2009

Analysis of the Arabidopsis Polyadenylation Factors PAP1, CstF64 and CstF77 and their characteristic inter-relationship

Amrita Bandyopadhyay
University of Kentucky, amritam29@uky.edu

ABSTRACT OF THESIS

"Analysis of the *Arabidopsis* Polyadenylation Factors PAP1, CstF64 and CstF77 and their characteristic inter-relationship"

3'-end modification by polyadenylation is a ubiquitous feature of almost all eukaryotic mRNA species and is catalyzed by a consortium of enzymes, the polyadenylation factors. Poly(A) polymerase (PAP), the enzyme catalyzing the addition of adenosine residues during the polyadenylation stage, exists in four isoforms within *Arabidopsis*. *In silico* and yeast two-hybrid studies showed that PAP1 has unique expression and interaction pattern in *Arabidopsis*, suggesting non-canonical functions of PAP1. Its exclusive interaction with PAP4 has not been reported in other living systems until now and hints at a difference in polyadenylation in plants with respect to mammals and yeast. Cleavage Stimulation Factor (CstF), a heterotrimeric complex of the polyadenylation factors CstF50, CstF64 and CstF77, plays a role largely in cleavage of pre-mRNA. This study highlights some aspects of the *Arabidopsis* homologs of CstF64 and CstF77, central to various cellular processes other than nuclear polyadenylation. *In silico* studies showed an elevated expression of CstF64 in the pollen while that of CstF77 remained fairly low. Yeast two-hybrid assays indicated a novel kind of interaction of CstF64 with Fip1(V). It is also speculated from sub-cellular localization techniques by agroinfiltration in tobacco leaves that CstF64 localizes in the cytoplasm and CstF77 in the nucleus, as found for the orthologs of CstF77 in other systems.

Keywords: mRNA 3'-end processing, Poly(A) polymerase (PAP), Cleavage Stimulation Factor (CstF), Yeast two-hybrid, Agroinfiltration.

Amrita Bandyopadhyay

April 15, 2009

"Analysis of the *Arabidopsis* Polyadenylation Factors PAP1, CstF64 and CstF77 and their characteristic inter-relationship"

By

Amrita Bandyopadhyay

Arthur G. Hunt
Director of Thesis

Charles T. Dougherty
Director of Graduate Studies

June 1, 2009
Date

RULES FOR THE USE OF THESES

Unpublished dissertations submitted for the Master's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements

.Extensive copying or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this thesis for use by its patrons is expected to secure the signature of each user.

Name Date

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

THESIS

Amrita Bandyopadhyay

The Graduate School

University of Kentucky

2009

"Analysis of the *Arabidopsis* Polyadenylation Factors PAP1, CstF64 and CstF77 and their characteristic inter-relationship"

THESIS

A Thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
College of Agriculture
at the University of Kentucky

By

Amrita Bandyopadhyay

Lexington, Kentucky

Director: Dr. Arthur G Hunt, Professor of Plant and Soil Sciences

Lexington, Kentucky

2009

Acknowledgements

First and foremost, I am deeply indebted to my advisor, Professor Arthur G Hunt, for his guidance, enthusiasm and encouraging advice throughout this time. Thank you for being with me and my family during times of crisis. I gratefully acknowledge the funding provided by the National Science Foundation (NSF). I would also like to acknowledge the support shown by the Department of Plant and Soil Sciences, College of Agriculture, University of Kentucky during the course of my study.

I would like to thank my committee members Dr. Sharyn Perry, Dr. Ling Yuan and Dr. Martha Peterson for their advice and guidance during my graduate career. I am grateful to Dr. Joseph Chappell and the members of his lab for giving me the opportunity to use their facilities and also for advice and opinion in times of need. I would like to express my gratitude to Dr. Randy Dinkins and Dr. Sharyn Perry for their assistance in the seedhouse and the microscopy room. I would also like to convey my sincere thanks to Dr. Michael Goodin for sharing his expertise with the agrobacterium infiltration performed in Chapter 3.

I would like to express my gratitude to the present and past members of the Hunt Lab, Carol Von Lanken, Lavanya Dampanaboina, Patrick Thomas, Carrie Merill, Angela Jones, Drs. Quinn Li, Srinivasa Rao, Balasubrahmanyam Addepalli, Kevin Forbes and Lisa Meeks. I would also like to thank Amy Crume for maintaining and providing me with the *N. benthamiana* plants and soil material for my experiments in Chapter 3.

There are many friends whom I would like to acknowledge for making my stay in Lexington a memorable one. I am also greatly indebted to my parents, my sister and my mother-in-law for their support in the completion of this thesis. Without their love and

constant attention I would not be where I am today. Last but not the least, my husband, Mainak, has not only provided me with encouragement, constructive criticism and emotional support during this period, but also brought structure and enrichment to my life.

Table of Contents

List of Tables

List of Figures

Chapter 1: Literature review on the 3'-end processing in prokaryotic and eukaryotic systems

1.1 Introduction

Polyadenylation is the covalent linkage of a polyadenine tract to a messenger RNA (mRNA) molecule. It is part of the route to producing mature messenger RNA for translation, in the larger process of protein synthesis to produce proteins (Fig.1.1). It takes place in just about all eukaryotic organisms at the end of transcription, as a part of post-transcriptional modification, which is comprised of 5'-capping, RNA splicing and 3'-end polyadenylation. While in the nucleus, the product of transcription from RNA polymerase II, known as pre-mRNA, is associated with a variety of proteins in complexes known as heterogeneous ribonucleoprotein particles (hnRNPs) (Wahle and Ruegsegger, 1999). It is at this point that post-transcriptional modifications take place to give rise to mRNAs from pre-mRNAs. The only known exceptions to polyadenylation are mRNAs coding for replication-dependent histone proteins in metazoans, which undergo endonucleolytic cleavage, but unlike normal eukaryotic organisms have different set of factors for cleavage and subsequent absence of polyadenylate tail (Wahle and Ruegsegger, 1999; Davila Lopez and Samuelsson, 2008).

1.2 Functions of mRNA polyadenylation

Every biological process has some beneficial aspect, however small it might be it can have a huge impact on a much broader context. Likewise, polyadenylation plays a huge role at various phases of the mRNA function and metabolism. The primary function being enhancement of translation, since mRNA transcripts are used for the purpose of

translation into proteins. It has been shown that the major targets of poly(A) tails in mRNA translation in eukaryotes are the binding of 40S ribosomal subunit and subsequent joining with the 60S subunit in synergistic association with the 5'-cap (Sachs et al., 1997). Although not indispensable, it has been shown to affect translational efficiency through mutational studies in temperature sensitive strains in yeast cells (Proweller and Butler, 1994).

Besides translation, 3'-end processing also plays vital roles in transcription from the early steps of initiation to termination. On one end, transcription initiation is said to be closely monitored by the CPSF and CstF, initially through TFIID and then via CTD of pol II. On the other end, transcription termination is shown to be dependent on cleavage and polyadenylation. Termination of transcription at its final stage is important to minimize unnecessary polymerase activity and also avoid transcriptional interference at downstream promoters for closely spaced genes and chromosomal elements such as centromeres and origins of replication (Zhao et al., 1999). It has been shown that not only does the strength of the poly(A) site affect termination but also the presence and the effectiveness of downstream "pause sites" signal termination of pol II (Aranda et al., 1998; Nag et al., 2007; Glover-Cutter et al., 2008). Transcriptional run-on analysis on mutant yeast strains revealed that the CFIA and Yhh1p (mammalian CPSF 160) were indispensable for termination (Proudfoot, 2004).

Another key component of regulation of gene expression in all eukaryotic organisms that is also influenced by polyadenyation is mRNA stability. At various points, mRNA stability is dependent on the amount of a certain mRNA transcript in the cytoplasm available for translation or on the translational capacity of the mRNA within

the nucleus itself. In the former case it maintains the basal level of gene expression, by controlling turnover of mRNA transcripts from signals acquired from translational products. In the later case it detects and degrades aberrant mRNA transcripts either in the nucleus or the cytoplasm, thus controlling quality control of mRNA biogenesis (Tucker and Parker, 2000; Doma and Parker, 2007). Generally, the mRNA degradation machinery follows two pathways: in one the targeted mRNA undergoes shortening of poly(A) tail (deadenylation) followed by decapping, both catalyzed by a set of proteins whereby leaving the transcript exposed for 5' to 3' exonuclease activity. In the second case, following deadenylation the target mRNA undergoes 3' to 5' cytoplasmic exonuclease activity (Parker and Song, 2004). While these two are the default pathways of mRNA degradations, there are evidences of other ways of mRNA degradation broadly classified as deadenylation-independent decay (Beelman and Parker, 1995). Although there might be multiple pathways of mRNA turnover, there seems to be a competition between the rate of normal reaction in the life of an mRNA and quality control event acting over the mRNA, marking it for degradation. A key point in this kind of kinetic competition is that any defect causing delay in normal forward reaction will trigger quality control (Doma and Parker, 2007). Within this, the decay rate of individual transcripts are again influenced by the susceptibility of the multiple turnover pathways that act on it (Beelman and Parker, 1995).

The mRNA poly(A) tail also serves to facilitate nucleocytoplasmic transport of the mRNA transcripts, although there have been mixed reports regarding this aspect. It was shown by Huang *et al* that mRNAs lacking the polyadenine tract were inefficiently exported from the nucleus to the cytoplasm. In fact, addition of the 90nt long poly(A) tail

immediately upstream to the ribozyme cleavage site was not sufficient to restore the export function in absence of the actual 3'-end processing (Huang and Carmichael, 1996). This indicated that the downstream events delineated by the 3'-end post-transcriptional events were equally important for mRNA export. Conversely, there have been reports in some cases of nuclear transport of mRNA lacking poly(A) tails in yeast by Duval et al, where the cleavage was also performed by hammerhead ribozyme, thus lacking proper cleavage/polyadenylation processes (Duvel et al., 2002). In mutational studies it has been particularly shown that certain pre-mRNA processing factors in yeast like the CFIA proteins, PAPI and Pab1p when mutated affect cleavage, termination and export, thus resulting in failure of nucleocytoplasmic transport and accumulation of mRNA transcripts in the nucleus (Brodsky and Silver, 2000; Hammell et al., 2002). Another study showed that faulty mRNA 3'-end processing leads to defective transcription termination which is responsible for disengaging mRNA export factors (Lei and Silver, 2002), indicating that 3'-end processing and nucleocytoplasmic export are mechanistically linked processes.

From earlier studies and later on from data supporting exon definition, polyadenylation machinery could be synergistically associated to splicing mechanism mainly found in the eukaryotic system. Depending on the situation, splicing factors can not only enhance or inhibit cleavage/polyadenylation and vice versa, but can also govern the decision of how or when to polyadenylate an mRNA precursor (Cooke et al., 1999; Zhao et al., 1999; Millevoi, 2006; Danckwardt, 2007). This kind of cross-talk between splicing and polyadenylation is further exemplified in the genes where alternative polyadenylation leads to the expression of more than one gene product (Lutz, 2008). Two

such instances are the expression of the IgM heavy chain during B lymphocyte differentiation and the differential expression of the Calcitonin gene (Proudfoot et al., 2002). In case of the former, membrane-bound IgM is produced from the usage of a much stronger downstream poly(A) site where the weaker upstream one located in the intron is spliced off. After differentiation though, secretory-specific IgM is produced using the upstream poly(A) site, resulting in two different kinds of IgM heavy chain in the development and maturation of B lymphocyte cells (Minvielle-Sebastia and Keller, 1999). In the later case of Calcitonin gene, a downstream enhancer element between exon 4 and 5, activates an weak internal poly(A)site in thyroid tissues, giving rise to calcitonin. But in neuronal cells this enhancer fails to act and the internal poly(A) site is spliced out resulting in CGRP using the poly(A) site downstream (Colgan and Manley, 1997; Proudfoot et al., 2002).

Polyadenylation has also been implicated to regulate translational levels of mRNA in the cytoplasm of cells during early development of some vertebrates and invertebrates. Most of the mRNAs in oocytes of vertebrates like *Xenopus* and mouse are translationally dormant with shorter poly(A) tails. During maturation, these are activated as per use by increasing the length of the polyadenine tails to almost 150nt while existing mRNAs are de-activated or repressed by deadenylation. In case of invertebrates like *Drosophila* and *C. elegans*, regulation of the length of poly(A) tail in mRNAs is essential for correct embryonic patterning and sex determination. Although not much have been established with respect to cytoplasmic polyadenylation in adult tissues, reports show the possibility of it in the central nervous system due to the presence of CPEB, a factor that regulates cytoplasmic polyadenylation (Richter, 1999, 2007). Apart from CPEB, factors

that have been shown to participate in elongation of poly(A) tail during translational activation of oocytes and embryos are CPSF and PAP. In fact, results from studies on translational activation of *X. laevis* oocytes demonstrate that untimely addition of PAP can cause early activation in dormant mRNAs without the putative CPSF interaction domain, suggesting the role of CPSF and / or CBEB merely as trasporters of PAP to the targeted mRNA (Dickson et al., 2001).

1.3 3'- end mRNA processing signal sequences

Certain sequences, flanking the site of the endonucleolytic attack within the mRNA precursors, regulate the 3'-end processing efficiency under certain cellular condition. Although much has been established about mammalian polyadenyalation signals, research in recent years have led to much better understanding of the same in yeast and plants. Highlights of each group are discussed briefly.

Mammalian polyadenylation signals are basically composed of 3 major elements: the AAUAAA motif, the downstream elements (DSE) and the poly(A) site (Fig. 1.2A). The hexanucleotide AAUAAA sequence is highly conserved among higher eukaryotes and found 10 to 30 nucleotides upstream of the cleavage site (Zhao et al., 1999). It is indispensable for both cleavage and polyadenylation. The only other frequent variant to this sequence is AUUAAA and all other mutations essentially cripple the whole process. The second sequence, which is often a U-rich / GU-rich element, is located around 30 nucleotides downstream of the cleavage site, although it can be functional even further downstream (Wahle and Ruegsegger, 1999). Mutational studies prove it to be more diffuse and poorly conserved if not possibly redundant. However, it was observed that the distance between the two sequence elements is important not only in defining the

cleavage site but also the strength of the poly(A) signal (Colgan and Manley, 1997). The third important element in the core polyadenylation signal sequence is the poly(A) site itself. It is composed of a dinucleotide, adenine (A) followed preferably by a cytosine (C). Apart from these there have been reports of auxiliary sequence elements that can modify the efficiency of 3' end processing in a positive or negative manner (Zhao et al., 1999). Many viral and some cellular genes have such sequences as enhancers (USE) upstream of the AAUAAA motif. Unlike the USE, it is still not clear how the downstream auxiliary elements work due to the diverse nature of the downstream element.

Likewise mature mRNAs are also generated in yeast by cleavage and polyadenylation of pre-mRNA precursors, but unlike mammals the signal sequences are far too complex to define (Guo and Sherman, 1996). It also involves 3 core elements in the process but they are quite different from their mammalian counterparts (Fig. 1.2B). The first one is the efficiency element (EE), working to increase the efficiency of the 3'-end formation as the name suggests. It is positioned at a variable distance upstream from the positioning element, although optimally it is located at a distance of 10 to 20 nts (Wahle and Ruegsegger, 1999). Mainly composed of alternating UA-dinucleotides or U-rich stretches, it has been observed from various studies that a U-residue is critical at the first and the fifth position of the sequence. The following element is the positioning element (PE), which directs the positioning of the cleavage factors 20 nts. downstream of this sequence at the selected poly(A) site. It consists of mainly 2 variants of A-rich sequences: AAUAAA and AAAAAA, although other related sequences have also shown equal function, excepting when a G-residue was at the start (Guo and Sherman, 1996).

Both the efficiency elements and the positioning elements are not only degenerate but also redundant. Typically in yeast they are followed downstream by a third kind of sequence known as the poly(A) site which can be more than one, that is in a cluster. Coincident with the mammals, cleavage preferably occurs 3' to an adenosine residue, thus most of them have a $T/C(A)_n$ sequence as poly(A) site (Wahle and Keller, 1996). The major difference between the mammalian and the yeast signaling system is the absence of any downstream element from the yeast poly(A) site. This is maybe because of closely placed genes in yeast and the convergent nature of transcription, for which the polyadenylation signals can sometimes function in both orientation (Zhao et al., 1999).

Plants, introduced at a later stage in the study of 3' end processing, have shown a lot of commonality with respect to both higher eukaryotes and yeast, yet have their own distinction from both. The *cis*-acting elements, indispensable in the process of signaling, can be grouped in three classes: a far-upstream element (FUE), one or more near-upstream element (NUE) and the respective poly(A) site (CS) (Li and Hunt, 1997) (Fig. 1.2C). A FUE lies 13 to almost 150 nts. upstream of a NUE and is generally a UG-rich sequence, much like that of the DSE of the mammals. It is required for the efficient usage of the downstream poly(A) sites through interaction with proteins of the processing factors (Rothnie, 1996). FUEs of different plant poly(A) signals are interchangeable and a single FUE can have control over a number of downstream elements (Hunt, 1994; Rothnie, 1996). Mutational studies have indicated a high degree of functional redundancy; small deletions have hardly any effect on polyadenylation. The FUE is followed by one or more NUE, which is a 6 to 10 nucleotides sequence lying 10 to 40 nucleotides upstream to the associated poly(A) cleavage site (Li and Hunt, 1997). It

generally consists of AAUAAA or related sequences. Directed mutagenic studies revealed that although the NUE had apparent functional analogy to the mammalian AAUAAA-motif, yet they had much tolerance to point mutations, indicating the flexibility with regards to sequence as long as the location of the sequence was maintained with respect to other *cis*-acting elements. The assembly of the specific processing complex at the poly(A) site was driven by the unique secondary structure resulting from the mutual interactions of the *cis*-elements, last but not the least of which is the polyadenylation site (CS) itself (Rothnie HM, 1994). It is generally situated in a U-rich region of the 3'-UTR and has a consensus sequence of Y(C, A) dinucleotide at the cleavage site. Mutations within this sequence changes the position of the poly(A) site and also the efficiency in some cases suggesting the independence of the sequence as *cis*-element (Li and Hunt, 1997). Also in plant genes there can be multiple cleavage sites, where the usage of a particular site is defined by the distance between the NUE and the CS, hence a particular CS can be used with more than one NUE when they are in a specified position (Hunt, 1994). In recent years some other features have also been shown to play a role in polyadenylation like, sequence composition and secondary structure of the pre-mRNA of yeast and plant genes (Rothnie, 1996). In general polyadenylation in mammals, yeast and plants, to this day optimally consists of an A-rich sequence, a U-rich element and a T/C(A) cleavage site (Zhao et al., 1999) although dissection of each group leads to unique differences among them.

1.4 Factors involved in the 3'- end processing

It is evident from the functional aspect that polyadenylation is a ubiquitous process in all eukaryotes and even in some prokaryotes however different their overall

function might be. Generally, 3' end processing takes place in a stepwise but tightly coupled manner where cleavage at the poly(A) site is followed by subsequent addition and elongation of the polyadenine tract in a non-templated fashion (Fig. 1.1). The whole process is facilitated by a plethora of protein factors playing the basic two roles; yet their nuances in characteristics provide for the distinction in the process from organism to organism. This section individualy deals with the process of polyadenylation in a gross manner in the case of prokaryotes like *E.coli* and then eukaryotes like mammals, yeast and finally plants. Thus, this serves to bring forward to the reader the similarities and dissimilarities in the process of polyadenylation in different living organisms.

1.4.1 Polyadenyaltion in prokaryotes (bacteria): The presence of polyadenine tracts in case of some bacterial genes has definitely brought to light the importance of polyadenylation. On further investigation it was found to have functional and structural differences. Unlike eukaryotic mRNA transcripts, which tend to have long polyadenyalate tracts, bacterial mRNA transcripts have much shorter poly(A) tracts ranging from 14 to 60 adenine residues. Obviously, differences among the eukaryotes and the prokaryotes are the basis of such disparity in the structure and process of polyadenyaltion. For instance, translation of mRNA occurs co-transcriptionally in bacteria and the mRNAs are utilized very rapidly and efficiently, without any further modifications unlike eukaryotes. Since transcription occurs in a polycistronic fashion it obviates the need of long half-lives of the mRNAs, which is a must due to the spatial and temporal constraints of eukaryotic mRNAs (Sarkar, 1997). The lack of compartmentalization in prokaryotes also makes nucleocytoplasmic export of the transcribed mRNA unnecessary.

Analysis of mutant strains of *E.coli* lacking 3' exonucleases, revealed six different classes of mRNA polyadenylation in bacteria (Fig. 1.3). The monocistronic *lpp* transcripts correspond to class I and class II types (Cao and Sarkar, 1992; Sarkar, 1997). In the former case the poly(A) tract was attached to the end of the primary transcript defined by the rho-independent transcription terminator. While in the later one, the mRNA was truncated at the stem-loop structure and the poly(A) tract attached (Cao and Sarkar, 1992; Sarkar, 1997). In the *crp* locus encoding cyclic AMP receptor proteins, the polyadenylation site was found to be downstream of the translation termination site at the extreme end of the putative rho-dependent transcription terminator, giving rise to class III poly(A) mRNA (Sarkar, 1997). The much known *lacZYA* operon contains an intercistronic stem-loop structure resulting in termination in the *lacZY* region with a poly(A) tail just distal to the intercistronic stem-loop structure. This is referred to as the class IV poly(A) mRNA (Sarkar, 1997). The *rpsO* mRNA encoding ribosomal protein S15 the polyadenine tail was attached to a truncated coding region distinctive of class V poly(A) mRNA (Sarkar, 1997). In the last group of class VI poly(A) mRNA, the *rho* gene encoding the transcription terminator is controlled by 2 attenuation sites in the untranslated leader region. Polyadenylation can occur at either of the sites giving rise to a transcript with poly(A) tail attached even to the *rho* promoter. Thus, from all these different classes of mRNA polyadenylation in *E.coli* it can be deduced that process is relatively indiscriminate and occurs at any unprotected 3'-end of a mRNA molecule, irrespective of sequence and secondary structure (Sarkar, 1997).

In all the above cases, however different as it might be, the protein factor that catalyzes the template independent sequential addition of adenylate residues to the 3'

hydroxyl termini of mRNA transcripts is a poly(A) polymerase or a member of the poly(A) polymerase family. Continued polyadenylation to some extent even in deletion mutant strains proved that more than one PAP existed. The major poly(A) polymerase (PAP I) of *E.coli* was found to be encoded by a region of *pcnB* locus and the product a 52 kDA protein with no significant homology with eukaryotic PAP. Any level of overexpression was deleterious to cells and deletion or disruption mutants of *pcnB* gene reduced growth rates by less than 50%. Besides polyadenylation, it had the unique property of controlling ColE1 plasmid copy number (Xu and Cohen, 1995; Sarkar, 1997). Another kind of poly(A) polymerase (PAP II) was also identified to be a relatively hydrophobic protein weighing approximately 36-kDA. It is encoded by *f310*, having sets of two paired cysteine and histidine residues resembling the RNA binding motif. The product has no significant sequence homology whatsoever to either *E.coli* PAP I or to any viral or eukaryotic poly(A) polymerase (Cao et al., 1996; Sarkar, 1997). This indicated that bacterial polyA) polymerases have evolved independently with convergent evolution with respect to function. The significant functional overlap between these two polymerases was likely in part to defend the cell against loss of a vital function. While eukaryotic PAPs are closely related proteins arising from a single gene undergoing alternative splicing or post-transcriptional modification, prokaryotic PAPs appears to originate from a single gene (Sarkar, 1997). Apart from *E.coli* there are a number of bacterial genes whose products have sequence homology to *E.coli* PAP I.

Some eukaryotic organelles like mitochondria and chloroplasts are believed to have originated from endosymbiotic prokarytes. Thus, polyadenylation of mRNA in these organelles were also a subject of interest in this field. Mitochondrial pre-mRNA

transcripts undergo cleavage at the polyadenylation site followed by addition of the poly(A) tail. The sequence of events resembles the polyadenylation of the *E.coli lacZY* gene and hence mitochondrial mRNAs also look like class IV type bacterial mRNAs. Unlike eukaryotes, the major function of polyadenylation in mitochondrial genes is to complete the translation termination codon, in the absence of specific recognition sequence, indiscriminately at all mRNA ends (Sarkar, 1997; Nagaike et al., 2008). The average lengths of polyadenylate tracts in mitochondrial transcripts range from 35 to 55 nucleotides, very much like the prokaryotes. On the other hand in plant chloroplasts the poly(A) tract is several hundred nucleotides in length ranging somewhere between prokayotes and eukaryotes (Sarkar, 1997). In fact, unlike most poly(A) tract sequence composition, chloroplastidial poly(A) tails can often contain adenylate clusters interspersed with guanylate and even sometimes cytidylate and uridylate residues much like those in the bacteriophage T7 mRNA (Sarkar, 1997). Most of the mRNA in chloroplast have poly(A) tail attached to truncated coding regions, corresponding to class V of bacterial poly(A) mRNA. It has been observed that in mitochondria and chloroplasts mRNAs which are polyadenyalated are degraded at a much faster rate than those which are not (Slomovic et al., 2006). This indicated an mRNA turnover mechanism by polyadenylate tails in chloroplasts, analogous to bacterial cells (Li and Hunt, 1997; Sarkar, 1997; Dreyfus and Régnier, 2002).

The functions of mRNA polyadenyalation in prokaryotes can be much different from that in eukaryotes. The poly(A) polymerase of *E.coli* is shown to control plasmid copy numbers as stated before (Sarkar, 1997). Polyadenylation of RNA I targets it for degradation by PNPase, thus inactivating the inhibitor of plasmid replication(Xu and

Cohen, 1995; Sarkar, 1997). There has been some ambiguity with regards to RNA stabilization. It has been observed that when the 3' terminus has a stem loop structure, polyadenylation acts synergistically with 3' exonucleases to promote degradation. On the other hand in a different set of conditions, when the 3' terminus of the RNA is not stabilized with secondary structures, polyadenylation competes with the same exonucleases as before to promote mRNA stability (Sarkar, 1997). Also the binding of the S1 protein to poly(A) tails during the mRNA recruitment to the 30S ribosome in *E.coli*, suggests a possible role of polyadenylation in the stimulation of translation initiation by S1 protein (Sarkar, 1997).

