The role of a canonical Poly(A) Polymerase in organ-identity dependent size regulation in *Arabidopsis thaliana*

Lang Son Vi

A thesis submitted to the University of East Anglia
for the degree of Doctor of Philosophy

John Innes Centre
Norwich

October 2011

Abstract

Different organs of plants and animals grow to characteristic sizes and shaped that depend on their identity. We still know surprisingly little about how organ identity modified patterns of growth within the primordium to achieve the correct final size.

Here, I report on a novel mutant in *Arabidopsis thaliana* that forms smaller leaves, but bigger flowers. Intriguingly, the effect of the mutation strictly depends on the identity of the organ, marking out this mutation as a promising entry point to studying the relation between organ identity and growth. The phenotypes are due to a partial loss of function of one of three canonical nuclear poly(A) polymerases (PAP) in *A. thaliana*, PAPS1. By contrast, even complete loss of function of the other two gene-family members only causes subtle phenotypes, indicating functional specialization amongst the different genes. Using promoter-swap and chimaeric-protein constructs, this functional specificity is shown to be largely encoded in the C-terminal non-catalytic domains of the different proteins. Using a mRNA fractionation coupled with microarray, I have defined a set of mRNAs whose poly(A)-tails and stability are changed in the mutant compared to wild type. The microarray analysis suggests that the reduced leaf-growth results from an ectopic upregulation of pathogen-reponse genes. Analysis of plants that are chimaeric for the *paps1-1* mutation shows that *PAPS1* acts cell autonomously to control petal size and *paps1-1* chimerism in the internal layer of the meristem invariably causes the meristem to split into two or more meristems. I propose that different pre-mRNAs are preferentially processed by different canonical PAPs with distinct outcomes in terms of their poly(A) tail lengths. This would open up an additional level of gene expression whereby cells could co-regulate large numbers of transcripts in response to developmental or environmental changes by modulating the balance of canonical PAP isoform activities. I also discuss possible reasons for the identity-dependence of the *paps1-1* mutant phenotypes.

Acknowledgements

My greatest thanks are to my supervisor Michael Lenhard for providing a fantastic work atmosphere, invaluable scientific advices and for his interest in my work. Michael gives me the full freedom to work and create new ideas. I have learnt a lot from the way he analyzes and criticizes the ideas both on this project and on science in general, which he often openly and honestly shares with the lab. His creativity and vision in future research motivate me a lot.

I am grateful to my advisors Robert Sablowski, who later becomes my primary supervisor, and John Doonan for constructive suggestions during my supervisory committee meetings.
I thank Simon Swiezewski and Fuquan Liu for fantastic discussion and motivation on this project. I thank Cornelia De Moore, James Manley, Nishta Rao, Ghanasyam Rallapalli for collaborating and teaching new techniques on this project. My funding for the PhD is supported by the John Innes Rotation PhD program , BBSRC and the German research foundation DFG.

A very big thanks to all past and present members of the Lenhard group and the Baurle group for their help and discussion during my PhD. Special thanks to Holger Breuninger, who always helps me with computing; Adrien Sicard whom I share bench with and for making the lab less lonely during the weekends; Nicola Stacey for her invaluable English support; Peggy Lange, Gerda Trost and Hjördis Czesnick and Christian Kappel who directly involve in some of the work in this project; the new lab members in Potsdam, Germany for making my movement to the new lab in Potsdam, Germany very enjoyable; and two very likable students in the lab at JIC, Lena Stransfeld and Nikolai Adamski.

My special thanks to Tung Le and Ngat Tran, my badminton mates Dora Szakonyi, Quynh Le, Thanh Luong, Frank Bauer and many other friends in Norwich and Potsdam with whom I have had so much fun with during my free time here.

Lastly, I would like to thank my whole family including my parents, my sister, my cousins and my aunts and my dearest friends Nguyen Thi Thuy Hong, Pham Thi Hong Van, An Dau for their constant communication and encouragements throughout my PhD.

Contents

List of Figures

List of Tables and Boxes

1 Chapter 1: Introduction

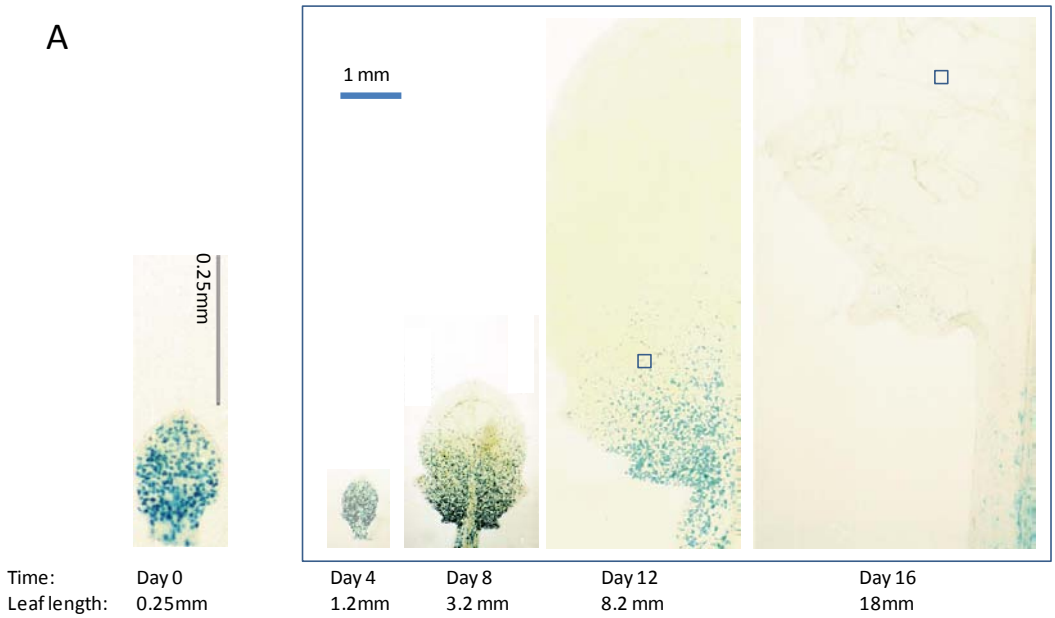
1.1 Organ size control in plants:

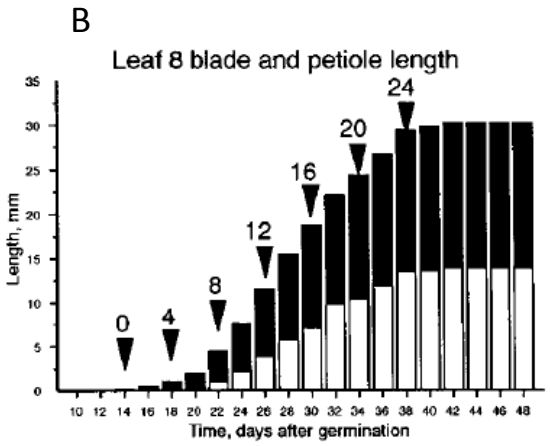
Plant organ size is under strong genetic control. Studies over the last 20-30 years in Arabidopsis have identified a number of genes that play a role in organ size control. In this section, I first introduce the organ grow process in plants. Secondly, I introduce the definition of a size regulator. I thirdly suggest a characterization pipeline that, in my opinion, is useful in understanding how the size regulator acts molecularly. Using these bullet points, I fourthly summarize the studies about the genes involved in size control in Arabidopsis. Fifthly, I pinpoint the general deduction from these studies that contributes to our current knowledge about how plant organ size is regulated. Finally, for the interest of this project, I emphasized what is known about how different organs can reach different sizes.

1.1.1 The organ growth process:

An organ is first inniated in the peripheral zone of the meristem (see further introduction in Chapter 5). After initiation, two basic cellular processes are tightly associated with the organ growth: cell proliferation and cell expansion, the patterns of which change spatially within an organ and also change temporally during the course of growth. Cell proliferation here implies the process of cell division (the forming of a new cell wall) and the growth of two daughter cells to reach the size of their mother cell. Cell division alone (e.g. the first few divisions in the early embryonic development after fertilization) does not cause growth. Our understanding of this pattern in leaves comes from a very thorough study by Donnelly and colleagues, (Donnelly et al., 1999) and reference therein, who used a cyclin promoter and a modified *GUS* (beta-glucuronidase) reporter gene (*pCYC1At::CDBGUS, the CDBGUS is a GUS* gene fused to a Cyclin Destruction Box) to specifically mark the dividing cells (**Figure 1.1**).

Initially, cell proliferation occurs evenly throughout the entire primordium (**Figure 1.1A**). Gradually, cell proliferations become restricted in a basipetal pattern. Proliferation is arrested first at the distal tip of the organ, and is then progressively arrested at the more basal parts, until all cells have stopped dividing. Cell expansion (the enlargement of cell size due to water uptake to vacuole) happens when a cell exits from the proliferating/cycling state. At high cycling rate, cells remain at a constant size of approximately 75 μm^2. At exit from proliferating, i.e. day 12 in leaf 8, mesophyll cells are about 200 μm^2 in area and continue to expand 4- to 5- fold. Since cell expansion starts right after the exit from the proliferating state, cell expansion also shows a basipetal pattern starting first at the tip and reaching the base of the leaf last. However, this basipetal pattern is very weak; instead, it seems that after day 12, there is a burst in cell expansion rate across all regions in leaves (Donnelly et al., 1999).





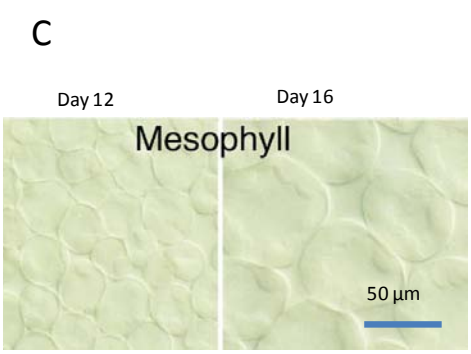


Figure 1.1. The patterns of cell division and cell expansion in Arabidopsis wild-type leaf 8.

Figure is modified from Donnelly et al., 1998.

A. Time/developmental series of cell division in the 8th leaf. The 8th Leaf is sampled at idicated days after emergence and stained with GUS to visualize dividing cells (blue staining). Numbers below the image indicate the length of the leaf blade at the corresponding time point. Note that the magnification of the image at day 0 is 10x higher than the magnification of images of later dates. The small squares indicate the positions where cell size is examined in panel C.

B. Time/developmetal series of blade length (black columns) and petiole length (white columns) of leaf 8. Arrows point at the time when samples were taken for GUS stain in panel A.

C. Palisade mesophyll cells start cell expansion after day 12.

To sum up, leaves grow mainly by cell proliferation at early stages of growth; later, cell divisions ceases locally accompanied by a gradual increase in cell size. Cell divisions stop completely when a leaf reaches about half of its final size (Day 16 **Figure 1.1A** and **B**), after that the organ grows solely by cell expansion. Cell expansion is normally coupled with endoreduplication, the process in which cells undergo several rounds of DNA replication without entering mitosis.

For the interest of this project, it is of note that there is a difference between leaves and petals in the endoreduplication during cell expansion. Cells in floral organs and especially in petals seldom endoreduplicate, while leaf cells undergo several rounds of endoreduplication.

1.1.2 A definition: What is a size regulator?

To be a size regulator, reduction and enhancement of the activity of the gene should show opposite effects on final organ size. Changes in shapes, which are caused by local growth regulations that do not results in overall size change is not considered here. The enhanced or reduced activity of the gene can be classified into four classes (See **Box 1.1**).This criterion is denoted as the 'size change in both directions' criterion.

Box 1.1. Kinds of perturbations to a gene in Arabidopsis thaliana:

Loss-of-function mutants: mostly point mutations, T-DNA insertions, deletion mutants or transgenic RNAi knockdown.Gain-of-function mutants: This kind of perturbation can have very diverse causes e.g. point mutations that make the protein constitutively active. Over-expressors: can be achieved by three ways:

iiia) Constitutively expressed: highly expressed, constitutive and spatially non-restricted expressors using 35S promoter (even though 35S promoter is not fully/constitutively active in the centre of the meristem).

iiib) Endogenous overexpressor: increase the mRNA level although still under control of the endogenous promoter such that the temporal and spatial expression patterns of the gene are largely unchanged. This can be achieved by supplementing the plant with extra copies of the genomic version of the gene or by using a two component system, e.g. LhG4/pOp, with the gene's own promoter.

Mis-expression using a different promoter than the gene's own promoter, yet not 35S promoter. Mis-expression, may but does not necessarily result in an increased in mRNA level compared to WT.

Care should be taken when applying this definition. Firstly, false negatives: what if the gene can only regulate organ size in one direction no matter how it is manipulated genetically?

This could be for several reasons, which are related to structure of the pathway in which the gene acts. Firstly, plant genes often have homologs, so if overexpression shows a phenotype in one direction (e.g. smaller organs), disruption of the gene, however, may not cause a phenotype in a reverse direction (e.g. larger organs) unless the activity from redundant homologs is also disrupted. Secondly, genes often act in pathways, if plants lose the activity of one gene, they might display a phenotype in one direction (e.g. smaller organs), but in order to elicit the opposite phenotype (e.g. bigger organs), the whole pathway needs to be upregulated. So clearly, single gene overexpression might not be sufficient. Even if in theory, single gene mis-regulation can create a 'bigger size' phenotype; in practice it may prove to be difficult to up-regulate the gene activity in a well co-ordinated way to give a well developed and bigger organ. For example, mutation in *jag* leads to a smaller organ, but over-expressing *JAG* by either *35S* promoter or *AP3* promoter failed to create bigger organs; rather, some strong expression lines disrupted with the organ development and caused an outgrowth of snake-like stumps in the place of flowers (Dinneny et al., 2004, Ohno et al., 2004).

Secondly, false positives: this is because there is a strong correlation between late flowering phenotype and larger flowers. Also mutants that are infertile tend to have larger petals/flowers (SV, ML personal observation). It is questionable; whether the increased growth in floral organs is the indirect effect of having more photo-synthetic organs (late flowering mutants produced more leaves at bolting).

Additionally, mis-regulations that increase size are much more likely to identify genes with a rather direct role in growth than are mis-regulations that give small organs. This is because growth/organ-size phenotypes have a very strong one-directional correlation with gene-misregulation, i.e. it is very likely that a perturbation to the activity of genes results in unhealthy, small plants. We found a lot of 'smaller size' mutants in any mutagenesis screens. In contrast, the other consequence of gene misregulation(s), which is the creation of normally developed and larger plants, happens rarely.

Taken together, it will be a clear case for a size regulator if one can show the size change can happen in both directions by different means of genetic mis-regulation of the gene. In other cases, especially, for genes whose mutants are smaller and whose over-expressors are not larger, other evidence needs to be gathered to demonstrate the direct role of the gene in organ size control.

1.1.3 Properties to be examined for organ size mutants:

For easier browsing through the literature of organ size regulators, I list which information I consider interesting for understanding the mechanism of organ size regulation (**Box 1.2**).

Box 1.2. Characterization pipelines for a size regulators:

i. *Cell number and cell size: Is the change in the final organ size due to cell number of cell size?*

ii. *Growth patterns: Are there changes in the temporal and spatial patterns of cell proliferation and/or cell expansion?*

iii. *Growth rate or growth period: Is the rate of growth or the period of growth is changed?*

iv. *Directions of the regulation: Do disruption and over-expression have opposite effects on organ size?*

v. *Pathway: Is there a genetic interaction of the gene with other known pathway(s) involved in size regulation?*

vi. *Organ type: Does the gene affect the size of more than one organ type? If yes, how do the degree of change in different organs compare?*

vii. *Expression domain: Where is the gene expressed?*

1.1.4 The players: Genes involved in organ size control in *Arabiodopsis thaliana:*

Using the seven points above, I summarized the genes that are described as organ size regulators in **Figure 1.2**, **Figure 1.4** and **Table 1.1.**

1.1.5 General conclusions about the mechanism of organ size control in *Arabiodopsis thaliana*:

First conclusion: Size control pathways are mostly genetically independent with several master regulators that do not converge (**Figure 1.2** (Breuninger and Lenhard, 2010)) suggesting size control is a complex trait.

Second conclusion: To make large organs, plants must positively regulate cell proliferation and cell expansion in a well co-ordinated way. Ectopically overexpressing components of the cell cycle alone (e.g. *35S::CYCD3;1* (Dewitte et al., 2003)) increases the cell division but fails to give bigger organs. In fact, *35S::CYCD3;1* transgenic plants are smaller because cells fail to expand properly. Similarly, ectopic upregulation of the endoreduplication pathway, which is necessary for cell expansion, (*35S::FZR2* (Larson-Rabin et al., 2009)) increases the size of individual cells, but fails to give bigger organs, because cells fail to proliferate properly.

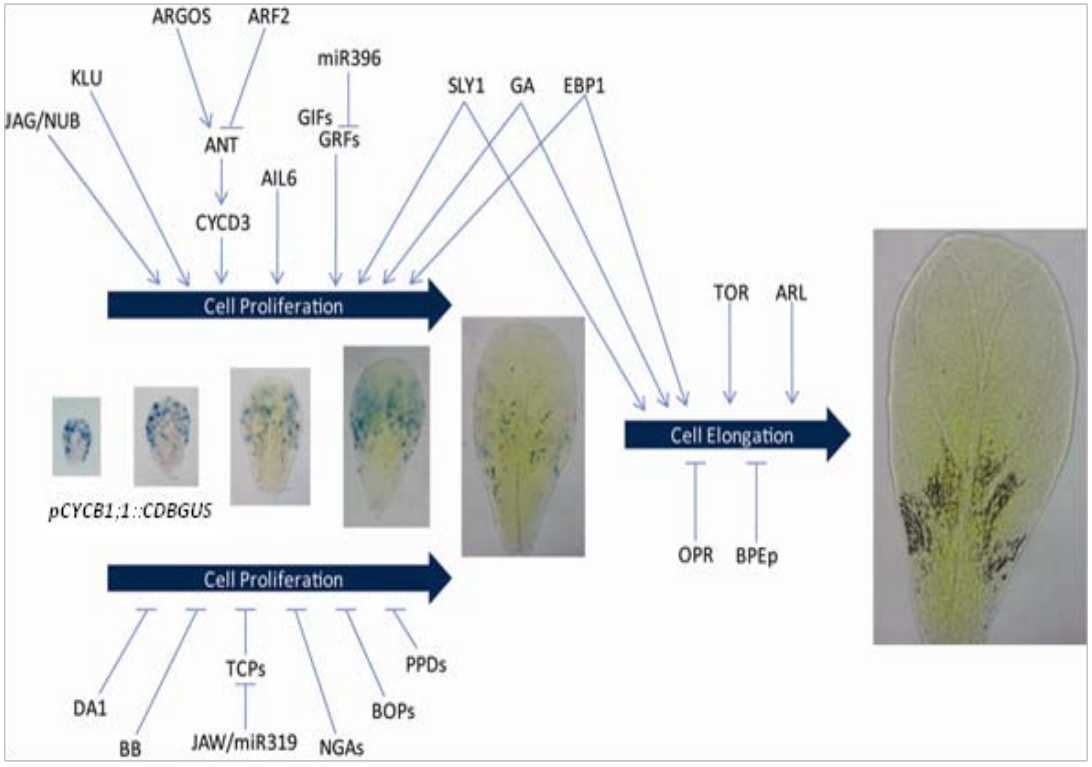


Figure 1.2. Genetic pathways that regulate organ growth in *Arabidopsis.*

(from (Breuninger and Lenhard, 2010)).

The petals are GUS (beta-glucuronidase)-stained to visualize the cell division pattern. See text for details on the size regulators.

Third conclusion: The interaction between cell division and cell expansion is very complex. Illustrated in **Figure 1.3** are all the theorically possible combinations between changes in cell division and cell expansion and their outcome to final organ size. In practice, if one project the known size regulators to this figure, almost all the theoretical possibilities are backed up with experimental phenotypes (**Figure 1.4**). The only one theoretical possibility that has not yet been backed up by experimental evidence is 'bigger' triangle 1 possibility (**Figure 1.4**), in which bigger organs result from an increase in cell divisions surpassing a decrease in cell expansion. Many size regulators change final organ size by affecting only cell division or by affecting only cell expansion (the genotypes that lie on line 3, 4 (cell expansion only) and line 5, 6 (cell division only) in **Figure 1.4**). This suggests that cell number and cell size can be regulated indendently. So co-ordination between cell division and cell expansion is necessary for organ size regulation, a misregulation of a gene may affect cell division alone or cell expansion alone or affecting both (the two squares and four triangles and line 1,2 in **Figure 1.3**).

Fourth conclusion: Mis-regulating the timing of the transition from cell proliferation to cell expansion seems to be a common defect for many size-regulating mutants (*klu*, *bb*, *da1*, *tcps*, *miR319*, *grf*, *an3*, *ppd1/2*, *argos*, and *arf2*) see **Table 1.1** for reference and more details.

Fifth conclusion: changes in final organ size is mostly observed as the result of a prolong duration of growth rather than a change in the growth rate. This property is not well documented for most of the genes. However for the cases where it has been analyzed, it seems that most of the factors regulate the period of growth *(*e.g. *KLUH, BB, DA1,ANT)* rather than rate of growth (e.g. proteasomal proteins).

Sixth conclusion: as mentioned above (in section 1.1.2. A definition: what is a size regulator?), a size regulator can act in one direction (size regulators that are in Blue colour in **Figure 1.4**) or in both directions (size regulators that are in Orange, Pink and Red colour in **Figure 1.4**. If in the other direction of genetic manipulation, the phenotype is indistinguishable from wild-type (for example *ail5* single mutant) the gene might be still accept to be a size regulator because one can argue that phenotype in one direction was masked by genetic redundancy. However, for genes that showed an unexpected phenotype in the other direction of genetic manipulation, it is not trivial to explain and draw conclusions about the specific function of this gene in organ size control. For example, the *jag* mutant has smaller organs, but overexpressing *JAG* under the control of the *35S* promoter or the petal specific *AP3* promoter interfered with organ development and failed to give bigger organs. It is subjective to say how specific is the role of JAG is in controlling organ size compared to its role in organ development. Additionally, for genes that act in both directions,

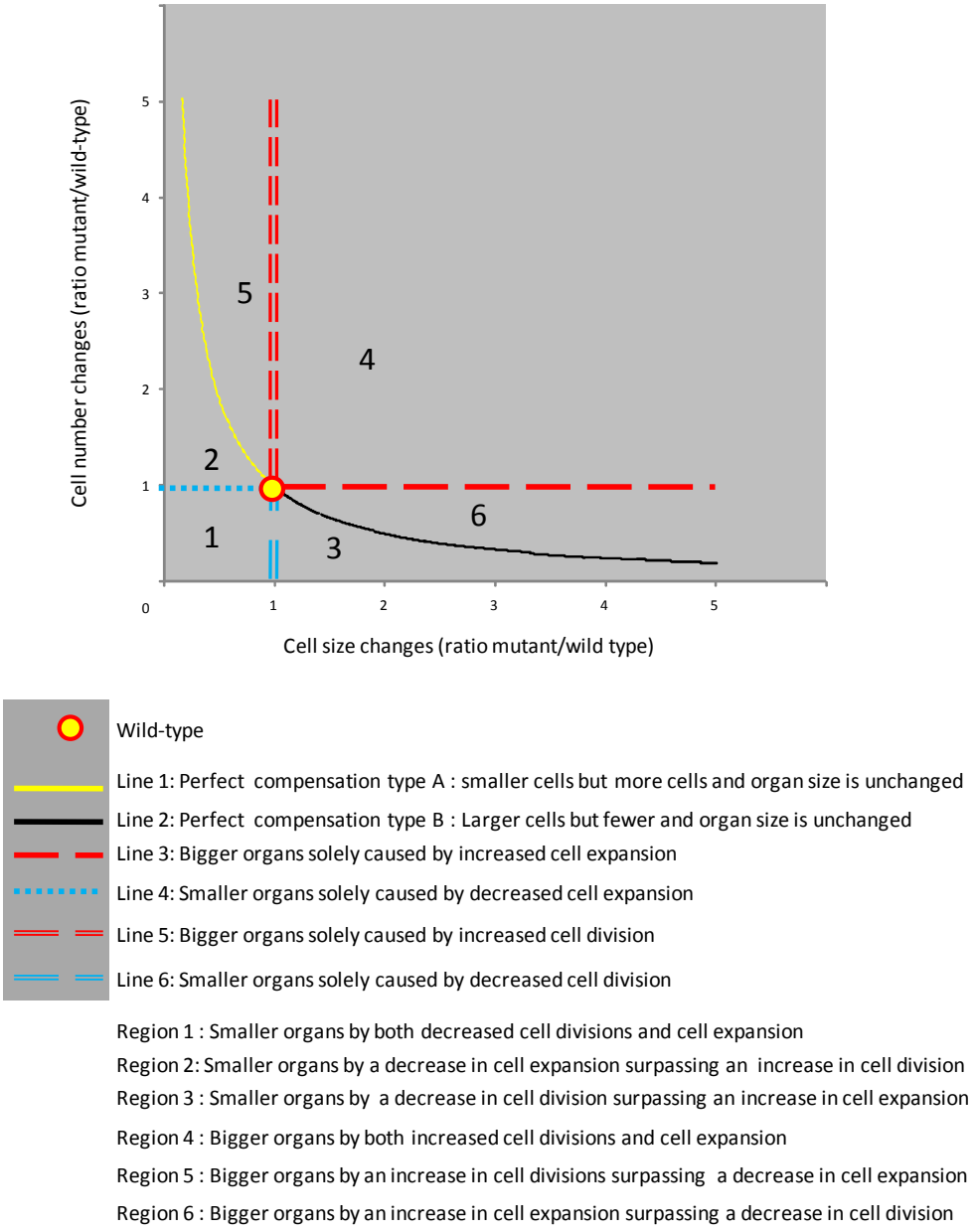


Figure 1.3. All theorically possible combinations between changes in cell number and in the cell size and their outcomes in terms of final organ size.

The six lines (number 1-6) divide the morphological space, generated by the combination of changes in cell size and cell number, into six region (number 1-6).

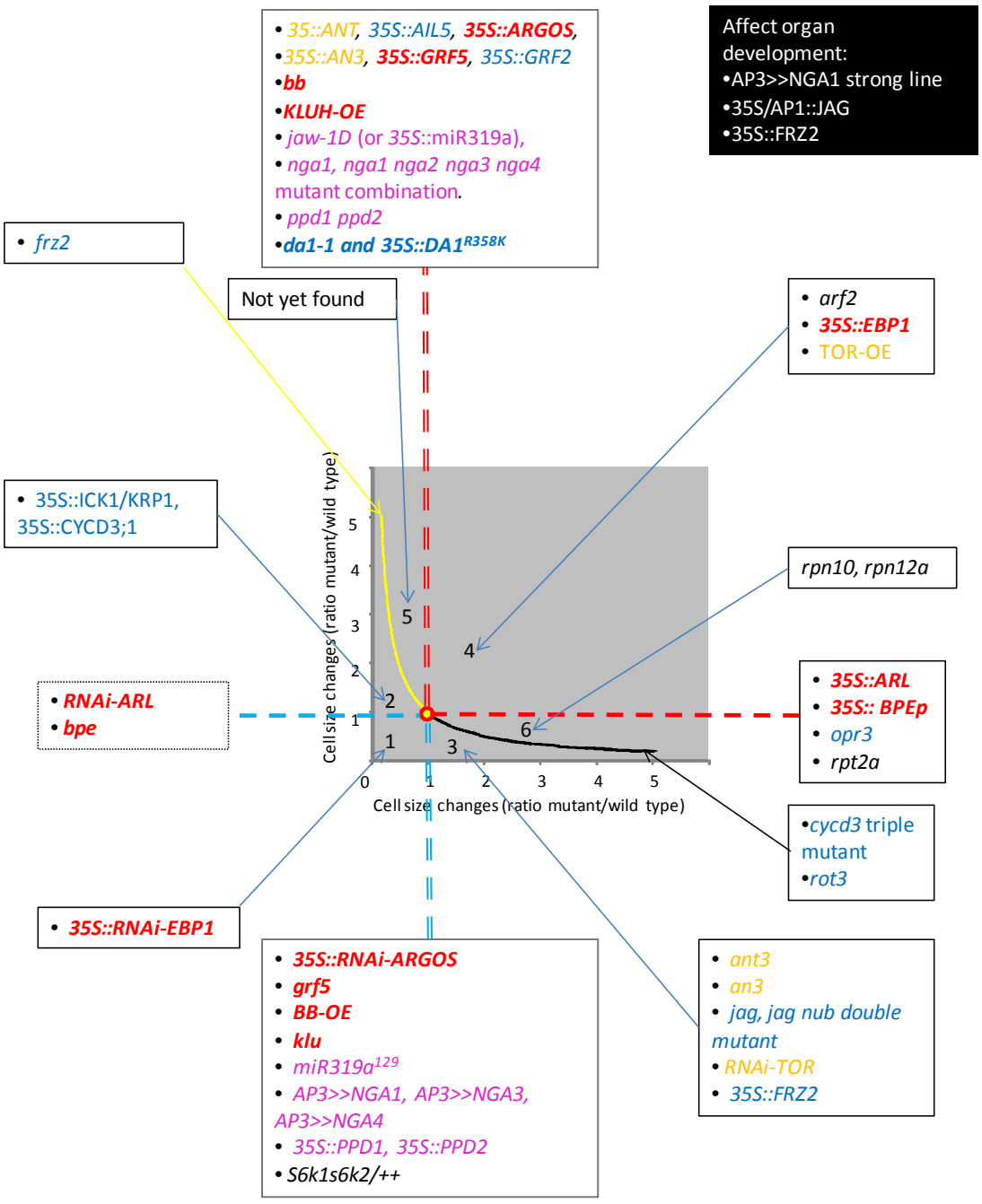


Figure 1.4. The effects of known size regulators on cell number and cell size and the outcomes in terms of final organ size.

The lines and triangle annotations are adapted as for figure 1.3. OE means overexpressor. Font/Colour codes of the text of the name of the size regulators:

Red and Bold = genes that can act in both directions (increase or decrease the final organ size) and display mirrored changes in cell number and cell size in opposite directions.

Pink = genes that can act in both directions (increase or decrease the final organ size) but cell number and cell size were not determined for one direction.

Figure 1.4. The effects of known size regulators on cell number and cell size and the outcomes in terms of final organ size.

Orange = genes that can act in both directions (promote or repress the final organ size), but exhibit partial compensation.

Black = genes that were found to act in one direction but have not been examined in the other direction.

Blue = genes that change organ size in one direction but in the other direction affect organ development or do not have an effect on organ size

In the black background box with white text : genes whose mis-regulation causes unexpected effects on organ development.

Table 1.1. Arabidopsis genes involved in organ size control:

Genes are sorted using three orders: the first order is cell division vs. cell expansion: genes affecting mainly cell divisions are listed first; the second order is positive vs. negative regulation: the positive regulators comes first; the third order is alphabetical order of the gene name. The phenotype of mutant and other kinds of mis-regulations are described (when possible) including the constructs for transgenic plants or types of mutations. The effects on cell division (+/- : more cells/less cells), effects on cell expansion (+/-: bigger cells/smaller cells) are shown. The known genetic/physical interactions are described. The spatial expression domains are described with a focus on leaves and flowers. The organs in which the effects are observed are mentioned. ND means not determined or no data available, EdOX: endogenous overexpressor.

[1] *ant ail6* (ANTERGUMENTA LIKE 6-At5g10510) double mutant display floral organ identity defect. [2] gain of function construct: rGRF2 resistant to miR396. [3] AP1::JAG causes malformed flowers, 35S::JAG causes bract like structure that subtend flowers and fusion cotyledons. [4] The double mutant *bop1 bop2* also produced bract-like organs which subtended flowers. The phenotype that resembles the weakly expressed 35S::JAG plants. [5] *jaw1-D* mutants cause larger leaves with an excessive out growth of the leaf margins. [6] *ppd1ppd2* double mutants mainly affect the dispersed meristematic cells like meristemoid, precursors of guard cells. [7] resistant form of TCP4 to miR319. [8] AP3::mTCP4 causes malformed flowers. [9] Mutants have decreased polyploidy level in leaf cells. [10] s6k1 s6k2/++ hemizygous mutants or the RNAiS6K plants showed high degree of phenotypic instability and size variation. All mutants however have smaller leaves with smaller cells and larger flowers, in which cell size were not measured. The larger flowers have almost double the DNA content, suggesting the defect in S6K causes chromosomal instability. [11] It is suggested that JA post-transcriptionally regulate the expression of the BPEp mRNA as pBPE::GUS do not response to JA, but BPEp mRNA appears to be more abundant when inducing plants with JA. [12] *opr3* mutants also affect the vein patterning in petals. [13] *rot3* mutants have wider leaf blades but shorter petioles and shorter pistils and stamens. [14] *rpn12a-1* and *rpn10-1* mutants have larger flowers with reduced cell number but larger cells without changes in polyploidy. The rosette of the mutant grows more slowly and the leaves are pointed. [a] Cell size changes without change in ploidy level; [b] Cell size changes with changes in polidy level. [pnd] the ploidy was not determined.

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell <u>division</u>:

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
1	*AINTEGUMENTA LIKE 5 (AIL5)*- At5g57390	AP2-domain TF	single	0	0	0	*35S::AIL5*	+	0	+	Homologous to ANT	Strong in developing floral primordia	Leaves and flowers similarly	(Nole-Wilson et al., 2005)
2	*AINTEGUMENTA (ANT)*- At4g37750	AP2-domain TF	single	-	+	-	*35S::ANT*	+	0	+	Up regulates *CYCD3;1* Synergistic interaction with AIL6[1]	growing domain of developing organs	Petals (stronger effect) and leaves	(Krizek, 2009, Mizukami and Fischer, 2000) (Nole-Wilson et al., 2005)
3	*AUXIN RESPONSE FACTOR 2 (ARF2)*- At5g62000	ARF TF	Single mutant	+	+	+	ND	ND	ND	ND	Upregulate *ANT, CYCD3;1* expression in mature leaves	Leaves, young floral organs, siliques, funiculus of ovule, embryo.	Seeds, leaves, sepals, gynoecia, but not petals.	(Schruff et al., 2006)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
4	*AUXIN-REGULATED GENE INVLOVED IN ORGAN SIZE (ARGOS)*-At3g59900	Novel-plant specific	*35S::antisense ARGOS*	-	0	-	*35S::sense ARGOS*	+	0	+	Upregulates *ANT* and induced by auxin	Young leaves and flowers.	Leaves, ND in flowers	(Hu et al., 2003)
5	*ANGUSTIFOLIA3 (AN3)/GRF INTERACTING FACTOR 1 (GIF1)*-At5g28640	SYT TF	single mutant	-	+	-	ND	+	0	+	In teract with *GRF5* in Y2H	Strong in basal region of young leaf primodia and floral buds	Leaves and flowers similarly.	(Horiguchi et al., 2005, Kim and Kende, 2004, Rodriguez et al., 2010)
6	*GROWTH REGULATING FACTOR 2(GRF2*)-At4g37740	QLQ/WRC TF	ND	ND	ND	ND	*35S::rGRF2*[2]	+	0	+	Is a target of *miR396b*	Young tissues, and in leaf has a basipetal expression.	Leaves, ND in flowers	(Kim et al., 2003, Horiguchi et al., 2005)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
7	*GROWTH REGULATING FACTOR 5(GRF5)*-At3g13960	QLQ/WRC TF	Single mutant	-	0	-	Not mentioned	+	0	+	In teract with *GIF1/AN3* in Y2H	Leaf primordial, NA in flowers.	Leaf, ND in flowers	(Kim et al., 2003)
8	*JAGGED (JAG)*-At1g68480	Zn finger TF	Single mutant	-	+	-	*AP1::JAG* and *35S::JAG*	causes bract-like structure that subtend flowers. Fusion cotyledon and Malformed floral organs			*NUB*	young primodia of lateral organs	Leaf and flowers [3]	(Dinneny et al., 2006)
9	*KLUH/CYP78A5* (KLU)-At1g13710	Cytochrom P450	Single mutant	-	0	-	EdOX	+	0	+	*CYP78A7*	Young organs primodia and restricted to the peripheral area at later stage of growth.	Leaves and flowers similarly.	(Adamski et al., 2009, Anastasiou et al., 2007, Eriksson et al., 2010)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
10	*miR396b*-At5g35407	miRNA	ND	ND	ND	ND	*35S::mir396b*	-	+	-	Targets GRF2. Act together with AN3 in meristem development, Is up-regulated by TCP4	Leaves, the mRNA level increases as leaves get older. Expressed more strongly at the tip compared to the base of the leaf . ND in flowers	Leaves, ND in flowers.	(Rodriguez et al., 2010)
11	*BIGBROTHER (BB)*-At3g63530	E3-ubiquitin ligase	Single mutant	+	0	+	EdOE	-	0	-	ND	Young organs, expression decreases during growth.	Flowers only, leaves not affected	(Disch et al., 2006)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
12	*BLADE ON PETIOLE 1 (BOP1)*-At3g57130) and *BLADE ON PETIOLE 2 (BOP2*-At2g41370)	TF	*bop1/bop2* double mutant	ND	ND	+	*35S::BOP 1* *35S::BOP 2*	ND	ND	-	Upregulates *JAG/NUB*	incipient of young leaf/flower primordia	Leaves and flowers [4]	(Norberg et al., 2005)
13	*DA1 (DA1)*-At1g19270	Ubiquitin interaction domain protein	Single, gain of function allele	+	0	+	*35S::DA1*	0	0	0	ND	Young organs	Leaves and flowers similarly	(Li et al., 2008)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
14	*miRNA319a/ (JAW)*-At4g23713	miRNA	Single *miR319a*[129]	ND	ND	-	*EdOE: jaw1-D* and *35S::miR319a*	+	0	+	Targets *TCP2,3,4, 10,24*	lowly expressed in both leaves and flowers	Leaves[5], ND on flowers.	(Palatnik et al., 2003, Nag et al., 2009)
15	*NGATHA 1 (NGA1)* (At2g46870)	B3 TF	Single mutant	+	ND	+	*AP3>>NGA1*	Weak line caused reduced petal growth. Strong line caused transformation of floral organs in to carpeloid organs.			ND	At the distal region of lateral organs (peripherals of leaves and tips of developing gynoecium)	Floral organs, ND on leaves	(Alvarez et al., 2009, Trigueros et al., 2009)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
16	*NGATHA2 (NGA2)*-AT3g61970); *NGATHA3 (NGA3)*-AT1g01030 *NGATHA4 (NGA4)*-At4g01500	B3 TF	Triple and quadruple mutant with NGA1	+	ND	+	*AP3>>NGA3* *AP3>>NGA4*	Both transgenic lines caused reduced petal and sepal growth.			ND	*NGA4* expressed similarly to *NGA1*	Floral organs, ND on leaves	(Alvarez et al., 2009, Trigueros et al., 2009)
17	*NUB*-At1g13400	TF	double mutant *jag nub*	-	NA	-	*FIL::NUB*	Increase number of cell layers in floral organs, however size were not examined.			JAG	Young primodia of Lateral organs	Leaves and flowers similarly	(Dinneny et al., 2006)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
18	*PEAPOD1 (PPD1)*-At4g14713 and *PEAPOD2 (PPD2)*-At4g14720	Plant specific putative DNA binding protein	double mutant *ppd1* *ppd2*	+	0	+	EdOx	ND	ND	-	ND	Basipetal pattern in leaves. Strong in vascular tissues	Leaves[6], ND in flowers	(White, 2006)
19	*TCP4*-At3g15030	TF	NA	NA	NA	NA	*AP3::mTCP4* [7]	Ectopic expression causes malformed flowers			Is a target of miR319. Positively regulates miR396 expression	ND	Leaves, flowers[8].	(Nag et al., 2009)

Table 1.1A. Arabidopsis genes involved in organ size control: Genes that affect mainly cell division (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
20	Other TCP proteins *TCP2*-At4g18390, *TCP3*-At1g53230, *TCP10*-At2g31070 *TCP24*-At1g30210	TF	Evidence that these proteins negatively regulate growth are: they are all targets of *miR319*, and they are homology to the *CINCINNATA* (Crawford et al., 2004) protein in Anthirinum. No mis-regulation studies have yet been reported.								Targets of *miR396*	ND	Leaves, flowers	(Palatnik et al., 2003, Nag et al., 2009)

Table 1.1B. Arabidopsis genes involved in organ size control: Genes that affect mainly cell expansion :

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
1	*ARGOS-LIKE* *ARL*-At2g44080	Novel plant specific	Single RNAi line	0	-[a]	-	*35S::ARL*	0	+[a]	+	Induced by Brassinosteroid (BR), rescue partially *br1-119* mutant	Strongly expressed in mature region of leaves/cotyledons, not observed in leaf primodia, strong in sepals and stamen.	Leaves, ND in flowers.	(Hu et al., 2006)
2	*AtTOP6B*-At3g20780;*AtSPO11*-At3g13160;*BIN4* - At5g24630	Topoisomerase VI components	Single mutant in each gene.	ND	-[b]	-	ND	ND	ND	ND	All three mutants are BR insensitive	ND	Leaves[9], ND in flowers	(Sugimoto-Shirasu et al., 2005, Sugimoto-Shirasu et al., 2002)

Table 1.1B. Arabidopsis genes involved in organ size control: Genes that affect mainly cell expansion (continued) :

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
3	*S6 kinases1 (S6K1)*- At3g08730 and *S6 kinases2 (S6K2)*- At3g08720	Protein kinase	*s6k1 s6k2*/+ + hemizygous and RNAi line	0	-/+[b,10]	-/+[10]	ND	ND	ND	ND	*TOR E2F/RBR,* Controlling chromosomal instability	ND	Smaller leaves but larger flowers[10].	(Henriques et al., 2010)
4	BIGPETAL (BPE)- At1g59640	TF	RNAi and TDNA single	0	-[a]	-	*35S::BPEp* in *opr3* background	0	+[a]	+	Upregulated by *35S::PI* and *35S::AP3* and *35S::SEP3* but downregulated in *35S::AGAMOUS*, Upregulated by Jasmonate[11]	Two forms of transcript:*BPEub* express everywhere. *BPEp* expresses specifically in floral organs	Only petals no effect on leaves	(Szecsi et al., 2006, Brioudes et al., 2009)

Table 1.1B. Arabidopsis genes involved in organ size control: Genes that affect mainly cell expansion (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
5	*ORP3*-At2g06050	12-oxophytodienoate reductase	Single mutant	0	+ [a]	+	*35S::OPR3*	0	0	0	JA biosynthesis enzyme,positively regulate *BPEp*	ND	Petals[12], ND in leaves	(Brioudes et al., 2009)
6	*RPT2A*-At4g29040	Proteosomal subunits	Single mutant	0	+ [a]	+	ND	ND	ND	ND	Proteosomal subunits	Expressed in both leaves and flowers.	Affect leaves cotyledons and petals.	(Sonoda et al., 2009, Kurepa et al., 2009)
7	*ROTUNDIFOLIA (ROT3)*-At4g36370	oxygen binding / steroid hydroxylase	Single mutant	ND	+ [pnd]	+/- [13]	ND	ND	ND	ND	BR biosynthesis	ND	Leaves and floral organs[11]	(Tsuge et al., 1996, Ohnishi et al., 2006)

Table 1.1C. Arabidopsis genes involved in organ size control: Genes that affect both cell division and the cell expansion:

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
1	*ErbB-3 epidermal growth factor receptor binding protein1* (*EBP1*-At3g51800)	nucleolus dsRNA binding protein	RNAi line	-	-[nd]	-	*35S::EBP1*	+	+[nd]	+	Affect RBR protein level; positively regulates cyclins level in an auxin dependent manner	ND	Leaves. Flowers ND	(Horvath et al., 2006)
2	*FIZZY-RELATED 2 (FRZ2)*-At4g22910	WD40 gene	Single mutant	+	-[b]	0	*35S::FRZ2*	-	+[b]	-	Affect endoreduplication	ND	Leaves, ND on flowers.	(Larson-Rabin et al., 2009)

Table 1.1C. Arabidopsis genes involved in organ size control: Genes that affect both cell division and cell expansion (continued):

	Gene name (Abbreviation)	Protein type	Mutant phenotype				Over-expression phenotype				Interaction/pathway	Expression domain	Organ affected	Reference
			Type of mutant	Cell division	Cell expansion	Final organ size	Construct	Cell division	Cell expansion	Final organ size				
3	*RPN12A*-At1g64520	Proteasomal subunits	Single mutant	-	+[a]	+	ND	ND	ND	ND	Proteasomal subunits	ND	Flowers and rosettes in opposite direction[14]	(Kurepa et al., 2009)
4	*RPN10*-At4g38630	Proteasomal subunits	Single mutant	-	+[a]	+	ND	ND	ND	ND	Proteasomal subunits	ND	Flowers and rosettes in opposite direction[14]	(Kurepa et al., 2009)
5	*Target of rapamycin (TOR)*-At1g50030	Protein kinase	RNAi	-	+[pnd]	-	EdOx	+	+	+	Phosphporylate S6 kinase	ND	Leaves, flowers ND	(Menand et al., 2002, Deprost et al., 2007)

i.e. the final organ size mirrors each other in different mis-regulations, the underlying molecular changes to the cell number and the cell size does not always mirrors each other. If a gene displayed mirrored changes in the cell number and the cell size in opposite directions of genetic manipulations, it gets red colour in figure 1.4. Otherwise, it gets Orange colour: despite the final size change mirrors each other, the cell number and cell size changes do not. If the other direction is not yet investigated it gets the pink colour. This conclusion may be related to the third conclusion in part: The ability of genes to act in both directions in controlling size, which depends on the structure of the genetic pathways (as explained in section 1.1.2.), also depends on on its ability to overcome the compensation mechanism between cell number and cell size. Therefore, it may be hard to tell in some cases whether a gene is a size regulator or not.

Seventh conclusion: The size effect on different organs, e.g., leaves vs petals, is similar in most of the mutants. I will further discuss this conclusion in more detail in the section below.

1.1.6 How can different organs reach different size? What is the relationship between organ identity and size regulation?

Final size of different organs within one plant is often very different. Understanding the mechanism of this is an important aspect in organ size regulation, yet at the moment, it is largely unknown what makes different organs grow to different sizes. In *Arabidopsis thaliana*, the effects of a size regulator on different organs were not always examined. Many studies analyzed only one type of organ (most often leaves (L)) without information about the effect on other organs such as petals (P), for example in the studies *AIL5* (L), *ARGOS* (L), *GRF5* (L), *GRF2* (L), *PPD1/2* (L), *miR396b* (L), *EBP1* (L), *ARL* (L), *TOR* (L), *Topoisomerase VI* components (L), *FRZ2* (L), *NGA* (P). I tried to extract this information and summarized it in **Table 1.1**. In many other studies, in which this aspect was examined, the effect of the size regulators on leaves and flowers was found to be mostly similar (*KLUH, DA1, ANT, ARF2, AN3, JAG, JAW/miR319a, TCP4, BOP1/2, ROT3*). If a mutant has smaller leaves, it is often found to have smaller flowers. The exceptions are i) *bpep* and *bb* mutants where the effect seemed to be only specific to floral organs; ii) *S6 kinase* mutants (*s6k1 s6k2/*++ heterozygous), proteasomal mutants (*rpt2, rpn10, rpn12a)* where leaves are smaller, but flowers are larger than WT and iii) *cincinnata* mutants in *Anthirinum* where leaves are bigger but flowers are smaller than WT. *BPEp* is expressed only in floral organs, so it is a floral specific size regulators (Szecsi et al., 2006). For most other cases, it seems that largely, the same size regulation pathways operate in different organs resulting in similar effect to the final organ growth. Organs like petals/sepals and leaves are homologous. The transformation of identity from one organ to the other requires as few genes as one, e.g. *ap2-1* mutant (Bowman et al., 1989) and *35S::UFO::SRDX* (Chae et al.,

2008). It is therefore formally possible that organ patterning genes interact with the organ-non-specific size regulators to fine tune the growth of the organ.

In animals, there is one piece of evidence from a study in fly wings that supports the above hypothesis. In wings and halteres (similar structures, but much smaller organs than wings), a Decapentalegic (Dpp)-dependent growth pathway is differentially modified to achieve different final organ sizes. The *Ultrabithorax (Ubx)* identity gene is specifically expressed in haltere and controls both its size and identity. *Ubx* controls haltere size by restricting both the transcription and the mobility of the morphogen Decapentaplegic (Dpp), which mediates size control. *Ubx* also restricts Dpp's distribution in the haltere by increasing the amounts of the Dpp receptor, *thickveins* (Crickmore and Mann, 2006).

Here I study a novel mutant that has bigger flowers but smaller leaves than WT. These reverse phenotypes in the two organs mark this mutant as a promising entry to study the relationship between the organ identity and size regulation. The causal gene encodes for a Poly(A) Polymerase, an enzyme that adds poly(A) tail to pre-mRNAs. Polyadenylation is part of the 3' end formation process of pre-mRNAs. Hence, below I give introduction about the 3' end processing of eukaryotic mRNAs with focus on polyadenylation.

1.2 Pre-mRNA 3' end formation in plants with comparison to yeast and human:

1.2.1 3' end formation in general:

In the nucleus of all eukaryotic cells, all newly synthesized pre-mRNAs, with the exception of histone mRNAs, are cleaved at their 3' end and then polyadenylated (**Figure 1.5**). This process, so called 3' end formation happens co-transcriptionally (Proudfoot, 2004) and can affect all aspects of mRNA metabolism: stability, translational efficiency, export, and generation of mRNA variants. Because of its universal importance in basic gene regulations, this process has been extensively studied in yeast, animals (reviewed in (Colgan and Manley, 1997, Millevoi and Vagner, 2010, Wahle and Ruegsegger, 1999, Zhao et al., 1999)) and plants (Rothnie, 1996, Hunt, 2008). Directly involved in this process are the *cis*-elements surrounding the cleavage site of the mRNAs and the *trans*-factors (**Figure 1.6**).

1.2.2 Roles of 3' end formation:

3' end formation can affect all aspects of mRNA metabolism: stability, translational efficiency, export, and generation of mRNA variants (see above reviews).

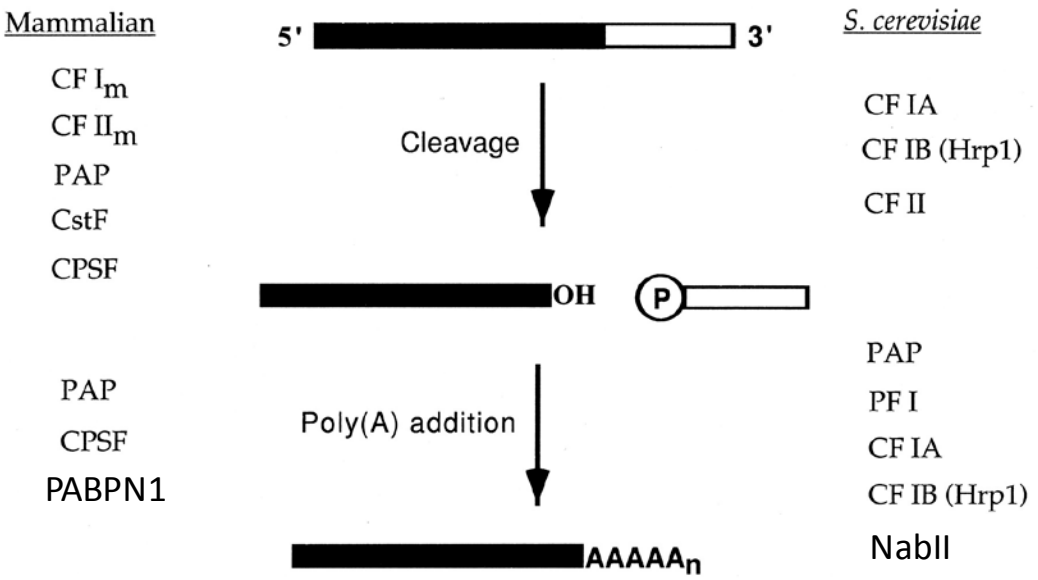


Figure 1.5. 3' end processing of eukaryotic mRNAs.

This figure is taken from (Zhao et al., 1999).

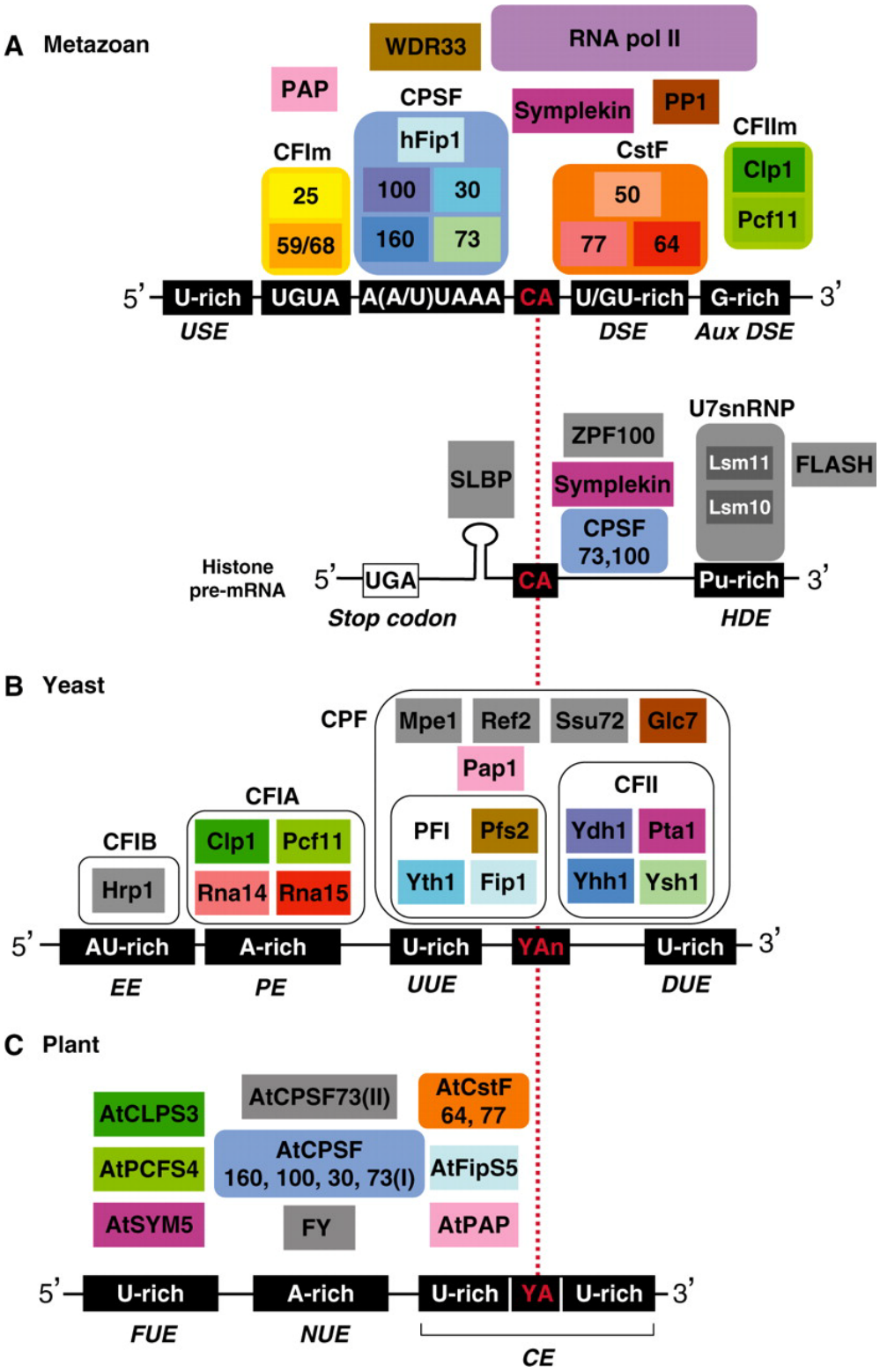


Figure 1.6. Schematic drawing of the eukaryotic 3′ end processing machineries.

This figure is taken from (Millevoi and Vagner, 2010).

1.2.1 Components of the 3’ end formation:

1.2.1.1 *cis*-elements :

There are *cis*-elements at the 3’ end of a pre-mRNA to signal where the cleavage site, which is also the polyadenylation site, is. The function of these elements is to recruit the 3’ end formation complex to the pre-mRNA and to specify the polyadenylation site. Animals, yeast and plants possess different *cis*-elements. These elements are conserved to a different degree in each kingdom.

***cis*-elements in human:**

In mammalian cells, three elements define the core polyadenylation signal: the upstream AAUAAA element, the downstream sequence elements (DSE) and the cleavage site itself (**Figure 1.6**). The highly conserved AAUAAA element is found 10 to 30 nucleotides upstream of the cleavage site.

Any single point mutation in this hexanucleotide (apart from A to U at the second position) strongly inhibits polyadenylation *in vitro* using HeLa nuclear extract and synthetic mRNA precursors (Sheets et al., 1990). The DSE is U-rich or GU-rich and is found within 30 nucleotides downstream of the cleavage site. DSE is poorly conserved, however small internal deletion (10bp) and and linker replacement mutation study on some mRNAs (simian virus 40 late (SV40) and adenovirus L3 pre-mRNAs, beta globin mRNAs) revealed that the proximity of the DSE to the poly (A) site can affect the effiency and/or the position of cleavage ((Mason et al., 1986), (MacDonald et al., 1994). The cleavage site, which is also the poly (A) addition site, is determined mainly by the distance between the USE and DSE (Chen et al., 1995) based on *in vitro* studies on SV40 pre RNA. The local sequence surrounding the cleavage site is not conserved. However, the two last nucleotides at the 3’ end of the 5’ half of cleaved product are very often 5’-CA-3’ with C found in the penultimate position of 59% of 63 vertebrate mRNAs analyzed, the value for A at the last position (denoted position +1) is 70% (Sheets et al., 1990). Converting this A to U or C shifted the terminus of the 5’ half-molecule to the adjacent adenosine downstream.

These early studies which only analyzed a small number of mRNA substrates, gave the conclusion that the AAUAAA elements is absolutely conserved. However, this view has been challenged by recent large scale studies revealing that the AAUAAA element is not so conserved; only 50%-70% of the analyzed mRNAs carry this elements in their 3’ end (Beaudoing and Gautheret, 2001, MacDonald and Redondo, 2002, Tian et al., 2005).

In addition to the three core pA signals, other sequence elements have been found in the mammalian system to enhance or repress 3’ end processing efficiency. Some enhancer/suppressor elements have been found, mostly in virus genes, although some were also found in cellular genes as summarized in (Zhao et al., 1999).

There are also examples of *cis*-elements that limit the poly(A)-tail length (discussed further in section 1.4.1.1

***cis*-elements in yeast:**

Yeast polyadenylation signals are less highly conserved compared to mammalian cells(Graber et al., 1999). However, synthetic minimal yeast 3' end signals exist (Guo and Sherman, 1996). The minimal yeast mRNA 3' end region requires at least three elements: The efficiency element, which ensures the efficiency of the positioning elements; the positioning element which positions the cleavage site; and the actual cleavage site.

***cis*-elements in plants:**

Plants polyadenylation signals are more complex and less well-conserved compared to mammalian or yeast cells. Both functional studies on a handful of *cis*-elements from some plant mRNAs (reviewed in (Rothnie, 1996)) and large scale sequence analysis of 3' end of cDNAs have been carried out (see below).

Firstly, based on functional studies some transcripts (including the Cauliflower mosaic virus (CaMV) *35S* transcript (Rothnie et al., 1994, Sanfacon and Hohn, 1990), the pea small subunit of Rubisco (*rbcS*) gene (Mogen et al., 1992), the maize 27kDa protein gene(Wu et al., 1993), the rice tungro bacilliform virus (Rothnie et al., 2001), the Agrobacterium *T-DNA* encoded octopine synthase gene (*ocs)*(MacDonald et al., 1991), and the nopaline synthase (*nos*) genes (Rothnie et al., 1994)), the general conclusion is that cis-elements in plants are not well conserved; the spacing and base composition are more important than the actual sequences. These studies agreed on a polyadenylation motif consisting of three components: i) the near upstream element (NUE) which is located 13 to 30 nt upstream of the cleavage site (CS). This NUE sequence is 6-10 nt in length, its composition is A rich and in some cases it resembles the highly conserved hexa-sequences AAUAAA in the human polyadenylation signal. ii) The far upstream element (FUE) is located at least 29 nucleotides upstream of the CS and is 60-100 nucleotides in length. The FUE sequence is rather diffuse. Apart from the general property of being U rich, no conserved motifs can be defined.iii) the cleavage site itself is commonly after 5'-YA-3' residues (Y is pyrimidine), similar to the 5'CA3' motif in mammals.

Mutational studies on those three elements of different genes gave different and sometimes contradictory conclusions about the importance/the conservatory of these cis-element sequences. For example, only deletion, but none of 18 possible base mutations in the AAUAAA motif, had a dramatic effect on 3'-end processing effiency of the CaMV/*35S* transcription unit (Rothnie et al., 1994). (Wu et al., 1993)reported a somewhat contrasting conclusion with maize 27kDa protein mRNA, where an AAUGAA motif is used as NUE. One kind of mutated NUE, AAUAAA, reduced the amount of mRNA polyadenylated at the normal pA site by half whereas other kinds, like AAUCAA or AAUUAA, decreased polyadenylation effiency to a very low level.

There is also evidence that *cis*-elements are also not conserved between monocotyledonous (monocots) and dicotyledonous (dicots) plants. When a wheat gene is expressed in tobacco plants, it is not well polyadenylated (Keith and Chua, 1986), and the poly A site was mislocated. Supporting this non-conservation in mRNA processing between monocots and dicots, (Keith and Chua, 1986) reported that an intron included in wheat *rbcS*gene was found to be inefficiently spliced in tobacco plants. The FUE cis-elements are sometimes interchangeable: CaMV FUE could partially activate an otherwise silent AAUAAA signal in nos mRNA (Rothnie et al., 1994). Another noteworthy result arising from this study is that the distance between USE and DSE is important. A 17-18 nucleotide increase in the distance is sufficient to alter the pre-mRNA poly adenylation profile (Rothnie et al., 1994). The authors suggested that the distances reflect the physical sizes of the trans-factors.

In summary, based on only 5 gene annotation units, the general conclusion is that cis-elements in plants are not well conserved; the spacing and base composition is more important than the actual sequences.

Since the release of the *Arabidopsis thaliana* genome (2000), the poly(A) signals have been studied on a large scale by aligning the sequences in the 3' end of cDNAs to the genomic sequences. Based on a data set of approximately 8000 Expressed Sequence Tags (EST) with authenticated poly(A) sites, a new model of the poly(A) signal has been compiled (**Figure 1.7**) (Loke et al., 2005). In this study, the most highly conserved motif found in the examined 3'UTRs was identical to the mammalian hexanucleotide AAUAAA. Yet, this hexamer was found in only 10% of the 8000 3'UTRs analyzed while the figure is approximately 53% for a human database of 13,000 authenticated poly(A) sites (Tian et al., 2005). The author also found some motifs in NUE and FUE (Loke et al., 2005) but with very low occurrences. For example, the most common motif in FUE is present in only less than 1% of 8000 3'UTR analyzed. To summarize, a large scale sequence alignment of 8000 3'UTRsidentified the conserved enrichment of base compositions (mainly U rich or A rich) at some locations in the 3'UTRs, yet no sequence motif was found to occur more than 10% of the transcripts analyzed. These results agree with the previous functional studies (above) and together they suggest that the cis-elements in plants are not very well-conserved compared to human and yeast.

Another noteworthy result from this paper (Loke et al., 2005),is that the authors also suggested the importance of secondary structure of 3'UTR for the choice of CS. The authors predicted the secondary structure of the 3'UTRs of *rbcS* and *CaMV*. They then mapped the published mutations in *cis*-elements from previous functional studies of *rbcS* and *CaMV* to the structures. These mutations not only change the sequence, but also disrupt the secondary structure. Interestingly, the resulting new choices of CS discovered in the previous studies (Rothnie et al., 1994, Sanfacon and Hohn, 1990, Mogen et al., 1992) are in agreement with their newly proposed hypothesis, that the secondary structure of

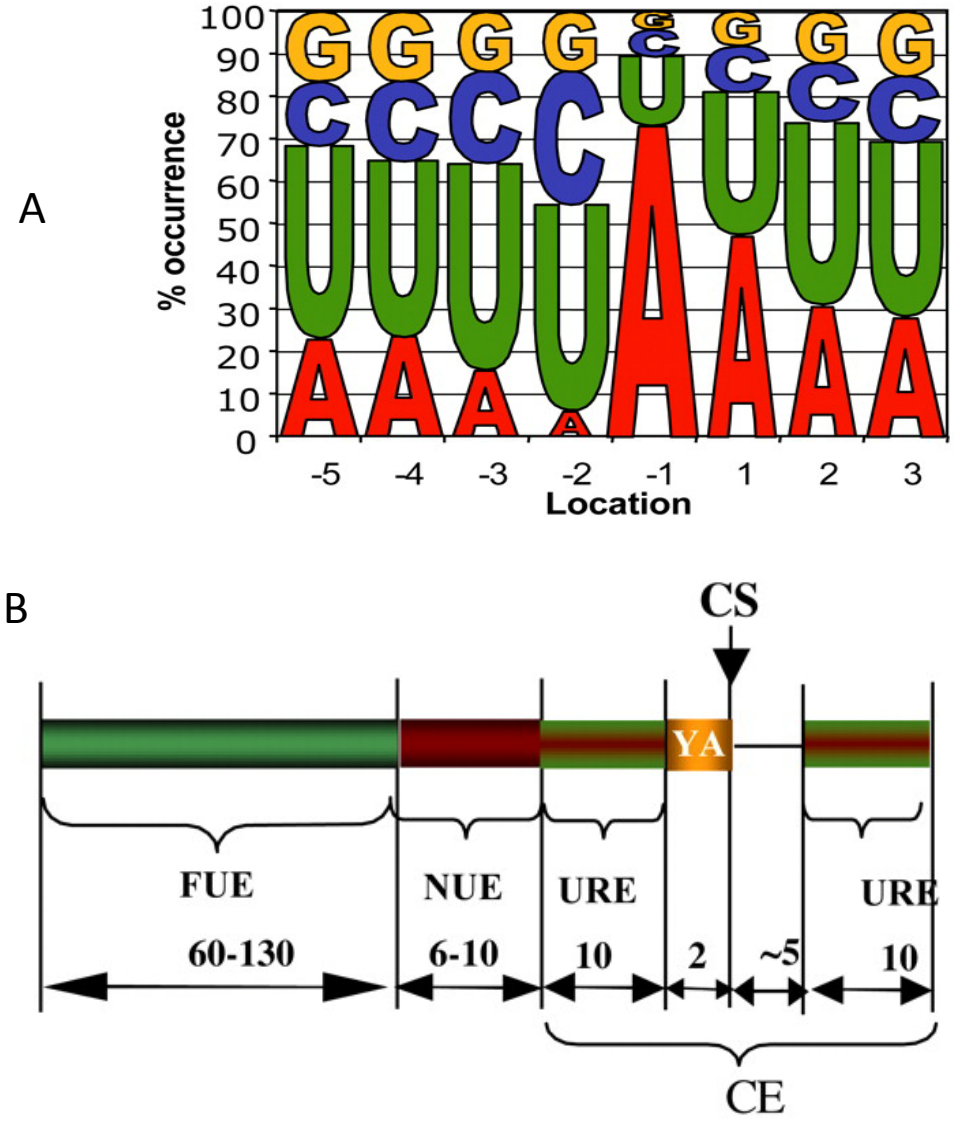


Figure 1.7. Proposed models of the *cis*-elements for 3' end processing in plants.

The figure and figure legend is modified from Loke et al., 2005.

Single-nucleotide profile of 3'-UTR and a current model of plant poly(A) signals. Distinct profiles flanking the CS are now named CEs. A, Sequence logo generated from the actual percentage of each of the four nucleotide's occurrence in the 8-K dataset, indicating preferred nucleotides flanking the CS (–5 to +3 nt). B, A current model for Arabidopsis mRNA poly(A) signals. URE, U-rich regions, which are found flanking both up and downstream of the CS.

3'UTRs can direct CS choice. However, there is currently no evidence that this secondary structure alone without the actual sequences is sufficient to determine CS.

To sum up, plant polyadenylation signals, which consist of near upstream elements (NUE), far upstream elements (FUE) and the cleavage site (CS), are much less well-conserved than mammalian polyadenylation signals. It is remains to be answer whether the less well conserved *cis*-elements have any biological functions, but it is formally possible that this diversity can allow 3' end processing of pre-mRNAs in plants to be more finely regulated, and the regulation pathway may be very different from animals and yeast.

1.2.1.2 *Trans*-factors:

The players:

So far, our understanding about *trans*-factors in 3' end complexes comes mostly from the *in vitro* biochemical studies on purified or recombinant individual factors or sub-complexes and yeast two-hybrid studies. These studies defined a polyadenylation complex as a large structure containing four main sub-complexes: the Cleavage Polyadenylation Specificity Factor complex (CPSF), the Cleavage Stimulating Factor complex (CstF), the Cleavage Factor (CF), and the Poly(A) Polymerase enzyme (PAP) (**Figure 1.6**) (Millevoi and Vagner, 2010)). Evidence from *in vitro* studies in mammalian systems indicates that proteins CPSF is the first subunit that forms the complex; CSPF160 binds AAUAAA elements, and CPSF73 has endonuclease activity (Mandel et al., 2006) and supposedly cleaves the RNA. CstF binds to the DSE and aids CPSF to bind strongly to the RNA substrates (Millevoi and Vagner, 2010). CFs bind to auxillary elements and increase the effiency of 3' end processing. A fully processive polyadenylation complex *in vitro* additionally requires poly(A) binding proteins (PABP) (Bienroth et al., 1993).

The assembly and interactions in the 3' end formation complex:

How do these individual *trans*-factors assemble? (Shi et al., 2009) used 3xMS2hairpin-tag-SV40RNA and 3xMS2hairpin–tag-adenovirus L3 pre-mRNAs as substrates to isolate the complexes. SV40 and adenovirus L3 pre-mRNAs are the two commonly used RNA substrates for *in vitro* 3' end processing studies; MS2 hairpin is an RNA sequence that forms a hairpin secondary structure and is bound specifically by coat proteins of MS2 bacteriophages. The authors isolated 85 proteins that are shared when using both substrates as bait and are not isolated using the mutated versions of substrates. The authors found the four main sub-complexes CPSF, CstF, CFI, CFII (though Clp1 was lacking), PAP, RNA polymerase II (RNA pol II) subunits, PAF (transcription enlongation complexes), components of splicesosomes, intergrators (that is involved in the 3' end processing of small RNAs), translational enlongation factors, DNA–damage repair proteins, and some novel components that had not been known to be associated with 3' end complexes before. The biological function of this interaction remains to be investigated, but

the data suggest that the interactions between 3' end process and other cellular processes are more complicated than anticipated.

***Trans*-factors in plants:**

Factors: Homologs of yeast and human *trans*-factors exist in plants (summarized in **Table 1.2**). Some of the factors have been studied biochemically; mutations have been isolated for some factors, and most of them are essential genes. However there are a handful of weak/hypomorphic mutations in some factors, for which strong mutants are lethal. These weak mutants surprisingly displayed specific phenotypes: late flowering and defects gene silencing or oxidative stress repsonse. The *cpsf100*/*esp5* mutation suppresses gene silencing (Herr et al., 2006*); cstf64-1* and *esp1* are two mutation alleles in *CstF64* gene, both of which results in late flowering and also suppress gene silencing (Liu et al., 2010), *Symplekin/esp4* is also a suppressor of gene silencing(Herr et al., 2006*);pfs2*/*fy-2* is late flowering (Simpson et al., 2003); *cstf77-1* is late flowering.

Assembly:

Using cell suspension cultures, CPSF components in Arabidopsis were identified by a proteomic approach (Zhao et al., 2009) using TAP-tagged version of CPSF100, CPSF73-I and II, CPSF30, AtFY and AtCLPS3. Note that antibodies against AtCPSF30, 73-I, -100, -160 (Xu et al., 2006); AtFY (Simpsons et al., 2003) and AtCLPS3 (Xing et al., 2008a) exists. In a different studies,(Herr et al., 2006) reported that FLAG-tagged AtCPSF100 co-purified with AtCPSF160, AtCPSF73-I, AtFY. (Delaney et al., 2006). Hunt et al., 2008 used yeast two-hybrid assays to investigate all possible pairwise interactions among trans-factors (Hunt et al., 2008), summarized in **Figure 1.8**. Using TAP-AtFY constructs, Manzano et al., 2009 pulled down FY complexes (using FY-C terminal TAP tag construct) and reported two distinct complexes at different sizes (Manzano et al., 2009). Over-expressing *FCA* promotes the formation of the high molecular weight complex. AtCPSF160 is more abundantly found in the low molecular weight complex. However, the composition of these two complexes was not analyzed by mass-spectrometry. FY interacts physically with AtCPSF100 and AtCPSF160. The authors showed that in the late flowering *fy-2* hypomorphic mutant, the complex where AtCPSF160 resides changes, but the complex where FCA resides does not change. The fact that *Arabidopsis thaliana* has several homologs for some singular *trans*-factors in human and yeast, and the results from *in vivo* pull down and yeast two hybrids experiments, suggest that *Arabidopsis thaliana* may have several distinct 3' end processing complexes.

It is concluded that plants have a more diverse *trans*-factors and *cis*-elements involved in 3' end processing and there are different 3' end processing complexes with perhaps different functions.

Table 1.2. The Arabdiopsis homologs/orthorlogs of yeast/human core 3' end processing factors:

The blank cells in the table indicates the information is not available.

