, Halliday KJ, van Zanten M. 2016. Molecular and genetic control of plant thermomorphogenesis. Nature plants **2**, 15190.

Ratcliffe OJ, Kumimoto RW, Wong BJ, Riechmann JL. 2003. Analysis of the Arabidopsis MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. The Plant cell **15**, 1159–1169.

Reddy ASN, Marquez Y, Kalyna M, Barta A. 2013. Complexity of the alternative splicing landscape in plants. The Plant cell **25**, 3657–3683.

Roeder AHK, Ferrándiz C, Yanofsky MF. 2003. The role of the REPLUMLESS homeodomain protein in patterning the Arabidopsis fruit. Current biology: CB **13**, 1630–1635.

Rosbash M, Séraphin B. 1991. Who's on first? The U1 snRNP-5' splice site interaction and splicing. Trends in biochemical sciences **16**, 187–190.

Rosloski SM, Singh A, Jali SS, Balasubramanian S, Weigel D, Grbic V. 2013. Functional analysis of splice variant expression of MADS AFFECTING FLOWERING 2 of Arabidopsis thaliana. Plant molecular biology **81**, 57–69.

Salomé PA, McClung CR. 2004. The Arabidopsis thaliana clock. Journal of biological rhythms **19**, 425–435.

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G. 2000. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science **288**, 1613–1616.

Sanchez SE, Petrillo E, Beckwith EJ, *et al.* 2010. A methyl transferase links the circadian clock to the regulation of alternative splicing. Nature **468**, 112–116.

Schlaen RG, Mancini E, Sanchez SE, Perez-Santángelo S, Rugnone ML, Simpson CG, Brown JWS, Zhang X, Chernomoretz A, Yanovsky MJ. 2015. The spliceosome assembly factor GEMIN2 attenuates the effects of temperature on alternative splicing and circadian rhythms. Proceedings of the National Academy of Sciences of the United States of America **112**, 9382–9387.

Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D. 2006. Highly specific gene silencing by artificial microRNAs in Arabidopsis. The Plant cell **18**, 1121–1133.

Scortecci KC, Michaels SD, Amasino RM. 2001. IdentiÆcation of a MADS-box gene, FLOWERING LOCUS M, that represses Øowering. The Plant journal: for cell and molecular biology **26**, 229–236.

Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, Amasino RA, Coupland G. 2006. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Genes & development **20**, 898–912.

Seo PJ, Park M-J, Lim M-H, Kim S-G, Lee M, Baldwin IT, Park C-M. 2012. A Self-Regulatory Circuit of CIRCADIAN CLOCK-ASSOCIATED1 Underlies the Circadian Clock Regulation of Temperature Responses in Arabidopsis. The Plant cell **24**, 2427–2442.

Sharbel TF, Haubold B, Mitchell-Olds T. 2000. Genetic isolation by distance in Arabidopsis thaliana: biogeography and postglacial colonization of Europe. Molecular ecology **9**, 2109–2118.

Sharp PA. 1994. Split genes and RNA splicing. Cell **77**, 805–815.

Sharp PA. 2005. The discovery of split genes and RNA splicing. Trends in biochemical sciences **30**, 279–281.

Sheldon CC, Finnegan EJ, Peacock WJ, Dennis ES. 2009. Mechanisms of gene repression by vernalization in Arabidopsis. The Plant journal: for cell and molecular biology **59**, 488–498.

Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL. 1999. Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. Development **126**, 4117–4128.

Somerville C, Koornneef M. 2002. A fortunate choice: the history of Arabidopsis as a model plant. Nature reviews. Genetics **3**, 883–889.

Soneson C, Love MI, Robinson MD. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research **4**, 1521.

Song J, Angel A, Howard M, Dean C. 2012. Vernalization - a cold-induced epigenetic switch. Journal of cell science **125**, 3723–3731.

Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T. 2015. Photoperiodic flowering: time measurement mechanisms in leaves. Annual review of plant biology **66**, 441–464.

Spanudakis E, Jackson S. 2014. The role of microRNAs in the control of flowering time. Journal of experimental botany **65**, 365–380.