1.4.2 Polyadenylation in eukaryotes (mammals): Polyadenylation in mammals has been studied intensively for years. Hence, nowadays, studies on polyadenylation in other organisms are based on the basic knowledge of polyadenylation in mammals. In mammals there are many protein factors that can cause either the endonucleolytic cleavage or polyadenylation, but sometimes are required for both (Fig. 1.4). The cleavage and polyadenylation specificity factor (CPSF), cleavage-stimulation factor (CstF), cleavage factor I_m and II_m (CFI_m and $CFII_m$), poly(A) polymerase (PAP) are involved in cleavage, while poly(A)-binding protein II (PabII) along with PAP, CPSF are needed for the polyadenylation step. From earlier studies it was shown that the AAUAAA sequence was crucial in both the stages of polyadenylation. In the initial stages of cleavage this *cis*-element was recognized by a plethora of *trans*-acting factors mainly by CPSF. The purified CPSF consists of four subunits weighing approximately about 160 KDa, 100 KDa, 73 KDa and 30 KDa out of which the 160 subunit is of primary activity. It was not only responsible for the sequence recognition but also cooperative interaction with other

cleavage and polyadenylation subunits to stabilize the whole assembly (Manley, 1995; Colgan and Manley, 1997; Zhao et al., 1999). The CPSF 100 and 73 subunits have shown close relatedness to each other, while the smallest subunit CPSF 30, although dispensible for the process, have shown quite interesting characteristics as the potential endonuclease for cleavage. But the best characterized one among these is CPSF 160. While CSPF recognizes and binds to AAUAAA sequence, another set of proteins, the CstF independently recognizes the downstream GU-rich sequences and binds to it through CstF 64. In fact the interaction of CPSF 160 with the upstream signal sequence is weak and tolerant to mutations (Colgan and Manley, 1997). The presence of other CPSF and the CstF subunits make it specific and thus strong. The CstF protein consists of CstF 50, CstF 64 and CstF 77. Out of these CstF 64 has the RNA recognition motif. CstF 77 makes all the connections among the other two CstF subunits and interacts with CPSF 160, thus stabilizing the initial cleavage complex through co-operative binding. Much of the properties of CstF will be discussed in some detail in chapter 3. CF Im and CF IIm are also factors that are important in establishing the stability of the cleavage complex through protein-protein interactions with other factors in the complex. In fact it has been thought that CF I_m prepares the pre-mRNAs for proper recognition by CstF complex. It consists of 72 kDa, 68 kDa, 59 kDa and 25 kDa subunits, of which 68 and 59 ones are closely related. CF II_m is an additional factor working in tandem with PAP in the formation of the cleavage-competent complex. Since it has yet to be purified and analyzed much remains unknown about its function. PAP plays a key role in both cleavage and polyadenylation. Cleavage efficiency is also affected by the presence of PAP in the cleavage complex either acting in cleavage or as a stimulatory factor. After

cleavage, addition of the adenylate residues at the 3' mRNA ends is catalyzed by PAP. There have been reports of various alternatively spliced variants of PAP, the largest and the catalytically active ones being 82 kDa and 77 kDa. Although PAP is very crucial in the cleavage and polyadenylation reaction, the protein has very little and unspecific affinity for RNA. Hence, the interaction with CPSF 160 tethers the protein to the mRNA substrates, without which it can add adenine residues to any random mRNA primer. Some of the crucial aspects of PAP in 3' end processing will be discussed in chapter 2. Although the presence of PAP with CPSF is sufficient in the synthesis of the poly (A) tail, the process is not only slow but also distributive. After about 10 A-residues in this manner, the addition of 33 kDa PAB II molecule changes the rate of the process by not only making it faster but also making it processive. PAB II has a high affinity for poly (A) and binds to the short tail to form a stable quaternary complex with CPSF, PAP and the RNA substrate (Colgan and Manley, 1997). But this property drastically changes again to distributive manner once the tract is 200 to 250 nts in length. The same PAB II molecule is said to measure and control the length of the growing poly(A) tract.

1.4.3 Polyadenylation in eukaryotes (yeast): As it has been already mentioned, there exists a major difference among the mammalian and yeast signal sequences, but the factors catalyzing the polyadenylation process has more or less similar functional aspects. Like mammalian polyadenylation, in yeasts too the process is divided in two steps- endonucleolytic cleavage and polyadenylation. Cleavage is carried out by cleavage factor IA (CF IA), cleavage factor IB (CF IB) and cleavage factor II (CF) and polyadenylation is catalyzed by polyadenylation factor I (PF I), poly(A)-binding protein 1 (Pab 1), poly(A) polymerase 1 (Pap 1) along with CF IA and CF IB (Fig. 1.5). The primary

components of the initiation complex in polyadenylation are CF IA and CF II since the factors have a high affinity for the signal sequences. The components of CF IA are Rna 14, Rna 15, Pcf 11, Pab1 and a 50 kDa polypeptide. Rna 14 is a 76 kDa polypeptide having 24% sequence homology to mammalian CstF 77. On the other hand Rna 15 is 38 kDa polypeptide having 43% homology in the RBD to mammalian CstF 64. Both the subunits are tightly bound to each other like their mammalian counterparts, but unlike the CstF units they are required in both cleavage and polyadenylation in yeast. Hence, they behave more like CPSF than CstF. Mutants in yeast are defective in both cleavage and polyadenylation. Pcf 11 is a 70 kDa polypeptide interacting with both Rna 14 and Rna 15. Extracts from mutants of Pcf 11 are similarly defective in cleavage and polyadenylation. Until now no mammalian homolog has been uncovered, although functionally it shares similarity to mammalian CPSF 160 (Shatkin and Manley, 2000). Yeast Pab1 is the major RNP in yeast which remains associated with the poly (A) tails in the cytoplasm and mediates mRNA translation and turnover. It is a 70 kDa polypeptide having nearest mammalian homology to PAB II (Zhao et al., 1999). It acts along with poly(A) specific nuclease (PAN) and nuclear poly(A)-binding protein (Nab2p) (Viphakone et al., 2008) to regulate the length of poly(A) tails, which is generally 50-90 nts long. Lastly, a 50 kDa subunit which is said to be a product of the Clp 1 gene, but not much has been known about it in yeast. The components of CF II are Cft1/Yhh1 (150 kDa), Cft2/Ydh1 (105 kDa), Brr5/Ysh1 (100 kDa) and Pta1 (90 kDa). The first subunit to be purified and identified was Yhh1 due to its sequence homology (24% identity and 51% similarity) to mammalian CPSF 160. Inactivation led to loss of both cleavage and polyadenylation, whereas reversal led to partial activation with revival of only cleavage

but not poly(A) addition (Zhao et al., 1999). Ydh1 had 24% identity and 43% similarity to Mammalian CPSF 100 and bind to pre-mRNA substrates either at the efficiency elements or at the poly(A) sequences in an ATP-dependent manner (Wahle and Ruegsegger, 1999; Zhao et al., 1999). The third subunit, Ysh1 is 23% identical and 48% similar to mammalian CPSF 73. A cold-sensitive mutant was shown to be defective in the in vivo splicing of the mRNA. The 90 kDa Pta1 polypeptide, shares limited similarity to mammalian protein symplekin, which recently have been implicated in to be a part of the CPSF complex. Pta1 is an essential gene playing a role in pre-tRNA processing and like symplekin helps in the assembly or stabilization of the polyadenylation complex (Zhao et al., 1999; Shatkin and Manley, 2000). The next important subunit for cleavage and polyadenylation is CF IB which is represented by the single 73 kDa polypeptide Hrp1/Nab4 (Wahle and Ruegsegger, 1999; Zhao et al., 1999). It is a shuttling protein and unlike mammalian export proteins, have possible roles in polyadenylation and subsequent nucleocytoplasmic transport. Although it is not essential for cleavage, yet the high affinity towards UA-rich polyadenylation sites in yeast plays a qualitative role by regulating the choice of cleavage site utilization (Zhao et al., 1999; Shatkin and Manley, 2000). Polyadenylation Factor I (PF I) has been purified from a co-purification of CF II-PF I and consists of Fip1, Pap1, Yth1, all the subunits of CF II, Pfs1 and Pfs2 (Wahle and Ruegsegger, 1999; Zhao et al., 1999). Factor Interacting with Pap1 (Fip1) interacts with Pap1 in yeast-two hybrid interactions and has a molecular weight of 35 kDa. Although for a long time there was no mammalian counterpart identified, but in 2004 sequence similarity showed that a human Fip 1 existed as an integral part of CPSF and acts in concert with CPSF 160 in RNA recognition (Shatkin and Manley, 2000; Kaufmann I et

al., 2004). Apart from interacting with Pap1 it also interacts with Yth1 and weakly with Rna14 (Zhao et al., 1999). Thus functionally it is similar to mammalian CPSF 160 subunit in linking the cleavage holoenzyme to the polyadenylation complex. Mutants are generally defective in polyadenylation but not in cleavage. The interacting protein, Yth1, similar to CPSF 30 in the mammals, is also shown to be a part of this complex. Any mutation within a zinc finger of the Yth1 reduces cleavage activity (Barabino et al., 1997; Zhao et al., 1999). Finally among Pfs1 and Pfs2 not much is known, except that Pfs1 has zinc knuckle while Pfs2 has seven WD-40 repeats and an N-terminal extension similar to mammalian CstF 50 and interacts with Rna14. The most important polyadenylation protein is Pap. In yeast it was the first factor to be elucidated due to its simplicity as a single 64 kDa polypeptide, 47% identical to its mammalian counterpart within its 400 amino acid residues. Similar to the mammalian PAP factor, it too, does not have any sequence specificity for the RNA substrates. As a part of the PF I holoenzyme, interaction with Fip 1 directs it towards the catalytic core (Wahle and Ruegsegger, 1999; Addepalli et al., 2004), but unlike the mammalian protein plays a role only in polyadenylation.

1.4.4 Polyadenylation in eukaryotes (plants): In comparison to the other mentioned organisms, study of the process of polyadenylation was quite recent but a rapidly growing area in plants. As mentioned earlier, plants shared similarities to mammals and also to yeast with respect to the *cis*- acting elements. Not only this, like mammals, plants also showed a great amount of alternative polyadenylation in its genome. For example, in rice more than 50% of the genes out of the 55,000 genes analyzed had more than one unique poly(A) sites (Shen et al., 2008). Hence it was an

obvious direction to delve deeper into the process and hypothesize a model for the polyadenylation machinery along with the *trans*- acting factors. The earliest component in this effort to be isolated, purified and elucidated was Poly (A) polymerase (PAP). Although there were differences in size and structure due to alternatively modified forms of this enzyme, all the PAP activities had not only similar biochemical properties among plants but also among mammals and yeast system. The primary role of PAP in mRNA production at all stages of development and translation regulation remains more or less conserved in all living organisms (Rothnie, 1996). The broader aspects of PAP in higher plants will be discussed in chapter 2. With further screening of the databanks and different libraries homologies among various mammalian / yeast and plant polyadenylation factors was also found. To date almost 28 subunits have been found to be expressed in *Arabidopsis* which are homologous to mammalian / yeast systems, excepting CFIm68 in mammals and HrpI in yeast. All of the CPSF and CstF protein factors are expressed from single genes while, PAP, Fip I, Clp I, Pcf I I, Pab and symplekin are expressed from multiple genes (Hunt et al., 2008). Further investigation with each gene product will elucidate the characteristics and properties of each, thus establishing a plant polyadenylation model.

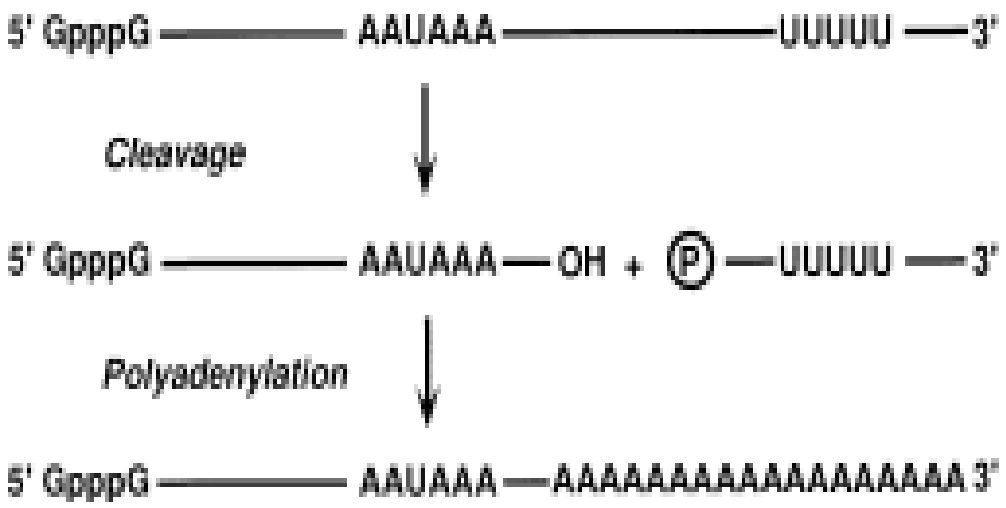


Figure 1.1: The two steps of 3' end processing (Wahle and Ruegsegger, 1999)

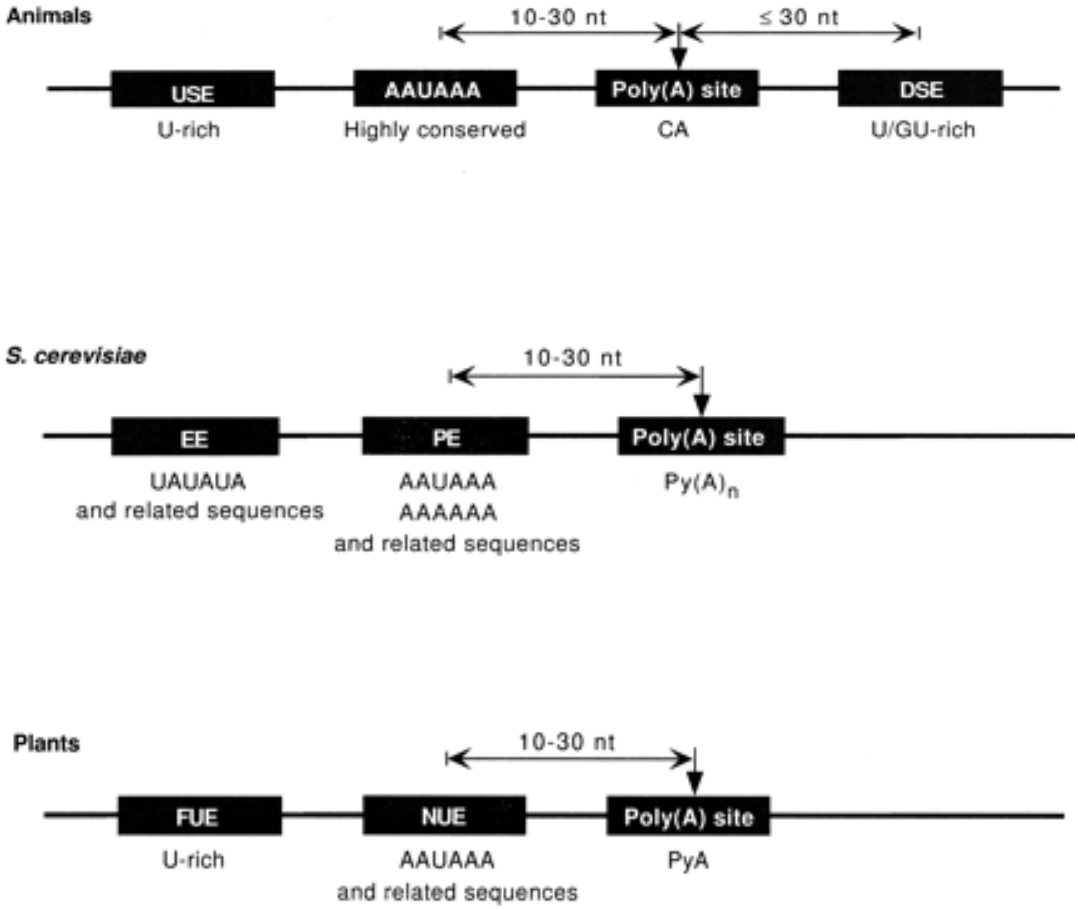


Figure 1.2: Schematic diagram of polyadenylation signals in 3 different living systems. (A) mammals (B) yeast (*S.cerevisiae*) (C) plants (Zhao et al., 1999)

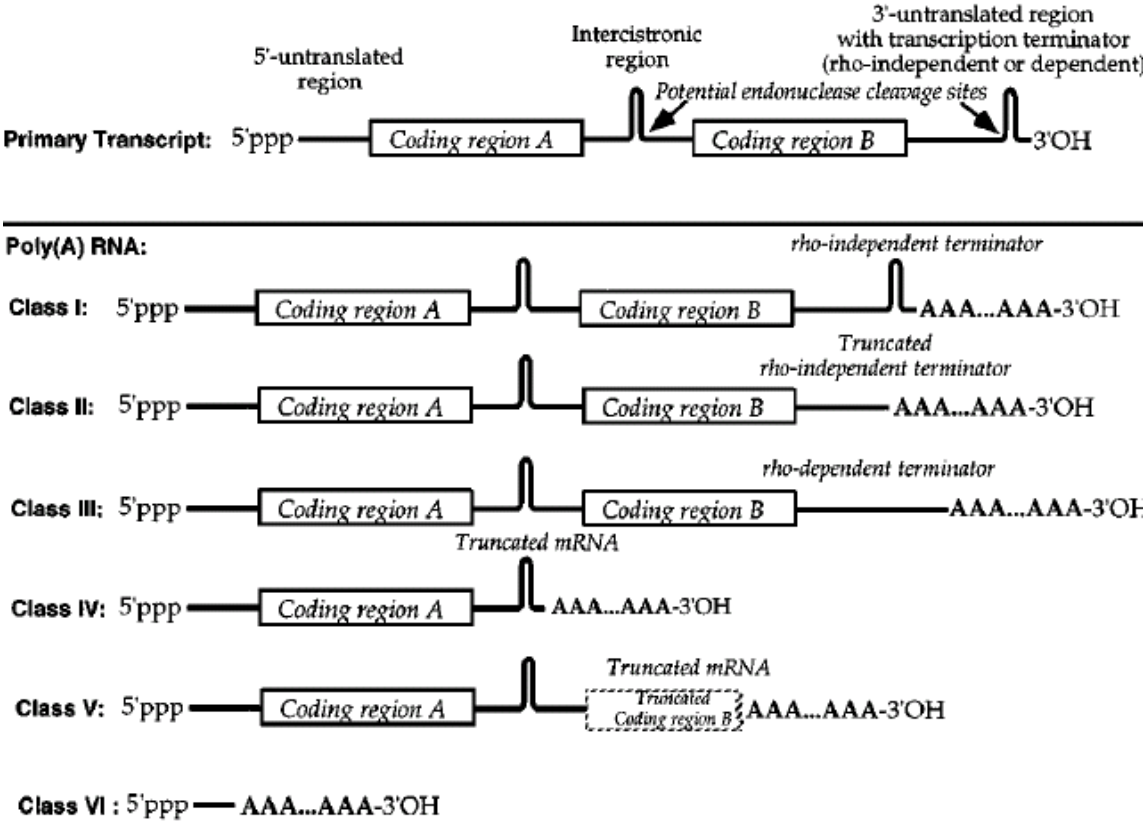


Figure 1.3: Six classes of mRNA polyadenylation in bacteria (Sarkar, 1997).

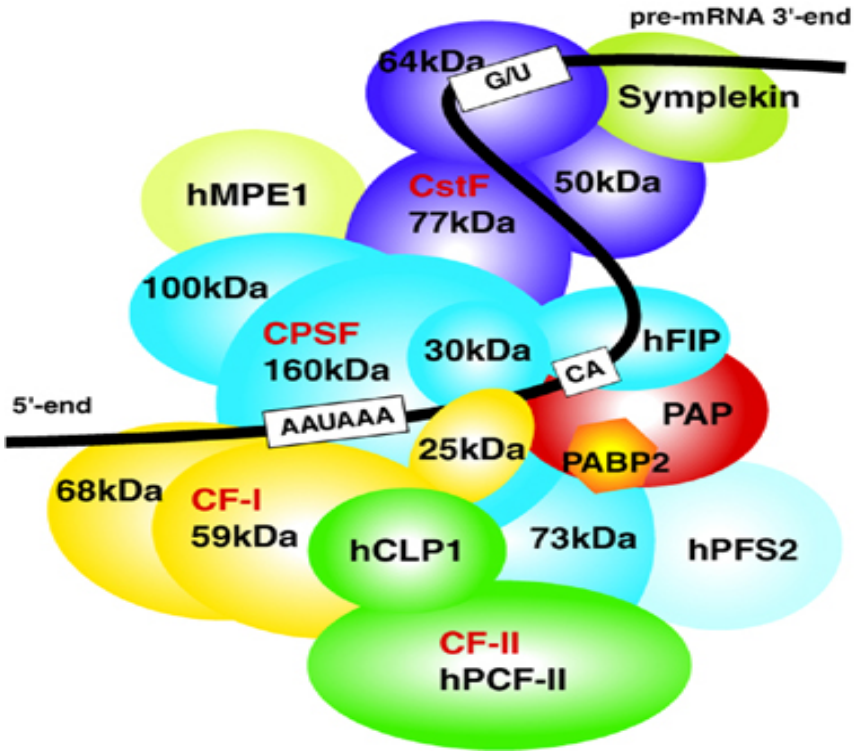


Figure 1.4: Model for the polyadenylation machinery in mammals (Adapted from the Walter Keller Lab website)

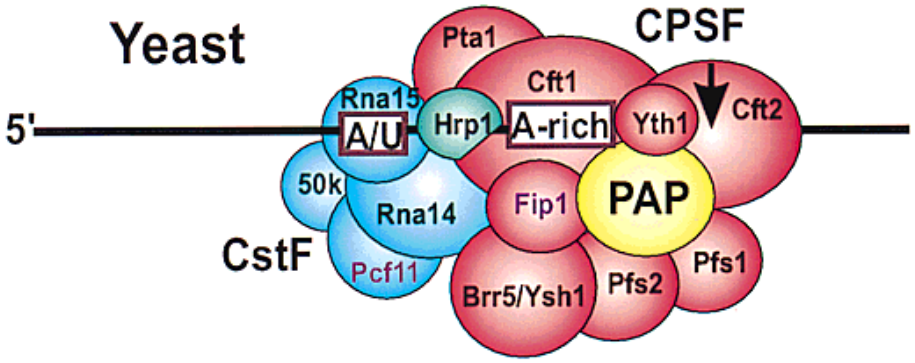


Figure 1.5: Model for the polyadenylation machinery in yeast (Shatkin and Manley, 2000)

Chapter 2: Interactions of PAP1 with the polyadenylation factors in *Arabidopsis*

2.1 Introduction

The process of polyadenylation involves two stages as mentioned in Chapter 1: an endonucleolytic cleavage of the pre-mRNA transcript at the signal sequence subsequently followed by the addition of a polymerized tail of adenine residues. The whole process engages a battery of proteins at various stages and one of the most important one being polynucleotide adenylyltransferase (EC 2.7.7.19). Polynucleotide adenylyltransferase, or poly(A) polymerase (PAP) as the name suggests, is an indispensable protein factor responsible for adding the adenine residues, thus resulting in the polymerized adenine tail at the 3' end of most mRNAs. This factor has been mentioned in some contexts in Chapter 1, but a detailed discussion will be presented in this chapter.

Poly(A) polymerase has been studied in great detail in living organisms and one of the first factors to be discovered in the study of polyadenylation. The work in poly(A) polymerase characterization was pioneered as early as 48 years ago by Mary Edmonds (Edmonds and Abrams, 1960; Edmonds et al., 1976). It was initially observed that the enzyme polymerized AMP residues in a non-specific manner, which can then be added to any RNA primer (Edmonds, 1982; Raabe et al., 1991). Later on the discovery of polyadenine tails in mammalian mRNA precursors and their varied metabolic functions fueled the interest in this protein factor. With subsequent research it was found that the non-specific PAP behaved in a specific manner with mRNA precursors having a precleaved AAUAAA sequence at its 3' ends, while in the presence of other

polyadenylation specific factors (Christofori and Keller, 1988; Christofori and Keller, 1989; Raabe et al., 1991; Wahle, 1991).

In 1991 the primary structure of bovine PAP was elucidated by Raabe et al (Raabe et al., 1991). The general structure features 3 domains: the N-terminal catalytic domain, the central domain and the C-terminal RNA binding domain (Fig. 2.1). PAP is a member of the nucleotidyltransferase superfamily. It houses 3 aspartate residues in its catalytic domain, characteristic to the members of this superfamily. This is typically the binding sites for ATP and metal ions (Martin, 2000). It also possesses an RNA binding site (RBS), downstream of the catalytic domain. The N-terminal domain and the central domain are more or less conserved among mammals, yeast and plants (Fig. 2.3). PAP is unique among nucleotidyltransferases in that it possesses an RNA recognition and binding domain, distinct from the active site, near its C-terminal end (Martin, 2000). It is this domain which mediates the primer recognition in mRNA precursors undergoing non-specific 3'-end processing.

Apart from RNA recognition and binding through interactions in a AAUAAA-specific manner, PAP engages in a host of protein-protein interactions. In mammals it has been shown that CPSF, PabN and CFI-25 interacts with PAP and aids in the assembly of the processing unit of mRNA polyadenylation (Keller and Minvielle-Sebastia, 1997; Kim and Lee, 2001). Not only this, but the very C-terminal regulatory domain also links polyadenylation to pre-mRNA splicing events via interactions with splicing factors like U1A and U2A (Fig. 2.1) (Gunderson et al., 1997; Forbes, 2005; Meeks, 2005). The C-terminal region also possesses 2 nuclear localization signals (NLS-1 and NLS-2) which efficiently localize PAP to the nucleus (Raabe et al., 1991; Raabe et al., 1994). Other than

nuclear localization, the C-terminal especially the NLS interacts with CPSF via its 160kDa subunit (Thuresson et al., 1994; Manley, 1995; Murthy and Manley, 1995). Most of the C-termini of PAPs are not evolutionarily conserved, but are generally very rich in serine and threonine residues (Fig. 2.3). The S/T-rich region lying in the C-terminal domain has multiple consensus and non-consensus sites for phosphorylation/ dephosphorylation (Raabe et al., 1994; Ballantyne et al., 1995; Colgan DF et al., 1998). Both specific and non-specific poly(A) polymerase activity is inhibited during phosphorylation, generally carried out by $p34^{cdc2}$/cyclin B complex, also known as mitosis promoting factor (MPF) (Colgan et al., 1996). This activity was hypothesized to regulate the formation of polyadenylation complex at the initiation stage (Raabe et al., 1994). Thus, downregulation of gene expression by controlling the enzymatic activity of PAP, takes place during the G_2/M-phase in somatic cells and also during meiotic maturation of oocytes (Colgan et al., 1996). Apart from regulation of PAP via splicing and phosphorylation, hormones have also played a part in regulating PAP expression in mammals if necessary (Jacob et al., 1975; Orava et al., 1979; Raju and Reddy, 1983; Xu et al., 1983).