Factors	Human/yeast counterpart	At code	Overexpression	Mutants phenotype	Available material	Enzymatic activity	Reference
CPSF160	CPSF160/Yhh1	At5g51660	NA	NA	Peptide antibody		(Xu et al., 2006)
CPSF100	CPSF100/Ydh1	At5g23880		+ some alleles are lethal + *esp5* enhances silencing and affects flowering time.	Peptide antibody		(Elliott et al., 2003, Tzafrir et al., 2004, Herr et al., 2006, Xu et al., 2006)
CPSF73	CPSF73/Ysh1	At1g61010 (CPSF73-I)	+ 35S::cDNA all transformations failed to give transformants +pCPSF73::cDNA transgenic plants are male sterile and have big flowers	+ Plants carrying inducible RNAi constructs are dead after induction,	Peptide antibody		(Xu et al., 2006)
		At2g01730 (CPSF73-II)		Embryonic lethal,	Peptide antibody		(Xu et al., 2006, Xu et al., 2004)

Table 1.2. The Arabdiopsis homologs of yeast/human core 3' end processing factors (continued):

Factors	Human/yeast counterpart	At code	Overexpression	Mutants phenotype	Available material	Enzymatic activity	Reference
CPSF30	CPSF30/Yth1	At1g30460		*oxt6* mutants show enhanced sensitive to oxidative stress (a *T-DNA* allele)		Endonuclease, interacts with calmodulin	(Addepalli and Hunt, 2007, Xu et al., 2006, Delaney et al., 2006)
FY	hPfs2/Pfs2	At5g13480		*fy-1* and *fy-2* are late flowering.			(Simpson et al., 2003)
CstF77	CstF77/Rna14	At1g17760		hypomorphic *cstf77-1* is late flowering; cstf77-2 causes female gametophytic lethality.		Bind RNA	(Yao et al., 2002, Bell and Hunt, 2010, Liu et al., 2010)
CstF64	CstF64/Rna15	At1g71800		*cstf64-1* reduced fertility, flowering time delayed; *cstf64-2* small, sterile; *esp1* enhances gene silencing			(Liu et al., 2010)

Table 1.2. The Arabdiopsis homologs of yeast/human core 3' end processing factors (continued):

Factors	Human/yeast counterpart	At code	Overexpression	Mutants phenotype	Available material	Enzymatic activity	Reference
CstF50	CstF50	At5g60940		?			No study
Fip1	hFip1/Fip1	At3g66652(FIPS3)					
		At5g58040(FIPS5)		a *T-DNA* insertion allele is lethal			(Forbes et al., 2006)
CFIm25	hCFIm25	At4g25550 (CFIS2)					No study
		At4g29820 (CFIS1)					No study
CFIIm	hClp1/Clp1	At3g04680 (CLPS3)	CLPS3::TAP Small deformed flowers, other morphological changes,	a *T-DNA* insertion allele is lethal	CLPS3-TAP	RNA kinase	(Xing et al., 2008a)
		At5g39930 (CLPS5)		*T-DNA* no phenotype			(Xing et al., 2008a)
Symplekin	Symplekin/Pta1	At1g27590 (SYM1)					
		At1g27595 (SYM2)					
		At5g01400 (SYM5)		*esp4* enhances gene silencing			(Herr et al., 2006)

Table 1.2. The Arabdiopsis homologs of yeast/human core 3' end processing factors (continued):

Factors	Human/yeast counterpart	At code	Overexpression	Mutants phenotype	Available material	Enzymatic activity	Reference
CFIIm	hPcf11p/Pcf11	At1g66500 (PCFS1)					(Xing et al., 2008b)
		At4g04885 (PCFS4)		*pcfs4* is late flowering			
		At5g43620 (PCFS5)					
Cytoplasmic Poly(A) Binding protein (PAB)	Pab1(yeast)	At1g34140 (PAB1)					(Belostotsky, 2003, Belostotsky and Meagher, 1993, Bravo et al., 2005, Chekanova and Belostotsky, 2003, Belostotsky and Meagher, 1996, Chekanova et al., 2001, Palanivelu et al., 2000)
		At4g34110 (PAB2)*					
		At1g22760 (PAB3)*					
		At2g23350 (PAB4)					
		At1g71770 (PAB5)*					
		At3g16380 (PAB6)					
		At2g36660 (PAB7)					
		At1g49760 (PAB8)					

Table 1.2. The Arabdiopsis homologs of yeast/human core 3' end processing factors (continued):

Factors	Human/yeast counterpart	At code	Overexpression	Mutants phenotype	Available material	Enzymatic activity	Reference
Nuclear Poly(A) Binding protein (PAB)	Nab2/PABPN1	At5g65260 (PABN1)					No study
		At5g51120 (PABN2)					
		At5g10350 (PABN3)					

* The PAB that can rescue *pab1* yeast mutant.

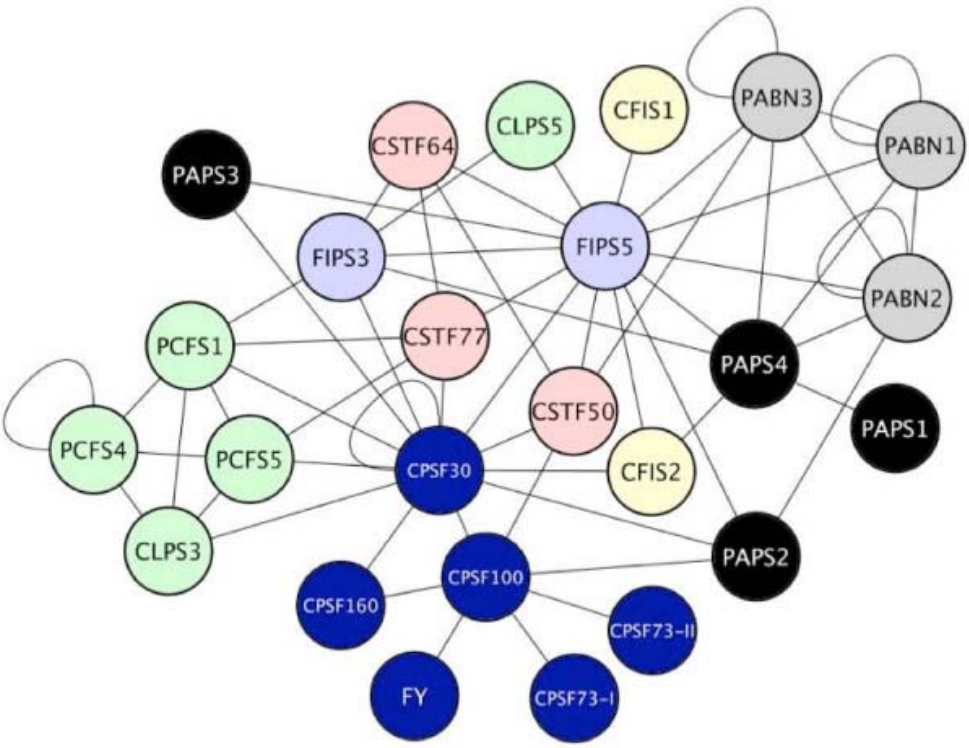


Figure 1.8. Interactions between 3' end processing factors identified by yeast two hybrid.

This figure is taken from (Hunt, 2008).

1.3 Poly (A) polymerase:

1.3.1 General properties of canonical PAPs in yeast and animals:

A PAP is classified as canonical PAP (cPAP) if it shares a high amino acid sequence identity 30-40% with, and are the most similar proteins in that organism genome to, the human PAP alpha or the yeast PAP1. cPAPs have a steady-state localization in the nucleus. cPAPs add adenosine to the 3' end of cleaved mRNAs to form poly(A) tails of mRNAs.

The number of cPAPs varies with species: human has three, *Drosophila melanogaster*(fruit fly) has one, *Saccharomyces cerevisiae* (budding yeast) has one, chicken has one, mammals has three, *Caenorhabditis elegans*(worm) has three, *Daniorerio* (fish) has two, *Amoeba* has one. Plants have more: *Physcomitrella patens* (moss) has two, *Chlamydomonas* (algae) has one, *Selaginella moellendorffii* has two, *Oryza sativa* (rice) has six, *Sorghum* has six, *Populus* has three, grapevine has four, and *Arabidopsis thaliana* has four. Canonical PAPs that have been studied includes bovine, human, mouse, yeast, fly, and *Arabidopsis*. In species whose genome has not been fully sequenced, there could be more canonical PAPs. The phylogenetic tree of 45 cPAPs, the alignment file of amino acid sequence and the % ID table are attached as an electronic version. A canonical PAP has two domains, the N-terminal catalytic domain and the C-terminal domain. A phylogenetic trees of cPAPs was published Meeks et al., 2009 (**Figure 1.9**). The N-terminal catalytic domain is highly conserved in eukaryotes but the C-terminal domain is not conserved between plants and animals. Yeast PAP does not have a C-terminal domain.

In vitro biochemical studies on mammalian cPAPs revealed that cPAP on its own does not have any preference to RNA substrates *in vitro* (Bienroth et al., 1993). The cPAP itself is a distributive enzyme *in vitro*, that is it only adds one or very few adenosines at one binding event (the K_m is quite high). Also on its own, *in vitro*, the cPAP does not give a specific poly(A)-tail length, i.e. the tail length in the final product depends on the amount of the enzyme and the time of the reaction (Wahle, 1995). With the help of the other components of the 3' end processing complex including CPSF and PABP, cPAP becomes a processive enzyme, and this three-component complex also regulates poly(A) tail length, such that inone binding event of the fully assembled complex to a pre-mRNA can lead to the formation of the final product of a default 200 to 250 A's (in mammalian systems) (Wahle, 1995).

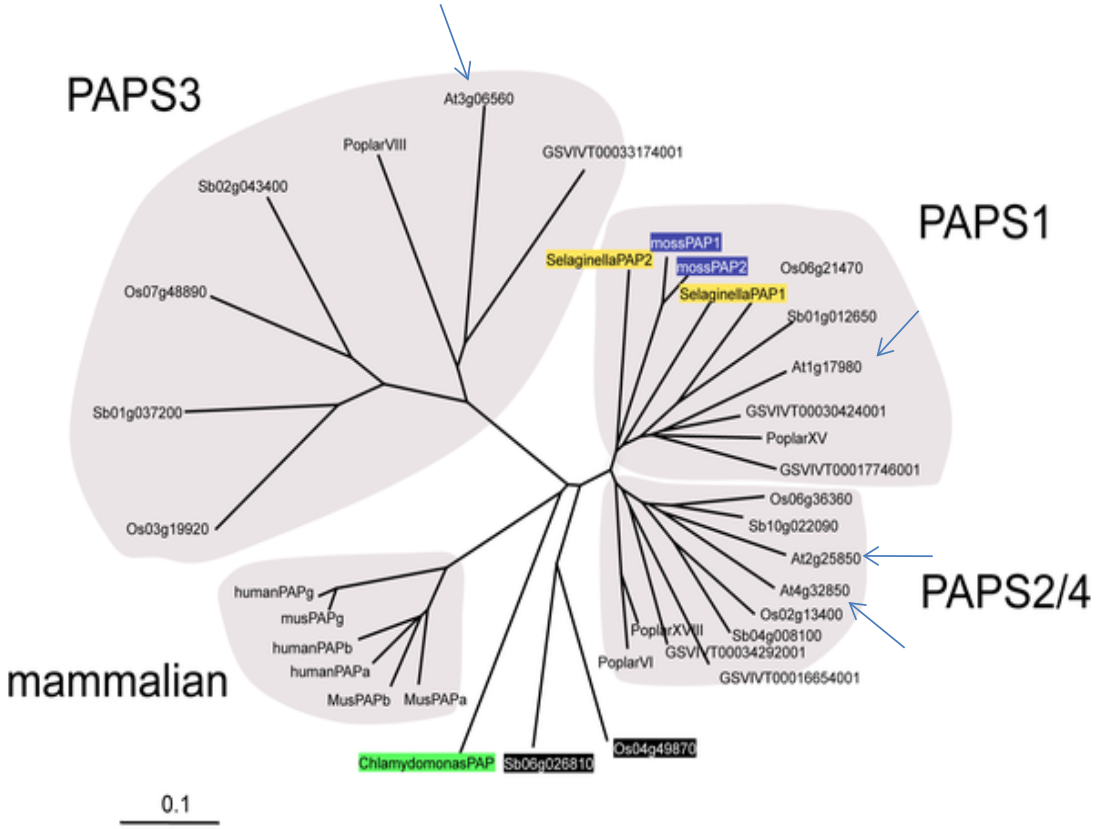


Figure 1.9. Phylogenetic tree of plant and mammalian cPAP proteins based on the proteins' enzymatic core.

Arrows indicate position of *Arabidopsis* proteins. This figure is from (Meeks et al., 2009)

1.3.2 Detailed properties of canonical PAPs in yeast and animals:

1.3.2.1 X-ray crystal structure:

Several structures of PAP have been solved, yet for only the N-terminal domain and not for the C-terminal domain. The Bohm group published the structure of the yeast PAP with 3' dATP, 3' deoxy adenosine ribonucleotide (also called cordycepin, which is an analog of ATP, and it is also a chain terminator) (Bard et al., 2000). In 2007, a structure of yeast PAP was published also by this group with the other three components that are required and sufficient for *in vitro* non-specific polyadenylation: RNA substrates, ATP and Mg^{2+} (Balbo and Bohm, 2007). Additionally, using other structures and steady state fluorescence studies (Balbo et al., 2007) revealed the protein domain movements during catalytic cycles. The structure of bovine PAP was published in 2000 by a different group (Martin et al., 2000) showing similar structure to yeast PAP. (Meinke et al., 2008) revealed the structure of yeast PAP in the complex with a peptide of Fip1.

These structures shed light on the molecular mechanism of how PAP acts, providing an explanation of specificity towards ATP but not other nucleotides and the reason why PAP is distributive Balbo et al., 2007. The N-terminal and C-terminal domains of the yeast PAP move independently as rigid bodies along two well defined axes of rotation during the catalytic cycle. This C-terminal domain of yeast PAP is not to be mistaken with the C-terminal domain of bovine canonical PAP, as yeast PAP lacks the similar domain to the bovine C-terminal domain; the whole yeast PAP, which is equivalent to the N-terminal catalytic domain of bovine PAP, can be subdivided into three parts the N-terminal, middle and C-terminal domains). (Balbo et al., 2007) proposed a model of the catalytic cycle of PAP (**Figure 1.10**).

In this model, a cycle begins with an open conformation allowing sampling of nucleotide substrates; only a correct nucleotide (ATP) can induce the closed enzyme conformation, thereby enabling efficient catalysis. After catalysis, i.e. addition of one AMP, the enzyme opens again to release the product which then must bind again for the next round of adenylyl transfer. The fact that the product (the poly(A) tail) is released after every cycle explains why the enzyme on its own is distributive.

1.3.2.2 Biochemical properties as revealed by site-directed mutagenesis:

By site-directed mutagenesis, many single amino acid (aa) substitutions have been made to the PAP protein to study the contribution of these aa to the kinetic of the enzyme (K_m, substrate specificity, V_{max}). (Martin et al., 1999, Balbo and Bohm, 2007, Martin et al., 2004, Zhelkovsky et al., 2004). These studies identified the three critical catalytic Asp residues (D113A, D115A, D157A in bovine PAP), mutations in any of which inactivate the enzyme. Also, residues affecting the substrate binding of: pre-mRNA, ATP Mg, were identified.

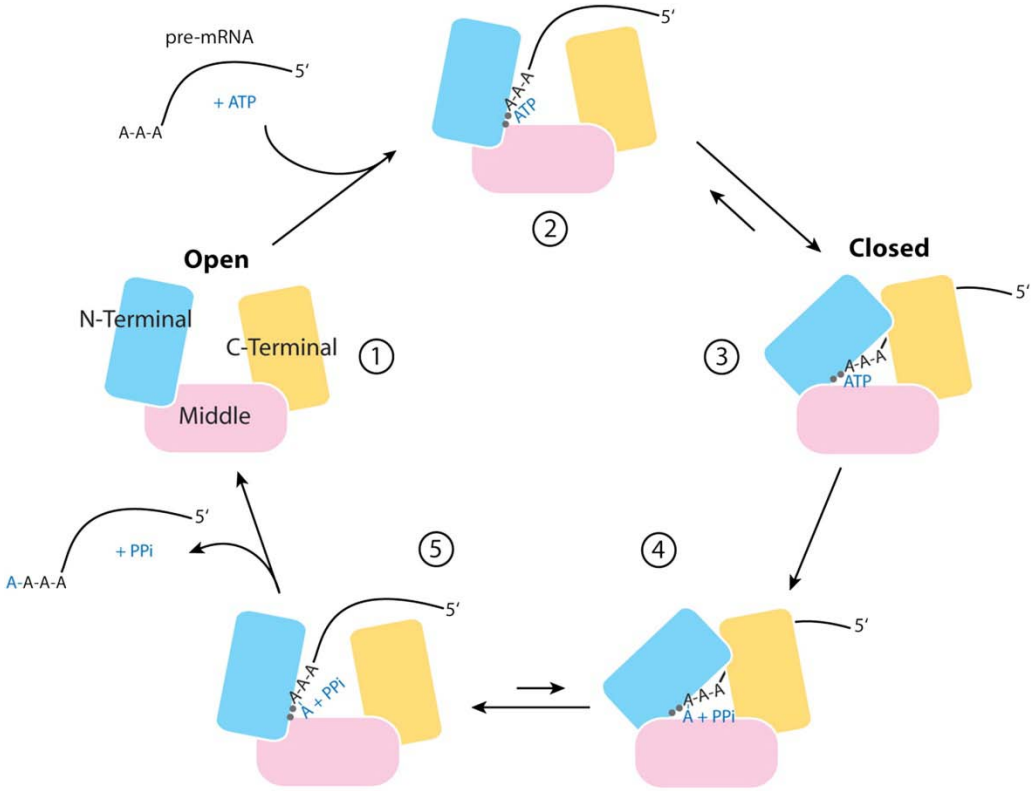


Figure 1.10. A Model for the catalytic cycle of a poly(A) polymerase.

The free enzyme exists in an open form. Upon binding both substrates (in random order), the enzyme undergoes a conformational change to the closed form, where catalysis occurs. After catalysis, the enzyme returns to the open form to release the products. The two metal ions (Mg^{2+} or Mn^{2+}) in the active site are shown as gray spheres.

The figure and figure legend are taken from (Mandel and Tong, 2007).

Finally there are residues that when changed can drive the specificity of the enzyme towards GTP/CTP rather than ATP.

1.3.2.3 Known physical and genetic interactions

A mutant that can partially suppress yeast *pap1-1* defect is the RNA polymerase III *pds1* mutant (Briggs and Butler, 1996). Other mutants in the RNA processing machinery also showed genetic interaction with *pap1-1*, e.g. the exosome mutants *rrp6* (van Hoof et al., 2000). Synthetic lethal mutants with *pap1-1* at permissive temperature are *lcp5* mutants, affecting a protein that is involved in maturation of 18S rRNA (Wiederkehr et al., 1998).

1.3.2.4 Post-translational modification at the C-terminal domain (CTD):

The mammalian cPAP protein can be inactivated by phosphorylation in the C-terminal domain during the M-phase of the cell cycle (Colgan et al., 1998). Mammalian Poly(A) polymerase alpha (PAPOLA) is also subject to other protein modifications at its C-terminal domain such as sumoylation (Vethantham et al., 2008) (**Figure 1.11**). PAPOLA protein was also shown to be acetylated at the nuclear localization signal (Shimazu et al., 2007). This modification promotes the cytoplasmic localization of PAPOLA and reduces the interaction between PAPOLA and CF.Im.

1.3.3 The genes encoding cPAPs and their mutant phenotypes in yeast and mammalians cPAPs:

1.3.3.1 Human cPAPs:

In human, three canonical PAPs that are encoded by three different genes have been discovered. They are named: PAPOLA (also called PAP alpha or hsPAPII), PAPOLB (also called PAP beta or testis PAP) and PAPOLG (also called PAP gama or neo-PAP).

PAPOLA (Swissprot P51003) and PAPOLG (Swissprot Q9BWT3)

Identification of PAPOLA and PAPOLG were identified in the early experiment by using antibodies and Western blot. In HeLa cell nuclear extracts, a monoclonal antibody NN2 (raised against a purified recombinant PAP protein from E.coli expressing a human cDNA of PAP (Thuresson et al., 1994)) detected three proteins of 90kDa, 100 kDa and 106kDa. There is a different antibody, anti-PAPIIex22 (Kyriakopoulou et al., 2001) that only recognizes only two larger forms.

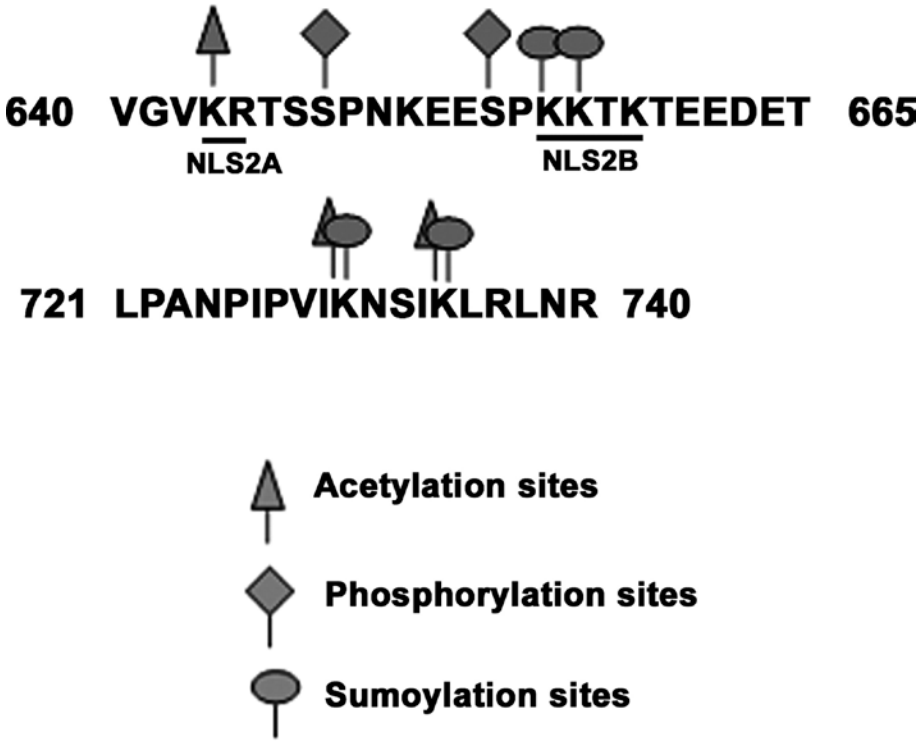


Figure 1.11. Modification of mamalian PAP C-terminal domain.

Figure is taken from (Vethantham et al., 2008)

PAPOLA The 100kDa and 106kDa forms are proteins encoded by human PAPOLA (or also called humanPAPII) (Thuresson et al., 1994, Kyriakopoulou et al., 2001). The protein PAPOLA (Swissprot P51003)has 745 amino acids. The 106kDa is most likely to be the phosphorylated form of PAPOLA (Thuresson et al., 1994). Both the 100kDa and 106kDa proteins are present in both nuclear and cytoplasmic fractions of HeLa cells. The gene encoding PAPOLA in mouse has been demonstrated to produce five different alternative spliced/alternative polyadenylated transcripts falling into two classes, the long forms and the short forms, which lack the C-terminal half of the gene (Zhao and Manley, 1996). Only the long transcripts seem to be translated as deducted from the western blot experiment using yet another antibody that is raised against purified bovine PAP I. The relative abundance of these spliced forms is tissue specific (Hela cells, mouse brain tissues, intestine tissues and liver tissues were examined). One of the long forms, PAPII, is expressed most strongly amongst PAPOLA transcript isoforms and is found in all tissues tested. In *in vitro* assay using purified recombinant proteins, the short forms were found to be inactive. It seems that the spliced form II (PAPII) of PAPOLA is the main PAP in human.

PAPOLG The 90kDa protein is the product of a different gene, called PAPOLG or also neo-PAP (Kyriakopoulou et al., 2001, Topalian et al., 2001).PAPOLG shared 76% identity in amino acid sequence with PAPOLA. PAPOLG seemed to be exclusively localized in the nucleus as demonstrated by transfection experiment (Topalian et al., 2001) and by western blot comparing the nuclear vs. cytoplasmic fractionations of HeLa cells (Thuresson et al., 1994).The protein is 736 amino acid long, has a calculated molecular weight of 82.8kDa. PAPOLG is an authentic canonical PAP (Topalian et al., 2001). PAPOLG has non-specific polyadenylation activitiy *in vitro* (when assayed on pre-cleaved synthetic RNA substrates and without CPSF addition), and this activity is comparable to bovine PAPI. PAPOLG also has specific polyadenylation and cleavage activity *in vitro* (assayed with pre-cleaved synthetic RNA substrates with added purified CPSF). In contrast to PAPOLA, PAPOLG seems not to be phosphorylated, as it is not sensitive to phosphatase treatment. The expression of PAPOLG varies in different tissues.

PAPOLB or testis PAP, is specifically expressed in mouse testis (Kashiwabara et al., 2002)..PAPOLB is shared 86% identity in amino acid sequence with PAPOLA. It is the shortest of the three cPAPs with a molecular weight of 70kDa. PAPOLB only contains 636 aa, it lacks the 100 aa at the C-terminal domain. The other unique properties of PAPOLB compared to the PAPOLA and PAPOLG is that it is cytoplasmic. Mice deficient in PAPOLB are male sterile, but otherwise have normal growth and behaviour. Female mice deficient for PAPOLB are fertile. Specific substrates of PAPOLB are RNAs from that are the haploid-specific genes that are important for the morphogenesis of the sperm cells. Poly(A)-tail tests showed that disruption of PAPOLB shortens the poly(A) tails of these mRNAs. Interestingly, poly(A)-tail shortening did not destabilize these mRNAs or repress translation of many of these genes as judged from Western blot results. The only one transcript that the authors

found to have shortened poly(A) tails and reduced protein levels in the nuclear, but not the cytoplasmic fraction in *tpap* mutant compared to WT is the TAF10 transcript. The author argued that as TAF10 is a component of the TFIID complex, which is dispensable for general RNA polymerase II–mediated transcription yet essential for selective expression of specific genes, disruption of TAF10 protein level in the nucleus is likely to contribute in part to the mutant phenotype.

1.3.3.2 Yeast cPAP and fly cPAP:

Yeast has only one canonical PAP, PAP1.Yeast PAP1 is nuclear. Yeast PAP1 protein is lacks C-terminal domain. *pap1-1* yeast mutant is temperature sensitive, and growth is inhibited at high temperatures(Patel and Butler, 1992).

Drosophila has only one canonical PAP, it is 659 amino acids long, has both N and C terminal domains, and is 56% identical in amino acid sequence to bovine PAP. The null mutant is lethal. A weak mutant allele, *hiiragi,* causes a very specific phenotype of defects in morphology of the wing margins (Murata et al., 2001). The molecular mechanism of this defect is unknown. Interestingly, HIIRAGI protein resides in both the nucleus and the cytoplasm and is essential for the polyadenylation of certain mRNAs in both compartments (Juge et al., 2002). Overexpression of hiiragi protein in female germ line causes hyperadenylation of mRNAs that are (*bicoid* mRNAs) and are not (*sop* mRNAs) normally regulated by cytoplasmic polyadenylation, and causes embryonic lethality. Yet, overexpressing *HIIRAGI* in somatic tissues did not cause any changes to the poly(A) tail length of the ubiquitously expressed *sop* mRNAs. The authors suggested that overexpressing HIIRAGI protein in germlines causes increased levels of cytoplasmic protein hence can cause non-specific hyperadenylation via increasing the interaction between PAP and mRNAs in the cytoplasm.

1.3.4 Plant canonical poly(A) polymerases:

1.3.4.1 Phylogenetics of cPAPs across kingdoms:

(Meeks et al., 2009) shows the phylogenetics of PAPs across kingdoms (**Figure 1.9**) (a more complete comparision with more species is attached as an electronic version). Plant species other than *Arabidopsis* possess several canonical PAPs, functional studies of which have not yet been reported.

1.3.4.2 *Arabidopsis thaliana* cPAPs:

The studies on canonical and non canonical PAPs in Arabidopsis are (Addepalli et al., 2004, Lange et al., 2009, Meeks et al., 2009) and reference therein.

Identity. Arabidopsis has four genes that encode highly similar proteins to human/yeast canonical PAPs (The % of identity and the alignment of the protein sequences are included

as electronic version). They are on four different chromosomes hence are named *PAPS1* (At1g17980), *PAPS2* (At2g25850), PAPS3 (At3g06560) and PAPS4 (At4g32850).

Domain structure. The exon and intron patterns of the four genes are similar (**Figure 1.12A**) (Addepalli et al., 2004). Amongst the four, PAPS3 is the most divergent because it encodes a smaller protein of 506 amino acid (aa). PAPS3 protein lacks the C-terminal domain and lacks the nuclear localization signal. This property of PAPS3 resembles human PAPOLB(Kashiwabara et al., 2002). The other three proteins PAPS1, PAPS2 and PAPS4 are larger in size (713, 800 and 738 amino acids, respectively) and contain both the highly conserved N-terminal domain and the variable C-terminal domain. The C-terminal domains of PAPS1, PAPS2 and PAPS4 do not align to each other and neither do they align to any known proteins either within or outside of the plant kingdom. The C-terminal domain of human PAP has been proposed to be the regulatory domain of PAP, and is subjected to extensive post translational modification (Colgan et al., 1998, Vethantham et al., 2008), yet whether this is also true for Arabidopsis cPAPS is not known.

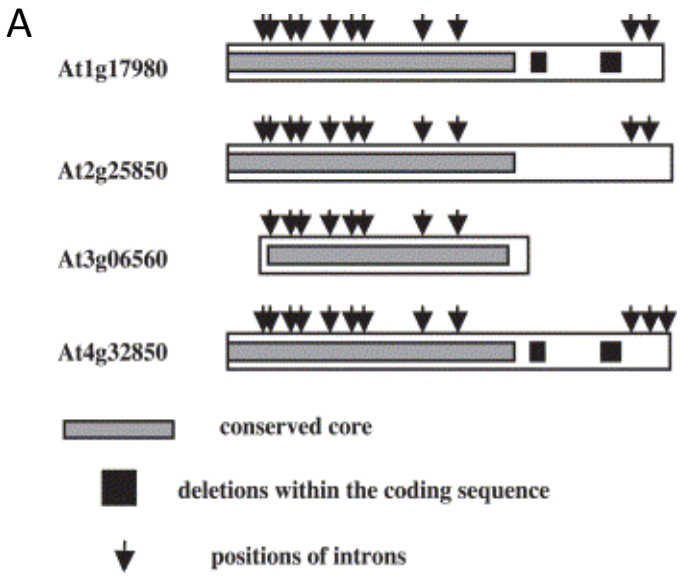
Mutants. Meeks *et al.,* 2009 reported the failure to obtain homozygous single mutants for all four cPAPS from the heterozygous plants carrying *T-DNA*s inserted in each locus. The authors suggested that a disruption to any single cPAPS causes lethality and hence every cPAPS is essential.

In vitro activity. All four recombinant cPAPS proteins have been shown to possess non-specific polyadenylation *in vitro* (Hunt et al., 2000, Addepalli et al., 2004). The activity of three proteins *in vitro* are not equal with PAPS1 = PAPS3 > PAPS4 (Addepalli et al., 2004). However, the differences in activity among them are not large, less than two fold. The activity of PAPS2 has been demonstrated in a different study (Hunt et al., 2000) and therefore is not suitable for comparisons with those of PAPS1, PAPS3 and PAPS4.

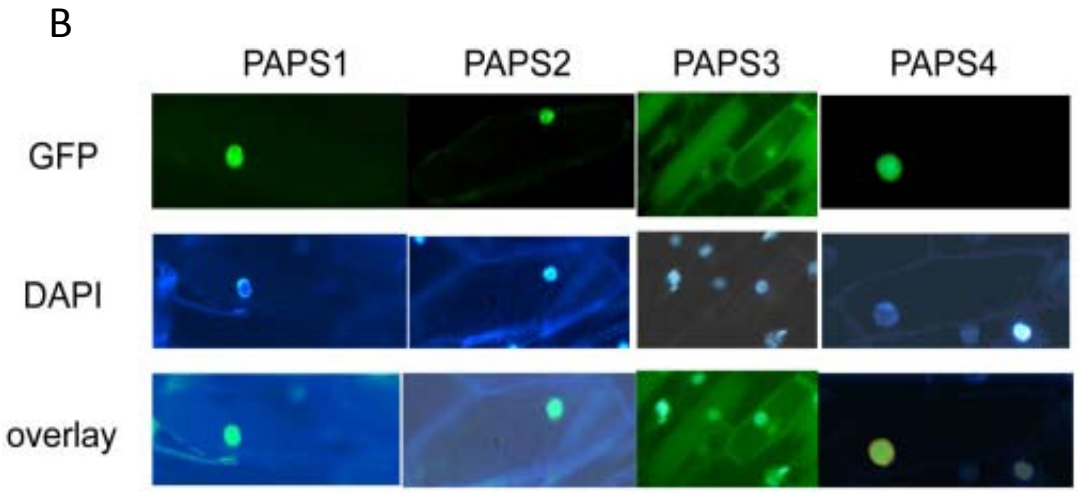
Protein localization. By bombardment of onion cells with construct over-expressing a translational fusion of AtPAPS and GFP (C-terminal fusion: *35S::PAPS-CGFP*), (Meeks et al., 2009) showed that PAPS1, PAPS2 and PAPS4 are exclusively localized in the nuclei while PAPS3 is exclusively localized in the cytoplasm (**Figure 1.12B**).

Expression domain. Using promoter: GUS constructs (Meeks et al., 2009) and Northern blots (Addepalli et al., 2004), three cPAPSs were shown to be widely expressed in largely overlapping domains (**Figure 1.13**). PAPS3 is highly expressed in pollen. Again this resembles PAPOLB in mammals (Kashiwabara et al., 2002).

Alternative splicing pattern. It has been known in animals that PAPOLA produce at least six different spliced forms (Zhao and Manley, 1996). (Addepalli et al., 2004) reported that Arabidopsis cPAPSs also share this property with human PAPOLA. For PAPS1, there are two spliced forms published (Addepalli et al., 2004). Spliced form ‘b’ is caused by an



Addepalli et al., 2004



Meeks et al., 2009

Figure 1.12. Gene structure and intracellular localization of PAPSs in *Arabidopsis.*

Figures were modified from the reference indicated underneath the figures.

A. The organization of the four cPAP genomic loci.

B. The cellular localization of cPAPs. Onion cells were bombarded with constructs encoding GFP fused to the C-terminal domains of cPAPs.

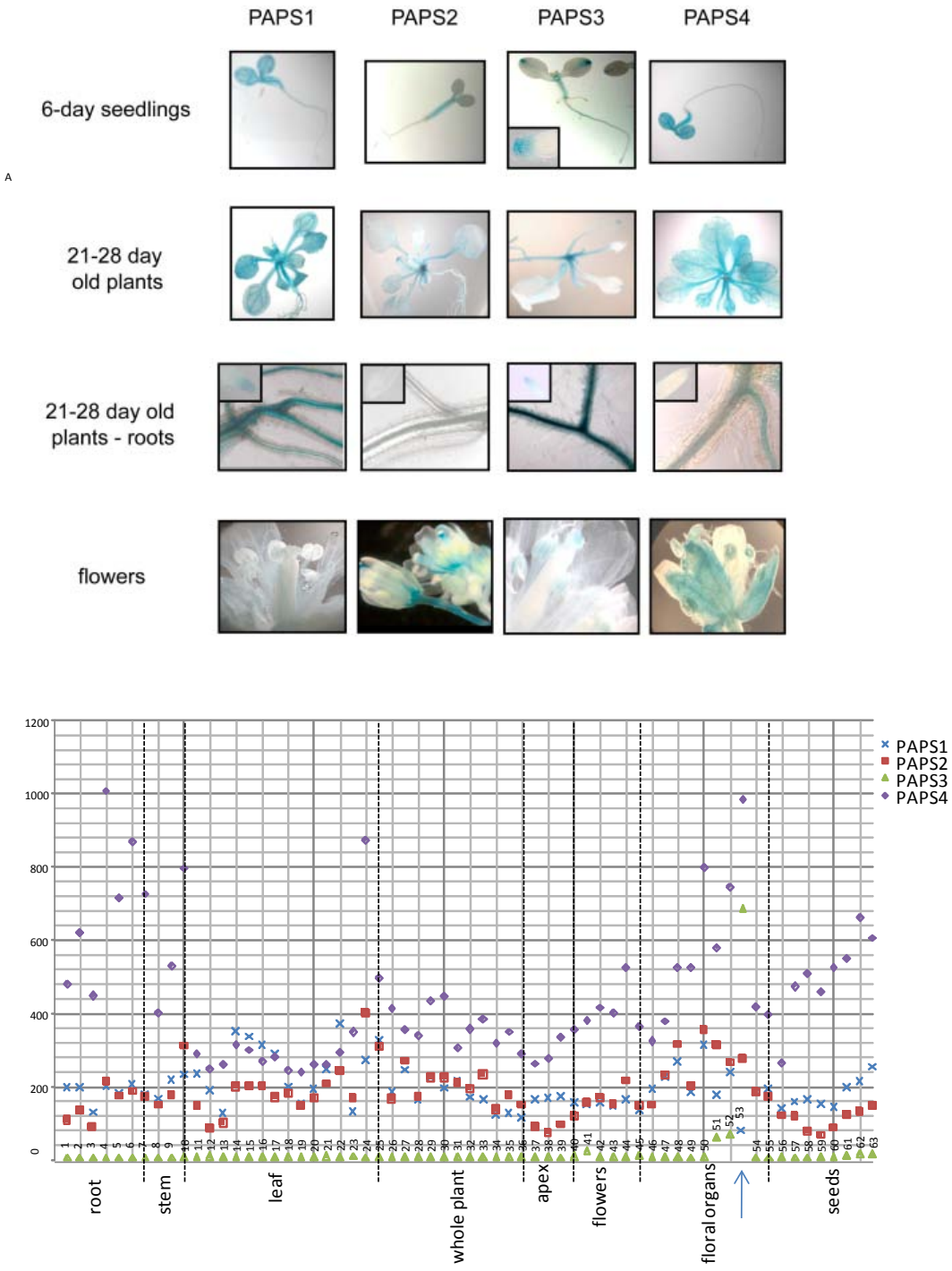


Figure 1.13. Expresssion patterns of *PAPSs*

A. Expression domains visualized by pPAPS::GUS constructs. From (Meeks et al., 2009)

B. Absolute expression levels of *PAPS1, PAPS2, PAPS3* and *PAPS4* in different *Arabidopsis* tissues. Data are from AtGenExpress . (http://jsp.weigelworld.org/expviz/expviz.jsp). Arrow indicates stamen and pollen samples (Sample 56). The 63 sample types are as follow:

Sample	Tissue cluster	Tissue	Age	AtGenexpress sample ID
1	root	roots	7 days	ATGE_3

2	root	roots	17 days	ATGE_9
3	root	root	15 days	ATGE_93
4	root	root	8 days	ATGE_94
5	root	root	8 days	ATGE_95
6	root	root	21 days	ATGE_98
7	root	root	21 days	ATGE_99
8	stem	hypocotyl	7 days	ATGE_2
9	stem	1st node	21+ days	ATGE_28
10	stem	stem, 2nd internode	21+ days	ATGE_27
11	leaf	cotyledons	7 days	ATGE_1
12	leaf	leaves 1 + 2	7 days	ATGE_5
13	leaf	rosette leaf #4, 1 cm long	10 days	ATGE_10
14	leaf	rosette leaf # 2	17 days	ATGE_12
15	leaf	rosette leaf # 4	17 days	ATGE_13
16	leaf	rosette leaf # 6	17 days	ATGE_14
17	leaf	rosette leaf # 8	17 days	ATGE_15
18	leaf	rosette leaf # 10	17 days	ATGE_16
19	leaf	rosette leaf # 12	17 days	ATGE_17
20	leaf	leaf 7, petiole	17 days	ATGE_19
21	leaf	leaf 7, proximal half	17 days	ATGE_20
22	leaf	leaf 7, distal half	17 days	ATGE_21
23	leaf	leaf	15 days	ATGE_91
24	leaf	senescing leaves	35 days	ATGE_25
25	leaf	cauline leaves	21+ days	ATGE_26
26	whole plant	seedling, green parts	7 days	ATGE_7
27	whole plant	seedling, green parts	8 days	ATGE_96
28	whole plant	seedling, green parts	8 days	ATGE_97
29	whole plant	seedling, green parts	21 days	ATGE_100
30	whole plant	seedling, green parts	21 days	ATGE_101
31	whole plant	developmental drift, entire rosette after transition to flowering, but before bolting	21 days	ATGE_22
32	whole plant	as above	22 days	ATGE_23
33	whole plant	as above	23 days	ATGE_24
34	whole plant	vegetative rosette	7 days	ATGE_87
35	whole plant	vegetative rosette	14 days	ATGE_89
36	whole plant	vegetative rosette	21 days	ATGE_90
37	apex	shoot apex, vegetative + young leaves	7 days	ATGE_4
38	apex	shoot apex, vegetative	7 days	ATGE_6
39	apex	shoot apex, transition (before bolting)	14 days	ATGE_8
40	apex	shoot apex, inflorescence (after bolting)	21 days	ATGE_29
41	flowers	flowers stage 9	21+ days	ATGE_31
42	flowers	flowers stage 10/11	21+ days	ATGE_32
43	flowers	flowers stage 12	21+ days	ATGE_33
44	flowers	flowers stage 15	21+ days	ATGE_39
45	flowers	flower	28 days	ATGE_92
46	floral organs	flowers stage 15, pedicels	21+ days	ATGE_40
47	floral organs	flowers stage 12, sepals	21+ days	ATGE_34

48	floral organs	flowers stage 15, sepals	21+ days	ATGE_41
49	floral organs	flowers stage 12, petals	21+ days	ATGE_35
50	floral organs	flowers stage 15, petals	21+ days	ATGE_42
51	floral organs	flowers stage 12, stamens	21+ days	ATGE_36
52	floral organs	flowers stage 15, stamen	21+ days	ATGE_43
53	floral organs	mature pollen	6 wk	ATGE_73
54	floral organs	flowers stage 12, carpels	21+ days	ATGE_37
55	floral organs	flowers stage 15, carpels	21+ days	ATGE_45
56	seeds	siliques, w/ seeds stage 3; mid globular to early heart embryos	8 wk	ATGE_76
57	seeds	siliques, w/ seeds stage 4; early to late heart embryos	8 wk	ATGE_77
58	seeds	siliques, w/ seeds stage 5; late heart to mid torpedo embryos	8 wk	ATGE_78
59	seeds	seeds, stage 6, w/o siliques; mid to late torpedo embryos	8 wk	ATGE_79
60	seeds	seeds, stage 7, w/o siliques; late torpedo to early walking-stick embryos	8 wk	ATGE_81
61	seeds	seeds, stage 8, w/o siliques; walking-stick to early curled cotyledons embryos	8 wk	ATGE_82
62	seeds	seeds, stage 9, w/o siliques; curled cotyledons to early green cotyledons embryos	8 wk	ATGE_83
63	seeds	seeds, stage 10, w/o siliques; green cotyledons embryos	8 wk	ATGE_84

inclusion of intron 6 which is predicted to result in truncated non-functional protein (**Figure 1.14**). In flowers, this spliced form 'b' is predominantly detected. The longer spliced form 'a' should produce a functional protein. This spliced form is predominantly detected in leaves. Based on the prediction that 'b' produces a non-functional protein, flowers may not require as much catalytically functional PAPS1 (protein produced from spliced form 'a') compared to PAPS2/PAPS4 proteins. *PAPS2, PAPS4* and *PAPS3* genes also produce alternative spliced forms (**Figure 1.14**).

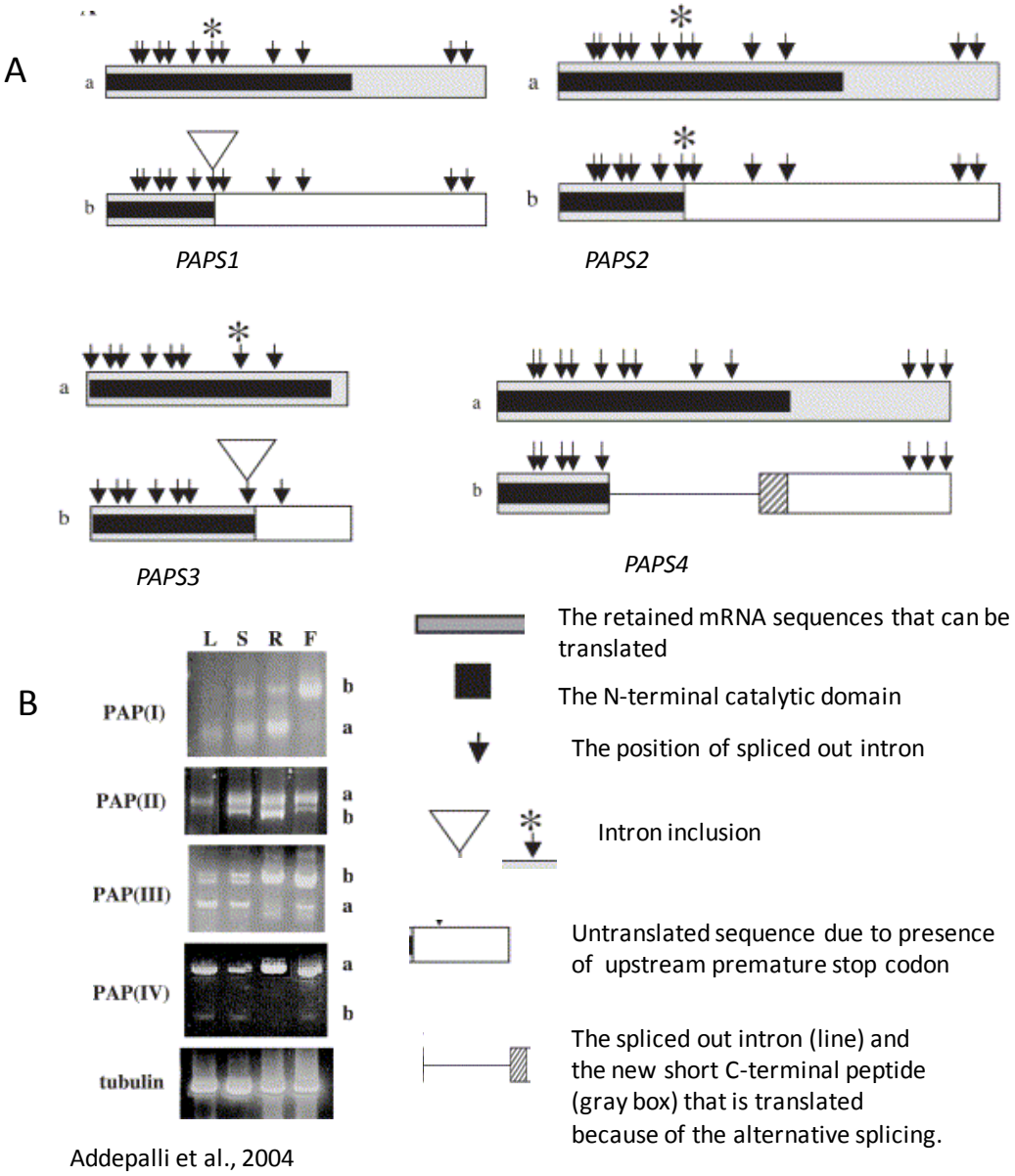
1.4 Regulation of canonical polyadenylation:

In this section, I summarize what is known about the regulation of polyadenylation with two main aspects: the poly(A)-tail length and the poly(A) site.

1.4.1 Poly(A)-tail length regulation

It is apparent that at the steady state level, cellular mRNAs have very different tail lengths that range from a few adenosines (A's) to ca. 90 A's in yeast (Brown and Sachs, 1998), ca. 250 A's in mammals (Sawicki et al., 1977, Wahle, 1995), ca. 150-200 A's in fly (Benoit et al., 2005) and ca. 150 A's in plants (this study). Measurement of the steady-state tail lengths on a genome-wide scale (Meijer et al., 2007, Beilharz and Preiss, 2007) showed that mRNAs can have very different averaged poly(A)-tail length. Theoretically, tail lengths can be controlled by three processes: i) the synthesis in the nuclei (*de novo* polyadenylation); ii) the shortening of poly(A) tail (deadenylation) in the cytoplasm or nuclei and iii) the extension of poly(A) tail, that is distinct from i) as this is not a *de novo* synthesis of the entire poly(A) tail, it is only addition of adenosines to a pre-existing poly(A) tail.

In practice, it is widely believed that the deadenylation generates most of the heterogeneity of the steady-state tail lengths. Deadenylation is mostly a cytoplasmic event. It is also widely believed that during *de novo* polyadenylation, most of mRNAs are synthesized with the same default tail length, which depends on species. Cytoplasmic extension of the poly(A) tail only applies to certain mRNAs, whose tails are deadenylated in the unfertilized oocytes but then re-adenylated in the cytoplasm after fertilization (Huarte et al., 1992). Below, I summarize the key experimental evidence that this generally accepted view is based on. As will be seen, much of the important evidence is based mainly on *in vitro* systems using a few mRNA substrates, and even *in vivo*, only a few cell types were tested. These experimental set up must be taken into account when extrapolating this generally accepted view to all cellular mRNAs and to mRNAs from different tissues.



Addepalli et al., 2004

Figure 1.14. The alternative spliced patterns of PAPSs.

Figures were modified from the reference indicated.

A. The alternative spliced patterns, two forms of transcripts a and b were shown.

B. The expression level differences between the two forms of transcripts a and b analyzed by RT-PCR. L: leaves, L—leaf, S—stem, R—root, F—flower.

1.4.1.1 *de novo* poly(A)-tail length control:

Mammals *de novo* poly(A)-tail length control

PABP1N ensures that polyadenylation reaches but does not exceed 250 A's (with few exceptions):

In mammals, it has been widely accepted that most mRNAs, after *de novo* synthesis, are homogeneous in poly(A)-tail length, which is approximately 250 adenosines. This conclusion is based on *in vitro* and *in vivo* evidence.

In vivo, (Sawicki et al., 1977) performed pulse chase experiments in HeLa cells using radioactively labeled adenosine. They analyzed the tail length of newly synthesized mRNAs in the nuclei and in the cytoplasm separately, examining the A tail length after various labeling times (from 1 minute to several hours and up to 3 days). The results indicated that in the nuclei, within one minute of labeling, most of the newly synthesized poly(A) tails have a homogenous length of 200 to 280 nucleotides. This pattern did not change over time even after 3 days of labeling, especially no incomplete "nascent" poly(A) could be detected. In contrast, in the cytoplasm the pattern changes over time such that the poly(A)-tail length rapidly becomes homogeneous at 200 to 280 nucleotides in length during the initial 40 minutes but these became heterogeneous in size after labeling periods of longer than two hours. The data strongly suggested that both newly synthesized and steady state mRNAs from the nucleus reach a tail of approximately 230 A's. The results however do not exclude the possibility that some pre-mRNAs with a shorter or longer than 230 A' could have escaped the detection due to their low abundance. In fact, poly(A) tails of such pre-mRNAs with less than 20 nucleotides have been discovered ((Rao et al., 1996) and see below). Also, the results do not exclude that in a different cell type, the nascent poly(A) tail can be different.

In vitro, using proteins from mammals, (Wahle, 1995) could faithfully reconstituted the same *de novo* A tail regulation using purified Cleavage Polyadenylation Specificity Factor (CPSF), and Poly(A) binding proteins (PAB II) and Poly(A) Polymerase (PAP). Precleaved L3-preRNA substrates, which carried poly(A) tails of different length, was used. In the presence of all the three proteins, a homogeneous poly(A) tails of 200-250 A's were found in the final products regardless of how many A's the substrates had before the reaction. PAP alone was not sufficient for this tail-length regulation and PABII and CPSF were necessary for the regulation. Bovine/yeast PAP alone, or yeast PAP combined with the non-compatible bovine CPSF and PABII, failed to synthesize a fixed tail length of 250 nucleotides *in vitro*. In latter *in vitro* study by the same research group using a modified substrate (Kuhn et al., 2009), the authors showed that only As residues were counted as part of the tail. The authors suggested a model for the role of PABII in controlling the poly(A) tail length. In this model, PABII, (or renamed as PABPN1 for poly(A) binding protein nuclear 1) binds poly(A) and forms a folded back structure within the 3' end complex. Once the poly(A) tails goes beyond

250 nucleotides, the poly(A)-PABPN1 structure no longer supports the interaction between PAP and CPSF. So it seems that during *de novo* synthesis of the poly(A) tail, the PABPN1, by binding directly to adenosine residues (one PABPN1 per 15 A's), served as a ruler to measure the poly (A) tail length.

Disruption of the nuclear PABPN1 in mammals and flies causes reduced level of the very long poly(A) species (Apponi et al., 2010, Benoit et al., 2005). However, the effect is quite subtle in the bulk poly(A)-tail analysis. Probably this is because of the mild disruption to the PABPN1 gene. The homolog of mammalian PABPN1 in fly is PABP2. Homozygous null *pabp2* mutants are lethal (Benoit et al., 2005). It is intriguing that PABP2 acts to promote polyadenylation of ubiquitously expressed genes including *sop* and *rp49* mRNA, but contrastingly PABP2 acts to restrict the polyadenylation of *cycB* and *osk* mRNA, which undergo cytoplasmic polyadenylation (Benoit et al., 2005).

Exceptions to the "250 A's" rule have been reported. (Rao et al., 1996) discovered that albumin mRNA in liver has a very short poly(A) tail ranging from 12 to 17 residues. The author showed that both intron-containing pre-mRNA and mature steady-state mRNA of albumin have the same short poly (A) tail. Later, a *cis* element, namely Poly(A) limiting element (PLE), was identified as being responsible for this regulation(Das Gupta et al., 1998). PLEs have been mapped to the terminal exons (Das Gupta et al., 1998, Gupta et al., 2001). Splicing factor U2FA binds to this element (Gu and Schoenberg, 2003). Surprisingly, the presence of this PLE in the 3' end of reporter genes stimulates faster 3' end processing, resulting in increased accumulation of mRNA compared to controls without PLE elements (Peng et al., 2005). These PLE-containing short-tail reporter mRNAs are efficiently recruited to polyribosome, and the final protein levels are comparable to those translated from the control PLE-lacking, long-tail mRNAs (Peng and Schoenberg, 2005). The group also found conserved motifs of PLEs in many mRNAs that have short (<20 A) tails suggesting that nuclear regulation of poly (A) tail length is a feature of many mRNAs (Gu et al., 1999).

In sum, the *in vitro* and *in vivo* evidence here strongly support the default *de novo* poly (A) tail length of 250 nucleotides for most pre-mRNAs in mammals. However, the discoveries of extremely short oligo(A) tail, PLE containing pre-mRNAs suggest that, in different cell types or with different pre-mRNA substrates, the *de novo* poly(A) tail can deviate from this rule. Apart from PLE, there is little evidence for regulation of the length of poly(A) tail added during pre-mRNA processing.

Yeast *de novo* poly(A)-tail length control

In yeast, *in vivo* evidence using conditional lethal mutants that disrupt core components of 3' end processing machinery such as *pap* (Proweller and Butler, 1994), *hrp1*(Kessler et al., 1997), rna15 (Kessler et al., 1997), *fip1* (Preker et al., 1995), supported the roles of these factor in poly(A) tail length control. However, in my opinion, it is questionable whether the reduction in poly(A)-tail length reflects the direct effect of *de novo* polyadenylation or the

indirect effects of inhibiting transcription and other cellular processes. In all of the studies, the authors did not uncouple the detrimental effect of disrupting these components to general transcription and cell growth. Apart from *fip1* and *pap* in which the effect on bulk poly(A) was found shortly after switching to the non-permissive temperature (after 5 minutes and 30 minutes, respectively), for other mutants, the poly(A) tail was already found to be different before the temperature switch and was examined again 3 or 6 hour after the switch. During this long period of time, transcription could have been inhibited hence the general reduction in poly(A) tail could reflect the reduction in newly synthesized nascent poly(A) tails and compared to the steady-state of deadenylated poly(A) tails.

Yeast lacks a homolog of human nuclear PABPN1, so the PABPN1-dependent poly(A)-tail regulation is not present in yeast. Instead, two other poly(A) binding proteins were shown to regulate poly(A) tail length. Both proteins function to suppress polyadenylation rather than to promote polyadenylation like in mammals. They are the cytoplasmic poly A binding protein Pab1p (Sachs and Davis, 1989)and the nuclear poly(A) binding protein Nab2p (Hector et al., 2002).

pab1 deletion is lethal. Pab1p was found to interact with cleavage factor IA (CFIA) and suppress polyadenylation *in vitro* (Minvielle-Sebastia et al., 1997). Despite the steady-state cytoplasmic location, Pab1p can also be present in the nucleus. Later experiments confirmed that Pab1p can shuffle between nucleus and cytoplasm and suggest that the N terminal domain carries the nuclear targeting signal (Dunn et al., 2005, Brune et al., 2005). Based on bulk poly(A)-tail analysis of mutants and studies on physical and genetic interactions of Pab1p, the following model was proposed. In the nucleus, Pab1p binds to poly(A) tails and recruits PAN (Poly(A) ribonuclease) complexes to trim off the poly(A) tail of newly synthesize mRNAs, which are 70-90 A's by default, to messenger specific poly(A) tails ranging from 55-70 nucleotides (Brown and Sachs, 1998). The model is based on *in vivo* poly(A)-tail length measurements, however, only three pre-mRNAs were analyzed (RPL46, PGK1 and MFA2). By using a GAL-inducible promoter to analyze only newly synthesize mRNAs, the authors also showed that the poly(A) tails were trimmed very rapidly (4 to 12 minute after induction) and the trimming could not be uncoupled from the *de novo* 3' end processing of the pre-mRNAs. Later experiments identified other positive and negative regulators of the Pab1p-PAN pathway in controlling poly (A) tail length (Mangus et al., 2004). I summarized these findings in **Figure 1.15**. The exosome mutant *rrp6* can suppress *pab1* lethality (Dunn et al., 2005, Brune et al., 2005).

PAN is one of the two deadenylation complexes in *S.cerevisae* (reviewed in (Parker and Song, 2004)). Its activity is stimulated by Pab1p. PAN contains 2 subunits, PAN2 and PAN3. PAN2 is a member of the RNaseD family and therefore is the catalytic subunit of PAN complex (Boeck et al., 1996). PAN3 interacts with and is thought to enhance and regulate PAN2 (Brown et al., 1996).

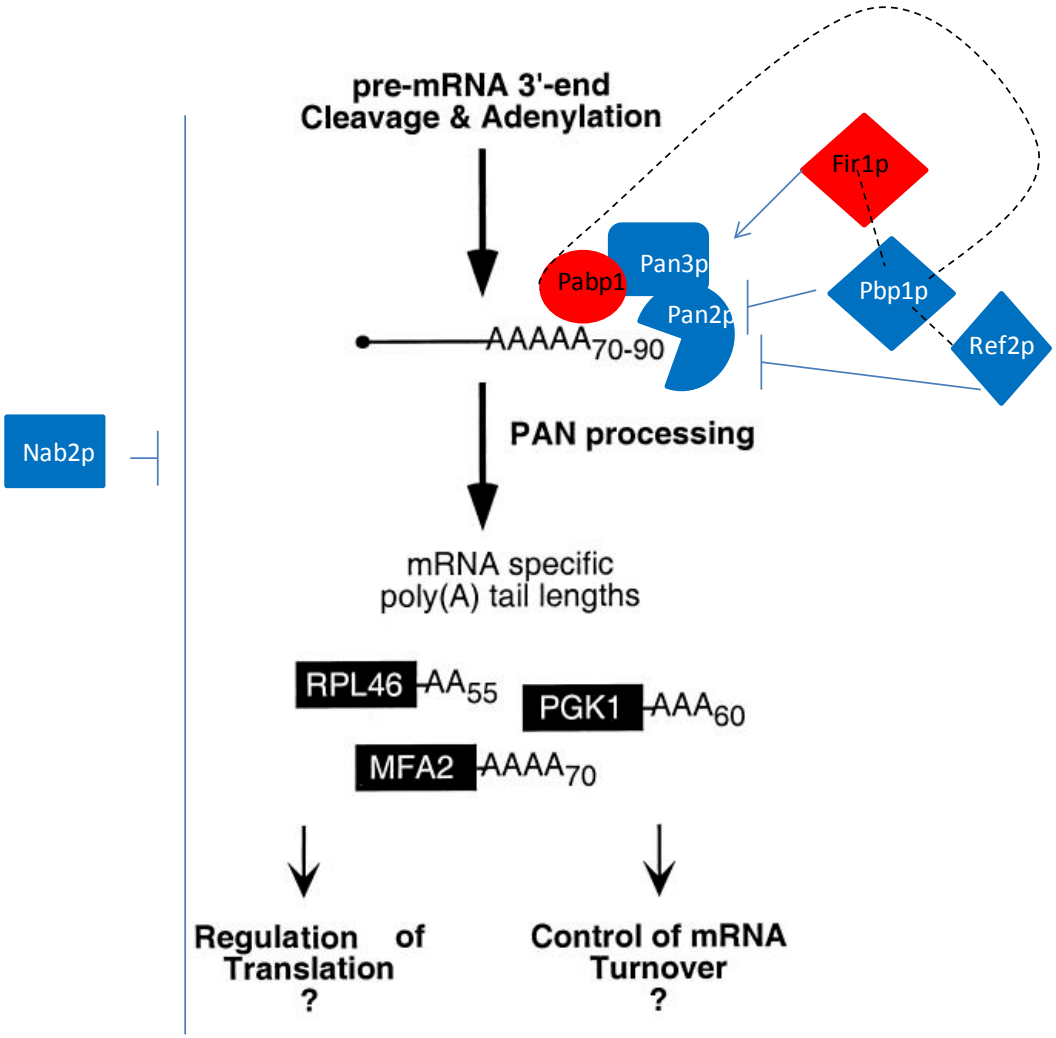


Figure 1.15. The model of poly(A)-tail length regulation in yeast.

Figure is modified from Brown and Sachs,1998 and Mangus et al., 2004. Proteins that interact physically with each other are connected with dash line.

PAN3 also interacts with Pab1p and with Mex67p(Ito et al., 2001), a shuttling protein involved in mRNA export. *pan2* and *pan3* single mutants and *pan2 pan3* double mutant show hyperpolyadenylation but in contrast to *pab1*, these mutants survive and display no growth defects (Brown et al., 1996, Boeck et al., 1996).

Nab2p is another poly(A) binding protein in yeast. Unlike Pab1p, Nab2p is constitutively nuclear. Nab2p is an essential nuclear poly(A) binding protein that plays a role in both poly(A) tail length control and nuclear export (Hector et al., 2002). Disruption of *NAB2* causes hyperadenylation of poly(A) tail and accumulation of poly(A)+ mRNA in the nucleus. The effect on mRNA export could cause synthesis-independent hyperadenylation, because the export defect keeps mRNAs in the nucleus hence protects them from the deadenylase enzyme in the cytoplasm, potentially increasing the average length ofpoly(A) tails.To rule out this possibility, the authors grew a cold sensitive mutant *nab2-21* at 30^0C to uncouple export defects and the poly(A)-tail length regulation. The experiments confirmed the direct role of Nab2p in the poly(A)-tail length regulation. The authors suggested that Nab2p functions to terminate polyadenylation. Also in this study, overexpressing Pab1p or targeting Pab1p to the nucleus rescued the lethality of *nab2* mutant, suggesting their function is interchangeable. Yet intriguingly, at the molecular level, the PAB1 overexpression strain actually enhances the hyperadenylation defect in *nab2* mutants. However, the export defect in the *nab2* mutant was partially rescued in the transgenic yeast, suggesting this is the explanation for the rescue of the lethal phenotype. In my interpretation, alternative explanations, for example, Pab1p can alleviate the hyperadenylation defects on only a few mRNAs that is essential for growth, can not be ruled out.

To sum up, in yeast *de novo* poly(A) tail of 70-90 nucleotides is subjected to two pathways for controlling tail length in the nucleus during synthesis. They are the Pab1p-PAN pathway, which trims off the A-tail to 55-70 nucleotides depending on transcripts, and the Nab2p, which terminates polyadenylation.

Finally, more mutants affect poly(A) tail length in yeast, but their molecular basis is not clear as reveal in a study on cordycepin sensitive mutants(Holbein et al., 2009).

Plants *de novo* poly(A)-tail length control

It is not known how plants regulate poly(A) tail length during synthesis. It is not known whether newly synthesized poly(A) tails have a fixed tail length. As shown in **Table 1.2,** homologs of human and yeast core 3' end processing factors exist in *Arabiodopsis*. *Arabidopsis* also has eight cytoplasmic poly(A) binding proteins and three nuclear poly(A) binding proteins (**Table 1.2**) expressed in an organ-specific manner (Belostotsky and Meagher, 1993). One of them, PAB3 can rescue the lethality of both *pab1,* and *nab2* mutant when expressed in yeast. However, similar to yeast Pab1p overexpression, overexpressing PAB3 did not rescue the hyperadenylation in yeast *nab2* mutants (Chekanova and

Belostotsky, 2003). It is formally possible that plants use similar *de novo* poly(A)-tail length regulation as mammals.

1.4.1.2 Deadenylation:

Once the mRNAs are exported from the nucleus to the cytoplasm, the homogenously long poly(A) tails are subjected to deadenylation. Deadenylation is the first step in the 3' to 5' degradation of eukaryotic mRNA (**Figure 1.16** reviewed in (Parker and Song, 2004, Goldstrohm and Wickens, 2008) and reference here on. A factor that can influence the steady-state average poly(A)-tail length is the variation in deadenylation rate during a mRNA life span and how short the poly(A) tails need to get in order to trigger rapid degradation. For example, mRNA X, which is rapidly degraded when its poly(A) tail reaches 10-15 nucleotide in length, will have shorter steady-state poly(A) tail length compared to mRNA Y, which is rapidly degraded when its poly(A) tail reaches 40-50 nucleotide in length. The transcript-specific poly(A)-tail-length threshold that primes the degradation of the mRNA is not well explored in the literature.

There are three deadenylation complexes: CCR4/POP2/NOT, PAN and PARN. PAN has been introduced above and its activity is dependent on PAB1P and it deadenylates pre-mRNA in the nucleus in yeast. PAN also acts in mammals (Yamashita et al., 2005). In the cytoplasm, CCR4 is the predominant deadenylase complex.

Again, it is not clear how important the role of cytoplasmic/nuclear deadenylation is to poly(A)-tail length control in plants. Plants have homologs to proteins of all three deadenylation complexes. Functional studies exist for only AtPARN and AtCCR4.

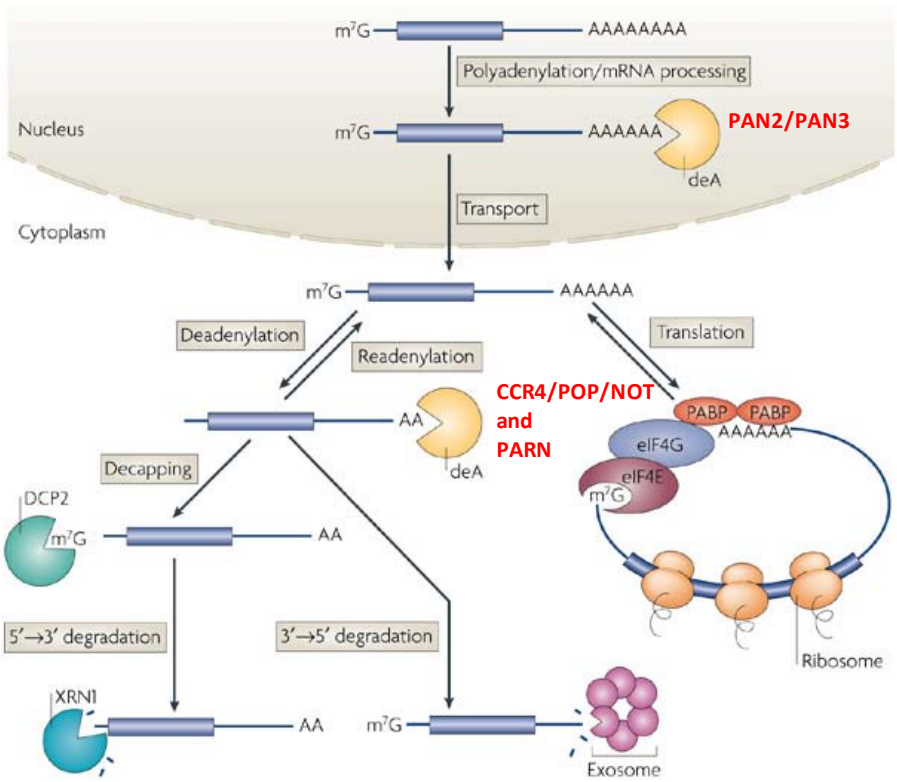
AtCCR4 is transiently upregulated by three plant hormones (jasmonic acid JA, salicyclic acid SA and abscisic acid ABA) and pathogen (Liang et al., 2009). Mutants and overexpression of AtCCR4 reduced and enhance PR1/2 (PATHOGEN RELATED 1/2)-mediated pathogen responses respectively, resulting more or less susceptible to *Pseudomonas syringae*.

The strong AtPARN mutant allele is embryo lethal (Reverdatto et al., 2004), the weaker allele *ahg2-1* (acid abscisic (ABA) hypersensitive germination 2) displays a specific phenotype: ABA hypersensitive (Nishimura et al., 2009, Nishimura et al., 2005). However, poly(A)-tail lengths were not analyzed, and no PARN-sensitive targets were identified.

1.4.2 Regulation of cleavage and the site of polyadenylation (Alternative polyadenylation)

For some genes, there can be more than one poly(A) sites, which are also the cleavage sites of pre-mRNAs. Alternative polyadenylation (APA), and the preference of which poly(A) site(s) to use are under regulation and have diverse roles in gene regulation (reviewed

A. mRNA degradation pathways in eukaryotes



B. Deadenylase complexes in eukaryotes.

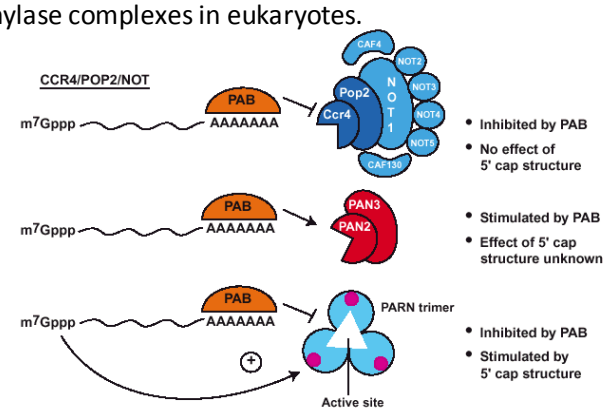


Figure 1.16. Deadenylation and mRNA degradation pathways in eukaryotes.

A. From Goldstrohm and Wickens, 2008

B. From Parker and Song, 2004.

extensively in (Lutz and Moreira, 2011, Lutz, 2008, Xing et al., 2010, Wang et al., 2008). I summarize the kinds of APA and highlight the functions of APA from recent studies.

Kinds of alternative polyadenylations:

For coding genes, alternative polyadenylation can be classified into three types depending on where the poly(A) sites are (**Figure 1.17**). If the gene has only one poly (A) site, no alternative polyadenylation happens.

In type I, the different poly(A) sites lie in the 3'UTRs, hence the coding sequences amongst the alternatively polyadenylated mRNAs are not changed, but the 3'UTR sequences are changed. Many RNA *cis*-acting elements that are present in 3' UTRs can regulate mRNA polyadenylation (via cytoplasmic polyadenylation elements), mRNA stability and translational effiency (via binding sites for miRNAs, translation factor etc…). In type II, the different poly(A) sites lie within exons. If the exon is the last exon, the pre-mRNA is not differently spliced; but if the exon is the middle exon, alternative polyadenylation can be coupled to alternative splicing. In type III, the different poly(A) sites lie within introns. In this type, the splicing patterns of pre-mRNAs change.

In plants, studies on a dataset of 55,000 rice authentic poly(A) sites revealed 50% of the genes have more than one poly(A) site, excluding microheterogeneity of 30 nucleotides (Shen et al., 2008). A recent study using a method for sequencing only the poly(A) site suggested that alternative polyadenylation is even more wide-spread and occurs not only in sense mRNAs but also in antisense RNA and intergenic cryptic transcripts (Wu et al., 2011). The authors identified 57,000 poly(A)-site clusters (PAC) (sites that are within 30 nucleotides away from each other are clustered as one PAC) that mapped to the sense strand of annotated transcripts, 24,000 PAC that mapped to the antisense strand of the annotated transcripts and 14,000 PAC that mapped to intergenic regions. The 57,000 PACs in sense orientation mapped to 18,000 genes, over 70% of which possess more than one poly(A) sites.

Recent discoveries give a hint about the biological role of alternative polyadenylations of antisense transcripts. For example, it was suggested that two proteins in the autonomous pathway that control flowering time in Arabidopsis, FCA, an RNA binding proteins and FPA, a Spen family protein, promote the selection of the proximal poly(A) site in the antisense transcript of the *FLOWERING LOCUS C* (*FLC*) gene, which is a key repressor of flowering (Liu et al., 2010, Hornyik et al., 2010). FCA does that by targeting three components of 3' end processing, FY, CstF64 and CstF77, to the distal poly(A) site of FLC antisense mRNAs. Interestingly, this regulation of the 3' end processing of antisense transcripts helps down regulate the *FLC* sense transcript, hence promoting flowering (Liu et al., 2010).