Sparks E, Wachsman G, Benfey PN. 2013. Spatiotemporal signalling in plant development. Nature reviews. Genetics **14**, 631–644.

Srikanth A, Schmid M. 2011. Regulation of flowering time: all roads lead to Rome. Cellular and molecular life sciences: CMLS **68**, 2013–2037.

Suárez-López P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G. 2001. CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature **410**, 1116–1120.

Sugliani M, Brambilla V, Clerkx EJM, Koornneef M, Soppe WJJ. 2010. The conserved splicing factor SUA controls alternative splicing of the developmental regulator ABI3 in Arabidopsis. The Plant cell **22**, 1936–1946.

Sundell D, Mannapperuma C, Netotea S, Delhomme N, Lin Y-C, Sjödin A, Van de Peer Y, Jansson S, Hvidsten TR, Street NR. 2015. The Plant Genome Integrative Explorer Resource: PlantGenIE.org. The New phytologist **208**, 1149–1156.

Sung S, Amasino RM. 2004. Vernalization and epigenetics: how plants remember winter. Current opinion in plant biology **7**, 4–10.

Sung S, Schmitz RJ, Amasino R. 2007. The Role of VIN3-LIKE Genes in Environmentally Induced Epigenetic Regulation of Flowering. Plant signaling & behavior **2**, 127–128.

Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PloS one **6**, e21800.

Sureshkumar S, Dent C, Seleznev A, Tasset C, Balasubramanian S. 2016. Nonsense-mediated mRNA decay modulates FLM-dependent thermosensory flowering response in Arabidopsis. Nature plants **2**, 16055.

Swiezewski S, Liu F, Magusin A, Dean C. 2009. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature **462**, 799–802.

Syed NH, Kalyna M, Marquez Y, Barta A, Brown JWS. 2012. Alternative splicing in plants--coming of age. Trends in plant science **17**, 616–623.

Talbert PB, Henikoff S. 2014. Environmental responses mediated by histone variants. Trends in cell biology **24**, 642–650.

Tamada Y, Yun J-Y, Woo SC, Amasino RM. 2009. ARABIDOPSIS TRITHORAX-RELATED7 is required for methylation of lysine 4 of histone H3 and for transcriptional activation of FLOWERING LOCUS C. The Plant cell **21**, 3257–3269.

Taoka K-I, Ohki I, Tsuji H, *et al.* 2011. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. Nature **476**, 332–335.

Team RC. 2015. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2014.

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G. 2004. Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science **303**, 1003–1006.

Verhage L, Severing EI, Bucher J, Lammers M, Busscher-Lange J, Bonnema G, Rodenburg N, Proveniers MCG, Angenent GC, Immink RGH. 2017. Splicing-related genes are alternatively spliced upon changes in ambient temperatures in plants. PloS one **12**, e0172950.

Wagner EJ, Carpenter PB. 2012. Understanding the language of Lys36 methylation at histone H3. Nature reviews. Molecular cell biology **13**, 115–126.

Walles B. 1991. Steeves, T. A. & Sussex, I. M. 1989 Patterns in plant development. Nordic journal of botany **11**, 204–204.

Wang B-B, Brendel V. 2004. The ASRG database: identification and survey of Arabidopsis thaliana genes involved in pre-mRNA splicing. Genome biology **5**, R102.

Wang Z, Ji H, Yuan B, Wang S, Su C, Yao B, Zhao H, Li X. 2015*a*. ABA signalling is fine-tuned by antagonistic HAB1 variants. Nature communications **6**, 8138.

Wang Y, Wang Z. 2015. Efficient backsplicing produces translatable circular mRNAs. RNA **21**, 172–179.

Wang X, Wu F, Xie Q, *et al.* 2012. SKIP Is a Component of the Spliceosome Linking Alternative Splicing and the Circadian Clock in Arabidopsis. The Plant cell **24**, 3278–3295.

Wang Z-P, Xing H-L, Dong L, Zhang H-Y, Han C-Y, Wang X-C, Chen Q-J. 2015*b*. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome biology **16**, 144.