The study of cytoplasmic polyadenylation in *Xenopus* (Gebauer and Richter, 1995) provided some interesting insights into the functional variation of PAP. Cytoplasmic polyadenylation requires a cytoplasmic form of PAP. Most (about 70%) of this protein is homologous to mammalian PAP but it lacks the NLS at the C-terminal domain. The expression of the cytoplasmic PAP was developmentally regulated, thus the concentration of the mRNA declined as the oocytes matured. Regulation of PAP level is crucial for controlled cytoplasmic polyadenylation and cell viability (Juge et al., 2002).

Although the canonical function of cytoplasmic PAP is similar to the nuclear one, association with CPEB, a factor in the cytoplasm and / or other cytoplasmic factors closely resembling CPSF, hint at a possibly different mechanism of polyadenylation in the cytoplasm (Hake and Richter, 1994; Gebauer and Richter, 1995). For instance, unlike nuclear polyadenylation, association of cytoplasmic PAP with these additional cytoplasmic factors is necessary even for polyadenylation of mRNA already 25-75 nucleotides long. Even though PAP alone is sufficient for polyadenylation and stimulation of translation, its premature recruitment to resting oocytes may cause anomalies in maturation (Dickson et al., 2001). At this point the CPSF-like cytoplasmic polyadenylation factors act as signal receptors and transporters of PAP to the mRNA, for efficient polyadenylation and oocyte maturation.

Functionally active forms of PAP can vary not only according to the location of action but also due to alternative splicing, competition and choice of poly (A) sites on the PAP gene. PAP I and PAP II, products of alternative splicing of PAPOLA gene (Raabe et al., 1991; Wahle, 1991; Zhao and Manley, 1996), are the major poly(A) polymerases. PAP II (~83 kDa) is the most common isozyme found in the vertebrates. The other significant one, PAP I (~77 kDa), is less commonly found with comparison to PAP II. Both of the PAPs perform similar basic function of adenine residue polymerization, but differ only by their C-terminal end sequences (Raabe et al., 1991; Raabe et al., 1994). It was hypothesized that phosphorylation/dephosphorylation activities at the C-terminal domain sequences of these 2 PAPs gave rise to difference in interaction pattern for PAP and other processing factors, facilitating the switch between different phases of 3'-end processing (Thuresson et al., 1994).

Poly(A) polymerase in yeast (PAP1p) is functionally similar to its mammalian counterpart (Fig. 2.1). It is a 64 kDa polypeptide and is a product of a single copy gene, *PAP I*. Structurally it is 47% identical to mammalian PAP within the first 400 amino acids (Fig. 2.3). This conserved region typically contains the catalytic domain with an amino terminus RNA Binding Site (N-RBS) and a carboxy terminus RBS (C-RBS) (Zhelkovsky et al., 1995; Zhao et al., 1999). Two specificity domains (SpD 1 and SpD 2) delineate both ends and confer protein-protein interactions with various factors of the polyadenylation machinery, like CF I and Fip1 (Kessler, 1995; Keller and Minvielle-Sebastia, 1997). Both of these domains are unique to yeast Pap1 and are required for enzyme activity (Zhelkovsky et al., 1995). Another distinction of Pap1p is that it is only essential during the polyadenylation phase but not during cleavage (Zhao et al., 1999). Although yeast Pap1 lacks the S/T-rich regions, it still undergoes regulation via phosphorylation and ubiquitination during cell cycle as control of enzymatic activity (Lingner and Kellerman, 1991; Mizrahi and Moore, 2000). Phosphorylation is mediated by a protein kinase different from Cdc28 homologue, during the S/G_2-phase unlike in mammals, where it happens during the G_2/M-phase (Colgan et al., 1996; Mizrahi and Moore, 2000). Phosphorylation generally precedes ubiquitination as a mode of yPAP1 regulation without involving proteolysis. Yeast Pap1 interacts, possibly via its N-terminal domain, with Uba2 and Ufd1 proteins that have been linked to the ubiquitin mediated protein degradation pathway (Dohmen et al., 1995; Johnson et al., 1995; del Olmo et al., 1997). Thus, Pap1p too undergoes post-translational modification like its mammalian counterpart but in ways that differ temporally and mechanistically.

Investigations in the recent past have lent much insight into the plant polyadenylation process. This includes the identification and characterization of the factors involved in this process, which bear an outstanding likeness to other eukaryotic polyadenylation factors. Plant Poly(A) polymerase is one of such factors which have been identified and characterized from a number of plant sources like wheat, pea, *Arabidopsis*, maize, and tobacco (Mans and Huff, 1975; Berry and Sachar, 1982; D'Alessandro and Srivastava, 1985; Dasgupta et al., 1995; Addepalli et al., 2004). In *Arabidopsis*, it is encoded by a family of 4 genes, with some predicted isoforms being very similar to their mammalian counterpart (Fig. 2.2). Although the plant PAPs so far isolated, have varying molecular weights ranging from 60-120 kDa, 3 PAPs out of 4 in Arabidopsis are ~83-95 kDa (Rothnie, 1996; Hunt et al., 2000). PAPS1, PAPS2 and PAPS4 (Fig. 2.2) (named according to their respective position in the chromosomes) are among the larger ones with ~800 amino acids, whereas PAPS3 (Fig. 2.2) has only 482 amino acids and is significantly shorter. Excepting PAPS3, all the 3 PAPs have N-terminal domain showing high degree of conservation with the mammalian PAP (Fig. 2.3), while their C-terminal domain differs even among themselves (Addepalli et al., 2004). Even though there is much sequence dissimilarity within the C-terminal domain, presence of S/T-rich regions within this domain suggests that phosphorylation/dephosphorylation phenomenon may exist in plants, much like yeast and mammalian PAP (Verma and Sachar, 1994; Hunt et al., 2000; Addepalli et al., 2004). Theoretically speaking, the conformational change brought about by the protein kinase (PK)/protein phosphatase (PP) acting upon the target protein, is a classic way of regulating enzymatic action (Sopory and Munshi, 1998). Plant PAPs also undergo similar

post-translational modification by their intrinsic PKs, although the spatio-temporal mechanism of such regulation in plants still needs to be explored (Sharma et al., 2002). Another distinctive feature of plant PAP is in its hormonal regulation during embryogenesis (Berry and Sachar, 1981; Berry and Sachar, 1982; Lakhani and Sachar, 1985; Rothnie, 1996). The four genes are expressed in a tissue specific manner (Addepalli et al., 2004). For example, PAPS1 is predominantly expressed in the roots, stem and flowers but not in the leaves where the predominant one is PAPS3 (Addepalli et al., 2004; Meeks, 2005). Like mammalian PAP genes, alternative splicing also has a high occurrence within *Arabidopsis* PAP genes (Zhao and Manley, 1996; Addepalli et al., 2004). But unlike mammals, the alternatively spliced products are highly truncated, and may retain some of the functionalities of the full-length catalytically-active enzyme (Addepalli et al., 2004; Meeks, 2005). Although much of the functional need for such novel mRNAs remains unknown, it has been suggested that the relative abundance of the mRNA and the PAP enzyme are controlled via stabilization of the alternatively spliced forms both at nuclear and cytoplasmic levels (Rothnie, 1996; Zhao and Manley, 1996; Addepalli et al., 2004).

It is quite noticeable from the facts outlined in this chapter and from the previous chapter that the plant 3'-processing machinery has similarities as well as differences with both the mammalian and yeast counterparts. Previous and recent studies have shown that PAP has non-specific distributive activity which is even true for plants. But in presence of other protein factors especially those implicated in the polyadenylation process, it becomes specific and progressive (Raabe et al., 1991). These protein interactions may have some similarities or some unique differences among the different organisms. For

instance *Arabidopsis* PAP1 (Fig. 2.2) does not interact with Fip1, unlike mammalian or yeast PAP or for that matter the other isoforms of AtPAP, which have been shown to interact though yeast two-hybrid assays (Forbes et al., 2006). Thus, the difference in the polyadenylation process from mammals to plants is the manifestation of these characteristic interactions among the protein factors like PAP and other 3'-processing proteins.

It has been demonstrated earlier that PAP has a propensity to exist as multiple isoforms mostly as products of alternative splicing (Zhao and Manley, 1996; Sharma et al., 2002; Addepalli et al., 2004). The significance of this hasn't been fully unearthed, but one hypothesis suggests that PAP has functions beyond polyadenylation within and outside of the nucleus. For example, in *Arabidopsis* itself, the PAPS3 isoform (Fig. 2.2) has intriguing spatio-temporal expression levels (Addepalli et al., 2004). It dominates in PAP expression more within the foliar extremities, while the lack of NLS at the C-terminal domain clues at possible functions in cytoplasmic polyadenylation. PAPS2 protein (Fig. 2.2), on the other hand, has the canonical PAP activity within the nucleus, but has been also shown to associate with PNP, a chloroplastidial enzyme involved in RNA metabolism, thus suggesting at a non-canonical role of PAP (Hunt et al., 2000). Hence, it is tempting to hypothesize that PAPS1 may also have functions beyond polyadenylation, since structurally it shows an intron inclusion event in its alternatively spliced form. It has almost imperceptible or no expression within the leaves, but shows varying levels of presence in stems, roots and flowers of *Arabidopsis* (Addepalli et al., 2004) This chapter delineates some efforts in exploring and understanding the characteristics of PAPS1 in plants by conducting a systematic analysis of the microarray

data obtained from NASC and protein-protein interactions among PAPS1 and other *Arabidopsis* polyadenylation factors. Through such basic studies, it will be possible to analyze the process of polyadenylation and fill the gaps in the web of interaction network involving *Arabidopsis* PAPS1.

2.2 Experimental procedures

2.2.1 In silico expression analysis for PAPS1, PAPS2, PAPS3 and PAPS4

The *in silico* gene expression analysis was performed using the data available in the Additional file 1: microarray keys and data from Hunt *et al* (Hunt et al., 2008), which was compiled from the NASC (Nottingham *Arabidopsis* Stock Centre) microarray database (Craigon et al., 2004). The datasets for *PAPS1, PAPS2, PAPS3 and PAPS4* were obtained from the 4 different experimental conditions (developmental stages, abiotic stress conditions, chemical and hormonal responses and biotic and differential light responses) (Appendix: A1). The data was plotted in a XY-scatter and also in a bar graph and analyzed.

2.2.2 Interaction assay for PAPS1 and other Arabidopsis polyadenylation factors in a GAL4 based two-hybrid system

A refined and elegant yeast-two hybrid assay was performed for the interactions between *Arabidopsis* PAPS1 (At1g17980) and the different *Arabidopsis* polyadenylation factors enlisted in Table 1. A Gal4-reporter based two-hybrid system with the yeast strain PJ69-4A (MATa *trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*) was used. The expression vectors were pGAD-C(1) for the activation domain and pGBD-C(1) for the binding domain (James et al., 1996; Forbes et al., 2006). The entire protein coding region or a part of it in question for

the respective protein factors, were cloned firstly into pGEM-T™ vector system (Promega) and then excised with *BglII*. These fragments were then cloned in the pGAD-C(1) and pGBD-C(1) vectors and then sequenced to confirm the correct reading frames of the gene fusions thus created (Delaney et al., 2006; Hunt et al., 2008). Competent yeast cells (PJ69-4A) were then transformed with the desired plasmid DNA using the polyethylene glycol and lithium acetate method (Gietz et al., 1992; Forbes et al., 2006; Xu et al., 2006). The transformants were plated in a synthetic complete medium containing glucose as the carbon source and lacking the nutritional supplements leucine (L) and tryptophan (W) (SC-LW) (Forbes et al., 2006; Xu et al., 2006). After a period of incubation at 30°C for approximately 4 days, the dual transformants from the colonies in SC-LW medium were plated in the selective synthetic growth medium lacking histidine (H) leucine (L) and tryptophan (W) along with the controls and incubated for 2-3 days approximately at 30°C. The analog 3-amino-1,2,4-triazole (3-AT) was used in the case of the histidine-lacking media (SC-HLW) plates (Forbes et al., 2006; Xu et al., 2006; Hunt et al., 2008). One of the empty plasmid DNA ("AD" and "BD") as well as one of the test plasmid DNA, were co-transformed with a complementary empty plasmid DNA and used as negative control, other than the empty AD and BD vectors. The co-transformation using *Arabidopsis* orthologue of CstF64 and CstF77 was used as positive control (Yao et al., 2002; Forbes et al., 2006). Positive interactions were those where 50-100% colonies from the dual transformants grew on the SC-HLW plated in comparison to the positive controls or else less than 10% were scored as negative.

2.3 Results

2.3.1 Expression characteristics of Arabidopsis PAPS genes from the in silico expression analysis

It has been observed from earlier RT/PCR and Northern blot studies that that isoforms of *Arabidopsis* PAP shows difference in tissue-specific expression levels and also in the mode of alternative splicing of the mRNA within the tissues (Addepalli et al., 2004; Meeks, 2005). Thus, PAPS3 was the dominant form in the leaves, PAPS1, PAPS2 and PAPS4 were expressed in varying degrees in roots and PAPS1, PAPS2 and PAPS3 were the major species of mRNA in the stems of *Arabidopsis* (Addepalli et al., 2004). Hence, the expression levels of PAPS1 especially with respect to other PAP isoforms in *Arabiodopsis*, was a matter of interest here and investigated though *in silico* studies. Expression data sets were obtained for 4 kinds of experimental variations in *Arabidopsis*; developmental stage, response to abiotic stress, response to chemical stress and response to biotic stress and differential light conditions. The data were obtained from a previous compilation and plotted as described in the following (Appendix: A1). The results from the analysis give some novel facts about the expression of PAPS1 gene in *Arabidopsis* with respect to the other 3 PAP genes. For instance, Fig. 2.3 (A) shows that most of the PAP expression is normal excepting around the pollen development stage where PAPS3 is dominantly expressed and PAPS1 is rather repressed. This is highlighted more in Fig. 2.3 (B) which shows the ~60 fold increase in PAP3 expression and rather ~4 fold decrease in PAP1 expression, with normalized data. Fig.2.4 shows that wounding in seedling can cause a increase in PAP1 expression compared to other PAP isoforms in *Arabidopsis*. Cyclohexamide treatment to seedlings (Fig.2.5) can also cause a huge

change in expression, almost a 6 fold increase. In the same experimental condition (Fig.2.5) imbibition in seeds, can also cause a gradual ~3 fold increase in PAPS1 with respect to other PAPs. The effect of different elicitors (Fig.2.6) on the expression of PAPS1 is also notable. With each 4h treatment the expression jumps by almost 2-3 folds from normal expression at 1h duration.

2.3.2 PAPS1 shows interactions only with PAPS4

Interactions among the polyadenylation factors are quite prevalent in other organisms, as it has been mentioned previously. Whether similar kind of interactions are also part of the plant polyadenylation complex is a matter of question. The ability of *Arabidopsis* PAPS1 homologue to interact with other *Arabidopsis* 3'-processing factors was tested using a yeast-two hybrid assay. In most cases the entire protein-coding region was fused to create an activation domain (AD) or a binding domain (BD) (James et al., 1996; Hunt et al., 2008). In some rare cases (eg. AtFip1(V)-NTD containing the first 137 amino acids and AtFip1(V)-CTD containing the last 500 amino acids) fused AD and BD plasmid DNA were made from partial protein-coding regions (Forbes et al., 2006). Negative controls were made from one of the "empty" AD or BD vectors and the fused gene product. The highly reproducible interaction between AtCstF64 and AtCstF77 was considered as a positive control in these tests (Yao et al., 2002). Co-transformation into yeast cells (PJ69-4A) gave rise to dual transformant colonies on the SC-LW plates, which were again plated on SC-HLW selection medium. Interactions for a pair in both combinations (eg.ADPAPS1+BDCstF64 and AD CstF64+BD PAPS1) were tested. In the case of proteins with an inherent activation domain (e.g. AtCPSF30), both the combinations were tested but only one set was scorable.

From previous studies in yeast and mammals, it has been observed that PAP interacts primarily with Fip1 (Preker et al., 1995; Kaufmann I et al., 2004). In yeast it has also been observed to have interactions with the CF I subunit (Kessler, 1995). On the other hand, in mammals it interacts positively with CPSF, CFI-25 and PabN (Thuresson et al., 1994; Murthy and Manley, 1995; Kim and Lee, 2001; Kerwitz et al., 2003). Even in plants like *Arabidopsis* interactions have been detected though a number of two-hybrid assays between PAP isoforms and CPSF100, CPSF30, Fip1(V), PabN and CFIS (Elliott et al., 2003; Forbes et al., 2006; Hunt et al., 2008). But before now, no interactions have been reported within the different PAP isoforms in any animal or plant system. Hence, it was very surprising to find that out of the following 26 interactions (Table 2.1) tested PAPS1 had positive results only with PAPS4 (Fig. 2.7). Although the results here were not affirmed by a different test, it might not be too optimistic to add that the results of the protein-protein yeast-two hybrid assays performed here are authentic.

2.4 Discussion

The role of poly(A) polymerase in 3'-end processing is pivotal and it is functionally conserved among a vast range of living organism. This kind of conservation leads to functional redundancy of the PAP protein resulting into multiple isoforms with a difference in their amino acid composition, especially at the C-terminal domain. Its functional importance and ubiquity can be testified in a number of living organisms (Niessing and Sekeris, 1974; Ryner et al., 1989; Zhao and Manley, 1996; Sharma et al., 2002). A similar picture is also present within *Arabidopsis* which also has 4 isoforms of the PAP (Addepalli et al., 2004). With respect to this, the tissue specific expression of each of the isoforms (PAPS1, PAPS2, PAPS3 and PAPS4) is remarkable. Thus, PAPS1

which has been largely expressed in stems, roots and flowers but not in leaves has some interesting connotation. *In silico* microarray studies show that its expression is comparably low in pollen (0.3535) compared with PAPS3 expression (35.435) (Fig. 2.3). This might suggest at a different mechanism of polyadenylation in pollen where PAPS3 is essential and PAPS1 is hence, dispensable, such as during spermatogenesis in mice where CstF64 is replaced functionally by τCstF64 (Wallace et al., 1999). Hence, whatever function is played by PAPS1 in *Arabidopsis* is performed by PAPS3 in pollen or the process is non-active or absent from the pollen tissue, rendering PAPS1 unnecessary. On the other hand yeast two-hybrid test results show that PAPS1 interacts positively with PAPS4 only out of all polyadenylation protein factors (Table 2.1). In other living organisms, poly(A) polymerase interacts with factors like Fip1, CPSF160, CFI-25_m, and PabN (Thuresson et al., 1994; Murthy and Manley, 1995; Preker et al., 1995; Kim and Lee, 2001; Kerwitz et al., 2003; Kaufmann I et al., 2004) and in *Arabidopsis* other PAP isoforms interact mostly with FipS5, PabN, CPSF30, CFIS and CPSF100 (Elliott et al., 2003; Forbes et al., 2006; Hunt et al., 2008). But to date, there have been no reports on self-interaction or interactions within the PAP isoforms themselves, even in plants. Hence, an obvious question is why PAPS1 shows this deviation from other PAPs and interacts only with PAPS4? Maybe, it can be hypothesized that PAPS1 is expressed only in the presence of PAPS4 along with other polyadenylation factors, thus controlling the expression of the whole PAPS1-PAPS4 subunit as a whole. But with this hypothesis a number of questions arise. Like, how is the property of PAPS4 altered by rendering PAPS1 non-functional and vice versa? Does this have any effect on the overall process of polyadenylation? The interactions of PAPS2,

PAPS3 and PAPS4 with FipS5, reminiscent of the yeast system, facilitate the recruitment of the PAPs to the rest of the polyadenylation complex. Hence, altering one of the connections in this intricate network of interactions might have some effect on the overall process. This adverse effect has been demonstrated by insertional mutation and RNA interference studies on all 4 isoforms of PAP (Meeks, 2005). This study shows that, although gene duplication might have caused redundancy in the enzymatic properties of PAPs, the functional aspects of the genes have changed with evolution and now their expression is necessary for plant viability. Hence, each of the PAP isoforms have unique tissue or developmental stage specific expression pattern (this study; Addepalli et al, 2004), hinting at distinct sub-functions performed by each of the gene products, in its local realm. The various functions of these isoforms may overlap with each other or they may work in concert with other protein factor(s) to perform a greater task. This might be the case for PAPS1-PAPS4 interaction, although the physiological significance of this interaction has yet to be ascertained.

Most of these PAP isoforms differ biochemically from each other especially at their C-terminal end which undergoes phosphorylation, other than post-transcriptional modification of the whole gene product by alternative splicing. Alternative splicing may give rise to an altogether different gene product, whereas phosphorylation/dephosphorylation may cause structural/conformational changes in the protein. This might finally influence the interactions of PAP with the RNA substrate, ATP, polyadenylation factors or other factors within and out of the nucleus. Therefore, this might be how the isoforms are not only differentially expressed, but their specific functions are also regulated at different steps of the cell cycle according to their available

substrate. It has been observed that the differential phosphorylation activity at the C-terminal end causes the difference in substrate specificity of PAP in SRP RNA and hence, the difference in the polyadenylation mechanism (Perumal et al., 2001). On the other hand, PAP can also associate with factors outside the canonical 3'-end processing unit for polyadenylation of non-coding mRNAs (Vaňáčová et al., 2005). PAP has also been implicated to play a role in the metabolism of "cryptic" unstable mRNA transcripts (Wyers et al., 2005; Lykke-Andersen and Jensen, 2006). A variant of PAP in animals, known as star-PAP interacts with PIPK I, involved in mRNA splicing, export and other cell signaling processes (Mellman et al., 2008). Even in plants, a chloroplastidial form of PAP is shown to be involved in RNA degradation pathway, opening up the idea of PAP activity in cell processes other than 3'-end processing (Burkard and Keller, 1974; Dasgupta et al., 1995; Dasgupta et al., 1998). Hence, it is imperative to find out the binding partners for PAPS1, so as to acquire an idea about the possible functions of PAPS1 not only in polyadenylation of mRNA but also other nuclear and extra-nuclear activities.

2.5 Conclusion

The *in silico* expression analysis exhibits an intriguing expression profile for PAPS1 in *Arabidopsis*. While it is under expressed in pollen, it shows moderately higher levels of expression during different abiotic, chemical and biotic stresses. The functional significance of such an expression pattern is yet to be determined. The yeast two-hybrid interaction assay between PAPS1 and other *Arabidopsis* polyadenylation protein factors clearly depict that only PAPS4 interacts positively with PAPS1. These results give us an opportunity to explain and understand the possible role played by PAPS1 in various

cellular processes in *Arabidopsis*, but also opens up an avenue of questions for further analysis and characterization of PAPS1. This might lay the basic foundation towards understanding an important aspect of polyadenyaltion and uncovering the mechanism of other cellular functions not only in *Arabidopsis*, but also the plant kingdom, in greater context.

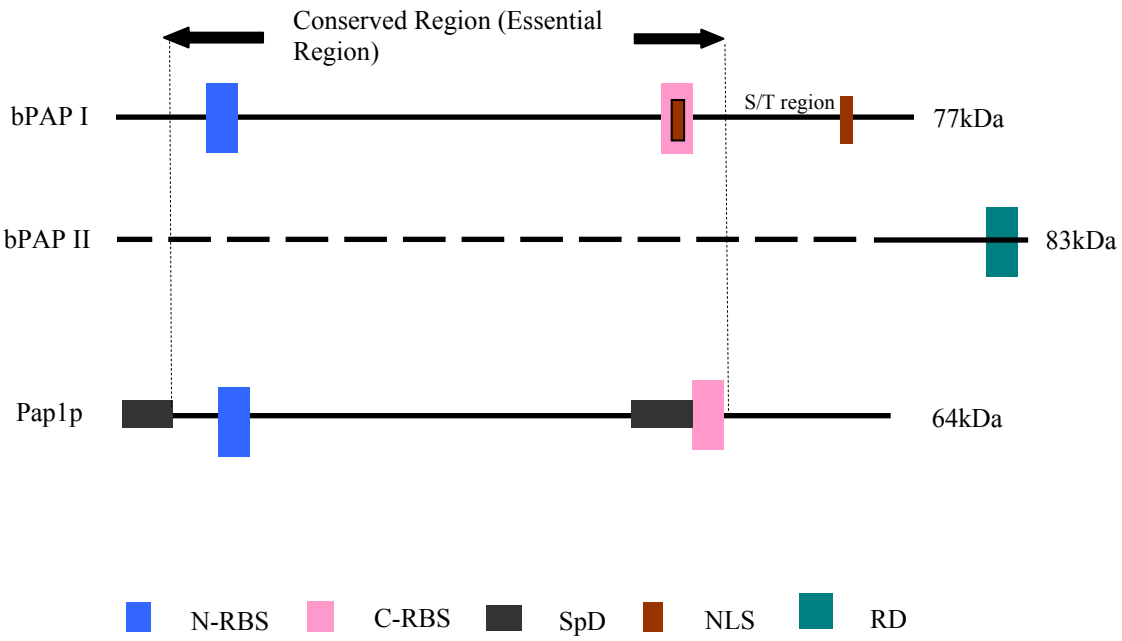


Figure 2.1: Schematic representation of poly(A) polymerases from calf (bPAP I and II) and yeast (Pap1p). The numbers at the right denote approximate molecular weights of the polypeptides. N-RBS, amino terminus RNA binding site; C-RBS, carboxy terminus RNA binding site; SpD, specificity domain; NLS, nuclear localization site; RD, regulatory domain for splicing via U1A/polyadenylation activities. Adapted from (Raabe et al., 1991; Raabe et al., 1994; Zhelkovsky et al., 1995).