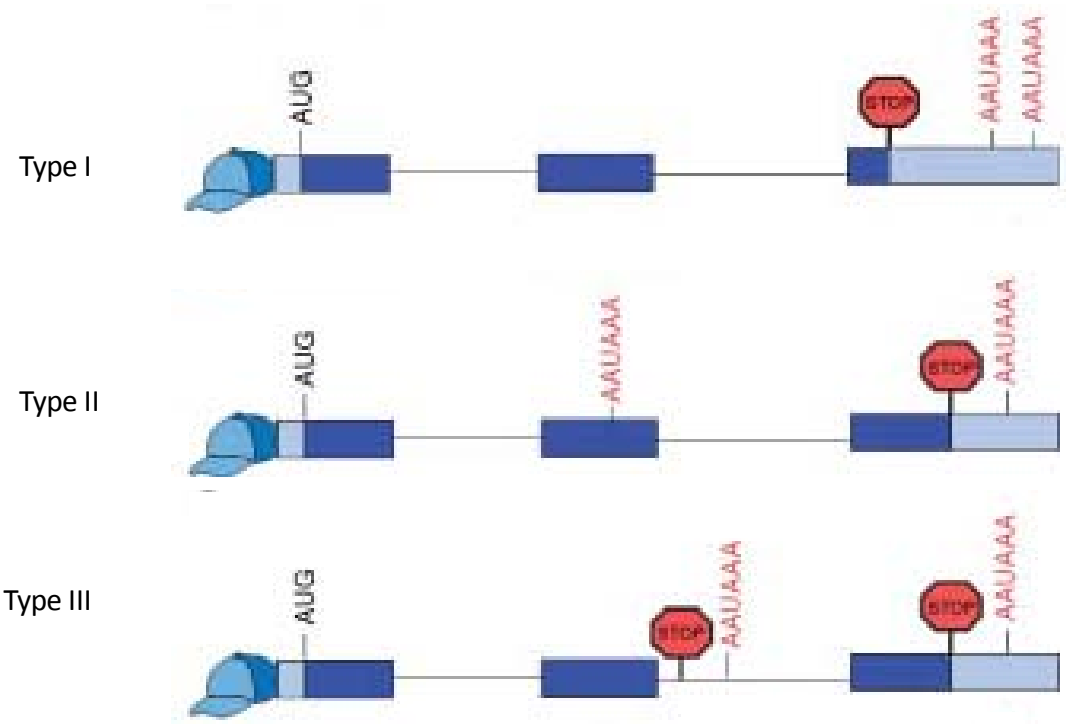


Figure 1.17. Types of alternative polyadenylation.

Figure is modified from Lutz and Moreira, 2011.

Type I. The alternative poly(A) sites lie within the 3'UTR.

Type II. The alternative poly(A) sites lie within exons.

Type III. The alternative poly(A) sites lie within introns.

1.5 Non-canonical PAPs:

Eukaryotes possess several other PAP-like proteins but with very low identity to the human PAPOLA, or yeast PAP1. These proteins are non-canonical PAPs (extensive reviewed in Schmidt and Norbury, 2010). Their roles are diverse. The most well studied is the TRAMP complex that adds short poly(A)tails (adenylation activity is suppressed after 3-4 adenosine (Jia et al., 2011)) to mRNAs, which mark them for degradation. ncPAPs often only target specific mRNAs like mis-folded and other short-lived nuclear RNAs, centromeric siRNAs and their precursors, rRNAs, miRNAs, U6 snRNAs, histone mRNAs, , miRNAs, HO-1 mRNAs (reviewed in Schmidt and Norbury, 2010). Some ncPAPs (star-PAP) can even add oligo(U) tail to mRNAs.

Non canonical PAPS (ncPAP): Arabidopsis possesses nine non-canonical PAPs (**Figure 1.18** (Lange et al., 2009)). These non-canonical PAPs are indentified *insilico* by the sequence similarity to the charaterized ncPAPs in yeast. The function of these geneshave not yet been reported, apart from one, which is At4g00060 (or MEE44) whose disruption causes embryonic lethality (Pagnussat et al., 2005). In addition, *Arabidopsis* also prossesses five bacterial-type PAPs that are targeted to organelles like mitochondria and chloroplasts (Zimmer et al., 2009).

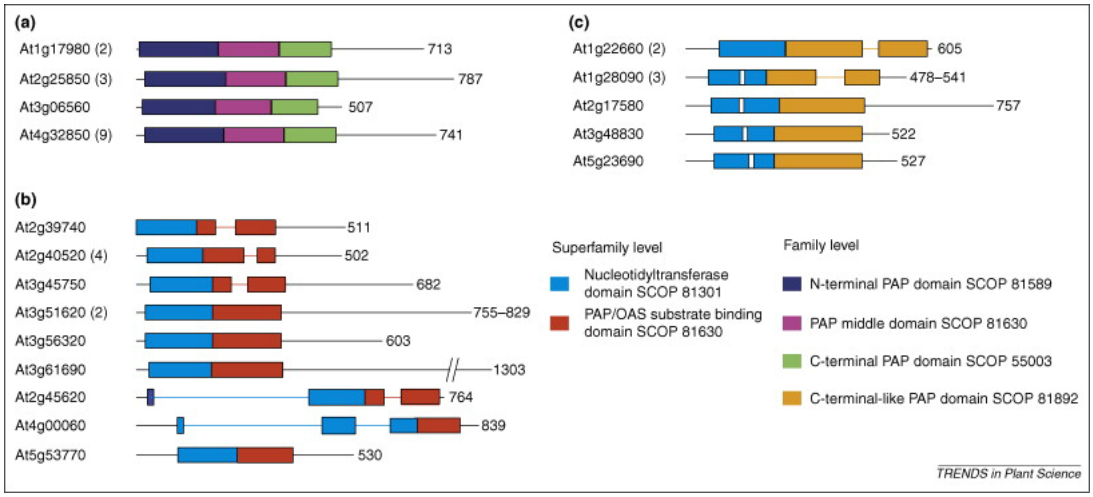


Figure 1.18. Canonical and non-canonical PAPs in *Arabidopsis*.

This figure is taken from (Lange et al., 2009)

Domain architecture of PAPs in *Arabidopsis thaliana.* Boxes represent conserved structural domains identified using the structural classification of proteins (SCOP) according to the superfamily database. Non-conserved regions are drawn as lines. AGI (*Arabidopsis* Genome Initiative) gene numbers are also given. The number of predicted gene models is indicated in parentheses. Protein length is indicated by the number of amino acids on the right.(a) cPAPs contain a characteristic combination of three well-conserved domains and are involved in mRNA 3′ end formation in the nucleus.(b) Nine ncPAPs are encoded in the *Arabidopsis* genome. Compared with cPAPs, the domains of the ncPAPs are more heterogeneous and do not belong to a corporate domain family. Members of the ncPAP family, such as the protein encoded by *At2g45620*, probably function as PUPs rather than PAPs.(c) Bacterial-type PAPs and tRNA-nucleotidyltransferases are characterized by the poly(A) polymerase C terminal region-like fold (yellow boxes), which is distinct from the PAP/OAS domain (poly(A) polymerase and/or oligo adenylate synthase substrate binding domain, red boxes) found in eukaryotic cPAPs and ncPAPs. Bacterial-type PAPs and tRNA-nucleotidyltransferases have the same basic structure but can be distinguished by a 10-amino acid motif (white boxes) that is present in the nucleotidyltransferase domain of bacterial PAPs, but absent in bacterial tRNA-nucleotidyltransferases. In fact, only one tRNA-nucleotidyltransferase is annotated in *A. thaliana* and, whereas the other four members of the family are putative PAPs.

1.6 Project Aims and Objectives:

1.6.1 Preliminary results:

From an Ethyl Methane Sulfonate (EMS)–mutagenized population, D.S noticed a mutant which has reduced leaf size but enlarged flower size. Because of the Lenhard group's interest in size control, the mutant was given to us, initially called *ds39*. The opposite size effects of this mutation on leaves and flowers are particularly exciting. As explained in section 1.1.6 most size regulators affect the size of leaves and flowers in similar ways. There are a few mutants (proteasomal mutants, *cincinnata* mutants and *s6 kinase* mutants) that regulates size of leaves and flowers in opposite directions. The basis of these reverse effects is not yet known. *ds39* therefore provides opportunities to tackle this elusive and important problem: how organ identity modifies the general growth patterns.

In this project, I identified the causal gene by map-based cloning and found that the gene is At1g17980, which encodes for one of the four canonical Poly(A) Polymerases in Arabidopsis.

1.6.2 Aims and objectives:

The project has three aims and several objectives are drawn for each aim:

For the first aim, to understand molecularly how ds39 affects organ size in general, the objectives are:

- To identify the causal gene underlying the *ds39* phenotype.
- To characterize the growth phenotype in detail, including measuring growth parameters: organ size, cell number, cell size of both leaves and petals in a final stage of the organ and also during organ development
- To determine when the gene is required for growth control.
- To determine whether the gene acts autonomously or non-autonomously.
- To study the genetic interactions between the gene and other genetic pathways that control organ size.
- To identify targets of PAPS1 that cause the growth phenotype in leaves and flowers (also for the second aim and third aim).

For the second aim, to clarify why the sizes of leaves and flowers are regulated in the reverse directions in this mutant, the objectives are:

- To clarify whether the reverse growth defect is dependent on organ identity or organ position.

- To identify direct mRNA targets of PAPS1 that cause the growth phenotype in leaves and also the mRNAs that causes the phenotype in flowers (this is also an objectives in the third aim).

For the third aim, to understand why a mutation in such a house keeping gene like canonical Poly(A) Polymerases can have such a specific phenotype, the objectives are:

- To analyze different mutant alleles of PAPS1.
- To analyze mutants in other canonical PAPs and also mutant combinations with PAPS1.
- To identify the specificity in the protein factors (i.e. domains of PAPS) that is responsible for the functional specification amongst PAPS.
- To identify the *cis*-elements (i.e. RNA motifs) that make mRNAs sensitive to PAPS1 defects.

2 .Chapter 2. *ds39* is a peculiar mutant with smaller leaves and larger flowers than wild-type.

In this chapter, I show the detailed phenotypic characterization of *ds39* looking at final organ size, final cell size, growth rate, growth period, cell division vs. time were compared between *ds39* and wild-type (WT) L*er*.

2.1 *ds39* mutant plants have larger floral organs, but smaller leaves:

ds39 mutants make larger flowers and smaller leaves than WT (**Figure 2.1** and **Figure 2.2**). Not only petals are larger (89% increase) but also the other floral organs: sepals (62% increase) anthers (27% increase) and gynoecium (20% increase) are larger, though to a lesser extent. In contrast, leaves of ds39 are only a third of the size of WT leaves. The height of *ds39* mutant plants is also reduced (**Figure 2.1A**).

To compare, the increase of 89% in petal size is a large effect compared to the effect of the other known negative regulators of size: the petal increase in *bpe-1* is 24%, (Szecsi et al., 2006); in the *da1-1* mutant it is 60% (Li et al., 2008) , and in the *bb-1* mutant it is 60% (Disch et al., 2006).

2.2 Final cell size is increased in petals, but decreased in leaves of *ds39* mutants:

I next compared the cell size of petals and leaves of the mutant with wild-type plants.

Epidermal cells on the adaxial face of mature petals at a position a third of the petal towards the tip were measured using the low melt agarose method (see section 9.2.4.8 for description). Using this experimental setting, the cells size in mutant petals are 21% larger than wild-type petals **Figure 2.2A** Compared to the overall increase of 86% in area of mutant petals, it is likely that the mutant petals are larger because of both more cells and larger cells. Although no difference in the spatial pattern of the change in cell-size between mutant and WT petals was observed, in the future, more measurements at various positions in the petals should be examined to statistically confirm this.

Mesophyll cells in the mature forth and fifth leaves were measured by using the chloral hydrate method (see section 9.2.4.8 for description). Cells were sampled at three positions: one third of the leaf towards the tip, the middle of the leaf and one third of the leaf towards the base; all of the positions are the middle point between the midvein and the outer border of the leaf. These measurements at the three positions were averaged for each genotype and the data is presented in **Figure 2.2A**. Using this experimental setting, in leaves of *ds39* mutants, the cell size reduces more than 2 fold compared to WT leaves. Representative

images of the cells at position a third of the leaf towards the tip, away from the midvein are shown in **Figure 2.2C**.

In sum, petals of the mutant are larger because of more cells and larger cells, while the leaves of the mutant are smaller mainly because of smaller cells, with little change in cell number. This suggests that DS39 wild-type gene function restricts both cell division and cell expansion in petals, and promotes cell expansion in leaves.

Compared to other negative size regulator mutants, the effects of the *ds39* mutation in promoting both cell size and cell number in petals is unique. *da1-1* and *bb-1* mutations only increase cell number; cell size remains unchanged (Disch et al., 2006, Li et al., 2008). The *bpe-1* mutation increases cell size and reduces cell division in petals (Szecsi et al., 2006). This suggests that the DS39 size regulator is probably a novel regulator.

2.3 *ds39* petals grow at the same rate but for longer period:

The enhanced petal growth in the mutant can be caused by either an increased growth rate or an extended growth period or both. To answer this question, a developmental series on growing petals was analyzed. Petals were dissected from the flowers/buds in the inflorescences and areas were measured then plotted against their growth periods. The growth period of the youngest bud, which contained the smallest dissectible petals (approximately 50.000 μm^2 in area) was set to be 0. The growth period of the next bud was calculated by adding the plastochron to the growth period of the previous bud. The plastochron (the time difference between two sequential buds) was estimated by counting how many flowers that were made in six days. The plastochron of the wild-type flower is 9.86 hours (SEM= 0.37 hours, n=18 plants); of the *ds39* mutant flower is 19.80 hours (SEM=0.75 hours, n=19 plants). The resulting growth curve is shown in **Figure 2.3A**. In the petals of the mutant, growth rate is the same as WT but the growth period is extended resulting in an overall increase in final petal size. This growth curve is better analyzed when the petal area is transformed to log2 value **Figure 2.3B**. Using log2 petal area vs time data, a straight line can be fitted with a high correlation (r^2>=0.9 for all six lines from six inflorescences analyzed) suggesing that the petal growth is exponential for both mutant and wild-type. The average slope caculated from three WT inflorescences is 0.034 (SEM=0.002, n=3)) and from three *ds39* inflorescences is 0.031 (SEM=0.001, n=3). ttest showed the difference between these two average slopes from mutant and WT is not significant. It is therefore concluded that the mutant and WT petals grow at the same rate but the mutant petals grow for a longer period of time resulting in a final increase in size compared to WT petals. This characteristic, i.e. having the same growth rate, but an extended growth period in *ds39* is shared with *da1-1* and *bb* petals.

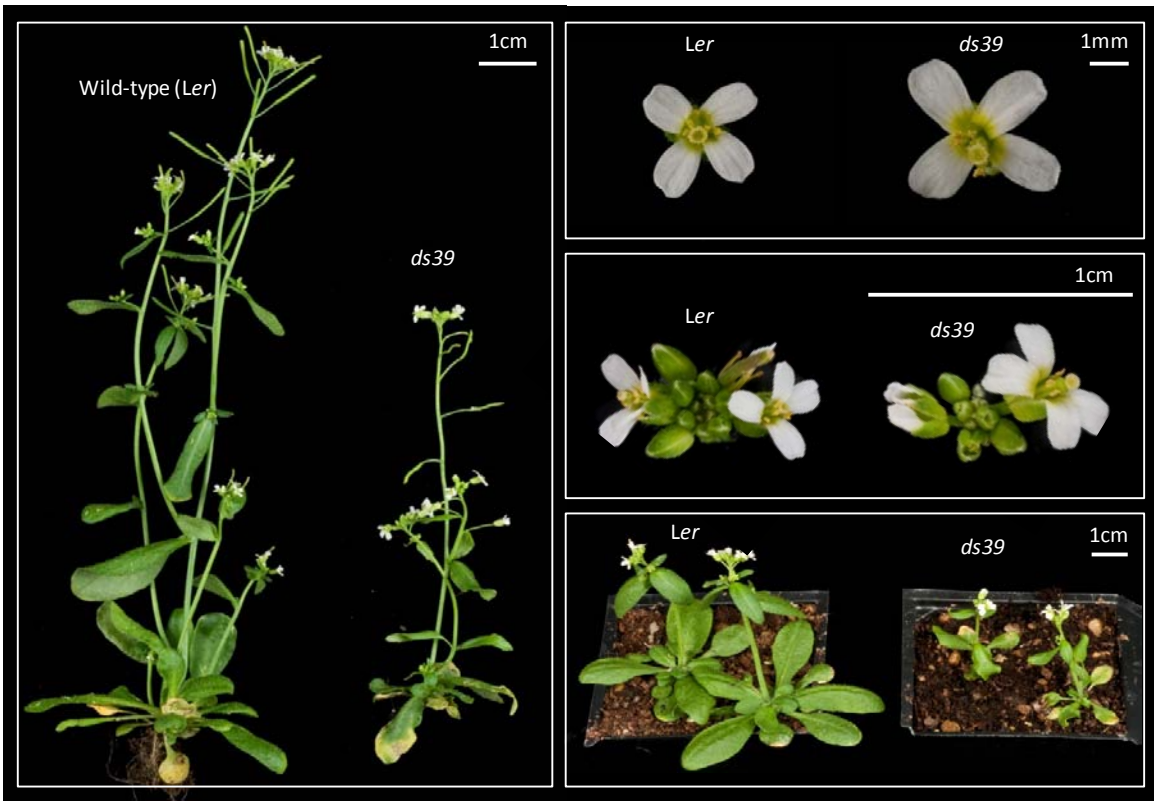


Figure 2.1. *ds39* mutant plants have smaller leaves but bigger flowers than wild-type.

Plants were grown at 21^0C.

A. Plants at 40 days after germination.

B. Mature flowers.

C. Inflorescences.

D. Plants at 30 days after germination.

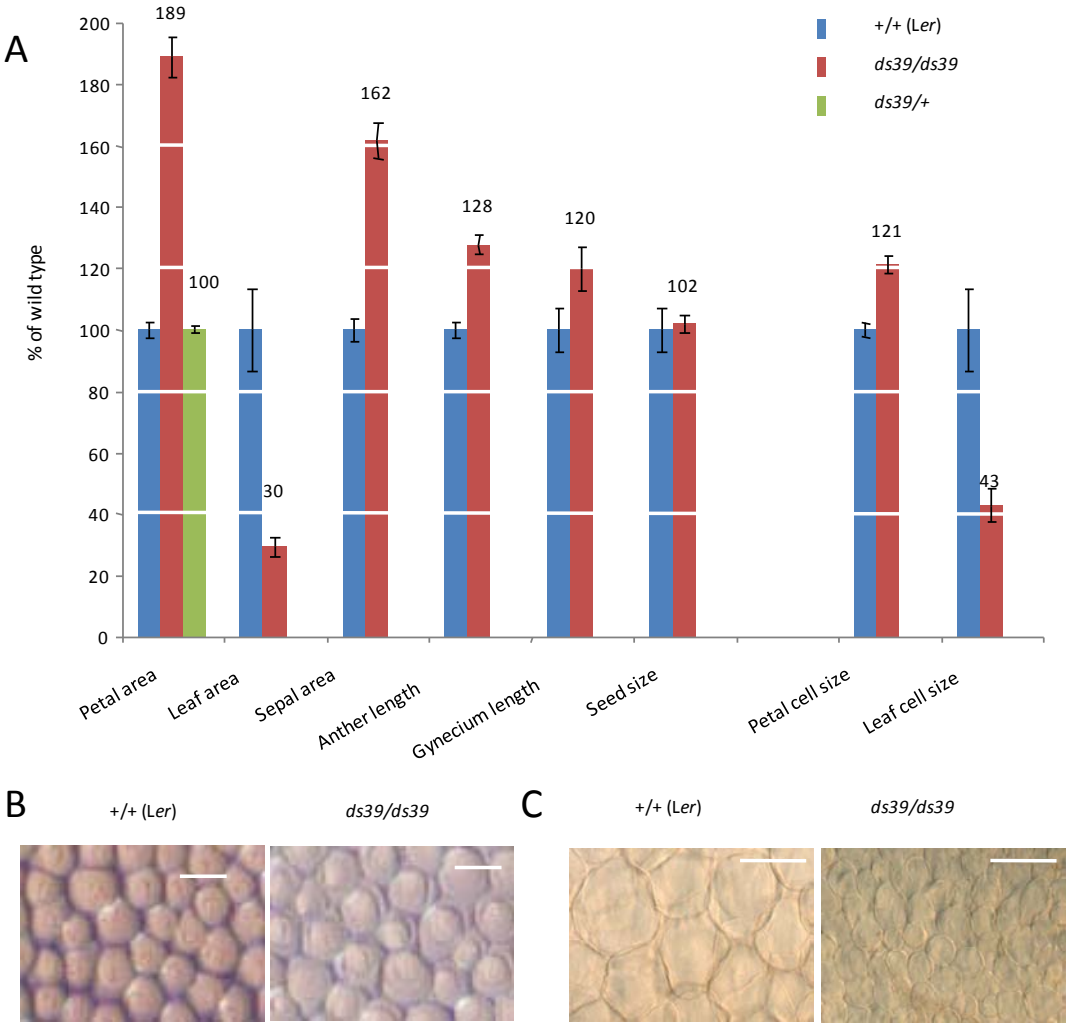


Figure 2.2. Final organ size and organ cell size of *ds39* mutants and WT.

Plants were grown at 21^{0}C.

A. Organs from mature flowers (petals, sepals, anthers, gynoecium) and mature leaves (the 4th and 5th leaves) of plants when they have produced 5 to 10 flowers (approximately 5 week old) were used for measurements. Numbers above the red bar showed the ratio (in %) of mutant/wt for that measurement. Value is mean ± s.e.m; for measurement on petals, sepals and anthers n ≥ 20; for gynoecium n = 7; for leaves n≥5; for seeds n ≥ 55; for petal cell size, at least 20 petals were examined (n ≥ 20); for leaf cell size, at least four leaves were measured (n≥4). All measurements are significantly different between WT and mutant (unpaired t-test p-value <0.01), except the seed size (not significantly different).

B. Cells from petals. They are images of conical epidermal cells on the adaxial face of mature petals at a position a third of the petal towards the tip.Scale bar 100 µm.

C. Mesophyll cells in leaves. Mature fourth and fifth leaves were cleared by chloral hydrate for 2 days, and then images were taken at a position a third of the leaf towards the tip, away from the midvein. Scale bar 50 µm

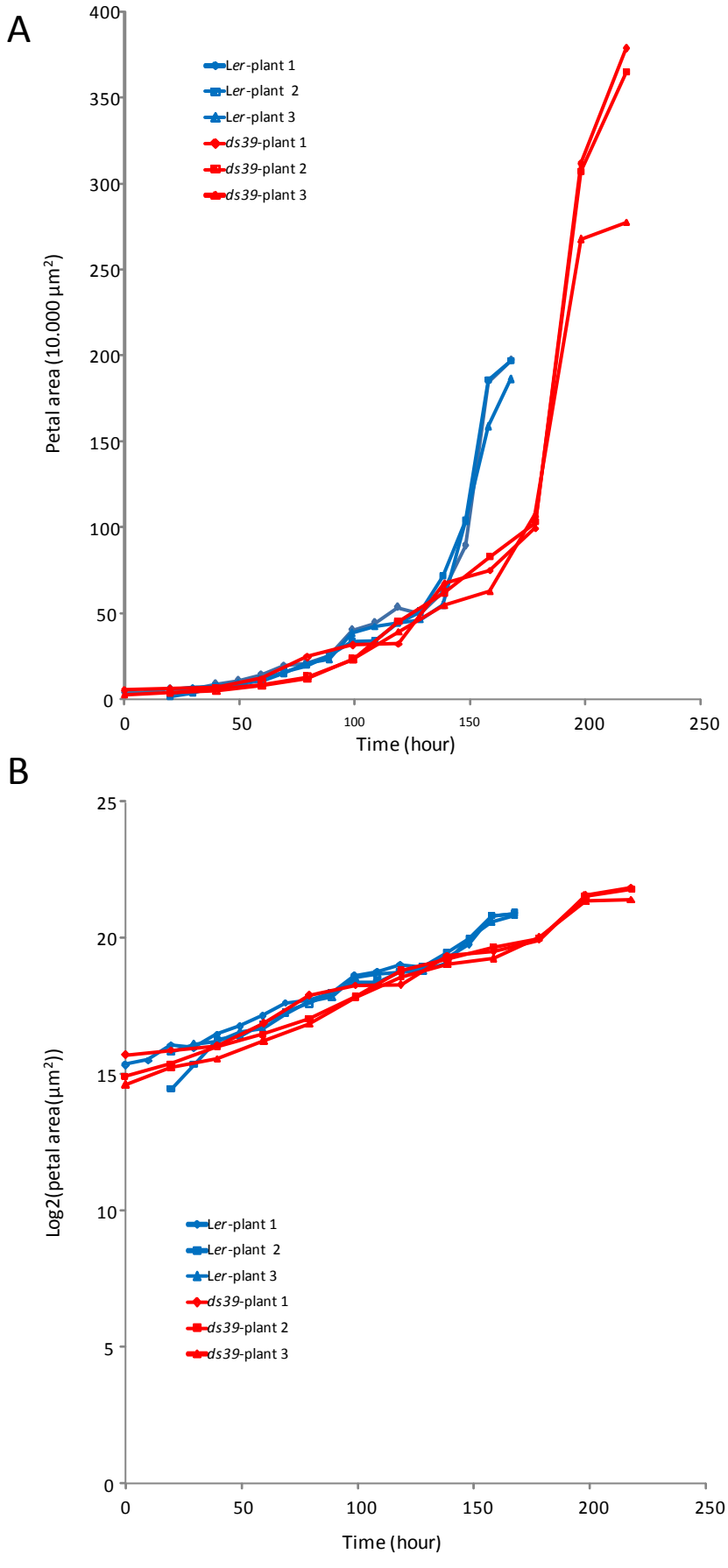


Figure 2.3. The *DS39* gene regulates the duration and not the rate of petal growth.

Figure 2.3. The *DS39* gene regulates the duration and not the rate of petal growth. (continued):

Developmental series of petal growth : Petals were dissected from the flowers/buds in the inflorescences and areas were measured then plotted against their growth periods. The growth period of the youngest bud, which contained the smallest dissectible petals (approximately 50.000 μm^2 in area) was set to be 0. The growth period of the next bud was calculated by adding the plastochron to the growth period of the previous bud. The plastochron (the time difference between two sequential buds) was estimated by counting how many flowers that were made in six days. Two to four petals per bud and three inflorescences per genotype were measured.

A. Petal area vs time.
B. Log2(Petal area) vs time. The petal area data in A. was transformed in to log2 scale.

2.4 Cell division is enhanced in *ds39*:

The increase in cell number in mature petals of the mutant suggests that cell division is enhanced. To understand how this enhancement occurs temporally and spatially, we directly visualized and traced the dividing cell using the mitotic reporter construct *pAtCycB1;1::CDBGUS*. This construct when transformed into plants allows only dividing cells to be stained blue after GUS staining. GUS only accumulates during the G2/M phases of the cell cycle because of the combined transcriptional restriction due to the *pAtCycB1;1* promoter, which is only active around the G2/M transition, and protein degradation at the end of mitosis that results from translationally fusing GUS to the Cyclin Destruction Box of CycB1;1(de Almeida Engler et al., 1999).

As shown in **Figure 2.4A**, cell division in petals is enhanced in the mutant compared to WT. Cell division is also arrested later in mutant compared to wild-type (arrest happens when *ds39* petals reach a size of 0.4 mm^2, while the number for Ler is 0.25 mm^2). This result suggests that the cell division arrest in *ds39* petals is delayed, then resulting in a longer period of growth which contributes to the increase in final size. Additionally, this 60% difference in the area, at which cell division is arrested (0.4 mm^2 for WT vs 0.25 mm^2 for mutant), combined with a 20% increase in cell size predict the total difference in final petal area of 92% (i.e.160% x 120% = 192%). This number is very close to the measured difference in final petal size i.e. 89%.

Spatially, the cell division pattern is not affect in the *ds39* mutant petals (**Figure 2.5**). Similar to wild-type, cell division is arrested first at the very tip and base of the petals and last at the central of the petals.

In leaves, the cell division activity is arrested later in the mutant, similar to what happened in petals of the mutant (**Figure 2.4B**) but to a lesser degree. This increased cell division is not able to compensate for the defect in cell expansion. However, quantitatively, the data does not fully match, cell size of mutant is 43% of cell size of WT, given the same number of cells, the over all size of mutant should be 43% of WT. In fact, the overall leaf size of the mutant is 30% of the WT, meaning that the cell number must also decrease in the mutant. This contradicts with the CDBGUS results. The resolution of GUS staining in this experiment is not at single cell level. So perhaps, in the 'blue' area marked by CDBGUS of the *ds39* mutant, a fewer number of cells divide.

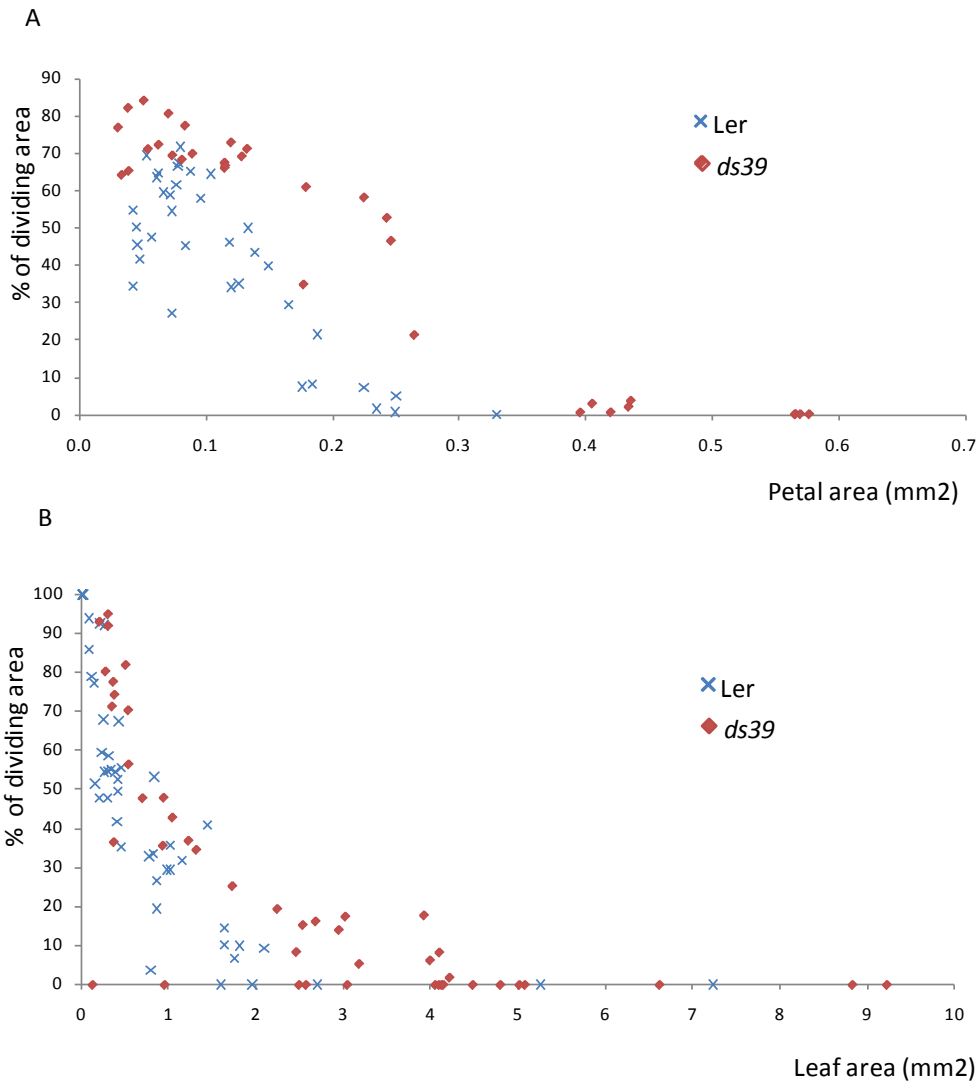


Figure 2.4. Cell division is enhanced in *ds39* mutants compared to WT.

Cell division area was visualized by GUS staining the organs from plants harbouring the mitotic marker construct pAtCycB1;1::CDBGUS. Dividing area is the area that showed blue GUS staining.

A. Developmental series of cell division in petals.

B. Developmental series of cell division in leaves.

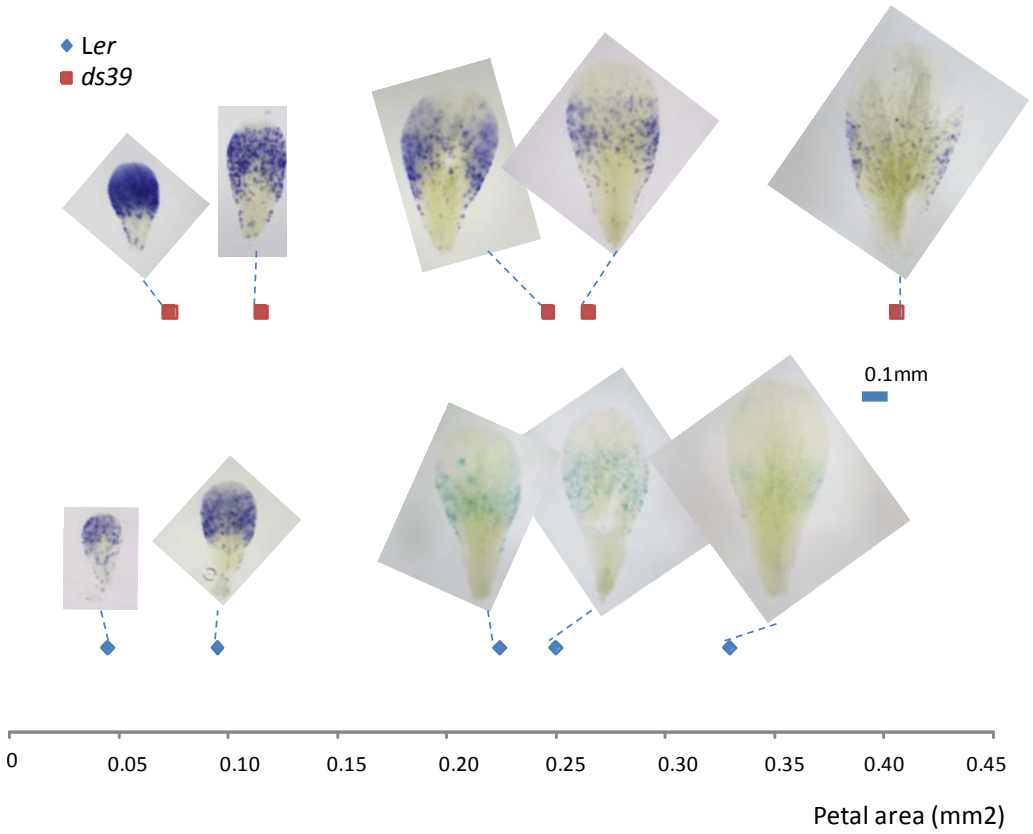


Figure 2.5. Cell division pattern does not change spatially between *ds39* mutants and WT

The positions of red square (*ds39*) and blue dots (WT) represent the petal size. The dashed lines connect the corresponding GUS stained images of the petals to their areas. All images are at the same magnification.

2.5 Other phenotype of *ds39*:

ds39 plants grow more slowly, form fewer leaves, fewer flowers and fewer side shoots (**Figure 2.1C** and **D**). Siliques of *ds39* mutants are shorter. Seed size is not changed (**Figure 2.2A**). Total seed number is much reduced. The seeds also have a reduced germination rate.

2.6 The opposite phenotypes of *ds39* on flower size and leaf size are dependent on the organ identity rather than organ position:

I ask whether the opposite phenotype of *ds39* on flower size and leaf size are dependent on the organ identity rather than organ position. The organ position scenario can happen for example: there could be a mobile growth inhibitor that travels from roots to leaves and flowers to control size; in *ds39* mutant, the mobility may be blocked so that the inhibitors may not be able to reach the flowers. In such case, leaves accumulate the inhibitors and grow less well but flowers are free of the inhibitors and enlarge.

To test this possibility, I combine *ds39* with a mutant background that changes an organ identity but keeping its position intact. *ap2-1* (Bowman et al., 1989) mutant transform sepals in to leaf-like organs. Sepals of *ds39* mutants are larger than L*er* WT but 'leaf-like' organs in the second whorles of the flower s of *ds39 ap2-1* double mutants are smaller than the 'leaf-like' organs in the second whorles of the flowers of single *ap2-1* single mutants (**Figure 2.6**). Therefore, the effect of *ds39* on petal enlargement is identity dependent rather than position dependent.

The conclusions in this chapter are: i) in petals, the *ds39* mutation enhances growth by delaying cell division arrest and also by increase cell expansion; ii) By contrast, in leaves, the *ds39* mutation represses growth by inhibiting cell expansion. iii) the opposite size effects of *paps1-1* mutations on leaves and sepals are dependent on organ identity.

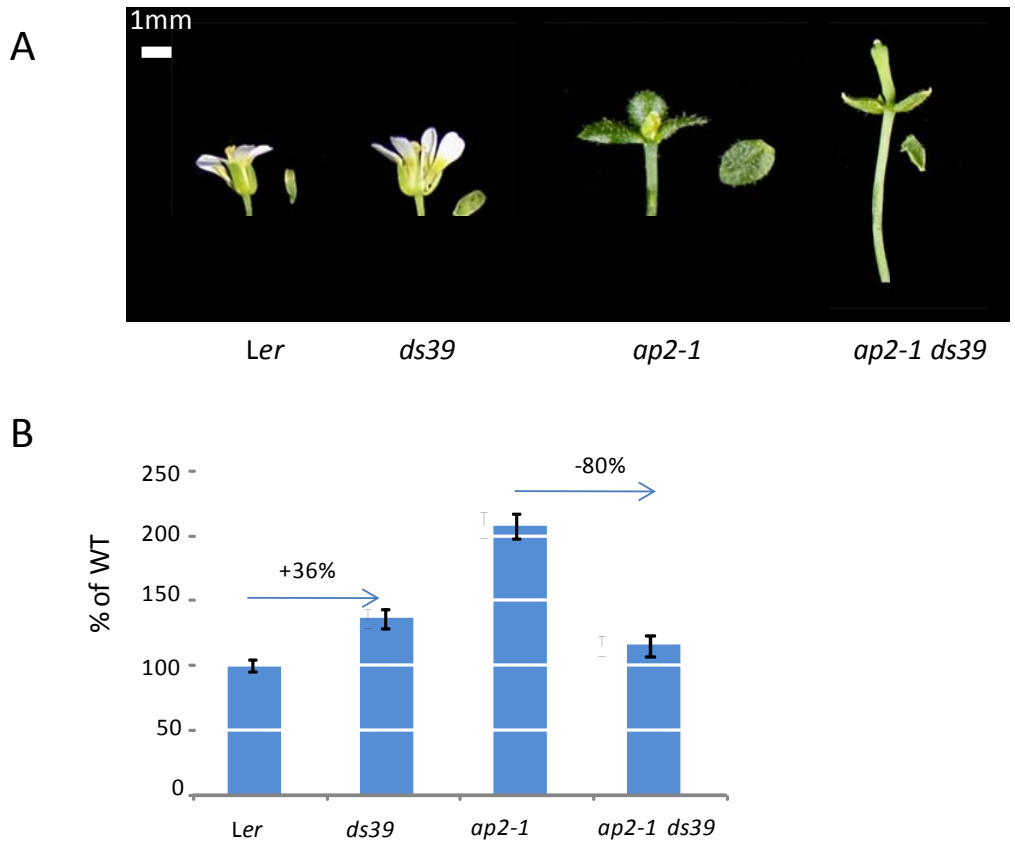


Figure 2.6. The opposite effect of ds39 on sepals and leaves is dependent on the organ identity rather than the position.

A. Floral organs of different genotypes were shown. Note that *ap2-1* mutant transforms the sepals (second whorl) in to leaf-like organs.

B. Size of the organs at the second whorl were measured for different genotype, error bars are S.E.M, n≥10.

3 .Chapter 3.The identification of the *ds39* mutation as *paps1-1* and the phenotypes and interactions among four *paps1* mutant alleles.

3.1 Mapping of the *ds39* mutation:

The gene causing the *ds39* phenotype was identified by map-based cloning (Lukowitz et al., 2000). The mutation is fully recessive as the petal size of F1 plant from a cross between WT and *ds39* is indistinguishable from WT (**Figure 2.2**). In the F2 population from the selfed F1 plants, the phenotypic segregation ratio is 33 mutant:232 wild-type (i.e. mutant ratio is 0.14). This ratio is much higher and statistically different from a ratio of one in 16 (the p-value is less than 0.001 in a χ^2 –test), so it rules out the possibility that the phenotype is caused by two mutations. This ratio is also lower and statistically different from a 1:3 ratio (pvalue is less than 0.001 in χ^2 –test). However, homozygous mutant seeds showed a low germination rate, so the simplest explanation is that the phenotype is caused by one mutation, with the distortion due to reduced germination of the homozygous mutant class.

After PCR-genotyping about 1000 phenotypically mutant individuals from F2 mapping population, the mutation is found to be tightly linked to two markers flanking an interval of 80kb (**Figure 3.1A**). 19 candidate genes (**Figure 3.1B**) in this interval were sequenced. *T-DNA* insertion lines which disrupt genes in this region were ordered. In the end, sequencing identified a C to A point mutation in At1g17980 (**Figure 3.1C**). Additionally, plants that are homozyogous for a *T-DNA* insertion that disrupts this gene were found to have smaller and pointed leaves. This suggests At1g17980 is the causal gene.

3.2 3.2. The causal gene is At1g17980, which encodes a canonical poly(A) polymerase 1 (PAPS1):

We have three lines of evidence to support that the *ds39* mutant is in At1g17980, which encodes for PAPS1. First evidence is from the mapping of *ds39* and the identification of a point mutation in At1g17980 gene in *ds39* mutant. Secondly, transforming *ds39* mutant with a genomic version of wild-type PAPS1 fully rescued the *ds39* phenotype in both leaves and flowers (**Figure 3.2A**). Thirdly, two independent *T-DNA* insertion lines which have the *T-DNA* inserted in PAPS1 locus are allelic to *ds39 (***Figure 3.2B** and **C**).F1 of a cross between *ds39* and these TDNA lines did not show a rescued phenotype as judged from their petal sizes (**Figure 3.2A**). One of them called *pap1-4* also showed similar phenotype to *ds39*, i.e. smaller leaves but larger flowers (**Figure 3.2B** and **Figure 3.3A**). Taken together, these three lines of evidence confirm that the mutation in PAPS1 is responsible for the *ds39* phenotype. From now on the *ds39* mutant allele is called *paps1-1*.

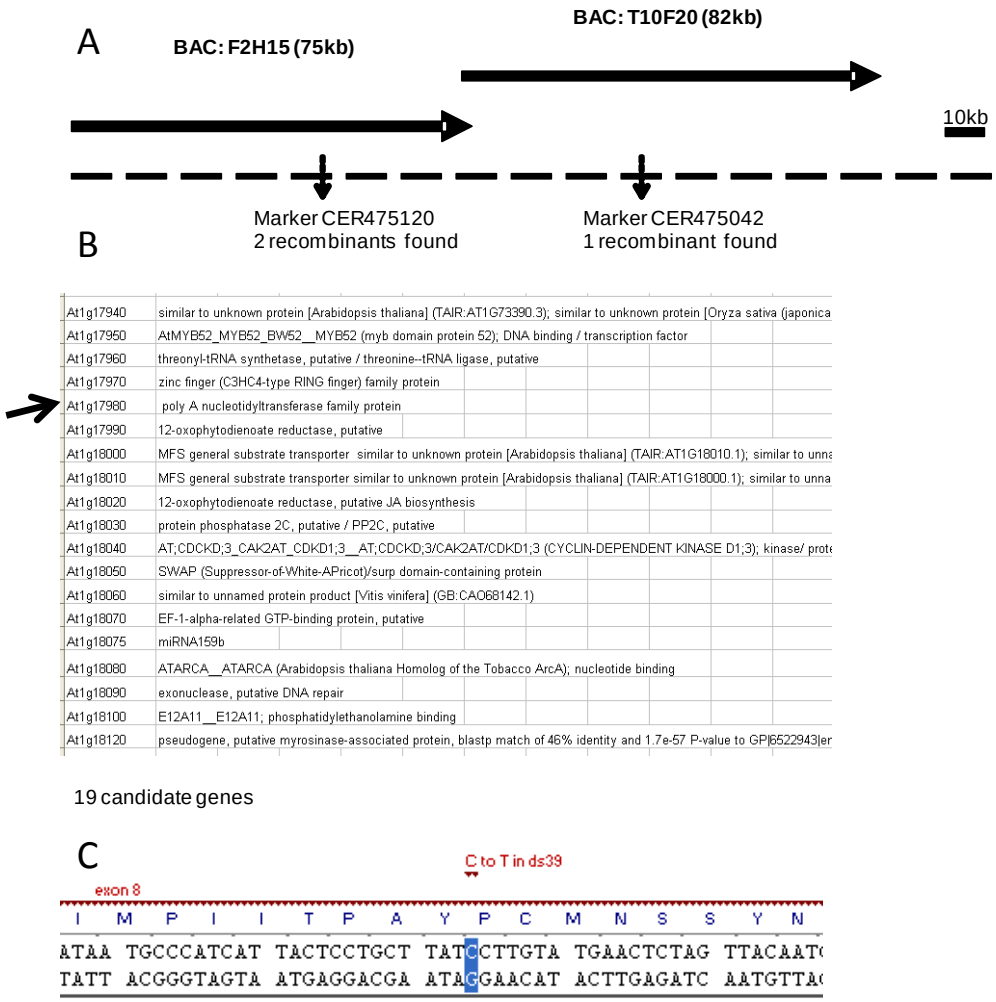


Figure 3.1. Fine mapping of the *ds39* mutation.

A. The *ds39* mutation was mapped to the interval of 60 kb between two markers indicated. Dashed lines represent the chromosome, horizontal thick arrows represent the two Bacterial Artificial Chromosome s (BACs) spanning the chromosomal region, vertical arrows represent the marker position.

B. Genes within the mapping interval (the annotation is from TAIR website). Arrow head points at the gene from which a mutation is found in *ds39* plants.

C. Part of the wild-type nucleotide sequence in exon 8 of At1g17980, where a C to T mutation is found in the *ds39* mutant. Above is the coding strand, below is the complementary strand. The translation of the coding strand is shown on top of the nucleotide sequence. One letter amino-acid code is used.

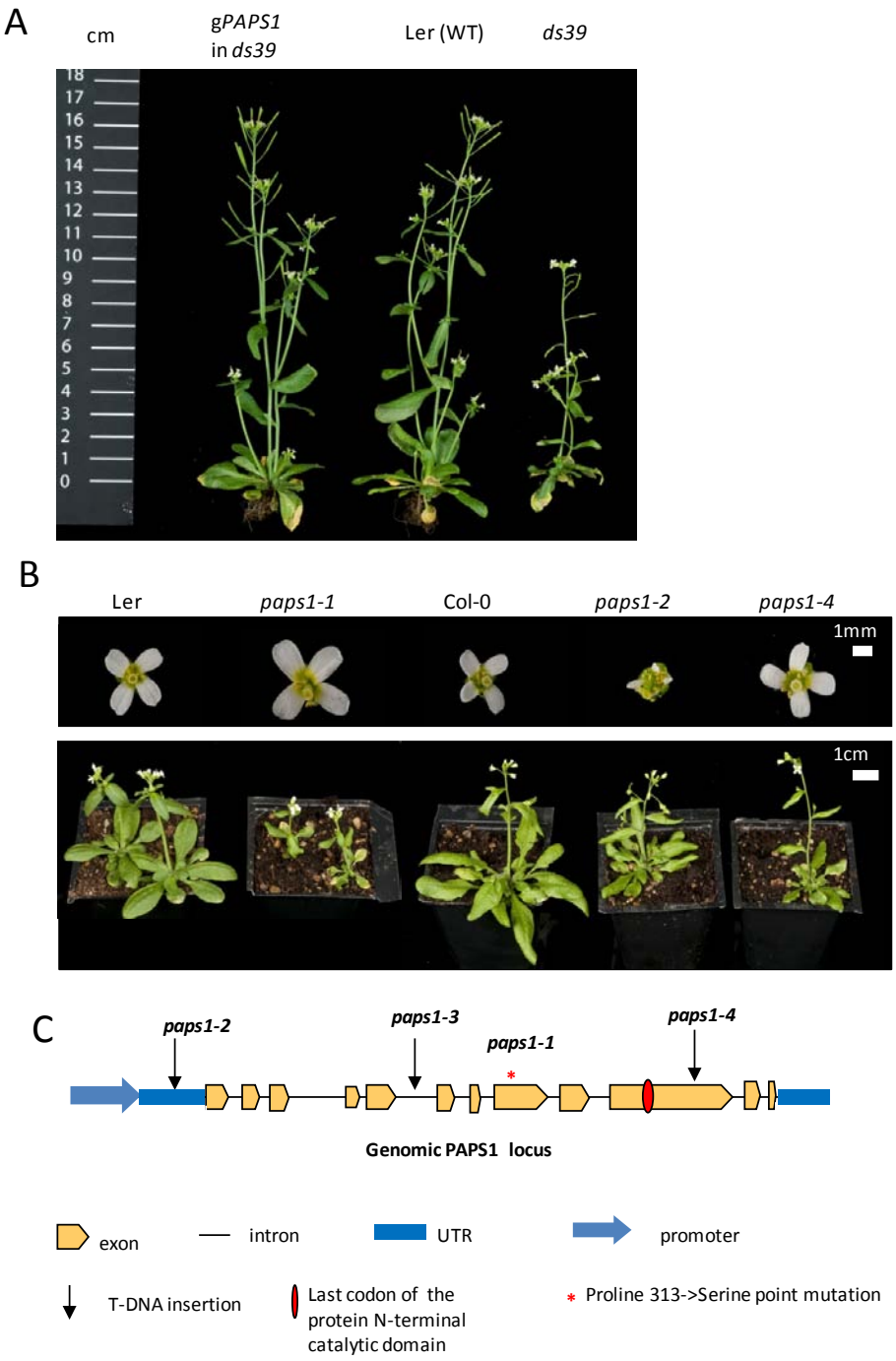


Figure 3.2. The causal mutation in *ds39* affects *PAPS1* locus.

Plants are grown at 21^0 C.

A. Full length wild-type *PAPS1* genomic fragment (g*PAPS1)* rescues *ds39* mutant. From here, *ds39* is called *paps1-1.* Plants are 5 week-old.

B. Whole plants and flower pictures of WT and mutants carrying different *paps1* alleles. Note that *paps1-1* is in L*er* background; *paps1-2*, *paps1-3* and *paps1-4* are in Col-0 background.

C. Genomic *PAPS1* locus showing gene structure, protein domains and positions of the *paps1* allele polymorphisms.

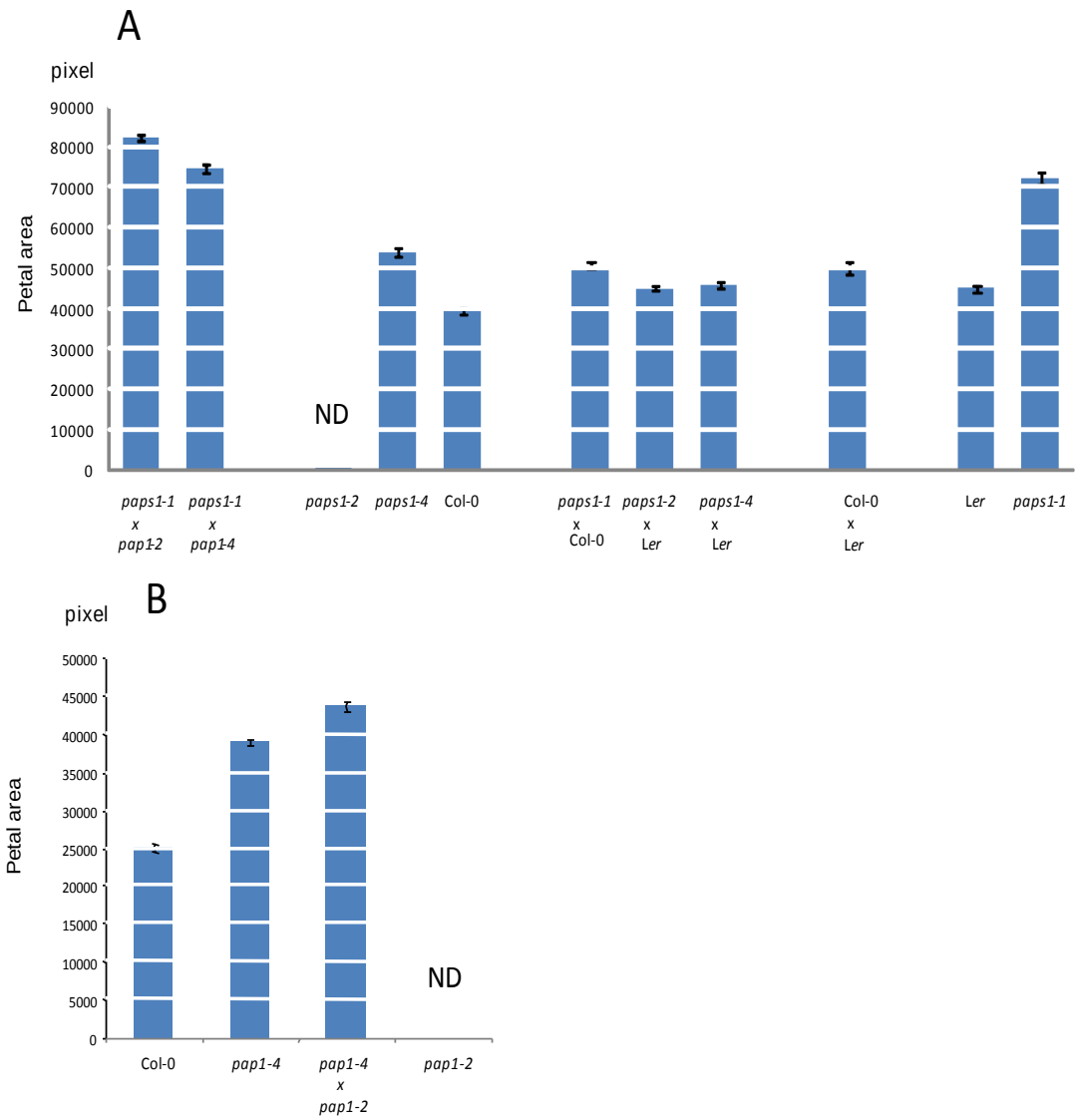


Figure 3.3 Allelic interactions between *paps1* alleles.

Petal size of plants with different genotype, which is shown below the bar, e.g., *paps1-1* x *paps1-2* means F1 plants between *paps1-1* and *paps1-2* homozyogous mutant. Note that here *paps1-1* is in L*er* background and other alleles *paps1-2* and *paps1-4* is in Col-0 background hence control crosses (Col-0 x L*er*) were also measured. *paps1-2* petals were deformed and much smaller than WT so were not measured (ND).

All plants in A were grown and measured together; all plants in B were grown and measured together in a different experiment from A.

3.3 The different *paps1* alleles show a range of different phenotypes:

Because of the peculiar phenotype of *paps1-1* mutant, more *paps1* mutant alleles were isolated. In addition to *paps1-1*, more alleles were isolated, all of which are *T-DNA* insertion alleles (**Figure 3.2B** and **C**). In the next sections, the phenotypes of these four mutant alleles and the defects in their *PAPS1* transcripts were examined in detail.

3.3.1 The *paps1-1* mutation probably changes the conformation of the protein.

The point mutation in *paps1-1* causes an amino acid change from proline at the amino acid (aa) number 313 to Serine (P313S). This point mutation lies in the N-terminal catalytic domain of PAPS1, which is highly conserved across kingdoms (**Figure 3.4**).The only PAPS-like protein that does not have proline at this position is Arabidopsis PAPS3 and Sorghum_Sb02g043400. PAPS3 is quite divergent amongst plants (see **Figure 1.9** and **Figure 1.12**). By alignment, Sorghum_Sb02g043400 was found to to have a large deletion at this region of the protein (**Figure 3.4**).

To know where exactly the amino acid is in the protein, the structure of the N-terminal catalytic domain of bovine PAP (Martin et al., 2000), which shares 43% identity with PAPS1 N-terminal catalytic domain, was examined (**Figure 3.5**). This structure only contains the first 513 aa of bovine PAP, which is the N-terminal catalytic part of the protein and lacks the C-terminal domain (residues 514 to 739). The N-terminal catalytic part is further subdivided in to three domains: the catalytic domain (residues 60-173), the middle domain (residues 20-59 and 174 to 352) and the C-terminal RNA binding domain (residues 353-498) (Martin et al., 2000). By sequence alignment, the position of the *Arabiodopsis* PAPS1-P313 is equivalent to bovine PAP-P321. As seen in **Figure 3.4**, P321 is at the loop between helix K and helix L in the middle domain. So P321 is neither in the catalytic nor RNA binding domain. It is very far away from the active site, i.e. the position of the 3'ATP, the analog of the substrate used in this crystallization. Therefore, the P321S mutated amin oacid should not make any direct contact with either the ATP substrate or pre-mRNA substrate. A closer look at position P321 reveals its potential structural role. In bovine PAP, P321 together with the other close by and highly conserved P318 form a bend in the linking peptide (**Figure 3.5B**). P321S mutation perhaps affects the movement of N-terminal and C-terminal domain during the catalytic cycle of PAP as suggested by (Balbo et al., 2007). This is also likely to be the reason why the *paps1-1* mutation is temperature-sensitive while the other *paps1* mutant alleles are not (see below). Probably, high temperature further enhanced the incorrect folding/conformation of the mutant protein.

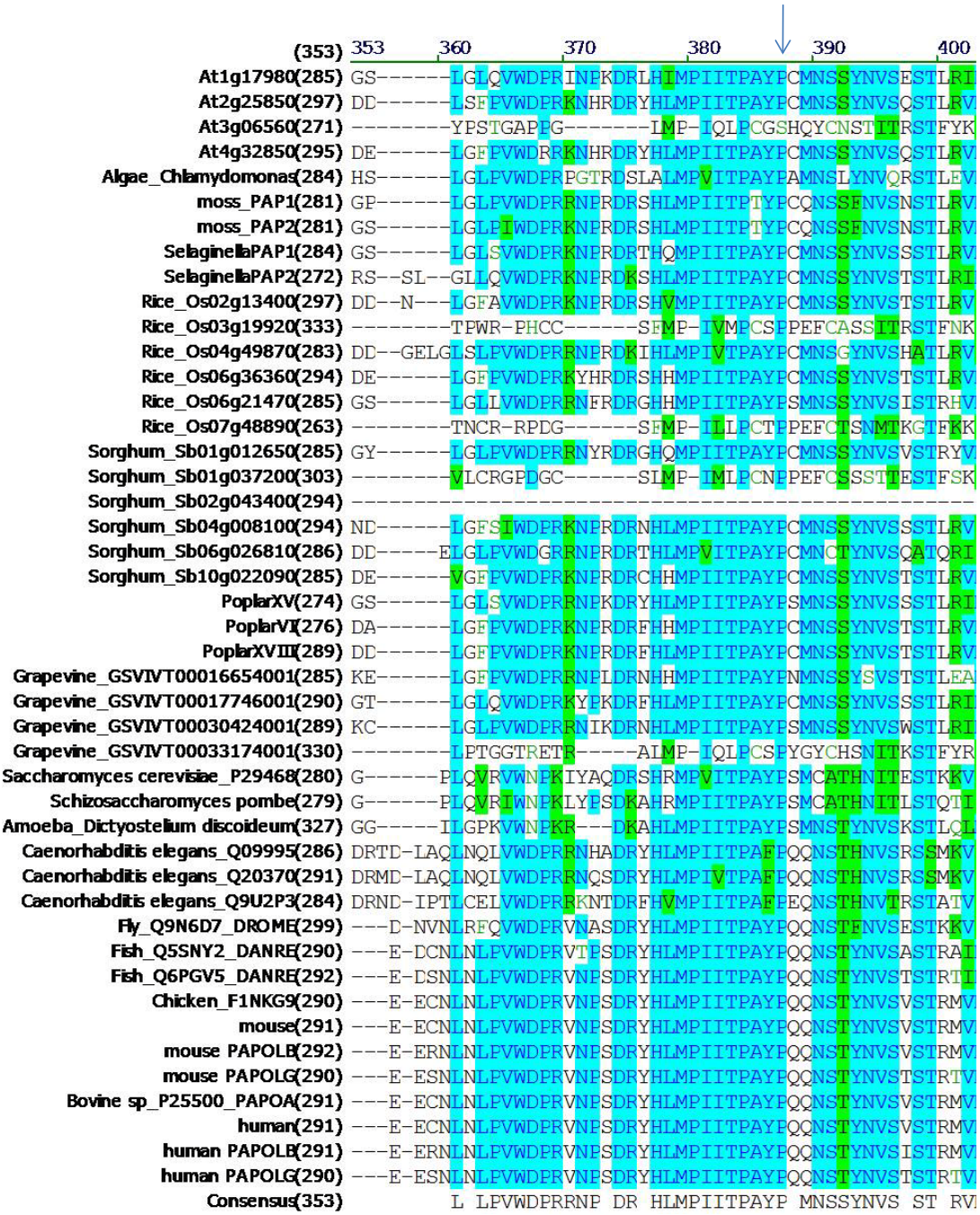


Figure 3.4. Position of the mutated proline residue (P313) in *paps1-1* mutant.

Apart from Arabidopsis PAPS3 (At3g06560) and Sorghum_Sb02g043400, this proline is conserved in all examined model organisms. The FASTA sequences used for the alignment and the full alignment are supplemented in electronic version. Alignment was carried out using alignX-vector NTI. The colour code of amino acid: Weakly similar: Dark green text on a white background ; Block of Similar: Black on a light green background; Conservative: Dark blue on a light blue background; Non-similar : Black on a white background .

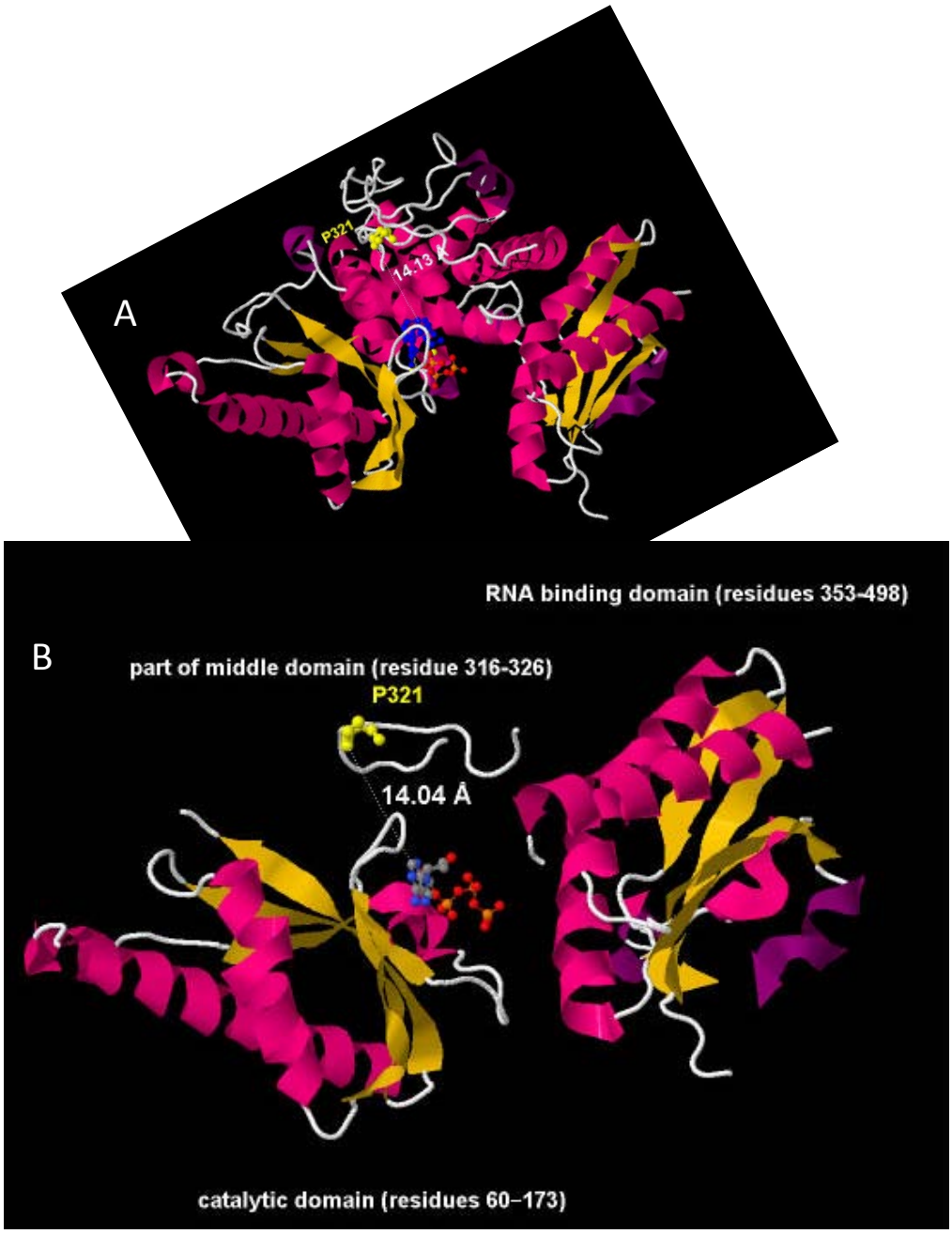


Figure 3.5 Position of the proline 321 (P321) in bovine PAPs

A. The whole structure of the N terminal catalytic domain of bovinPAP (structure 1F5A-protein data bank). Helices are in magenta ribbon, short helices are in purple ribbons, beta strands are in yellow ribbon, linker peptides are in white threads, P321 are in yellow ball and stick, d3'ATP, the substrate analog has its sugar and base colored in blue and gray, its phosphate group in red ball and sticks.

B. Simplified structure where parts of the protein, notably of the middle domain, are removed. Note the curve where P321 lies in the linker peptide of the middle domain. Distance from P321 to the base of d3' ATP is measured to be 14.04 Å.

To sum up, the position of the mutated proline in the protein structure suggests that the PAPS1-P313S mutated protein probably has a reduced function because of impaired conformation.

3.3.2 *paps1-1* mutants are temperature sensitive and this is not due to the higher requirement for *PAPS1* at higher temperature.

Another property of *paps1-1* is that the mutation is temperature sensitive. While WT plants grow faster at 28^0 C than at 23^0 C (**Figure 3.6A**), the *paps1-1* seedlings did not grow at 28^0C as well as they grows at 23^0 C. *paps1-1* seeds did germinated at 28^0 C; however, the seedlings produced very tiny leaves.

This growth depression behavior at high temperature is only found in *paps1-1* but not found in either of the other *paps1* alleles *paps1-2* or *paps1-4* (**Figure 3.6A** and **B**). This observation weakens the possibility that the growth deteriorates at high temperature because plants require more PAPS1 activity to cope with increasing temperature.

Strikingly, while the growth inhibition is exaggerated in *paps1-1* leaves by high temperature, the growth promotion in petals is not. Instead, at 25^0C, the *paps1-1* petals become narrower. They deform mildly (**Figure 3.7A**- compare *paps1-1* flowers at 23^0C and 25^0C). Moreover, when plants are moved from 23^0C , at the stage when they already produce some normal-large flowers, to even higher temperature (28^0C), the newly made flowers (examined one week after the switch) becomes severely deformed and infertile like *pap1-2* flowers (**Figure 3.7B**).

Taken together, the data suggests that the *paps1-1* mutation interferes with the normal conformation of the protein, and this structural defect is further enhanced by high temperature.

3.3.3 *paps1-4* mutants share the phenotype, but showed weaker phenotype compared to *paps1-1*

paps1-4 is a *T-DNA* insertion allele. The *T-DNA* inserted at an exon that ecodes for part of the C terminal domain of PAPS1 protein (**Figure 3.8**).

Similar to *paps1-1* phenotype, *paps1-4* also has smaller and pointed leaves, and bigger flowers, although both the petal size increase and leaf size decrease is to a lesser degree compared to *paps1-1* (**Figure 3.2B** and **Figure 3.3A**)

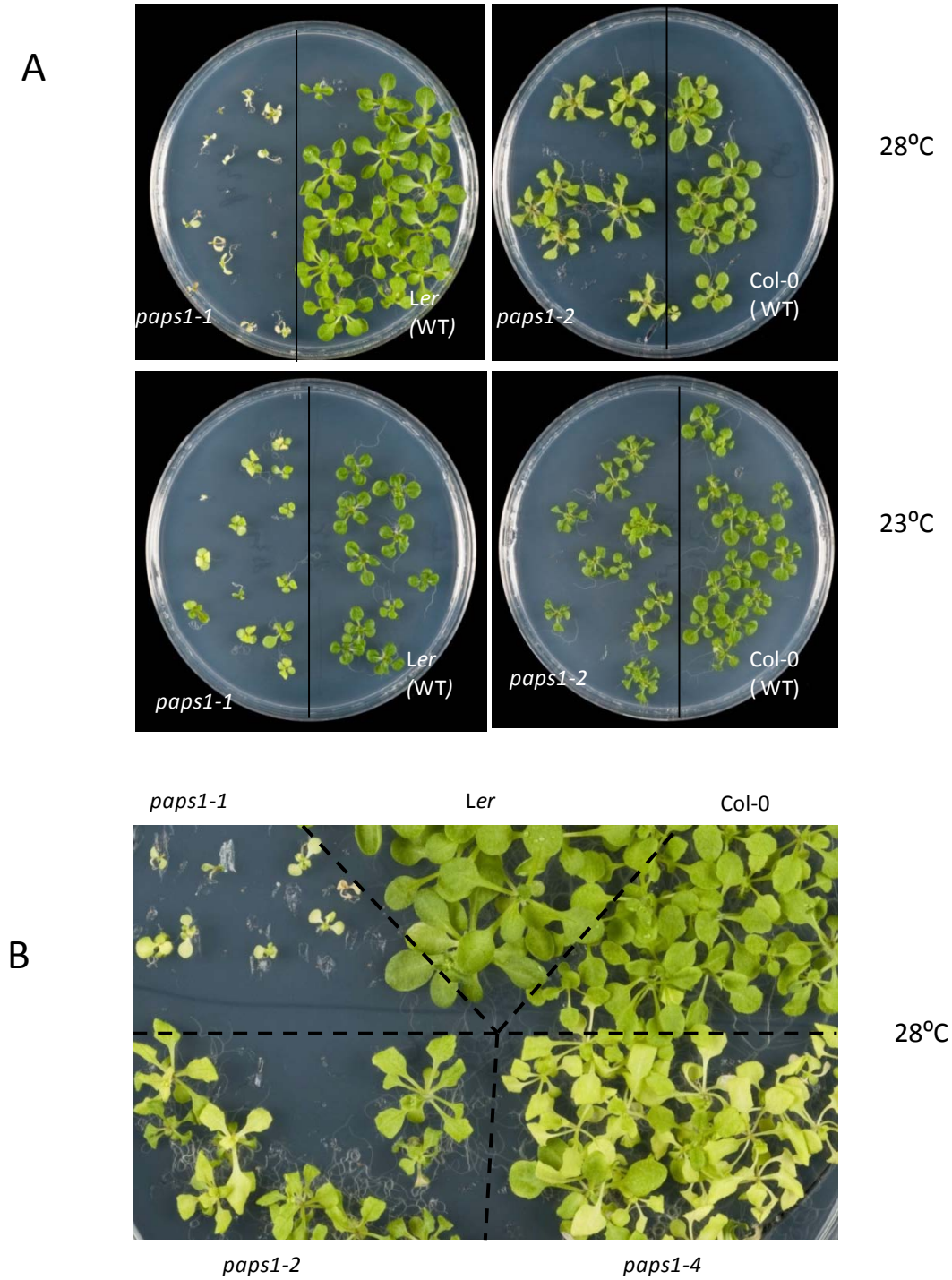


Figure 3.6. *paps1-1* is temperature sensitive but the requirement for PAPS1 activity is not higher at higher temperature

A. 15 day-old seedlings of *paps1-1* (in L*er* background) and *paps1-2* (in Col-0 background) grown at 23^{0}C and 28^{0}C.

B. 25 day-old seedlings of *paps1-1*, WT L*er*, WT Col-0, *paps1-2* and *paps1-4* grown at 28^{0}C.

Figure 3.7. Petals of *paps1-1* plants grown at high temperature (25-28^{0}C) but not at 23^{0}C resemble *paps1-2* petals.

A. Flowers of plants grown at different temperature: *paps1-1* at 23^{0}C and 25^{0}C; *paps1-2* at 23^{0}C; WT (Col-0) at 25^{0}C; *paps1-4* at 25^{0}C. Compare *paps1-1* flowers at 23^{0}C and 25 ^{0}C and *paps1-2* at 23^{0}C.

B. Inflorescenes of *paps1-1* and WT (L*er*) plants one week after being switched from 23^{0}C to 28^{0}C. Note that flowers become deformed and sterile.