Werner JD, Borevitz JO, Warthmann N, Trainer GT, Ecker JR, Chory J, Weigel D. 2005. Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural flowering-time variation. Proceedings of the National Academy of Sciences of the United States of America **102**, 2460–2465.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D. 2005. Integration of spatial and temporal information during floral induction in Arabidopsis. Science **309**, 1056–1059.

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA. 2006. The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proceedings of the National Academy of Sciences of the United States of America **103**, 14631–14636.

Wu G, Park MY, Conway SR, Wang J-W, Weigel D, Poethig RS. 2009. The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. Cell **138**, 750–759.

Wu G, Poethig RS. 2006. Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. Development **133**, 3539–3547.

Xiong L, Gong Z, Rock CD, Subramanian S, Guo Y, Xu W, Galbraith D, Zhu JK. 2001. Modulation of abscisic acid signal transduction and biosynthesis by an Sm-like protein in Arabidopsis. Developmental cell **1**, 771–781.

Xu M, Hu T, Zhao J, Park M-Y, Earley KW, Wu G, Yang L, Poethig RS. 2016. Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in Arabidopsis thaliana. PLoS genetics **12**, e1006263.

Xu F, Xu S, Wiermer M, Zhang Y, Li X. 2012. The cyclin L homolog MOS12 and the MOS4-associated complex are required for the proper splicing of plant resistance genes. The Plant journal: for cell and molecular biology **70**, 916–928.

Xu L, Zhao Z, Dong A, Soubigou-Taconnat L, Renou J-P, Steinmetz A, Shen W-H. 2008. Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in Arabidopsis thaliana. Molecular and cellular biology **28**, 1348–1360.

Yadav RK, Perales M, Gruel J, Girke T, Jönsson H, Reddy GV. 2011. WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. Genes & development **25**, 2025–2030.

Yamaguchi A, Abe M. 2012. Regulation of reproductive development by non-coding RNA in Arabidopsis: to flower or not to flower. Journal of plant research **125**, 693–704.

Ye C-Y, Chen L, Liu C, Zhu Q-H, Fan L. 2015. Widespread noncoding circular RNAs in plants. The New phytologist **208**, 88–95.

Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. 2013. Long non-coding RNA: a new player in cancer. Journal of hematology & oncology **6**, 37.

Zhang X-N, Mount SM. 2009. Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development. Plant physiology **150**, 1450–1458.

Zhu Q-H, Helliwell CA. 2011. Regulation of flowering time and floral patterning by miR172. Journal of experimental botany **62**, 487–495.

Zhu Q-H, Wang M-B. 2012. Molecular Functions of Long Non-Coding RNAs in Plants. Genes **3**, 176–190.

PUBLICATIONS

Capovilla G, Symeonidi E, Wu R, Schmid M: Contribution of major FLM isoforms to temperature-dependent mediated flowering in *Arabidopsis thaliana*. *Accepted for publication (J. Exp. Bot.)*

You Y, Sawikowska A, Neumann M, Posé D, **Capovilla G**, Langenecker T, Neher RA, Krajewski P, Schmid M: Temporal dynamics of gene expression and histone marks at the *Arabidopsis* shoot meristem during flowering. Nat. Commun. 2017, 8:15120.

Conn VM, Hugouvieux V, Nayak A, Conos SA, **Capovilla G**, Cildir G, Jourdain A, Tergaonkar V, Schmid M, Zubieta C, Conn SJ: A circRNA from *SEPALLATA3* regulates splicing of its cognate mRNA through R-loop formation. Nat. Plants 2017, 3:17053.

Capovilla G, Pajoro A, Immink RGH, Schmid M: Role of alternative pre-mRNA splicing in temperature signaling. Curr. Opin. Plant Biol. 2015, 27:97–103.

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

MANUSCRIPT IN PREPARATION

Capovilla G, Delhomme N, Collani S, Shutava I, Symeonidi E, Bezrukov I, Schmid M: PORCUPINE regulates development in response to temperature variations through alternative splicing in *Arabidopsis thaliana*. *In preparation*