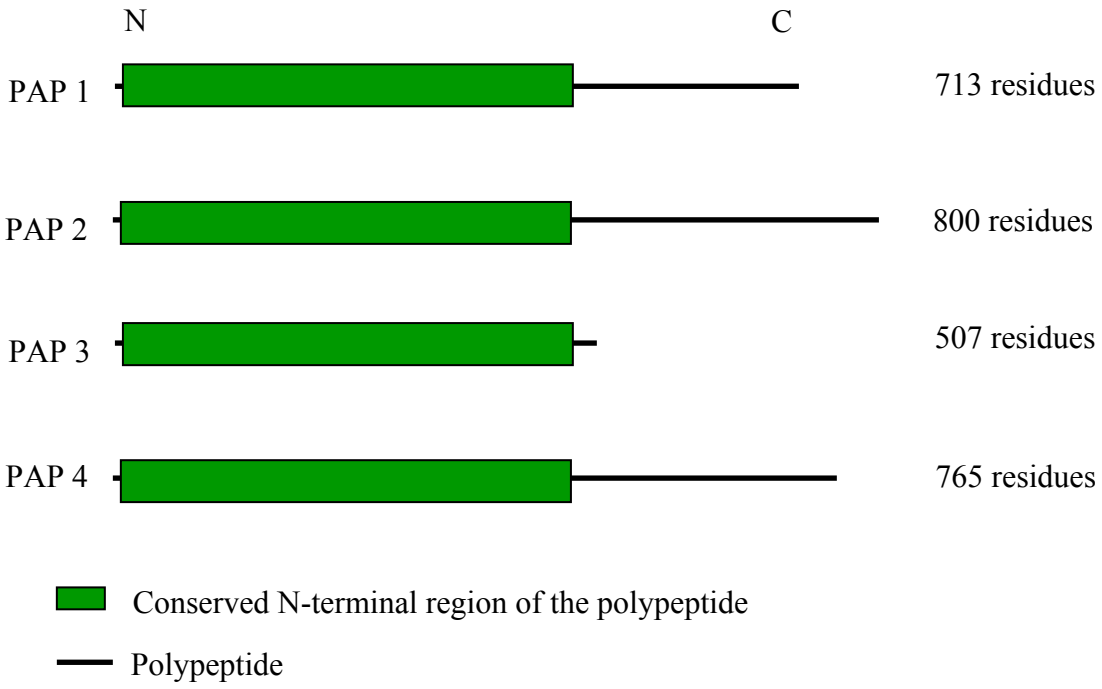
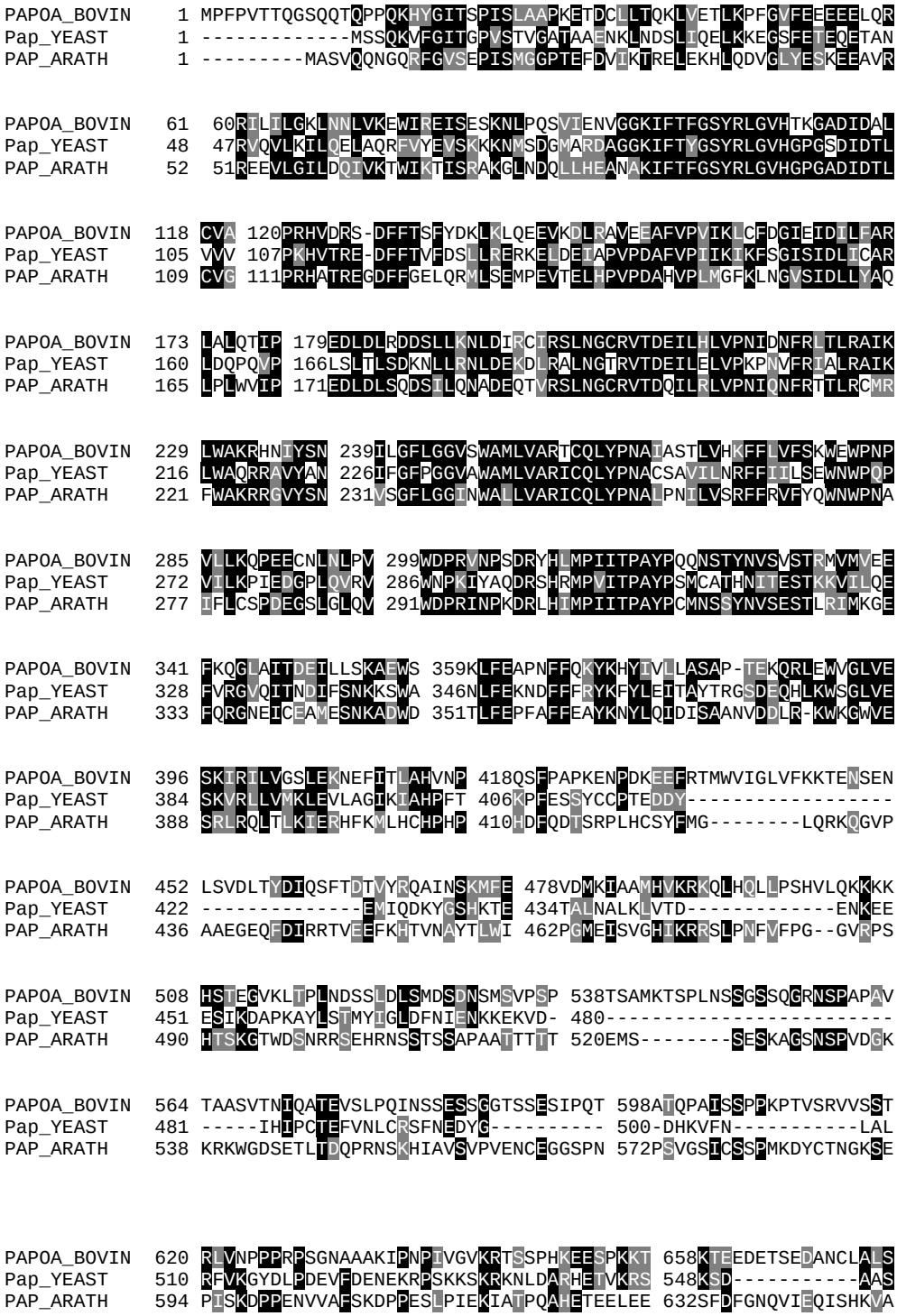


Figure 2.2: Schematic representation of the four poly (A) polymerases from *Arabidopsis thaliana*. The numbers at the right denote amino acid length of the polypeptides. Adapted from Addepalli *et al*, 2004.


PAPOA_BOVIN 1 MPFPVTTQGSQQTQPPQKHYGITSPISLAAPKETDCLLTQKLVETLKPFGVFEEEEELQR
Pap_YEAST 1 ----------MSSQKVFGITGPVSTVGATAAENKLNDSLIQELKKEGSFETEQETAN
PAP_ARATH 1 ---------MASVQQNGQRFGVSEPISMGGPTEFDVIKTRELEKHLQDVGLYESKEEAVR
PAPOA_BOVIN 61 60RILILGKLNNLVKEWIREISESKNLPQSVIENVGGKIFTFGSYRLGVHTKGADIDAL
Pap_YEAST 48 47RVQVLKILQELAQRFVYEVSKKKNMSDGMARDAGGKIFTYGSYRLGVHGPGSDIDTL
PAP_ARATH 52 51REEVLGILDQIVKTWIKTISRAKGLNDQLLHEANAKIFTFGSYRLGVHGPGADIDTL
PAPOA_BOVIN 118 CVA 120PRHVDRS-DFFTSFYDKLKLQEEVKDLRAVEEAFVPVIKLCFDGIEIDILFAR
Pap_YEAST 105 VVV 107PKHVTRE-DFFTVFDSLLRERKELDEIAPVPDAFVPIIKIKFSGISIDLICAR
PAP_ARATH 109 CVG 111PRHATREGDFFGELQRMLSEMPEVTELHPVPDAHVPLMGFKLNGVSIDLLYAQ
PAPOA_BOVIN 173 LALQTIP 179EDLDLRDDSLLKNLDIRCIRSLNGCRVTDEILHLVPNIDNFRLTLRAIK
Pap_YEAST 160 LDQPQVP 166LSLTLSDKNLLRNLDEKDLRALNGTRVTDEILELVPKPNVFRIALRAIK
PAP_ARATH 165 LPLWVIP 171EDLDLSQDSILQNADEQTVRSLNGCRVTDQILRLVPNIQNFRTTLRCMR
PAPOA_BOVIN 229 LWAKRHNIYSN 239ILGFLGGVSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWEWPNP
Pap_YEAST 216 LWAQRRAVYAN 226IFGFPGGVAWAMLVARICQLYPNACSAVILNRFFIILSEWNWPQP
PAP_ARATH 221 FWAKRRGVYSN 231VSGFLGGINWALLVARICQLYPNALPNILVSRFFRVFYQWNWPNA
PAPOA_BOVIN 285 VLLKQPEECNLNLPV 299WDPRVNPSDRYHLMPIITPAYPQQNSTYNVSVSTRMVMVEE
Pap_YEAST 272 VILKPIEDGPLQVRV 286WNPKIYAQDRSHRMPVITPAYPSMCATHNITESTKKVILQE
PAP_ARATH 277 IFLCSPDEGSLGLQV 291WDPRINPKDRLHIMPIITPAYPCMNSSYNVSESTLRIMKGE
PAPOA_BOVIN 341 FKQGLAITDEILLSKAEWS 359KLFEAPNFFQKYKHYIVLLASAP-TEKQRLEWVGLVE
Pap_YEAST 328 FVRGVQITNDIFSNKKSWA 346NLFEKNDFFFRYKFYLEITAYTRGSDEQHLKWSGLVE
PAP_ARATH 333 FQRGNEICEAMESNKADWD 351TLFEPFAFFEAYKNYLQIDISAANVDDLR-KWKGWVE
PAPOA_BOVIN 396 SKIRILVGSLEKNEFITLAHVNP 418QSFPAPKENPDKEEFRTMWVIGLVFKKTENSEN
Pap_YEAST 384 SKVRLLVMKLEVLAGIKIAHPFT 406KPFESSYCCPTEDDY----------
PAP_ARATH 388 SRLRQLTLKIERHFKMLHCHPHP 410HDFQDISRPLHCSYFMG--------LQRKQGVP
PAPOA_BOVIN 452 LSVDLTYDIQSFTDTVYRQAINSKMFE 478VDMKIAAMHVKRKQLHQLLPSHVLQKKKK
Pap_YEAST 422 --------------EMIQDKYGSHKTE 434TALNALKLVTD------------ENKEE
PAP_ARATH 436 AAEGEQFDIRRTVEEFKHTVNAYTLWI 462PGMEISVGHIKRRSLPNFVFPG--GVRPS
PAPOA_BOVIN 508 HSTEGVKLTPLNDSSLDLSMDSDNSMSVPSP 538TSAMKTSPLNSSGSSQGRNSPAPAV
Pap_YEAST 451 ESIKDAPKAYLSTMYIGLDFNIENKKEKVD- 480--------------------------
PAP_ARATH 490 HTSKGTWDSNRRSEHRNSSTSSAPAATTTTT 520EMS--------SESKAGSNSPVDGK
PAPOA_BOVIN 564 TAASVTNTQATEVSLPQINSSESSGGTSSESIPQT 598ATQPALSSPPKPTVSRVVSST
Pap_YEAST 481 -----IHIPCTEFVNLCRSFNEDYG---------- 500-DHKVFN----------LAL
PAP_ARATH 538 KRKWGDSETLTDQPRNSKHIAVSVPVENCEGGSPN 572PSVGSICSSPMKDYCTNGKSE
PAPOA_BOVIN 620 RLVNPPPRPSGNAAAKIPNPIVGVKRTSSPHKEESPKKT 658KTEEDETSEDANCLALS
Pap_YEAST 510 RFVKGYDLPDEVFDENEKRPSKKSKRKNLDARHETVKRS 548KSD----------AAS
PAP_ARATH 594 PISKDPPENVVAFSKDPPESLPIEKIATPQAHETEELEE 632SFDFGNQVIEQISHKVA

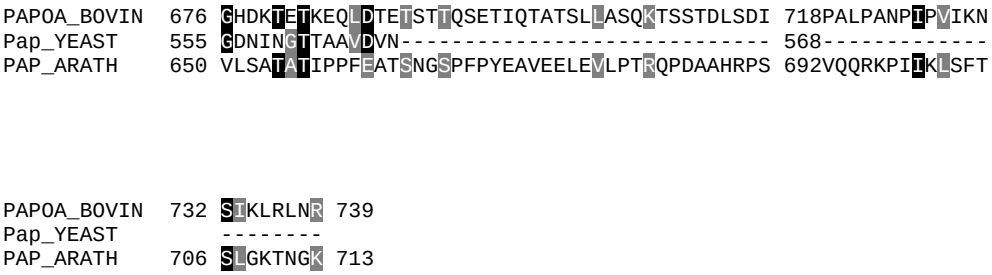
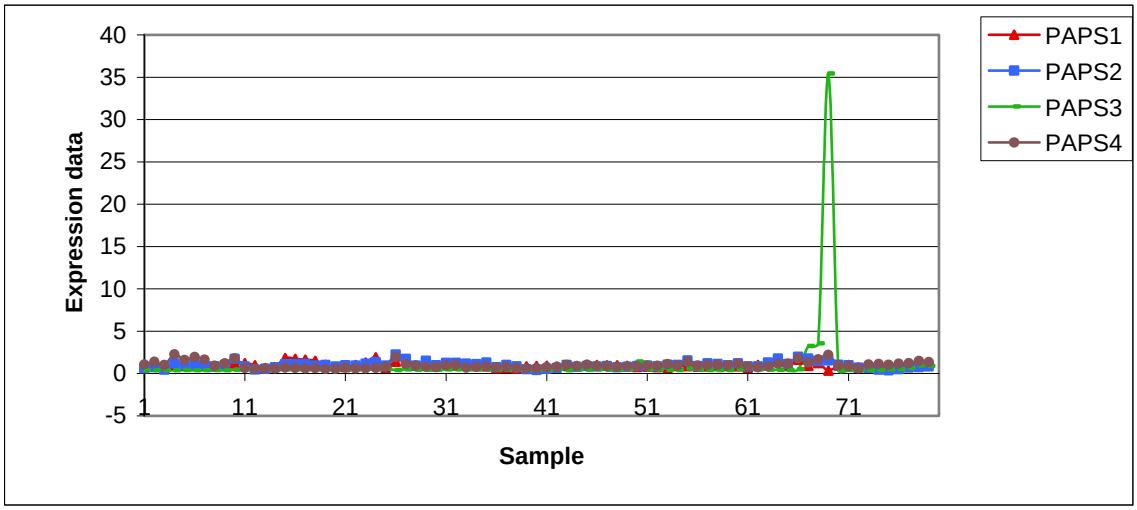


Figure 2.3: Amino acid sequence alignment and comparison of poly (A) polymerases. The open reading frames encoded by calf (PAPOA_BOVIN, UniProt P25500), yeast (Pap_YEAST, UniProt P29468) and Arabidopsis (PAP_ARATH, UniProt Q9LMT2) are compared using ClustalW and formatted using Boxshade. Residues which are identical in all three polypeptides are shaded in black, with white uppercase lettering. Positions that are similar are shaded in grey, with white uppercase lettering.

(A)



(B)

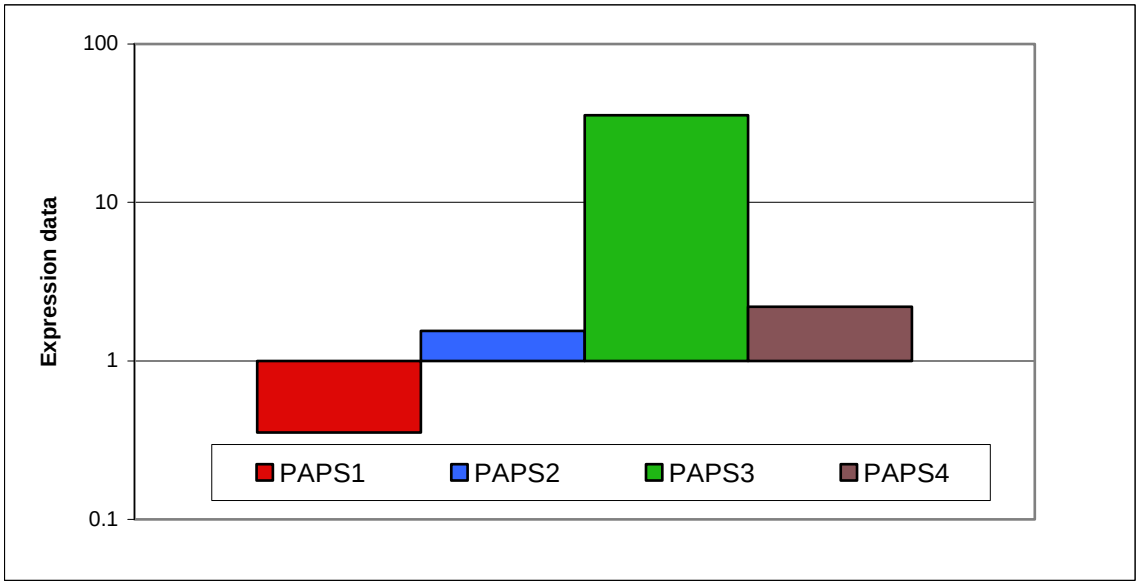


Figure 2.4: Expression data analysis for *Arabidopsis* PAP subunits during developmental stage. The data obtained from NASC was plotted as shown. The respective PAP subunits are color-coded in the legend. The expression values are represented along the Y-axes. The samples (including some WT and mutants) that are represented along the X-axes in (A): 1-7, root 7-21 days; 8-10, stem 7-21 days; 11-27, leaf 7-35 days; 28-38, whole plant 7-23 days; 39-49, shoot apex 7-21 days; 50-71, flowers and floral organs 21+ days; 72-79, 8 week seeds and siliques. The normalized expression in mature pollen is depicted in (B).

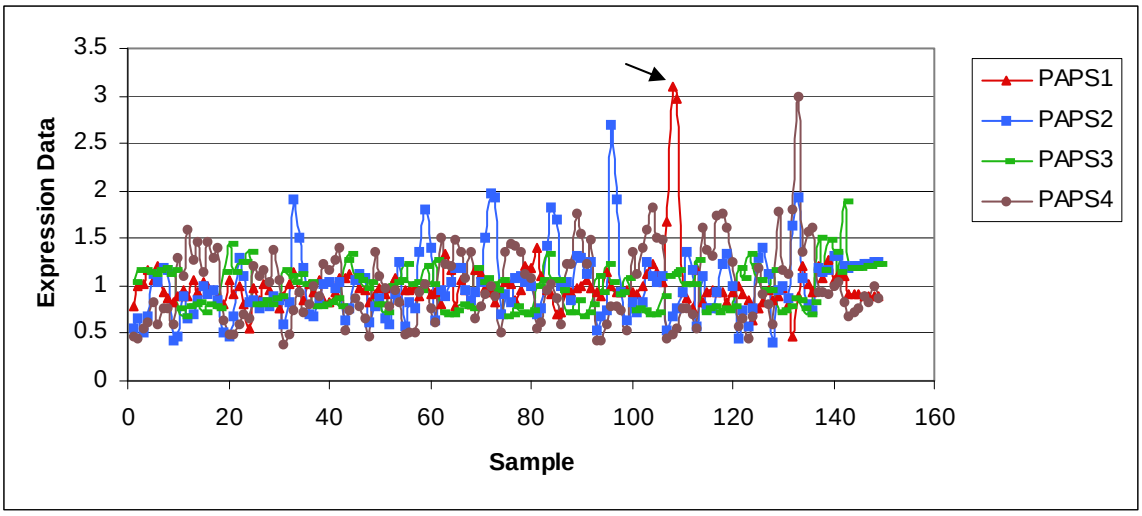


Figure 2.5: Expression data analysis for *Arabidopsis* PAP subunits during different abiotic stress conditions. The data obtained from NASC was plotted as shown. The respective PAP subunits are color-coded in the legend. The expression values are represented along the Y-axis. The samples that are represented along the X-axis: 1-18, control; 19-30, cold; 31-42, osmotic; 43-54, salt; 55-68, drought; 69-80, genotoxic; 81-92, oxidative; 93-106, UV-B; 107-120, wound; 121-136, heat; 137-141, cell culture control; 142-149, cell culture with heat. The arrow (→) shows the expression of PAP due to wounding in seedlings.

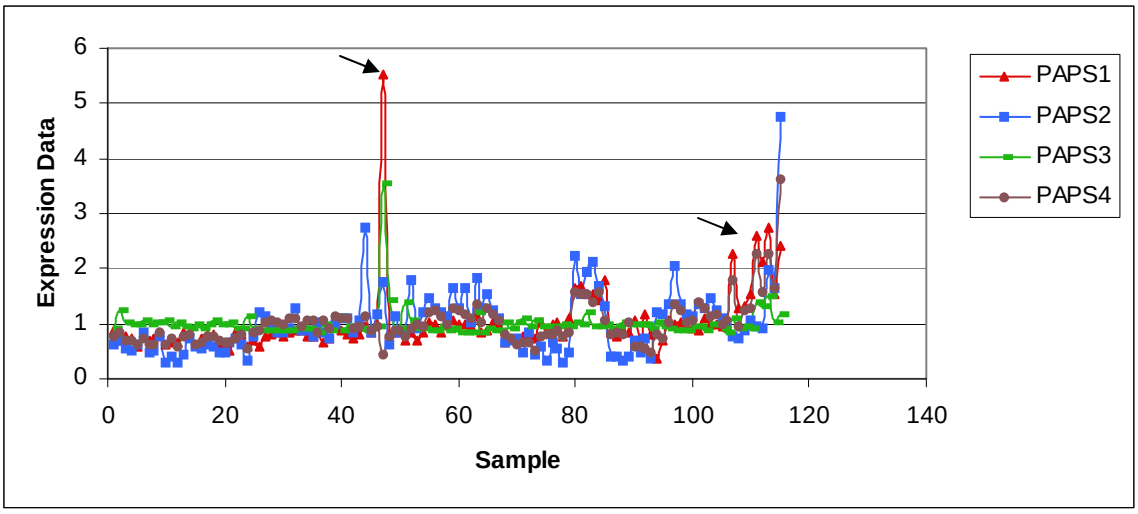


Figure 2.6: Expression data analysis for *Arabidopsis* PAP subunits during chemical and hormonal treatments. The data obtained from NASC was plotted as shown. The respective PAP subunits are color-coded in the legend. The expression values are represented along the Y-axis. The various samples for the chemical and hormonal treatments are represented along the X-axis. The arrows (→) show the expression of PAP during cyclohexamide treatment in seedlings and imbibition in seeds.

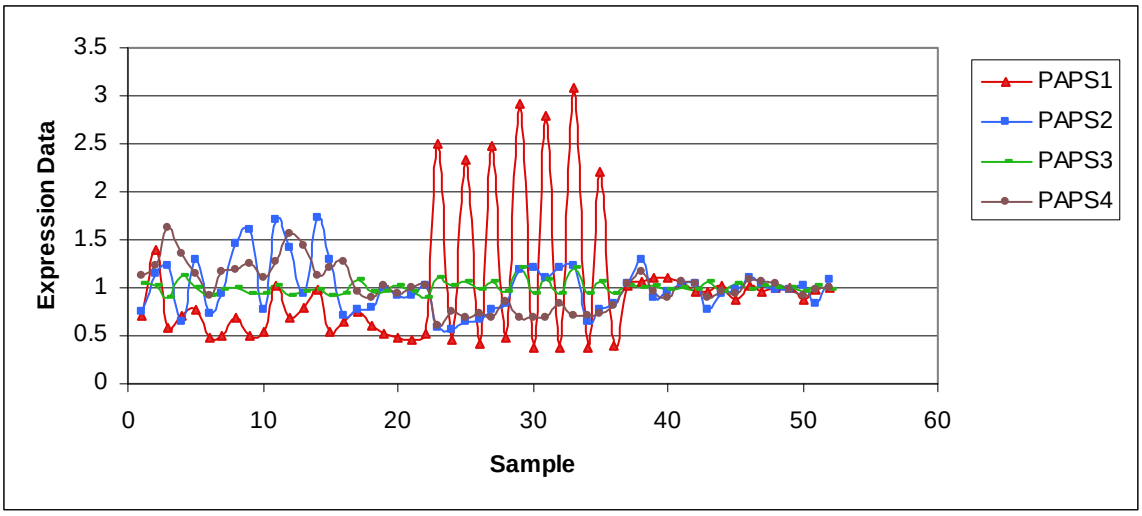
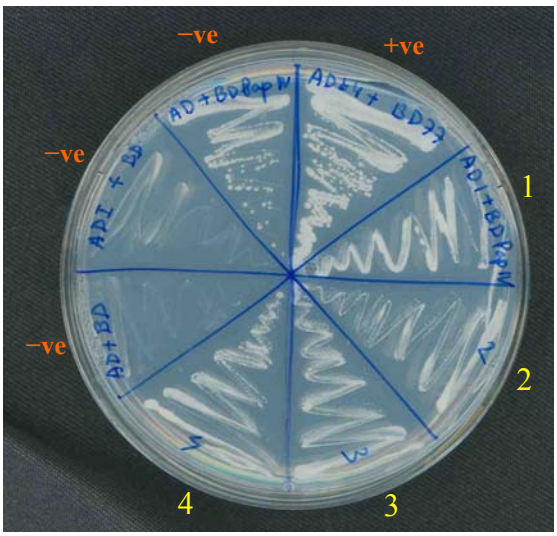


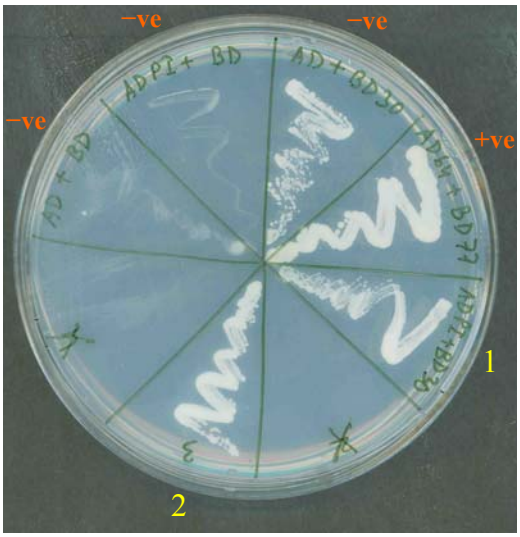
Figure 2.7: Expression data analysis for *Arabidopsis* PAP subunits during biotic stress and differential light conditions. The data obtained from NASC was plotted as shown. The respective PAP subunits are color-coded in the legend. The expression values are represented along the Y-axis. The samples that are represented along the X-axis: 1-16, control and *Pseudomonas syringae* infection; 17-22, control and *Phytophthora* infection; 23-36, control and other elicitors; 37-52, dark and different light conditions. Note the expression levels of PAPS1 (sample numbers 23-36) due to application of different elicitors at 1h and 4h durations.

Table 2.1: results from the yeast two-hybrid interactions between Arabidopsis PAPS1 (At1g17980) and other Arabidopsis polyadenylation factors. 'N'= negative interaction, 'Y'= positive interaction. '*' mark denotes incidence of self-activation which have been taken into account.