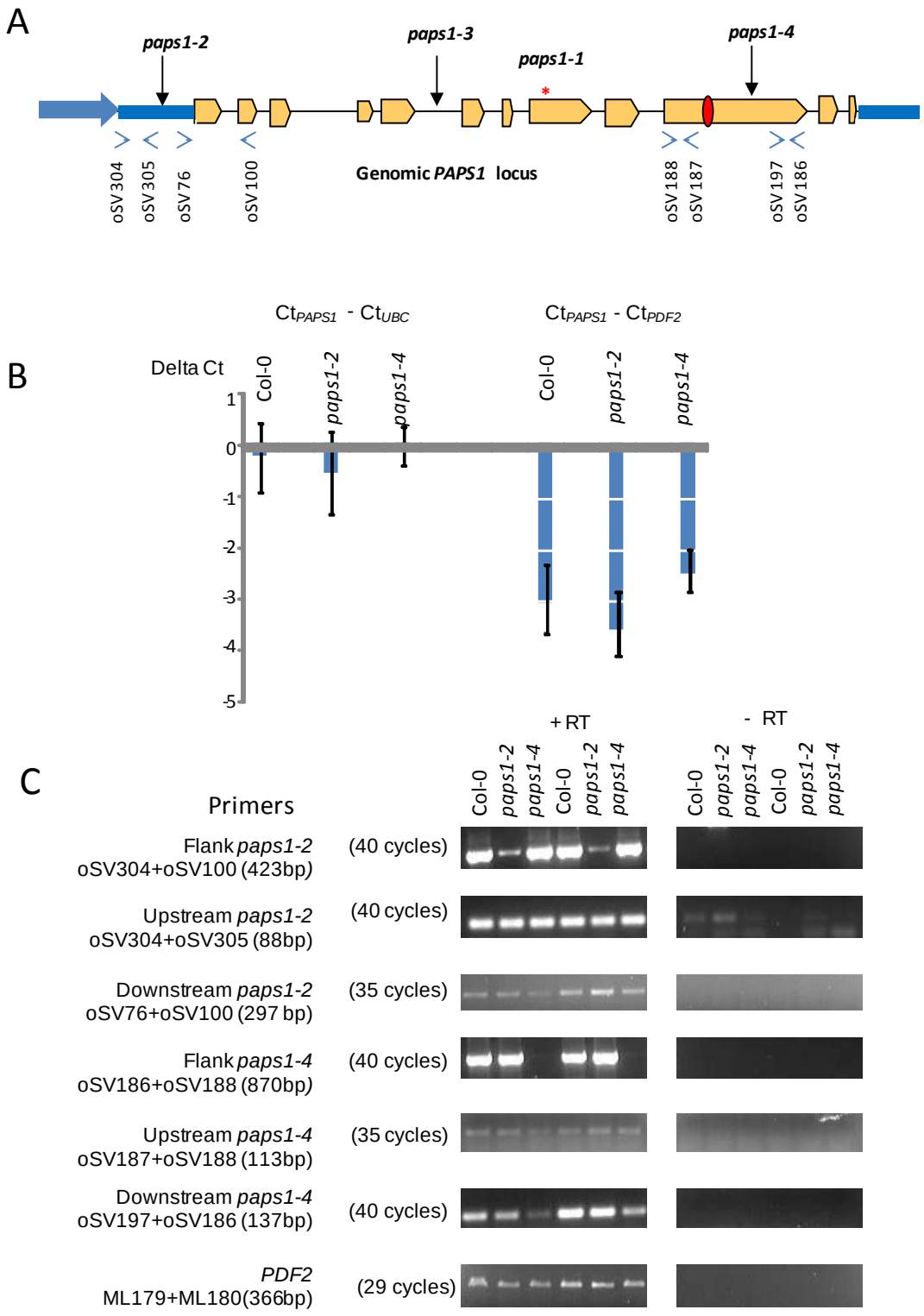


Figure 3.8 Defects of *PAPS1* transcripts in *paps1-2* and *paps1-4* mutant alleles

A. Schematic diagram of *PAPS1* locus showing position of the point mutation or *T-DNA* insertions for each alleles. The annotation is similar to **Figure 3.2**. The positions of primers for qPCR and RT-PCR in B are shown as blue arrows.

Figure 3.8 Defects of *PAPS1* transcripts in *paps1-2* and *paps1-4* mutant alleles (continued)

B. qPCR comparing abundance of PAPS1 mRNAs (using primers oSV187 and oSV188) in Wt (Col-0), *paps1-2* and *paps1-4* with two house keeping genes *PDF2* and *UBC*. Comparisions are based on the Ct values.

C. RT-PCR with different primers, positions of which relate to the insertion site of the *T-DNA* in indicated the mutant alleles.

RT-PCR analysis using primers flanking the *paps1-4* insertion site gave no PCR products, suggesting there is no intact *PAPS1* transcript in *paps1-4* (**Figure 3.8C**). RT-PCR analysis with primers downstream of the *paps1-4* insertion site showed a slight reduction in the abundance of transcript from this region of the gene (**Figure 3.8C**). However, both qPCR and RT-PCR analysis using primers upstream of the *paps1-4* insertion site suggest equal abundance of transcripts transcribed in this region of the gene in *paps1-4* compared to WT (**Figure 3.8B** and **C**). In the best scenario, this truncated transcript is able to give rise to a truncated protein. Since the insertion is after the catalytic domain, the hypothetical truncated protein is expected to retain some activity.

Based on these analysis of PAPS1 mRNA, in *paps1-4*, the PAPS1 protein is probably truncated hence activity of PAPS1 is probably reduced slightly resulting in a phenotype that resembles but is less strong than in *paps1-1*.

3.3.4 *paps1-2* mutants are allelic to *paps1-1* and *paps1-4* but cause petals to be deformed*:*

paps1-2 mutant plants has a *T-DNA* inserted in the *5' UTR* of the *PAPS1* locus. Similar to *paps1-1* and *paps1-4,* the *paps1-2* mutant plants also have smaller leaves (**Figure 3.2B**); however in contrast with the other alleles, the petals of the *paps1-2* mutant are not bigger (**Figure 3.2B** and **Figure 3.7A**). By contrast, the petals are all smaller and narrower than WT. There are usually only one or two petals per flowers in the *paps1-2* flowers instead of four petals per flowers in WT plants. The *paps1-2* flowers are female sterile, which is caused by the deformed gynecium. Pollen of *paps1-2* mutants has very reduced fertility but some were viable because F1 plant between homozygous *paps1-4* (female) plants and homozygous *paps1-2* (male) plants were obtainable. Contrastingly, the reciprocal crosses to this cross yeilded no seeds.

PAPS1 transcript was checked in *paps1-2* mutants. qPCR and RT-PCR analysis using primers upstream and downstream of the *T-DNA* insertion site showed that these two regions of the transcript is as abundant as WT (**Figure 3.8B** and **C**). However, RT-PCR analysis using the primers flanking the insertion site suggests a served reduction of the transcript in the mutant compared to WT. It is unexpected to have a PCR product, length of which is approximately the same WT (Figure 3.8C), using the primers flanking the T-DNA insertion site in *paps1-2* mutants. The presence of this PCR product are either due to PCR contamination, which I consider unlikely because the PCR experiments have been repeated twice with similar results, using two independent RNA samples. A second explanation for the presence of this PCR product is that somehow a small fraction of the transcript manages to splice out perfectly the *T-DNA*. The PCR product thefore should be sequenced in future experiment. Nevertheless, the important deduction from this RT-PCR analysis is that the full length coding region of the mRNA of *PAPS1* is likely to be intact and equally abundant in

paps1-2 compared to WT (based on the qPCR and RT-PCR with primers downstream of the insertion site). Because of the insertion in 5'UTR, this potentially intact coding region of PAPS1 can still be mis-translated for example: *T-DNA* insertion can results in a chimeric fusion of '*T-DNA-PAPS1*' transcripts that is translated in to a protein with a different reading frame. Therefore, it is not exactly clear about what could go wrong to the PAPS1 protein in *paps1-2* mutants. We are now generating antibody against PAPS1 in order to check the PAPS1 protein in *paps1-2*.

To sum up, *paps1-2* has *T-DNA* inserted in the *5' UTR*, which causes the petal to deform, a unique phenotype that is not shared by *paps1-1* and *paps1-4*. RT-PCR and qPCR suggest that the coding of region of PAPS1 mRNA is likely to be intact and the transcript is as abundant as wild-type, however as the insertion is in 5'UTR, the protein may not be translated normally. Notably, because of the phenotypic similarity between *paps1-2* flowers and *paps1-1* flowers at high temperature (28^0C), it is likely that the remaining PAPS1 activity in *paps1-2* plants is lowest amongst *paps1-1* (at 23^0C*)*, *paps1-2* and *paps1-4* plants.

3.3.5 *paps1-3* mutants are probably embryonic lethal or gametophytic lethal.

paps1-3 mutants has a *T-DNA* inserted in the intron 5 of the *PAPS1* gene (**Figure 3.2**). Genotyping 90 F2 plants from a heterozygous *paps1-3*/+ plant found 51 wild-type plants and 39 heterozygous plants and no homozygous plant. The segregation ratio (close to 1:1) suggests *paps1-3* is probably gametophytic lethal. Reciprocal crosses with Col-0 will have to be carried out to prove this.

3.3.6 The allelic relationship between *paps1* mutant alleles:

Figure 3.3 shows the petal size of homozygous *paps1* mutants and trans-heterozygous plants (i.e. F1 plants between two different *paps1* homozygous mutants, alleles of which are different). Both *paps1-2* and *paps1-4* are allelic to *paps1-1* because their trans-heterozyogous with *paps1-1* plants still have large petals (**Figure 3.3A**). The different in background (F1 of L*er* x Col-0) did cause hybrid vigor, i.e. bigger petals than either of parent, however this effect is very small compared to the increase (over the wild-type) in the petal size of the *paps1-1/paps1-4* and *paps1-1/paps1-2* trans-heterozygous.

In the same ecotype Col-0, the combination of *paps1-2* allele, which causes smaller petals than WT and *paps1-4* allele, which causes bigger petal than WT, resulted in trans-heterozygous plants with even larger petals than *paps1-4* homozygous mutant plants (**Figure 3.3A**). Using normal dominant/recessive relationship can not explain this further increase in petal size. I explain this in the general discussion in chapter 8.

In conclusion, the *ds39* mutation is in At1g17980, which encodes for poly(A) polymerase 1 (PAPS1), *ds39* is renamed *paps1-1*. In total, four *paps1* mutant alleles were isolated. *paps1-1* and *paps1-4* similarly have smaller leaves and bigger flowers than WT whereas *paps1-2* have smaller leaves but deformed and smaller petals than WT. One other allele, *paps1-3*, is embryonic or gametophytic lethal.

4 .Chapter 4. Properties of PAPS1 and the functional specialization among four PAPSs in *Arabidopsis.*

In this chapter, I show that *in vitro* PAPS1 WT protein possesses polyadenylation activity but the mutated protein in *paps1-1* (the PAPS1^{P313S}) is almost completely inactive. I showed however, *in vivo*; the bulk poly(A) tails of all the mRNAs in cells does not change dramatically in the *paps1-1* mutant. Later, I looked at the properties of all canonical PAPS in Arabidopsis with respect to: their knock out mutant phenotypes, the expression domains by using promoter:GUS constructs, their alternative splicing patterns and finally, the conditions where the balance amongst PAPSs changes. Lastly, I show evidence from transgenic plants to support that the functional specificity of different PAPSs lies with the C-terminal domains.

Together, these results suggest a novel layer of gene regulation whereby cell controls large group of genes in response to stimulus by adjusting the balance of different canonical PAPs.

4.1 *In vitro* poly(A) polymerase activity:

To explore the biochemical properties of PAPS1, the His-tag PAPS1 and His-tag PAPS1^{P313S} protein were expressed in E.coli, purified and assayed for activity *in vitro*. This experiment is done in collaboration with NishtaRao in James Manley's lab. My contribution is generating the two constructs ready for expression in E.coli.

The activity was determined in a non-specific polyadenylation assay, which measure the ability of the enzyme to add adenosines to radioactively-labelled precleaved mRNAs from SV40 virus. As shown in **Figure 4.1A**, the WT PAPS1 protein has activity, while the mutated PAPS1^{P313S} has almost no activity at 26°C, but may retain a very low level of activity at 16°C. This suggests that the mutated protein is not functional and the activity decreases at 26°C compared to 16°C.

To conclude, based on *in vitro* activity *paps1-1* seems to be a hypoactive/loss of function mutation rather than a hyperactive mutation.

4.2 *In vivo* bulk poly(A)-tail analysis:

To assess the effect of *paps1-1 in vivo*, the bulk poly(A) tail length was measured in mutant and in WT plants. Knowing the mutation is temperature sensitive, I analyzed RNA from seedlings grown at 28°C. **Figure 4.1B** shows that despite the dramatic effect on growth of seedlings at high temperature, the overall poly(A)-tail composition and patterns are almost unchanged in *paps1-1* compared to WT (**Figure 4.1B**). There could be a subtle reduction in the proportion of the longest (A) tails, but the results were not reproducible in other replicates in flowers (not shown).

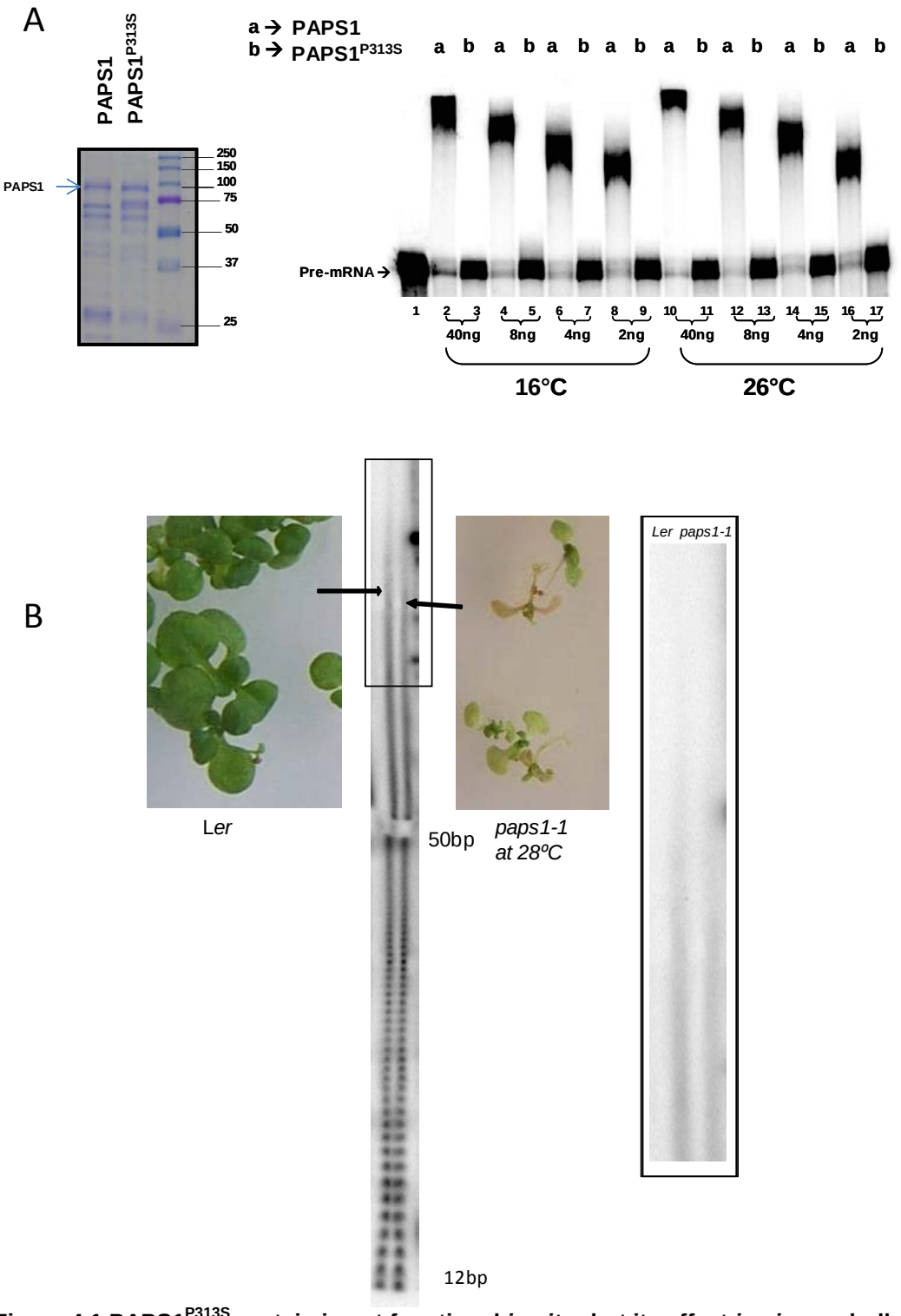


Figure 4.1.PAPS1^{P313S} protein is not functional *in vitro* but its effect *in vivo* on bulk mRNAs tail-length is subtle.

Figure 4.1. PAPS1^{P313S} protein is not functional *in vitro* but its effect *in vivo* on overall mRNAs tail-length is subtle (continued).

A. In vitro non-specific PAP assay Left: SDS-PAGE showing purified wt and mutated PAPS1 (PAPS1-^{P313S}) protein from E.coli, Right: Radiograph of RNA products resolved on agarose gel after mixing different concentration of purified PAPS1 proteins with the precleaved substrate.

B. In vivo bulk poly(A)-tail analysis. Total cellular RNAs were end-labelled with ^{32}P cordycepin then treated with RNaseT1 and RNaseA to degrade all RNAs but the poly(A) tails, which are then resolved in 8M urea 10% polyacrylamide gel.

The righter most picture is a magnified picture of the upper part of the gel (i.e., the square marked on the gel).

A yeast temperature sensitive *pap* mutant (*pap1-1*) showed a very strong reduction in poly(A) tail length after being shifted from permissive to non-permissive temperature (Proweller and Butler, 1994) . Compared to this result, clearly the effect in the *Arabidopsis paps1-1* mutant is very subtle.

Bulk poly(A)-tail analysis using RNA from inflorescences of Col-0, *paps1-2* and *paps1-4* showed similar results, i.e. the difference between mutant and WT was not detected (not shown).

4.3 Phenotypes of other *paps* single and double mutants:

To answer whether the phenotype of *paps1-1* is specific to *PAPS1* compared to other *PAPSs*, the knock out mutants, which are all *T-DNA* insertion lines, for the other three *PAPSs* were ordered from the seed stock centre and analyzed. There are several *T-DNA* insertion lines (shown in **Figure 4.2**) for *PAPS2* and *PAPS4* and *PAPS3*. The plan was first try to find null mutants for each *PAPS*. If this is not possible because some are essential, I would look for *T-DNA* insertions that inserted in the C-terminal domain or UTRs etc., hopefully to be able to obtain weak but surviving homozygous mutant plants.

Therefore, initially, for each *PAPS*, one line, *T-DNA* in which inserted in the exon of the gene and preferably in the N-terminal catalytic domain, was chosen for genotyping and phenotyping. With these criteria, three *T-DNA* insertion lines were chosen *paps2-3*, *paps4-3* and *paps3-1*, the insertions sites of which are shown in **Figure 4.2**. In the end, I was able to get plants that are homozygous for the *T-DNA* insertion in all three lines.

4.3.1 *paps3* single mutants:

Three *paps3* mutant alleles were isolated (**Figure 4.2**). *paps3-1* homozygous plants (confirmed by genomic DNA genotyping) did not show a phenotype. However, *paps3-3* and *paps3-4* homozygous mutant (confirmed by genomic DNA genotyping) clearly showed a phenotype. The overall stature, the leaves and flowers of *paps3-3* and *paps3-4* homozyogous mutant plants are smaller (**Figure 4.3**). The homozygous plants are also sterile.

The linkage of the phenotype to the homozygous state the T-DNA insertion was confirmed by analyzing approximately twenty T2 plants that were seggregating for the T-DNA. Because *paps3-3* and *paps3-4* are two independent T-DNA lines which share similar phenotype, it is unlikely that the phenotype is caused by an additional T-DNA that is unlinked to the PAPS3 locus. In the future, allelic test and transformation with *PAPS3* genomic rescue construct can be carried out to further confirm this linkage.

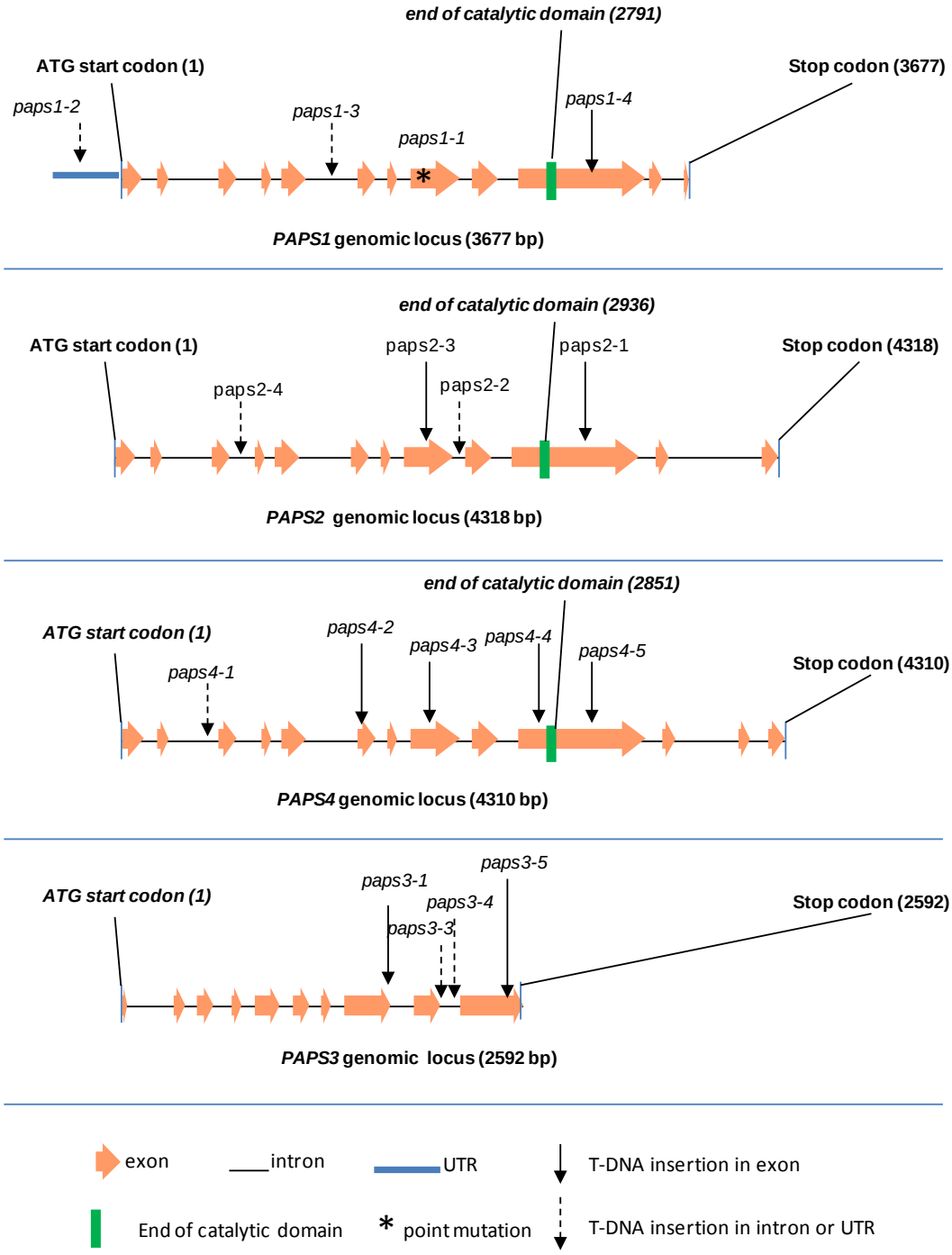


Figure 4.2. Mutant alleles and positions of *T-DNA* insertion in four PAPS genes.

The numbers in brackets indicate the nucleotide position from the 5' A (in the first ATG codon) of the gene genomic sequence.

paps3-3

paps3-4

Figure 4.3. *paps3-3* and *paps3-4* mutant showed deleterious phenotype.

Pictures of plants three weeks after germination that are segregating for the indicated *paps3 T-DNA* insertion alleles. The homozygous mutants, which were confirmed by PCR-genotyping, are marked with red arrows. Other plants are either WT or heterozygous for the *T-DNA* insertion.

paps3-1 has *T-DNA* inserted in exon while *paps3-3* and *paps3-4* has *T-DNA* inserted in the intron (**Figure 4.2**), *paps3-1* did not show a phenotype while *paps3-3* and *paps3-4* did. This is quite unexpected. The insertion position is based on the published sequence of the insertion positions in these lines, may be *paps3-1* mutant has a T-DNA inserted in an intron instead of an exon. RT-PCR needs to be done to check for mRNA of *PAPS3* in these three homozygous mutant alleles.

4.3.2 *paps2* and *paps4* single mutants:

The more interesting PAPS to us are PAPS2 and especially PAPS4 because they are more similar to PAPS1. I focused on *paps2-3* and *paps4-3* alleles. Molecularly, there is no full length/intact mRNA of the PAPS in its corresponding mutant (**Figure 4.4A**). Because the insertion is in the middle exon of the gene, it is very unlikely that the hypothetical fusion protein(s), which is created by translational fusion of *T-DNA* sequence to *PAPS* sequence upstream or downstream of the insertion site, are active. Therefore, *paps4-3* and *paps2-3* are probably null mutants.

Surprisingly, homozygous *paps2-3* and *paps4-3* plants are indistinguishable from wild-type in our standard growing condition. Their petals are not larger. A closer examination showed that *paps4-3* is late flowering (**Figure 4.4C** and **D**). I also checked the phenotype of the segregating families for other alleles of PAPS2 and PAPS4 (**Figure 4.2**). At least twenty plants were sown out per family, 25% of which should be homozyogous for the *T-DNA* insertion. However, all of these plants were, morphologically indistinguishable from WT.

4.3.1 *paps2 paps4* double mutants:

Since *PAPS2* and *PAPS4* are probably redundant, the double mutants were generated to reveal the phenotype. Double mutants were found but apart from being late flowering (even later than *paps4-3*), they showed no additional phenotype at our standard growing condition (**Figure 4.4C** and **D**). This data indicates indeed *PAPS2* and *PAPS4* are partially redundant at least to the flowering time phenotype. However, the petal size and leave size and overall growth of the double mutant are normal.

This suggests the role among *PAPS2*, *PAPS4* and *PAPS1* such that at our standard growth condition, *PAPS1* appears to be the only essential *PAPS*. When *PAPS1* is wild-type, *PAPS2* and *PAPS4* are both dispensable.

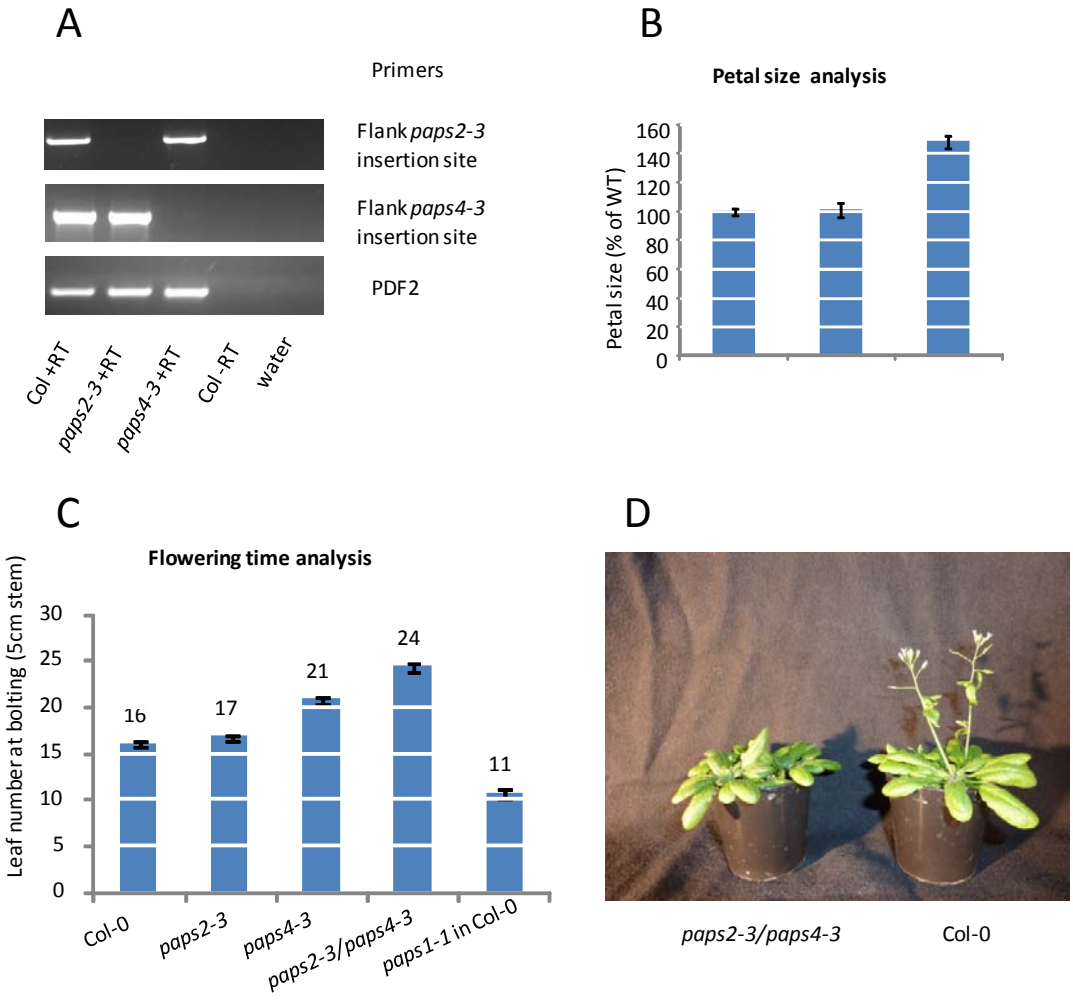


Figure 4.4. Phenotype of *paps2-3* and *paps4*-3 single mutants and *paps2-3 paps4-3* double mutant

A. RT-PCR on cDNA made from WT (Col-0), *paps2-3* and *paps4-3* single mutant. Primers are indicated. PDF2 is a house keeping gene used for control.

B. Petal size of WT (Col-0), *paps1-1* and *paps2-3/paps4-3* double mutants . *paps1-1* in Col-0 is the paps1-1 mutant, originally in Ler background, which was backcrossed three times to Col-0.

C. Flowering time (the number of leaves the plants have made at the time the inflorescences stem is 5cm long) of WT and mutants. These measurements were done by H.C.

D. Picture of WT and the *paps2-3/paps4-3* double mutant. Both genotypes were sown out on at the same time, and picture was taken 43 days after sowing.

4.4 Spatial expression pattern of *PAPSs*:

To visualize *PAPS* expression domains, the promoters of the four *PAPSs* were fused upstream of the *GUS* reporter gene. At least 10 independent transgenic lines for each constructs were analyzed to account for the variation in the insertion positions of the transgene. **Figure 4.5** showed the representing GUS-stained image of transgenic plants carrying the reporter constructs in WT background. In inflorescences, p*PAPS2::GUS* and *PAPS4::GUS* express strongly and overlappingly, *pPAPS1::GUS* did not show any detectable *GUS* staining, p*PAPS3::GUS* showed a weak staining. p*PAPS3::GUS* showed strong expression in in the trichome. In seedlings, the expressions are being analyzed on T2 plants as this thesis is being written so the results are not yet available.

4.5 Functional specializations of PAPSs:

4.5.1 Promoter is not likely to account for the functional differences between *PAPS1* and *PAPS4* gene.

As in section 4.4, the spatial *PAPS1* expression is different from *PAPS4*, but the temporal and dynamic of this regulation is difficult to determine by the *promoter::GUS* analysis. To directly test the role of promoters, we ask if *PAPS4* can rescue *paps1-1* if expressed by PAPS1 promoter. **Figure 4.7** and **Figure 4.8** shows that phenotype of neither leaves nor flower of *paps1-1* mutant was rescued in any of the T2 transgenic lines carrying the *pPAPS1::genomicPAPS4* transgene (the construct is illustrated in **Figure 4.6**) . This result suggests promoter is not likely to account for the functional differences between *PAPS1* and *PAPS4* gene.

4.5.2 The C-terminal domains are likely to be the specificity determinants between PAPS1 and PAPS4:

As *pPAP1::genomicPAPS4* did not rescue, the protein sequence and especially the very divergent C terminal domains sequence may encodes specificity amongst PAPS. To test this possibility, I modified the *pPAP1::genomicPAPS4* construct, replacing PAPS4 C terminal domain with PAPS1 C-terminal domain (*pPAP1::gNPAPS4::gCPAPS1*_construct shown in **Figure 4.6** and transformed this construct to *paps1-1* mutants. Interestingly, in contrast to *pPAP1::*genomicPAPS4, some of the T2 transgenic lines carrying *pPAP1::gNPAPS4::gCPAPS1* did rescue, though partially, the petal size phenotype (**Figure 4.5**). More over, all of the lines have bigger leaves than *paps1-1* (**Figure 4.6**). These results

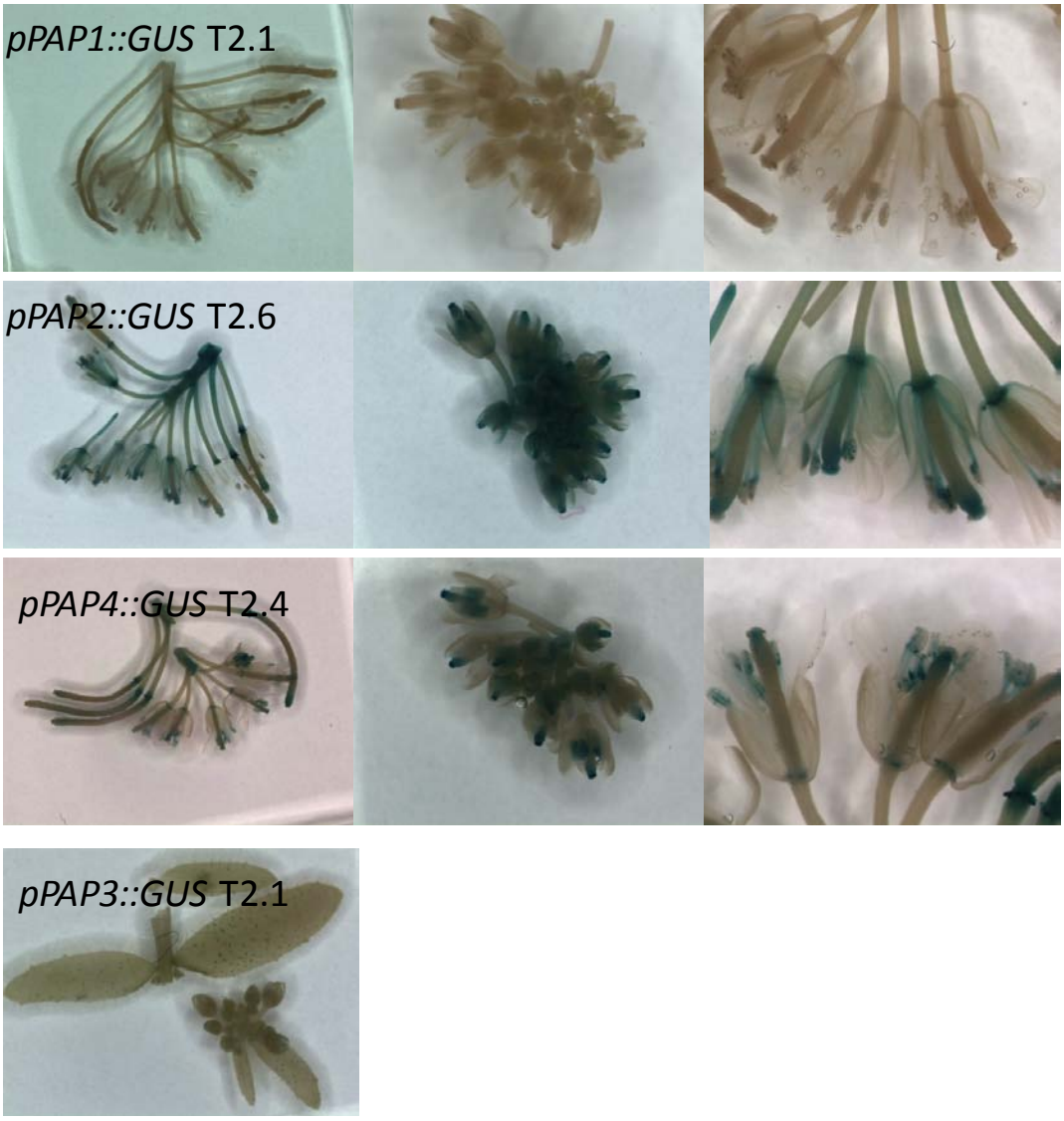


Figure 4.5. *PAPS* expression domain overlaps each other.

GUS stain pictures of transgenic plants expressing different *PAPS* promoter driving *GUS*. Each row are from one transgenic line indicated in the outer left pannel. From left to right of each row are pictures showing inflorescences with siliques, younger buds and zoomed in floral structures respectively.

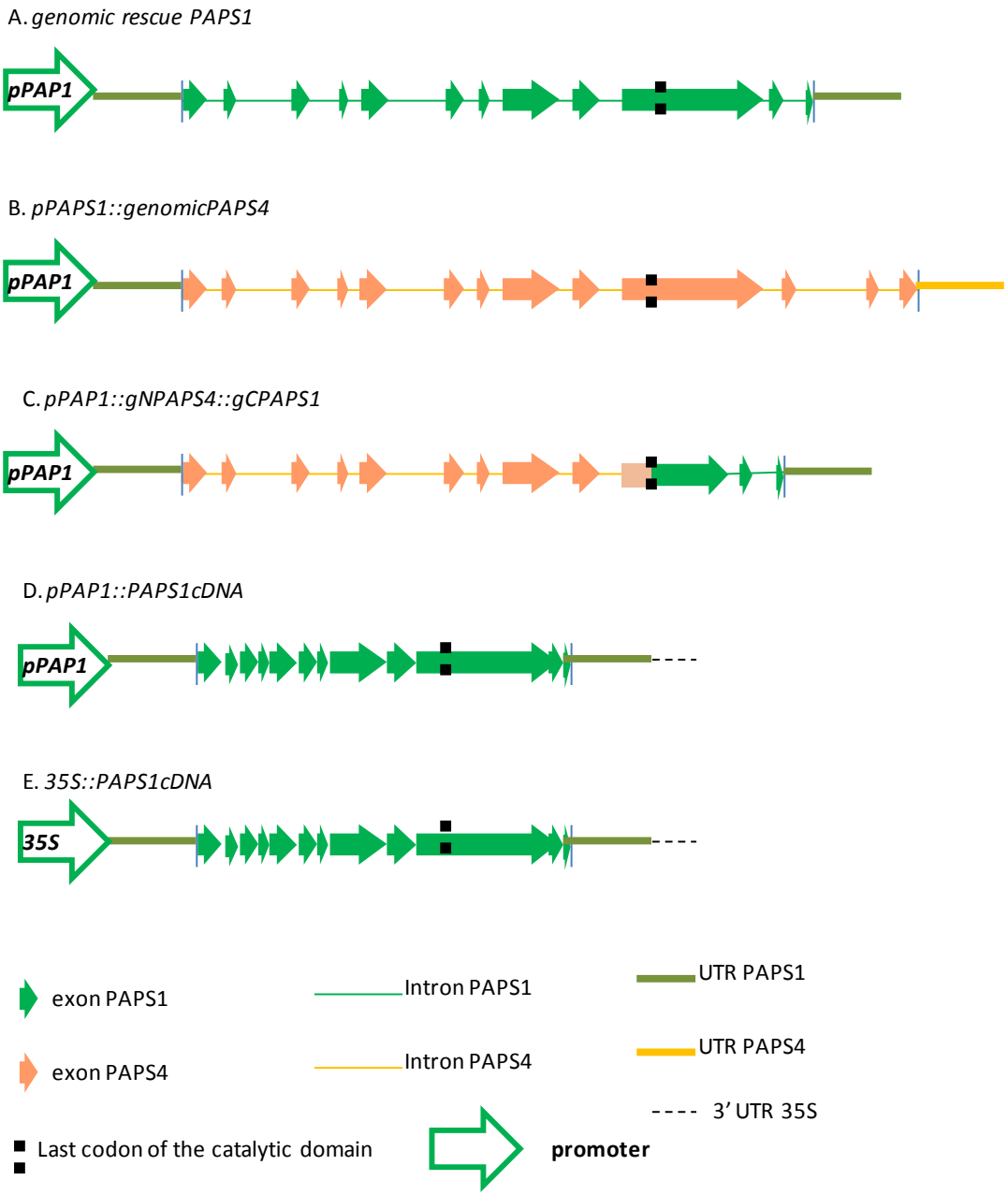


Figure 4.6. Constructs used to transform paps1-1 to check for rescued phenotype.

35S is a constitutive promoter from Cauliflower mosaic virus *35S* gene.

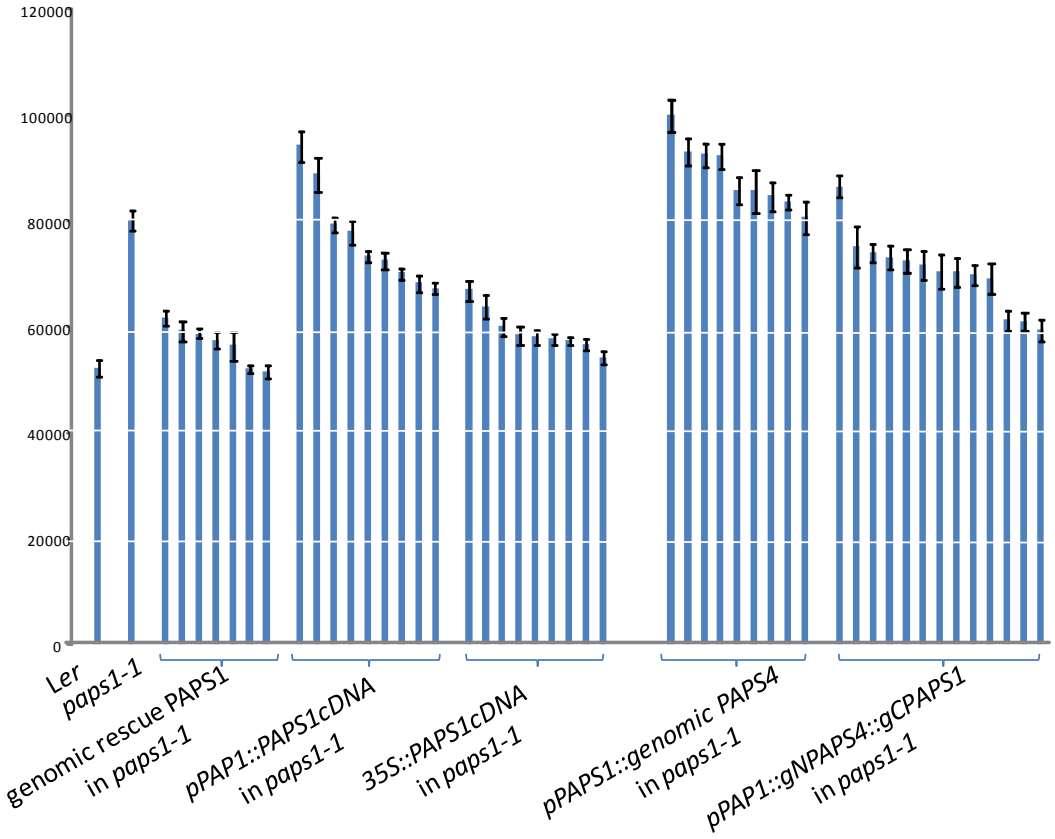


Figure 4.7. Petal size of different transgenic plants compared to *paps1-1* and WT.

In case of transgenic plants, bars that are grouped by a underneath bracket represent measurements from plants carrying the same construct indicated. Each bar represent an independent T2 transgenic family. At least 4 plants, which contain the transgene, were measured per family. This gives an average value for that family. Descriptions of the constructs are shown in **Figure 4.6.**

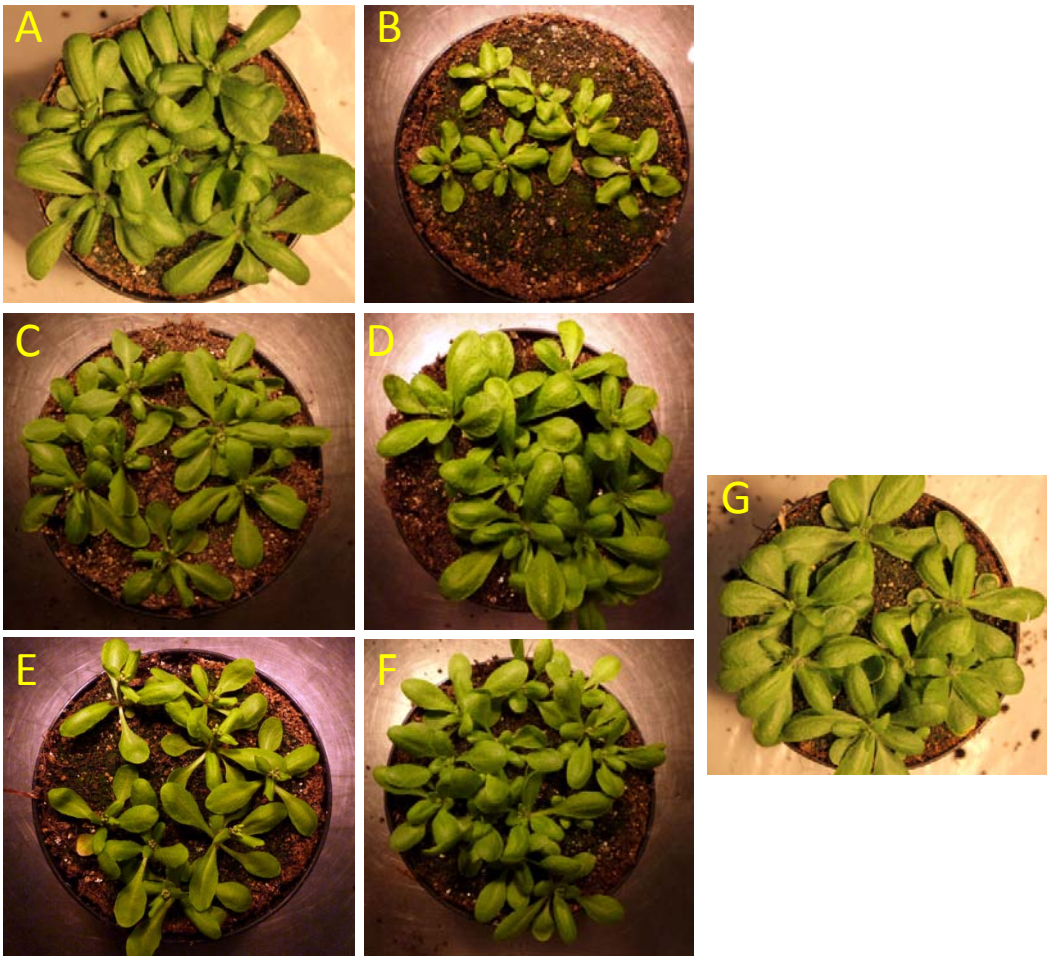


Figure 4.8 Leaf phenotype of WT, paps1-1 and different transgenic lines.

A. L*er* (WT)

B. *paps1-1*

From C-G: Basta selected T2 family plants, which are *paps1-1* mutant carrying one of the following transgene: (Details information about the constructs are described in **Figure 4.6**)

C. *pPAP1::PAPS1cDNA*

D. *35S::PAPS1cDNA*

E. *pPAPS1::genomic PAPS4*

F. *pPAP1::gNPAPS4::gCPAPS1*

G. *genomic rescue PAPS1*

showed that swaping the both the C-terminal domain and the promoter can largely (but not completely) make PAPS4 equivalent to PAPS1.

4.6 Conditions that change the balance among PAPS transcripts.

We hypothesize that mRNAs has different preference for different PAPSs, which perhaps give different A-tail length to its substrate. If this is used as an active mechanism to adjust the poly(A) tails of some transcripts, cell must actively control these preferences. The mechanism of this fine tune can be at multiple levels.

For example, control the abundance of a PAPS through transcription and translation control or through controlling its activity (e.g. phosphorylation inhibits the yeast PAP during the cell cycle (Colgan et al., 1998) or through controlling its recruitment to mRNA substrates.

One possibility, which we can easily check, is whether cells can change the balance in the mRNA abundance of different PAPSs. We therefore used bio-informatics to find which physiology conditions/genotypes/treatments change the balance of the mRNA abundance among *PAPS1*, *PAPS2* and *PAPS4*. Using GeneInvestigator, a public database of microarray experiments in many different conditions, we discovered that *PAPS1*, but not other two *PAPS*, is up-regulated when treating plants with FLG22 peptide (**Figure 4.9**). FLG22 is a 22 aa synthetic peptide derived from flagellin; the FLG22 treamentmimics a bacterial attack to plants. On the opposite direction of regulation, *PAPS1*, but not other two *PAPS*s, is down-regulated in plants that are treated with cold.

4.7 Overexpressing *PAPS1* did not results in any morphological changes:

paps1-1 mutants carrying the overexpression *35S::PAPS1cDNA* construct are indistinguishable from wild-type with respects to plant statures, leaf size and petal size (**Figure 4.7** and **Figure 4.8**). The transgenic plants showed no other visible morphological changes.

4.8 Discussion:

The bulk poly(A) tail analysis results indicate that in *paps1-1* mutants and other *paps1* mutants carrying *paps1-4* and *paps1-2* alleles, only a small number of transcripts were affected. This result again lends credit to the hypothesis that there is only a small proportion of the transcriptome that appears to be sensitive to the *paps1-1* defect.

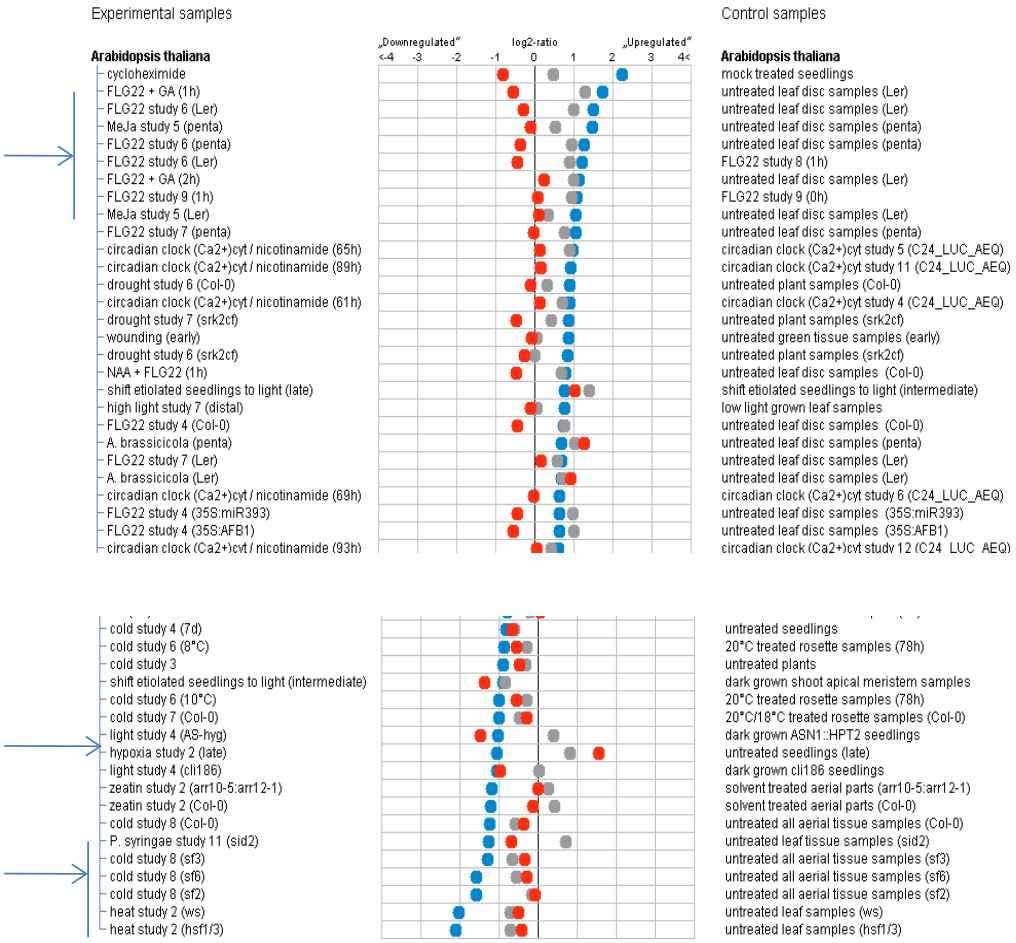


Figure 4.9 Conditions where *PAPS1* is up/down regulated while other *PAPS*s remains largely unchanged, or change in the reverse direction.

Interesting conditions are highlighted with arrow.

blue dot is *PAPS1*,

gray dot is *PAPS2*,

red dot is *PAPS4*

The fact that *PAPS3* is pollen specific explains the fertility problem in homozygous *paps3-3* and *paps3-4*. Mammals also have a canonical PAP that lacks C-terminal domain, which is testis specific (PAPLOB). *Arabidopsis* PAPS3 may be the PAPLOB equivalent in plants. Because the petals of *paps3-3* and *paps3-4* are not bigger, this mutant is of little immediate interest to us. Nevertheless, the important conclusion from this results is that it suggests that PAPS3 and PAPS1 are functionally non-redundant.

Judged by the expression domains, which are based on the pPAPS::GUS experiements, *PAPS1* is not the mainly expressed *PAPS* in inflorescences. However, some consideration must be taken in to account when analyzing the expression domains using *promoter::GUS* reporter gene. For examples, in contrast to the results from *promoter::GUS* experiments, microarray data from AtGenExpress, and previous study (Addepalli et al., 2004) using Northern blot indicates that *PAPS1*, *PAPS2* and *PAPS4* expression levels are similar in floral organs (**Figure 1.13**). This indicates that the staining patterns may not truly reflect the expression dynamic of PAPSs (including spatial, temporal and mRNA stability).It is also possible that there are other transcriptional regulatory elements outside of the promoter fragments that was used in driving GUS expression in these constructs. Did I capture the full upstream regulatory sequence of PAPS1? Probably, yes because: The *pPAPS1* promoter sequence used in *pPAPS1::GUS* construct is as long as the fragment that is present upstream of the ATG in the genomic rescue construct, which fully rescued the *paps1-1* phenotype. This suggests that the missing elements lie downstream of the ATG in *PAPS1* genomic locus. One can test this by either making translational fusion of *PAPS1* genomic to a reporter protein or directly carry out *in situ* hybridization probing for *PAPS1* mRNA. Nevertheless, results on Northern blot (Addepalli et al., 2004) and AtGeneExpress clearly indicates that *PAPS1/2/4* transcripts are equally abundant in inflorescences. This suggests that expression patterns alone can not be the reason for the differences in phenotypes in flowers of different *paps* mutants.

Knowing about the complication with *pPAPS1* promoter explained above, I reconsider the interpretation of the promoter swap *pPAPS1::genomicPAPS4* result. The question is: did *pPAPS1::genomicPAPS4* not rescue because of the PAPS4 protein could not replace PAPS1 protein or because I failed to reconstitute *PAPS1* expression dynamic in this construct. I believe the later is unlikely because of the results of the followed construct *pPAP1::gNPAPS4::gCPAPS1.* The very same promoter in *pPAPS1:: genomicPAPS4* was used in the *pPAP1::gNPAPS4::gCPAPS1* construct, which did show rescued phenotype, meaning that the promoter combined with genomic construct did work well. In the future experiments, one should check the expression of the PAPS4 transgene in *pPAP1::genomicPAPS4*.

One thing, though unlikely, that I reserve for the interpretation of the *pPAP1::gNPAPS4::gCPAPS1* construct is whether the protein sequence of C-terminal domain (CTD) of PAPS1 protein or the 3'UTR of the *PAPS1* gene, or both are responsible

for the rescue? Both CTD and 3' UTR were changed in creating *pPAP1::gNPAPS4::gCPAPS1* derived from *pPAP1::genomicPAPS4* (**Figure 4.6**). Nevertheless, what I can be sure from this experiment is the promoter plus the sequence starting from C terminal domain to the end of the gene are sufficient to rescue the *paps1-1* phenotype in petals.

For future experiments, a way to avoid the complications about *pPAPS1* promoters and *PAPS1-3' UTR* and focusing only on protein domain is to use *35S* promoter and *35S* terminator. This is because *35S::cDNAPAPS1::35S terminator* fully rescued the *paps1-1* without any additional morphological phenotype (**Figure 4.7, Figure 4.8)**. Next step would be to create *35S::genomicPAPS4* or *35S::cDNAPAP4* and transform *paps1-1* to check for rescued petal size.This construct will also answer whether supplying *paps1-1* with ample polyadenylation activity from a different *PAPS*, *PAPS4* is able to rescue the phenotype of the mutants. This experiment will clarify whether the total activity of *PAPS* in general or the specific activity of individual PAPS underlies *paps1-1* phenotype.

The identified conditions, where the balance of PAPSs expression changes, are fascinating because it shows that each *PAPS* is under active gene regulations. Hence, I speculate that there will be more conditions that cells can finetune *PAPS* expression to regulate poly(A) tail length of specific pools of transcripts in order to response to those conditions. It will even be more interesting if we can find some target genes that are changed in *paps1-1* and are also changed in WT which is treated with these conditions. Additionally, we can check whether subjecting WT plants to these conditions can induce the change in poly(A) tail lengths of the target genes. Several *PAPS1* regulated genes were identified and described in chapter 7.

Taken together, the data suggests the functional specializations of canonical PAPSs and the C-terminal domain of the protein is likely to be the specificity determinant.

5 Chapter 5. Analysis of the cell autonomy of *PAPS1*

5.1 The idea:

5.1.1 The concept of non-autonomous/autonomous effect:

An important question to be clarified in studying the function of a developmental regulatory gene is whether the gene acts autonomously or non-autonomously. This basic question can be answered by studying chimeric plants with genetically mutant and wild-type parts. If the cellular or organ-wide phenotype solely depends on the genotype of the cells or the organ, the effect is autonomous. Conversely, if the wild-type parts can rescue the phenotype of the neighboring mutant parts, or the mutant parts influence the phenotype of the wild-type parts, the effect is non-autonomous.

There are several examples of non-autonomous effects in developmental processes such as: florigen *FT (*Corbesier et al*., 2007)* in flowering time regulation, *WUS/CLV3* in stem-cell maintenance. (Schoof et al., 2000).

In the case of growth, this question is highly pertinent because growth must be co-ordinated amongst different parts of the organ and different organs in the organism in order to form the correct shape and size. Growth co-ordination can happen over a range of distance: for example between 'source' organs (photosynthetically active leaves) and 'sink 'organs (flowers), amongst petals in a flower to ensure species-specific floral symmetry or between anthers and gynoecium to ensure the physical architecture that allows self-pollination.

Several studies have shown that growth is under the control of both autonomously or non-autonomously acting factors (Galloni and Edgar, 1999). In plants, a growth promoting factor *KLUH* acts non-autonomously to co-ordinate growth of different organs in one flower, or different flowers within an individual inflorescence (Eriksson et al., 2010).

Another question involving growth co-ordination regards the contribution of the epidermal layer and internal layer in driving/restricting organ growth in plants (Savaldi-Goldstein et al., 2007). What happens if the growth of two layers is uncoupled: Can the growth of one layer drive the growth of the other layer (a cell non-autonomous effect)?

The specific question: non-autonomous/autonomous effect across a range of distance.

Here, I want to address the question of a possible non-autonomous aspect of *PAPS1* function across several scales. There are six levels that I will consider (**Figure 5.1**):

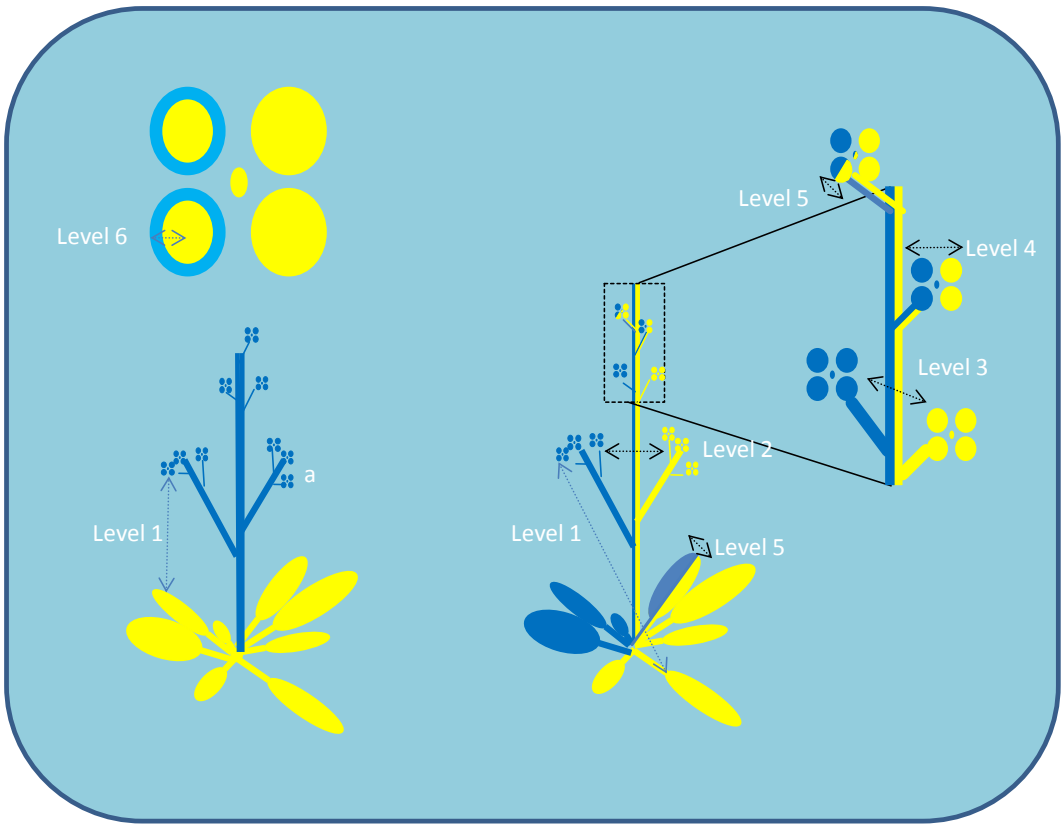


Figure 5.1 Sectored plants allow the analysis of growth co-ordination across different distances

Level 1: Two different organs in individual plants: leaves and flowers.

Level 2: Petals from different side inflorescences in individual plants.

Level 3: Petals from different flowers in one individual inflorescence.

Level 4: Petals within individual flowers.

Level 5: Two halves of individual petals or of leaves.

Level 6: Layers of individual petals: epidermis (L1), illustrated as the border of the circle, and internal (L2) layers.

Yellow color represents wild-type tissues. Blue color represents mutant tissues.

In each case, the specific question when examining the levels 2-6 is: In case of petals, is the size of the genotypically *paps1-1* mutant part increased compared to its wild-type counterpart? The logic would be reversed for leaves.

5.2 The method

5.2.1 A system for generating predictable chimeras:

The investigation of non-autonomous effects requires the generation of chimeric plants. To make sure that any non-autonomous effects on growth will be captured, one needs to consider at least two issues. The first issue is temporal. In my case, the phenotype of interest is growth, which is a continuous process starting from organ initiation until the fully mature organ reaches its final size. The non-autonomous effect may influence growth at any moment during this process. Therefore, if the chimeric situation is not established during the particular time window of gene action, the non-autonomous effect can be missed. The second issue is spatial, i.e. the range of communication that determines how close the two sectors of different genotype need to be in order for the non-autonomous effect to take place. These two issues combined require a temporal and spatial control over the induction of chimeras.

A method has been developed in the Lenhard lab to generate predictable chimeras (Adamski et al., 2009, Eriksson et al., 2010). The method has four components (illustrated in **Figure 5.2**). The first is a *Cre/loxP* system that ensures a non-reversible switch in genotype. Cre is an enzyme that specifically recognizes and recombines two directly repeated *loxP* sequences in such a way that after recombination, the sequence between the two *loxP* sites is circularized and excised from the original DNA molecule. This is used to excise a *loxP*-flanked genomic rescue fragment for the gene of interest from the genome of a corresponding homozygous mutant, thus uncovering the mutation. As the circularized fragment is not replicated or regularly distributed during mitosis anymore, essentially all daughters of a recombined cell will have the homozygous mutant genotype. The second component of the system is the ethanol inducible AlcR-*AlcA* system that allows the temporal induction of *Cre*, and hence allows the timing of induction of the chimera to be controlled (Roslan et al., 2001). In the set-up outlined above, plants will form wild-type organs in the absence of ethanol-induction, and only after induction will recombination occur and mutant organs will be formed. In the AlcR-*AlcA* system, the fungal transcription factor AlcR is activated only in the presence of ethanol; it then binds to the *cis*-element *AlcA* and transcribes the gene downstream of *AlcA*, in my case *Cre*. Another advantageous property of the ethanol inducible system is that it can be induced incompletely. Hence, after a mild induction some cells will express *Cre*, some others will not. The third component is a promoter of choice to control the expression of *AlcR*. These promoters add additional layers

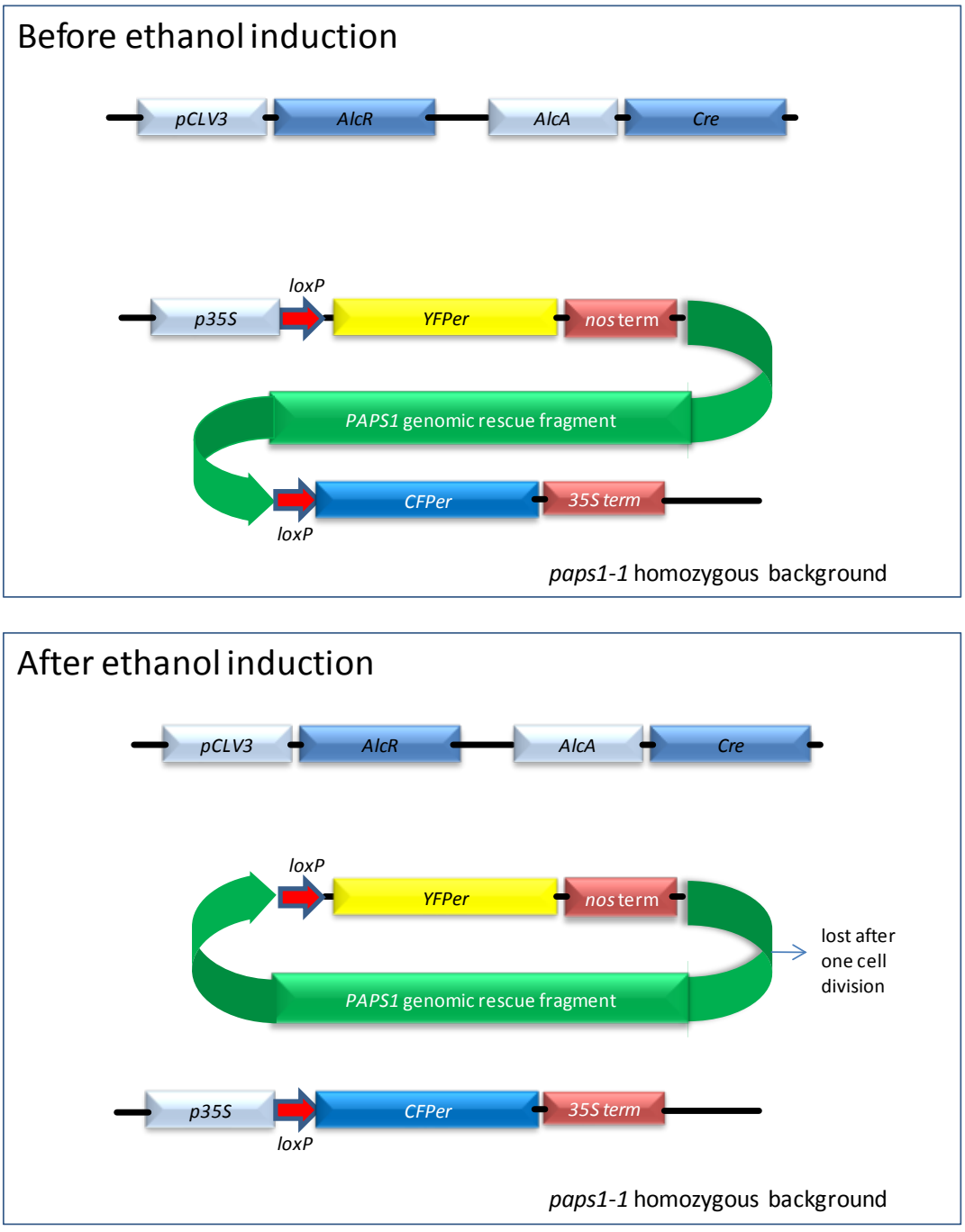


Figure 5.2 The system to generate predictable chimeras. See text for description.

term stands for terminator.

of control over the recombination including spatial and developmental control (see below). Fourthly, a dual-colour fluorescence system is used: Cyan Fluorescent Protein (CFP) and Yellow Fluorescent Protein (YFP) label the cells of the opposite genotypes as specified by the design of the *loxP* construct (**Figure 5.2**). The recombined/mutant sectors will be marked with CFP (blue) and the non-recombined/wild-type sectors will be marked with YFP (yellow). The net result of these four components is a temporally and spatially controllable induction of chimeras, the genotypes of which are tractable using a fluorescent microscope. The third component of the system provides flexible controls over recombination. By using a different promoter driving *AlcR*, we can choose when and where chimerism is induced. For example, to produce chimerism long before organs initiate, *pCLV3::AlcR-AlcA::Cre* (from now on abbreviated as *CLV3>>Cre*), which restricts the *Cre* expression domain to stem cells, can be used. *CLV3* is specifically expressed in the stem cells and immediate daughters of stem cells (Fletcher et al., 1999). To produce chimerism at or after organ initiation of petals and sepals, a petal/sepal-specific promoter *AP3* or *AP1* (Weigel and Meyerowitz, 1994).can be used (abbreviated as *AP1>>Cre* or *AP3>>Cre*) For the specific purposes of this project, two other promoter sources can be used to spatially regulate *Cre* expression. The first is *pKLU>>Cre*, in which *Cre* should be expressed only in the peripheral cells and not the meristem (Zondlo and Irish, 1999) (*pKLU>>Cre)*, therefore, should give rise to mutant rosettes and wild-type inflorescences, which is the reverse effect of *pCLV3>>Cre*. The other promoter is *pAtML1*, in which *Cre* should only be expressed in the epidermis L1 layer (Sessions et al., 1999).

5.2.2 How to analyze of growth of different sectors:

One of the aims I set out to investigate was to analyzethe co-ordination between different parts of petals or of leaves (level 5), and for this one needs to compare the size of different sectors of an organ. Chimeric organs can have sectors of different sizes, depending on the number of precursor cells that had one or the other genotype. This is true even when the sectors do not differ in their growth behaviour. How then do we know which sectors to compare with one another to be able to learn about the effect of the *paps1-1* mutation? To answer this, we need to understand how the sectors within individual organs are formed. To understand how sectors are formed, we need to consider the process of organ initiation in the *Arabidopsis* shoot meristem, and the concept of organ anlagen and of clonal analysis.

5.2.2.1 *Arabidopsis* shoot meristem structure

An *Arabidopsis* shoot meristem has three clonally distinct layers, L1, L2 and L3 (**Figure 5.3A**). In the L1, cells divide anticlinally to make the entire epidermis.The L2 layer, in which cells also divide anticlinally, and L3, in which cells divide in both anti- and periclinally, together form internal tissues, including the vasculature. Because of the strictly anticlinal divisions, cells in the L1 layer and L2 layer do not mix with each other or mix with cells from L3 layers. Therefore, three sets of stem cells are required.

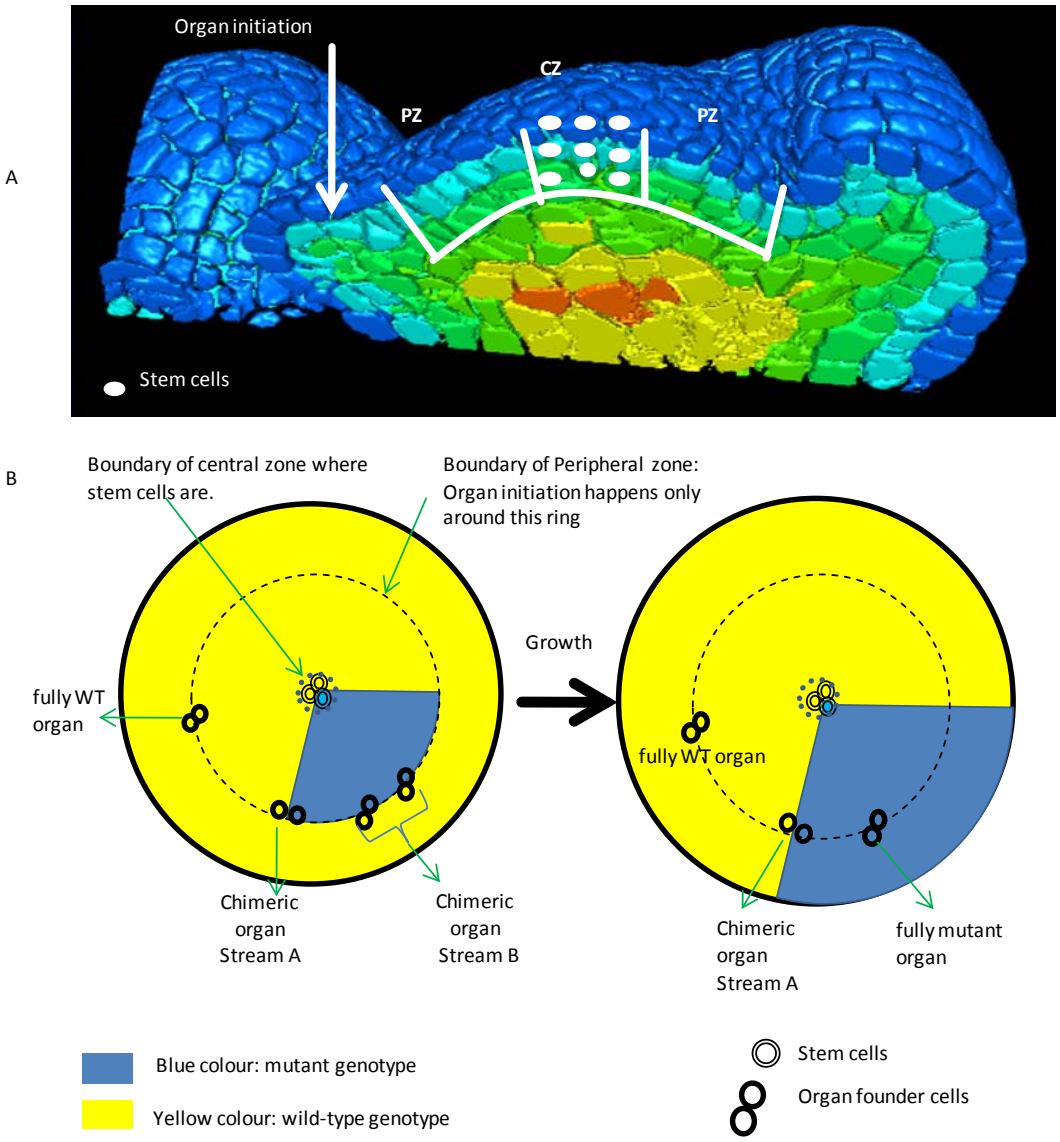


Figure 5.3 *Arabidopsis* meristem structure and the formation of chimeric organs

A. Structure of the *Arabidopsis* shoot meristem. CZ: central zone, PZ: peripheral zone. The image is modified from raweb.inria.fr/

B. Two ways to form chimeric organs (see text for more details).

Shoot apical meristem view from the top. Only the L1 layer is illustrated. There are three stem cells in the L1. Organ anlagen containing two founder cells in the L1 are illustrated. The two possible scenarios to form chimeric organs, stream A and stream B, are shown. Stream A: chimeric organs are formed at the boundary of the two sectors that are derived from two stem cells; stream B: chimeric organs are formed when the boundary between recent stem-cell daughter cells formed after recombination and the older daughters formed before recombination is displaced into the organ formation zone. When the meristem grows further, only stream A can create chimeric organs.

It is estimated that there are about three stem cells in each layer (Steward and Dermen, 1970). Each stem cell divides slowly and asymmetrically to give rise to two immediate daughters of stem cells. Depending on their position within the meristem, the daughters will either retain the stem-cell fate and take the mother cell's place, or they will enter differentiation. Ongoing cell division by the centrally located stem cells will push earlier stem-cell daughters towards the peripheral zone, where they will be incorporated into organ anlagen to make up organs.

5.2.2.2 Organ anlagen

The very early organ primordium at the time of its specification is called the organ anlage. The cells that make up the organ anlage are referred to as founder cells. Different organs have different number of founder cells. Each founder cell in one layer forms a sector in the mature organ. Thus, by determining how many "fractions" an organ can be subdivided by sectors, one can estimate the number of founder cells as the inverse of this fraction. If for example marked sectors can make up either ¼, ½, or ¾ of an organ's epidermis, the most likely number of epidermal founder cells for the organ is four. For this to be valid, however, the sectors must originate outside of the organ, so that one can be sure that one captured the founder cells at the earliest point when they are committing their fate.

5.2.2.3 Clonal sectors

If some of these founder cells are mutant, the resulting sectors in the mature organs will also be mutant. Depending on the number and the arrangement of the mutant founder cells in the organ anlage, the sector sizes and patterns can vary. The more mutant founder cells there are, the larger will be the mutant sector.

Now, we can answer the question of which sectors need to be compared to assess the effect of the *paps1-1* mutation. We must compare sectors that are derived from the same number of founder cells. Generally, it is not possible to trace back the number of founder cells that gave rise to a sector, when its size can be affected by the number of founder cells and a potential difference in the growth of cells of the two genotypes (wild-type or *paps1-1*). There are fortunately two exceptions. One case is when the organ anlage has only two founder cells in one layer. There can then be only one kind of chimeric organ: it consists of only two sectors, each derived from one founder cell. In the other case, the organ anlage has more than two founder cells, but there are also two sectors, the boundary of which coincides with the midvein. The midvein is assumed to always divide an organ into two halves with equal number of founder cells. From the studies (Bossinger and Smyth, 1996), the number of founder cells in one layer is estimated to be eight for leaves, eight for sepals, eight for gynoecium, four for anthers and two for petals.

Therefore, to analyze the growth co-ordination between sectors within one organ, I aimed to analyze chimeric petals and leaves with the sector boundary running along the midvein.

5.2.2.4 How to generate chimeric petals and leaves with a sector boundary along the midvein?

Having considered the structure and the dynamics of the stem cell population and the formation of the organ in the meristem, it is inferred that there are two ways to form chimeric organs using the*CLV3>>Cre* system. Firstly, if one uses a mild ethanol treatment, sometimes only one of the stem cells recombines. As the stem cells divide, they eventually give rise to two distinct clonal populations, one of which contains all mutant cells and the other contains all wild-type cells. If at the border of these two populations an organ is formed, it will be chimeric. This is the main stream from which chimeric organs arise (denoted as stream A in **Figure 5.3B**). However this is not the only possibility. Another way(denoted as stream B) to create a chimeric organ is when the boundary between the stem cells daughters that were formed after the focal stem cell underwent recombination and the daughters that were formed before comes to lie within the region of organ initiation in the meristem periphery, giving rise to a chimeric organ (**Figure 5.3.B**). Unlike the stream A type, which can operate during the life time of the chimeric meristem, the stream B-type formation of chimeric organs happens in a very restricted time window. This is because after recombination, the organ initiation zone will soon be filled with only the newly made descendants of stem cells. In contrast to stream A, stream B can also happen even if all three stem cells are recombined i.e. a homogenous stem cell population.

In the approach presented here I therefore applied only mild ethanol induction on the *paps1-1* mutants carrying the *CLV3>>Cre* and one copy of the *loxP::*genomic rescue *PAPS1::loxP* (termed *floxPAPS1*)construct in order to obtain sectored meristems, and as a result sectored petals.

5.3 The results

In total, 13 F1 plants from a cross of a *paps1-1* mutant that is homozygous for the *CLV3>>Cre* transgene and a *paps1-1* that is homozygous for the *floxPAPS1* transgene were induced with ethanol at 7 days after sowing and then analyzed. **Figure 5.4** shows a typical sectored rosette. Most of the rosette is wild-type, though some younger leaves have recombined sectors. These chimeric rosettes later produced several kinds of inflorescences. These inflorescences can be classified by two criteria: firstly, whether they are sectored or non-sectored and secondly, whether the recombination happens in the epidermis, in the internal layers or in all layers. All possible classes of inflorescences are listed in **Box 5.1**.

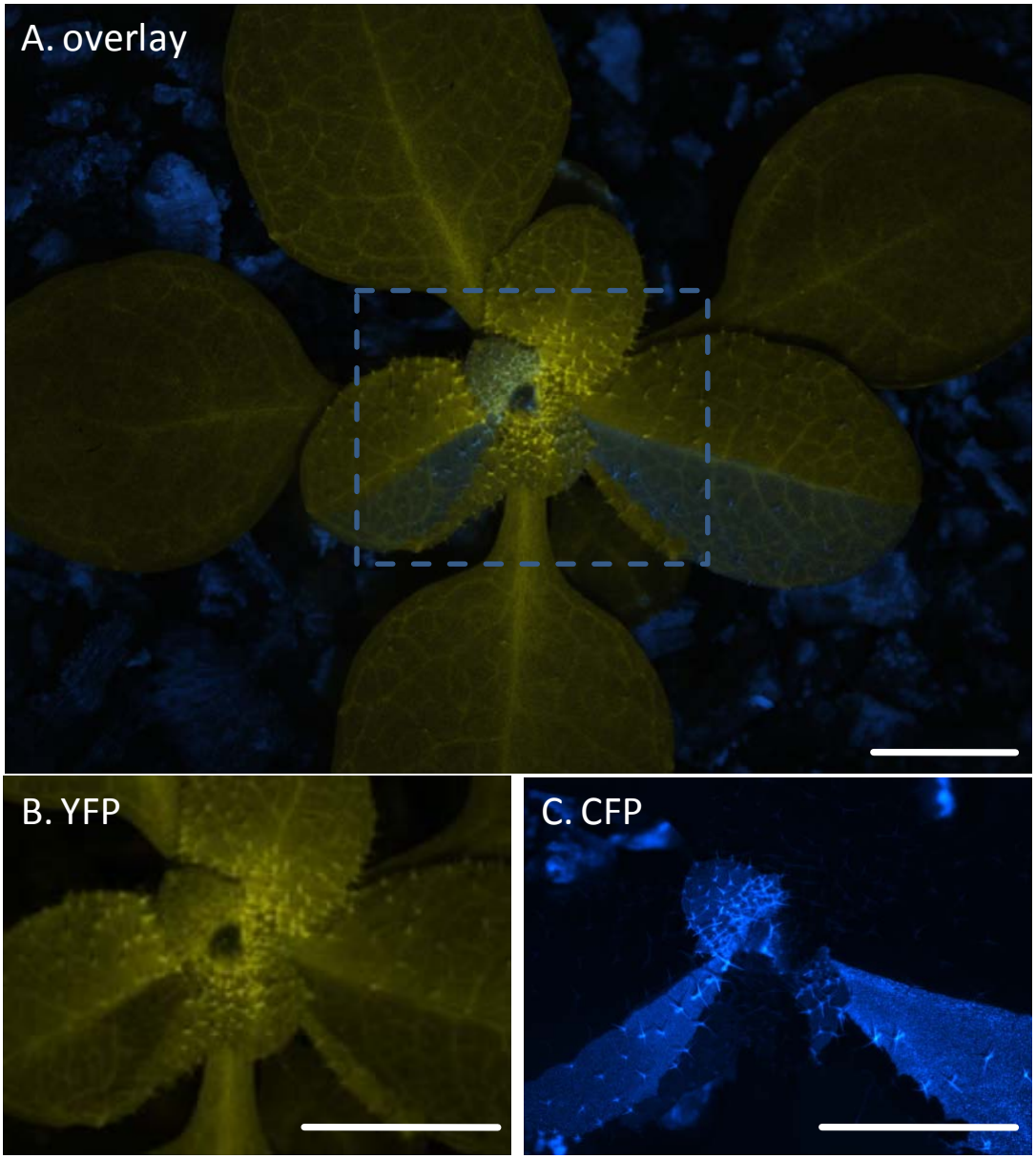


Figure 5.4 A sectored rosette

paps1-1 homozygous mutant carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes were induced with ethanol vapour 7 days after germination and imaged 13 days later. Scale bar 0.5 cm.

A. Overlay of CFP and YFP fluorescence micrographs. YFP (yellow) signal indicates non-recombined, i.e. rescued wild-type tissues, CFP (blue) signal indicates recombined, i.e. *paps1-1* mutant tissues. This is true for all following fluorescence images in this chapter.

B. YFP fluorescence micrograph of the marked square in A.

C. CFP fluorescence micrograph of the marked square in A

Box 5.1 The nine possible classes of inflorescences that can be observed when using the *CLV3>>Cre* system.

The border of the square represents the L1 layer, the inside of the square represent the L2 and L3 layers. Note that for the split inflorescences, the proportion of the two sectors can varies from a third to two third of the inflorescences. Chimeras that have the L1 and L2 of one genotype, but the L3 of the opposite genotype are possible, but difficult to detect using a fluorescence stereomicroscope (see text), so are not considered here.

5.3.1 Level 1 and Level 2: The interaction between two different organs in individual plants: leaves and flowers and the interaction between petals from different side inflorescences in individual plants.

The analysis of growth co-ordination at the first two levels (level 1: leaves versus flowers, and level 2: between flowers from different inflorescences) requires only analysis of non-sectored inflorescences. Petals were compared from two different kinds of inflorescences, as shown in **Figure 5.5** (inflorescence 1: fully mutant and inflorescence 2: fully wild-type). Most rosette leaves of these plants are wild-type (**Figure 5.4**). The fully recombined, i.e. mutant, petals are 50% larger than the fully non-recombined, i.e. wild-type petals (**Figure 5.6**-Level 1 and 2-Whole flower-Both). These results indicate that the petal enlargement caused by the *paps1-1* mutation does not require the whole rosette to be mutant and is neither influenced by, nor does it influence the growth of wild-type organs from a different side shoot. Hence, over different inflorescences *PAPS1* acts autonomously.

With regards to the contribution of cell layers, the results indicate that the genotype of the epidermis largely determines the growth behaviour of the organ; petals with a mutant epidermis were 44% larger than wild-type petals (**Figure 5.6**-Level 1 and 2-Whole flower-Epi), while loss of *PAPS1* function in the internal tissue did not lead to a significant size increase (**Figure 5.6**-Level 1 and 2-Whole flower-Int). Therefore, across cell layers, the paps1-1 effect seems to be non cell-autonomous.

Levels 3 to 5 require the analysis of sectored inflorescences**.**

5.3.2 Level 3: the interaction between petals from different flowers in individual inflorescences.

Only one epidermally sectored inflorescence was obtained in this experiment (**Figure 5.7**). The flowers with a mutant epidermis had 37% bigger petals than the wild-type flowers that are derived from this very same sectored inflorescence (**Figure 5.6-**Level 3-Whole flower-Epi). The results again support a cell-autonomous effect of the *paps1-1* mutation at the level of different flowers in one inflorescence; however I did not have data to conclude about the effect of a chimera in the internal layer at this level.

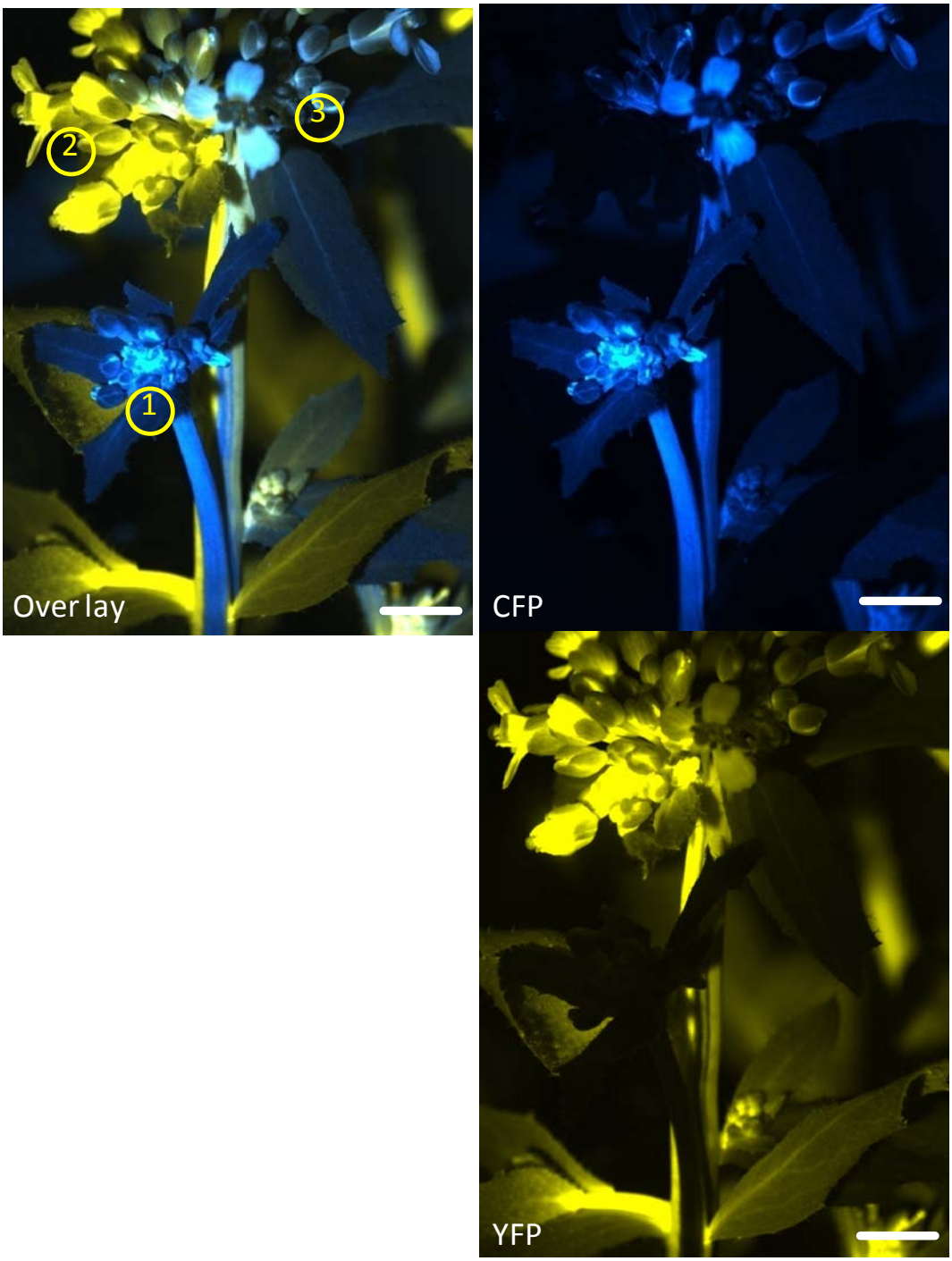


Figure 5.5 Chimeric inflorescences generated by *CLV3>>Cre.*

paps1-1 homozygous mutant carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes were induced with ethanol vapour 7 days after germination and imaged 30 days later. Scale bar 5 mm.

Inflorescence 1: a non-sectored inflorescence that was recombined in all cell layers.

Inflorescence 2: part of the sectored and fasciated inflorescence that was non-recombined in all cell layers.

Inflorescence 3 : part of the sectored and fasciated inflorescence that was internally recombined, but non-recombined in the L1.

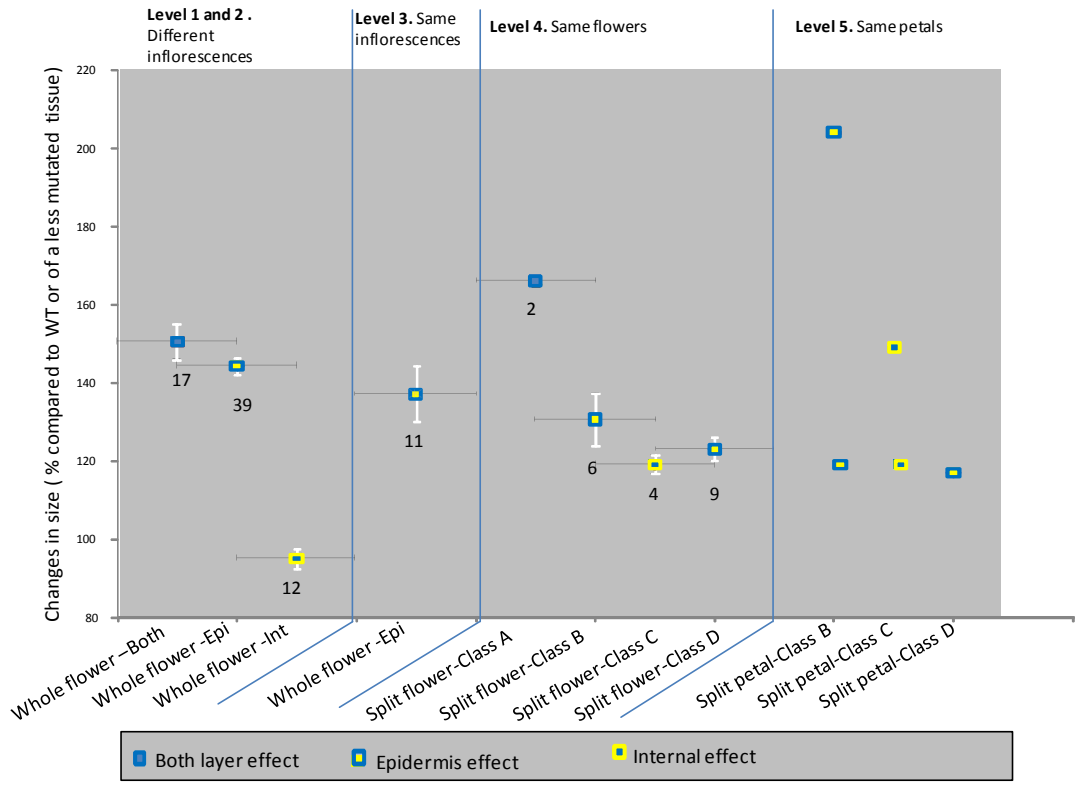


Figure 5.6 The *paps1-1* mutation act autonomously to promote the growth of petals.

paps1-1 homozygous mutant carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes were induced with ethanol vapour 7 days after germination and analysed 30 days later. Refer to Figure 5.1. for a definition of the five levels. Level 1 to level 4, each data point represents the mean ratio ± s.e.m, and the number of petals (n) that were used to calculate the mean is indicated by the number beneath. For level 5, each data point represents the measurement for one petal; in total only 5 split petals were obtained in this experiment.

Level 1 and 2. Petals were dissected from flowers from different inflorescences. These inflorescences were non-sectored and recombined in both layers (Whole flower - Both), in the epidermal layer only (Whole flower - Epi) or in internal layers only (Whole flower - Int). The size of the petals in each category was normalized to non-recombined petals from other inflorescences on the same plants.

Level 3. Petals were dissected from different flowers from within one inflorescence. This inflorescence was sectored. Only epidermal recombination was obtained for this level. Petal size was normalized to the non-recombined petals from flowers within that inflorescence.

Level 4. Petals were dissected from split flowers and normalized to the other petals in the same flower. The petals with least mutated tissue were set to be 100%. Refer to **Figure 5.8** and **Figure 5.9** (below) and text for the definitions of the split-flower classes.

Figure 5.6 The *paps1-1* mutation act autonomously to promote the growth of petals (continued).

Level 5. Split petals. Both halves of split petals were measured, and the size of the half with least mutant tissue was set to 100%. Refer to **Figure 5.10** and **Figure 5.11** (below) and text for the definitions of the split-petal classes.

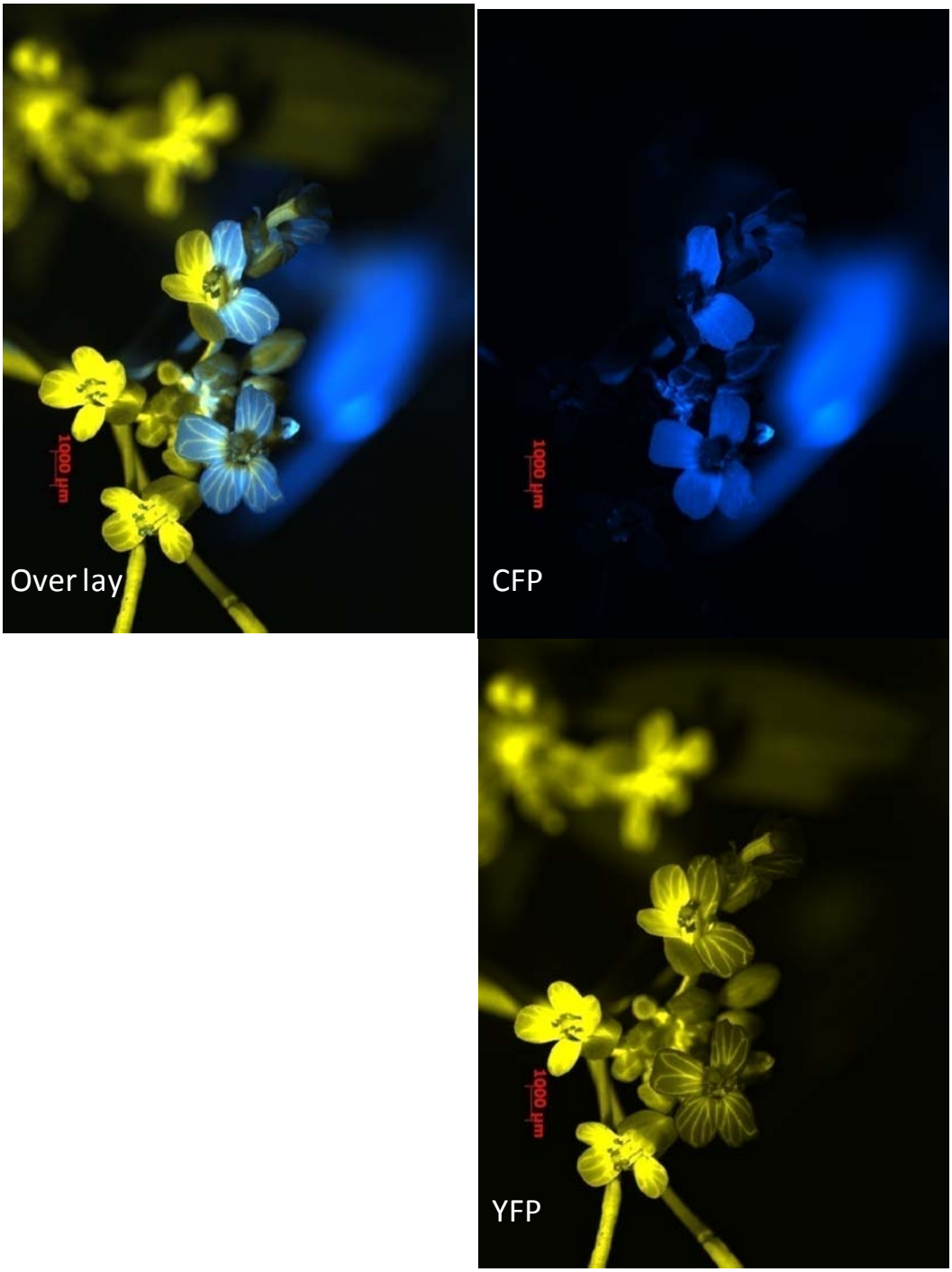


Figure 5.7 The inflorescence with a sector boundary in the epidermis overlying wild-type internal tissue.

paps1-1 homozygous mutant carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes were induced with ethanol vapour 7 days after germination and and imaged 30 days later. Scale bar 1mm.

5.3.3 Level 4: the interaction between petals within individual flowers.

At level 4 of this analysis, split or sectored flowers can be classified to five classes using two criteria. The first criterion is which layer(s) was recombined: L1, L2 or both (the genotype of the L3 layer in flowers is difficult to ascertain using only a stereomicroscope, as the L3 only contributes a relatively small amount of the final tissue to floral organs; for example, it only contributes to the internal tissue in a small region at the base of sepals, and not at all to petals). The second criterion is the genotype of the remaining parts of the flowers (wild type or mutant). Using these two criteria, all five theoretically possible split flowers are illustrated in **Figure 5.8** and **Figure 5.9.** In Class E, no flower was found in this experiment. The size comparisons of petals from each class of split flowers are shown in Figure 5.6-Level 4. The percentages were calculated as the ratio of the petals with more mutant tissues over the petals with the least mutant tissues from the same sectored flower. Clearly, in all classes, there is an increase in the size of the petals with more mutant tissue relative to the control petals, albeit only a moderate one of about 20% in some of the cases. This is again consistent with an autonomous effect of the mutation on growth. Regarding the contribution of each layer, loss of *PAPS1* function in either the epidermis or the internal tissue of the petal led to a comparable increase in petal size (**Figure 5.6**)

5.3.4 Level 5: the interaction between different parts of individual petals.

Split petals can also be classified into five classes using the two criteria similar to split flowers (i.e., recombination happened in which layer (L1 or L2) and the genotype of the remaining parts of the petals) as illustrated in **Figure 5.10**. Petals have four founder cells, two of which form the epidermis and the other two form the internal tissue including the vasculature. The size comparisons of these five classes of petals are shown in **Figure 5.6**-level 5 and **Figure 5.11**. Note that for two classes (split in both layers [class A] and split internally, but fully mutant in the epidermis [class E] no petals were found in this experiment. In total, five split petals were found, all of which, regardless of the particular combination of mutant and wild-type tissue, showed increased growth of the petal half with more mutant tissue relative to the other half, again supporting a cell-autonomous effect of the *paps1-1* mutation.

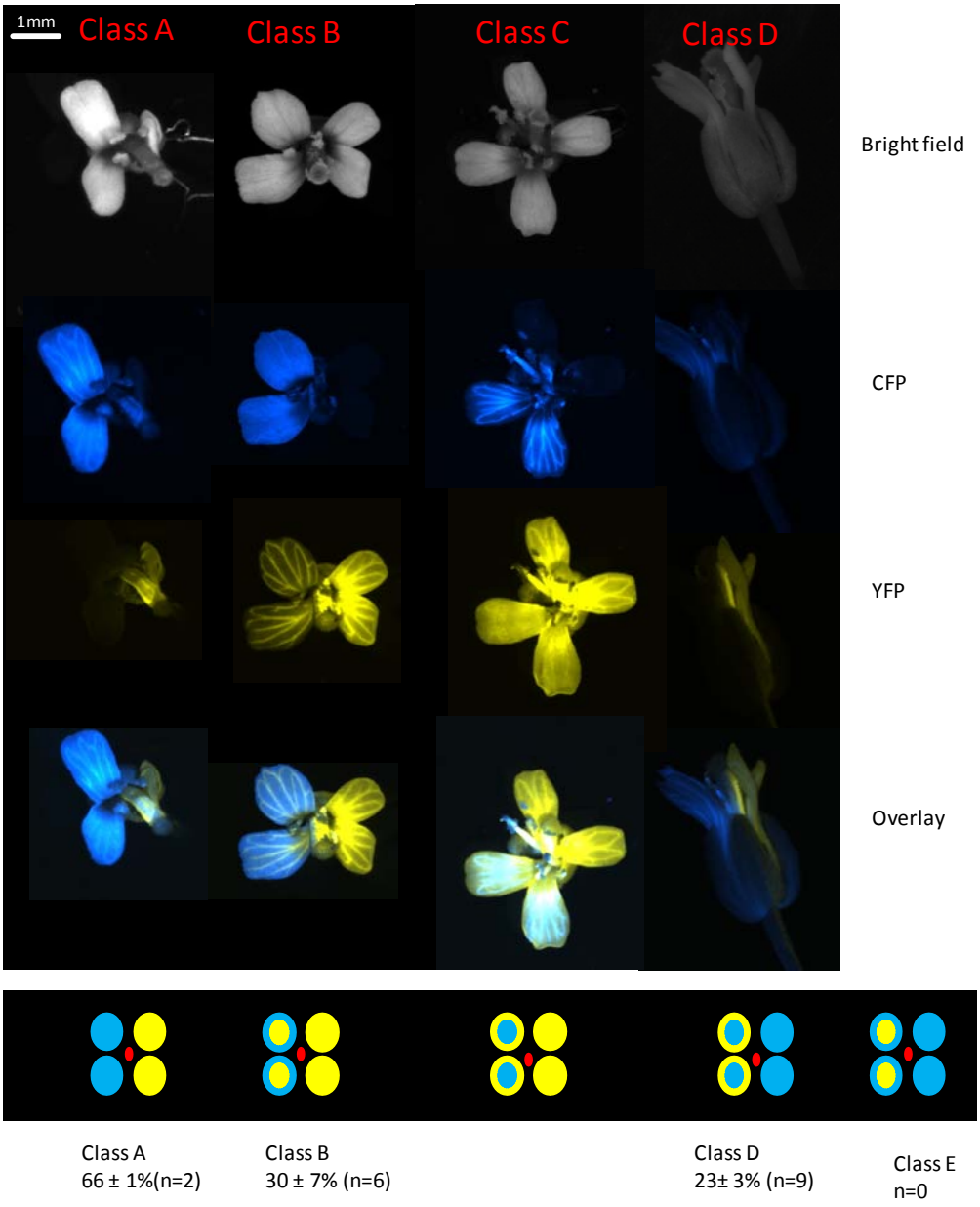


Figure 5.8 Five different theorically possible classes of split flowers and the four classes found in this experiment.

Class A. Fully split flowers in both layers. Two petals are fully WT and two petals are fully mutant.

Class B. Split in the epidermal layer. Two petals are fully WT and two petals are mutant in the epidermis, WT in the internal layer.

Class C. Split in the internal layer. Two petals are fully WT and two petals are mutant in the internal layer, WT in the epidermis.

Class D. Split in the epidermal layer. Two petals are fully mutant and two petals are mutant in the internal layer, WT in the epidermis.

Figure 5.8. Five different theorically possible classes of split flowers and the four classes found in this experiment (continued).

Class E. Split in the internal layer. Two petals are fully mutant and two petals are mutant in the epidermis, WT in the internal layer.
Number in % indicates the increase in size of the petals with more mutated tissues over the other petals with less mutated tissued from the same flowers. Values are average ± s.e.m, and the number of petals (n) that were used to calculate the mean is indicated.

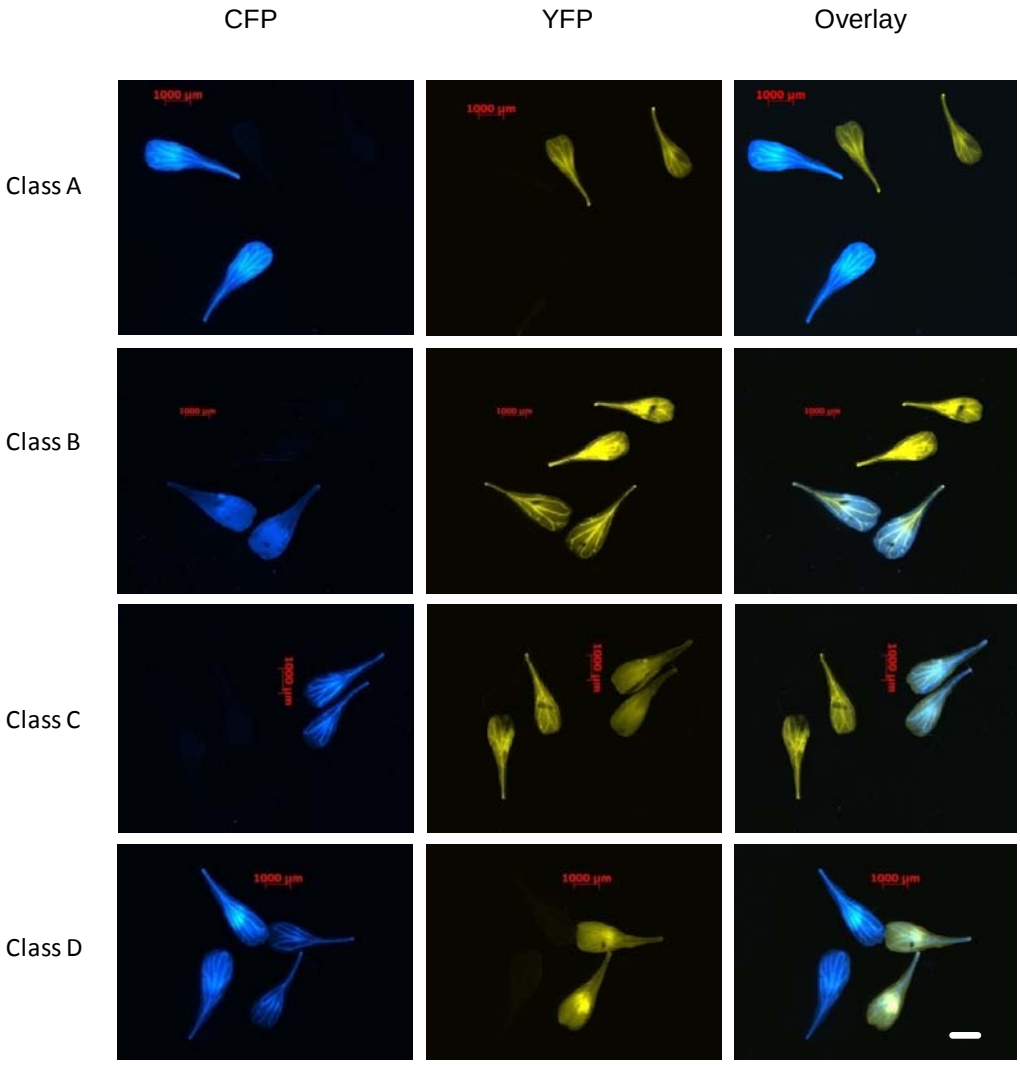


Figure 5.9 Petals from sectored flowers.

Each panel shows the four petals from a split flower in **Figure 5.8.** Refer to **Figure 5.8** for the class definitions. Scale bar 1mm.

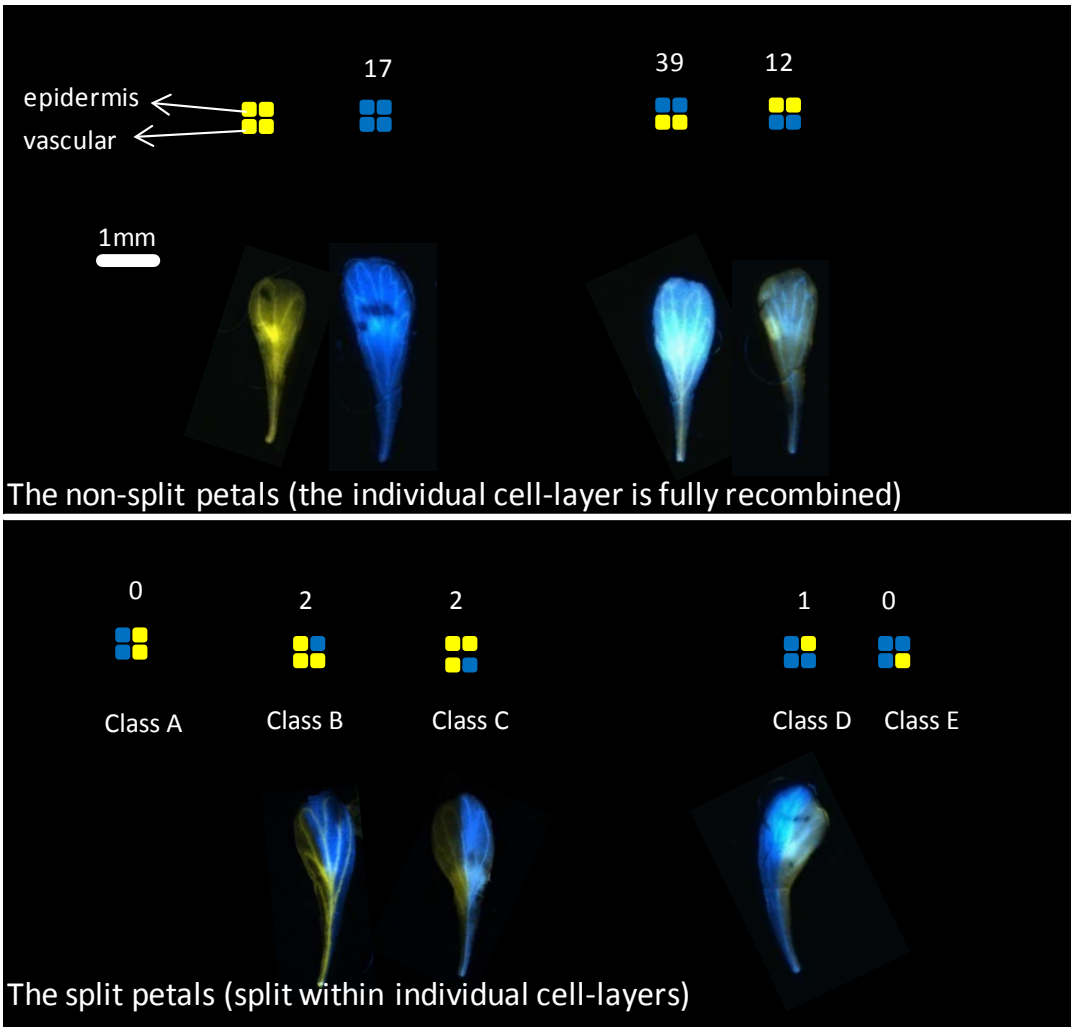


Figure 5.10 *paps1-1* mutation is autonomous within different parts of a petal.

Eight classes of theoretically possible recombined petals generated by ethanol inducing *paps1-1* homozygous mutants carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes are shown. The four squares represent the organ anlage of a petal consisting of clonally distinct cells arranged in two rows of two cells each. The two upper cells form the epidermis (L1), while the two lower ones form inner tissue (L2). The genotype of the inner tissue is recognized in the chimeric petals by the presence of fluorescence signal from vasculature. The number above the squares indicates the number of recombined petals found in this experiment for the corresponding class.

Upper pannel : From left to right: WT petals and the three classes of non-split petals: Fully recombined in both layers, fully recombined in L1 only, fully recombined in L2 only.
Below pannel: The five classes of split petals (Class A to E) were illustrated.
Class A: split in both layers, Class B: split in L1; the other half is fully WT; Class C: split in L2, the other half is fully WT; Class D Split in L1, the other half is fully mutant; Class E: split in L2, the other half is fully mutant.

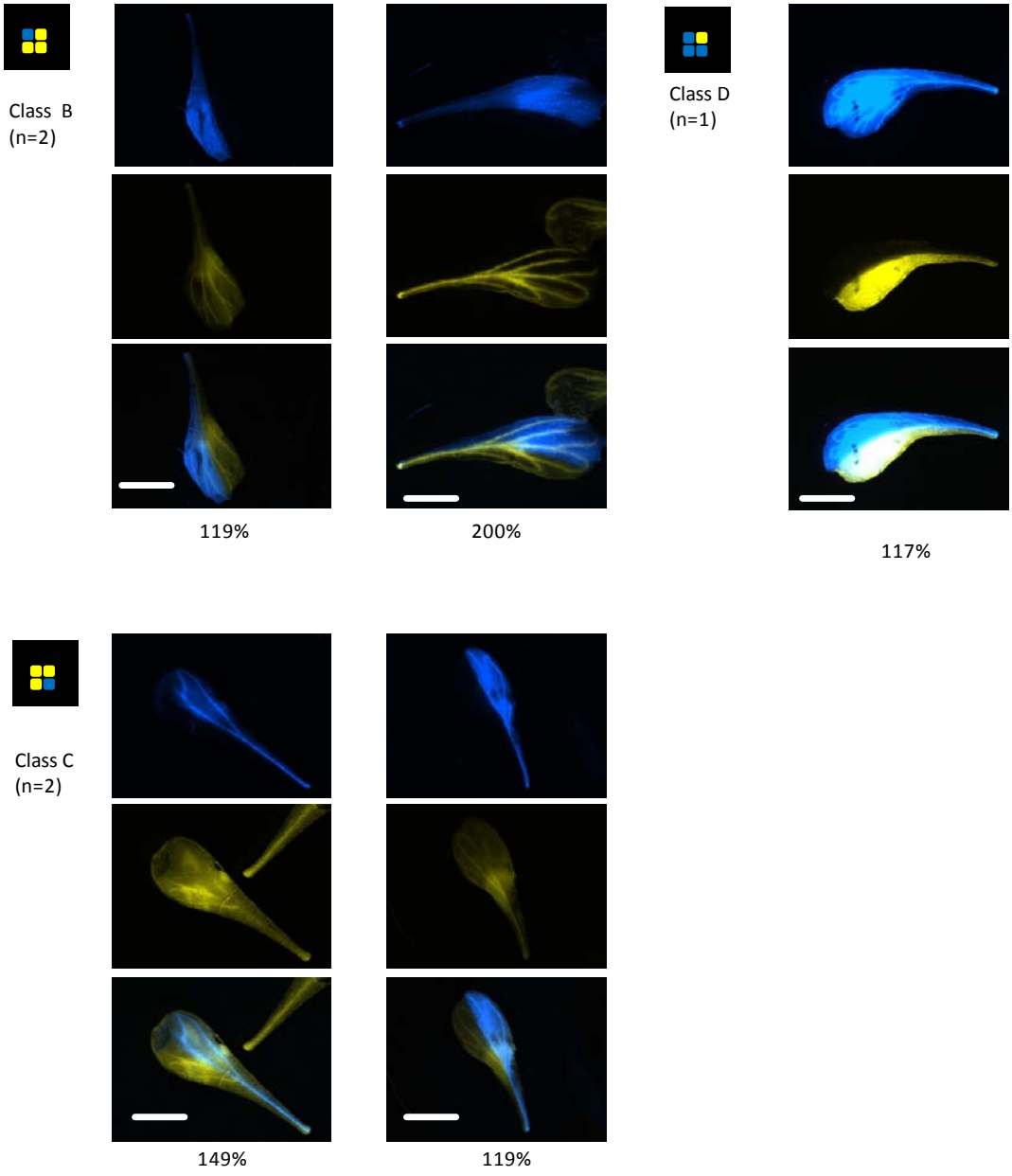


Figure 5.11 Split petals show asymmetric growth.

Magnified pictures of petals shown in Figure 5.10. The genotype of the inner tissue is recognized in the chimeric petals by the presence of fluorescence signal from vasculature. The percentage shown under the pictures is the size of the mutant half normalized to the size of the control half with less mutant tissue.

Scale bar is 1mm.

5.3.1 The effect of having *paps1-1* and wild-type tissues in the the internal layer: the internal chimera induces inflorescences meristem to split:

From the analysis of chimeric plants, out of at least 20 inflorescences observed, no inflorescence was found to be split in both layers. I found one inflorescence that contained sectors in the epidermal layer only, while the internal tissue was fully wild-type (**Figure 5.7**); one inflorescence that contained sectors in the epidermis only, while the internal tissue was mutant. However, I did not find inflorescences that were split in both layers or were split in the internal layer only. All plants analyzed contained internally-sectored stems, which in theory should give internally split inflorescences. However, invariably all of these internally split stems later fasciated and split into two or more stems (**Figure 5.12**, **Figure 5.13**, **Figure 5.14**, **Figure 5.15**). Some of these derived stems that still contained mutant and wild-type tissue internally, split again and again, until eventually the two split stems were homogeneous in the internal tissue (either mutant or wild-type). Therefore, no inflorescences with a sector boundary running through the internal tissue or through both epidermis and internal tissue were obtained except one inflorescence shown in **Figure 5.16**. This inflorescence produced a flower with a sector boundary in both layers, but the flowers that are made later in the fully CFP/mutant sector were wild-type (**Figure 5.16 B**).This suggested that the sector was terminated, possibly because it arose not in the long-term stem cells themselves, but in the immediate daughters of stem cells fated to differentiate that also express the *CLV3>>Cre* construct. Taken together, these results suggest that the presence of *paps1-1* mutant and wild-type tissues in the internal layer of a shoot/inflorescence meristem induces the meristem to split in to two or more meristems.

5.4 Discussion:

The growth effect of the *paps1-1* mutation is cell-autonomous, and a mutant epidermis is sufficient to drive excess growth in petals.

The results showed that the growth effect of the *paps1-1* mutation is autonomous at the level of different organs within one plant and within one flower, and also at the level of the two halves of one petal. When mutant and wild-type tissues were juxt aposed within one petal, the mutant half grew larger than the wild-type half. However, within one half of the petal, the *paps1-1* mutation seemed to act in a non cell-autonomous manner between the epidermis and the internal layer. The results here suggest that the epidermis plays a dominant role, as a mutant epidermis can promote growth of a wild-type L2 layer underneath. This non cell-autonomous effect on growth between layers is

Figure 5.12 The fasciated meristem phenotype was caused by chimerism in meristems, not by ethanol induction of *CLV3>>Cre*.

All plants were treated with ethanol vapour for 1 hour at when 7 days old. Pictures were taken at 31 days after induction. Scale bar 1cm.

Left: *paps1-1* mutant carrying only the *CLV3>>Cre* transgene. Note that no fasciated meristems were observed. Right, upper: *paps1-1* mutant carrying both *CLV3>>Cre* and *floxPAPS1* transgenes. Note some strongly fasciated meristems (arrows). Right, lower: another *paps1-1* mutant carrying *CLV3>>Cre* and *floxPAPS1* transgenes. Note the mildly fasciated meristems (arrows).

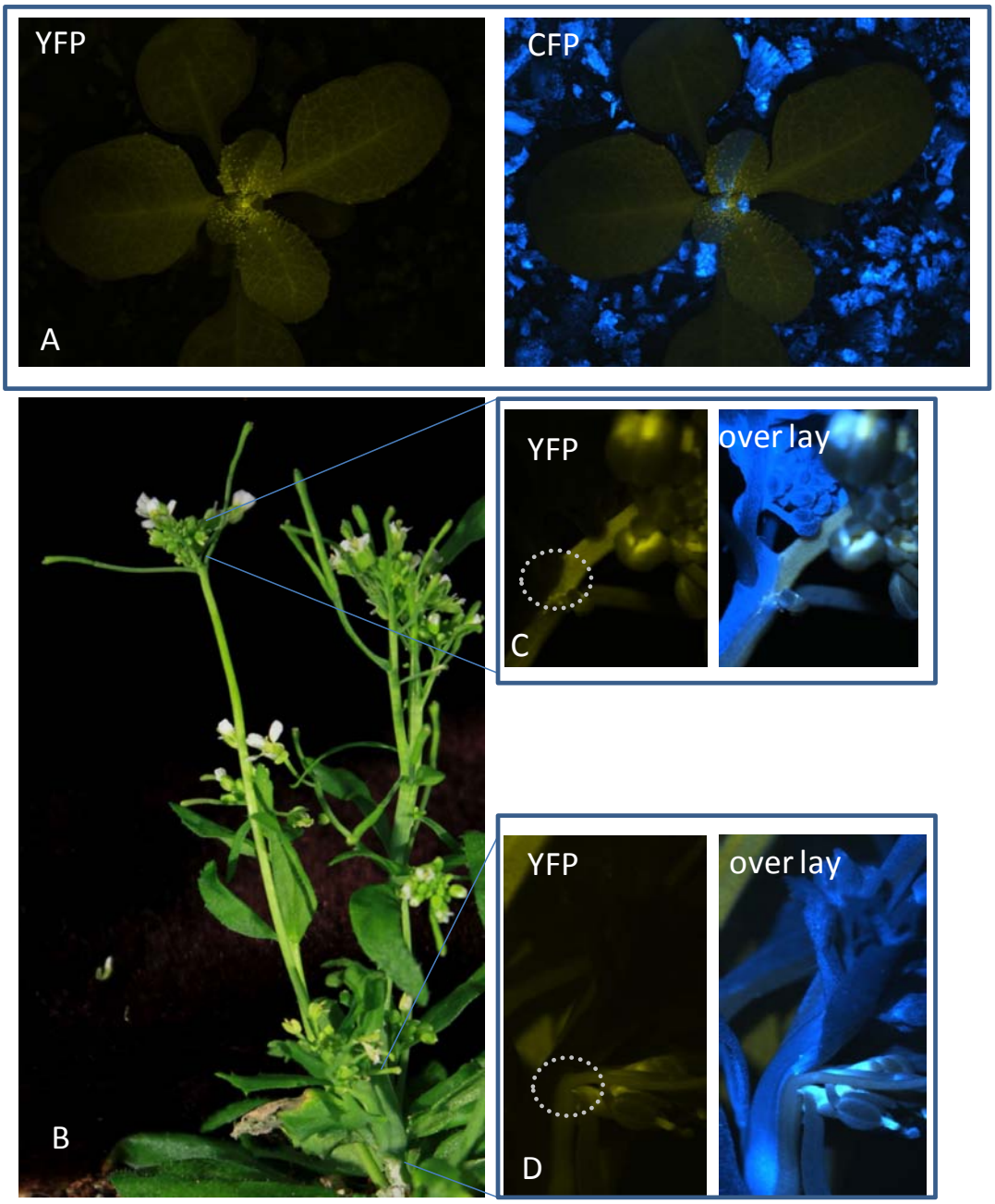


Figure 5.13 Internal chimerism causes inflorescences to split.

A. Magnified fluorescence micrographs of the same plant shown in Figure 5.12 right, upper panel. *paps1-1* homozygous mutant carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes were induced with ethanol vapour 7 days after germination. Image of a plant taken 13 days after induction.

B. The same plant in A imaged at 37 days after induction showing fasciated shoot.

C. YFP micrograph and YFP/CFP/brightfield overlay of the enlarged section in B showing that the internally sectored side shoot was split into two side shoots, one of which is fully recombined in both layer.

D. YFP micrograph and YFP/CFP/brightfield overlay of another section in B showing the internal chimerism inducing a split of the inflorescence.

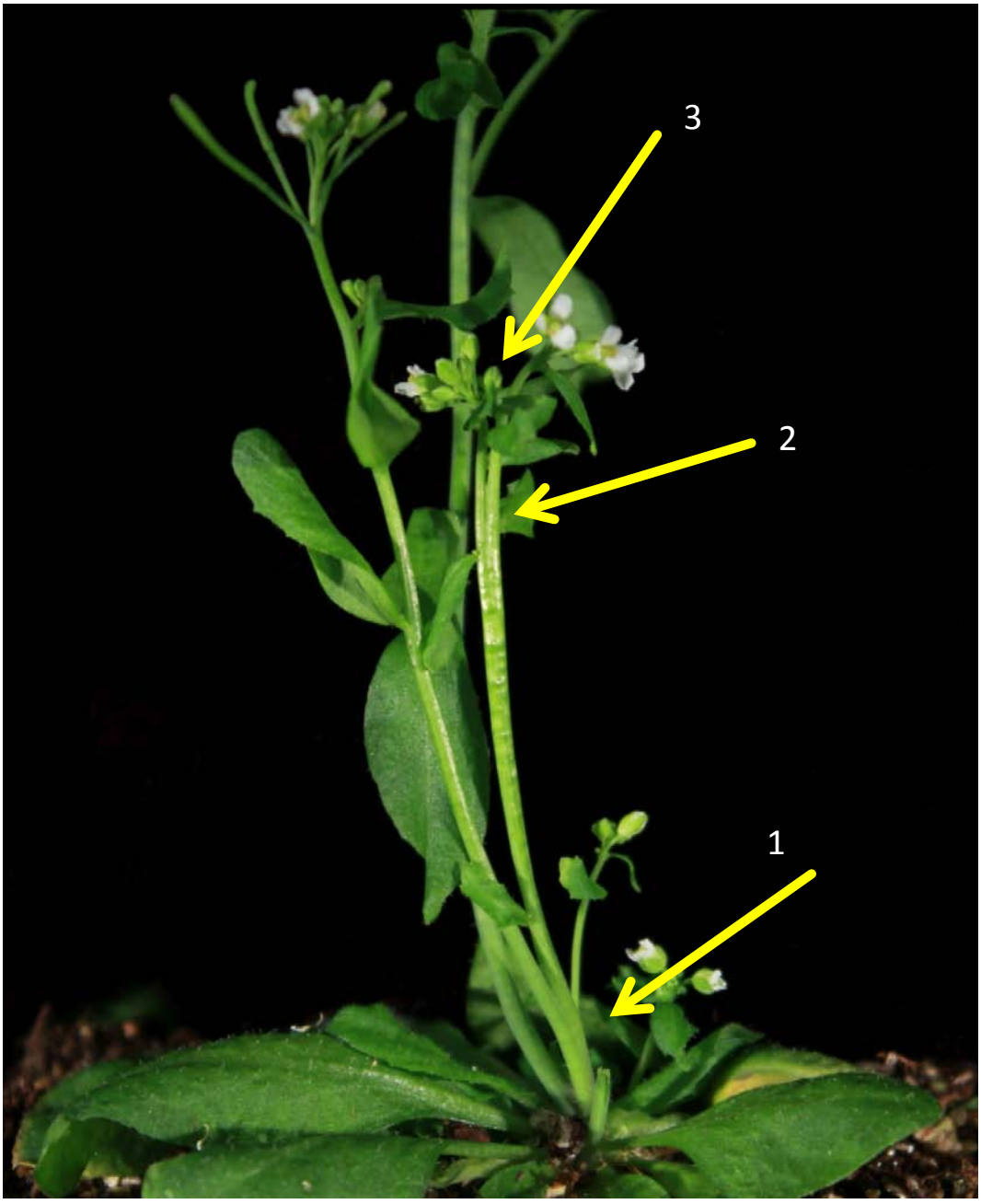


Figure 5.14 Another plant with *paps1-1* and wild-type sectors showed fasciated meristem.

Magnified image of plant on the right, lower half of Figure 5.12. Arrow points at three positions where a meristem was split into two or three new meristems. The numbers indicates the splitting positions that are examined with fluorescence microscope as shown in **Figure 5.15.**

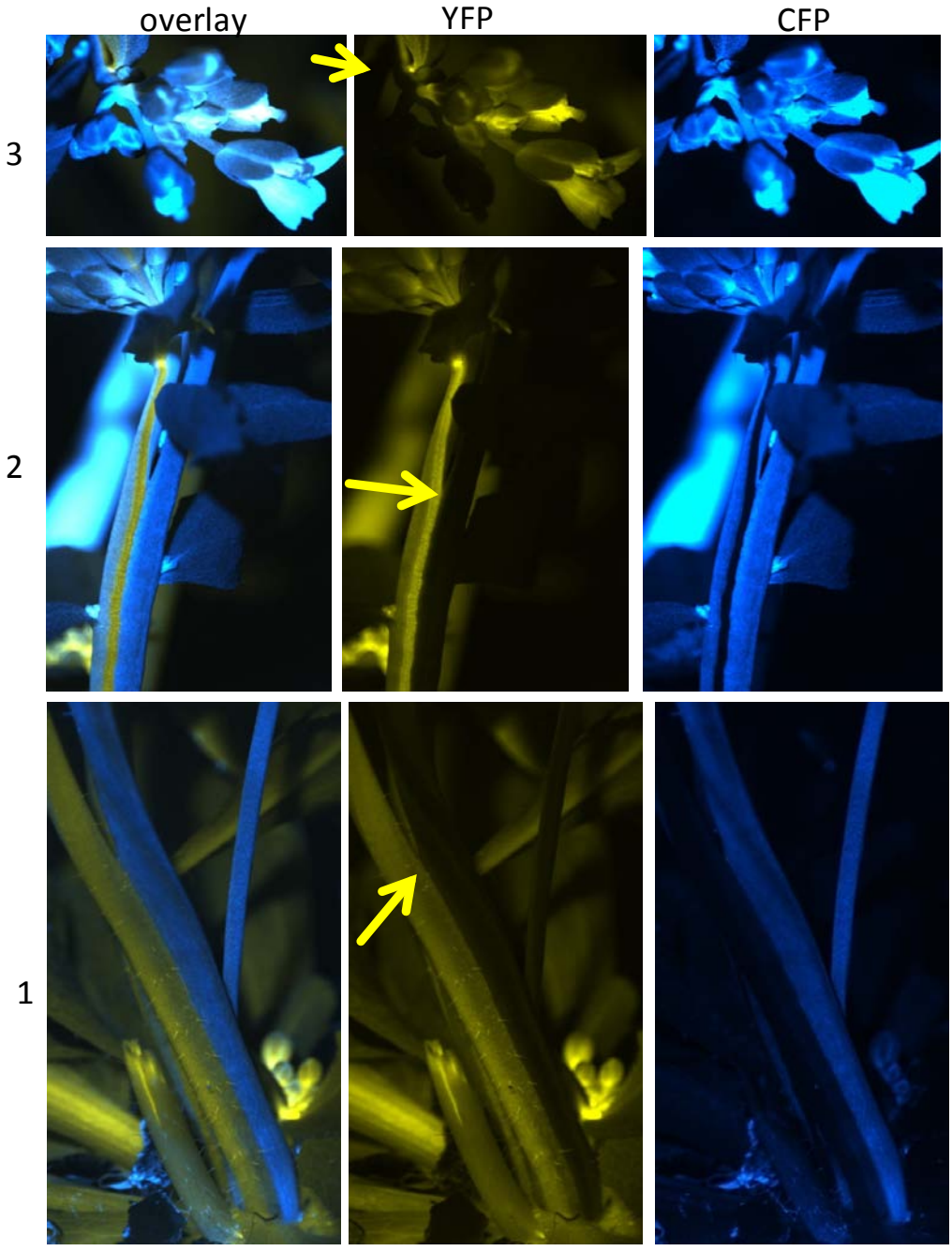


Figure 5.15 *paps1-1* and wild-type chimerism in the internal layers causes meristems to split.

Magnified fluorescence images of plant shown in Figure 5.14. Arrows point at positions where a meristem was split into two new meristems. Numbers on the left correspond to sections numbered in Figure 5.14.

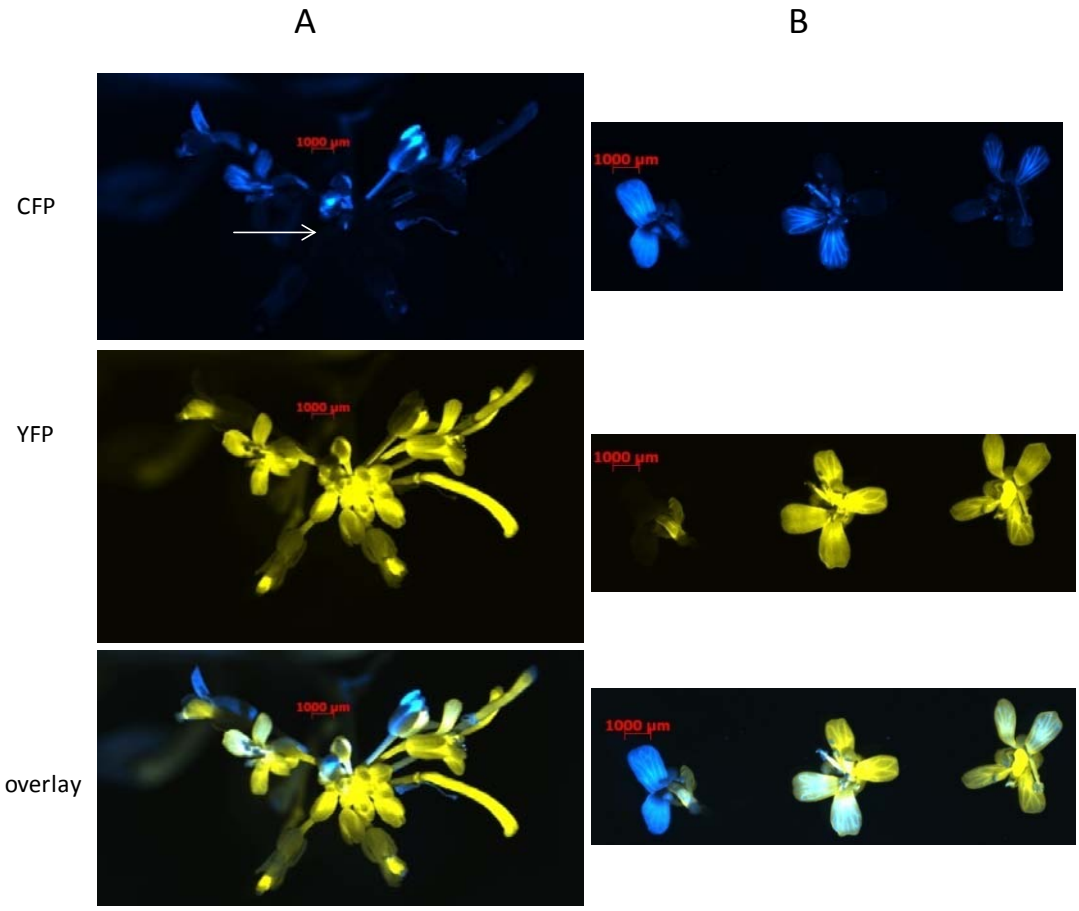


Figure 5.16 An internally sectored inflorescence.

paps1-1 homozygous mutant carrying one copy each of the *CLV3>>Cre* and the *floxPAPS1* transgenes were induced with ethanol vapour 7 days after germination and imaged 30 days later. YFP, CFP and overlay micrograph of inflorescence (A) and of the flowers (B) dissected from the inflorescence in (A). Note the terminated CFP sectors in (A)-arrow. Scale bar 1mm.

likely to be a general rule in plants because chimera layers with a different growth regulator such as *BRI1* (*BRassinosteroid Insensitive 1*) (Savaldi-Goldstein et al., 2007) also follow this rule that the epidermis can drive the growth of the whole organ.

Do the internal layers have any role? The answer is less straightforward. One result suggests no contribution of internal tissue. There was no difference inthe size of petals from two flowers that were taken from two different inflorescences, with one inflorescence being internally recombined, the other inflorescence wild-type (**Figure 5.6**-Level 1 and 2- whole flower-Int). On the other hand, another result suggests a contribution of the internal tissue. The internally recombined petals were 20% larger than fully wild-type petals derived from the same flowers (**Figure 5.6** Level 3-split flower class C). These two results contradict each other on the role of *paps1-1* mutant internal tissues in promoting petal growth. Supporting a role also for the internal tissue, at level of individual split petals (level 5), the data suggest a 20% increase when the epidermis is mutant, and a 20% increase also when the internal tissue is mutant (**Figure 5.6** Level 5). At present, the reason for this discrepancy is unclear; potentially the analysis of a larger number of chimeric inflorescences could indicate which of the two scenarios is supported by the majority of observations and whether there are different trends at different levels of analysis.

The results here also suggest that the presence of *paps1-1* mutant and wild-type tissues in the internal layer of a shoot/inflorescence meristem induces the meristem to split in to two or more meristems. It is important to point out that this effect is unlikely to be the effect of the *CLV3>>Cre* transgene or the effect of ethanol induction. This is because the *paps1-1* mutant plants containing *CLV3>>Cre*, which were induced together with these chimeric plants, did not show this fasciation and split phenotype (**Figure 5.12)**.

The evidence presented here support the hypothesis that internal chimerism in meristems can induce meristems to split. This is a novel finding. The stem cell population must be under tight control to maintain a certain size, despite the fact that stem cells do divide. A potential cause for this split is an asynchronised cell division between wild-type stem cells and *paps1-1* mutant stem cells. This would also suggest that the difference in growth caused by the *paps1-1* mutation acts at a single-cell level and also in stem cells, long before organ initiation. It remains an open question why only internal chimerism seems to induce a split, while epidermal chimeras form nicely sectored inflorescences without splitting. Additional explanation for the meristem splitting can be the conflict of the two parts of the chimera.

6 Chapter 6. Other genetic analysis of *PAPS1*

6.1 Ethanol inducible *PAPS1*:

In order to answer when *PAPS1* activity is required for organ size regulation, I generated plants carrying an ethanol inducible PAPS1 construct (*pPAPS1::AlcR-AlcA::genomicPAPS1*) and the negative control (*pPAPS1::AlcR-AlcA::YFP*) in the *paps1-1* background.

After a continous ethanol induction (by vapour or irrigation every other day) starting from bolting for 20 days, petals were measured on the 12^{th} to 16^{th} flowers (**Figure 6.1**). The experiment results proved that the ethanol inducible system worked and some line of *pPAPS1::AlcR-AlcA::PAPS1* is able to fully rescue the petal size defect in the *paps1-1* while none of the negative control is.

A future experiment is to induce only once and then measure the petals of all the flowers which are made after the induction to see which flowers show the most rescued phenotype. From that, one can deduce which stage of the flower development is most sensitive to *PAPS1* activity.

6.2 *cstF64* mutants do not have bigger flowers like *paps1-1* mutants:

To ask whether the organ size phenotype is caused by a general reduction in 3' end processing activity or by specific reduction of the enzyme poly(A) polymerase 1 activity only, the phenotype of *cstf64-1* mutants (Liu et al., 2010) is examined. *Arabidopsis* has only one homolog of CstF64, which is a component in the 3' end processing machinery of *Arabidopsis*. As shown in **Figure 6.2**, the *cstf64-1* mutants have the same petal size as wild-type.

These results suggest the petal size phenotype is specific to the polyadenylation defect and it is not caused by a defect in cleavage of pre-mRNA or a general reduction in 3' end processing.

6.3 Genetic interactions between PAPS1 and other known size regulators.

To test whether PAPS1 has genetic interactions with other size regulators, double mutants of *paps1-1* with other size regulators which affect cell division and/or cell expansion were

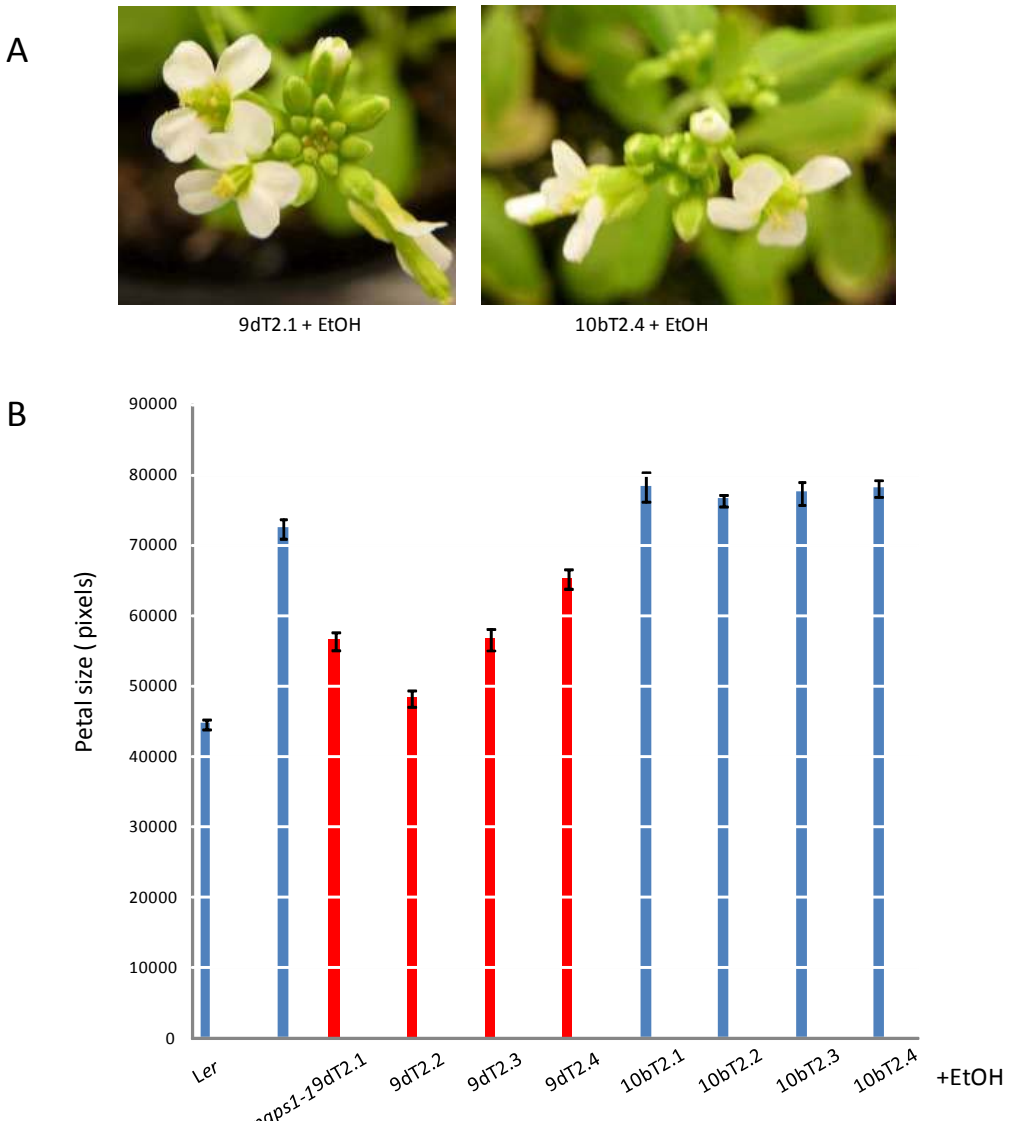


Figure 6.1 Ethanol inducible *PAPS1* constructs rescue *paps1-1* phenotype.

9dT2.X are the independent T2 families of paps1-1 mutants carrying the pSV9d (*pPAPS1::AlcR-AlcA::genomicPAPS1*), the ethanol inducible rescue construct; with X is the number of the family. 10bT2.X are the independent T2 families of *paps1-1* mutants carrying the pSV10b (*pPAPS1::AlcR-AlcA::YFPer*), the negative control construct; with X is the number of the family. All plants were treated with ethanol once every other day (by watering plants with 2% Ethanol) immediately after bolting and flowers were taken for measure 20 days after induction (Flower number 12th-16th).

A. Inflorescences after 20 days of ethanol induction. Note pSV9dT2.1 inflorescences have more flower buds than pSV10bT2.4.

B. Petal sizes measured from plants. For 9dT2.X and 10bT2.X, size was averaged amongst at least four T2 plants per family. All plants were treated with Ethanol.