	Arabidopsis gene	PAPS1 (At1g17980)	
		Protein factor tested	Interaction (N/Y)
1	At5g51660	CPSF160	N
2	At5g23880	CPSF100	N
3	At1g61010	CPSF73-I	N
4	At2g01730	CPSF73-II	N
5	At1g30460	CPSF30	N *
6	At1g17760	CstF77	N
7	At1g71800	CstF64	N
8	At5g60940	CstF50	N
9	At5g13480	FY	N
10	At1g17980	PAPS1	N
11	At2g25850	PAPS2	N
12	At3g06560	PAPS3	N
13	**At4g32850**	**PAPS4**	**Y ***
14	At3g66652	FIPS3	N
15	At5g58040	FIPS5- NTD	N
		FIPS5- CTD	N
16	At4g25550	CFIS2	N
17	At4g29820	CFISI	N
18	At3g04680	CLPS3	N
19	At5g39930	CLPS5	N
20	At1g66500	PCFS1	N
21	At4g04885	PCFS4	N
22	At5g43620	PCFS5	N
23	At5g10350	PABN3	N
24	At5g51120	PABN2	N
25	At5g65260	PABN1	N



PAPS1 + PAPS4 (Y)*



PAPS1 + CPSF30 (N)*

Figure 2.8: Pairwise protein-protein interaction plates. The AD fusion protein and the BD fusion protein were co-transformed and dual transformants from SC-LW plates (not shown) were grown on SC-HLW selection medium (as seen) with controls. The positive control (+ve) is the interaction between CstF64 + CstF77 with 200% colony growth, negative controls (-ve) are the interaction between empty AD + BD vectors or AD-Test Gene + empty BD / empty AD + BD-Test Gene with almost 0% colony growth. Positive interaction plates are scored as "Y" and negative as "N". '*' mark denotes incidence of self-activation which have been taken into account. 1- 4 indicate number of samples tested.

Chapter 3: Characteristics of Cleavage Stimulation Factors 64 and 77 (CstF64 and CstF77) in *Arabidopsis*

3.1 Introduction

3'-end processing has sparked interest in the scientific world and has been investigated for more than 40 years now. This interest has been broadly generated for the diverse roles played by cleavage and polyadenylation process in all living organisms. Although much has been uncovered now, a cornucopia of facts regarding 3'-end processing is yet to be discovered.

It is now known that the process is catalyzed by a cumulative co-operation of a group of protein factors, working either as a sole subunit or a group of subunits (multimeric complex). Cleavage and polyadenylation can be uncoupled, *in vitro*, and hence has been studied in great detail. From these studies, it has been repeatedly observed that a consortium of protein factors, in mammals, is the key to the recognition of the DSE and subsequent stimulation of the cleavage process in the mRNA precursors. Thus, this key subunit came to be known as 'Cleavage stimulation Factor (CstF)' in mammals. From then on functionally conserved homologs of mammalian CstFs have been found in other organisms, some of which will be covered in this chapter.

Mammalian CstF subunit was the first to be studied in detail with respect to other organisms. In fact, the CstF64 protein factor was the second to be researched about, after PAP (discussed in Chapter 2)(Wilusz and Shenk, 1988; Takagaki et al., 1992). Besides the 64kDa polypeptide, mammalian CstF subunit also consists of 77kDa and 50kDa protein factors (Takagaki et al., 1990; Gilmartin and Nevins, 1991). UV-crosslinking

studies led to the detection of a protein, which later was proved to be the 64kDa subunit of mammalian CstF by immunoprecipitation by monoclonal antibodies raised against CstF64 (Moore CL, 1988; Wilusz and Shenk, 1988; Takagaki et al., 1990). Later on cloning and characterization of this subunit by Takagaki et al (Takagaki et al., 1992) shed light on the broad structural features (Fig. 3.1). The amino terminus, approximately 80 residue region, contains a ribonucleoprotein type RNA binding domain (RBD). It was hypothesized and later proven that the N-terminal RBD was necessary (Pérez Cañadillas JM, 2003) and sufficient (Takagaki and Manley, 1997) to bind mRNA precursors at their downstream element (DSE). The binding of CstF64, more so by the N-terminal RBD at the G/U-rich sequences at DSE, was corroborated by mutational studies (Takagaki et al., 1992) of AAUAAA and RasH (MacDonald et al., 1994) mapping later on. The N-terminal RRM domain (RBD) is followed by a stretch of about 100 residues known as the "hinge domain" (Takagaki et al., 1992; Takagaki and Manley, 2000; Qu et al., 2007). Next to this, the protein is composed of a stretch of 5-amino acid repeats (MEARA/G consensus sequence) which is repeated in a helical fashion 12 times in most mammalian CstF64. This region is embedded typically within a flexible proline-glycine rich region (Takagaki et al., 1992; Zhao et al., 1999). Much of the functional aspect of this region is unkown and the variable nature is the cause of the differential characteristics of CstF64 and its homologs in other organisms (Richardson et al., 1999; Hatton et al., 2000). The C-terminal region of about last 100 residues are much more conserved (Fig. 3.2) than the N-terminal RBD. Apart from protein-protein interactions for the 3'-end processing system, it plays a vital role in RNA termination and transcription (Aranda and Proudfoot, 2001; Qu et al., 2007).

Another important subunit of the cleavage stimulation factor complex is CstF77, which shares a surprising homology to *Drosophila su(f)* (suppressor of forked protein) gene (Takagaki and Manley, 1994). Takagaki et al showed that a 717-residue protein not only interacted with the 160kDa subunit of CPSF complex but also established links with CstF50 and CstF64, which themselves did not interact (Murthy and Manley, 1992; Takagaki and Manley, 1994). In the same study, they also showed that all 3 components of CstF complex were essential for activity and CstF77 acted as the bridging subunit (Takagaki and Manley, 1994). Structurally, (Fig. 3.3) CstF77 consists of amino terminus "Half a TPR" domain, otherwise known as HAT domain which closely resembles the tetratricopeptide (TPR) repeat (Preker and Keller, 1998; Bai et al., 2007). This region is followed by a proline rich domain spanning roughly 70 residues towards the carboxy terminus end of the protein. Embedded within the 9 repeat of the HAT domain is a bipartite nuclear localization signal (NLS), which is of considerable significance since CstF77 is the only CstF subunit having a NLS (Takagaki and Manley, 1994). Although the protein contains 2 hydrophobic regions, the overall protein is hydrophilic in nature (Takagaki and Manley, 1994).

The CstF50 subunit plays a vital role by interacting with CstF77 and Pol II-CTD, thus linking 3'-end processing machinery with transcription (McCracken et al., 1997). It contains 7 transducin or WD-40 repeats which are thought to mediate important protein-protein interactions (Takagaki and Manley, 1992; Manley, 1995). Other than 3'-end processing, CstF50 also interacts with BARD1 protein, associated to BRCA1 in DNA repair and tumor suppression (Kleiman and Manley, 1999, 2001).

CstF plays a very vital role in 3'-end processing by binding to the DSE and defining the poly (A) site for cleavage (Takagaki et al., 1989). It has been extrapolated to signal the end of cleavage by dissociating after its completion and before the start of polyadenylation, hence playing almost no role in the later stage of 3'-end processing (Zhao et al., 1999). In the initial stages of cleavage, CstF interacts with CPSF to form a multi-protein complex before binding to the mRNA-precursors at the polyadenylation signal. The CPSF160 subunit is sufficient for binding to the upstream AAUAAA signal sequence (Takagaki et al., 1992), but this otherwise weak binding is stabilized and strengthened by the co-operative binding of CstF through its 77kDa subunit and the RNA-binding 64kDa subunit (Wilusz et al., 1990; Murthy and Manley, 1992). Thus CstF77 not only interacts with the CPSF160 subunit (Murthy and Manley, 1995) via its HAT domain, but also acts as a bridge of sorts for CstF64 and CstF50 (Takagaki and Manley, 1994, 2000; Bai et al., 2007). The interactions within the CstF protein complex are of special significance. Firstly, the self-association activity of CstF77 via sequences just N-terminal to the proline rich C-terminal domain hint at possible dimerization of the subunit which have been further narrowed down to the HAT-C domain (Takagaki and Manley, 2000; Benoit et al., 2002; Bai et al., 2007). Secondly, the interaction between CstF64 and CstF77 is hypothesized to be crucial for the functioning of the former subunit in RNA recognition. This crucial interaction was first stated by Tagakaki in 1994 and later on substantiated by many (Takagaki and Manley, 1994, 2000; Bai et al., 2007). The 100 residues of the hinge domain of CstF64 interact with the proline rich region of CstF77. This interaction has far reaching implications in the overall assembly of the cleavage complex. The RRM of the CstF64 remains occluded by the C-terminal helix

situated close to the hinge domain and this prevents CstF64 from binding at the G/U-sequence downstream (Pérez Cañadillas JM, 2003; Qu et al., 2007). Following interaction with CstF77, at the hinge domain the C-terminal helix is unfolded and destabilized due to the conformational change, facilitating the binding of CstF64 to the DSE. It has been explained by Bai et al (Bai et al., 2007) through a model (Fig. 3.5) that the CstF77 homodimer associated by their HAT-C domain interacts with another set of CstF64 and CstF50 subunits, thus forming the CstF cleavage complex with 2 copies of each subunit (Legrand et al., 2007). This may explain how G/U-sequences with more than 10 nucleotides are recognized stably by the RRM of CstF64 that has a binding capacity of only 4 to 5 nucleotides (Pérez Cañadillas JM, 2003; Bai et al., 2007).

The CstF77 subunits simultaneously bind to the large CPSF160 monomer. This model is in line with the biochemical evidence provided to show that the CPSF-CstF subunits physically associate and organize themselves before the poly(A) signal recognition as already mentioned before. Thus once the initial cleavage complex of CPSF-CstF forms, the bipartite signal sequences are read in one single interaction (Takagaki and Manley, 2000). This helps to lower the chances of false recognition; fine tuned all the more by sequences within the RRM of CstF64 and also expedites the processing efficiency of the 3'-end machinery (Pérez Cañadillas JM, 2003).

As it has been mentioned earlier, apart from nuclear polyadenylation, CstF complex also has functions in other cellular processes like transcription termination, splicing, alternative polyadenylation, nucleo-cytoplasmic transport, cytoplasmic polyadenylation and so on.

The role of CstF50 in transcription termination has already been discussed before (McCracken et al., 1997). Although CstF64 or CstF77 does not directly interact with RNA polymerase II, the association of CstF with CPSF, PC4 (a transcriptional co-activator) and assembly of the cleavage complex at the recognition sequence all signal the completion of the formation of a pre-mRNA transcript and disengages RNA pol II (McCracken et al., 1997; Calvo and Manley, 2001). In mammals, termination does not require pre-mRNA cleavage, but the assembly of the cleavage multiprotein complex is a pre-requisite for efficient transcription termination and 3'-end processing (Zorio and Bentley, 2004). As a consequence of faulty cleavage mechanism and subsequent aberration in transcription termination, nucleo-cytopalsmic transport may be suspended. This results in retention of mRNA transcripts in the nucleus, poor release from the site of transcription, failure of the exporting factors to engage and even degradation by the nuclear exosome (Brodsky and Silver, 2000; Lei and Silver, 2002, 2002; Libri et al., 2002; Torchet et al., 2002). This has been mentioned in some detail in chapter 1.

Niwa et al showed that 3'-end processing and splicing were also related in some cases (Niwa et al., 1992). In case of terminal exons, presence of a 5'-splice site 300 nucleotides or closer to 3'-site caused a depression of polyadenylation and also reduced the binding of CstF64 to mRNA. In some instances, splicing went hand in hand with alternative polyadenylation. Apart from the 64kDa protein, another form of CstF64 was reported in mouse and rat tissues, especially in the male gametic cells and to a lesser extent in brain cells (Wallace et al., 1999; Wallace et al., 2004; Huber et al., 2005). The 70kDa protein, named as τCstF64, was expressed from an autosomal chromosome (chromosome 19 in mouse) only at the meiotic and post-meiotic phases (Dass et al.,

2001). This difference in expression of an important gene in 3'-end processing suggested a different mode of polyadenylation during spermatogenesis showing a high prevalence of alternative polyadenylation and non-AAUAAA site usage (Monarez et al., 2007). A possible outcome of this characteristic is that τCstF64 has a differential and more non-specific binding affinity in comparison to CstF64 (Monarez et al., 2007). Another well researched instance, which has been mentioned in Chapter 1, is that of the IgM gene expression during the B-lymphocyte cell maturation. A distal, stronger poly (A) site is chosen over a weaker, proximal poly (A) site before maturation, but this reverses during the secretory stage after the B-cells are fully differentiated (Edwalds-Gilbert and Milcarek, 1995; Takagaki et al., 1996). The fact that the level of CstF64 is the limiting factor for the shift in poly (A) site usage has been a matter of controversy (Martincic et al., 1998; Takagaki and Manley, 1998), but it has been shown that CstF64 increases dramatically in differentiated cells and is vital for cell viability at initial stages.

Alternative polyadenylation has also been observed with the gene that codes for the *Dropsophila* homolog of CstF77, *su(f)*. In keeping with the function of CstF77 in 3'-end processing of pre-mRNAs so far, any mutation in *su(f)* gene similarly affects polyadenylation efficiency and mRNA stability (Mitchelson et al., 1993; Takagaki and Manley, 1994). The protein shows high homology (56.2% identity, 69.4% similarity) overall to CstF77 and conservation of 14 out of 15 pro-rich residues in the region implicated in many of the protein-protein interactions (Takagaki and Manley, 1994). It has 9 exons from which 3 transcripts are generated by alternative polyadenylation. The shortest transcript utilizes a poly(A) site within an intron and hence is a 1.3kb in length, and encodes a polypeptide lacking much of the HAT domain (Mitchelson et al., 1993). It

has been proposed that *su(f)* mRNA accumulation is autoregulated by CstF77 in a tissue specific manner (Audibert and Simonelig, 1998; Juge et al., 2000). Surprisingly, a similar kind of intronic poly(A) site has also been discovered in vertebrates and, although, they are not aberrantly degraded, their expression and functional characterization are yet to be investigated (Pan et al., 2006).

CstF77, an indispensable subunit in nuclear polyadenylation, has been also seen to have some function in cytoplasmic polyadenylation. It remains associated with cytoplasmic polyadenylation factors CPEB, CPSF 100, XGLD2 and eIF4E and helps to stabilize the cytoplasmic cleavage unit, much like the nuclear cleavage unit (Rouget et al., 2006). Although, it is not essential for cytoplasmic polyadenylation, impairement of *Xenopus* homologue of CstF77 function accelerates oocyte maturation and protein synthesis without modifying polyadenylation *in vivo*, while *in vitro* it represses mRNA translation. Thus, it was concluded that *Xenopus* CstF77 plays a role in mRNA masking (Rouget et al., 2006).

Research in the 3'-end processing system in yeast progressed in parallel as that of their mammalian counterparts. Initially cleavage/polyadenylation factor I (CFI) was identified and later on separated into CFIA and CFIB for the purpose of cleavage and polyadenylation in yeast (Chen and Moore, 1992; Kessler et al., 1996). CFIA consists of 4 polypeptides, 2 of which will be discussed to some detail in here since they have homologs in the mammalian CstF subunit. The 38kDa Rna15 polypeptide has a 42.5% identity and 62.5% similarity to its mammalian CstF64 counterpart (Fig. 3.2), especially within its RRM-type RBD at the amino terminus (Minvielle-Sebastia et al., 1991; Takagaki and Manley, 1994). The 76kDa Rna14 polypeptide is the yeast homolog of

mammalian CstF77 (Minvielle-Sebastia et al., 1994). It shares moderate homology (Fig. 3.4) to its mammalian counterpart (24.3% identity, 37.2% similarity). Although there are certain differences in the polyadenylation signal sequences and their respective trans-acting factors, functional and to some extent, the structural homology evidently proves that 3'-end processing is conserved from yeast to mammals (Takagaki and Manley, 1994).

Rna15 is considerably shorter than CstF64 and the structural homology ends at the amino terminal RBD and the hinge domain, 100 residues upstream of the carboxy terminus of the mammalian CstF64 protein. The RBD of Rna15 is followed by a stretch of glutamines and asparagines at the C-terminal end (Minvielle-Sebastia et al., 1991; Zhao et al., 1999) (Fig. 3.1). The RRM-type RBD, although can be UV-crosslinked to substrate RNA, behaves differently from that of CstF64 since the substrate RNAs do not interact with the RBD of mammalian CstF64 (Takagaki and Manley, 1997). This deviation might stem from the difference in the positions of the recognition sequences described in Chapter 1, in the yeast system with respect to the mammals. The UA-rich efficiency element (EE) which is functionally analogous to the DSE in mammals is placed upstream of the cleavage site, 10 nucleotides upstream to the A-rich positioning element (PE). Unlike CstF64, it was observed that Rna15, in the presence of Rna14 and Hrp1, specifically bound to the A-rich PE (Kessler et al., 1997; Gross and Moore, 2001). The strong interaction between Rna15 and Rna14 is conserved in yeast (Kessler et al., 1996; Kessler et al., 1997), but the interacting domains are slightly different from their mammalian counterpart due to differences in the structural motifs (Fig. 3.1 and Fig. 3.3) (Minvielle-Sebastia et al., 1991; Noble et al., 2004). The protein encoded by Rna14 (Fig.

3.3) contains 6 HAT domains roughly divided into the N-terminal HAT and the C-terminal HAT domains, but lacks the proline rich carboxy-terminal domain (Noble et al., 2004; Legrand et al., 2007). About 100 residues towards the C-terminal end are indispensable for interaction with Rna15 and other proteins in the 3'-end processing complex (Minvielle-Sebastia et al., 1994; Gross and Moore, 2001; Noble et al., 2004; Legrand et al., 2007). A NLS is embedded within the C-HAT domain, but occasionally Rna14 is also localized in the cytoplasm especially in the mitochondria (Minvielle-Sebastia et al., 1991). This suggests a possible role of Rna14 in mitochondrial metabolism (Minvielle-Sebastia et al., 1991; Rouillard et al., 2000). In 2004, Noble et al performed *in vitro* experiments to show that Rna14-Rna15 not only form heterodimers but also has a penchant to form tetramers (Noble et al., 2004). This was later confirmed to be the property of the C-HAT domain which do form stable heterotetramers (Bai et al., 2007). The dimerization not only helps in Rna15-RNA binding by increasing the number of RRMs available per PE, but also exposes an area of conserved residues for a number of protein-protein interactions (Bai et al., 2007; Legrand et al., 2007). Thus Rna14 plays a very important role in assembling the cleavage complex by bringing together Hrp1 and Rna15, bridging the various components of CFIA and PFI (through Pcf11 and Pfs2) and signaling the end of cleavage to the awaiting polyadenylation complex by interacting with Fip1 (Gross and Moore, 2001; Helmling et al., 2001). On the other hand it was observed by deletion mutation that unlike CstF64, the residues C-terminal to the hinge domain in Rna15 was responsible for the interaction with Rna14 and also crucial for interactions with some transcriptional activators and co-activators (Aranda and Proudfoot, 2001; Calvo and Manley, 2001). The C-terminal approximately 50 residues

form unique helical structures which are important for various protein-protein interaction in 3'-end processing and transcription termination (Birse et al., 1998; Aranda and Proudfoot, 2001). Disruption in this region causes defects in RNA processing and loss of interaction with Pcf11, an important subunit in the 3'-processing complex having interactions with RNA polymerase II (Meinhart et al., 2005; Qu et al., 2007). Hence, this region is indispensable for 3'-end maturing but also is important for transcription termination, through interactions with transcription specific factors (Proudfoot and O'Sullivan, 2002). Mutational studies led to the conclusion that Rna15, Rna14 and the other protein factors of CFIA are vital at both the steps of 3'-end processing in yeast (Minvielle-Sebastia et al., 1991). This feature of CFIA deviates much from the mammalian CstF subunit, which is important only for cleavage and in this vein rather mimics the CPSF subunit to some extent, having roles in both cleavage and polyadenylation.

Scientific investigations in animals and yeast system have always paved the path for further studies in other living systems like plants. Although the process of acquiring knowledge for plant system had a late start, yet it never lagged behind, for long. This has also been the case for the study in cleavage and polyadenylation machinery in plants. The subtle similarities and dissimilarities in the recognition sequences have already been discussed in Chapter 1. Now that the Pandora's box of trans-acting factors related to the process are being discovered, structural and functional uniqueness in plants, compared to the mammalian and yeast systems, are also being disclosed. Some of these features relating to *Arabidopsis* homolog of PAP have been discussed in Chapter 2. Experiments in the recent past showed that the *Arabidopsis* orthologs of CstF64 and CstF77 (AtCstF64

and AtCstF77, respectively) interacted very strongly, *in vitro* (Yao et al., 2002). Thus it can be safely assumed that this interaction is one of the most important interactions within the cleavage/polyadenylation machinery and conserved from yeast to mammals to plants. But from the same experiments it was found that the Arabidopsis ortholog of CstF50 (AtCstF50) does not interact with the remaining CstF subunits in *Arabidopsis*, which deviates from the mammalian CstF50. Extensive two-hybrid studies later on corroborated this and also showed that a possible interaction rather exists between AtCstF50 and AtCPSF100 (Forbes et al., 2006). The *Arabidopsis* homolog of Fip1 [AtFip1(V)] shows a moderate (26% identity, 40% similarity) similarity to hFip1, but more closely resembles to yFip1 (38% identity, 56% similarity) within its conserved domains (Forbes et al., 2006). In addition to RNA-binding and its interaction with PAP, CstF77 and a number of other 3'-end processing factors, AtFip1(V) interacts with AtCstF64; this interaction is unique to plants since it is absent in yeasts and human Fip1 homologs (Hunt et al., 2008). These facts raise a curious question as to the structural and functional characterization of AtCstF64 and AtCstF77 and their potential binding partners so far discovered. Detailed studies through *in silico* analysis, two hybrid system, *in vitro* and expression in plant system, *in vivo* brought to surface other features that give distinctiveness to the 3'-end processing in plants and abolishes the direct comparison of the system to yeast and mammals. This chapter largely focuses on some of the characteristic features of the *Arabidopsis* homologs of CstF64 and CstF77 (AtCstF64 and AtCstF77) through *in silico*, *in vitro* and *in vivo* studies.

3.2 Experimental procedures

3.2.1 In silico expression analysis for AtCstF64 and AtCstF77

The *in silico* gene expression analysis was performed using the data available in the Additional file 1: microarray keys and data from Hunt *et al* (Hunt et al., 2008), which was compiled from the NASC (Nottingham Arabidopsis Stock Centre) microarray database (Craigon et al., 2004). The datasets for AtCstF64, AtCstF77 and AtCstF50 were obtained from 4 different experimental conditions (Developmental stages, Abiotic stress conditions, Chemical and hormonal responses and Biotic and differential light responses) (Appendix: A2). The data was plotted in a XY-scatter graph in most cases for analysis. In case of the developmental stages, the data was also depicted in bar diagrams especially with the comparative expression values at the pollen developmental stage for ease of analysis.

3.2.2 Interaction assay for AtCstF 64 and AtFip1(V) in a GAL4 based two-hybrid system

A Gal4-reporter based two-hybrid system with the yeast strain PJ69-4A was used for this interaction assay. The expression vectors were pGAD-C(1) for the activation domain and pGBD-C(1) for the binding domain (James et al., 1996; Forbes et al., 2006). For a detailed description of the system refer to Chapter 2, section 2.2 Experimental procedures. For cloning the 3 portions of AtCstF64, the cDNA sequence information for AtCstF64 (At1g71800) was downloaded from "The Arabidopsis Information Resource (TAIR)" (http://arabidopsis.org/index.jsp). Based on the sequence information, oligonucleotides (Table 3.1) were designed with the intention of amplifying the N-terminal 450bp (~150 amino acid residues), the hinge region 250bp (~85 amino acid residues) and the C-terminal 300bp (~100 amino acid residues) of the cDNA for

AtCstF64. The amplified parts were cloned firstly into pGEM-T™ vector system (Promega) and then excised with *Bgl II*. These fragments were then cloned in the pGAD-C(1) and pGBD-C(1) vectors and then sequenced to confirm the correct reading frames of the gene fusions thus created. The pGAD and pGBD clones for AtFip1, AtCstF77 were obtained from Dr. Kevin P. Forbes (Forbes et al., 2006). The AtFip1 was cloned into 2 portions. The N-terminal 161 residues and the C-terminal 263 residues were individually tested for interactions (Forbes et al., 2006). Procedures for transformation and subsequent incubation into competent yeast cells were similar to the one used in Chapter 2, section 2.2 Experimental procedures. The process of choosing the controls for this study and the criteria for the observation of the positive interactions was also same to the one that was followed in Chapter 2, section 2.2 Experimental procedures.

3.2.3 Agrobacterium mediated infiltration of AtCstF64 and AtCstF77

Sequence information for AtCstF64 and AtCstF77 was downloaded from "TAIR" (http://arabidopsis.org/index.jsp) and the whole coding regions for each were amplified using the designed primers (Table 3.1) by standard PCR methods. Each of the PCR products were subcloned into pGEM-T™ vector system (Promega). AtCtF64 was excised with *Bgl II* and *Hind III* and AtCstF77 was excised with *Sal I*. Both were individually cloned into pGD-GFP (pGDG) and pDG-DsRed (pGDR) vectors. The clones were confirmed by sequencing. The controls for the experiment were the empty pGDG and pDGR vectors along with a nuclear signal cloned into the pGDG and pGDR (Nu_{GFP} and Nu_{DSR}) vectors. All the controls were kindly contributed by Dr. Michael M. Goodin from his work (Goodin M M et al., 2002). The C58C1 strain of *Agrobacterium tumefaciens* was transformed with the pGDG-64 (64_{GFP}), pGDR-64 (64_{DSR}), pGDG-77

(77_{GFP}) and pGDR-77 (77_{DSR}) plasmids using freeze-thaw method described by An et al (An et al., 1988). They were incubated on $LB^{RifTetKan}$ (LB + 100µg/ml Rifampicin + 5µg/ml Tetracycline + 100µg/ml Kanamycin) media plates at 28°C for 2 days. For the infiltration suspension, 2 day old transformed cells were mixed in MES buffer (10mM MES, pH 5.6, 10mM $MgCl_2$) and the OD_{600} was adjusted (generally 0.6-0.7) to the desirable limit. 100mM acetosyringone was added to this final suspension (1.5µL/ ml of suspension solution) and the bacterial preparation was incubated, undisturbed at RT (28°C) for 2-3 hrs. *Nicotiana benthamiana* plants were propagated under described greenhouse conditions (Martins et al., 1998) and used as plant material for all further agroinfiltrations. Normally 2-3 leaves (2-3 cm wide at mid-leaf) per plant were chosen for infiltration. Using 1 ml disposable syringe the bacterial cell suspension was gently infiltrated from the abaxial side of the chosen leaves. When correctly infiltrated, the leaves looked darker and water-soaked. It is a good idea to demarcate the region of infiltration with light pin-pricks. After the completion of infiltration the plants were kept under observation at 25°C, 8h/16h light/dark photoperiod for approximately 60 hrs before microscopic observations were performed. The infiltrated parts of the *N. benthaminana* plants were mounted in water and observed under epifluorescence microscope (Zeiss Axioplan2 HB100) using the GFP (ex: D470/40; em: D535/40; beamsplitter 500 DCLP) and DsRed (ex: HQ545/30X; em: HQ610/75M, Q570LP) filter sets. Images of the required frames were captured using the attached camera (Zeiss AxioCam MRc5) and adjusted with the provided software (AxioVision).

3.3 Results

3.3.1 Characteristics of AtCstF64 and AtCstF77 from the in silico expression analysis

The expression of certain genes may vary according to various states in organisms especially in plants. Whether the expression of polyadenylation factors in *Arabidopsis* is also variable was explored by analyzing microarray data. Expression data sets were obtained for 4 kinds of experimental variations in *Arabidopsis*: developmental stages, responses to abiotic stress, responses to chemical and hormonal treatment, and responses to biotic stress and differential light conditions. The data were downloaded from a public domain for microarray data (NASC) and plotted (Appendix: A2). Out of the 4 experimental conditions, results shown in Fig. 3.6 are the most remarkable. This figure shows that around the pollen development stage there is a gross difference in the expression of CstF64 and CstF77. This differential expression is emphasized in Fig. 3.6C, showing that the expression of CstF64 in pollen increases by almost 6 fold while the expression for CstF77 drops by almost 10 fold. Also noticeable to some extent (Fig 3.6B) is the increase in CstF77 expression in the seeds during embryo development by almost 3 fold. The expression data for the abiotic stress (Fig 3.7) shows a 2-3 fold increase in the expression of CstF64 during cold shock treatment to young seedlings. The expression data for the chemical and light conditions (Fig 3.8) shows a dramatic 8 fold increase in CstF77 expression in seeds undergoing imbibition, but gradually decreases with the concomitant increase in CstF64 (almost 4 fold) with time. The expression data for biotic stress (Fig 3.9) induced by pathogen and other elicitors in the leaf show no significant change in CstF64 and CstF77 expression.

3.3.2 C-terminal domain of AtCstF64 does not interact with AtFip1(V)

Few facts about the interaction between CstF64, CstF77 and Fip1 have already been brought out from past research (Table 3.2). For example, the conserved interaction between the hinge domain of CstF64 and pro-rich C-terminus of CstF77 has been established from yeasts to mammals (Takagaki and Manley, 1994; Kessler et al., 1996; Hatton et al., 2000; Takagaki and Manley, 2000; Bai et al., 2007). Also Forbes *et al* showed through another two-hybrid assay that the N-terminal domain of Fip1(V) does not interact with AtCstF64 but does so with the C-terminal domain of AtCstF77 (Forbes et al., 2006). So by filling in the gaps from these facts the interaction between the CTD of Fip1(V) with all the 3 parts of AtCstF64 is a matter of interest. Also to analyze in this process is the interaction between AtCstF64 and Fip1(III), another possible Arabidopsis homolog of Fip1 on the 3rd chromosome. All these interaction studies may give us a cursory insight into an important interaction "hub" in *Arabidopsis* polyadenylation machinery and hence can be extrapolated to the plant kingdom as well.

The C-terminal 300bp (100 residues) of the AtCstF64 coding region was introduced into the AD and BD vectors for use in two-hybrid assays. Fip1(V) clones (Forbes et al., 2006) were obtained for this study. The standards for positive and the negative controls for the assays were the same ones used in Chapter 2. After transformation into competent yeast PJ69-4A cells and subsequent plating, the plates were scored for positive interactions in the selective medium (SC-HLW). The results (Table 3.2) showed that there were no interactions between the C-terminal end of CstF64 and N-terminal domain of Fip1(V), C-terminal domain of Fip1(V), full length Fip1(III) and full length CstF77 (Fig 3.10). These results are preliminary, as controls that show

that the CstF64 individual domains are expressed in yeast and can interact with some other protein were not done. However, tentatively, we may conclude that the interactions of AtCstF64 do not involve the C-terminal part of the protein.

3.3.3 The subcellular localization of AtCstF64 and AtCstF77

Functional characterization is an important part in the delineation of a gene and its product. Determining the cellular targeting signals and regions of subcellular accumulation can hence play a major role in this direction. Specifically, co-localization studies can provide potential insights into the behavior of one protein in close association with another. With the growing list of genes, the need for quick, simple and cost-effective analytical methods are in high demand and widely used. One such novel system of studying protein localization in plants is through infiltration of *N.benthamiana*, mediated by *A.tumefaciens*, transformed with suitable desired vectors. This system has been successfully used for studying different kinds of protein localization and interaction (Goodin M M et al., 2002) and hence was the choice system for studying the localization of AtCstF64 and AtCstF77 in plants too. The controls that were used in this study were nuclear signal targeted pGD-GFP and pGD-DsRed (Nu_{GFP}, Nu_{DSR}) vectors and their respective empty vectors without any targeting sequences. Both AtCstF64 and AtCstF77 were individually cloned into the pGD vectors (pGDG and pGDR). All the controls and the test genes were tested in individual plant system. Expression of Nu_{GFP} and both the empty vectors (pGDG and pGDR) could easily be seen (Fig. 3.11) whereas the Nu_{DSR} construct yielded very poor expression (not shown). Readily-detectable levels of pGDR-64 (64_{DSR}) and pGDR-77 (77_{DSR}) could be seen as well (Fig. 3.12 and Fig. 3.13A). Interestingly, 64_{DSR} distribution was indistinguishable from pGDR alone (Fig. 3.12),

which might be due to the absence of any targeting sequence as predicted from its structural features. 77_{DSR} seems to accumulate in the nuclei (Fig. 3.13B, C, D), in accordance with the predicted nuclear targeting sequence present at the C-terminal end of the protein, much like that in other living systems (Bonneaud et al., 1994; Takagaki and Manley, 1994; Rouget et al., 2006). This was confirmed by co-expressing 77_{DSR} and Nu_{GFP}; the results showed a coincidence of distribution of the nuclear marker and 77_{DSR} (Fig. 3.13D).

3.4 Discussion

From the results described in this study, the functions of the CstF64 and CstF77 subunits of the CstF complex in 3'-end processing in plant pre-mRNAs, are starting to look intriguing. The *in silico* expression analysis has yielded some curious facts regarding the expression in pollen and seeds during the developmental phase. The first one that demands notice is the ~6 fold increase in CstF64 expression and ~10 fold decrease in CstF77 expression in pollen development. The comparatively higher expression of CstF64 is quite expected considering the major role of CstF64 in RNA binding and establishment of stability of the cleavage complex. But the drop in the expression of an equally important factor like CstF77 is an indication of an alternative mode of 3'-end processing in the male germ cells (pollen) much like the novel processing factor during spermatogenesis in mice (Wallace et al., 1999; Monarez et al., 2007). The same can be predicted about the sudden increase of CstF77 expression during embryonic development in seeds, where mobilization of seed storage proteins calls for heavy duty mRNA transcription and translation machinery. This might have some likeness to the role played by CstF77 during *Xenopus* oocyte maturation (Rouget et al., 2006), where the

cytoplasmic CstF77 masks the available mRNA till they are needed for translation. These facts necessitate the investigation of CstF77 and its function especially relating to pollen and seed development in plants.

The importance of CstF64 as a cleavage/polyadenylation factor has been established through various studies in other living systems, as already mentioned in the introducing literature. It not only affects cleavage complex assembly, stability and processivity, but also plays roles in other nuclear processes like transcription termination, nuclear transport, and stability of mRNA. The binding of CstF64 and its homologs in other organisms to RNA precursors and also interactions with other protein factors contribute to the functioning of the cellular processes. One such contact that seems unique to plants is with the Arabidopsis homolog of Fip1 [Fip1(V)] protein. Fip1 is one of the major proteins which harbors a great deal of protein interactions and links cleavage to polyadenylation. In some cases, its major interaction with poly(A) polymerase (PAP) has been assumed to go so far as to increase the RNA binding affinity of PAP by itself binding to mRNA precursors (Kaufmann I et al., 2004; Forbes et al., 2006). But the rationale behind the interaction between AtFip1(V) and AtCstf64 has gone unexplained. Hence, it was imperative to gather an idea about the interaction domains responsible for this contact. Previous similar experiments have shown that the N-terminal 137 residues of Fip1 had interactions with PAP and AtCstF77 (Forbes et al., 2006; Hunt and Addepalli, 2008), whereas the remainaing C-terminal 789 residues interact with AtCstF64 and bind to RNA (Forbes et al., 2006; Hunt et al., 2008). On the other hand, studies in two-hybrid system in here, show that the C-terminal (~100 residues) region of AtCstF64 failed to interact with not only the amino and carboxy end of AtFip1(V), but also AtFip1(III) and

AtCstF77. This brings us down to only 2 regions for interaction in AtCstF64 – the N-terminus RRM domain (~150 residues) and the hinge region (~85 residues), the latter of which have already been implicated to interact with CstF77 (Takagaki and Manley, 1994; Kessler et al., 1996; Hatton et al., 2000). Once solved, this will lend us some idea in defining the interaction pattern among these protein factors and bring forth one of the working models for it. Multiple interactions within the same domain will vote for a sequential model, where the factors interact at different stages of the process. Conversely, situation devoid of overlapping interaction domains will be more suggestive of a concurrent model, where the proteins exist in a multimeric complex and function at the same time as a unit.

One question that arises from the reported binding of AtFip1(V) to FUE (Forbes et al., 2006), is the possible candidate for binding at NUE and the endonuclease for the poly(A) site. Binding of AtFip1(V) to precursor RNA and its interaction with AtCstF77 is analogous to that of hFip1, but binding at the GU-rich FUE sequence of pre-mRNA mimics the binding of Hrp1 at the PE in yeasts (Kessler et al., 1997; Gross and Moore, 2001; Kaufmann I et al., 2004; Forbes et al., 2006). These diverse interaction properties add to the uniqueness of the processing unit in plants. But how does the interaction between AtFip1(V) and AtCstF64 fit in the bigger puzzle? This might be answered with future endeavors in detailed characterization of AtCstF64 and one of its major interacting partners, AtCstF77.

The other question that is very crucial to the functioning of the whole CstF unit in plants is the binding of CstF64 to the available signal sequences especially NUE or other sequences within the pre-mRNA via its RRM-like RBD. Although there are numerous

candidates having the RNA-binding property, functional and sequence comparison with other organisms predict CstF64 to be the likely occupant of NUE (MacDonald et al., 1994; Hatton et al., 2000; Gross and Moore, 2001). Since, it is known from other sources that CstF64 binds RNA only in close interaction with other protein factors, it was necessary to find out the likely binding partners for it and their possible mechanism of interaction. One of such factors which have a conserved interaction with CstF64 and its homologs is CstF77 (Yao et al., 2002). But other than AtCstF64, AtCstF77 also interacts with AtFip1(V), AtCPSF30, AtPCF1, and AtPCF5 (Forbes et al., 2006; Hunt et al., 2008). This indicates at possible role for AtCstF77 in controlling the endonuclease activity of AtCPSF30 with respect to its interaction with AtFip1(V). It might be an optimistic guess to assume that the interaction of AtFip1(V) with AtCstF77 releases its inhibitory effect on AtCPSF30 (Barabino et al., 1997; Tacahashi et al., 2003; Delaney et al., 2006; Forbes et al., 2006; Addepalli and Hunt, 2007) and thus, activates the nuclease property of the later for the cleavage action in times of need. Biochemical assays, *in vitro*, like EMSA could produce some probable answers to this type of interaction and also other interacting partners.

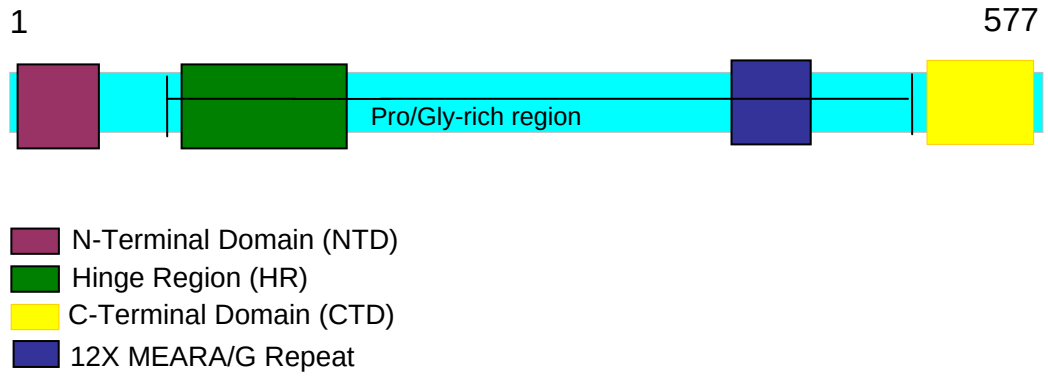
While *in vitro* studies say much about the canonical aspects of the proteins studied in here, experiments performed *in vivo* also demonstrate the actual functional aspects of AtCstF64 and AtCstF77. CstF64 lacks any organelle targeting sequence like NLS, but is nonetheless found in the nucleus (near Cajal bodies) in mammals (Schul et al., 1996). CstF77, which has a NLS, has been shown to be targeted to the nucleus in case of mammals and yeast (Rouget et al., 2006). In this study localization of AtCstF77 was

also found to be in the nucleus (Fig. 3.13). The Arabidopsis CstF64, expressed by itself (Fig 3.12), is located in the cytoplasm.

3.5 Conclusion

The dearth of knowledge about the cleavage/polyadenylation complex in plants opened a doorway to vast information resources. Scientists from then on have been working on unearthing these facts and building a coherent description of the mechanism behind the processing complex. From what has been known till now, it would be grossly wrong to draw a parallelism between the plant and the animal system. Studies in this discourse about *Arabidopsis* homologs of CstF64 and CstF77 add all the more to that existing uniqueness. Since the CstF complex, especially the CstF64 and CstF77 subunits, play a central role in 3'-end processing, any difference in their structural and/or functional characteristics suggests possible differences in the architecture of the processing unit in plants. Although the results in here are far from being conclusive, the discoveries made during the course of this study are pieces of a jigsaw puzzle that might help to unravel the 3'-end processing machinery within the plant kingdom. Even though the overall function of the system might be conserved in plants, structural differences starting from the cis-acting sequences to the trans-acting protein factors make cleavage and polyadenylation processes distinct in their own sense.

(A)



(B)

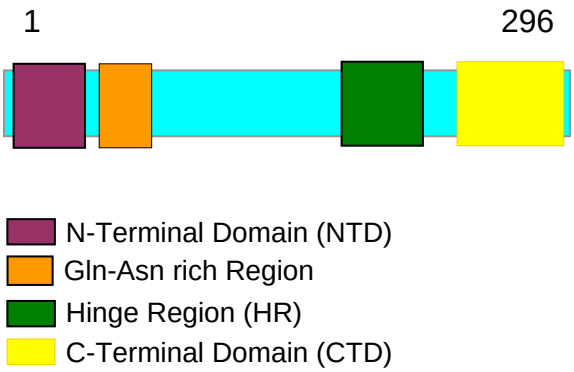


Figure 3.1: Schematic representation of CstF64 (A) from human and Rna15 (B) from yeast (*S.ceravisea*). The respective domains are depicted in the legend (for detailed functions of the domains refer to preceding text).

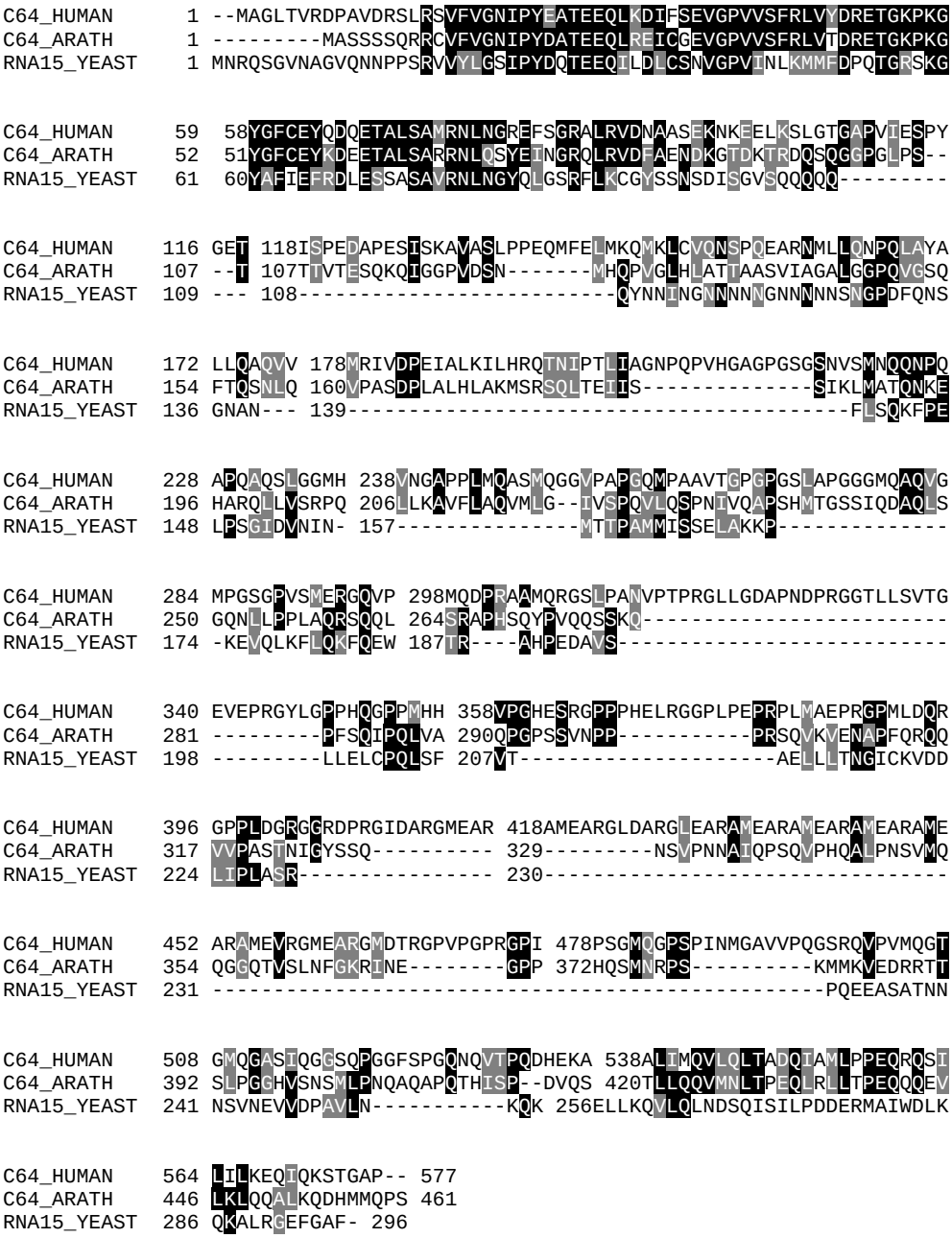
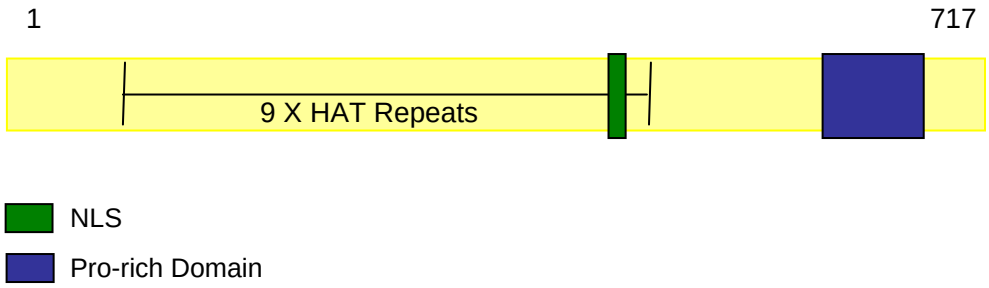


Figure 3.2: Amino acid sequence alignment and comparison of CstF64 and its homologs. The open reading frames encoded by human (C64_HUMAN, UniProtKB/TrEMBL P33240), yeast (RNA15_YEAST, UniProtKB/TrEMBL P25299) and Arabidopsis (C64_ARATH, UniProtKB/TrEMBL Q9M9G6) are compared using ClustalW and formatted using BoxShade. Residues which are identical in all three polypeptides are shaded in black, with white uppercase lettering. Positions that are similar are shaded in grey, with white uppercase lettering.

(A)



(B)

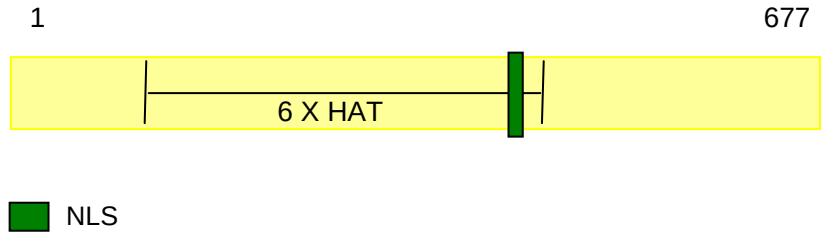
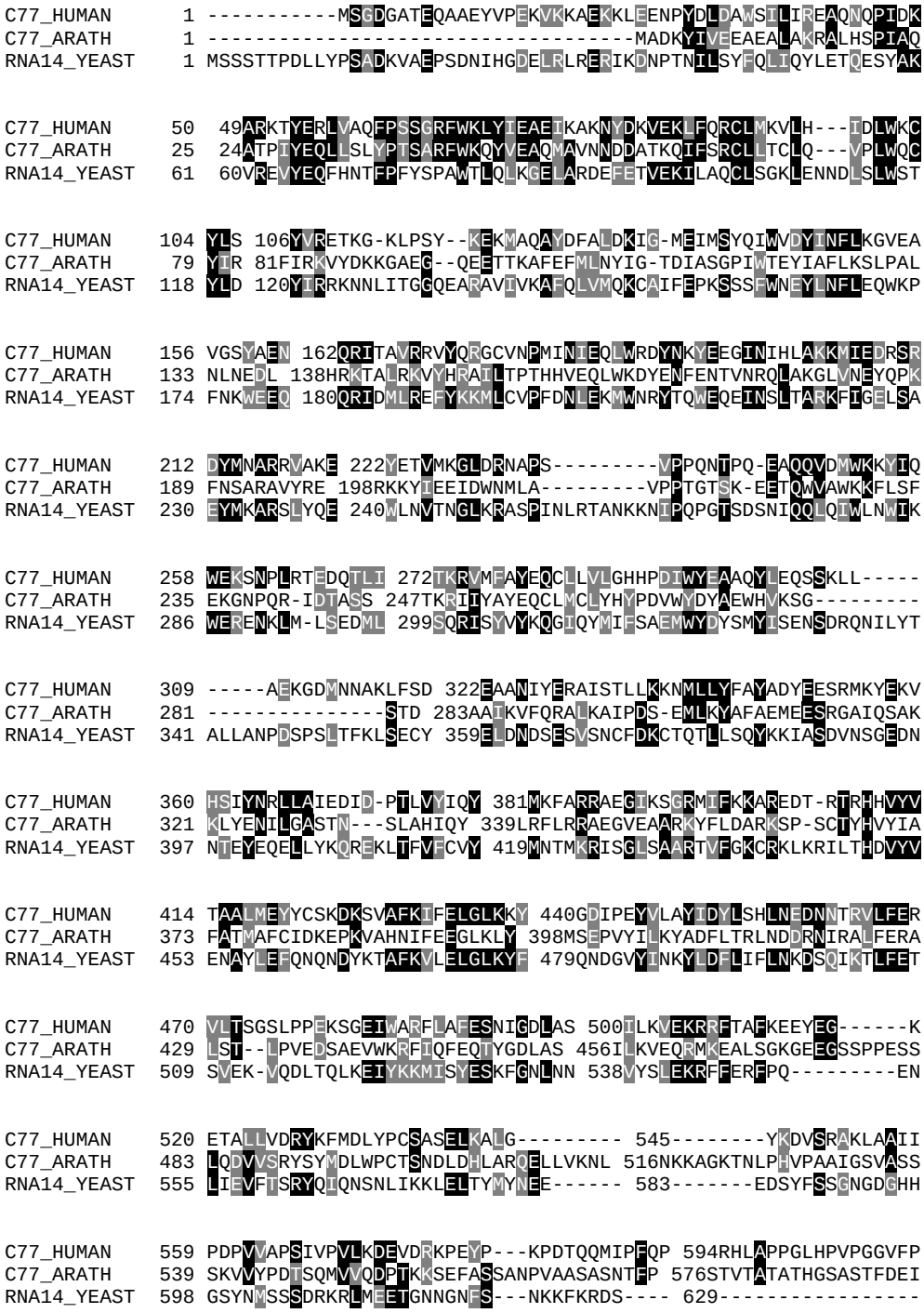


Figure 3.3: Schematic representation of CstF77 (A) from human and Rna14 (B) from yeast (*S.ceravisea*). The respective domains are shown in the legend (for details into the functional aspect of the domains refer to text).