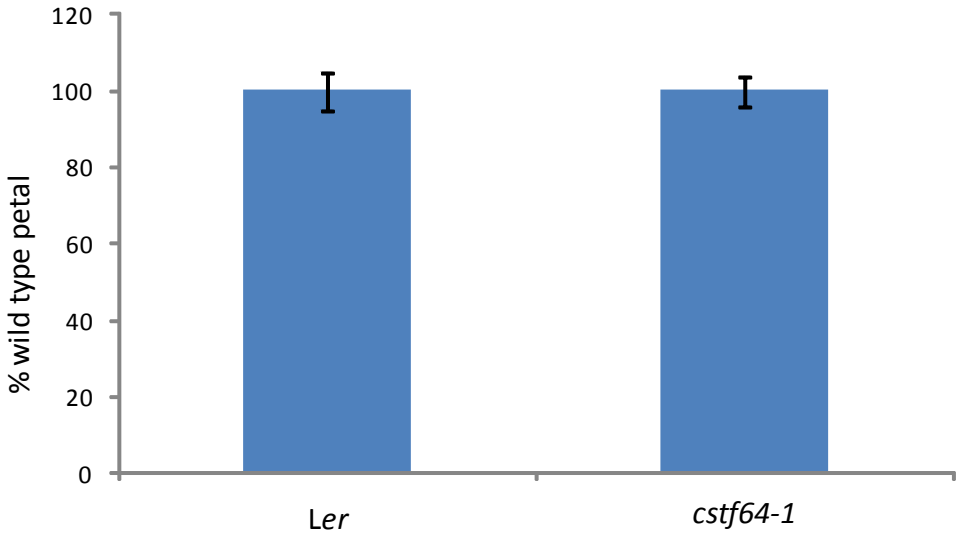


Figure 6.2 *cstf64-1* do not have bigger petals than WT

Petals of L*er* and *cstf64* mutant plants were measured. s.e.m, n=20

analyzed. Because of the ecotype/background problem, the *paps1-1* mutation in L*er* were crossed directly to *klu-2, bb-1, ant*^{7F5} ; and the *paps1-1* mutants that had been backcrossed to Col-0 three times were crossed to *35S::ARGOS*, *bpep*, *an3*, *jag*. More crosses are ongoing but at the end of my PhD, the double mutant of *paps1-1* and *klu*, *bb*, and *ant*^{7F5} were ready for analyzed. As shown in **Figure 6.3**, *paps1-1* is partially additive of *klu-2*, *bb-1* and *ant*^{7F5}. The effects of *paps1-1* in these double mutant backgrounds are not as strong as the effect in the *paps1-1* single mutant background. Shown in **Figure 6.3**, a closer look at the petal lengths and petal widths revealed that the effects on petal width of *paps1-1* are comparable in the double mutants *paps1-1 bb-1* (+ 21%) and *paps1-1 ant*^{7F5} (+25%) as to in single *paps1-1* mutant (+19%). The effects of *paps1-1* on petal length however, are much reduced in double mutants *paps1-1 bb-1* (+ 14%) and *paps1-1 ant*^{7F5} (+17%) compared to single *paps1-1* mutant (+32%). In *paps1-1 klu-2* mutant background, the enlargement effect of *paps1-1* on both petal length (ca. +25%) and petal width (ca. +12%) is reduced compared to that effect in single *paps1-1* mutant (ca. +32% and +19% respectively).

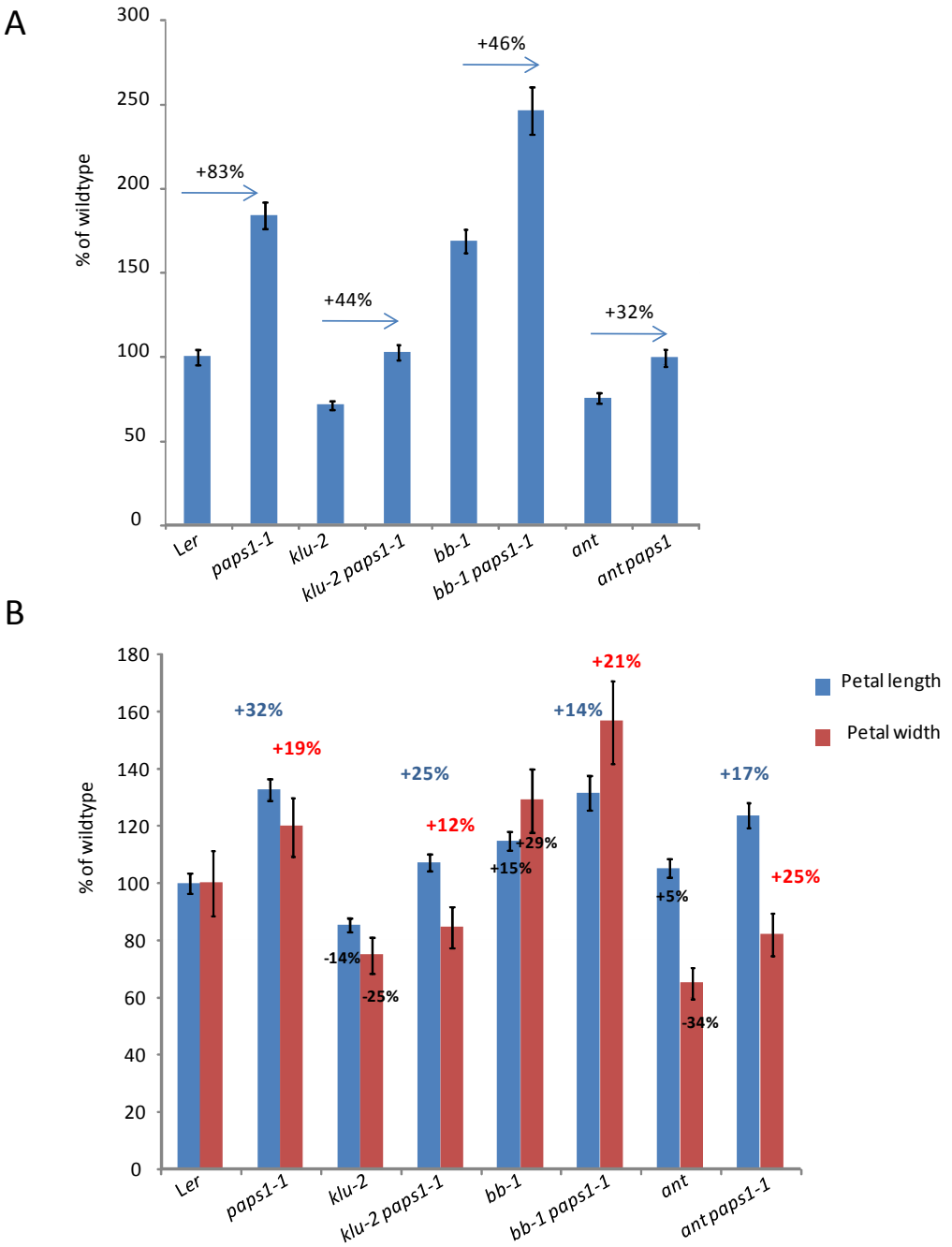


Figure 6.3 Genetic interactions between *PAPS1* and other size regulators that affects cell divisions: *KLUH, BB, ANT*.

s.e.m, n >= 20.

A. % Petal area compared to WT (L*er*) , numbers and arrows indicate the increase causes by *paps1-1* mutation in different background (i.e. equal to the ratio of value from double mutant divided by the value from single mutant *klu* or *bb* or *ant* respectively).

B. % Petal length and width compared to WT. The % numbers (above the bars) in Bold (red colour is for changes in petal width, blue colour is for changes in petal length) indicate the increase caused by *paps1-1* in different background (i.e. equal to the ratio of value from double mutant divided by the value from single mutant *klu* or *bb* or *ant* respectively).The %

Figure 6.3. Genetic interactions between *PAPS1* and other size regulators that affects cell divisions: *KLUH, BB, ANT (*continued).

numbers (below /in) the bars and in black colour indicate changes caused by single mutant *klu* or *bb* or *ant* respectively i.e. ratio of single mutant divided by WT.

7 Chapter 7. Whole genome microarray analysis

7.1 The design of the microarray:

To understand the molecular defect in leaves and flowers, I performed whole-genome microarray analysis. There are three layers in the design of the microarray (**Figure 7.1**)

The first layer is the genotype: wild-type (WT) (L*er*) vs. mutant (*paps1-1*).

The second is the tissue specificity: inflorescences (FL) or seedlings (LE). As *paps1-1* causes opposite effects on the growth of leaves and petals, we would like to indentify and then compare the lists of affected genes in leaves and floral organs. For this purpose, RNA was extracted from inflorescences of L*er* (WT) and *paps1-1* and from whole seedlings (including roots) of L*er* (WT) at 9 day-old and *paps1-1* at 11 day-old. The two different ages were chosen because *paps1-1* mutant seedlings grow more slowly than L*er*. At these ages, both seedlings were at a comparable early growth stage as judged by the size and cell division patterns of the first two leaves (**Figure 7.2A** and **C**). Also, at this age, the meristem is still vegetative as judged by the 1600X fold less expression of the *APETALA1*(*AP1*) gene, a marker for flower identity, in the seedlings compared to inflorescences (**Figure 7.2B**). Plants were grown at 21^0C so that the phenotype is strong in both leaves and flowers.

The third layer is the type of RNA used in hybridization: total RNA or RNA that was fractionated based on poly(A)-tail length. Total RNA is the standard microarray design using poly(A)+ RNA. Array hybridized with total RNA will give us information about the changes in the abundance (or steady-state level) of mRNAs. For the fractionated RNA, the idea is to identify the direct targets of PAPS1 based on the assumption that direct targets of PAPS1 should have a shorter poly(A) tail in the *paps1-1* mutant compared to WT. RNA was fractionated using an adapted protocol from mammalian system (Meijer et al., 2007) which results in the separation of total mRNA into two pools: the short poly(A)-tail pool which is enriched in mRNAs having a poly(A) tail of less than 50 nucleotides, and the long poly(A)-tail pool which is enriched for mRNAs having poly(A) tail of longer than 50 nucleotides.

Figure 7.3 showed that the fractionation was successful. After the fractionation, the long fraction and short fraction which comes from the same input RNA were differentially labeled (four biological replicates with dye swap) and hybridized to the same array (**Figure 7.1**). With this hybridization, each mRNA will have a ratio, referred to as the long/short ratio, which reflects how much of the mRNA is in the long poly(A)-tail form (over 50 A's) and how much is in the short poly(A)-tail form (less than 50 A's). If this long/short ratio changes between WT and mutant, the mRNA changes the poly(A)-tail length and therefore very likely to be the direct mRNA target of PAPS1.

To sum up, the microarray were designed combining three factors: mutant vs. wt, inflorescences vs. seedlings and total mRNA vs. fractionated mRNAs in order to determine

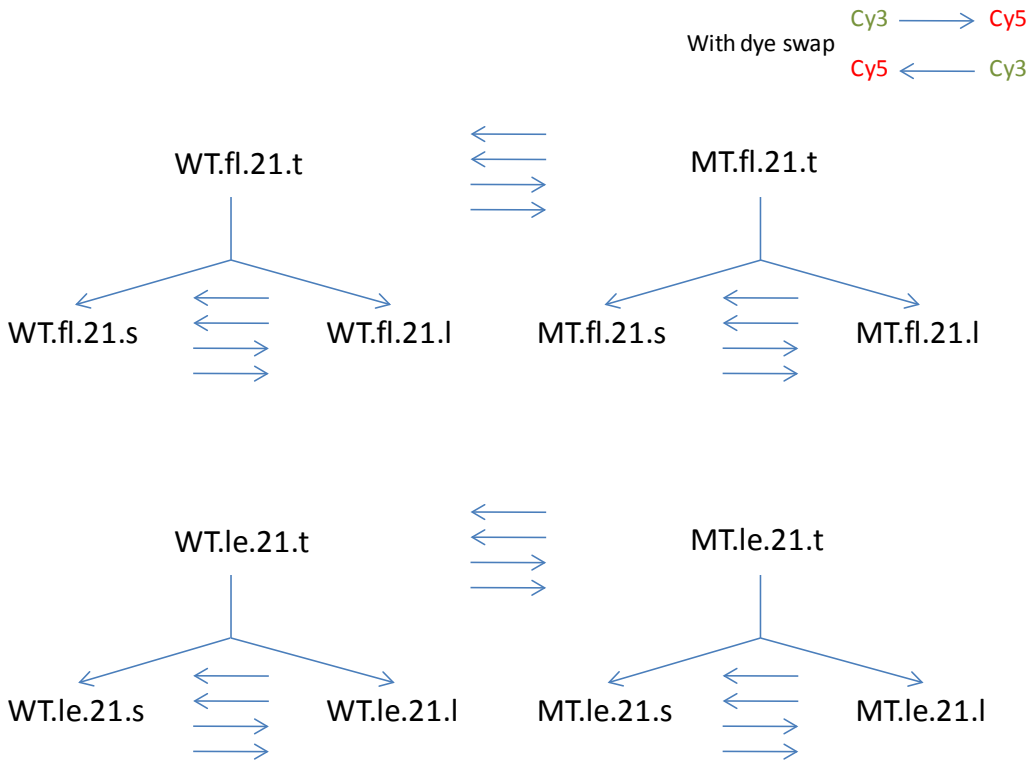


Figure 7.1 The design of the microarray .

WT: wild-type; MT: *paps1-1* mutant.

fl: inflorescences; le: 9-11 day-old seedlings.

21: plant grown at 21^0C; 16: plant grown at 16^0C.

t: total RNA; l: a long fraction (mRNAs with a poly(A) tail longer than 50 nucleotides; s: a short fraction (mRNAs with poly(A) tail shorter than 50 nucleotides)

The number of horizontal arrows indicates the number of biological replicates.

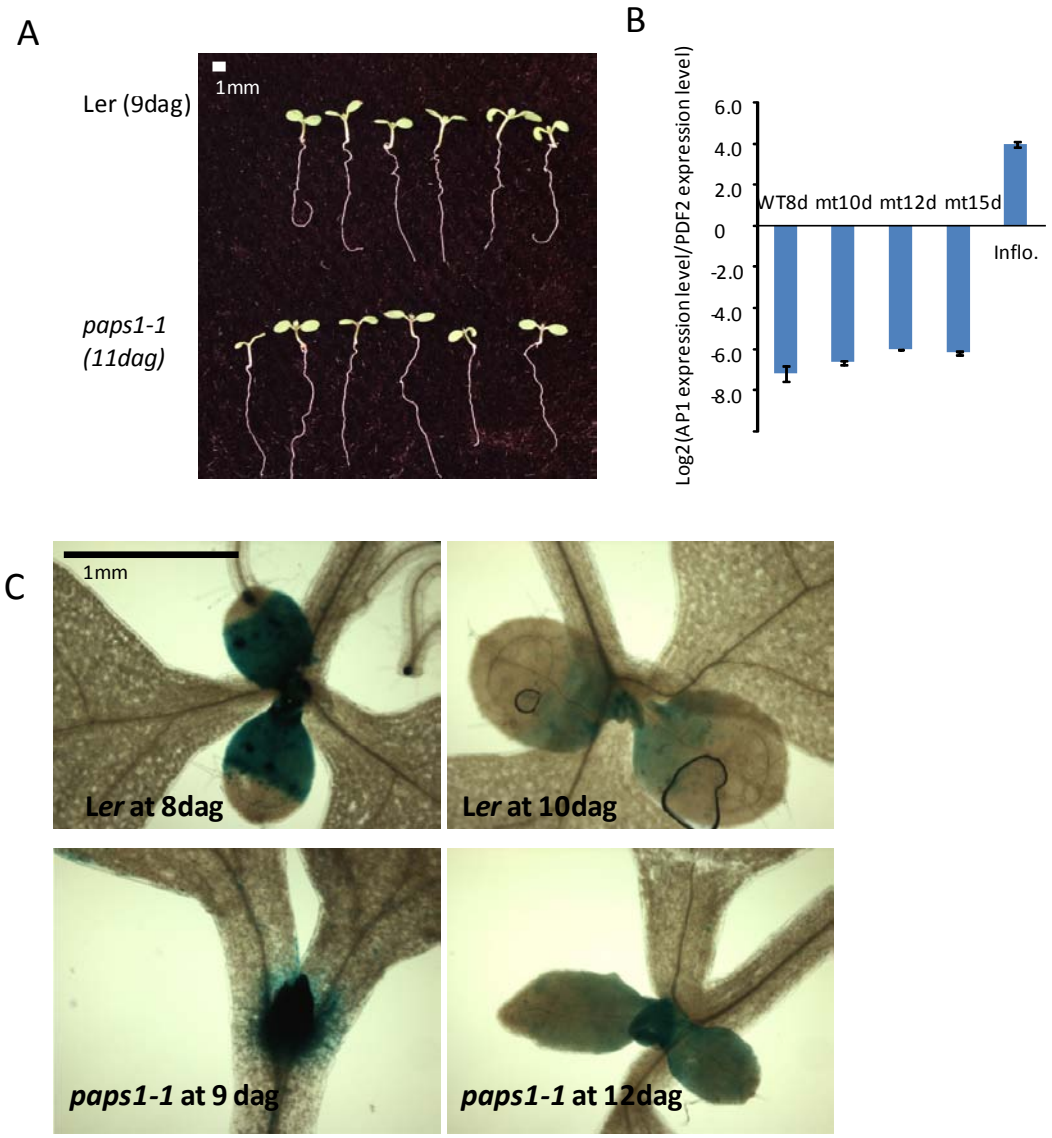


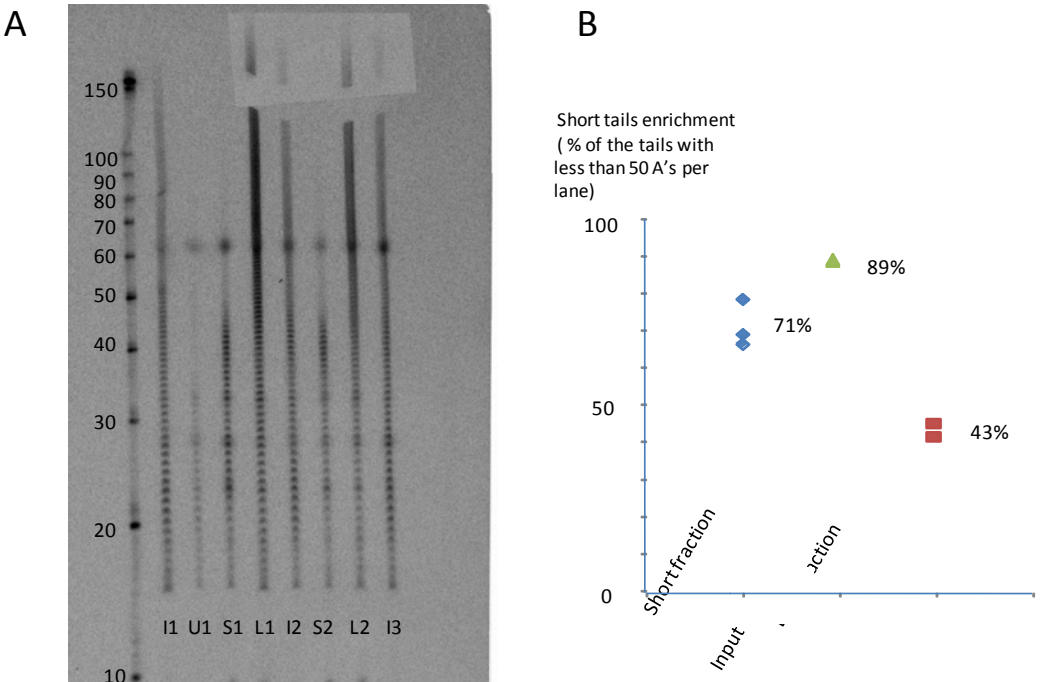
Figure 7.2 The seedlings used for microarray analysis.

dag is days after germination.

A. Pictures of seedlings at the stages used for microarray analysis.

B. qPCR showing the expresssion of *AP1*, a floral-meristem marker, in WT seedlings at 8 dag (WT8d) and mutant seedlings at 10,12 and 15 dag (mt10d, mt12d, mt15d) , i.e. before and after the stages where the samples were harvested for microarray. Expression was normalized to *PDF2* gene. Inflorescence RNA was used as a positive control for *AP1* expression in the qPCR.

C. Cell division patterns revealed by pCycB1;1::CDBGUS reporter constructs at one or two day before and one day after the stages of the seedlings used for microarray.



I: input
S: Short (0.085SSC elution)
L: Long (Water elution)
U: Unbound.
1,2,3 are three RNA samples from three different biological samples.

Figure 7.3 Poly(A)-tail length dependent fractionation of mRNA.

A. Radiograph showing the poly(A)-tail length patterns of different fractions. RNA fractions were end-labelled with ^{32}P cordycepin then treated with RNaseT1 and RNaseA to degrade all RNAs but the poly(A) tails which are then resolved in 8M urea 10% polyacrylamide gel. Amount used for labelling: I1: 1 µg total RNA; I2: 2 µg total RNA; I3: 3µg total RNA; S and L: a fifth of the short and long fractions that were precipitated from the fractionation of 80µg of total RNA . U: one thirty fifth of the unbound from the fractionation of 80µg of total RNA . The three biological replicates are: Sample 1: *paps1-1* inflorescences; sample 2: *paps1-1* seedlings; sample 3: *paps1-1* inflorescences. Plants were grown at 21^{0}C.

B. Quantitation of the graph in A.

the changes in both steady-state abundance and poly(A)-tail length between WT and mutant, and to compare these changes between seedlings and inflorescences.

At the first attempt for the microarray hybridization, I used inflorescence samples of plants grown at 16^0C. At this temperature, the flowers were still 80% bigger in the mutant compared to wild-type plants but the reduction in leaf size is mild. So supposedly there are fewer targets affected at this temperature in flowers yet the size regulator 'X' is still strongly affected. Using 16^0C therefore increase chances of identifying 'X' by reducing the other unrelated changes. Since initially we did not use dye swaps and the array results may contain many false positives, I do not present the analysis of this array here. Therefore all the analysis below is based on plants grown at 21^0C.

Regarding the idea to capture the direct PAPS1-sensitive targets that have altered polyadenylation by using RNA fractionation method, there are some false negatives that is noteworthy. These false negatives are mRNAs that do not change the steady state poly(A) tail length but can still be the direct PAPS1-sensitive target. If for these transcripts, the altered polyadenylation may result in immediate degradation of the pre-mRNA that have shorter poly(A) tails. In such case, although the transcript is a direct target but its steady state poly(A) tail length do not change but the abundance of the mRNA may change.

7.2 Raw microarray results- the lists of genes that change abundance and/or tail length in seedlings and/or inflorescences:

7.2.1 Total RNAs.

Using the cut-off criteria of a corrected p-value <0.05 and a magnitude of the log2-fold change $(\text{mutant/WT})^2$ >= 1, 616 differentially expressed genes (413 up-regulated, 203 down-regulated) between WT inflorescences and mutant inflorescences were identified. In seedling samples, more genes were differentially expressed (995 in total: 655 up-regulated and 340 down-regulated). After analyzing the overlap between differentially expressed genes in the two tissues, the genes similarly mis-regulated in both seedlings and inflorescences, the genes mis-regulated in seedlings only and in inflorescences only were summarized in **Figure 7.4A** and listed in supplemented excel files.

7.2.2 Fractionated RNAs.

As cut-off criteria, we used a corrected p-value<0.05 and a magnitude of $(\log2(\text{Mutant}^{\text{long/short ratio}}) - \log2(\text{WT}^{\text{long/short ratio}}))^2 >= 1$, we identified in the seedling samples four genes

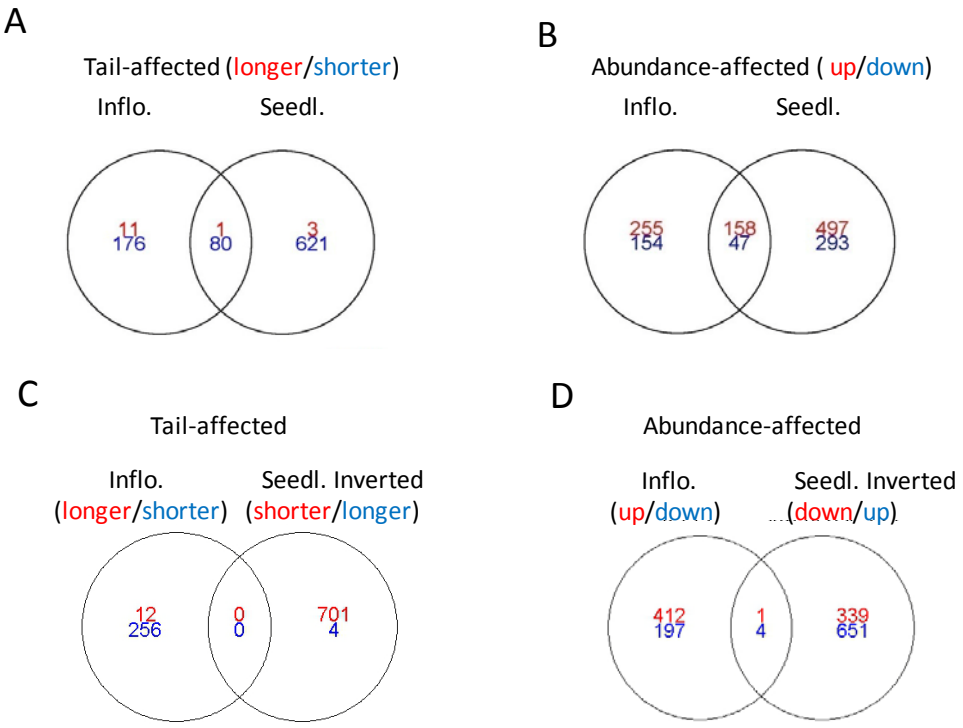


Figure 7.4 Overlap analysis between mis-regulated genes in seedlings (seedl.) vs. inflorescences (inflo.).

A. Tail-affected: transcripts that change poly(A)-tail length in *paps1-1* compared to wild-type. The number of genes whose mRNAs have a shorter poly(A) tail is shown in blue, a longer poly(A) tail length shown in red.

B. Abundance-affected: transcripts whose total abundance changes in *paps1-1* compared to wild-type. The number of down-regulated genes is shown in blue, of up-regulated genes is shown in red.

C. and D. inverse overlap analysis of A and B respectively, to reveal inversely regulated genes in inflorescences vs. seedlings. Note the colour code is reversed in seedling samples compared to inflorescence samples.

whose mRNAs have longer poly(A) tail and 701 genes whose mRNAs have a shorter poly(A)tail in mutant compared to WT tissue. The value for inflorescence samples were 12 (longer tail) and 256 (shorter tail). After analyzing the overlap between the poly(A)-tail affected genes in the two tissues, the genes whose poly(A) tail was similarly changed in both seedlings and inflorescences, the genes whose poly(A) tail was changed in seedlings only and in inflorescences only were summarized in **Figure 7.4B** and listed supplemented excel files. The abbreviation and nomenclatures of different pools of mis-regulated transcripts are summarized in **Box 7.1**

Box 7.1. Nomenclatures/Abbreviations for the mis-regulated transcript pools because of *paps1-1* mutation.

For simplicity, the mRNAs, identified by the total-RNA arrays, that change abundance between paps1-1 and WT, were referred to as ***abundance-affected*** *mRNAs. The mRNAs, identified by fractionated-RNA arrays, that change poly(A)-tail length between paps1-1 mutant and WT were referred to as* ***tail-affected*** *mRNAs. The seedling sample is abbreviated with LE and the inflorescences sample is abbreviated with FL.*

*The affected genes that were only mis-regulated in one organ but were unchanged in the other organ will be referred to as organ-specific tail-affected gene (****OSTAG****) and organ-specific abundance-affected gene (****OSAAG****). The affected transcript pool that was shared between both tissue types, i.e. inflorescences and seedlings, were abbreviated as* ***BOTAG*** *(both-organ tail-affected genes) and* ***BOAAG*** *(both-organ abundance-affected gene).*

*To show the organ-specific information, the organ name (LE or FL) will be added as prefix of these abbreviations. To show the direction of regulation (****Up*** *or* ***Down*** *with OSAAG and BOAAG;* ***Longer*** *and* ***Shorter*** *with OSTAG and BOTAG), a corresponding suffix will be added.*

For example, a full name for a pool FL-OSTAG-shorter indicates the transcripts that have a shorter poly(A) tail in the paps1-1 mutant inflorescences compared to WT inflorescences, but are unaffected in paps1-1 mutant seedlings.

7.2.3 Verifications of the microarray by qPCR and LM-PAT:

To verify the total RNA microarray, I performed quantitative PCR (qPCR) to check the expression of the abundance-affected genes identified by the array. The results are shown in **Figure 7.5**. Eight out of eight genes showed similar behavior in qPCR compared to array. For two of them (**Figure 7.5B**), the fold change determined by qPCR is quite different from the one determined by array, however the directions of regulation are similar between qPCR and array experiments. The result suggests that the total RNA microarray results are quite robust and accurate.

To verify the fractionated microarray, LM-PAT (Ligation-Mediated Poly(A) tail Test) (Salles et al., 1999) should be used but I encountered problems with LM-PAT. Many of the tail-affected genes that were chosen to be verified were too lowly expressed hence LM-PAT failed to amplify the ends of the mRNAs. In some other cases, the PCR gave many products with different sizes. The LM-PAT is now being optimized by another PhD student in the Lenhard's lab. Therefore, it must be kept in mind that all the analysis on tail-affected transcripts needs further verification independently by PCR based poly(A) tail test.

7.3 Detail-analysis of the gene lists and discussion

We want to answer two main questions with these microarray experiments:

1. How do the mis-regulated genes (with respect to both abundance and tail length changes) compare between seedlings and inflorescences taking into account the organ-specific expression factor (i.e. whether the expression of a gene is organ-specific)? Can this explain the opposite phenotypes in leaves and flowers?

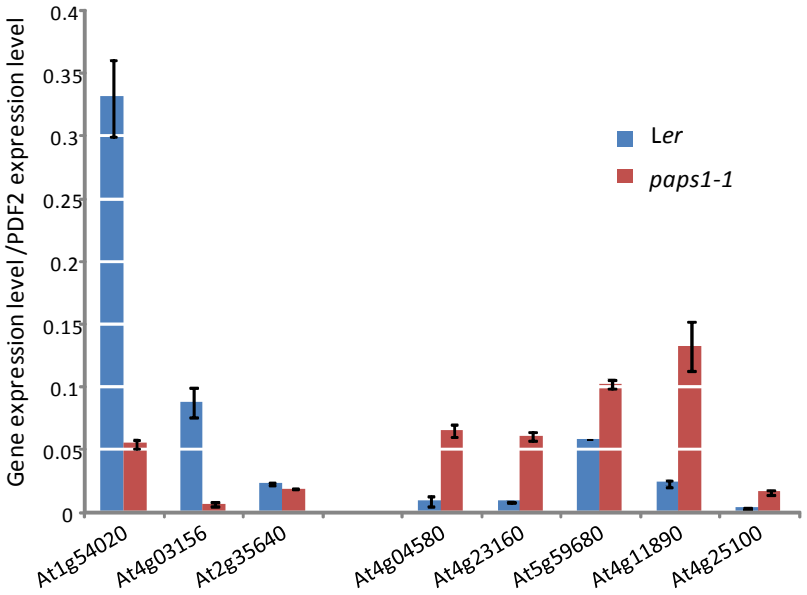
2. Are there any prominent pathways that are changed by the *paps1-1* mutation? Can these pathways explain the growth phenotypes in leaves and flowers?

To answer question 1, overlap analysis and 2D-contour map analysis were performed.

7.3.1 Inflorescences and seedlings have distinct PAPS1-dependent transcripts

First to compare the affected genes in LE vs. FL, overlapping analysis was performed on affected genes (both tail and abundance-affected genes) between the two tissues. The results revealed that the mis-regulated genes were largely non-overlapping between the two tissues (**Figure 7.4A** and **B**). Moreover, despite the opposite growth effects on seedlings

A



B

Gene\Data	Fold change (MT/WT) determined by qPCR	array	Mean expression value in inflorescences WT.16°C	MT.16°C	WT. 21°C	MT. 21°C	qPCR/array difference*
At1g54020	0.17	0.22	319	72	199	41	-0.4
At4g03156	**0.07**	**0.36**	**92**	**33**	**32**	**35**	**-2.4**
At2g35640	0.81	0.90	410	370	444	334	-0.2
At4g04580	7.72	4.48	45	201	44	158	0.8
At4g23160	7.36	4.06	32	129	35	85	0.9
At5g59680	**1.76**	**5.40**	**29**	**157**	**35**	**280**	**-1.6**
At4g11890	5.65	3.94	56	219	60	180	0.5
At4g25100	5.83	3.54	30	110	2370	1873	0.7

Figure 7.5 qPCR validations of several abundance-affected genes identified from total RNA microarray.

A. qPCR were performed using RNA from inflorescences of plants grown at 16^0C. Expression levels are normalized to *PDF2* gene.

B. Table comparing the qPCR results and the microarray results. * qPCR/array difference is calculated by log2(the value in column 2 divided by the value in column 3). Genes for which this difference is more than one is bold.

and flowers, there were only five abundance-affected mRNAs and no tail-affected mRNAs were regulated in opposite directions between seedlings and inflorescences (e.g. lower abundance in seedlings but higher abundance in inflorescences due to *paps1-1* mutation) (**Figure 7.4C** and **D**). This suggests that the opposite phenotypes are caused by different genetic pathways and not by the inverse regulation of one genetic pathway.

This overlapping analysis defined the pools of transcripts based on firstly the kind of mis-regulated (abundance/tail), secondly the tissues affected (LE, FL or both tissues) and thirdly direction of regulation (up/down and shorter/longer). For nomenclatures of transcript pools, see explanations in **Box 7.1**

7.3.2 Distinct *PAPS1*-dependent transcripts in inflorescences and seedlings are not due to the issue of organ specific expression.

I asked whether a gene only being affected in one organ was due to that gene only being expressed in this organ but not the other one. To answer this question, I employed a 2D-contour map with the total mRNA abundance, i.e. expression level, in FL (horizontal axis) and the total mRNA abundance in LE (vertical axis) as the two dimensions (**Figure 7.6**). The map has the resolution of 22x22=484 squares, each of which represents a specific combination of the two values in two dimensions. The number of genes falling into each square was calculated and then plotted (spectrum heat map) to produce what is called a 2D-contour map shown in **Figure 7.6**. If genes are expressed comparably in LE and FL, they will lie around the diagonal line of the 2D-contour map, while organ-specific genes will lie on either sides of the diagonal line. I use this 2D-contour map to analyze distribution in different transcript pools: the entire trancriptome (Bulk); OSAAG and BOAAG (**Figure 7.6**), OSTAG and BOTAG (**Figure 7.7**). There were some spots that were shifted away from the diagonal line (Marked with red arrow in **Figure 7.6** and **Figure 7.7**). Nevertheless, for most of transcript pools, the overall distribution of the spots was more or less centered along the diagonal line, arguing against organ-specific expression as the main reason for OSAAG and OSTAG.

7.3.3 The reason behind OSTAG?

The existence of distinct sets of OSTAG in LE and FL attracted my attention, because if one assumes that OSTAG are direct targets of PAPS1, this result suggests that organ identity changes the specificity PAPS1 to its mRNA substrates. I was therefore interested in what makes the LE-OSTAG and FL-OSTAG different from each other in terms of the mRNAs properties: tail length and abundance in FL and LE.

One pattern concerning the mRNA abundance stands out from the 2D-contour analysis. In the FL-OSTAG-shorter pools and BOTAG-shorter pools (refer to box 6.1 for nomenclature),

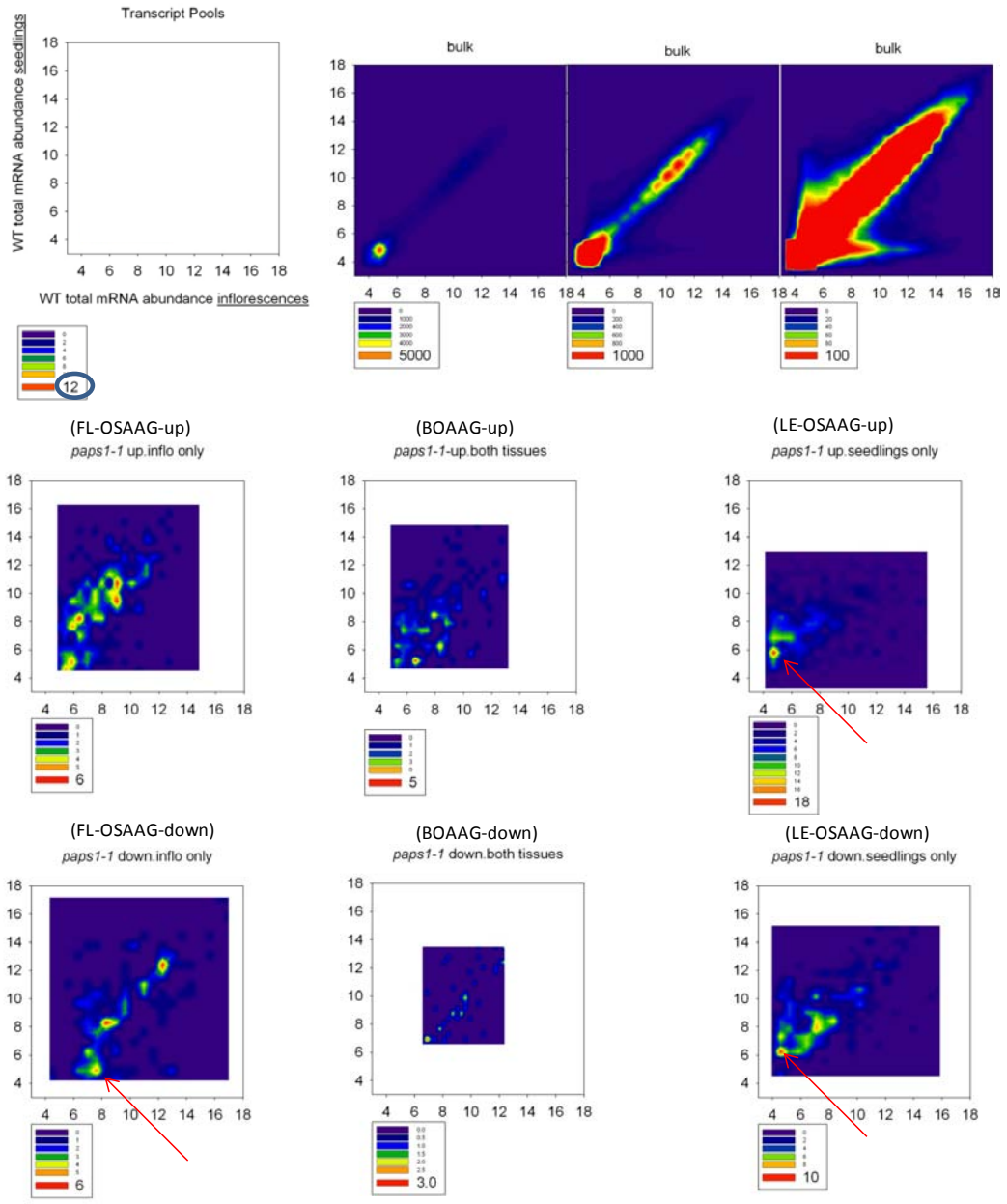


Figure 7.6 2D-contour: WT total mRNA abundance in inflorescences (inflo.) vs total mRNA abundance in seedlings (seedl.) analyzing abundance-affected transcripts.

See **Box 7.1** for explanation of the different transcript pools used in the analysis. In brief, different pools of transcripts (bulk: all mRNAs, abundance-affected transcripts (up/down in both tissues, in seedlings only and in inflorescences only) were separated in two dimensions: WT total mRNA abundance in inflorescences vs WT total mRNA abundance in seedlings.

WT total mRNA abundance is log2 of WT expression value extracted from microarray data. The number of genes that occupy an area in the 2D-map were spectrum-colour coded. The number on the right of each colour in the spectrum indicates a threshold, e.g. (black circle in the top left panel: if the area is occupied by more than 12 genes (the threshold), it will be

Figure 7.6 2D-contour: WT total mRNA abundance in inflorescences (inflo.) vs total mRNA abundance in seedlings (seedl.) analyzing abundance-affected transcripts (continued).

shown in red. The red arrow highlight gene clusters that is potentially interesting (see text for details)

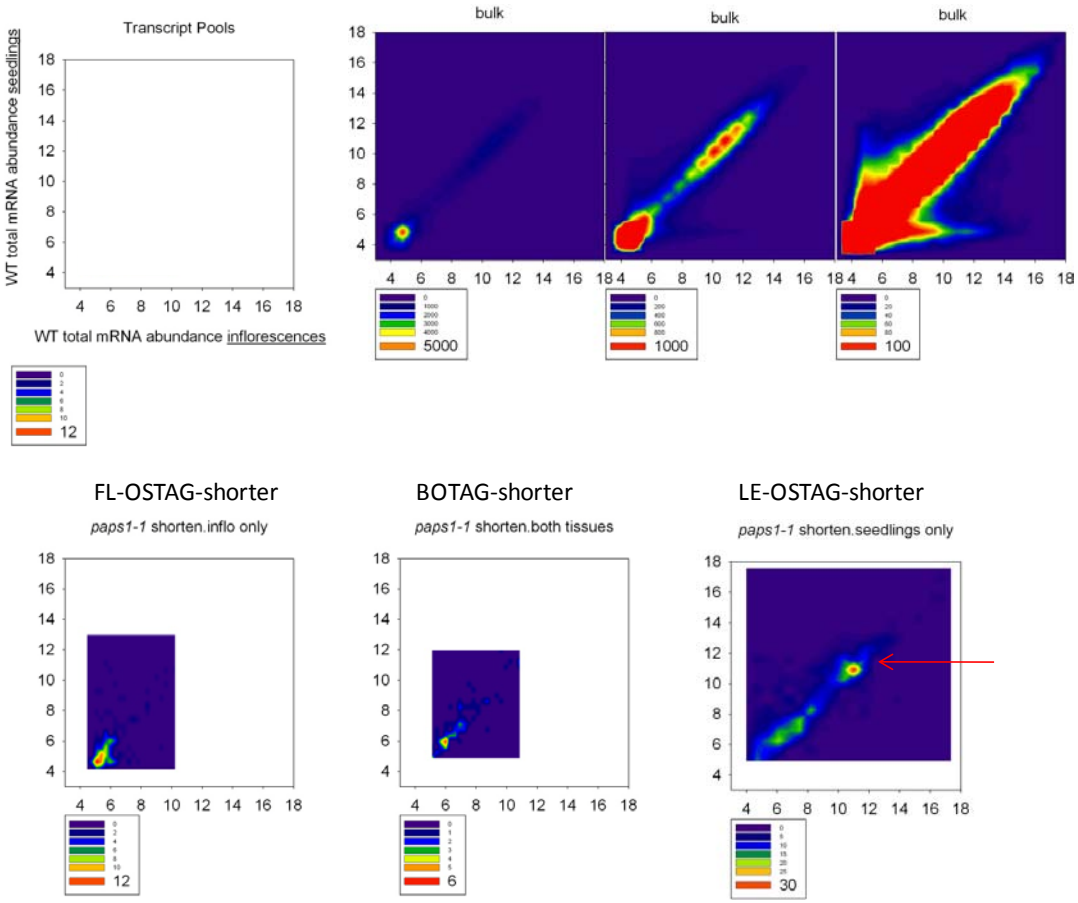


Figure 7.7 2D-contour: WT total mRNA abundance in inflorescences (inflo.) vs. seedlings (seedl.) analysing tail-affected transcripts

See **Box 7.1** for detailed explanation of the different transcript pools used in the analysis. In brief, different pools of transcripts (bulk: all mRNAs and tail-affected transcripts (shorter in both tissues, in seedlings only and in inflorescences only) were separated in two dimensions: WT total mRNA abundance in inflorescences vs WT total mRNA abundance seedlings.

WT total mRNA abundance is log2 of WT expression values extracted from microarray data. The number of genes that occupy an area in the 2D-map were spectrum-colour coded. The number on the right of each colour in the spectrum indicate a threshold, e.g. if the area is occupied by more than 12 genes (the threshold), it will be shown in red. The red arrows highlight a gene cluster that is potentially interesting (see text for details)

there were very few genes which are highly expressed (**Figure 7.7**) but in the LE-OSTAG-shorter, there seemed to be an enrichment of genes that are highly expressed in both organs. Therefore, it seems that the FL-OSTAG-shorter and the BOTAG-shorter tend to be lowly expressed but many of the LE-OSTAG-shorter do not. Therefore, I speculate that in the inflorescences, PAPS1 function is more important for the proper polyadenylation of lowly abundant transcripts than the highly abundant transcripts.

Additionally, I employed more 2D-contour mapping to compare the patterns in the OSTAG set of FL and LE in different two-dimension spaces: tail-length in FL vs. tail length in LE and tail-length vs. abundance in FL or LE. Firstly, comparing the poly(A)-tail length in FL vs. tail length in LE, **Figure 7.8** showed that there is a subpopulation of FL-OSTAG-shorter that have longer poly(A) tails in FL than in LE. This pattern is not observed for LE-OSTAG-shorter of BOTAG-shorter, yet it is observed, though to a lesser degree, in bulk RNA (arrow **Figure 7.8-**bulk RNA) suggesting the property of having longer poly(A) tail in FL than in LE is specific to FL-OSTAG.

Finally, 2D-separation by tail length vs. abundance in one tissue type probably discriminate FL-OSTAG and LE-OSTAG the best (**Figure 7.9 Figure 7.10 Figure 7.11**). FL-OSTAG-shorter set are enriched with lowly-expressed and long tail transcripts, which were also enriched in the bulk transcript pools but not in LE-OSTAG. By contrast, the LE-OSTAG set is enriched in both lowly-expressed, long-tail transcripts and highly-expressed normal-tail transcripts, with the later being more enriched.

To sum up, tail-affected and abundance-affected transcripts in inflorescences and seedlings do not overlap; suggesting that *PAPS1*-dependent regulation is modified in an organ-specific manner, despite the organ-nonspecific expression of the affected genes. FL-OSTAG-shorter tends to have a long tail and be lowly expressed while the LE-OSTAG-shorten pool is more enriched with mRNAs that are highly expressed and have a medium tail length.

7.3.4 Gene categories analysis linking differential expressed genes to cellular pathways:

Are there any prominent pathways changed by the *paps1-1* mutation? Can these pathways explain the growth phenotype in leaves and flowers? This analysis is not completed. But here I present the ideas and some preliminary results.

To find biological pathways that may be enriched in the misregulated genes, I employed MASTA (Reina-Pinto et al., 2010), and MAPMAN (Thimm et al., 2004). MASTA is a bioinformatic tool to compare a differential-expressed gene (DEG) list of interest with 600 other DEG lists that are identified from published microarray experiments analyzing various conditions/mutants. MAPMAN is another bioinformatic program, where genes that are

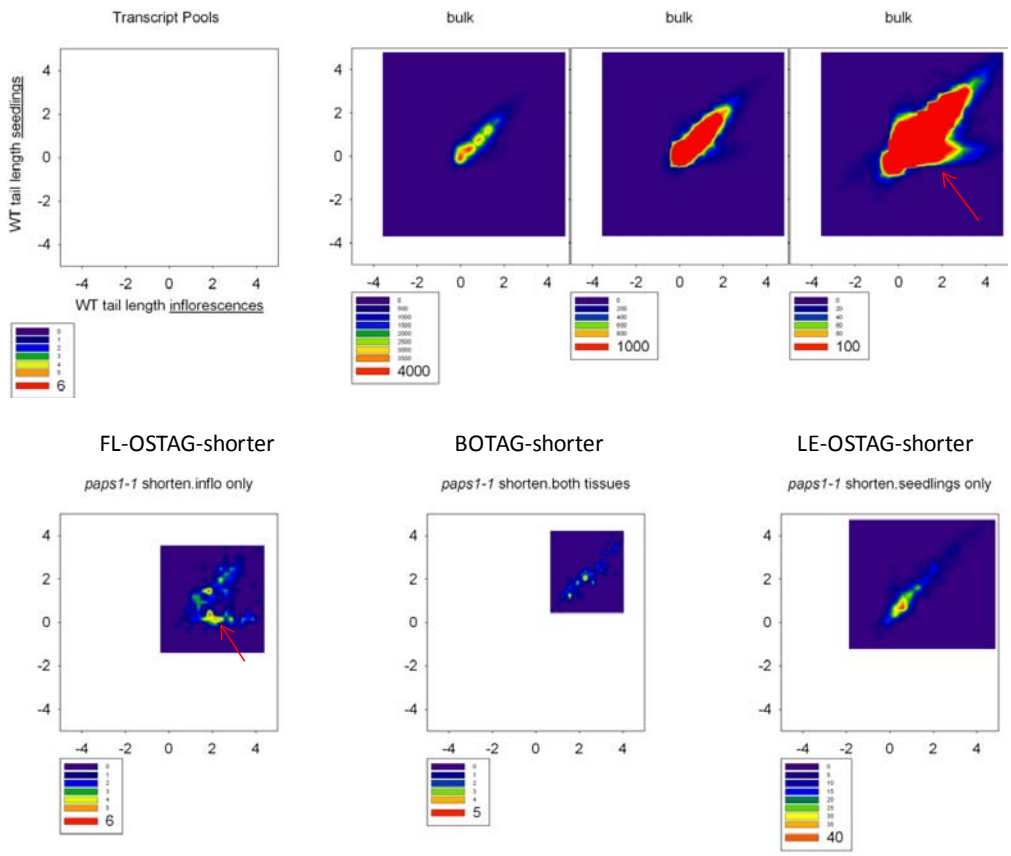


Figure 7.8 2D-contour: WT poly(A)-tail length in inflorescences (inflo.) vs WT poly(A)-tail length in seedlings (seedl.) analysing tail-affected transcripts.

See **Box 7.1** for detailed explanation of the different transcript pools used in the analysis. In brief, different pools of transcripts (bulk: all mRNAs and tail-affected transcripts (shorter in both tissues, in seedlings only and in inflorescences only) were separated in two dimensions: the WT poly(A)-tail length in inflo. and seedl.

WT poly(A)-tail length is log2 of the WT long tail/short tail ratio as determined from the fractionated microarray. Positive value indicate relatively long poly(A) tails.The number of genes that occupy an area in the 2D-map were spectrum-colour coded. The number on the right of each colour in the spectrum indicate a threshold, e.g. if the area is occupied by more than 6 genes (the threshold), it will be shown in red.

The red arrow highlights gene clusters that are potentially interesting (see text for details)

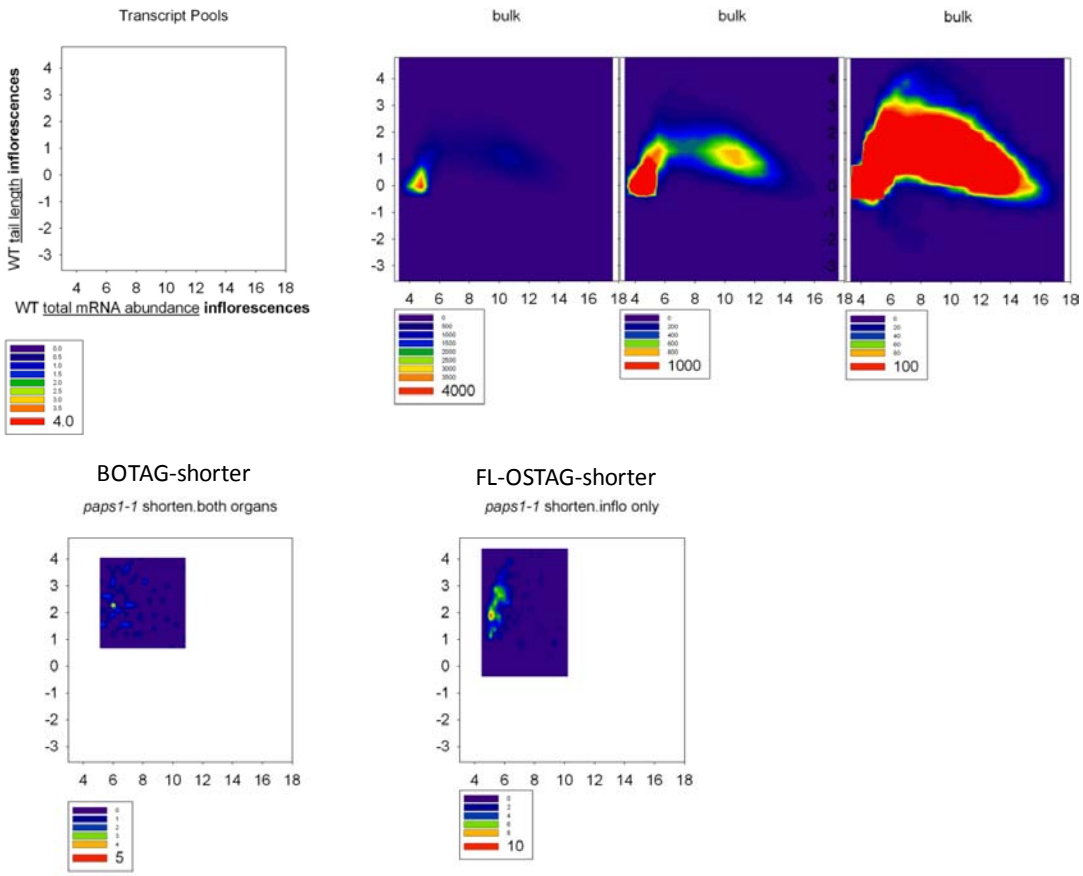


Figure 7.9 2D-contour: WT tail length in inflorescences (inflo.) vs WT total mRNA abundance in inflo. analysing tail-affected transcripts.

See **Box 7.1** for detailed explanation of the different transcript pools used in the analysis. In brief, different pools of transcripts (bulk: all mRNAs and tail-affected transcripts (shorter in inflorescences only and in both tissues) were separated in two dimensions: the WT poly(A)-tail length in inflo. and the WT total mRNA abundance in inflorescences.

WT poly(A)-tail length is log2 of the WT long tail/short tail ratio as determined from the fractionated microarray. Positive value indicate relatively long poly(A) tails. WT total mRNA abundance is log2 of WT expression values extracted from microarray data. The number of genes that occupy an area in the 2D-map was spectrum-colour coded. The number on the right of each colour in the spectrum indicate a threshold, e.g. if the area is occupied by more than 4 genes (the threshold), it will be shown in red.

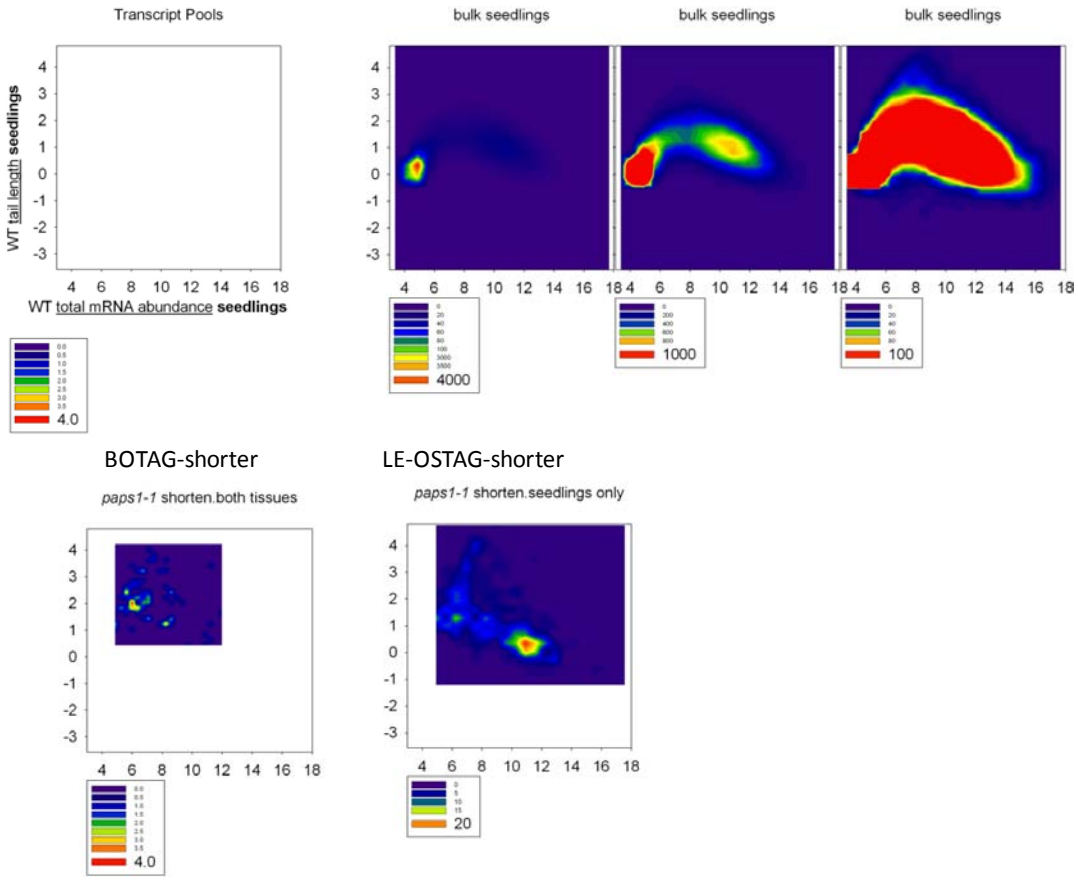


Figure 7.10 2D-contour: WT tail length in seedlings (seedl.) vs WT total mRNA abundance in seedlings analysing tail-affected transcripts.

See **Box 7.1**for detailed explanation of the different transcript pools used in the analysis. In brief, different pools of transcripts (bulk: all mRNAs and tail-affected transcripts (shorter in seedlings only and in both tissues) were separated in two dimensions: the WT poly(A)-tail length in seedlings and the WT total mRNA abundance in seedlings.

WT poly(A)-tail length is log2 of the WT long tail/short tail ratio as determined from the fractionated microarray. Positive value indicate relatively long poly(A) tails. WT total mRNA abundance is log2 of WT expression values extracted from microarray data. The number of genes that occupy an area in the 2D-map was spectrum-colour coded. The number on the right of each colour in the spectrum indicate a threshold, e.g. if the area is occupied by more than 4 genes (the threshold), it will be shown in red.

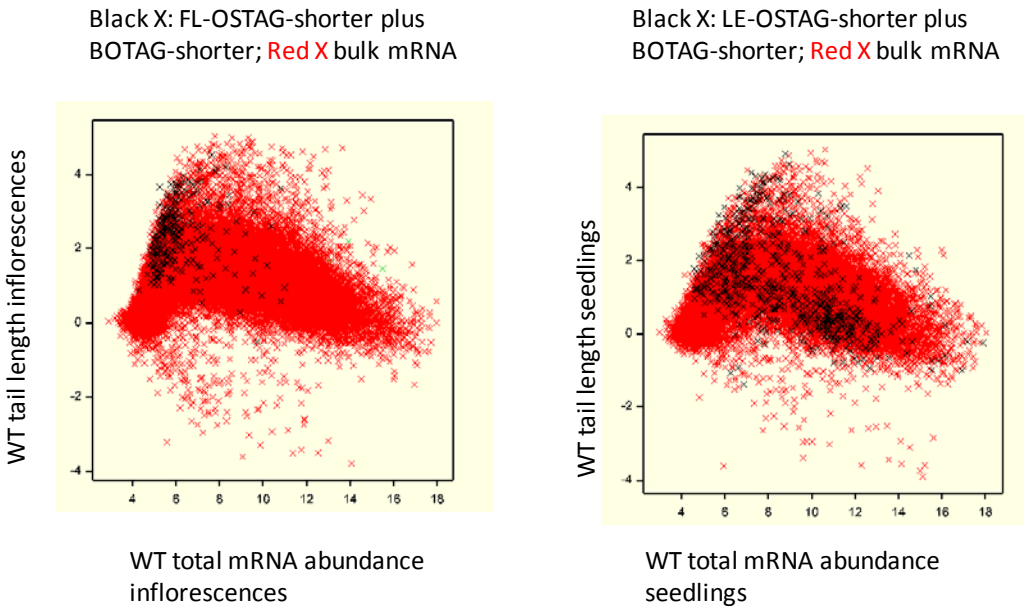


Figure 7.11 2D-map non-contour : separation of genes with tail length vs abundance in seedlings and inflorescences.

Similar to **Figure 7.9** and **Figure 7.10** but here the density in each area is not considered, and the different pools of transcripts were overlappingly drawn on top of each other. Red crosses represent bulk mRNA. Black crosses represent tail-affected transcripts in inflorescences (left) or seedlings (right). See **Box 7.1** for detailed explanation about pools of transcripts.

WT poly(A)-tail length is log2 of the WT long tail/short tail ratio as determined from the fractionated microarray. Positive value indicate relatively long poly(A) tails. WT total mRNA abundance is log2 of WT expression values extracted from microarray data.

known/predicted to be involved in one biological pathway are grouped together, which helps answer whether a DEG is enriched for a particular biological pathway.

Combining both of these programs, we found that the LE-OAASG is enriched with pathogen related genes while FL-OAASG is not (genes lists attached as an electronic supplementation file). In *paps1-1* seedlings, a pathogen response seemed to be ectopically elicited. By contrast, in *paps1-1* inflorescences, this is not the case. This suggests PAPS1 in seedlings resembles a negative regulator of pathogen response.

7.3.5 Other interesting observations from microarray analysis

There was strikingly not much correlation between tail length changes and abundance changes.

Overlap analysis between tail-affected transcripts and abundance-affected transcripts showed that poly(A)-tail length changes seldom correlate with mRNA abundance changes (**Figure 7.12**) This is striking, because it was well believed that poly(A) tail length is important for mRNA stability. The result also suggests that the changes in poly(A)-tail length by *paps1-1* mutants could leads to the instability of some mRNAs but not the others. Furthermore, looking at the extent of tail changes revealed that the distributions of the LFC tail changes (**Figure 7.13**) are similar between the two pools: (the tail-affected but abundance-unaffected mRNAs) vs. (the tail-affected but abundance-downregulated mRNAs). This suggests the degree of tail change cannot be the reason for whether a tail-affected mRNAs is unstable or not. It remains to be answered how the poly(A)-tail length changes in *paps1-1* plants result in different effects on the mRNA stability.

It is also an open question what the consequences of the tail-affected, but abundance-unaffected type of misregulation are for the target mRNAs? Does this type of mis-regulation affect translational effiency or whether does it not have any consequences to the regulation of the mRNAs?

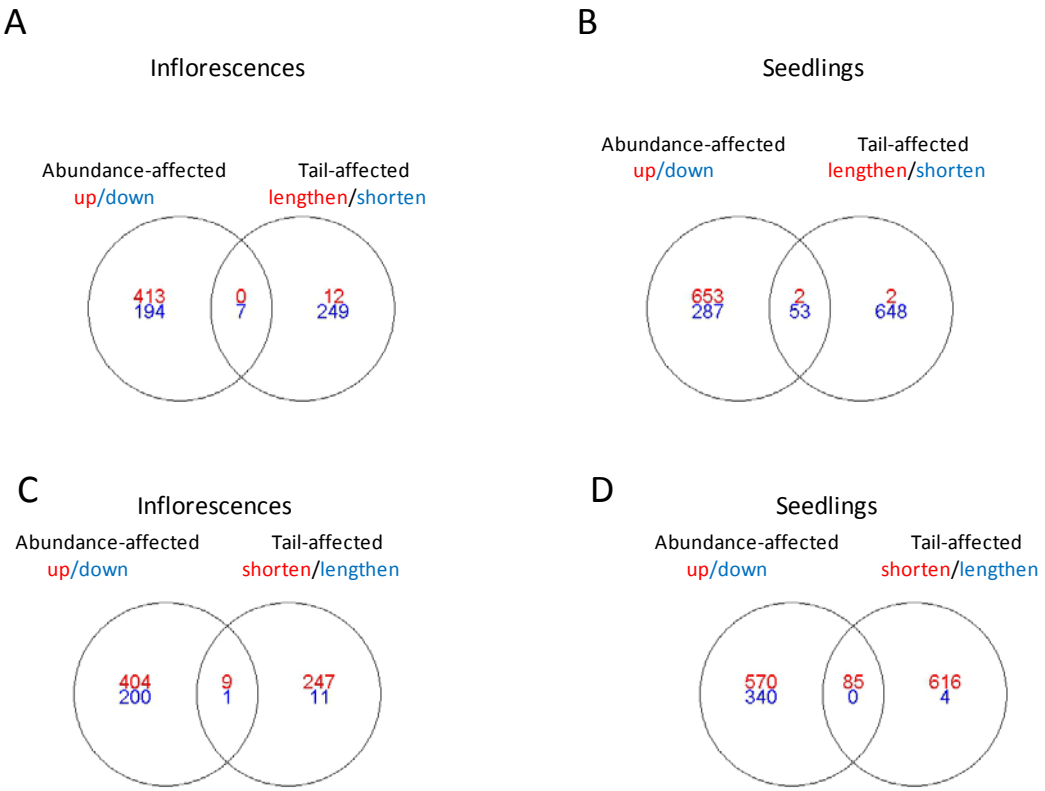


Figure 7.12 Overlap analysis between tail-affected and abundance-affected gene lists in each tissue type.

A and C: inflorescences; B and D: seedlings. A and B normal overlap analysis; C and D inverse overlap-analysis (note that the colour label is reversed compared to A and B for the tail-affected genes).

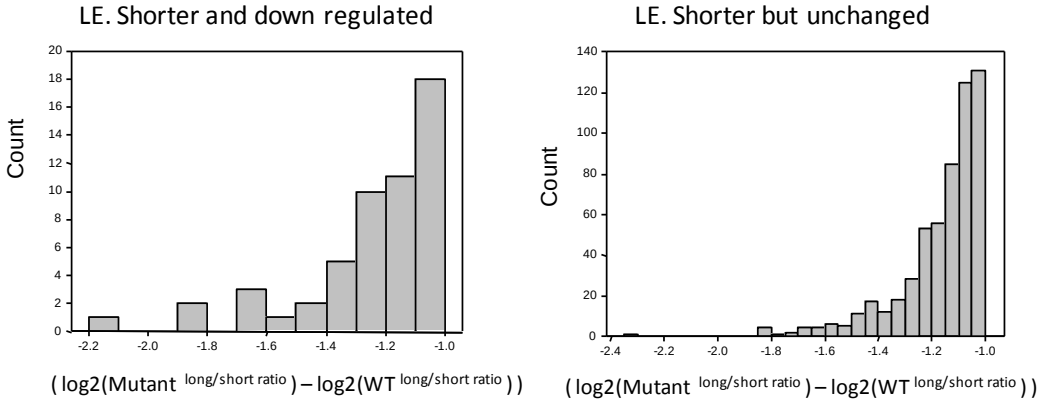


Figure 7.13 The degree of poly(A)-tail shortening in shorter-tail transcripts that were down regulated and transcripts that were unchanged in expression level because of the *paps1-1* mutation.

The degree of poly(A)-tail shortening is represented by (log2(Mutant $^{\text{long/short ratio}}$) – log2(WT $^{\text{long/short ratio}}$)). Left: transcripts that have shortened poly(A) tails and are down regulated in mutant seedlings (LE.) compared to WT seedlings (correspond to the 53 genes in **Figure 7.12**.) Right: transcripts that have shortened poly(A) tails but have the expression level unchanged in the mutant seedlings compared to WT seedlings (correspond to 648 genes in **Figure 7.12B** excluding the 85 genes in **Figure 7.12D**) .

8 Chapter 8. General discussion

8.1 Summary

Here I have characterized a mutant that affects sizes of flowers and leaves in opposite directions: i.e., smaller leaves but bigger flowers than wild-type.

Phenotypically, I showed that the opposite phenotypes are dependent on organ identity rather than organ position. I measured and compared the growth parameters of leaves and petals between mutants and wild-type showing that: i) Mutant leaves are smaller mainly due to reduced cell expansion ii) Mutant petals are larger due to both increased cell proliferation, whose spatial patterns remain unchanged, and increased cell expansion. iii) Mutant petals are larger because the growth period is prolonged while the growth rate is unchanged.

Molecularly, I identified the causal gene that is At1g17980, a gene encodes for PAPS1 protein , one of the four canonical poly(A) polymerases (PAPSs) in *Arabidopsis thaliana.* I isolated and characterized the mutants in the other three *PAPS1*-orthologs in *Arabidopsis* showing that i) The single null mutants of the other PAPS orthologs: *PAPS2* and *PAPS4* do not have a phenotype. ii) The *paps2 paps4* double mutants are late flowering, but otherwise morphologically normal; especially, the sizes of *paps2 paps4* petals and leaves are comparable to wild-type (WT). ii) Mutants in the third ortholog of PAPS1, i.e., *paps3* mutants, displayed a pleiotropic phenotype, i.e *paps3* plants are very small and sterile.

Different *paps1* alleles, which differ in the abundance and integrity of the PAPS1 transcripts have different phenotypes: *paps1-1* and *paps1-4* similarly have smaller leaves and bigger flowers than WT, whereas *paps1-2* have smaller leaves, but deformed and smaller petals than WT. One other allele, *paps1-3*, is most likely gametophytic lethal.

By transgenic experiments, I showed that a chimeric PAPS created by replacing the N-terminal domain of PAPS1 with the N-terminal domain of PAPS4 (the *pPAPS1::gNPAPS4::gCPAPS1*) can fully rescue the size phenotypes in both leaves and flowers of the *paps1-1* mutant. By contrast, promoter swapping alone (*pPAPS1::gPAPS4*) does not rescue *paps1-1* mutant, suggesting that the differences in the C-terminal domain makes different PAPSs functional non-redundant.

In vitro, the mutated protein $PAPS1^{P313S}$ (encoded by the *paps1-1* allele) has almost no activity while the wild-type protein has polyadenylation activity. *In vivo*, however, the bulk poly(A) tail change is very subtle between *paps1-1* mutant and WT.

Using an mRNA fractionation coupled with microarray, I have defined a set of mRNAs whose poly(A)-tails and stability are changed in the mutant compared to wild type. Independently by qPCR analysis, I could verified the abundance-affected transcripts identified in microarray experiments. Subsequent bioinformatic analysis suggests compared to WT, in *paps1-1* mutant seedlings but not *paps1-1* mutant inflorescences, ectopic

pathogen responses are elicited. The tail-affected transcripts identified by fractionation microarray, however, remains to be verified by LM-PAT test.

Based on the results suggested by the microarray, it seems that the affected transcripts composed of only small proportions of the transcriptome; and there is probably not much correlation between tail-length changes and abundance changes. Furthermore, seedlings and inflorescences seemed to have distinct pools of abundance-affected transcripts and perhaps also of tail-affected transcripts.

By genetic analysis, I showed that mutants in the *CstF64*, a component of the 3' end processing complex, do not share phenotype with *paps1-1*. *PAPS1* seems to be additive, albeit not completely, to other size regulators *bb*, *klu* and *ant* mutants.

Using a system to generate predictable *paps1-1* chimerism, I showed that the effect of *paps1-1* in petal growth is cell-autonomous within individual petals, but non cell-autonomous across the epidermis and internal (L2) layer in petals. Additionally, I showed that the *paps1-1* chimerism in the internal layer of a meristem caused the meristem to split.

Taken together, this study uncover an potentially additional layer of gene regulation: a transcript-specific synthesis-dependent regulation of poly(A) tail length by different canonical PAP isoforms. Some transcripts that are more sensitive to the defect in this regulation can serve as size regulators.

8.2 Opposite size regulations in leaves and flowers are caused by different, potentially novel pathways.

The opposite phenotypes in leaves and flowers of *paps1-1* mutants can be caused by the opposite regulation of the same size regulator or by affecting two independent size regulators. The results here support the later scenario. The results of microarray analysis, performed using total RNA, revealed that very few genes were oppositely regulated in the two organs (i.e. upregulated in seedlings but downregulated in inflorescences and vice versa). Moreover, microarray analysis also showed that all the known size regulators are not misregulated in the *paps1-1* mutant. Genetic interactions showed that the petal enlargement is largely additive to *BB*, *KLUH* and *ANT*. Therefore, the affected size regulating pathways in leaves and flowers are probably distinct and potentially novel.

8.3 What could be the reason for the identity dependent effect of *paps1-1* mutation?

Organ identity (leaf vs. flowers) influence the the effects of *paps1-1* mutation and causes several differences found in these two organs: i) phenotypically: size; ii) molecularly: mRNAs that are sensitive to polyadenylation defects.

8.3.1 Organ identity dependent size regulation in *paps1-1* mutants:

paps1-1 mutation promotes petal growth but suppresses leaf growth. Upregulated genes in the mutant seedlings are enriched for defense-related genes, whose abundance remains largely unchanged in the mutant inflorescences. It is well known that when a pathogen response is elicited, growth is inhibited (Chinchilla et al., 2007).

Therefore, the data suggest that *paps1-1* leaves are smaller because of the upregulation of pathogen-defense related genes. By contrast, in petals, the defense pathway is not upregulated, probably allowing petals to respond to some yet unknown hypothetical size regulator 'X', which might be misregulated in petals due to the *paps1-1* mutation.

A remaining question is whether wild-type inflorescences also innately do not response to pathogen attack or only *paps1-1* inflorescences do not mount an ectopic pathogen attack like the leaves. The future experiments are : i) to suppress pathogen response in *paps1-1* seedlings by combining *paps1-1* with some main pathogen-reponse mutants (e.g., *pr1*, *npr1*, *sid2-1*, *35S::NahG* etc.) and see whether this rescues the leaf-size defect; ii) to clarify how well the WT inflorescences and mutant inflorescences respond to pathogen attack.

Recently, (Winter et al., 2011) showed that when the plants enter reproductive phase, *LEAFY* is expressed and down-regulate genes that are involved in pathogen-reponse. Hence it seems likely that inflorescences of WT innately have a low response to pathogen attack.

8.3.2 Organ identity dependent regulation of the sensitivity of pre-mRNAs to the *paps1-1* defect:

Microarray analysis clearly showed that the pools of abundance-affected (and potentially also tail-affected) transcripts are largely non-overlapped between the two organs. More importantly this specificity is not the results of these genes being expressed in only one organ but not the other. It is also because *PAPS1* is expressed restrictly in one organ, either. Hence, organ identity must influence the sensitivity of some transcripts to *paps1-1*

defect. One potential explanation is that organ identity determines the preference of some mRNAs to be processed by one PAPS complex or the other complexes, hence making some mRNAs more sensitive to *PAPS1* defects than the other mRNAs. See further discussion on section 8.4

8.4 Why are some transcripts more sensitive to *paps1-1* defect than others?

Microarray analysis on mRNAs fractionated according to their poly(A)-tail length suggested that there is a subclass of mRNAs (ca. 1000 transcripts) that are more sensitive to polyadenylation defect.