C77_HUMAN 1 ----------MSGDGATEQAAEYVPEKVKKAEKKLEENPYDLDAWSILIREAQNQPIDK
C77_ARATH 1 ----------MADKYIVEEAEALAKRALHSPIAQ
RNA14_YEAST 1 MSSSTTPDLLYPSADKVAEPSDNIHGDELRLRERIKDNPTNILSYFQLIQYLETQESYAK
C77_HUMAN 50 49ARKTYERLVAQFPSSGRFWKLYIEAEIKAKNYDKVEKLFQRCLMKVLH---IDLWKC
C77_ARATH 25 24ATPIYEQLLSLYPTSARFWKQYVEAQMAVNNDDATKQIFSRCLLTCLQ---VPLWQC
RNA14_YEAST 61 60VREVYEQFHNTFPFYSPAWTLQLKGELARDEFETVEKILAQCLSGKLENNDLSLWST
C77_HUMAN 104 YLS 106YVRETKG-KLPSY--KEKMAQAYDFALDKIG-MEIMSYQIWVDYINFLKGVEA
C77_ARATH 79 YIR 81FIRKVYDKKGAEG--QEETTKAFEFMLNYIG-TDIASGPIWTEYIAFLKSLPAL
RNA14_YEAST 118 YLD 120YIRRKNNLITGGQEARAVIVKAFQLVMQKCAIFEPKSSFWNEYLNFLEQWKP
C77_HUMAN 156 VGSYAEN 162QRITAVRRVYQRGCVNPMINIEQLWRDYNKYEEGINIHLAKKMIEDRSR
C77_ARATH 133 NLNEDL 138HRKTALRKVYHRAILTPTHHVEQLWKDYENFENTVNRQLAKGLVNEYQPK
RNA14_YEAST 174 FNKWEEQ 180QRIDMLREFYKKMLCVPFDNLEKMWNRYTQWEQEINSLTARKFIGELSA
C77_HUMAN 212 DYMNARRVAKE 222YETVMKGLDRNAPS---------VPPQNTPQ-EAQQVDMWKKYIQ
C77_ARATH 189 FNSARAVYRE 198RKKYIEEIDWNMLA---------VPPTGTSK-EETQWVAWKKFLSF
RNA14_YEAST 230 EYMKARSLYQE 240WLNVTNGLKRASPINLRTANKKNIPQPGTSDSNIQQLQIWLNWIK
C77_HUMAN 258 WEKSNPLRTEDQTLI 272TKRVMFAYEQCLLVLGHHPDIWYEAAQYLEQSSKLL-----
C77_ARATH 235 EKGNPQR-IDTASS 247TKRIIYAYEQCLMCLYHYPDVWYDYAEWHVKSG---------
RNA14_YEAST 286 WERENKLM-LSEDML 299SQRISYVYKQGIQYMIFSAEMWYDYSMYISENSDRQNILYT
C77_HUMAN 309 -----AEKGDMNNAKLFSD 322EAANIYERAISTLLKKNMLLYFAYADYEESRMKYEKV
C77_ARATH 281 ---------------STD 283AAIKVFQRALKAIPDS-EMLKYAFAEMEESRGAIQSAK
RNA14_YEAST 341 ALLANPDSPSLTFKLSECY 359ELDNDSESVSNCFDKCTQTLLSQYKKIASDVNSGEDN
C77_HUMAN 360 HSIYNRLLAIEDID-PTLVYIQY 381MKFARRAEGIKSGRMIFKKAREDT-RTRHHVYV
C77_ARATH 321 KLYENILGASTN---SLAHIQY 339LRFLRRAEGVEAARKYFLDARKSP-SCIYHVYIA
RNA14_YEAST 397 NTEYEQELLYKQRELTFVFCVY 419MNTMKRISGLSAARTVFGKCRKLKRILTHDVYV
C77_HUMAN 414 TAALMEYYCSKDKSVAFKIFELGLKKY 440GDIPEYVLAYIDYLSHLNEDNNTRVLFER
C77_ARATH 373 FATMAFCIDKEPKVAHNIFEEGLKLY 398MSEPVYILKYADFLTRLNDDRNIRALFERA
RNA14_YEAST 453 ENAYLEFQNQNDYKTAFKVLELGLKYF 479QNDGVYINKYLDFLIFLNKDSQIKTLFET
C77_HUMAN 470 VLTSGSLPPEKSGEIMARFLAFESNIGDLAS 500ILKVEKRRFTAFKEEYEG------K
C77_ARATH 429 LST--LPVEDSAEVWKRFIQFEQTYGDLAS 456ILKVEQRMKEALSGKGEEGSSPPESS
RNA14_YEAST 509 SVEK-VQDLTQLKEIYKKMISYESKFGNLNN 538VYSLEKRFFERFPQ---------EN
C77_HUMAN 520 ETALLVDRYKFMDLYPCSASELKALG--------- 545--------YKDVSRAKLAAII
C77_ARATH 483 LQDVVSRYSYMDLWPCTSNDLDHLARQELLVKNL 516NKKAGKTNLPHVPAAIGSVASS
RNA14_YEAST 555 LIEVFTSRYQIQNSNLIKKLELTYMYNEE------ 583-------EDSYFSSCNGDGHH
C77_HUMAN 559 PDPVVAPSIVPVLKDEVDRKPEYP---KPDTQQMIPFQP 594RHLAPPGLHPVPGGVFP
C77_ARATH 539 SKVVYPDTSQMVVQDPTKKSEFASSANPVAASASNTFP 576STVTATATHGSASTFDEI
RNA14_YEAST 598 GSYNMSSSDRKRLMEETGNNGNFS---NKKFKRDS---- 629-----------------

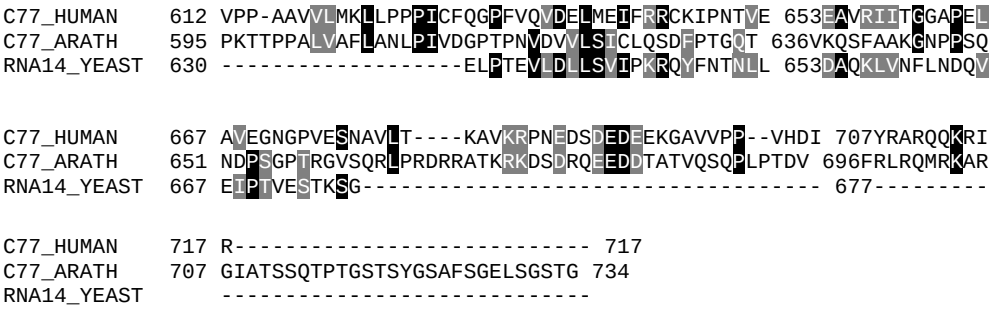


Figure 3.4: Amino acid sequence alignment and comparison of CstF77 and its homologs. The open reading frames encoded by human (C77_HUMAN, UniProtKB/TrEMBL Q12996), yeast (RNA14_YEAST, UniProtKB/TrEMBL P25298) and Arabidopsis (C77_ARATH, UniProtKB/TrEMBL Q8GUP1) are compared using ClustalW and formatted using BoxShade. Residues which are identical in all three polypeptides are shaded in black, with white uppercase lettering. Positions that are similar are shaded in grey, with white uppercase lettering.

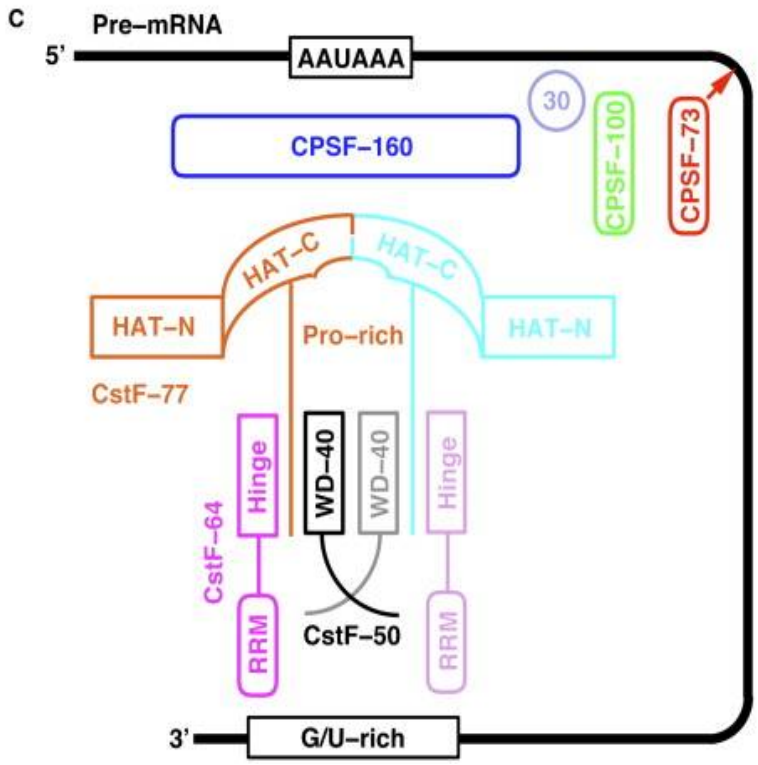
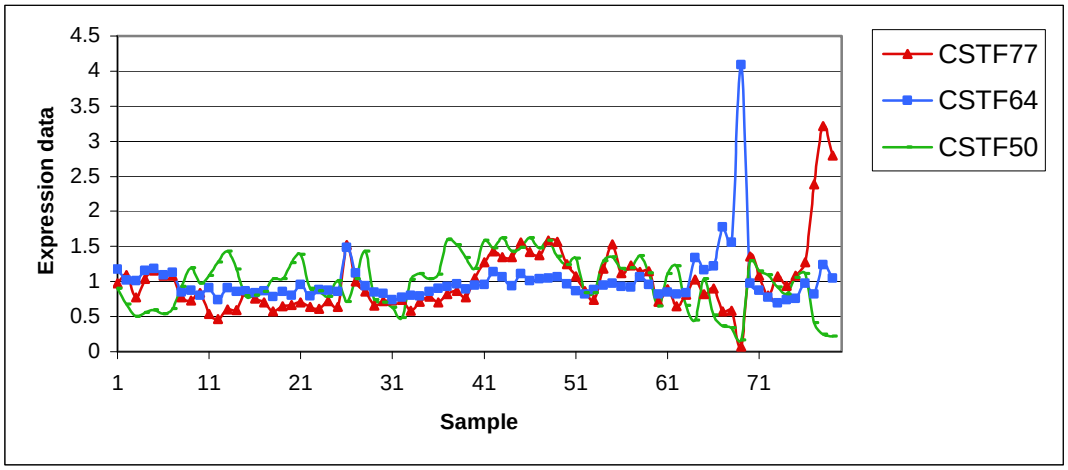


Figure 3.5: Schematic representation of the CstF subunit assembly with relation to other cleavage/polyadenylation factors and the pre-mRNA (detailed explanation to be found within text). CstF77 forms a dimer and bridges one unit each of CstF64 and CstF50 via its pro-rich domain and also makes contacts with the CPSF160 subunit via the HAT-C domain. (Adapted from Murthy and Manley, 1995; Takagaki and Manley, 2000; Bai *et al*, 2007).

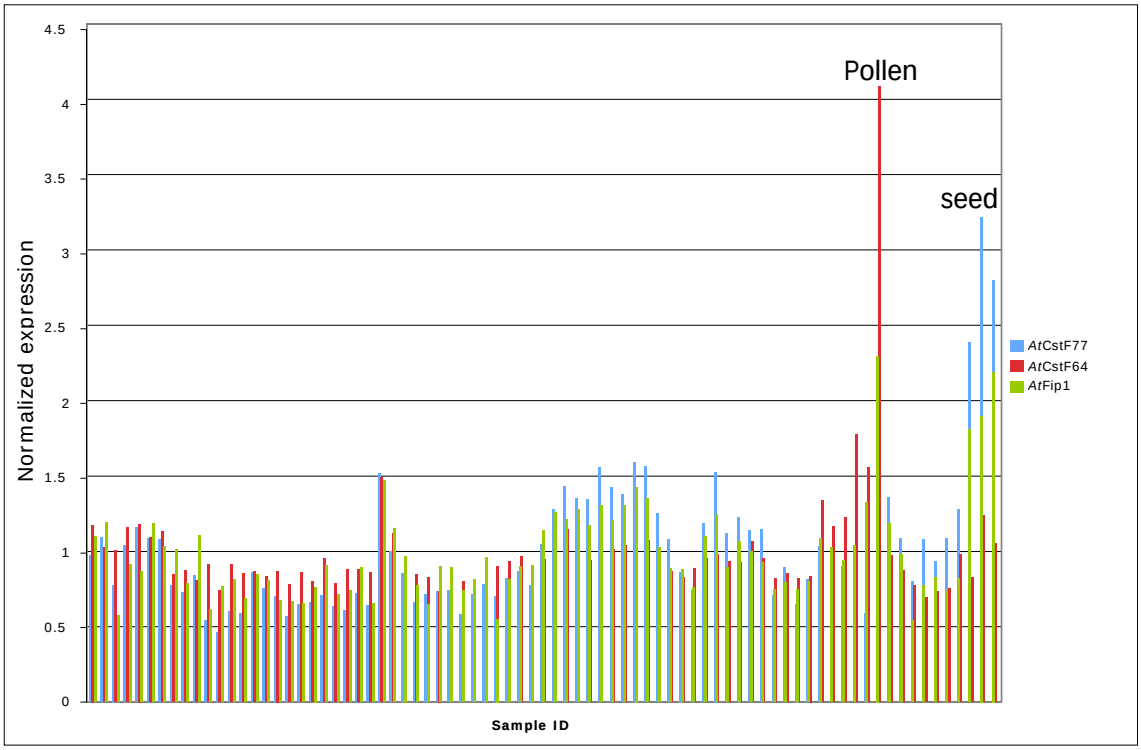
Table 3.1: DNA oligonucleotides used in this study.

Purpose used for	Sequence (5'→ 3')	Primer name
N-terminal end of CstF64	AGATCTATGGCTTCATCATCATCCCA	64NF
N-terminal end of CstF64	AGATCTTTAGACCTGAGGACCACCTAG	64NR
Hinge region of CstF64	AGATCTATGATAGCGGGTGCGCTA	64HF
Hinge region of CstF64	AGATCTTTAAATCCCAAGCATTACCTGC	64HR
C-terminal end of CstF64	AGATCTATGAATTTTGGCAAAAGAATAAACGAG	64CF
C-terminal end of CstF64	AGATCTTTATGAAGGCTGCATCATGTGG	64CR
Cloning of CstF64 for agroinfiltration	AGATCTATGGCTTCATCATCATCCCAACGAC	64F
Cloning of CstF64 for agroinfiltration	GTCGACTGAAGGCTGCATCATGTGGT	CstF645'R
Cloning of CstF77 for agroinfiltration	GCGGCCGCATGGCTGATAAGTACATCGTCGA	77F
Cloning of CstF77 for agroinfiltration	TCTAGAGCGCCAGTGCTACCAGAAAGCTCGC	CstF775'R

(A)



(B)



Normalized expression values were obtained from the NASC developmental series microarray set (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl), plotted as shown

(C)

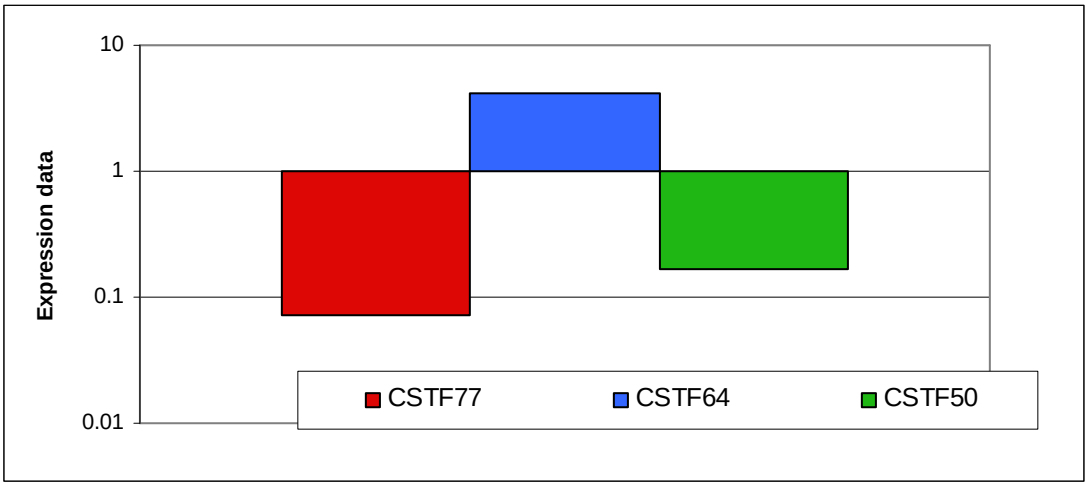


Figure 3.6: Expression data analysis for *Arabidopsis* CstF subunits during developmental stage. The data obtained from NASC was plotted as shown. The respective CstF subunits are color-coded in the legend. The expression values are represented along the Y-axes. The samples (including some WT and mutants) that are represented along the X-axes in (A) and (B): 1-7, root 7-21 days; 8-10, stem 7-21 days; 11-27, leaf 7-35 days; 28-38, whole plant 7-23 days; 39-49, shoot apex 7-21 days; 50-71, flowers and floral organs 21+ days; 72-79, 8 week seeds and siliques. The normalized expression in mature pollen exclusively is depicted in (C).

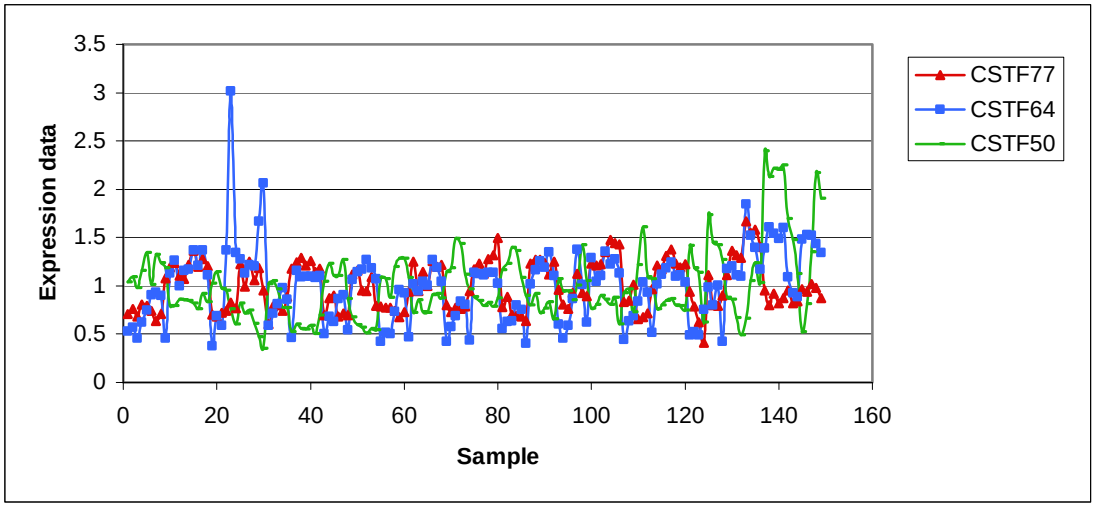


Figure 3.7: Expression data analysis for *Arabidopsis* CstF subunits during different abiotic stress conditions. The data obtained from NASC was plotted as shown. The respective CstF subunits are color-coded in the legend. The expression values are represented along the Y-axis. The samples that are represented along the X-axis: 1-18, control; 19-30, cold; 31-42, osmotic; 43-54, salt; 55-68, drought; 69-80, genotoxic; 81-92, oxidative; 93-106, UV-B; 107-120, wound; 121-136, heat; 137-141, cell culture control; 142-149, cell culture with heat.

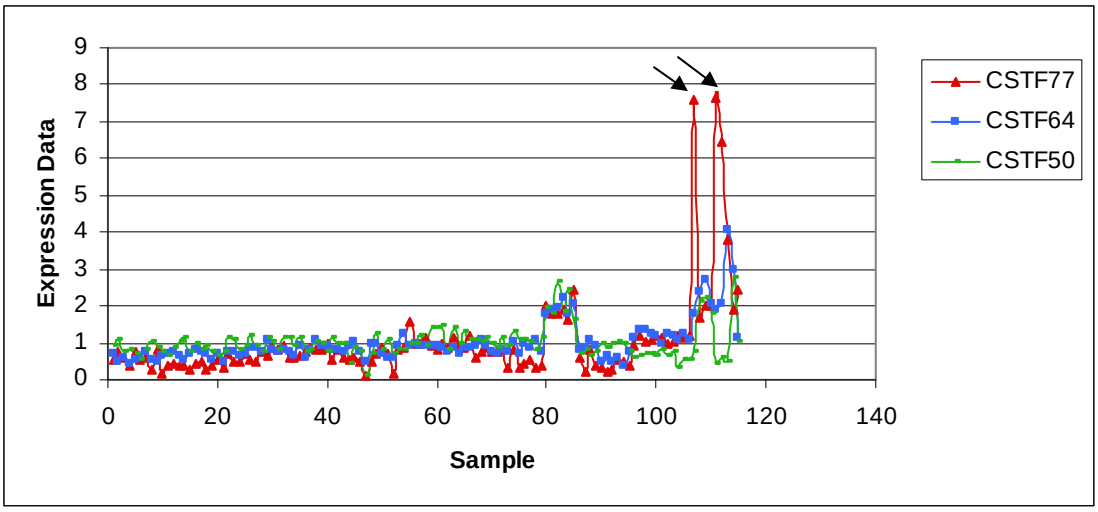


Figure 3.8: Expression data analysis for *Arabidopsis* CstF subunits during chemical and hormonal treatments. The data obtained from NASC was plotted as shown. The respective CstF subunits are color-coded in the legend. The expression values are represented along the Y-axis. The various samples according to the chemical and hormonal treatments are represented along the X-axis. The arrows (→) show the expression of CstF77 during imbibition of seeds.

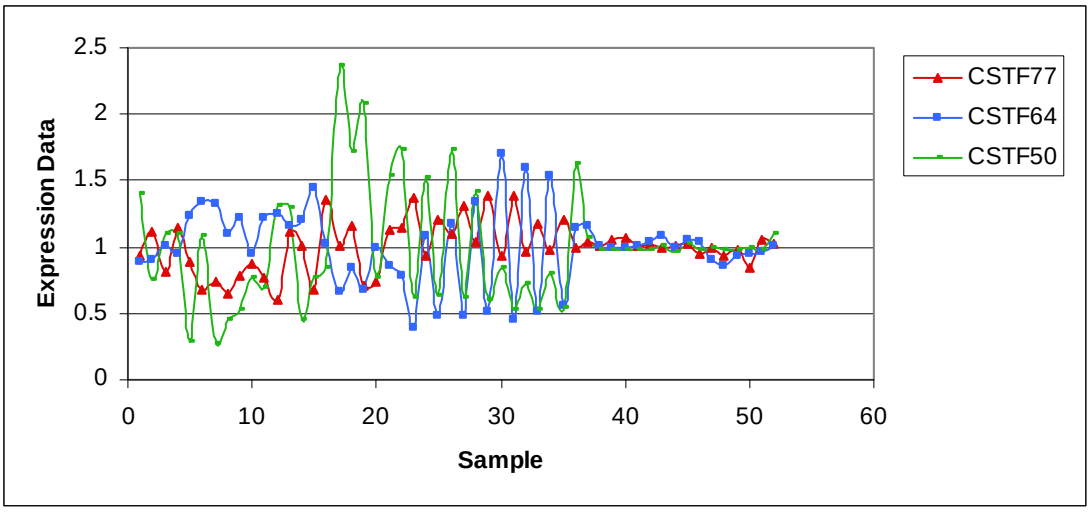
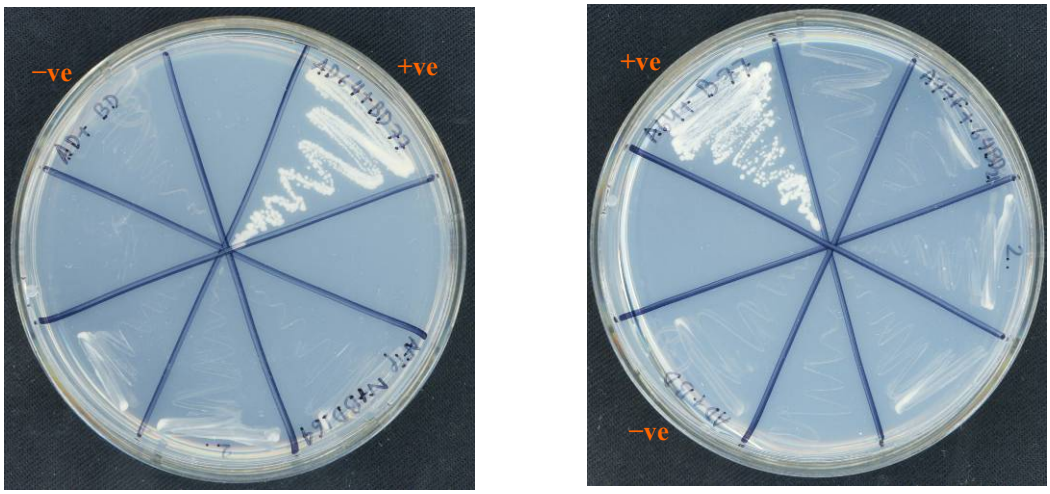


Figure 3.9: Expression data analysis for *Arabidopsis* CstF subunits during biotic stress and differential light conditions. The data obtained from NASC was plotted as shown. The respective CstF subunits are color-coded in the legend. The expression values are represented along the Y-axis. The samples that are represented along the X-axis: 1-16, control and *Pseudomonas syringae* infection; 17-22, control and *Phytophthora* infection; 23-36, control and other elicitors; 37-52, dark and different light conditions.

Table 3.2: Results from the yeast two-hybrid interactions between CstF64 (N, H, C) and other *Arabidopsis* polyadenylation factors. 'X'= negative interaction, '√'= positive interaction, '?'= interaction not yet performed. 'N'= N-terminal 150 amino acid residues, 'H'= Hinge domain 85 amino acid residues and 'C'= C-terminal 100 amino acid residues. Fip1(V)-N= N-terminal 161 amino acid residues, Fip1(V)-C= C-terminal 263 amino acid residues. (**Interaction with full length CstF64**, Adapted from Hunt *et al*, 2008)

Test No.	***Arabidopsis* gene**	**Factor**	**Interaction result with:**			**Interaction with full length CstF64**
			64N	64H	64C	
1	At5g58040	Fip1 (V)-N	?	?	X	X
2	At5g58040	Fip1(V)-C	?	?	X	√
3	At3g66652	Fip 1(III)	?	?	X	√
4	At1g17760	CstF77	?	?	X	√

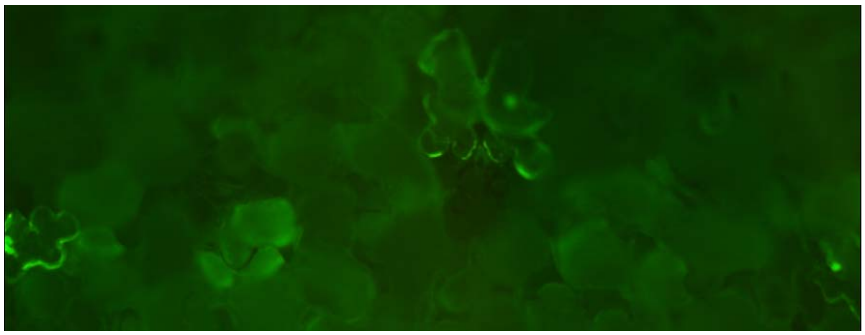


Fip1(V)-N + CstF64-C

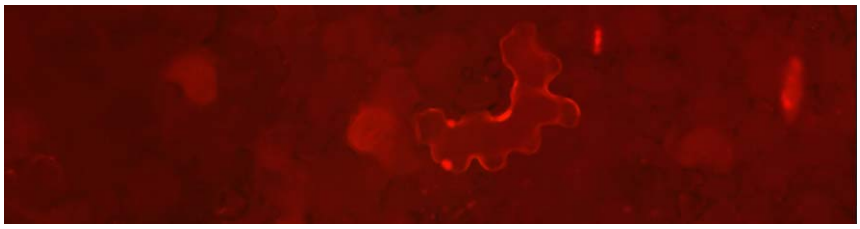
CstF77 + CstF64-C

Figure 3.10: Pairwise protein-protein interaction plates. The AD fusion protein and the BD fusion protein were co-transformed and dual transformants from the SC-LW plates (not shown) were grown on SC-HLW selection medium (as seen) with controls. The positive control (+ve) is the interaction between CstF64 + CstF77 with 200% colony growth, negative control (-ve) is the interaction between empty AD + BD vectors with 0% colony growth. Positive test interactions are scored as '√' and negative interactions are scored as 'X' with respect to the controls.