A plausible explanation for this sensitivity is that the different PAPSs that exist in plants may have different substrate specificity or different kinetics. In the first scenario, PAPS1-sensitive transcripts would be preferentially polyadenylated by PAPS1. This is surprising because PAPS has no substrate specificity *in vitro* on its own. Additionally, polyadenylation happens after the cleavage of pre-mRNA so CPSF should be the one which interacts with pre-mRNA first and therefore is more likely to determine the substract specificity of the 3'end formation reaction. So probably PAPS1 has substrate specificity by inheriting the substrate specificity of the protein complex it is in.

By contrast, in the second scenario there need not be substrate specificity; instead, all mRNAs are processed by all PAPS complexes, yet sensitivity would apply only to a subclass of transcripts that somehow require a particularly high polyadenylation activity. Results in chapter 7 showed that tail-affected mRNAs in inflorescences are enriched for ones with long poly(A) tails but low expression levels. So, one possibility is that these mRNAs are unstable, they undergo rapid transcription and degradation, hence they requires more polyadenylation than the others; at the same time PAPS1 is perhaps the most active PAPS. This possibility is consistent with the result that *paps2* and *paps4* single and even *paps2 paps4* double mutants showed only a very mild phenotype.

The two scenarios have a fundamental difference: one strongly implies there is a substrate preference, the other does not. Experiments to distinguish these two hypotheses require changing PAPS enzyme kinetics without affecting substrate specificity. One experiment could be to overexpress PAPS1 in *paps2 paps4* double mutants, and vice versa, to overexpress PAPS2 or PAPS4 in *paps1-1* mutants to see if there is any rescue phenotypically (organ size in case of *paps1-1* and flowering time in case of *paps2 paps4*) and molecularly (A-tail change). Other experiments to prove substrate preference can be to show that PAPS1-sensitive mRNAs associated more strongly with PAPS1 complexes than other PAPS complexes by RNA immunoprecipitation using PAPS antibody. A third

experiment is to identify transcripts that are PAPS2 and PAPS4 sensitive and check whether they overlap with PAPS1 sensitive transcripts. For example, if the most highly PAPS1-sensitive transcripts are also affected in *paps2 paps4* double mutant, albeit to a lesser degree, this would suggest that the sensitivity depends on how high the overall requirement for polyadenylation (by all PAPS) is for that transcript rather than the substrate specificity.

Future experiments to futher characterise PAPS1-sensitive transcripts are, for example, to determine if there are any *cis*-elements, likely in the *3' UTR* of the tail-affected transcripts, that encodes for the sensitivity to *PAPS1* using ethanol an inducible *GUS/YFP-3'UTR* reporter gene system.

8.5 The functional specialization amongst PAPSs

The fact that mutations in different *PAPS* genes have different phenotypes suggests that they have specific functions. At the protein level, I showed that the C-terminal domain is the specificity determinant between PAPS1 and PAPS4 proteins. A possible explanation is that by modifications/protein interaction at the CTD, plants can change the activity of PAPS or the recruiment of PAPS to hypothetical different PAPS complexes in plants thereby making PAPSs functionally different.

Literature have evidence to support this hypothesis. Manzano et al., 2009 pulled down FY, a componenent of the 3'end complex, complexes (using FY-C terminal TAP tag construct) and reported two distinct complexes at different sizes (Manzano et al., 2009).C-terminal domain of mammalian PAPOLA can be sumoylated or acetylated (Vethantham et al., 2008) (Shimazu et al., 2007). (see also section1.3.2.4)

The next experiments are to find the interactors of the C-terminal domain by yeast two hybrid or *in planta* pull-down using antibody against C-terminal domain.

8.6 An additional layer of gene regulations: PAPS-dependent transcript-specific *de novo* regulation of poly(A) tail length.

It is a widely accepted view that canonical PAPs are non-specific and add a poly(A) tail of the same length to every pre-mRNA (see section 1.4.1.1). The results here suggest that different plant canonical PAPSs are functionally distinct and perhaps by using different PAPSs, the *de novo* poly(A) tail synthesis are different for each transcript.

Moreover, *in silico* results suggest that *PAPS1* expression is specifically induced by pathogen attack; loss of *PAPS1* function causes upregulation of pathogen-related genes in

seedlings. Together, these results strongly suggest that plants actively use *PAPS1* as a break to prevent plants from overreacting to a pathogen, as an unchecked defense response will result in serve growth retardation. The expression of *PAPS2* and *PAPS4*, by contrast, are up-regulated during anoxia treatment. It is therefore tempting to hypothesize that for a certain stimulus, plants can respond not only by transcriptional changes, but also by altering the poly(A)-tail length patterns of specific transcripts, which consequently changes the post-transcriptional regulation of the mRNA. Given that different PAPSs are known to be also regulated post post-transcriptionally by alternative splicing and protein modifications at the C-terminal domain, the potential to modify PAPS activity in response to stimuli will be diverse.

The findings here that different mRNAs respond differently to polyadenylation defects opens up questions about how poly(A)-tail length patterns can be used for the regulation of genes, and why for some stimuli/some genes, the regulation at the poly(A)-tail level seems to employed. One scenario I imagine is that by choosing different poly(A) sites, plants can employ different PAPS complex which will give the pre-mRNA a different poly(A)-tail length. More than half of plant genes have more than one poly(A) sites, whose biological functions are largely unknown. Especially for those alternative poly(A) sites that are only 15-30 base pairs apart and do not have any binding sites for miRNAs in between, it is difficult to imagine what their functions could be. Much research on alternative polyadenylation (APA) using next-generation sequencing only focuses on the sequence of the APA, but none of the studies checked the poly(A)-tail length of these spliced forms. It is possible that in some cases, the sequences of the 3' UTR are not important; the important point may be that by choosing a different poly(A) site, the mRNA will have a different poly(A)-tail length pattern. Because the APA can be as small as 10 nucleotides apart, which can be very difficult to detect, this kind of regulation will be very likely missed out in conventional poly(A) tail tests, unless the 3' ends are subcloned and sequenced. Finally, one question that will be interesting to address is that how does this potential novel mode of gene regulation via poly(A)-tail intergrated with the promoter-driven transcriptional regulation. For example, for the PAPS1-dependent transcripts that change both their tail length and abundance, can we find any biological conditions where the abundance change is not coupled with the poly(A)-tail change ?

In sum, the most interesting questions to address in the future is why some transcripts are sensitive to polyadenylation defects and others not, and how different patterns of poly(A)-tail lengths can contribute to gene regulation.

8.7 Relationship between Poly(A) polymerases and Deadenylase CCR4.

There are three enzyme complexes that specifically degrade/shortens poly(A) tails: PAN, CCR4 and PARN. Interestingly, the *ccr4* mutant in Arabidopsis is also more sensitive to pathogen. *PR1* expression is downregulated and its mRNA has a shorten poly(A) tail in *ccr4* compared to WT(Liang et al., 2009). *PR1* is upregulated in *paps1-1* mutant. These two results suggests that the *PR1* pre-mRNA may have their poly(A) tail length tightly regulated by *PAPS1* and *CCR4*.

8.8 The cell-autonomous effect of *paps1-1* mutation in petal enlargement

So far the autonomy/non-autonomy has been studied with several size regulators: *KLUH* (non cell-autonomous)(Eriksson et al., 2010); *AN3* (cell non-autonomous)(Kawade et al., 2010); *BRI1* (non cell-autonomous) (Savaldi-Goldstein et al., 2007);. The autonomous activity of *PAPS1* suggests that the synchronised growth of different parts within an organ can be overridden.

8.9 *paps1-1* chimerism in the internal layer of a meristem induces the meristem to split.

The results in chapter 5 suggest that the presence of *paps1-1* mutant and wild-type tissues in the internal layer of a shoot/inflorescence meristem induces the meristem to split into two or more meristems. Two questions arise from these results.

Firstly, why do only the meristems that are chimeric for the *paps1-1* mutant, but not the fully *paps1-1* mutant meristems, split? One potential cause is that the *paps1-1* stem cells divide asynchronously compared with the neighbouring wild-type stem cells; this might cause the meristem to split.

The stem cell population must be under tight control to maintain a certain size, despite the fact that stem cells do divide. To maintain a constant number of stem cells, when the stem cells divide in wild-type plants, on average one of the daughters of stem cells must lose its fate, supposedly because it is displaced out of the location where stem cells can receive the *WUS*-mediated stem-cell inducing signal. In the *paps1-1* chimeric meristem, either both of the stem-cell daughters may maintain their stem-cell identity thus leading to fasciation, or somehow the stem-cell neighbouring cells ectopically acquire a stem cell fate. This

possibility would also suggest that the difference in growth caused by the *paps1-1* mutation acts at a single-cell level and also in stem cells, that is long before organ initiation. It is not yet known whether in WT all stem cells divide synchronously or not (live image can only track the stem cells in the epidermis (Reddy et al., 2004) but the study did not investigate this question). One other potential cause for the meristem split is that the *paps1-1* mutation may cause a reduction *CLV* signalling (in fact, total RNA microarray indicates that *CLV2* expression is reduced mildly (1.5 fold) in *paps1-1* mutant compared to wild-type). If all the stem cells have reduced activity of *CLV*, there may be no problem, but if stem cells have unequal *CLV* signaling amongst their population, the meristem might split. To test these two hypotheses, plants with chimeric meristem for *CLV3/CLV2* and also chimeric for different size regulators like *DA1* or *ANT* should be tested to see whether this can also cause meristem spliting. It will also be useful to do *in situ* hybridization with probes against *CLV3* and *WUS* in the split meristems.

Secondly, why do only the internal *paps1-1* chimeras but not the external *paps1-1* chimera induce the meristem splitting? Does this internal-layer-chimerism inducing split suggest a more important/direct role of the internal tissue in stem cell homeostasis?

8.10 Using the meristem splitting to increase crop yield

This way of using chimerism in the meristem to induce a meristem to split may have a potential to change the architecture of the plant for agriculture purposes. Inflorescence meristem splitting increases the number of inflorescences without the need of making new axillary leaves. Hence, one can increase the harvest index (ratio of harvestable material over total mass).

8.11 The modifier screen on mutagenized *paps1-4* mutant and the quest for the novel size regulator in flowers.

The *trans*-heterozygous *paps1-4/paps1-2* has bigger flowers than *paps1-4* mutants suggesting that the misregulation of the flower size can be enhanced beyond the level seen in *paps1-4*. One hypothesis is that in *paps1-4* flowers, there are size repressor(s) 'X' that are downregulated partially; if so 'X' null mutants would make even bigger flowers in a *paps1-4* background. Therefore, I mutagenized *paps1-4* by EMS and screen the M2 population for modifiers of petal size. At the same time, this screen can also identify modifiers that act together with PAPS1 to regulate the polyadenylation of pre-mRNAs. We identified a number of mutants that have bigger flowers than *paps1-4* and one mutant with deformed flowers, which resembles *paps1-2*. Future experiments involve crossing these

mutant out to WT to test whether the mutated genes have a genetic interaction with *PAPS1*, in which case they will be studied further. This will be recognized by comparing the novel single mutant to the double mutant with *paps1-4* and looking for more or less than additive effects in the double mutant.

8.12 Hypothesis to explain the opposite size-regulation in petals of different *paps1* alleles and opposite phenotype between leaves and flowers:

The hypothesis is based on four observations stemmed from the analysis of different *paps1* alleles and of *paps1-1* at increasing temperature: i) *paps1-1* phenotype changes from bigger flowers at 23^0C to smaller/deformed petals at 28^0C. ii) *paps1-2*, homozygous of which gives small/deformed petals but *trans*-heterozyogus of which with *paps1-4* gives bigger petals (even bigger than *paps1-4* homozygous mutant). iii) Leaves of *paps1-1*, in contrast to WT leaves' behavior at high temperature, showed a dramatic reduction in size in response to high temperature. iv) *paps1-3*, the most likely null *paps1* allele*,* is probably gametophytic lethal.

These four observations allow us to make three assumptions: i) The remaining PAPS activity in different mutant alleles is ranked in this descending order: (**Figure 8.1**): WT (100%), *paps1-4*, *paps1-1* at 23^0C, *paps1-1* at 28^0C or *paps1-2*, *paps1-3* (0%); ii) There are two classes of mRNAs with different sensitivities to PAPS1 defect. One class is highly sensitive to PAPS1 defect. A reduction in PAPS1 at *paps1-1* level is able to cause them to be mis-polyadenylated or mis-regulated. The other class is less sensitive such that at the PAPS1 reduction level in *paps1-1* mutants, they are not affected; they are only mis-regulated after PAPS1 activity is further reduced. This assumption is supported by the microarray data showing that there are 1% of the transcriptome are highly sensitive to *paps1-1* defect; iii) There are negative size regulator(s) X, which belongs to the highly sensitive class, and gets down-regulated by the *paps1-1* defect. There are positive regulator(s) Y (or genes that are essential for growth) which belongs to the less sensitive class and are not affected/or only mildly affected by *paps1-1* defect.

With these three assumptions, we can explain why some *paps1* alleles cause big flowers and some cause small flowers. For example, in *paps1-2*, the remaining PAPS1 activity is below the limit for the less sensitive classes, hence many genes that are essential for growth (genes Y) are affected. This in turns causes growth depression and smaller/deforms petals. Although in *paps1-2*, target X is also mis-regulated but its growth-promoting effect is masked by the mis-regulations of Ys. This epitasis effect is lifted in the trans-heterozygous of *paps1-2/paps1-4*. In this genetic background, or in the *paps1-1* mutants, the remaining PAPS1 activity is at the level such that Y mRNAs become un-affected while X mRNAs are

more mis-regulated than in *paps1-4* mutants. The final out come is that *paps1-2/paps1-4* trans-heterozygous plants form normally-shaped petals which are larger than *paps1-4* homozygous petals.

This hypothesis could also be used to explain why leaves are smaller in all *paps1* alleles as a second hypothesis in parallele with, and not mutually exclusive with the hypothesis in section 8.3.1. Perhaps, the sensitivity of both classes of transcript increases in leaves compared to petals because somehow in leaves PAPS1 is the more important PAPS. This results in the growth depression in leaves of all *paps1* alleles.

The hypothesis predicts for many of the PAPS1-dependent transcripts, the degree of mis-regulation is ranked such as the strongest mis-regulation is with *paps1-2* and mildest effect is with *paps1-4*. We have some preliminary qPCR data on expression level of several PAPS1-dependent transcripts in WT and different *paps1* mutant alleles that supports this hypothesis.

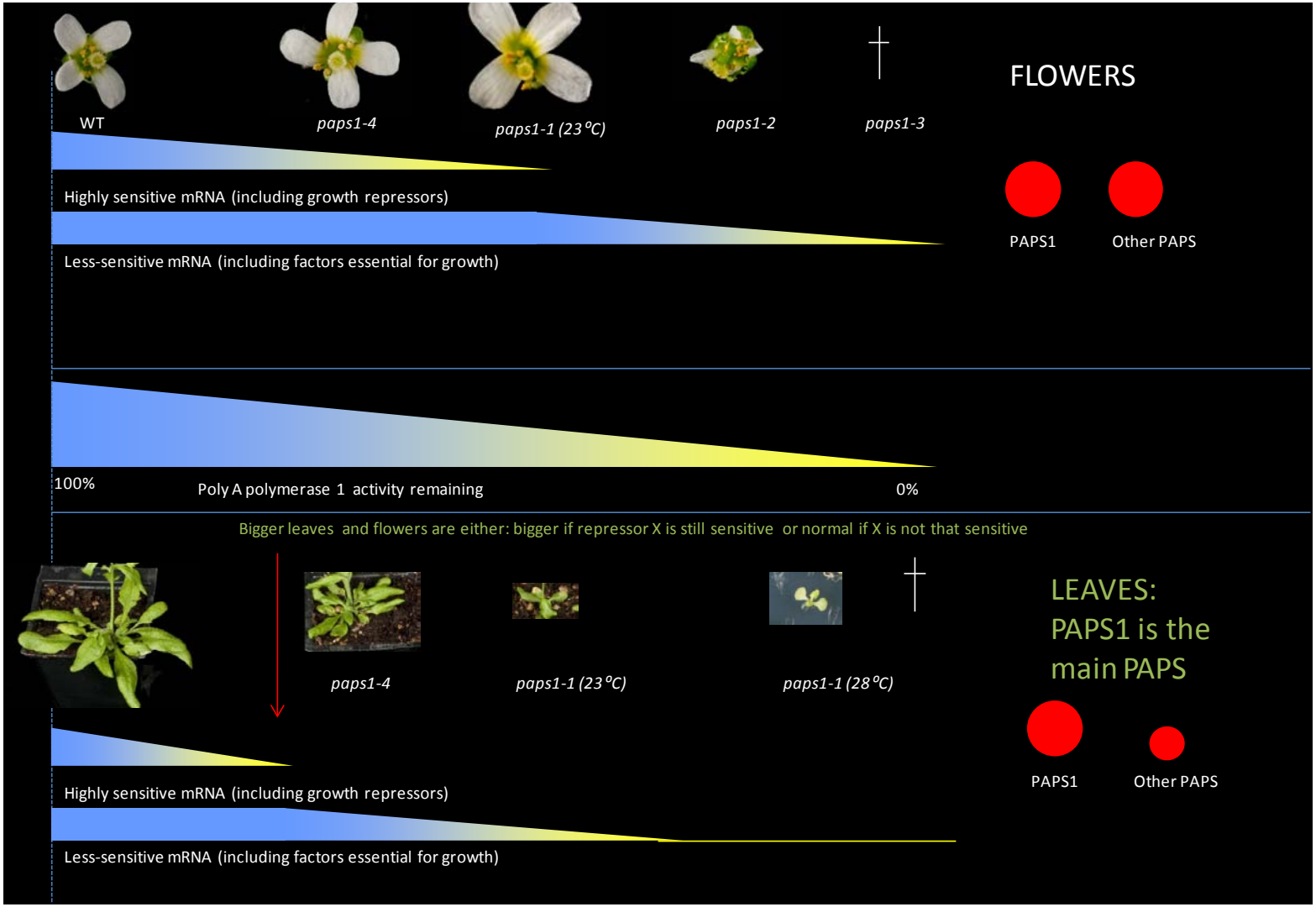


Figure 8.1 Hypothesis to explain for the differences in flower phenotypes of *paps1* alleles

9 Chapter 9. Materials and methods

9.1 Materials

9.1.1 Plant materials

The Landsberg *erecta* (L*er*) background was used as wild-type background for experiments involving mutants that are also in L*er* background including *paps1-1, paps3-1* (*T-DNA* line GT_5_7040), *klu-2, bb-1, ant^{7F5}, cstF64-1.* L*er* is a derivative of the Landsberg (La-0) accession that is mutated at the *ERECTA* locus (Redei, 1962). The *bb-1* (*bb*) and *klu-2* background alleles used in this study are described in Disch et al. (2006) and Anastasiou et al. (2007). The *ant^{72F5}* mutant was kindly provided by Kai Schneitz (University of Zürich, Switzerland). This allele displays a very strong ovule phenotype similar to the putative null allele *ant-1* (Klucher et al., 1996).

T-DNA insertion lines or *T-DNA* lines were obtained from NASC. Other mutants including *paps1-2* (*T-DNA* line *SAIL_172_F11)*, *paps1-3* (*T-DNA* insertion line *WiscDsLox413-416L14)*, *paps1-4* (*T-DNA* line *WiscDsLox441G5)*, *paps2-3* (*T-DNA* line *SALK_126395*), *paps4-3* (*T-DNA* line *SALK_007979*), *paps3-3* (*T-DNA* line *SALK_133557*), *paps3-4* (*T-DNA* line *SALK_133558*), are in the Columbia-0 (Col-0) background and therefore Col-0 was used as wild-type control for these plants. *35S::NahG* (Col-0 background) is obtained from Cyril Zipfel lab. For other plants, the background is specified in the text or in the seed stock lists (attached as electronic version).

cstf64-1 is kindly provided by Caroline Dean (Liu et al., 2010).

9.1.2 Bacterial and yeast strains

E. coli strain XL1-Blue (Stratagene, La Jolla, USA) was used for the cloning and propagation of plasmids. *Agrobacterium tumefaciens* (*Agrobacterium*) strain GV3101 ((Van Larebeke et al., 1974) was used for plant transformations.

9.1.3 Growth media

Compositions of bacteria, yeast and plant tissue culture media are described in Table 9.1**.**

9.1.4 Selective antibiotics and herbicides

The final concentrations of antibiotics and herbicides used in this study are listed in **Table 9.2**

Media	Composition
Bacterial growth media	
Lysogeny broth (LB) medium	10 g/l Bacto-tryptone, 5 g/l bacto-yeast extract, 10 g/l NaCl, pH was adjusted to 7.0, for solid medium 1.5% (w/v) Bacto agar (Difco) was added.
Super Optimal broth with Catabolite repression (SOC)	20 g/l Bacto-tryptone, 5 g/l bacto-yeast extract, 0.5 g/l NaCl, 2.5 mM KCl, pH was adjusted to 7.0, medium was autoclaved prior to addition of glucose to a final concentration of 20 mM and $MgCl_2$ to a final concentration of 2 mM. This media is used for incubation of bacterial after transformation by electroporation.
Plant growth medium	
½ MS medium	2.15 g/l Murashige and Skoog (MS) plant salt mixture (ForMedium, Hunstanton, UK), for solid medium 0.8% (w/v) Bacto agar (Difco) was added.

Table 9.1 Media used for bacteria and plant tissue cultures

Antibiotic/herbicide	Dissolved in	Final concentration
Selection of Bacteria		
Carbenicillin (Carb)	ddH_2O	50 µg/ml
Chloramphenicol (Cam)	ethanol	30 µg/ml
Gentamycin (Gent)	ddH_2O	25 µg/ml
Kanamycin (Kan)	ddH_2O	50 µg/ml
Spectinomycin (Spec)	ddH_2O	40 µg/ml
Tetracycline (Tet)	ethanol	10 µg/ml
Selection of Arabidopsis thaliana		
Basta	H_2O	0.1% (v/v)
Kanamycin (Kan)	ddH_2O	50 µg/ml
Phosphinothricin (PPT)	ddH_2O	10 µg/ml

Table 9.2 Concentrations of antibiotics and herbicides for selection of bacteria and *Arabidopsis*.

Antibiotic were made as 1000x concentration then aliquots were stored at -20°C.

9.1.5 Chemicals

Unless stated otherwise, chemicals were purchased from Sigma-Aldrich Ltd (Haverhill, UK), Melford Laboratories Ltd (Ipswich, UK) or Fisher Scientific (Loughborough, UK).

9.1.6 Enzymes

Restriction endonucleases were purchased from New England Biolabs Ltd (NEB, Ipswich, UK), Invitrogen Ltd (Paisley, UK) or Roche (Burgess Hill, UK). Modifying enzymes were purchased from Roche or NEB. Taq polymerase was purchased from Roche or Promega Ltd (Southampton, UK).

9.1.7 Oligonucleotides

Oligonucleotides used as primers for standard polymerase chain reaction (PCR) and sequencing reactions were ordered from Sigma-Aldrich Ltd (Haverhill, UK) and resuspended in sterile ddH_2O to obtain a 100 µM stock solution which is then stored at -20°C. Oligonucleotides used in this study are listed in APPENDIX A.

For RT-PCR primers in **Figure 3.8**: qPCR primers for *PAPS1* are oSV187 + oSV188. Primers upstream *paps1-2* are oSV304 + oSV305, primers flanking *paps1-2* are oSV304 + oSV100, primers downstream *paps1-2* are oSV76 + oSV100. Primers upstream *paps1-4* are oSV187 + oSV188, primers flanking *paps1-4* are oSV186 +oSV188, primers downstream *paps1-4* are oSV186 + oSV197. Primers for RT-PCR *PDF2* are ML 179 + ML180. Primers for qPCR of *UBC* are oSV167 + oSV168. Primers for qPCR of *PDF2* are oHB86+oHB87.

For RT-PCR primers in **Figure 4.4A**: Primers flanking *paps2-3* are oSV198 + oSV121. Primers flanking paps4-3 are oSV110 + oSV112. Primers for RT-PCR of *PDF2* are ML179 + ML180.

Primers for qPCR validation in **Figure 7.5**: At1g54020 (oSV205 and oSV206); At4g03156 (oSV242 and oSV243); At2g35640 (oSV232 and 233); At4g04580 (oSV211 and oSV212); At4g23160 (oSV213 and oSV214); At5g59680 (oSV215 and oSV216); At4g11890 (oSV234 and oSV235); At4g25100 (oSV236 and oSV237).

9.1.8 Plasmids

All plasmids used and generated in this study are listed in APPENDIX B. Detailed procedures on constructing these plasmids are listed in APPENDIX C.

9.1.9 Stock solutions

General stock solutions are lists below

TE	10 mM Tris-HCl pH 8.0, 1 mM EDTA
50x TAE	40 mM Tris base, 40 mM acetic acid, 1 mM EDTA
1x PBS	137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM KH_2PO_4
10% (w/v) SDS	100 g sodium dodecyl sulphate per litre dH_2O

Other specific solutions for each protocol are listed within the protocol.

9.2 Methods

9.2.1 Molecular biology - DNA-related methods

9.2.1.1 Preparation of electro-competent *E. coli* and *Agrobacterium* cells

The protocol below is for E.coli. For Agrobacterium, the volumes and growing temperature are given in parentheses. 50 ml LB-broth containing the appropriated antibiotic were inoculated with a single colony of *E. coli* strain XL1-Blue (*Agrobacterium* strain GV1301) and incubated overnight at 37 (28)°C with shaking at 200 rpm. 4 ml of this overnight culture were used to inoculate 400 ml of prewarm LB-broth (the OD at this stage of the whole culture was about 0.1). The culture was then grown to an OD_{600} of 0.4-0.5 (about 2 h for *E. coli* and 3-4 h for *Agrobacterium*) at 37 (28)°C with shaking and then placed on ice for 15 min. The following steps were carried out on ice or in the cold room and all centrifugation is at 4°C. The cells were harvested by centrifuging at 4.600 x g for 10 min The supernatant was discarded and the cells were gently resuspended in 200 ml of sterile ddH_2O (prechilled to 4°C) . The cells were centrifuged again at 4.600 x g for 10 min. The supernatant was discarded and the pellet was resuspended gently 200 ml of chilled sterile ddH_2O. After centrifugation at 4.600 x g for 10 min, the supernatant was discarded and the cells were gently resuspended in 10 ml chilled sterile 10% (v/v) glycerol. The cell suspension was transferred to a sterile 50 ml tube and centrifuged at 4.000 rpm for 10 min. The supernatant was removed and the cells were resuspended in 1 ml chilled sterile 10% (v/v) glycerol. Cells were aliquots in 40 (20) µl, frozen immediately in liquid nitrogen and stored at -80°C.

9.2.1.2 Transformation of *E. coli* and *Agrobacterium* cells

Growing temperature and times for *Agrobacterium* are given in parentheses.

Aliquots of electro-competent cells were thawed on ice for about 5 min, 1-4 µl ligation or 0.5 µl plasmid DNA were added and the cell suspension was transferred to a Bio-Rad 2 mm cuvette (prechilled on ice). The electroporation was carried out at 2.5 kV, 200 Ω resistance and 25 µF capacitance. Then 1 ml SOC medium was added to the cuvette and the cell suspension was transferred to a 1.5 ml tube. For transforming with a Amp^R plasmid, 0.2ml SOC medium was used. The cells were incubated at 37 (28)ºC with shaking at 200 rpm for 1 (2) h. This step is ommited if an Amp^R plasmid was used for transformtion. The cells were centrifuged in a microcentrifuge briefly and most of the supernatant was discarded. The remaining supernatant (200µl) was used to resuspend the cells and then spread on LB-agar plates containing the appropriate antibiotics. The plates were incubated at 37 (28)ºC overnight (48h).

9.2.1.3 Frozen bacterial stocks

600 µl of overnight culture was mixed with 400 µl of sterile 60% (v/v in water) Glycerol. The mixture was frozen immediately with liquid Nitrogen and stored at -80 ºC

9.2.1.4 Alkaline lysis miniprep of plasmid DNA from *E. coli*

Extraction buffer I	50 mM Tris-HCl pH 8.0, 10 mM EDTA
Extraction buffer II	1% (w/v) SDS, 0.2 M NaOH
Extraction buffer III	3.1 M KOAc pH 5.5

A single colony of *E. coli* was used to inoculate 5 ml LB-broth containing the appropriate antibiotic overnight at 37°C with shaking (200 rpm). The following procedure was carried out at room temperature. 2 ml of the overnight culture were transferred to a 2 ml tube and centrifuged at 13.000 rpm for 1 min. The supernatant was discarded, 150 µl extraction buffer I were added and vortexed to resuspend the cells completely. After adding 150 µl extraction buffer II, the tube was gently mixed by inverting. 200 µl extraction buffer III was added and the tube was again gently inverted 2-4 times. The sample was centrifuged at 13.000 rpm for 10 min. The supernatant was transferred to a fresh tube containing 500 µl isopropanol and mixed by inverting, then centrifuged at 13.000 rpm for 15 min. The pellet was washed with 70% (v/v) ethanol, air-dried and resuspended in 50 µl TE containing 0.1 mg/ml RNase A. For sequencing, the plasmids were resuspended in ddH_2O.

For low-copy number plasmid (the pBarMAP derivatives), a modification of this miniprep was used. 6 ml overnight culture was used with twice the amount of the solutions and the plasmid was finally elute with 20 µl TE.

9.2.1.5 Alkaline lysis midiprep of plasmid DNA from *E. coli*

Extraction buffer I 50 mM Tris-HCl pH 8.0, 10 mM EDTA

Extraction buffer II 1% (w/v) SDS, 0.2 M NaOH

Extraction buffer III 3.1 M KOAc pH 5.5

50 ml LB-broth containing the appropriate antibiotic were inoculated with a single colony of *E. coli* and incubated overnight at 37°C with shaking (200 rpm). The next day the overnight culture was transferred to a sterile 50 ml tube and centrifuged at 7500 rpm for 3 min (Sigma 4-15C, rotor: 12196-H). The supernatant was discarded, 5 ml extraction buffer I were added and vortexed thoroughly. After adding 5 ml extraction buffer II, the tube was gently mixed by inverting. 7 ml extraction buffer III were added and the tube was gently inverted 2-4 times. The sample was centrifuged at 10.000 rpm for 10 min. The supernatant was transferred to a fresh tube containing 17 ml isopropanol and inverted. The sample was centrifuged at 10.000 rpm for 15 min. The pellet was washed with 70% (v/v) ethanol, centrifuged for 5 min at 10.000 rpm and air-dried. The pellet was resuspended in 200 µl TE containing 0.1 mg/ml RNase A.

9.2.1.6 QIAprep Spin Miniprep Kit

For highquality DNA, plasmid DNA from bacterial cultures grown overnight was isolated using QIAprep Spin Miniprep Kit (QIAGEN Ltd #27104, Surrey, UK) according to the manufacturer's instructions.

9.2.1.7 Preparation of plasmid DNA from *Agrobacterium*

10 ml LB-broth containing the appropriate antibiotics were inoculated with a single colony and the culture was grown for 16-24 h at 28°C. 2 ml culture were transferred to a 1.5 ml tube and centrifuged in a microcentrifuge for 1 min at 13.000 rpm. The following steps were performed according to the manufacturer's instructions for the QIAprep Spin Miniprep Kit (QIAGEN) using 2x the volumes of buffer P1, P2 and N3. The plasmid DNA was eluted with 20 µl EB. 17.5 µl eluate were used in a restriction digest and the restriction pattern was compared to respective *E. coli* plasmid DNA.

For low copy number plasmid (pBarMAP derivatives) 10ml of overnight culture was used or the midiprep was used with a modification: A pinch lysozyme was added to the extraction buffer I, the mixture was incubated 5 min room temperature before extraction buffer II was added.

9.2.1.8 High-throughput plant DNA preparation method, to use in PCR applications for fragments up to 1 kb

Sucrose solution 50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 300 mM sucrose

The solution was newly prepared and used immediately for every experiment.

A piece of leaf (ca.0.5cm x 0.5cm) was collected into one of the collection microtubes (QIAGEN #19560) and one 3 mm tungsten carbide bead (QIAGEN #69997) was added to each tube. 300 µl sucrose solution were added to each tube and the tubes were sealed with collection microtube caps (QIAGEN #19566). The tissue was ground with 1200 strokes/min for 2 min using the SPEX SamplePrep Model 2000 Geno/Grinder (SPEX CertiPrep, Metuchen, USA). The samples were centrifuged briefly for 30s. The metal beads were removed using a magnet. The samples were incubated at 55°C to 60°C for 1h and 600 µl ddH_2O were added. 1 µl sample was used per 20 µl PCR reaction. These DNA preparations can be stored in -20°C and used within two weeks with less than 2 thawn-freeze cycles.

The metal beads were reused in the next experiments by cleaning with 1 M HCl for approximately 30 min then rinsing with water and 70% (v/v) EtOH and air-dried.

9.2.1.9 Plant DNA miniprep

Miniprep buffer 200 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% (w/v) SDS

All steps were at room temperature. Up to 1cm x 0.5 cm leaf material was collected into 1.5 ml tube and ground with a plastic pestle. 400 µl Miniprep buffer were added and grinding was continued. The samples can be stored on ice at these step until all the samples were ground. The samples were then centrifuged for 10 min at 13.000 rpm. The supernatant was transferred to a new tube containing 400 µl isopropanol. The samples were mixed by inverting and centrifuged for 15 min at 13.000 rpm. The supernatant was pour off and the pellet was washed with 500 µl 70% (v/v) EtOH. The pellet was air-dried for 20 min and dissolved in 150 µl ddH_2O. 1 to 3 µl of the sample were used in a PCR reaction.

9.2.1.10 Restriction digestion of plasmid DNA

To prepare restriction fragments for cloning, the restriction digests were performed in the buffer and at the temperature recommended by the manufacturer for 2h to overnight. To avoid star activity, the final volume of enzyme/glycerol was kept below 10% (v/v) of the final reaction volume (usually 20-50 µl). For preparing recipient vector in cloning, 2U of restriction endonuclease was added to 3 µg of plasmid (ca. 1µl to 2µl of the miniprep if the plasmid is high copy number e.g., ML939 derivatives), samples were incubated for one hour then another 2U of restriction endonuclease was added and incubated for another two hours.

9.2.1.11 Agarose gel electrophoresis of DNA

DNA loading dye 30% (w/v) sucrose, 0.2% (w/v) cresol red, 0.3% (w/v) tartrazine

DNA fragments were separated by electrophoresis in horizontal agarose gels. Gels were prepared with 1x TAE buffer and 0.7-3% (w/v) agarose. 0.01% (v/v) ethidium bromide (EtBr) was added to the molten agarose before the gels were casted. Gels were run in 1x TAE at 80C-110V. The DNA was visualised using a short-wavelength ultraviolet light transilluminator (610 nm) (UVP Inc., Cambridge, UK). For the isolation of DNA fragments

from gels, the DNA was visualised using long-wave ultraviolet light transilluminator to minimise DNA damage.

9.2.1.12 Isolation of DNA fragments from agarose gel

DNA fragments were isolated from agarose gel using the QIAquick Gel Extraction Kit (QIAGEN #28704) according to the manufacturer's instructions.

9.2.1.13 DNA dephosphorylation

If the vectors can be self-ligated (e.g., by cloning using one restriction enzyme), the 5′ phosphates of the vector were removed with rAPid Alkaline Phosphatase (rAP) (Roche #04-898-133-001). Up to 1 µg of the 5' ends of vector DNA were dephosphorylated using 1x rAP buffer and 1 µl of rAP in a total reaction volume of 20 µl. The sample was incubated for 30 min at 37°C and the enzyme was heat-inactivated for 2 min at 75°C.

9.2.1.14 DNA ligation

Both the insert and vector DNA were run on a agarose gel to check their concentration before using in ligation. 10 µl ligations were performed using 1x ligation buffer and 0.5 µl-1µl T4 DNA ligase (NEB, M0202L). Insert and linearised vector DNA were added to give a molar ratio of DNA ends of 1:3 vector:insert. Ligations were incubated at room temperature for 2 h or at 16ºC overnight. Up to µl of this ligation were used to transform E.coli electrocompetent cells.

9.2.1.15 Quantification of nucleic acid concentration

The optical density (OD_{260}) of a nucleic acid sample was measured in an Ultrospec 1000E UV/visible light spectrophotometer (Pharmacia, Kent, UK) and the concentration of the nucleic acid was calculated by assuming that: 1 OD_{260} unit equals 50 µg/ml double-stranded DNA and 40 µg/ml single-stranded DNA or RNA. The NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Thermo Fisher Scientific Inc.) was used for accurate quantifications of very small sample volumes. For approximate estimates the DNA sample was run in an agarose gel supplemented with 0.01% (v/v) EtBr with DNA standards of a range of known concentrations. The DNA was visualised on a UV transilluminator and the DNA concentration of sample was estimated by comparing the intensity of the band with those of the known standards.

9.2.1.16 Polymerase chain reaction (PCR)

The GoTaq Flexi DNA Polymerase (Promega #M8301) was used for amplification of DNA fragments from DNA plasmids, colony-PCR and genotyping. The Phusion Taq from NEB (M0530S/L) was in PCR for cloning to reduce PCR errors. For amplifications of DNA fragments using primers that introduce mutations, the PCR was run at an annealing temperature 5^{o}C below the actual annealing temperature for the first five cycles. To ensure

correct handling and mixing of the components, a negative and a positive control sample were run together with the samples.

Final concentration	Component
1 pg-40 ng 50-500 ng	plasmid DNA or genomic DNA
1x Or 5x	GoTaq Flexi Buffer Or HF Phusion PCR buffer
200 µM	deoxyribonucleotides (dNTPs)
0.25-0.5 µM	oligonucleotide A
0.25-0.5 µM	oligonucleotide B
0.5 U 0.4 U	GoTaq Flexi polymerase Phusion Taq polymerase
20 µl	Total volume

PCR amplifications were performed in either a Tetrade Thermal Cycler PTC-225 (MJ Research, Waltham, USA) or in a G-Storm GS1 thermal cycler (Gene Technologies Ltd, Essex, UK) using the following program.

Step	Temperature		Time	
	GoTaq	PhusionTaq	GoTaq	Phusion Taq
1 (Denaturation)	95°C	98°C	2 min	30 s
2 (Denaturation)	95°C	98°C	30 s	10 s
3 (Annealing)	T_m-3°C	T_m°C (must be >60°C)	30 s	30 s
4 (Elongation)	72°C		1 min/1 kb	30 s/1 kb
5	Repeat steps 2 to 4 for 24 to 39 times			
6 (Final elongation)	72°C		3 to 5 min	
7	25°C		1 min	

For overlap-PCR for cloning, Phusion Taq was used. In the first round of PCR, I used 25 cycles using conditions above. The PCR products from the first round were electrophoresis and bands of expected size were gel-purified and used as template for the second round of PCR (also 25 cycles).

9.2.1.17 Bacteria colony-PCR

A single colony was picked and transferred to a 1.5 ml tube containing 50 µl ddH_2O. As a negative control, an empty space in the petri agar plate was touched by the pipet tip and the tip was washed (by pipetting up and down) in 50 µl ddH_2O. The tube was vortexed, kept on ice and 5 µl of the suspension was used as template in a 20 µl PCR reaction using GoTaq. The initial denaturation was increased to 5 minutes. 40 cycles of PCR was used.

9.2.1.18 DNA sequencing

Cycle sequencing reactions were performed using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, USA) according to the manufacturer's instructions. 100-200 ng template DNA were used in a 10 µl reaction containing 3.5 µl of 1 µM sequencing primer, 1.5 µl Big Dye sequencing buffer and 1 µl ABI BigDye version 3.1 sequencing mix. Analysis of the reaction products was performed on AbiPrism 3730XL and 3730 capillary sequencers (Perkin and Elmer, Connecticut, USA) by the DNA Sequencing Service at the JIC. The data were downloaded and analysed using the Vector NTI Advance 11 software package (Invitrogen). The cycle sequencing reactions were performed using the following program.

Step	Temperature	Time
1	96°C	10 s
2	55°C	5 s
3	60°C	4 min
4	Repeat steps 1 to 3 25 times	
5	Cool down to 10°C	

Some of the sequencing were carried out using a service provided by LGC Genomics GmbH, Germany according to the company protocols.

9.2.1.19 DNA concentration

Plasmid/genomic DNA was diluted to the total volume of larger than 50 µl with ddH_2O. 0.1 volume of 3M NaOAc pH 5.2, two volumes of ethanol (96-100%) were added. The tube was mix and incubated at -20^oC for 30 min, then spun at 13.000 rpm for for 20 min at room temperature. The pellet was then washed with 70% ice cold ethanol, spun again for 5 min. The pellet was allowed to dry completely then elute in ddH_2O.

9.2.2 Molecular biology – RNA-related methods

9.2.2.1 Mini Hot Phenol *Arabidopsis* RNA extraction procedure

Homogenisation buffer 100 mM Tris pH 8.5, 5 mM EDTA pH 8.0, 100 mM NaCl, 0.5% (w/v) SDS

Unless otherwise stated, the steps were carried out at room temperature.

Hot phenol tubes were prepared by adding 250 µl phenol, 500 µl homogenisation buffer and 5 µl β-mercaptoethanol to a 1.5 ml tube and heated in a 60^oC water bath. Approximately 300 mg frozen *Arabidopsis* tissue (maximum half the volume of a 1.5 ml tube) were ground to fine powder in liquid nitrogen with a pestle. A pinch of glass sand was added to aid the grinding process. After the liquid nitrogen was completely evaporated but the sample was not thawn, the hot phenol mixture was added. The tube was shaken for 15 min, 250 µl chloroform were added and the tube was shaken for additional 15 min. The sample was centrifuged in a microcentrifuge at 13.000 rpm for 10 min and 550 µl aqueous layer were transferred to new tube. After adding 550 µl phenol:chloroform:isoamylalcohol (25:24:1), the mixture was shaken for 10 min. The sample was centrifuged at 13.000 rpm for 10 min and 500 µl aqueous layer were transferred to new tube. 50 µl 3 M sodium acetate and 400 µl isopropanol were added and the tube was incubated at -80^oC for 15 min. After centrifugation at 13.000 rpm for 30 min (4^oC), the supernatant was removed and the pellet was air-dried

for 10 minutes*. The pellet was then resuspended in 500 µl ddH_2O**. After adding 4 M LiCl, the sample was mixed and incubated overnight at 4°C. The next day the sample was centrifuged at 13.000 rpm for 30 min (4°C), the pellet was washed in 1 ml 80% (v/v) ethanol and centrifuged at 13.000 rpm for 5 min. The supernatant was removed carefully to not disturb the delicate pellet and the pellet was air-dried. The pellet was resuspended in 20 µl ddH_2O and the RNA was quantified using a spectrophotometer.

* It is important not to dry the phellet for too long or it will be very difficult to dissolve in the next step.

** It is important to dissolve the phellet fully to prevent carry over of genomic DNA.

9.2.2.2 DNase digestion of RNA samples

Traces of DNA that may be present following RNA isolation protocols were removed by digestion with Turbo-DNAse (Ambion Inc. #AM1907, Warrington, UK). A maximum of 10µg of RNA was used with 10U Turbo DNase in a 50 µl total reaction volume. 5 volume 10x Turbo DNase buffer and 1 µl Turbo DNase (10 U/µl) were added to the RNA and mixed gently. After incubation at 37°C for 30 min, 0.1 volume of DNase Inactivation Reagent was added, mixed and incubated for 5 min at room temperature. The sample was centrifuged for 1.5 min at 10.000 x g and the supernatant was transferred to a fresh tube.

For higher amount of RNA to be digested, it is important to adjust the reaction volume so that the concentration of RNA does not exceed 0.1 µg/µl.

9.2.2.3 cDNA synthesis of polyadenylated transcripts

For the synthesis of cDNA, the SuperScript III Reverse Transcriptase (RTase) kit (Invitrogen #18080-051) was used according to manufacturer's instructions.

2 µg of total plant RNA (up to 12.2µl) was mixed with 0.8 µl of 100µM oligo$(dT)_{17}$ primer, 1 µl of 10mM dNTPs and water up to 14 µl. The tube was incubated for 5 min at 65°C and then briefly chilled on ice. The following reaction mixture was prepared and added to each sample.

Volume (µl)	Component
4	5x First Strand Buffer
1	DTT (100mM)
1	SuperScript III RTase

The samples were mixed by pipetting up and down, then incubated at 50°C for 60 min followed by heat inactivation of the RTase enzyme at 75°C for 10 min. The cDNA was

stored at -20^{o}C. For qPCR and standard RT-PCR the cDNA was diluted one in 10 and 4 µl was used for a 20 µl PCR reaction.

9.2.2.4 RNA concentration:

For crude estimating the integrity and concentration of the extracted total RNA, RNA was electrophoresis in 1% Agarose gel in 1xTBE. To minizmize RNAse contamination, the gel apparatus was cleaned twice with soap and rise with ddH_2O before electrophoresis, TBE was autoclaved and agarose was used from freshly open bottle, which is designated to use only for RNA.

RNA was precipitated using 0.1 volume of 3M Sodium Acetate pH 5.2 and 1 volume of isopropanol. The mixture was incubated at -20ºC for at least 30 minutes or overnight, then centrifuged at 13,000 rpm at 4ºC for 45 minutes. The supernatant was removed and the pellet was washed with 200 µl of 80% ice-cold ethanol. The pellet was air dried and dissolved in 11 µl water, stored at -20ºC.

9.2.2.5 Poly(A) tail removal by RNase H/oligo dT treatment of total RNA:

Chemical/reagents	Amount added
Total RNA (DNase treated)	20 µg
10x RNase H buffer	10 µl
RNase H (5U/µl) (NEB)	0.4 µl
100 µM oligo dT*	3.2 µl
Water	Adjust to total volume reaction volume of 100 µl

* if poly(A) tail persists, can increase oligo dT amount up to 2 times.

The above mixture was prepared, in the sample without oligo dT, water was added in replacement of oligo dT. The reaction was incubated at 37ºC for one hour. Then 100µl phenol/chloroform/isoamyl alcohol (25:24:1) was added to inactivate the enzyme. The tube is mixed by inversing and then spin at 13,000rpm for 10 minutes at room temperature. The aqueous phase was transferred to a fresh tube and RNA was precipitated using 0.1 volume of 3M Sodium Acetate pH 5.2 and 1 volume of isopropanol. The mixture was incubated at -20ºC for at least 30 minutes or overnight, then centrifuged at 13,000 rpm at 4ºC for 45

minutes. The supernatant was removed and the pellet was washed with 200 µl of 80% ice-cold ethanol. The pellet was air dried and dissolved in 11 µl water, stored at -20ºC.

9.2.2.6 RNA Ligation Mediated Poly(A) test (LM-PAT).

The protocol was done according to (Meijer et al., 2007).

RNA ligation

4 µg of total RNA was mixed with with 50 picomoles of the RL-PAT-anchor DNA primer (oSV220), water added to 8 µl. The mixture was incubated at 65□C for 5 minutes then place immediately on ice. 1 µl of RNA ligation buffer and 1 µl of RNA ligase (Promega M1051) were added and incubate at 37ºC for 30 minutes, then inactivated at 65ºC for 15 minutes then put on ice.

cDNA synthesis

50 picomoles of the RL-PAT-Rev primer (oSV219), 1 µl of 10 mM dNTPs was then added to the tube from the ligation step, water was then added to adjust the total volume to 25 µl. The tube was incubated at 65ºC for 5 minutes then placed immediately on ice. The following master mix was then added to the tube: 10 µl of first strand buffer 5X, 2.5 µl of 0.1 M DTT, 11.5 µl of water and 1 µl of the Superscript III reverse transcriptase (Invitrogen). The tube was then incubated at 55ºC for one hour, heat inactivate at 70ºC for 15 minutes. The resulting cDNA was stored at -20ºC.

PCR

PCR was done using GoTaq (Promega) with the following mix:

Chemical/reagents	Amount added
Water	33.75 µl
5x GoTaq Flexi buffer	10 µl
25 mM $MgCl_2$	3 µl
10 mM dNTPs	1 µl
20 µM RL-PAT-Rev primer (oSV219)	1 µl
20 µM Gene specific forward primer	1 µl
GoTaq	0.25 µl
cDNA template from the above step	1 µl

Cycle conditions

The following condition was used: initial denaturation: 5 minutes at 95ºC; 40 cycles of: 1 minutes at 95ºC, 1 minute at 58 ºC, 2 minutes at 72ºC; final extension: 10 minutes at 72ºC.

Nested PCR

For transcripts that have low abundance, nested PCR was carried out using the same PCR conditions as above with modifications. In the first PCR: the first gene specific primer and oSV219 were used at half the concentration listed above (i.e. final concentration was 0.2 µM). In the second round of PCR: a second gene specific primer (the nested primer) and oSV219 was used at full concentration (i.e. 0.4µM); the template was 2 µl of the 20x dilution of the first PCR product.

9.2.2.7 Bulk poly(A) tail-length analysis:

The protocol is modified from (Preker et al., 1995). In summary, RNA samples (2 ug) were 3'- end labelled with ^{32}P cordycepin and subjected to RNase A and RNase T1 digestion. The Rnase treatment was stopped and poly(A) tails were precipitated and separated by ectrophoresis on a gel of 10% polyacrylamide and 8.3 M urea.

3'end label with cordycepin

Chemicals/Reagents	Concentration of stock	Amount added	Final concenctration /amount
PAP Buffer	5x	2µl	1x
Rnase-free water		5 µl	
Yeast PAP enzyme	5U/ µl	1 µl	5U
^{32}P cordycepin	10 µCi/µl, 5000Ci/mmol	1 µl	10 µCi
Total RNA	2 µg/ µl	1 µl	2 µg

Total RNA was extracted using hot phenol method and dilute in water to 2µg/µl. 2µg of total RNA were 3'- end labelled with ^{32}P cordycepin (From PerkinElmer, Product No. NEG026250UC) using yeast poly(A) polymerase (USB Affymetrix Product No. 74225Y). The reaction was incubated at 37ºC for 20 minutes followed by heat inactivation at 70ºC for 10 minutes.

Rnase treatment

Chemicals/Reagents	Concentration of stock	Amount added	Final concenctration/amount
Tris-HCl (pH 8)	100mM	8 µl	10mM
NaCl	3M	8 µl	300mM
Rnase T1 *	125U/ µl	1 µl	125U
Rnase A*	10mg/ml	0.5 µl	5 µg
Rnase-free water		51.5 µl	

***** Rnase T1 was purchased from Ambion 1000U/µl then diluted in 10mM Tris-HCl (pH 8)-300mM NaCl to make 125U/ µl. Rnase A was purchased from Sigma.

The above master mix (70 µl) was added to the tube (10 µl) in the previous step and was incubated at 37□C for 60 minutes.

Stopping RNase reaction

Chemicals/Reagents	Concentration of stock	Amount added	Final concenctration/amount
proteinase K	20 mg/ml	5 µl	1 mg/ml
SDS	10 %	10 µl	0.5 %
EDTA	500 mM	2 µl	10 mM
Water		3 µl	

The above master mix (20 µl) was added to the tube in the previous step (80 µl) and was incubated at 37ºC for 30 minutes.

Precipitation of poly(A) tails

Chemicals/Reagents	Concentration of stock	Amount added	Final concenctration/amount
tRNA	10 mg/ml	5.2 µl	52 µg

Glycogen	23 mg/ml	5.4 µl	125 µg
Ammonium acetate*	6M	83 µl	2.5 M
$MgCl_2$	1M	3 µl	15 mM
Rnase-free water		3.4 µl	

Ammonium acetate* make fresh with autoclaved water, can store in the fridge for several weeks.

The above master mix (100 µl) was added to the tube in the previous step (100 µl) and then 500 µl absolute Ethanol was added. The tube was mixed by inversing several times then centrifuge for 1 hour at 13,000 x g at 4ºC to 10ºC. The pellet was washed with ice-cold 80% Ethanol and allowed to dry at room temperature.

Resuspend in loading buffer

Chemicals/Reagents	Concentration of stock	Amount added	Final concentration
Formamide	98%	960 µl	94%
EDTA	0.5M	40 µl	20 mM
Bromphenol blue		2 mg	0.2%
Xylene cyanol		2 mg	0.2%

Loading dye was prepared and stored in small aliquots in -20ºC.

The dried pellet in the previous step was resuspended in the 10 µl of the loading buffer (formula above). The RNA can be stored in - 20ºC or immediately used for electrophoresis.

Gel preparation and electrophoresis

Sequencing gel apparatus were used. The glass plates were cleaned thoroughly with a final rinse with ethanol. The glass plates then siliconized with dimethyl dichlorosilane (Sigma) on one side. 0.4 mm comb were used. The gel mixture were prepared with the following recipe.

Chemicals/Reagents	Concentration of stock	Amount added	Final concentration/ amount
TBE	5x	18 ml	1x

Water		10 ml	
Urea		45 g	8.3M
Acrylamide	30%	33 ml	10%
TMED	100%	60 µl	
APS	10%	330 µl	

This is the formula for a final volume of 90ml, enough for one gel. APS 10% was made fresh or can be stored in -20□C for 2 months. TBE was autoclaved and keep free of Rnase.

TBE, water and urea were mixed and heat until fully dissolved then acrylamide was added. The mixture was cooled with icy water to slow down polymerization. APS and TMED was added last and quickly pour in to the gel apparatus. The polymerized gel together with glass plates were then stored in the cold room overnight and used next morning. RNA samples (3-5 µl) were heated at 70ºC for 2 minutes then load on the gel together with RNA ladder (3 µl) which was heated at 95 C for 5 minutes. The gel was run at 1.5kV 43mA at 55ºC for about 3 hours until the faster running dye (bromephenol blue, which co-migrates with 10bp RNA) is 7 cm above the bottom of the gene. The gel was carefully peeled off from the glass plates on to blotting paper then exposed to the phosphoimage screen overnight.

RNA ladder preparation

The RNA ladder was the Decade Marker system (Ambion AM7778), it was prepared according to the manufacture protocol. ^{32}P Gama ATP was ordered from PerkinElmer. Once produce the ladder can be stored in -20ºC in small aliquots and used within 4 weeks.

9.2.2.8 Poly(A) tail-length dependent RNA fractionation.

The protocol is based on (Meijer and de Moor, 2011) with modifications: Total RNA was fractionated without the spiked in radioactive synthetic poly(A) probes. Instead, the poly(A) tail length of the final fractions were analyzed using the bulk poly(A) tail length analysis protocol (section 1.2.2.7). Below is the protocol for fractionation.

Materials

80 µg total RNA in RNase free water

PolyA Tract System 1000, containing paramagnetic streptavidin beads, GTC extraction buffer, dilution buffer, biotinylated oligo (dT), β-mercaptoethanol (BME), 0.5x SSC and magnetic separation stand, Promega cat. no. Z5400

Extra 50 pmole/µl biotinylated oligo(dT), Promega cat. no. Z5261

RNase free water

3M Sodium acetate pH 5.2

100% isopropanol

10 µg/µl yeast tRNA

70% ethanol

Important: non-stick tubes (2ml) (Ambion AM12475) were used in all steps to prevent sticking of RNA to the tubes which results in significant loss of RNA from the short fractions.

Methods

The GTC, BME, biotinylated oligo(dT), 0.5xSSC and H_2O were allowed to reach room temperature. 41 µl BME was added to each 1 ml of GTC (GTC/BME). 20.5 µl BME was added to each 1 ml dilution buffer (DIL/BME). The DIL/BME was preheated to 70°C. The 0.085x SSC was prepared from DEPC-treated water and 0.5x SSC provided in the poly(A) Tract kit. Note that this concentration was critical in order to obtain the 50 nucleotide threshold of separation i.e. the short fraction was depleted of of RNA having a longer tail than 50 nucleotides. The paramagnetic streptavidin beads were prepared as follow: The beads were completely resuspened by gently rocking the bottle. The resuspended beads (600 µl) were transferred to a 2 ml tube for each sample. The tube was placed on the magnetic stand and the stand was slowly moved to a horizontal position until the beads were collected at the tube side (takes 30 seconds to 1 minute). The storage buffer was carefully poured off by tilting the tube so that the solution run over the captured beads. The beads were resuspended in 600 µl 0.5xSSC and captured again using the stand. This wash step was repeated twice. The beads were resuspended in 600 µl 0.5xSSC. 80µg of total RNA (2 µg/µl) was mixed with 400 µl GTC/BME in a 2 ml non-stick tube. 15 µl biotinylated oligo(dT) and 816 µl DIL/BME were added and incubated at 70°C for 5 minutes. The tube was spun at 12,000 g for 10 minutes at room temperature. Note that the cooling centrifuge was used otherwise the temperature will go to 40ºC. After the spin the supernatant was added to the beads. The biotinylated oligo(dT) was allowed to bind to the beads by rotation at room temperature for 15 min. The beads were then captured using the magnetic stand. The supernatant was transferred to a new tube and kept on ice (400 µl unbound fraction). The beads were washed three times with 0.5x SSC, supernatant was discarded. The tube was rotated for 5-15 minutes between each wash. The beads were resuspended in 400 µl 0.085x SSC, then rotate for 15 minutes at room temperature. The beads were captured using the magnetic stand. The eluate (400 µl) was transferred to a new non-sticky tube. This eluation step using 0.085x SSC was repeated one more time in ordered to fully elute the short fractions. The resulting pooled 800 µl eluate (the short fraction) was kept in a tube ice. Note care was taken not to disturb the beads and also to remove the last drops of eluate

from the beads. After the short fraction was eluted, the beads was resuspended in 400 µl water and rotated for 15 minutes at room temperature. The eluation with water was repeated one more time resulting in a combined 800 µl eluate (the long fraction). The short fraction, long fraction and unbound fraction were spun at 13,000 rpm for 5 minutes at 4ºC to remove any transferred beads. To the supernatant from the unbound fraction (400 µl), 400 µl of isopropanol was added, mixed and precipitated overnight. To the supernatants from the long (800 µl) and short fraction (800 µl), 20 µg of tRNA, 72 µl of Sodium acetate, and 720 µl of isopropanol were added, mixed and precipitated overnight.

Next day, all the samples were spun at 13,000 rpm for 30 minutes at 4ºC to precipitate the RNA. The supernatant was removed carefully, and washed with 800 µl of 80% ethanol. The samples were spun again at 13,000 rpm for 20 minutes at 4ºC and supernatant was carefully removed leaving 20 µl. Samples were spun again at 13,000 rpm for 5 minutes at 4ºC. Note that it is important that this extra spin was carried out to make sure that the pellet sticks to the tube wall. The pellet was allowed to dry and finally resuspended in 12 µl water. RNA concentration was measured and samples were stored in -20ºC for no longer than 2 weeks.

All the fractions were analyzed by bulk poly(A) tail analysis as described in section 215, in which 2 µl (approximately 2 µg) of each fractions were used.

9.2.2.9 Non-specific polyadenylation assay:

(Adopted from ((Vethantham et al., 2008)), performed by Nishta Rao.

^{32}P-labeled SVL precleaved substrate (pre-SVL) was prepared as described (Ryner et al., 1989). A nonspecific polyadenylation assay reaction mixture consisted of 2.5% PVA, 1 mM $MnCl_2$, 100–125 ng of BSA, 1 mM ATP, 0.5 U of RNasin, 10 mM HEPES (pH 7.9), 25 mM NH_4(SO4)2, 0.2 mM PMSF, 0.2 mMDTT, and the indicated amounts of PAP from both reaction mixtures. The nonspecific assays were incubated for 30 min at 30°C, and RNAs were resolved by denaturing PAGE and subjected to autoradiography.

9.2.2.10 Quantitative PCR (qPCR):

qPCR was carried out using SensiMix SYBR Low-ROX Kit (Bioline Cat No. QT625-05) and AB7500 Fast Real time PCR machine according to the standard manufacturer's instructions. The master mix is modified as follow:

SensiMix SYBR Low-ROX 2x	5 µl
Primer forward (10 µM)	0.25 µl
Primer reverse (10 µM)	0.25ul
Water	0.5 µl

cDNA (1 in 20 dilution*) 4 µl

*cDNA was synthesized as in 9.2.2.3 and diluted 1 in 20 and store in small aliquots in -20°C C.

The data is analyzed using LinRegPCR program (Ramakers et al., 2003).

9.2.3 Molecular biology – protein-related work

9.2.3.1 Expresssion and purification of recombinant *Arabidopsis* PAPS1 from E.coli:

This procedure was provided and performed by Nishta Rao.

pET-PAPS1^{wt} or pET-PAPS1^{P313S} was transformed to Rosetta 3 cells. A single colony was picked and inoculated in 50 ml LB with 200µg/ml Kanamycin. Cells were grown for 4-5 hours at 37°C. 20 ml of this culture was added into 400 ml of LB+antibiotic and grown at 37°C for 1-2 hours. The cultures were shifted to 16°C and let to cool for 15-20 min before IPTG to a final concentration of 1mM was added and these cultures were grown overnight in the shaker (16 hours). After 16 hours, the cultures were cool briefly on ice and harvested by centrifugation at 4°C at 5000 rpm for 10 min. The pellets from 1x400 ml flask were resuspended in 12 ml of cold binding buffer (PMSF at final concentration of 0.5mM). Pellet can be frozen at this stage if needed at -80C. The frozen pellets were thawn in a 37°C water bath and placed on ice. 50µl of protease inhibitor cocktail was added. The mixtures were sonicated at an output between 2-4 , 20 times while on ice and made sure that there were no bubbles. The mixtures were transferred to clean tubes, spun at 9000 rpm in cold for 20 min. While the spin was going, 1mL of Ni agarose per tube of lysate was spun down at 5000 rpm. Supernatant was removed. This agarose beads were washed 2-3 times with binding buffer, then resuspended with 0.5 ml of binding buffer. The supernatants of cell lysates were transfered to new falcon tubes. 12ml of binding buffer was added and then the mixtures were filtered using 0.8µM filter. Ni-agarose slurry was added to these mixtures and the tubes were placed on rotator in a cold room for 2 hours. The tubes were spun 6 min in a cold room.

The protein column (BioRad) was set up in the cold room. The supernatant was decanted and the slurry was mixed with some binding buffer and poured into column It was made sure that there were no bubbles. The column was washed with wash buffers 1, 2 (increasing Imidazole concentration). Upon finishing the second wash, an elution buffer (high Imidazole concentration) was added to the column and eluates was collected.

9.2.4 Plant-related methods

9.2.4.1 Plant growth conditions on soil

All plants were grown either in glasshouses, in controlled environment rooms (CER) or in Sanyo growth cabinet. All plants were grown under long day conditions consisting of a photoperiod of 10 hours lit with 400 W metal halide power star lamps supplemented with 100 W tungsten halide lamps, providing a level of photosynthetically active radiation (PAR) of 113.7 µmol photons m-2s-1 and a red:far red (R:FR) ratio of 2.41 and a 6 hour extension of tungsten halide lamps alone giving PAR 14.27 µmol photons m-2s-1 and R:FR 0.66.

Humidity was around 50-70%. The standard temperature is from 22°C to 25° C. For *paps1-1* mutant, various growth temperatures were used ranging from 16°C to 28°C. The exact temperature is mentioned in each specific experiment. Control plants were grown side by side with the plants to be tested.

Plants were grown on a loam-based soil consisting of John Innes Potting Compost Number 1, vermiculite and grit in the ratio 1.5:1:1 by volume or a mixture of Levingtons Potting Compost Number 3 and grit in the ratio 6:1 by volume. The seeds were sown on water-saturated soil, and covered with transparent lids. To break seed dormancy, the trays were stratified in at 4°C for 2 to 3 days before transferring them to a glasshouse or CER. After 10-14 days, the lids were removed. The plants were watered once or twice daily. To harvest seeds, the inflorescences were bagged when the mature siliques turn yellow. After the siliques had dried up, the seeds were harvested and dried at RT or in a closed nylon bag with silica gel for at least 2 weeks before sowing.

9.2.4.2 Growing conditions on MS plates

All sterile procedures involving plant material were performed in laminar flow hoods in a tissue culture room. Autoclaved media were cooled to ~50°C and the appropriate antibiotic or PPT was added prior to pouring the medium into 90 mm or 140 mm Petri dishes (Sterilin, Caerphilly, UK). After sowing the sterilised seeds onto the medium, the plates were sealed with micropore surgical tape (3M). The seeds were stratified at 4°C for 2 to 3 days and then transferred to a CER.

For transferring of seedlings from MS plates to soil: Seedlings were normally tranfered to soil 10-14 days after germination. The seedlings were covered with a propagator lid for about 5 days only after transferering.

9.2.4.3 Seed sterilisation

Wet sterilization

Sterilisation solution: 5% (v/v) sodium hypochlorite solution, 0.1% (v/v) Tween-20

Seeds were transferred to 1.5 ml tube and 1 ml 70% (v/v) EtOH was added. After the seeds had settled, the EtOH was decanted and 1 ml sterilisation solution was added. An incubation step for not more than 4 min followed. Brief centrifugation were occationally required to bring the seeds to the bottom of the tube in the following washing procedures. The sterilisation solution was removed and the seeds were washed five times with sterile ddH_2O and rinsed with 1 ml 100% (v/v) EtOH. The seeds were then air-dried in a laminar flow hood.

Gas sterilization

Tubes or bags containing a small amount of seeds (less than 50µl) were fumed with chlorine gas. The tubes/bags of seeds were put in a rack next to an open container which contains 3ml concentrated HCl mixed with 100ml bleach. The open container and tubes/bags of seeds were put into a bigger container and sealed for four hours. The tubes/bags containing seeds were then left under a sterile flow hood overnight to make sure all the chlorine gas has evaporated.

9.2.4.4 Mapping of the *ds39* (i.e. *paps1-1*) mutation

ds39 homozygous mutants after being backcrossed to L*er* twice were crossed to Col-0. The F1 plants were allowed to self and seeds were collected in pools. In the F2 population, the mutants were identified as seedlings which small and pointed leaves, which were later checked for the bigger-flower phenotype. Initially, bulk segregation analysis was carried out using a pool of 20 mutant plants using markers across the genome of Arabidopsis. This identified a location between marker *nga64* and *nga392* which is tightly linked to the mutation. Another 140 mutant plants identified in this F2 population were genotyped for left inner markers: *CER474022*, *CER481865* and right inner markers: *F19G10*, *CER470358*. Fine mapping was carried out using about 1000 F2 mutant-looking individuals in pools of three. These approximately 335 pools of DNA were PCR-screened with the two flanking markers *CER481865* and *CER470358* to identify recombinants. New markers were designed to finally narrow down the mapping interval to 80kb between 2 markers: *CER475120* (2 recombinants found) and *CER475042* (1 recombinant found). There were 19 genes in this inteval. *T-DNA* insertion lines that have *T-DNAs* predicted to have inserted in each of these 19 genes were looked up and ordered. At the same time, the cDNA of these 19 genes from the mutant were sequenced. A *T-DNA* insertion lines (*SAIL_172_F11*) i.e., the later called *paps1-2* allele, segregated for the pointed leaves and small leaves phenotype. Sequencing of cDNA for At1g17980 found a mutation C to A.

9.2.4.5 *Agrobacterium* floral dip transformation of *Arabidopsis*

Infiltration medium 5% (w/v) sucrose, 0.05% (v/v) Silwet L-77 (De Sangosse Ltd #45290, Cambridge, UK)

The floral dip transformation protocol was carried out (Clough and Bent, 1998). *Arabidopsis* plants were grown in the glasshouse until they have produced three inflorescences with young buds. To prepare the Agrobacterium, 10 ml LB-broth containing the appropriate antibiotic were inoculated with a single colony of *Agrobacterium* strain GV1301 transformed with the desired plasmid and grown to stationary phase at 28°C with shaking at 200 rpm. 400 ml LB-broth were inoculated with 2 ml overnight culture and the culture was grown overnight at 28°C. The next day, the cells were harvested at 4.600 x g for 15 min at room temperature and the cell pellet was then resuspended in 10 ml infiltration medium. The cell suspension was transferred to beakers containing 1 l infiltration medium. The plants were inverted into the suspension (all aboveground tissues submerged) and left submerged for 20 seconds, during which the pots were pulling up and down gently, still keeping the inflorescences in the suspension. Infiltrated plants were placed horizontally in trays, covered with propagator lids and left in low light or dark overnight. The following day the plants were placed into the glasshouse and left to set T1 seeds.

9.2.4.6 Cross-pollination of *Arabidopsis* plants

Young flower buds whose stigmas were still fully enclosed within the sepals were emasculated (removing all floral organs but the gynoecium). The emasculated bud was then left untouched for one day then pollinated with pollen from the respective donor plant. The pollinated buds were then allowed to set siliques and the matured siliques were harvested shortly before dehiscence.

9.2.4.7 Measurement of organ size using ImageJ

Petals were dissected from the 6th to 15th flowers and used for measurements. For leaves, the 4th and 5th leaf of plants at the bolting stage were taken for measurements. To measure organ size, the organs were placed with forceps onto Sellotape. Once all organs were collected, the tape was stuck onto a black Perspex sheet. The organs were scanned with a resolution of 3600 dpi in greyscale (8-bit) using the HP Scanjet 4370. The organ size was then measured using the Image Processing and Analysis in Java (ImageJ) software (http://rsbweb.nih.gov/ij/).

9.2.4.8 Measurement of cell size:

Two methods were used: For petals, low-melt agarose was used to analyse epidermal cells; for leaves Chloral hydrate method was used to analyse palisade cells in the subepidermal layer.

a. Low-melt agarose method to analyze size of petal epidermal cell:

The protocol is adapted from (Horiguchi et al., 2006). A drop of 2% low-melt agarose containing 0.01% bromophenol blue pre-warmed at 50°C was placed on a pre-warmed glass slide. The droplet was smeared by a pipet tip to get a thin layer of melting agarose. A petal was immediately gently placed on it. Once the gel solidified, the petal was carefully

peeled off, and the remaining gel cast was left to dry for about 10 min. The gel cast was then observed without a cover glass under a differential phase contrast microscope.

b. Chloral hydrate method:

FAA solution: Ethanol, 50ml; Glacial acetic acid, 5ml; Formaldehyde (37-40%), 10ml; distilled water, 35ml.

Chloral hydrate solution: chloral hydrate, 200 g; glycerol 20 g; distilled water, 50 ml.

Leaves are fixed in FAA solution (by immersing in FAA solution and vaccum infiltration). Then FAA solution was replaced with Chloral hydrate solution and the tube was incubated overnight. Samples were mount on a glass slide and observe under a differential phase contrast microscope.

9.2.4.9 Ethanol induction of plants

Induction by vapour

Ten to twelve pots, each of which is 10 cm in diameter, 10cm high and contains 4-8 plants, were put into a tray and covered in a nylon bag. Ten to twelve 1.5 ml tubes, each of which contains 1ml of absolute ethanol, were open and distributed evenly in the nylon bag containing the pots and plants. The nylon bag was sealed during incubation time, which varies depending on the experiments. This is the standard for plants grown on soil. For seedlings that were grown on 10cm-diameter petri dishes, 5-10 µl of absolute ethanol was pipetted on to the cap of a PCR tube and put on to the agar surface, the dish was closed with a lid and sealed with parafilm. Two days after induction, plants can be pricked out on to soil.

For induction of *Cre*, the incubation time was 20-30 minutes. This created chimeric meristems. An incubation of up to four hour has been tried, at which both fully recombined and chimeric meristems were obtained. Plants were induced at 7 days after germination (resulting in only a few leaves in the rosette were mutant, and most of the rosette are wild-type).

For induction of *PAPS1* in *pPAPs1::AlcR-AlcA::genomicPAPS1,* I used an incubation time of two hours in each time, and once every other day.

Induction by irrigation

This induction method was used for the induction of *PAPS1* in *pPAPs1::AlcR-AlcA::genomicPAPS1* only. 15ml of 2% ethanol was used to irrigate a pot containing approximately 300ml of soil. 8 plants were grown per such pot. Induction was carried out once every other day.

9.2.4.10 GUS staining:

Staining buffer (100ml) : 2 ml 10% Triton X-100; 10 ml 0.5 M NaPO4 pH 7.0; 500 µl 100 mM Potassium ferrocyanide; 500 µl 100 mM Potassium ferricyanide; 87 ml dH_2O.

Staining buffer was prepared and can be stored at 4°C for 2 weeks. Just before the experiment, 40 µl of X-Gluc Stock 50 mM (store at -20°C) was added to 1ml staining buffer to make the GUS staining solution (1X). Tissues were immersed in this GUS staining solution and vaccum infiltrated 5 times until all tissues sunk to the bottom of the tube. The tube was incubated at 37°C for either 2 hours, 4 hours or overnight depending on the strength of the GUS expression in the plants. The reaction was stopped by exchanging the solution with 70% Ethanol. Several rinses with 70% Ethanol (20 minute incubation then exchanged with new 70% Ethanol solution) were carried out until the tissue was cleared of chlorophyll.

9.2.4.11 Developmental series of petal growth:

Petals were dissected from the flowers/buds in the inflorescences and areas were measured then plotted against their growth periods. The growth period of the youngest bud, which contained the smallest dissectible petals was set to be 0. The growth period of the next bud was calculated by adding the plastochron to the growth period of the previous bud. The plastochron (the time difference between two sequential buds) was estimated by counting how many flowers that were made in six days.

9.2.5 Microscopy

9.2.5.1 Clonal analysis using fluorescence microscopy

The fluorescence expression in transgenic plants was analysed using a Zeiss SteREO Lumar stereomicroscope fitted with a Zeiss AxioCam MRm digital camera. Images taken were analysed using the Zeiss Axiovision Software 4.6. For fluorescence analysis the following filter sets for the respective chromophores were used.

Chromophore	Zeiss Filter set	Excitation (nm)	Emission (nm)
CFP	47 HE	BP 436/20	BP 480/40
YFP	46 HE	BP 500/20	BP 535/30

9.2.6 Microarray analysis:

9.2.6.1 Plant materials for microarray:

For inflorescences collection, main inflorescences excluding the open flowers/buds were collected from 30 day-old plants grown on soil. Plants at the stage of harvest, have around 7-12 siliques (WT) or 2-5 siliques (mutant). Plants were grown in long day condition and materials were collected at 5 or 7 hours after the light period starts.

For seedlings collection, the whole seedlings including roots were harvested from 11 days old (mutant) or 9 day old (WT) grown on MS plates. The mutants were sown out 2 days before so that they can be harvested at the same time as WT. Plants were grown in long day condition and materials were collected at 5 or 7 hours after the light period starts.

9.2.6.2 RNA hybridization:

Total RNA was prepared by hot phenol method and cleaned up with QIAGEN RNAeasy Kit, digested with TURBO Dnase (Ambion).

Fractionated RNA was prepared as in section 9.2.2.8. From 80 µg of total RNA input for each RNA sample, 10 µl of short fraction and 10 µl of long fraction were obtained (concentration is around 1µg/µl).

RNA were shipped on dryice to CRX Biosciences Ltd (James Lindsay Place DUNDEE DD1 5JJ, UK) for subsequent labelling and array hybridization.

Two colour array platform was used: Agilent Arabidopsis 4x44K oligo microarray (G2519F, AMADID 15059), and Agilent 4x44K gasket slides (G2534-60011). 1 µg of total RNA samples, and 2 µl the short fraction and 8 µl the long fraction were used for labelling with the Quick Amp Labelling Kit – Two Colour (Agilent# 5190-0444).

9.2.6.3 Bioinformatic analysis:

<u>For array using total RNA:</u>

Analyses were done using the R/Bioconductor (Gentleman et al., 2004) package Limma Background corrections were done using the backgroundCorrect function with the minimum method : background intensities were subtracted from foreground intensities, resulting values of zero or below were replaced by half the minimum foreground intensity. Normalization was done using quantile normalization method.

Differential gene expression were identified using ttest with Benjamini and Hochberg (BH) corrected p-value (Benjamini et al., 2001). The thresold was absolute log2-fold change (mutant/WT) of 1 and pBH <0.05.

<u>For array using fractionated RNA:</u>

Analyses were done using the R/Bioconductor (Gentleman et al., 2004) package Limma. Background corrections were done using the backgroundCorrect function with the minimum method: background intensities were subtracted from foreground intensities, resulting values of zero or below were replaced by half the minimum foreground intensity.

Normalization was done for each microarray using negative and positive control spots with known quantities of RNA. This was done using the normalizeWithinArrays function with the control method : a loess curve is fitted through a set of control spots and then applied to all other spots (Oshlack et al., 2007).

A quantile normalization using the normalizeBetweenArrays function with the tquantile method was done for the following groups of all different samples: MT long, MT short, WT long, WT short. Expression ratios between long- and short-tail RNA fractions (long/short ratio) were calculated using the Limma method. ttest with with Benjamini and Hochberg (BH) corrected p-value was used to identify genes that change the ratio long/short between mutant and wild-type. Genes with absolute log2 expression ratios (log2(Mutant $^{\text{long/short ratio}}$) – log2(WT $^{\text{long/short ratio}}$)) above 1 and BH-corrected p-values below 0.05 were considered significant.

APPENDIX A: Oligonucleotide lists

1. For genotyping of PAPS alleles: For TDNA specific amplification, use RP and BP primers (of the respective *T-DNA*) combinations. For locus specific amplification, use LP and RP primers combinations.

Oligonucleotide name	Sequence	Description
oSV166	TAATGCCCATCATTACTCCTGCGAAT	genotype *paps1-1*
oSV126	GCTTTGTTTGATTCCATAGC	genotype *paps1-1*
oSV100	TCTCGTACAATCCAACATCTTG	genotype *paps1-2* LP primer
oSV91	AGTGTCCAACTCTCCAAGTTTC	genotype *paps1-2* RP primer
oSV126	GCTTTGTTTGATTCCATAGC	genotype *paps1-3* LP primer
oSV78	TGGGACCTAGACATGCAACTAG	genotype *paps1-3* RP primer
oSV77	TGTGAAGTAAACTCAACCCAGAC	genotype *paps1-4* LP primer
oSV79	GGTCTTCTATCAATGGAATTG	genotype *paps1-4* RP primer
oSV120	ACATGGAGATGTTGAACTGCC	genotype *paps2-3* LP primer
oSV121	CCACTGTTCCACGTATATCAAAC	genotype *paps2-3* RP primer
oSV110	TGCATCTGCTGCCACTATATC	genotype *paps4-3* LP primer
oSV111	TTGCTGAAGCTGTAGGGTCTG	genotype *paps4-3* RP primer
oSV116	TGGATACCAACCGAATGCAAC	genotype *paps3-4* LP primer
oSV117	CTGCAAACAAACATCTCACAG	genotype *paps3-4* RP primer
ML437	TGGTTCACGTAGTGGGCCATCG	BP primer for SALK-TDNA
oSV139	AACGTCCGCAATGTGTTATTAAGTTGTC	BP primer for Ws-TDNA
ML438	TTCATAACCAATCTCGATACAC	BP primer for SAIL-TDNA

2. For RT-PCR

Oligonucleotide name	Sequence	Description
oSV187	GTTGCAGACCCATGAAGTAAGAG	qPCR primer for PAPS1
oSV188	GCTCACACTGAAGATCGAGAGAC	qPCR primer for PAPS1
oSV76	AACAAGCACGCGGCTGATTC	RT-PCR PAPS1
oSV100	TCTCGTACAATCCAACATCTTG	RT-PCR PAPS1
oSV186	GTGGGAAGAACCTCAAGTTC	RT-PCR PAPS1
oSV197	GGAAACCAAGTGATTGAGCAG	RT-PCR PAPS1
oSV304	CCAACATTCGTTGAGAATCTTAC	RT-PCR PAPS1
oSV305	ATCTCATTCTTTGCGACGGTG	RT-PCR PAPS1
ML179	CGTTACTGCCAGCCATTGTAGAA	RT-PCR and qPCR for PDF2
ML180	CCGCAGGTAAGAGTTTGGAACAT	RT-PCR and qPCR for PDF2
oSV167	CTGCGACTCAGGGAATCTTCTAA	qPCR primer for UBC
oSV168	TTGTGCCATTGAATTGAACCC	qPCR primer for UBC
oSV198	CCTAGTATGTTGGTTTCTCGA	RT-PCR PAPS2
oSV121	CCACTGTTCCACGTATATCAAAC	RT-PCR PAPS2
oSV110	TGCATCTGCTGCCACTATATC	RT-PCR PAPS4
oSV112	CAATCGTGCCATGGTGGTGGGTAC TCAAAATTTAGG	RT-PCR PAPS4

3. For qPCR validation:

Oligonucleotide name	Sequence	Description
oSV205	CGAAATGGCGGAAACTAAAC	At1g54020 qPCR
oSV206	CTTTGTGTCCTGCGAAGAATAAC	At1g54020 qPCR
oSV242	CAAGCATCTGTGGTAGTAGTGAAGAC	At4g03156 qPCR
oSV243	GCTATCGTTCTTATTCTCTGTGTTGTTC	At4g03156 qPCR
oSV232	GGACAAGAGAGGAGGCATAAG	At2g35640 qPCR
oSV233	CGTCTACCAAACCATTCATACC	At2g35640 qPCR
oSV211	CAACAAAGGCTACAACTGGATAATAC	At4g04580 qPCR
oSV212	GGACAATTCTTCTTCACCAGTTG	At4g04580 qPCR
oSV213	GCAAAGCGTCCGACAATATC	At4g23160 qPCR
oSV214	CGGGTCCTTTACGGGACTAC	At4g23160 qPCR
oSV215	AGACCACTTGGCCCTTATCTAC	At5g59680 qPCR
oSV216	GGTGTGCATCCAATGTGTAAG	At5g59680 qPCR
oSV234	TGGAACAGAGGAGAAGCCATAG	At4g11890 qPCR
oSV235	TCAGCATTCTCATCCACACATAG	At4g11890 qPCR
oSV236	CCAGAACCGAAGACCAGATTAC	At4g25100 qPCR
oSV237	GCCTCAAGTCTGGCACTTACAG	At4g25100 qPCR
oSV167	CTGCGACTCAGGGAATCTTCTAA	UBC qPCR
oSV168	TTGTGCCATTGAATTGAACCC	UBC qPCR
oHB86	TGGCTCCAGTCTTGGGTAAG	PDF2 qPCR
oHB87	GCCTGTCTTCAGCAAGTTCTAC	PDF2 qPCR

2. For other genotyping:

Oligonucleotide name	Sequence	Description
ML822	GGTGAAGTTGAAGTGAAAAAGATCGTTC	genotyping klu-2
ML823	CTTGGAAGTCAAACCAACGAAGGAA	genotping klu-2
cstf64.f7	ATTCAGATTAGTTACGGATAGAGA	genotyping of cstf64-1
cstf64.R2	ACGGGTTTTGTCAGTGC	genotyping of cstf64-1
ML5	AACCCACTGCTAGATTCTCCT	genotyping of bb-1
ML6	TAAAGTATAGAAGTCCACCCAAG	genotyping of bb-1
ML7	GAGAAGGCTCTGAAGGCTT	genotyping of bb-1

Other oligos (for cloning, sequencing) can be found in the oligo list (electronic excel file).

APPENDIX B. Plasmids

pSV1 (ML939_XbaI_rescue_gPAPS1)
pSV1b (pSV1XbaI gPAPS1 inserted in rev direction)
pSV1c (pSV1b_SDM_SalI_UTR5')
pSV2 (ML1297_AscI_gPAPS1_AscI
pSV3 (ML996_AscI_gPAPS1)
pSV4a (ML939_PstI_cPAPS1_KpnI)
pSV4b (ML939_PstI-cPAPS1_KpnI_35SUTR3'_EcoRI)
pSV4c (ML939_pPAPS1_cDNA_PAPS1)
pSV4d (pBarMAP AscI/pPAPS1UTR5-PAPS1cDNA-UTR3-35Sterm)
pSV5a (ML939-NcoI_gPAPS4_NcoI)
pSV5a rev (ML939-NcoI_gPAPS4_NcoI) gPAPS4 inverse
pSV5b (ML939_SDM_NcoIatATG_gPAPS4_NcoI)
pSV5c (pSV1_SDM_NcoI at ATG_gPAPS1)
pSV5d (ML939-pPAPS1::ATGgPAPS4-PAPS4UTR3')
pSV5e (pBarMAP/AscI/pPAPS1UTR5::ATGgPAPS4-PAPS4UTR3')
pSV6a (ML939-SEP3cDNA)
pSV6b (Op::SEP3 in PPpzp222)
pSV7a (Op::PI-Op::AP3)
pSV7b (Op::SEP3-Op::PI-Op::AP3)
pSV9a (AlcA::UTR5'gPAPS1UTR3'::35Sterm in ML1004)
pSV9b (pPAPS1::UTR5::EcoRI::AlcR)
pSV9c (pBarMAP_AscI_AlcA::UTR5gPAPS1::UTR3-35Sterm)

pSV9d.1 (pBarMAP:pPAPS1::AlcR-AlcA-genomic rescue PAPS1)
pSV9d.2 (pBarMAP:pPAPS1::AlcR-AlcA-genomic rescue PAPS1)
pSV9x (pPAPS1::UTR5 in ML939)
pSV10a (pBarMAP AlcA::YFP)
pSV10b.1 (pBarMAP:pPAPS1::AlcR-AlcA::YFP)
pSV10b.2 (pBarMAP:pPAPS1::AlcR-AlcA::YFP)
pSV11a (ML939-SDM cDNA NdeI PAPS1)
pSV11b (pET28a-PAPS1cDNAwt)
pSV11c (ML939-SDM cDNA NdeI PAPS1 SDM P313S - the ds39 mutated)
pSV11d (pET28a_cDNAPAPS1-P313S-the ds39 mutated version)
pSV12 (pSV5c SDM to have StuI site at Cterm ML939 genomic PAPS1)
pSV14(pPAPS1::UTR5PAPS1::NPAPS4::CtermPAPS1 (genomic)
pSV15 (pBarMAP/AscI/pPAPS1::UTR5PAPS1::ATGNPAPS4-CtermPAPS1 (genomic))
pSV16 (pPAPS1::GUS)
pSV17 (pPAPS2::NLSGUS)
pSV18a (pPAPS3::NLSGUS)
pSV18b (pPAPS3::NLSGUS)
pSV19a (pPAPS4::NLSGUS)
pSV19b (pPAPS4::NLSGUS)
pSV20 (pBarMAP-pPAPS2::GUS)
pSV21 (pBarMAP-pPAPS1::GUS)
pSV22a (pBarMAP-pPAPS3::GUS)
pSV22b (pBarMAP-pPAPS3::GUS)

pSV23a (pBarMAP-pPAPS4::GUS)
pSV23b (pBarMAP-pPAPS4::GUS)
pSV24(p35S::UTR5PAPS1-cDNAPAPS1-UTR3PAPS1::35Sterm in pML595)

APPENDIX C. Cloning procedures:

Unless otherwise stated, most of the plasmids are stored as vector NTI file. The Database of vector NTI is attached as an electronic folder.

1. pSV1 (pSV1 (ML939_XbaI_rescue_gPAPS1).

Genomic fragment of PAPS1 (5877bp) was isolated from from a BAC clone JAtY72B09 by digesting the BAC with XbaI and gel purify the fragment around 5.8 kb. The fragment was cloned into ML939 using XbaI site.

2. pSV1b (pSV1 XbaI gPAPS1 inserted in reverse direction)

pSV1b is similar to pSV1 but the XbaI-genomic PAPS1-XbaI fragment was inserted in the reverse direction in ML939 digested XbaI vector.

3. pSV1c (pSV1b_SDM_SalI_UTR5')

Introduce SalI site in front of the 5' UTR of genomic PAPS1 (in pSV1b) by overlap PCR using primer oSV114 and oSV115. Introduce the modified fragment back in to pSV1 by BstEII and KpnI site.

4. pSV2 (ML1297_AscI_gPAPS1_AscI)

pSV1 digested AscI insert + ML1297 digested with AscI vector.

5. pSV3 (ML996_AscI_gPAPS1)

pSV1 digested AscI insert + ML996 digested with AscI vector.

6. pSV4a (ML939_PstI_cPAPS1_KpnI)

cDNA of PAPS1 was amplified using primer oSV89 and oSV97 (introduce PstI site and KpnI sites to the PCR product). Clone this KpnI/PstI digested PCR product in ML939 digested KpnI/PstI vector.

7. pSV4b (ML939_PstI-cPAPS1_KpnI_35SUTR3'_EcoRI)

Insert the the 35S-3'UTR from ML1297 to pSV4a vector by EcoRI/KpnI site.

8. pSV4c (ML939_pPAPS1_cDNA_PAPS1)

Insert the promoter pPAPS1 from pSV1 (ScaI-BstEII) to vector pSV4b (ScaI-BstEII).

9. pSV4d (pBarMAP AscI/pPAPS1UTR5-PAPS1cDNA-UTR3-35Sterm)

Combine the PAPS1 insert from pSV4c to pBarMAP by AscI site .

10. pSV5a (ML939-NcoI_gPAPS4_NcoI)

Isolate the genomic PAPS4 fragment from BAC clone: JAt49G10 and JAt59J22 by NcoI digestion and gel purifying the 9906bp NcoI-genomicPAPS4-NcoI fragment. Clone the fragment into ML939 using NcoI site.

11. pSV5a rev (ML939-NcoI_gPAPS4_NcoI) gPAPS4 inverse

pSV5a rev has the NcoI-genomicPAPS4-NcoI inserted in to ML939/NcoI in the reverse direction compared to pSV5a.

12. pSV5b (ML939_SDM_NcoIatATG_gPAPS4_NcoI)

Introduce NcoI site at ATG of gPAPS4 by overlap PCR using primer oSV112 and oSV113. Clone back in to pSV5a using PmeI/XhoI site.

13. pSV5c (pSV1_SDM_NcoI at ATG_gPAPS1)

Introduce NcoI site at ATG of gPAPS1 in pSV1 by overlap PCR using primer oSV104 and oSV105. Clone back in to pSV1 using ClaI/BstEII sites.

14. pSV5d (ML939-pPAPS1::ATGgPAPS4-PAPS4UTR3')

Combine the insert genomic PAPS4 from pSV5b/NcoI and vector pSV5c/NcoI using NcoI site.

15. pSV5e (pBarMAP/AscI/pPAPS1UTR5::ATGgPAPS4-PAPS4UTR3')

Introduce the fragment in pSV5d in to pBarMAP by AscI site.

16. pSV6a (ML939-SEP3cDNA)

PCR amplify SEP3 cDNA using primer oSV122 and oSV123 introducing BamHI site at one end, the other end is blune. Clone in to ML939 using BamHI/SmaI site.

17. pSV6b (Op::SEP3 in PPpzp222)

The vector PPpzp222 do not have a full sequence map. The putative map of OP::PI and Op::AP3 are obtained from Robert Sablowski's lab:

Op::PI vector : LB-ECORI-SACI-SACII-NOTI-XBAI-OPERATOR-SPACER-OPERATOR-XBAI-SPEI CAMV TATA-NHEI-CLAI-SACI-XHOI-APOI-KPNI-NHEI-**PI**-NHEI-BAMHI-SPEI-XBAI-NOS-PSTI-HINDIII-RB

The SEP3 cDNA was digested with BamHI and then partially with KpnI (SEP3 has an internal KpnI site). Ligate this fragment in to Op::PI vector using KpnI and BamHI site.

18. pSV7a (Op::PI-Op::AP3)

Op::PI vector: LB-ECORI-SACI-SACII-NOTI-XBAI-OPERATOR-SPACER-OPERATOR-XBAI-SPEI CAMV TATA-NHEI-CLAI-SACI-XHOI-APOI-KPNI-NHEI-**PI**-NHEI-BAMHI-SPEI-XBAI-NOS-PSTI-HINDIII-RB

Op::AP3 vector: LB-ECORI-SACI-SACII-NOTI-XBAI-OPERATOR-SPACER-OPERATOR-XBAI-SPEI CAMV TATA-NHEI-CLAI-SACI-XHOI-APOI-KPNI-KPNI-**AP3 MINIGENE**-BAMHI-BAMHI-SPEI-XBAI-NOS-PSTI-HINDIII-RB

Cut Op:AP3 with NotI , blunt end, then cut with EcoRI, purify the backbone plasmid. Cut Op::PI with PstI, blunt the end, cut with EcoRI, purify fragment. Ligate.

19. pSV7b (Op::SEP3-Op::PI-Op::AP3)

Cut pSV7a with NotI, blunt end, then cut with EcoRI, purify the backbone plasmid. Cut pSV6b with PstI, blunt the end, cut with EcoRI, purify fragment. Ligate to give pSV7b.

Putative map of pSV7b

LB- ECORI SACI-SACII-NOTI-XBAI-OPERATOR-SPACER-OPERATOR-XBAI-SPEI CAMV TATA-NHEI-CLAI-SACI-XHOI-APOI- **SEP3**-BAMHI-SPEI-XBAI-NOS-PSTI blunt-bluntNOTI-XBAI-OPERATOR-SPACER-OPERATOR-XBAI-SPEI CAMV TATA-NHEI-CLAI-SACI-XHOI-APOI-KPNI-NHEI-**PI**-NHEI-BAMHI-SPEI-XBAI-NOS-PSTI-blunt-blunt-NOTI-XBAI-OPERATOR-SPACER-OPERATOR-XBAI-SPEI CAMV TATA-NHEI-CLAI-SACI-XHOI-APOI-KPNI-KPNI-**AP3 MINIGENE**-BAMHI-BAMHI-SPEI-XBAI-NOS-PSTI-HINDIII-RB

20. pSV9a (AlcA::UTR5'gPAPS1UTR3'::35Sterm in ML1004)

ML1004 digested with SalI/SmaI (vector) ligate to pSV1c digested with BsaAI/SalI (insert).

21. pSV9x (pPAPS1::UTR5 in ML939)

PCR amplify the promoter PAPS1 from pSV1 using primer T7 and oSV141 , digest PCR product with NcoI. Clone this insert into ML939 (NcoI/EcoRIblunt)vector.

22. pSV9b (pPAPS1::UTR5::EcoRI::AlcR)

pSV9x digested with EcoRI (insert) + EA6 digested with EcoRI (vector)

22. pSV9c (pBarMAP_AscI_AlcA::UTR5gPAPS1::UTR3-35Sterm)

pSV9a digested with AscI insert + pBarMAP digested with AscI vector. Remembered to select for reverse direction .

23. pSV9d.1 (pBarMAP:pPAPS1::AlcR-AlcA-genomic rescue PAPS1)

pSV9d.2 (pBarMAP:pPAPS1::AlcR-AlcA-genomic rescue PAPS1)

pSV9b digested with PacI insert + pSV9c digested with PacI vector. There are two directions of the pPAPS1::AlcR and AlcA::Genomic rescue PAPS1 (number 9d.1 and 9d.2)

24. pSV10a (pBarMAP AlcA::YFP)

ML1043 digested PacI/AscI insert + pBarMAP digested with PacI/AscI vector.

25. pSV10b.1 (pBarMAP:pPAPS1::AlcR-AlcA::YFP)

pSV10b.2 (pBarMAP:pPAPS1::AlcR-AlcA::YFP)

pSV9b digested with PacI insert + pSV10a digested with PacI vector; 10b.1 and 10b.2 refer to two directions of pPAPS1::AlcR and AlcA::YFP.

26. pSV11a (ML939-SDM cDNA NdeI PAPS1)

Introduce NdeI site in front of the ATG site in PAPS1 by PCR using oSV128 and oSV101. Digest PCR product with EcoO109I then purify 333bp fragment (insert). Digest pSV4a with PstI, blunt then digest with EcoO109I (vector). Ligate.

27. pSV11b (pET28a-PAPS1cDNAwt)

pSV11a digested with NdeI/EcoRI insert + pET28a NdeI/EcoRI vector.

28. pSV11c (ML939-SDM cDNA NdeI PAPS1 SDM P313S - the ds39 mutated)

Digest pSV11a with BglII/PmeI (vector). PCR using oSV78 and oSV80 using cDNA from *PAPS1-1* mutant. Digest PCR product with BglII/PmeI, isolate the 875bp-fragment (insert). Ligate.

29. pSV11d (pET28a_cDNAPAPS1-P313S-the ds39 mutated version)

pET28a (EcoRI/NdeI) vector + pSV11c (EcoRI/NdeI) insert.

30. pSV12 (pSV5c SDM to have StuI site at Cterm ML939 genomic PAPS1)

Introduce StuI site by overlap PCR using primer oSV202 and oSV203. Clone the insert into pSV5c by PmeI/Bsp1407I sites.

31. pSV14(pPAPS1::UTR5PAPS1::NPAPS4::CtermPAPS1 (genomic)

pV5d digested with StuI/PstI vector 7380 bp + pSV12 StuI/PstI insert (1950bp fragment).

32. pSV15 (pBarMAP/AscI/pPAPS1::UTR5PAPS1::ATGNPAPS4-CtermPAPS1 (genomic))

pSV14 digested with AscI insert + pBarMAP digested with AscI vector

33. pSV16 (pPAPS1::GUS)

pSV9x digested with KpnI + pKJ1 digested with KpnI vector.

34. pSV17 (pPAPS2::NLSGUS)

PCR amplify promoter PAPS2 by primer oSV189 and oSV190; digest with KpnI (insert) + pKJ1 digested with KpnI vector.

35. pSV18a (pPAPS3::NLSGUS); pSV18b (pPAPS3::NLSGUS)

PCR amplify promoter pPAPS3 by primer oSV246 and oSV247; digest with KpnI (insert) + pKJ1 digested with KpnI vector. a and b are clones derived from the pPAPS3 from two independent PCR reactions.

36. pSV19a (pPAPS4::NLSGUS) ; pSV19b (pPAPS4::NLSGUS)

PCR amplify promoter pPAPS4 by primer oSV244 and oSV245, digest with KpnI (insert) + pKJ1 digested with KpnI (vector). a and b are clones derived from the pPAPS3 from two independent PCR reactions.

37. pSV20 (pBarMAP-pPAPS2::GUS)

pSV17 digested with AscI insert + pBarMAP digested with AscI vector.

38. pSV21 (pBarMAP-pPAPS1::GUS)

pSV16 digested with AscI insert + pBarMAP digested with AscI vector.

39. pSV22a (pBarMAP-pPAPS3::GUS)

pSV18a digested with AscI insert + pBarMAP digested with AscI vector.

40. pSV22b (pBarMAP-pPAPS3::GUS)

pSV18b digested with AscI insert + pBarMAP digested with AscI vector.

41. pSV23 (pBarMAP-pPAPS4::GUS)

pSV19a digested with AscI insert + pBarMAP digested with AscI vector.

42. pSV23b (pBarMAP-pPAPS4::GUS)

pSV19b digested with AscI insert + pBarMAP digested with AscI vector.

43. pSV24(p35S::UTR5PAPS1-cDNAPAPS1-UTR3PAPS1::35Sterm in pML595)

pSV4a digested with PacI insert + pML595 digested with Pac I vector.

Abbreviations

35S	Promoter of the 35S gene from CaMV
aa	Amino acid(s)
ABA	Abscisic acid
ABI	ABSCISIC ACID-INSENSITIVE
ADP	Adenosine 5’ diphosphate
Agrobacterium	Agrobacterium *tumefaciens*
AIL	AINTEGUMENTA-LIKE
Ala	Alanine
AMP	Adenosine monophosphate
Amp	Ampicillin
AN3	ANGUSTIFOLIA3
ANT	AINTEGUMENTA
AP3	APETALA3
APE	APETALA
APS	Ammonium peroxide sulfate
Arabidopsis	*Arabidopsis thaliana*
ARF	AUXIN RESPONSE FACTOR
ARGOS	AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE
ARL	ARGOS-LIKE
At	*Arabidopsis thaliana*
ATP	Adenosine 5’ triphosphate
BB	BIG BROTHER
BOAAG	both organs (seedlings and inflorescences) abundance-affected genes
BOP	BLADE-ON-PETIOLE
BOTAG	both organs (seedlings and inflorescences) tail-affected genes
bp	Base pairs
BPE	BIGPETAL
BR	Brassinosteroid
BSA	Bovine serum albumin
Ca	Calcium
Cam	Chloramphenicol
CaMV	Cauliflower Mosaic Virus
Carb	Carbenicillin
CDB	Cyclin-destruction box
CDK	Cyclin-dependent kinase
cDNA	Complementary deoxyribonucleic acid
CDS	Coding sequence
CER	Controlled environment room
CFP	Cyan fluorescent protein
CIN	CINCINNATA
CLV	CLAVATA
cm	Centimetre
Col	Columbia, *Arabidopsis thaliana* accession
cPAP	canonical Poly(A) Polymerase
CTD	C-terminal domain
CYC	CYCLIN
CYP	Cytochrome P450
CZ	central zone
d	Deionised
dCTP	2’-deoxycytosine 5’-triphosphate
DELLAs	DELLA proteins

DEG	Differential expressed gene list
DMCs	Dispersed meristematic cells
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxynucleoside triphosphates
down	downregulated
ds	Double-stranded
DTT	Dithiothreitol
E. coli	*Escherichia coli*
e.g.	Exempli gratia
EDTA	Ethylendiaminetetraacetic acetate
EIL1	EIN3-Like
EIN	ETHYLENE INSENSITIVE
ER	ERECTA
er	Endoplasmatic reticulum
ERL	ERECTA-LIKE
EtBr	Ethidium bromide
EtOH	Ethanol
EYFP	Enhanced yellow fluorescent protein
F	Faraday
FAE	FASCIATE
FL	flowers/inflorescence tissue
FZR	Fizzy-related
g	Gramme
GA	Gibberellin
Gent	Gentamycin
GFP	Green fluorescent protein
GIF	GRF-INTERACTING FACTOR
GRF	GROWTH REGULATING FACTOR
GUS	ß-glucuronidase
h	Hour(s)
H_2O	Water
HCl	Hydrochloric acid
His	Histidine
i.e.	Id est
IAA	Indole-3-acetic acid, auxin
ICK	Inhibitor/interactor with cyclin-dependent kinase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JA	Jasmonic acid
JAG	JAGGED
Kan	Kanamycin
kb	Kilo base pairs
kDa	Kilodalton
KLU	KLUH / CYP78A5
kV	Kilo-volt
l	Liter
L1	Layer 1 (epidermal layer)
L2	Layer 2 (subepidermal layer)
L3	Layer 3 (layer forming vasculature and pith of plants)
LB	Lysogeny broth or Left border of the *T-DNA*
LE	leaves/seedling tissue
L*er*	Landsberg *errecta*, *Arabidopsis thaliana* accession
LFY	LEAFY
M	Molar

m	Milli
min	Minute(s)
mm	Millimeter
mRNA	Messenger RNA
MS-salts	Murashige and Skoog basal salt mixture
N	Amino
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NASC	Nottingham *Arabidopsis* stock centre
NLS	Nuclear localisation sequence
NUB	NUBBIN
OD	Optical density
ORF	Open reading frame
OSAAG	organ specific abundance-affected genes
OSTAG	organ specific tail-affected genes
PAP	Poly(A) Polymerase
PAPS	canonical Poly(A) Polymerase like protein in Arabiodopsis
PBS	Phosphate-buffered Saline
PCR	Polymerase chain reaction
PI	PISTILLATA
PMSF	Phenylmethylsufonyl fluoride
PPD	PEAPOD
PPT	Phosphinothricin (BASTA)
Pro	Proline
PZ	Peripheral zone
RNA	Ribonucleic acid
RNase	RNA nuclease
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase PCR
s	Seconds
S. cerevisiae	*Saccharomyces cerevisiae*
SAM	Shoot apical meristem
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SEP	SEPALLATA
Ser	Serine
SLY	SLEEPY
SOC	Super optimal broth with catabolite repression
Spec	Spectinomycin
ss	Single-stranded
SSC	Standard saline citrate
Str	Streptinomycin
SV40	Simian virus 40
TAE	Tris-acetate EDTA
TAP	Tandem affinity purification
Taq	DNA polymerase from *Thermus aquaticus*
TCP	TEOSINTE BRANCHED1, CYCLOIDEA, PCF
T-DNA	transferred DNA
TE	Tris/EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tet	Tetracycline
Tm	Melting temperature of primers

TOR	TARGET OF RAPAMYCIN
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
Tx	Generation x after transformation
U	Unit(s)
up	upregulated
UTR	Untranslated region
UV	Ultraviolet
V	Volt
v/v	Volume by volume
vol	Volume
w/v	Weight by volume
wt	Wild-type
WUS	WUSCHEL
X-Gal	Bromo-chloro-indolyl-galactopyranoside
YFP	Yellow fluorescent protein
µ	Micro
Ω	Ohm

Reference

2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature,* 408**,** 796-815.

ADAMSKI, N. M., ANASTASIOU, E., ERIKSSON, S., O'NEILL, C. M. & LENHARD, M. 2009. Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. *Proc Natl Acad Sci U S A,* 106**,** 20115-20.

ADDEPALLI, B. & HUNT, A. G. 2007. A novel endonuclease activity associated with the Arabidopsis ortholog of the 30-kDa subunit of cleavage and polyadenylation specificity factor. *Nucleic Acids Res,* 35**,** 4453-63.

ADDEPALLI, B., MEEKS, L. R., FORBES, K. P. & HUNT, A. G. 2004. Novel alternative splicing of mRNAs encoding poly(A) polymerases in Arabidopsis. *Biochim Biophys Acta,* 1679**,** 117-28.

ALVAREZ, J. P., GOLDSHMIDT, A., EFRONI, I., BOWMAN, J. L. & ESHED, Y. 2009. The NGATHA distal organ development genes are essential for style specification in Arabidopsis. *Plant Cell,* 21**,** 1373-93.

ANASTASIOU, E., KENZ, S., GERSTUNG, M., MACLEAN, D., TIMMER, J., FLECK, C. & LENHARD, M. 2007. Control of plant organ size by KLUH/CYP78A5-dependent intercellular signaling. *Dev Cell,* 13**,** 843-56.

APPONI, L. H., LEUNG, S. W., WILLIAMS, K. R., VALENTINI, S. R., CORBETT, A. H. & PAVLATH, G. K. 2010. Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet,* 19**,** 1058-65.

BALBO, P. B. & BOHM, A. 2007. Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure,* 15**,** 1117-31.

BALBO, P. B., TOTH, J. & BOHM, A. 2007. X-ray crystallographic and steady state fluorescence characterization of the protein dynamics of yeast polyadenylate polymerase. *J Mol Biol,* 366**,** 1401-15.

BARD, J., ZHELKOVSKY, A. M., HELMLING, S., EARNEST, T. N., MOORE, C. L. & BOHM, A. 2000. Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. *Science,* 289**,** 1346-9.

BEAUDOING, E. & GAUTHERET, D. 2001. Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. *Genome Res,* 11**,** 1520-6.

BEILHARZ, T. H. & PREISS, T. 2007. Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. *RNA,* 13**,** 982-97.

BELL, S. A. & HUNT, A. G. 2010. The Arabidopsis ortholog of the 77 kDa subunit of the cleavage stimulatory factor (AtCstF-77) involved in mRNA polyadenylation is an RNA-binding protein. *FEBS Lett,* 584**,** 1449-54.

BELOSTOTSKY, D. A. 2003. Unexpected complexity of poly(A)-binding protein gene families in flowering plants: three conserved lineages that are at least 200 million years old and possible auto- and cross-regulation. *Genetics,* 163**,** 311-9.

BELOSTOTSKY, D. A. & MEAGHER, R. B. 1993. Differential organ-specific expression of three poly(A)-binding-protein genes from Arabidopsis thaliana. *Proc Natl Acad Sci U S A,* 90**,** 6686-90.

BELOSTOTSKY, D. A. & MEAGHER, R. B. 1996. A pollen-, ovule-, and early embryo-specific poly(A) binding protein from Arabidopsis complements essential functions in yeast. *Plant Cell,* 8**,** 1261-75.

BENJAMINI, Y., DRAI, D., ELMER, G., KAFKAFI, N. & GOLANI, I. 2001. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res,* 125**,** 279-84.

BENOIT, B., MITOU, G., CHARTIER, A., TEMME, C., ZAESSINGER, S., WAHLE, E., BUSSEAU, I. & SIMONELIG, M. 2005. An essential cytoplasmic function for the nuclear poly(A) binding protein, PABP2, in poly(A) tail length control and early development in Drosophila. *Dev Cell,* 9**,** 511-22.

BIENROTH, S., KELLER, W. & WAHLE, E. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J,* 12**,** 585-94.

BOECK, R., TARUN, S., JR., RIEGER, M., DEARDORFF, J. A., MULLER-AUER, S. & SACHS, A. B. 1996. The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. *J Biol Chem,* 271**,** 432-8.

BOSSINGER, G. & SMYTH, D. R. 1996. Initiation patterns of flower and floral organ development in Arabidopsis thaliana. *Development,* 122**,** 1093-102.

BOWMAN, J. L., SMYTH, D. R. & MEYEROWITZ, E. M. 1989. Genes directing flower development in Arabidopsis. *Plant Cell,* 1**,** 37-52.

BRAVO, J., AGUILAR-HENONIN, L., OLMEDO, G. & GUZMAN, P. 2005. Four distinct classes of proteins as interaction partners of the PABC domain of Arabidopsis thaliana Poly(A)-binding proteins. *Mol Genet Genomics,* 272**,** 651-65.

BREUNINGER, H. & LENHARD, M. 2010. Control of tissue and organ growth in plants. *Curr Top Dev Biol,* 91**,** 185-220.

BRIGGS, M. W. & BUTLER, J. S. 1996. RNA polymerase III defects suppress a conditional-lethal poly(A) polymerase mutation in Saccharomyces cerevisiae. *Genetics,* 143**,** 1149-61.

BRIOUDES, F., JOLY, C., SZECSI, J., VARAUD, E., LEROUX, J., BELLVERT, F., BERTRAND, C. & BENDAHMANE, M. 2009. Jasmonate controls late development stages of petal growth in Arabidopsis thaliana. *Plant J,* 60**,** 1070-80.

BROWN, C. E. & SACHS, A. B. 1998. Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. *Mol Cell Biol,* 18**,** 6548-59.

BROWN, C. E., TARUN, S. Z., JR., BOECK, R. & SACHS, A. B. 1996. PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. *Mol Cell Biol,* 16**,** 5744-53.

BRUNE, C., MUNCHEL, S. E., FISCHER, N., PODTELEJNIKOV, A. V. & WEIS, K. 2005. Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA,* 11**,** 517-31.

CHAE, E., TAN, Q. K., HILL, T. A. & IRISH, V. F. 2008. An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development,* 135**,** 1235-45.

CHEKANOVA, J. A. & BELOSTOTSKY, D. A. 2003. Evidence that poly(A) binding protein has an evolutionarily conserved function in facilitating mRNA biogenesis and export. *RNA,* 9**,** 1476-90.

CHEKANOVA, J. A., SHAW, R. J. & BELOSTOTSKY, D. A. 2001. Analysis of an essential requirement for the poly(A) binding protein function using cross-species complementation. *Curr Biol,* 11**,** 1207-14.

CHEN, F., MACDONALD, C. C. & WILUSZ, J. 1995. Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res,* 23**,** 2614-20.

CHINCHILLA, D., ZIPFEL, C., ROBATZEK, S., KEMMERLING, B., NURNBERGER, T., JONES, J. D., FELIX, G. & BOLLER, T. 2007. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature,* 448**,** 497-500.

CLOUGH, S. J. & BENT, A. F. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J,* 16**,** 735-43.

COLGAN, D. F. & MANLEY, J. L. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev,* 11**,** 2755-66.

COLGAN, D. F., MURTHY, K. G., ZHAO, W., PRIVES, C. & MANLEY, J. L. 1998. Inhibition of poly(A) polymerase requires p34cdc2/cyclin B phosphorylation of multiple consensus and non-consensus sites. *EMBO J,* 17**,** 1053-62.

CORBESIER, L., VINCENT, C., JANG, S., FORNARA, F., FAN, Q., SEARLE, I., GIAKOUNTIS, A., FARRONA, S., GISSOT, L., TURNBULL, C. & COUPLAND, G. 2007. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science,* 316**,** 1030-3.

CRAWFORD, B. C., NATH, U., CARPENTER, R. & COEN, E. S. 2004. CINCINNATA controls both cell differentiation and growth in petal lobes and leaves of Antirrhinum. *Plant Physiol,* 135**,** 244-53.

CRICKMORE, M. A. & MANN, R. S. 2006. Hox control of organ size by regulation of morphogen production and mobility. *Science,* 313**,** 63-8.

DAS GUPTA, J., GU, H., CHERNOKALSKAYA, E., GAO, X. & SCHOENBERG, D. R. 1998. Identification of two cis-acting elements that independently regulate the length of poly(A) on Xenopus albumin pre-mRNA. *RNA,* 4**,** 766-76.

DE ALMEIDA ENGLER, J., DE VLEESSCHAUWER, V., BURSSENS, S., CELENZA, J. L., JR., INZE, D., VAN MONTAGU, M., ENGLER, G. & GHEYSEN, G. 1999. Molecular markers and cell cycle inhibitors show the importance of cell cycle progression in nematode-induced galls and syncytia. *Plant Cell,* 11**,** 793-808.

DELANEY, K. J., XU, R., ZHANG, J., LI, Q. Q., YUN, K. Y., FALCONE, D. L. & HUNT, A. G. 2006. Calmodulin interacts with and regulates the RNA-binding activity of an Arabidopsis polyadenylation factor subunit. *Plant Physiol,* 140**,** 1507-21.

DEPROST, D., YAO, L., SORMANI, R., MOREAU, M., LETERREUX, G., NICOLAI, M., BEDU, M., ROBAGLIA, C. & MEYER, C. 2007. The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. *EMBO Rep,* 8**,** 864-70.

DEWITTE, W., RIOU-KHAMLICHI, C., SCOFIELD, S., HEALY, J. M., JACQMARD, A., KILBY, N. J. & MURRAY, J. A. 2003. Altered cell cycle distribution, hyperplasia, and inhibited differentiation in Arabidopsis caused by the D-type cyclin CYCD3. *Plant Cell,* 15**,** 79-92.

DINNENY, J. R., WEIGEL, D. & YANOFSKY, M. F. 2006. NUBBIN and JAGGED define stamen and carpel shape in Arabidopsis. *Development,* 133**,** 1645-55.

DINNENY, J. R., YADEGARI, R., FISCHER, R. L., YANOFSKY, M. F. & WEIGEL, D. 2004. The role of JAGGED in shaping lateral organs. *Development,* 131**,** 1101-10.

DISCH, S., ANASTASIOU, E., SHARMA, V. K., LAUX, T., FLETCHER, J. C. & LENHARD, M. 2006. The E3 ubiquitin ligase BIG BROTHER controls arabidopsis organ size in a dosage-dependent manner. *Curr Biol,* 16**,** 272-9.

DONNELLY, P. M., BONETTA, D., TSUKAYA, H., DENGLER, R. E. & DENGLER, N. G. 1999. Cell cycling and cell enlargement in developing leaves of Arabidopsis. *Dev Biol,* 215**,** 407-19.

DUNN, E. F., HAMMELL, C. M., HODGE, C. A. & COLE, C. N. 2005. Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev,* 19**,** 90-103.

ELLIOTT, B. J., DATTAROY, T., MEEKS-MIDKIFF, L. R., FORBES, K. P. & HUNT, A. G. 2003. An interaction between an Arabidopsis poly(A) polymerase and a homologue of the 100 kDa subunit of CPSF. *Plant Mol Biol,* 51**,** 373-84.

ERIKSSON, S., STRANSFELD, L., ADAMSKI, N. M., BREUNINGER, H. & LENHARD, M. 2010. KLUH/CYP78A5-dependent growth signaling coordinates floral organ growth in Arabidopsis. *Curr Biol,* 20**,** 527-32.

FLETCHER, J. C., BRAND, U., RUNNING, M. P., SIMON, R. & MEYEROWITZ, E. M. 1999. Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science,* 283**,** 1911-4.

FORBES, K. P., ADDEPALLI, B. & HUNT, A. G. 2006. An Arabidopsis Fip1 homolog interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. *J Biol Chem,* 281**,** 176-86.

GALLONI, M. & EDGAR, B. A. 1999. Cell-autonomous and non-autonomous growth-defective mutants of Drosophila melanogaster. *Development,* 126**,** 2365-75.

GENTLEMAN, R. C., CAREY, V. J., BATES, D. M., BOLSTAD, B., DETTLING, M., DUDOIT, S., ELLIS, B., GAUTIER, L., GE, Y., GENTRY, J., HORNIK, K., HOTHORN, T., HUBER, W., IACUS, S., IRIZARRY, R., LEISCH, F., LI, C., MAECHLER, M., ROSSINI, A. J., SAWITZKI, G., SMITH, C., SMYTH, G., TIERNEY, L., YANG, J. Y. & ZHANG, J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol,* 5**,** R80.

GOLDSTROHM, A. C. & WICKENS, M. 2008. Multifunctional deadenylase complexes diversify mRNA control. *Nat Rev Mol Cell Biol,* 9**,** 337-44.

GRABER, J. H., CANTOR, C. R., MOHR, S. C. & SMITH, T. F. 1999. Genomic detection of new yeast pre-mRNA 3'-end-processing signals. *Nucleic Acids Res,* 27**,** 888-94.

GU, H., DAS GUPTA, J. & SCHOENBERG, D. R. 1999. The poly(A)-limiting element is a conserved cis-acting sequence that regulates poly(A) tail length on nuclear pre-mRNAs. *Proc Natl Acad Sci U S A,* 96**,** 8943-8.

GU, H. & SCHOENBERG, D. R. 2003. U2AF modulates poly(A) length control by the poly(A)-limiting element. *Nucleic Acids Res,* 31**,** 6264-71.

GUO, Z. & SHERMAN, F. 1996. Signals sufficient for 3'-end formation of yeast mRNA. *Mol Cell Biol,* 16**,** 2772-6.

GUPTA, J. D., GU, H. & SCHOENBERG, D. R. 2001. Position and sequence requirements for poly(A) length regulation by the poly(A) limiting element. *RNA,* 7**,** 1034-42.

HECTOR, R. E., NYKAMP, K. R., DHEUR, S., ANDERSON, J. T., NON, P. J., URBINATI, C. R., WILSON, S. M., MINVIELLE-SEBASTIA, L. & SWANSON, M. S. 2002. Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J,* 21**,** 1800-10.

HENRIQUES, R., MAGYAR, Z., MONARDES, A., KHAN, S., ZALEJSKI, C., ORELLANA, J., SZABADOS, L., DE LA TORRE, C., KONCZ, C. & BOGRE, L. 2010. Arabidopsis S6 kinase mutants display chromosome instability and altered RBR1-E2F pathway activity. *Embo J,* 29**,** 2979-93.

HERR, A. J., MOLNAR, A., JONES, A. & BAULCOMBE, D. C. 2006. Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis. *Proc Natl Acad Sci U S A,* 103**,** 14994-5001.

HOLBEIN, S., WENGI, A., DECOURTY, L., FREIMOSER, F. M., JACQUIER, A. & DICHTL, B. 2009. Cordycepin interferes with 3' end formation in yeast independently of its potential to terminate RNA chain elongation. *RNA,* 15**,** 837-49.

HORIGUCHI, G., FUJIKURA, U., FERJANI, A., ISHIKAWA, N. & TSUKAYA, H. 2006. Large-scale histological analysis of leaf mutants using two simple leaf observation methods: identification of novel genetic pathways governing the size and shape of leaves. *Plant J,* 48**,** 638-44.

HORIGUCHI, G., KIM, G. T. & TSUKAYA, H. 2005. The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. *Plant J,* 43**,** 68-78.

HORNYIK, C., TERZI, L. C. & SIMPSON, G. G. 2010. The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev Cell,* 18**,** 203-13.

HORVATH, B. M., MAGYAR, Z., ZHANG, Y., HAMBURGER, A. W., BAKO, L., VISSER, R. G., BACHEM, C. W. & BOGRE, L. 2006. EBP1 regulates organ size through cell growth and proliferation in plants. *Embo J,* 25**,** 4909-20.

HU, Y., POH, H. M. & CHUA, N. H. 2006. The Arabidopsis ARGOS-LIKE gene regulates cell expansion during organ growth. *Plant J,* 47**,** 1-9.

HU, Y., XIE, Q. & CHUA, N. H. 2003. The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. *Plant Cell,* 15**,** 1951-61.

HUARTE, J., STUTZ, A., O'CONNELL, M. L., GUBLER, P., BELIN, D., DARROW, A. L., STRICKLAND, S. & VASSALLI, J. D. 1992. Transient translational silencing by reversible mRNA deadenylation. *Cell,* 69**,** 1021-30.

HUNT, A. G. 2008. Messenger RNA 3' end formation in plants. *Curr Top Microbiol Immunol,* 326**,** 151-77.

HUNT, A. G., MEEKS, L. R., FORBES, K. P., DAS GUPTA, J. & MOGEN, B. D. 2000. Nuclear and chloroplast poly(A) polymerases from plants share a novel biochemical property. *Biochem Biophys Res Commun,* 272**,** 174-81.

HUNT, A. G., XU, R., ADDEPALLI, B., RAO, S., FORBES, K. P., MEEKS, L. R., XING, D., MO, M., ZHAO, H., BANDYOPADHYAY, A., DAMPANABOINA, L., MARION, A., VON LANKEN, C. & LI, Q. Q. 2008. Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. *BMC Genomics,* 9**,** 220.

ITO, T., CHIBA, T., OZAWA, R., YOSHIDA, M., HATTORI, M. & SAKAKI, Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A,* 98**,** 4569-74.

JIA, H., WANG, X., LIU, F., GUENTHER, U. P., SRINIVASAN, S., ANDERSON, J. T. & JANKOWSKY, E. 2011. The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. *Cell,* 145**,** 890-901.

JUGE, F., ZAESSINGER, S., TEMME, C., WAHLE, E. & SIMONELIG, M. 2002. Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in Drosophila. *EMBO J,* 21**,** 6603-13.

KASHIWABARA, S., NOGUCHI, J., ZHUANG, T., OHMURA, K., HONDA, A., SUGIURA, S., MIYAMOTO, K., TAKAHASHI, S., INOUE, K., OGURA, A. & BABA, T. 2002. Regulation of spermatogenesis by testis-specific, cytoplasmic poly(A) polymerase TPAP. *Science,* 298**,** 1999-2002.

KAWADE, K., HORIGUCHI, G. & TSUKAYA, H. 2010. Non-cell-autonomously coordinated organ size regulation in leaf development. *Development,* 137**,** 4221-7.

KEITH, B. & CHUA, N. H. 1986. Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. *EMBO J,* 5**,** 2419-25.

KESSLER, M. M., HENRY, M. F., SHEN, E., ZHAO, J., GROSS, S., SILVER, P. A. & MOORE, C. L. 1997. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes Dev,* 11**,** 2545-56.

KIM, J. H., CHOI, D. & KENDE, H. 2003. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. *Plant J,* 36**,** 94-104.
KIM, J. H. & KENDE, H. 2004. A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. *Proc Natl Acad Sci U S A,* 101**,** 13374-9.
KLUCHER, K. M., CHOW, H., REISER, L. & FISCHER, R. L. 1996. The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. *Plant Cell,* 8**,** 137-53.
KRIZEK, B. 2009. AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. *Plant Physiol,* 150**,** 1916-29.
KUHN, U., GUNDEL, M., KNOTH, A., KERWITZ, Y., RUDEL, S. & WAHLE, E. 2009. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J Biol Chem,* 284**,** 22803-14.
KUREPA, J., WANG, S., LI, Y., ZAITLIN, D., PIERCE, A. J. & SMALLE, J. A. 2009. Loss of 26S proteasome function leads to increased cell size and decreased cell number in Arabidopsis shoot organs. *Plant Physiol,* 150**,** 178-89.
KYRIAKOPOULOU, C. B., NORDVARG, H. & VIRTANEN, A. 2001. A novel nuclear human poly(A) polymerase (PAP), PAP gamma. *J Biol Chem,* 276**,** 33504-11.
LANGE, H., SEMENT, F. M., CANADAY, J. & GAGLIARDI, D. 2009. Polyadenylation-assisted RNA degradation processes in plants. *Trends Plant Sci,* 14**,** 497-504.
LARSON-RABIN, Z., LI, Z., MASSON, P. H. & DAY, C. D. 2009. FZR2/CCS52A1 expression is a determinant of endoreduplication and cell expansion in Arabidopsis. *Plant Physiol,* 149**,** 874-84.
LI, Y., ZHENG, L., CORKE, F., SMITH, C. & BEVAN, M. W. 2008. Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes Dev,* 22**,** 1331-6.
LIANG, W., LI, C., LIU, F., JIANG, H., LI, S., SUN, J. & WU, X. 2009. The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. *Cell Res,* 19**,** 307-16.
LIU, F., MARQUARDT, S., LISTER, C., SWIEZEWSKI, S. & DEAN, C. 2010. Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. *Science,* 327**,** 94-7.
LOKE, J. C., STAHLBERG, E. A., STRENSKI, D. G., HAAS, B. J., WOOD, P. C. & LI, Q. Q. 2005. Compilation of mRNA polyadenylation signals in Arabidopsis revealed a new signal element and potential secondary structures. *Plant Physiol,* 138**,** 1457-68.
LUKOWITZ, W., GILLMOR, C. S. & SCHEIBLE, W. R. 2000. Positional cloning in Arabidopsis. Why it feels good to have a genome initiative working for you. *Plant Physiol,* 123**,** 795-805.
LUTZ, C. S. 2008. Alternative polyadenylation: a twist on mRNA 3' end formation. *ACS Chem Biol,* 3**,** 609-17.
LUTZ, C. S. & MOREIRA, A. 2011. Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. *WIREs RNA,* 2**,** 23-31.
MACDONALD, C. C. & REDONDO, J. L. 2002. Reexamining the polyadenylation signal: were we wrong about AAUAAA? *Mol Cell Endocrinol,* 190**,** 1-8.
MACDONALD, C. C., WILUSZ, J. & SHENK, T. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol Cell Biol,* 14**,** 6647-54.
MACDONALD, M. H., MOGEN, B. D. & HUNT, A. G. 1991. Characterization of the polyadenylation signal from the T-DNA-encoded octopine synthase gene. *Nucleic Acids Res,* 19**,** 5575-81.
MANDEL, C. R., KANEKO, S., ZHANG, H., GEBAUER, D., VETHANTHAM, V., MANLEY, J. L. & TONG, L. 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature,* 444**,** 953-6.
MANDEL, C. R. & TONG, L. 2007. How to get all "A"s in polyadenylation. *Structure,* 15**,** 1024-6.
MANGUS, D. A., EVANS, M. C., AGRIN, N. S., SMITH, M., GONGIDI, P. & JACOBSON, A. 2004. Positive and negative regulation of poly(A) nuclease. *Mol Cell Biol,* 24**,** 5521-33.
MANZANO, D., MARQUARDT, S., JONES, A. M., BAURLE, I., LIU, F. & DEAN, C. 2009. Altered interactions within FY/AtCPSF complexes required for Arabidopsis FCA-mediated chromatin silencing. *Proc Natl Acad Sci U S A,* 106**,** 8772-7.

MARTIN, G., JENO, P. & KELLER, W. 1999. Mapping of ATP binding regions in poly(A) polymerases by photoaffinity labeling and by mutational analysis identifies a domain conserved in many nucleotidyltransferases. *Protein Sci,* 8**,** 2380-91.

MARTIN, G., KELLER, W. & DOUBLIE, S. 2000. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J,* 19**,** 4193-203.

MARTIN, G., MOGLICH, A., KELLER, W. & DOUBLIE, S. 2004. Biochemical and structural insights into substrate binding and catalytic mechanism of mammalian poly(A) polymerase. *J Mol Biol,* 341**,** 911-25.

MASON, P. J., ELKINGTON, J. A., LLOYD, M. M., JONES, M. B. & WILLIAMS, J. G. 1986. Mutations downstream of the polyadenylation site of a Xenopus beta-globin mRNA affect the position but not the efficiency of 3' processing. *Cell,* 46**,** 263-70.

MEEKS, L. R., ADDEPALLI, B. & HUNT, A. G. 2009. Characterization of genes encoding poly(A) polymerases in plants: evidence for duplication and functional specialization. *PLoS One,* 4**,** e8082.

MEIJER, H. A., BUSHELL, M., HILL, K., GANT, T. W., WILLIS, A. E., JONES, P. & DE MOOR, C. H. 2007. A novel method for poly(A) fractionation reveals a large population of mRNAs with a short poly(A) tail in mammalian cells. *Nucleic Acids Res,* 35**,** e132.

MEIJER, H. A. & DE MOOR, C. H. 2011. Fractionation of mRNA based on the length of the poly(A) tail. *Methods Mol Biol,* 703**,** 123-35.

MEINKE, G., EZEOKONKWO, C., BALBO, P., STAFFORD, W., MOORE, C. & BOHM, A. 2008. Structure of yeast poly(A) polymerase in complex with a peptide from Fip1, an intrinsically disordered protein. *Biochemistry,* 47**,** 6859-69.

MENAND, B., DESNOS, T., NUSSAUME, L., BERGER, F., BOUCHEZ, D., MEYER, C. & ROBAGLIA, C. 2002. Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene. *Proc Natl Acad Sci U S A,* 99**,** 6422-7.

MILLEVOI, S. & VAGNER, S. 2010. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res,* 38**,** 2757-74.

MINVIELLE-SEBASTIA, L., PREKER, P. J., WIEDERKEHR, T., STRAHM, Y. & KELLER, W. 1997. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc Natl Acad Sci U S A,* 94**,** 7897-902.

MIZUKAMI, Y. & FISCHER, R. L. 2000. Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci U S A,* 97**,** 942-7.

MOGEN, B. D., MACDONALD, M. H., LEGGEWIE, G. & HUNT, A. G. 1992. Several distinct types of sequence elements are required for efficient mRNA 3' end formation in a pea rbcS gene. *Mol Cell Biol,* 12**,** 5406-14.

MURATA, T., NAGASO, H., KASHIWABARA, S., BABA, T., OKANO, H. & YOKOYAMA, K. K. 2001. The hiiragi gene encodes a poly(A) polymerase, which controls the formation of the wing margin in Drosophila melanogaster. *Dev Biol,* 233**,** 137-47.

NAG, A., KING, S. & JACK, T. 2009. miR319a targeting of TCP4 is critical for petal growth and development in Arabidopsis. *Proc Natl Acad Sci U S A,* 106**,** 22534-9.

NISHIMURA, N., KITAHATA, N., SEKI, M., NARUSAKA, Y., NARUSAKA, M., KUROMORI, T., ASAMI, T., SHINOZAKI, K. & HIRAYAMA, T. 2005. Analysis of ABA hypersensitive germination2 revealed the pivotal functions of PARN in stress response in Arabidopsis. *Plant J,* 44**,** 972-84.

NISHIMURA, N., OKAMOTO, M., NARUSAKA, M., YASUDA, M., NAKASHITA, H., SHINOZAKI, K., NARUSAKA, Y. & HIRAYAMA, T. 2009. ABA hypersensitive germination2-1 causes the activation of both abscisic acid and salicylic acid responses in Arabidopsis. *Plant Cell Physiol,* 50**,** 2112-22.

NOLE-WILSON, S., TRANBY, T. L. & KRIZEK, B. A. 2005. AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant Mol Biol,* 57**,** 613-28.

NORBERG, M., HOLMLUND, M. & NILSSON, O. 2005. The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. *Development,* 132**,** 2203-13.

OHNISHI, T., SZATMARI, A. M., WATANABE, B., FUJITA, S., BANCOS, S., KONCZ, C., LAFOS, M., SHIBATA, K., YOKOTA, T., SAKATA, K., SZEKERES, M. & MIZUTANI, M. 2006. C-23 hydroxylation by Arabidopsis CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. *Plant Cell,* 18**,** 3275-88.

OHNO, C. K., REDDY, G. V., HEISLER, M. G. & MEYEROWITZ, E. M. 2004. The Arabidopsis JAGGED gene encodes a zinc finger protein that promotes leaf tissue development. *Development,* 131**,** 1111-22.

OSHLACK, A., EMSLIE, D., CORCORAN, L. M. & SMYTH, G. K. 2007. Normalization of boutique two-color microarrays with a high proportion of differentially expressed probes. *Genome Biol,* 8**,** R2.

PAGNUSSAT, G. C., YU, H. J., NGO, Q. A., RAJANI, S., MAYALAGU, S., JOHNSON, C. S., CAPRON, A., XIE, L. F., YE, D. & SUNDARESAN, V. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. *Development,* 132**,** 603-14.

PALANIVELU, R., BELOSTOTSKY, D. A. & MEAGHER, R. B. 2000. Conserved expression of Arabidopsis thaliana poly (A) binding protein 2 (PAB2) in distinct vegetative and reproductive tissues. *Plant J,* 22**,** 199-210.

PALATNIK, J. F., ALLEN, E., WU, X., SCHOMMER, C., SCHWAB, R., CARRINGTON, J. C. & WEIGEL, D. 2003. Control of leaf morphogenesis by microRNAs. *Nature,* 425**,** 257-63.

PARKER, R. & SONG, H. 2004. The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol,* 11**,** 121-7.

PATEL, D. & BUTLER, J. S. 1992. Conditional defect in mRNA 3' end processing caused by a mutation in the gene for poly(A) polymerase. *Mol Cell Biol,* 12**,** 3297-304.

PENG, J., MURRAY, E. L. & SCHOENBERG, D. R. 2005. The poly(A)-limiting element enhances mRNA accumulation by increasing the efficiency of pre-mRNA 3' processing. *RNA,* 11**,** 958-65.

PENG, J. & SCHOENBERG, D. R. 2005. mRNA with a <20-nt poly(A) tail imparted by the poly(A)-limiting element is translated as efficiently in vivo as long poly(A) mRNA. *RNA,* 11**,** 1131-40.

PREKER, P. J., LINGNER, J., MINVIELLE-SEBASTIA, L. & KELLER, W. 1995. The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell,* 81**,** 379-89.

PROUDFOOT, N. 2004. New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr Opin Cell Biol,* 16**,** 272-8.

PROWELLER, A. & BUTLER, S. 1994. Efficient translation of poly(A)-deficient mRNAs in Saccharomyces cerevisiae. *Genes Dev,* 8**,** 2629-40.

RAMAKERS, C., RUIJTER, J. M., DEPREZ, R. H. & MOORMAN, A. F. 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett,* 339**,** 62-6.

RAO, M. N., CHERNOKALSKAYA, E. & SCHOENBERG, D. R. 1996. Regulated nuclear polyadenylation of Xenopus albumin pre-mRNA. *Nucleic Acids Res,* 24**,** 4078-83.

REDDY, G. V., HEISLER, M. G., EHRHARDT, D. W. & MEYEROWITZ, E. M. 2004. Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana. *Development,* 131**,** 4225-37.

REDEI, G. P. 1962. Supervital Mutants of Arabidopsis. *Genetics,* 47**,** 443-60.

REINA-PINTO, J. J., VOISIN, D., TEODOR, R. & YEPHREMOV, A. 2010. Probing differentially expressed genes against a microarray database for in silico suppressor/enhancer and inhibitor/activator screens. *Plant J,* 61**,** 166-75.

REVERDATTO, S. V., DUTKO, J. A., CHEKANOVA, J. A., HAMILTON, D. A. & BELOSTOTSKY, D. A. 2004. mRNA deadenylation by PARN is essential for embryogenesis in higher plants. *RNA,* 10**,** 1200-14.

RODRIGUEZ, R. E., MECCHIA, M. A., DEBERNARDI, J. M., SCHOMMER, C., WEIGEL, D. & PALATNIK, J. F. 2010. Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. *Development,* 137**,** 103-12.

ROSLAN, H. A., SALTER, M. G., WOOD, C. D., WHITE, M. R., CROFT, K. P., ROBSON, F., COUPLAND, G., DOONAN, J., LAUFS, P., TOMSETT, A. B. & CADDICK, M. X. 2001. Characterization of the ethanol-inducible alc gene-expression system in Arabidopsis thaliana. *Plant J,* 28**,** 225-35.

ROTHNIE, H. M. 1996. Plant mRNA 3'-end formation. *Plant Mol Biol,* 32**,** 43-61.

ROTHNIE, H. M., CHEN, G., FUTTERER, J. & HOHN, T. 2001. Polyadenylation in rice tungro bacilliform virus: cis-acting signals and regulation. *J Virol,* 75**,** 4184-94.

ROTHNIE, H. M., REID, J. & HOHN, T. 1994. The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3'-end formation in plants. *EMBO J,* 13**,** 2200-10.

RYNER, L. C., TAKAGAKI, Y. & MANLEY, J. L. 1989. Multiple forms of poly(A) polymerases purified from HeLa cells function in specific mRNA 3'-end formation. *Mol Cell Biol,* 9**,** 4229-38.

SACHS, A. B. & DAVIS, R. W. 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell,* 58**,** 857-67.

SALLES, F. J., RICHARDS, W. G. & STRICKLAND, S. 1999. Assaying the polyadenylation state of mRNAs. *Methods,* 17**,** 38-45.

SANFACON, H. & HOHN, T. 1990. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. *Nature,* 346**,** 81-4.

SAVALDI-GOLDSTEIN, S., PETO, C. & CHORY, J. 2007. The epidermis both drives and restricts plant shoot growth. *Nature,* 446**,** 199-202.

SAWICKI, S. G., JELINEK, W. & DARNELL, J. E. 1977. 3'-Terminal addition to HeLa cell nuclear and cytoplasmic poly (A). *J Mol Biol,* 113**,** 219-35.

SCHOOF, H., LENHARD, M., HAECKER, A., MAYER, K. F., JURGENS, G. & LAUX, T. 2000. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell,* 100**,** 635-44.

SCHRUFF, M. C., SPIELMAN, M., TIWARI, S., ADAMS, S., FENBY, N. & SCOTT, R. J. 2006. The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. *Development,* 133**,** 251-61.

SESSIONS, A., WEIGEL, D. & YANOFSKY, M. F. 1999. The Arabidopsis thaliana MERISTEM LAYER 1 promoter specifies epidermal expression in meristems and young primordia. *Plant J,* 20**,** 259-63.

SHEETS, M. D., OGG, S. C. & WICKENS, M. P. 1990. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic Acids Res,* 18**,** 5799-805.

SHEN, Y., JI, G., HAAS, B. J., WU, X., ZHENG, J., REESE, G. J. & LI, Q. Q. 2008. Genome level analysis of rice mRNA 3'-end processing signals and alternative polyadenylation. *Nucleic Acids Res,* 36**,** 3150-61.

SHI, Y., DI GIAMMARTINO, D. C., TAYLOR, D., SARKESHIK, A., RICE, W. J., YATES, J. R., 3RD, FRANK, J. & MANLEY, J. L. 2009. Molecular architecture of the human pre-mRNA 3' processing complex. *Mol Cell,* 33**,** 365-76.

SHIMAZU, T., HORINOUCHI, S. & YOSHIDA, M. 2007. Multiple histone deacetylases and the CREB-binding protein regulate pre-mRNA 3'-end processing. *J Biol Chem,* 282**,** 4470-8.

SIMPSON, G. G., DIJKWEL, P. P., QUESADA, V., HENDERSON, I. & DEAN, C. 2003. FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. *Cell,* 113**,** 777-87.

SONODA, Y., SAKO, K., MAKI, Y., YAMAZAKI, N., YAMAMOTO, H., IKEDA, A. & YAMAGUCHI, J. 2009. Regulation of leaf organ size by the Arabidopsis RPT2a 19S proteasome subunit. *Plant J,* 60**,** 68-78.

SUGIMOTO-SHIRASU, K., ROBERTS, G. R., STACEY, N. J., MCCANN, M. C., MAXWELL, A. & ROBERTS, K. 2005. RHL1 is an essential component of the plant DNA topoisomerase VI complex and is required for ploidy-dependent cell growth. *Proc Natl Acad Sci U S A,* 102**,** 18736-41.

SUGIMOTO-SHIRASU, K., STACEY, N. J., CORSAR, J., ROBERTS, K. & MCCANN, M. C. 2002. DNA topoisomerase VI is essential for endoreduplication in Arabidopsis. *Curr Biol,* 12**,** 1782-6.

SZECSI, J., JOLY, C., BORDJI, K., VARAUD, E., COCK, J. M., DUMAS, C. & BENDAHMANE, M. 2006. BIGPETALp, a bHLH transcription factor is involved in the control of Arabidopsis petal size. *EMBO J,* 25**,** 3912-20.

THIMM, O., BLASING, O., GIBON, Y., NAGEL, A., MEYER, S., KRUGER, P., SELBIG, J., MULLER, L. A., RHEE, S. Y. & STITT, M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J,* 37**,** 914-39.

THURESSON, A. C., ASTROM, J., ASTROM, A., GRONVIK, K. O. & VIRTANEN, A. 1994. Multiple forms of poly(A) polymerases in human cells. *Proc Natl Acad Sci U S A,* 91**,** 979-83.

TIAN, B., HU, J., ZHANG, H. & LUTZ, C. S. 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res,* 33**,** 201-12.

TOPALIAN, S. L., KANEKO, S., GONZALES, M. I., BOND, G. L., WARD, Y. & MANLEY, J. L. 2001. Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors. *Mol Cell Biol,* 21**,** 5614-23.

TRIGUEROS, M., NAVARRETE-GOMEZ, M., SATO, S., CHRISTENSEN, S. K., PELAZ, S., WEIGEL, D., YANOFSKY, M. F. & FERRANDIZ, C. 2009. The NGATHA genes direct style development in the Arabidopsis gynoecium. *Plant Cell,* 21**,** 1394-409.

TSUGE, T., TSUKAYA, H. & UCHIMIYA, H. 1996. Two independent and polarized processes of cell elongation regulate leaf blade expansion in Arabidopsis thaliana (L.) Heynh. *Development,* 122**,** 1589-600.

TZAFRIR, I., PENA-MURALLA, R., DICKERMAN, A., BERG, M., ROGERS, R., HUTCHENS, S., SWEENEY, T. C., MCELVER, J., AUX, G., PATTON, D. & MEINKE, D. 2004. Identification of genes required for embryo development in Arabidopsis. *Plant Physiol,* 135**,** 1206-20.

VAN HOOF, A., LENNERTZ, P. & PARKER, R. 2000. Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol Cell Biol,* 20**,** 441-52.

VAN LAREBEKE, N., ENGLER, G., HOLSTERS, M., VAN DEN ELSACKER, S., ZAENEN, I., SCHILPEROORT, R. A. & SCHELL, J. 1974. Large plasmid in Agrobacterium tumefaciens essential for crown gall-inducing ability. *Nature,* 252**,** 169-70.

VETHANTHAM, V., RAO, N. & MANLEY, J. L. 2008. Sumoylation regulates multiple aspects of mammalian poly(A) polymerase function. *Genes Dev,* 22**,** 499-511.

WAHLE, E. 1995. Poly(A) tail length control is caused by termination of processive synthesis. *J Biol Chem,* 270**,** 2800-8.

WAHLE, E. & RUEGSEGGER, U. 1999. 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev,* 23**,** 277-95.

WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.

WEIGEL, D. & MEYEROWITZ, E. M. 1994. The ABCs of floral homeotic genes. *Cell,* 78**,** 203-9.

WHITE, D. W. 2006. PEAPOD regulates lamina size and curvature in Arabidopsis. *Proc Natl Acad Sci U S A,* 103**,** 13238-43.

WIEDERKEHR, T., PRETOT, R. F. & MINVIELLE-SEBASTIA, L. 1998. Synthetic lethal interactions with conditional poly(A) polymerase alleles identify LCP5, a gene involved in 18S rRNA maturation. *RNA,* 4**,** 1357-72.

WINTER, C. M., AUSTIN, R. S., BLANVILLAIN-BAUFUME, S., REBACK, M. A., MONNIAUX, M., WU, M. F., SANG, Y., YAMAGUCHI, A., YAMAGUCHI, N., PARKER, J. E., PARCY, F., JENSEN, S. T., LI, H. & WAGNER, D. 2011. LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev Cell,* 20**,** 430-43.

WU, L., UEDA, T. & MESSING, J. 1993. 3'-end processing of the maize 27 kDa zein mRNA. *Plant J,* 4**,** 535-44.

WU, X., LIU, M., DOWNIE, B., LIANG, C., JI, G., LI, Q. Q. & HUNT, A. G. 2011. Genome-wide landscape of polyadenylation in Arabidopsis provides evidence for extensive alternative polyadenylation. *Proc Natl Acad Sci U S A*.

XING, A., MOON, B. P., MILLS, K. M., FALCO, S. C. & LI, Z. 2010. Revealing frequent alternative polyadenylation and widespread low-level transcription read-through of novel plant transcription terminators. *Plant Biotechnol J,* 8**,** 772-82.

XING, D., ZHAO, H. & LI, Q. Q. 2008a. Arabidopsis CLP1-SIMILAR PROTEIN3, an ortholog of human polyadenylation factor CLP1, functions in gametophyte, embryo, and postembryonic development. *Plant Physiol,* 148**,** 2059-69.

XING, D., ZHAO, H., XU, R. & LI, Q. Q. 2008b. Arabidopsis PCFS4, a homologue of yeast polyadenylation factor Pcf11p, regulates FCA alternative processing and promotes flowering time. *Plant J,* 54**,** 899-910.

XU, R., YE, X. & QUINN LI, Q. 2004. AtCPSF73-II gene encoding an Arabidopsis homolog of CPSF 73 kDa subunit is critical for early embryo development. *Gene,* 324**,** 35-45.

XU, R., ZHAO, H., DINKINS, R. D., CHENG, X., CARBERRY, G. & LI, Q. Q. 2006. The 73 kD subunit of the cleavage and polyadenylation specificity factor (CPSF) complex affects reproductive development in Arabidopsis. *Plant Mol Biol,* 61**,** 799-815.

YAMASHITA, A., CHANG, T. C., YAMASHITA, Y., ZHU, W., ZHONG, Z., CHEN, C. Y. & SHYU, A. B. 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol,* 12**,** 1054-63.

YAO, Y., SONG, L., KATZ, Y. & GALILI, G. 2002. Cloning and characterization of Arabidopsis homologues of the animal CstF complex that regulates 3' mRNA cleavage and polyadenylation. *J Exp Bot,* 53**,** 2277-8.

ZHAO, H., XING, D. & LI, Q. Q. 2009. Unique features of plant cleavage and polyadenylation specificity factor revealed by proteomic studies. *Plant Physiol,* 151**,** 1546-56.

ZHAO, J., HYMAN, L. & MOORE, C. 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev,* 63**,** 405-45.

ZHAO, W. & MANLEY, J. L. 1996. Complex alternative RNA processing generates an unexpected diversity of poly(A) polymerase isoforms. *Mol Cell Biol,* 16**,** 2378-86.

ZHELKOVSKY, A., HELMLING, S., BOHM, A. & MOORE, C. 2004. Mutations in the middle domain of yeast poly(A) polymerase affect interactions with RNA but not ATP. *RNA,* 10**,** 558-64.

ZIMMER, S. L., SCHEIN, A., ZIPOR, G., STERN, D. B. & SCHUSTER, G. 2009. Polyadenylation in Arabidopsis and Chlamydomonas organelles: the input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. *Plant J,* 59**,** 88-99.

ZONDLO, S. C. & IRISH, V. F. 1999. CYP78A5 encodes a cytochrome P450 that marks the shoot apical meristem boundary in Arabidopsis. *Plant J,* 19**,** 259-68.