(A)



(B)



(C)

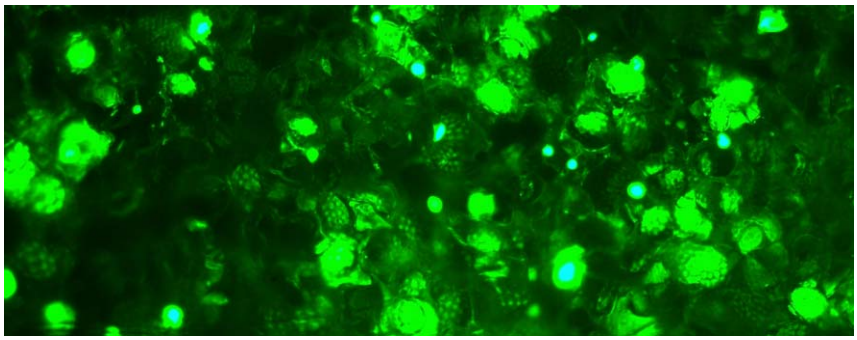
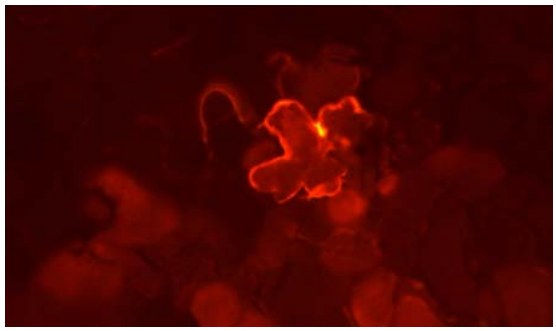


Figure 3.11: Epifluorescence micrographs showing Agrobacterium mediated infiltration of N.benthamiana leaves with control plasmids expressing autofluorescent proteins GFP and DsRed. The controls are: (A) GFP expression in whole cell from pGDG (100X magnification). (B) DsRed expression in whole cell from pGDR (100X magnification). (C) Nuclear GFP expression from Nu_{GFP} (100X magnification).

(A)



(B)

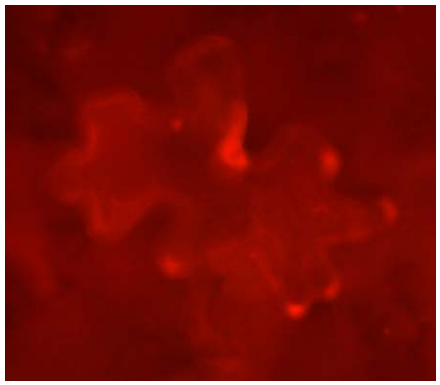
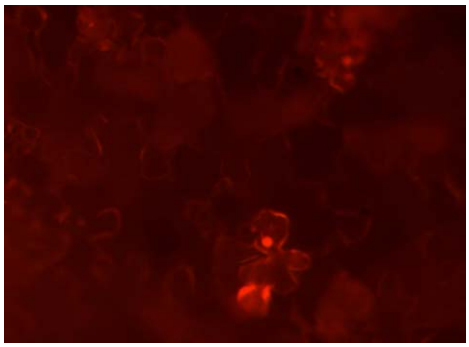
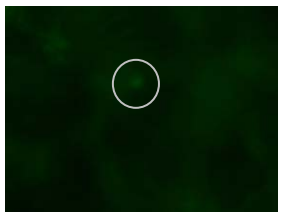
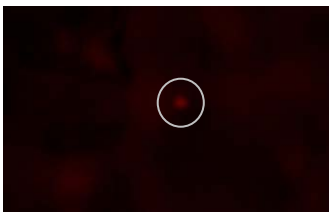


Figure 3.12: Epifluorescence micrographs showing Agrobacterium mediated infiltration of N.benthamiana leaves with 64_{DSR} plasmids expressing autofluorescent protein DsRed. (A) DsRed expression in whole cell from 64_{DSR} (100X magnification). (B) DsRed expression in whole cell from 64_{DSR} (200X magnification).

(A)







(B) (C) (D)

Figure 3.13: Epifluorescence micrographs showing Agrobacterium mediated infiltration of N.benthamiana leaves with 77_{DSR} plasmids expressing autofluorescent protein DsRed. (A) DsRed expression in nucleus of cell from 77_{DSR} (100X magnification) (B) Nuclear GFP expression from Nu_{GFP} (100X magnification) (C) Nuclear DsRed expression from 77_{DSR} (100X magnification) (D) superimposed frame of (B) and (C). All of the snapshots are at the same frame.

A1: NASC keys and data for figures 2.4 – 2.6 in Chapter 2

Microarray sample key - developmental stage

sample #	tissue cl	tissue	genotype	age	NASC sa	PAPS1	PAPS2	PAPS3	PAPS4
1	root	roots	Wt	7 days	ATGE_3	1.0134	0.6095	0.3756	1.0743
2	root	roots	Wt	17 days	ATGE_9	1.015	0.7526	0.3654	1.3867
3	root	root	Wt	15 days	ATGE_93	0.6668	0.4334	0.3925	1.0022
4	root	root	Wt	8 days	ATGE_94	1.0379	1.2001	0.3715	2.252
5	root	root	Wt	8 days	ATGE_95	0.9438	0.9815	0.4202	1.5959
6	root	root	Wt	21 days	ATGE_98	1.0742	1.0463	0.3799	1.9449
7	root	root	Wt	21 days	ATGE_99	0.9308	0.959	0.392	1.6232
8	stem	hypocotyl	Wt	7 days	ATGE_2	0.8686	0.8437	0.4119	0.9013
9	stem	1st node	Wt	21+ days	ATGE_28	1.1237	0.9983	0.3901	1.1831
10	stem	stem, 2nd	Wt	21+ days	ATGE_27	1.2069	1.7407	0.4504	1.7788
11	leaf	cotyledon:	Wt	7 days	ATGE_1	1.2172	0.834	0.5054	0.6504
12	leaf	leaves 1 +	Wt	7 days	ATGE_5	0.9753	0.4864	0.5509	0.5587
13	leaf	rosette le:	Wt	10 days	ATGE_10	0.6634	0.5674	0.4825	0.5833
14	leaf	rosette le:	gl1-T	10 days	ATGE_11	0.7655	0.7371	0.4801	0.6112
15	leaf	rosette le:	Wt	17 days	ATGE_12	1.8177	1.1156	0.4672	0.7036
16	leaf	rosette le:	Wt	17 days	ATGE_13	1.7425	1.1229	0.4589	0.6718
17	leaf	rosette le:	Wt	17 days	ATGE_14	1.6258	1.1201	0.5119	0.6037
18	leaf	rosette le:	Wt	17 days	ATGE_15	1.4995	0.9529	0.46	0.6269
19	leaf	rosette le:	Wt	17 days	ATGE_16	1.0384	1.0069	0.5134	0.5483
20	leaf	rosette le:	Wt	17 days	ATGE_17	0.787	0.8328	0.4789	0.5355
21	leaf	rosette le:	gl1-T	17 days	ATGE_18	0.7534	0.9798	0.4785	0.6903
22	leaf	leaf 7, pet	Wt	17 days	ATGE_19	0.9997	0.936	0.4652	0.5839
23	leaf	leaf 7, prc	Wt	17 days	ATGE_20	1.2746	1.1624	0.5508	0.585
24	leaf	leaf 7, dis	Wt	17 days	ATGE_21	1.9226	1.3592	0.4928	0.6579
25	leaf	leaf	Wt	15 days	ATGE_91	0.6827	0.9428	0.5731	0.7812
26	leaf	senescing	Wt	35 days	ATGE_25	1.4141	2.2325	0.4002	1.9512
27	leaf	cauline le:	Wt	21+ days	ATGE_26	1.6881	1.7299	0.4712	1.1133
28	whole pla	seedling,	Wt	7 days	ATGE_7	0.9601	0.9348	0.4433	0.9245
29	whole pla	seedling,	Wt	8 days	ATGE_96	1.2678	1.5137	0.4291	0.7952
30	whole pla	seedling,	Wt	8 days	ATGE_97	0.8415	0.9607	0.437	0.7609
31	whole pla	seedling,	Wt	21 days	ATGE_10(	1.1608	1.2561	0.4571	0.9728
32	whole pla	seedling,	Wt	21 days	ATGE_101	1.0088	1.2567	0.4734	1.001

33	whole plai	developm	Wt	21 days	ATGE_22	1.1144	1.1732	0.4805	0.6822
34	whole plai	as above	Wt	22 days	ATGE_23	0.8891	1.0913	0.4739	0.799
35	whole plai	as above	Wt	23 days	ATGE_24	0.8494	1.3004	0.5049	0.8616
36	whole plai	vegetative	Wt	7 days	ATGE_87	0.633	0.7698	0.472	0.7137
37	whole plai	vegetative	Wt	14 days	ATGE_89	0.6633	0.9926	0.4363	0.7836
38	whole plai	vegetative	Wt	21 days	ATGE_90	0.6049	0.8401	0.4703	0.6475
39	apex	shoot ape	Wt	7 days	ATGE_4	0.8575	0.5114	0.4611	0.5872
40	apex	shoot ape	Wt	7 days	ATGE_6	0.8817	0.4186	0.4324	0.6181
41	apex	shoot ape	Wt	14 days	ATGE_8	0.9004	0.5408	0.3999	0.7514
42	apex	shoot ape	Wt	21 days	ATGE_29	0.8151	0.6744	0.5257	0.7959
43	apex	shoot ape	clv3-7	21+ days	ATGE_46	0.9531	0.9932	0.3763	1.0128
44	apex	shoot ape	lfy-12	21+ days	ATGE_47	0.8961	0.7957	0.4672	0.8866
45	apex	shoot ape	ap1-15	21+ days	ATGE_48	1.0332	0.8941	0.446	1.0475
46	apex	shoot ape	ap2-6	21+ days	ATGE_49	1.0125	0.8289	0.5121	0.8955
47	apex	shoot ape	ufo-1	21+ days	ATGE_52	0.9846	0.8894	0.4442	0.8392
48	apex	shoot ape	ap3-6	21+ days	ATGE_50	0.9788	0.7366	0.43	0.789
49	apex	shoot ape	ag-12	21+ days	ATGE_51	0.9128	0.7724	0.4954	0.8364
50	flowers	flowers st	Wt	21+ days	ATGE_31	0.7776	0.8701	1.4071	0.8514
51	flowers	flowers st	Wt	21+ days	ATGE_32	0.8151	0.9464	0.4887	0.9298
52	flowers	flowers st	Wt	21+ days	ATGE_33	0.7582	0.8513	0.4585	0.8977
53	flowers	flower sta	clv3-7	21+ days	ATGE_53	0.6779	1.0331	0.4396	1.103
54	flowers	flower sta	lfy-12	21+ days	ATGE_54	0.9507	1.0117	0.3966	0.9232
55	flowers	flower sta	ap1-15	21+ days	ATGE_55	0.8753	1.5487	0.6009	1.3961
56	flowers	flower sta	ap2-6	21+ days	ATGE_56	0.8247	0.8707	0.4628	0.9532
57	flowers	flower sta	ufo-1	21+ days	ATGE_59	0.8632	1.1895	0.4157	1.0032
58	flowers	flower sta	ap3-6	21+ days	ATGE_57	1.016	1.1181	0.3976	1.038
59	flowers	flower sta	ag-12	21+ days	ATGE_58	0.9465	0.9787	0.4152	0.9808
60	flowers	flowers st	Wt	21+ days	ATGE_39	0.8611	1.2084	0.4921	1.1704
61	flowers	flower	Wt	28 days	ATGE_92	0.6934	0.8361	0.6875	0.8108
62	floral orga	flowers st	Wt	21+ days	ATGE_40	0.9949	0.8407	0.4998	0.7248
63	floral orga	flowers st	Wt	21+ days	ATGE_34	1.1773	1.2916	0.4485	0.8491
64	floral orga	flowers st	Wt	21+ days	ATGE_41	1.3845	1.7621	0.4294	1.1765
65	floral orga	flowers st	Wt	21+ days	ATGE_35	0.9602	1.1246	0.4116	1.1777
66	floral orga	flowers st	Wt	21+ days	ATGE_42	1.6198	1.9815	0.4952	1.7812

67	floral orga	flowers st	Wt	21+ days	ATGE_36	0.9165	1.7558	3.2388	1.2974
68	floral orga	flowers st	Wt	21+ days	ATGE_43	1.2391	1.4903	3.5525	1.6657
69	floral orga	mature pc	Wt	6 wk	ATGE_73	0.3535	1.5462	35.435	2.1979
70	floral orga	flowers st	Wt	21+ days	ATGE_37	0.9528	1.0285	0.4156	0.9343
71	floral orga	flowers st	Wt	21+ days	ATGE_45	0.9995	0.9622	0.416	0.8869
72	seeds	siliques, v	Wt	8 wk	ATGE_76	0.7266	0.6769	0.4509	0.5947
73	seeds	siliques, v	Wt	8 wk	ATGE_77	0.8189	0.671	0.3698	1.0576
74	seeds	siliques, v	Wt	8 wk	ATGE_78	0.8572	0.4476	0.3966	1.1393
75	seeds	seeds, sta	Wt	8 wk	ATGE_79	0.7838	0.3799	0.4594	1.0246
76	seeds	seeds, sta	Wt	8 wk	ATGE_81	0.7446	0.4968	0.4923	1.1746
77	seeds	seeds, sta	Wt	8 wk	ATGE_82	1.0358	0.688	0.6539	1.2304
78	seeds	seeds, sta	Wt	8 wk	ATGE_83	1.1032	0.7393	0.9296	1.4786
79	seeds	seeds, sta	Wt	8 wk	ATGE_84	1.3089	0.8294	0.8887	1.355

Microarray sample key - abiotic stress

sample #	tissue cl	tissue	genotype	age	NASC sa	stress	time	treatmer	PAPS1	PAPS2	PAPS3	PAPS4
1	control, a	seedling,	not applic	not applic	1	untreated	0 h	Control: N	0.79	0.5498	1.0473	0.4645
2	control, a	seedling,	not applic	not applic	71	mock	0.25 h	Control: T	1.006	0.6489	1.16	0.4499
3	control, a	seedling,	not applic	not applic	11	mock	0.5 h	Control: T	1.0195	0.5018	1.1707	0.5455
4	control, a	seedling,	not applic	not applic	21	mock	1 h	Control: T	1.172	0.6685	1.1539	0.618
5	control, a	seedling,	not applic	not applic	31	mock	3 h	Control: T	1.0684	1.118	1.134	0.8334
6	control, a	seedling,	not applic	not applic	81	mock	4 h	Control: T	1.2003	1.0383	1.1732	0.6004
7	control, a	seedling,	not applic	not applic	41	mock	6 h	Control: T	0.928	1.1859	1.1973	0.7651
8	control, a	seedling,	not applic	not applic	51	mock	12 h	Control: T	0.8714	1.1606	1.1301	0.7697
9	control, a	seedling,	not applic	not applic	61	mock	24 h	Control: T	0.8188	0.4258	1.1744	0.5917
10	control, rc	root	not applic	not applic	2	untreated	0 h	Control: N	0.8921	0.4679	0.7571	1.295
11	control, rc	root	not applic	not applic	72	mock	0.25 h	Control: T	0.9545	0.8963	0.6774	1.095
12	control, rc	root	not applic	not applic	12	mock	0.5 h	Control: T	0.8965	0.6475	0.7918	1.586
13	control, rc	root	not applic	not applic	22	mock	1 h	Control: T	1.0649	0.7006	0.7982	1.2623
14	control, rc	root	not applic	not applic	32	mock	3 h	Control: T	0.9234	0.8442	0.8284	1.4721
15	control, rc	root	not applic	not applic	82	mock	4 h	Control: T	1.0344	1.0064	0.7304	1.1513
16	control, rc	root	not applic	not applic	42	mock	6 h	Control: T	0.9925	0.9177	0.7968	1.464
17	control, rc	root	not applic	not applic	52	mock	12 h	Control: T	0.9565	0.9545	0.787	1.2978
18	control, rc	root	not applic	not applic	62	mock	24 h	Control: T	0.87	0.849	0.7709	1.3898
19	cold, aeria	seedling,	not applic	not applic	111	cold	0.5 h	Cold stres	0.7984	0.5187	1.1541	0.6334
20	cold, aeria	seedling,	not applic	not applic	121	cold	1 h	Cold stres	1.0551	0.4736	1.438	0.4958
21	cold, aeria	seedling,	not applic	not applic	131	cold	3 h	Cold stres	0.907	0.677	1.1487	0.4827
22	cold, aeria	seedling,	not applic	not applic	141	cold	6 h	Cold stres	0.9971	1.2997	1.2462	0.589
23	cold, aeria	seedling,	not applic	not applic	151	cold	12 h	Cold stres	0.812	1.0969	1.244	0.6976
24	cold, aeria	seedling,	not applic	not applic	161	cold	24 h	Cold stres	0.5492	0.8184	1.3567	0.6481
25	cold, root	root	not applic	not applic	112	cold	0.5 h	Cold stres	0.9779	0.8579	0.8153	1.2034
26	cold, root	root	not applic	not applic	122	cold	1 h	Cold stres	0.8433	0.7606	0.7969	1.1063
27	cold, root	root	not applic	not applic	132	cold	3 h	Cold stres	1.0257	0.8276	0.8474	1.1646
28	cold, root	root	not applic	not applic	142	cold	6 h	Cold stres	0.9583	0.7772	0.8042	1.048
29	cold, root	root	not applic	not applic	152	cold	12 h	Cold stres	0.8615	0.8981	0.8712	1.3883
30	cold, root	root	not applic	not applic	162	cold	24 h	Cold stres	0.7532	0.8476	0.8911	1.0653
31	osmotic, a	seedling,	not applic	not applic	211	osmotic	0.5 h	Osmotic s	0.8976	0.5916	1.1658	0.372
32	osmotic, a	seedling,	not applic	not applic	221	osmotic	1 h	Osmotic s	1.0157	0.8315	1.1041	0.4803

33	osmotic, a	seedling,	not applic	not applic	231	osmotic	3 h	Osmotic s	1.1516	1.9016	1.0409	0.734
34	osmotic, a	seedling,	not applic	not applic	241	osmotic	6 h	Osmotic s	0.9443	1.4991	1.1248	0.932
35	osmotic, a	seedling,	not applic	not applic	251	osmotic	12 h	Osmotic s	0.8545	1.1933	1.0096	0.7119
36	osmotic, a	seedling,	not applic	not applic	261	osmotic	24 h	Osmotic s	1.0173	0.6915	1.0456	0.7998
37	osmotic, r	root	not applic	not applic	212	osmotic	0.5 h	Osmotic s	0.9272	0.6725	0.7759	0.997
38	osmotic, r	root	not applic	not applic	222	osmotic	1 h	Osmotic s	1.0559	0.8217	0.7912	0.8851
39	osmotic, r	root	not applic	not applic	232	osmotic	3 h	Osmotic s	0.8508	1.0239	0.8155	1.231
40	osmotic, r	root	not applic	not applic	242	osmotic	6 h	Osmotic s	0.8232	1.0397	0.8199	1.1696
41	osmotic, r	root	not applic	not applic	252	osmotic	12 h	Osmotic s	0.9402	0.9846	0.8617	1.2636
42	osmotic, r	root	not applic	not applic	262	osmotic	24 h	Osmotic s	1.0804	1.0339	0.7811	1.3974
43	salt, aeria	seedling,	not applic	not applic	311	salt	0.5 h	Salt stres:	1.0841	0.6282	1.2626	0.522
44	salt, aeria	seedling,	not applic	not applic	321	salt	1 h	Salt stres:	1.1214	0.7706	1.3288	0.7383
45	salt, aeria	seedling,	not applic	not applic	331	salt	3 h	Salt stres:	1.0823	1.0612	1.1053	0.8643
46	salt, aeria	seedling,	not applic	not applic	341	salt	6 h	Salt stres:	0.9692	1.1137	1.0505	0.7907
47	salt, aeria	seedling,	not applic	not applic	351	salt	12 h	Salt stres:	0.9533	1.0789	0.9858	0.6477
48	salt, aeria	seedling,	not applic	not applic	361	salt	24 h	Salt stres:	0.836	0.6087	1.0474	0.4761
49	root, aeria	root	not applic	not applic	312	salt	0.5 h	Salt stres:	0.9112	0.7891	0.804	1.3608
50	root, aeria	root	not applic	not applic	322	salt	1 h	Salt stres:	0.9817	0.8622	0.8066	1.1126
51	root, aeria	root	not applic	not applic	332	salt	3 h	Salt stres:	0.9077	0.6599	0.7244	0.9667
52	root, aeria	root	not applic	not applic	342	salt	6 h	Salt stres:	0.9793	0.5921	0.991	0.7809
53	root, aeria	root	not applic	not applic	352	salt	12 h	Salt stres:	1.0778	0.9057	1.0318	0.9442
54	root, aeria	root	not applic	not applic	362	salt	24 h	Salt stres:	0.857	1.2506	1.0591	0.835
55	drought, a	seedling,	not applic	not applic	471	drought	0.25 h	Drought s	0.9621	0.5654	1.2366	0.4783
56	drought, a	seedling,	not applic	not applic	411	drought	0.5 h	Drought s	0.9554	0.8283	1.0238	0.4985
57	drought, a	seedling,	not applic	not applic	421	drought	1 h	Drought s	0.9762	0.7666	1.0297	0.5061
58	drought, a	seedling,	not applic	not applic	431	drought	3 h	Drought s	0.8848	1.3626	0.95	0.9246
59	drought, a	seedling,	not applic	not applic	441	drought	6 h	Drought s	1.0035	1.8031	1.1999	1.0099
60	drought, a	seedling,	not applic	not applic	451	drought	12 h	Drought s	0.9148	1.4005	1.0219	0.7613
61	drought, a	seedling,	not applic	not applic	461	drought	24 h	Drought s	0.9283	0.627	1.2688	0.6156
62	drought, r	root	not applic	not applic	472	drought	0.25 h	Drought s	0.7962	0.9616	0.7289	1.4957
63	drought, r	root	not applic	not applic	412	drought	0.5 h	Drought s	1.3296	0.8899	0.7094	1.2344
64	drought, r	root	not applic	not applic	422	drought	1 h	Drought s	1.1765	1.0366	0.7027	1.1992
65	drought, r	root	not applic	not applic	432	drought	3 h	Drought s	0.7915	1.19	0.741	1.4764
66	drought, r	root	not applic	not applic	442	drought	6 h	Drought s	1.0557	1.1907	0.8047	1.3602

67	drought, r	root	not applic	not applic	452	drought	12 h	Drought s	0.9593	0.9632	0.7899	1.0885
68	drought, r	root	not applic	not applic	462	drought	24 h	Drought s	0.8104	0.8554	0.7736	1.3595
69	genotoxic	seedling,	not applic	not applic	511	genotoxic	0.5 h	Genotoxic	1.1711	0.9271	1.1903	0.6538
70	genotoxic	seedling,	not applic	not applic	521	genotoxic	1 h	Genotoxic	1.1271	1.0486	1.0476	0.7794
71	genotoxic	seedling,	not applic	not applic	531	genotoxic	3 h	Genotoxic	0.9783	1.499	1.0896	0.91
72	genotoxic	seedling,	not applic	not applic	541	genotoxic	6 h	Genotoxic	0.9248	1.9715	1.0126	0.9906
73	genotoxic	seedling,	not applic	not applic	551	genotoxic	12 h	Genotoxic	0.8321	1.9272	0.9485	0.892
74	genotoxic	seedling,	not applic	not applic	561	genotoxic	24 h	Genotoxic	0.931	0.7056	1.058	0.5125
75	genotoxic	root	not applic	not applic	512	genotoxic	0.5 h	Genotoxic	1.0335	0.911	0.6798	1.3585
76	genotoxic	root	not applic	not applic	522	genotoxic	1 h	Genotoxic	1.0113	0.8172	0.6987	1.4516
77	genotoxic	root	not applic	not applic	532	genotoxic	3 h	Genotoxic	0.8558	1.0916	0.786	1.427
78	genotoxic	root	not applic	not applic	542	genotoxic	6 h	Genotoxic	0.9649	1.1068	0.7158	1.355
79	genotoxic	root	not applic	not applic	552	genotoxic	12 h	Genotoxic	1.2063	1.0334	0.6902	1.1201
80	genotoxic	root	not applic	not applic	562	genotoxic	24 h	Genotoxic	1.1927	0.9868	0.7415	1.0731
81	oxidative,	seedling,	not applic	not applic	611	oxidative	0.5 h	Oxidative	1.3916	0.6944	1.0013	0.5441
82	oxidative,	seedling,	not applic	not applic	621	oxidative	1 h	Oxidative	1.1092	0.7738	1.0559	0.6129
83	oxidative,	seedling,	not applic	not applic	631	oxidative	3 h	Oxidative	0.9483	1.4298	1.3336	0.954
84	oxidative,	seedling,	not applic	not applic	641	oxidative	6 h	Oxidative	0.9146	1.8182	1.0506	1.0127
85	oxidative,	seedling,	not applic	not applic	651	oxidative	12 h	Oxidative	0.7007	1.6908	1.0603	0.8683
86	oxidative,	seedling,	not applic	not applic	661	oxidative	24 h	Oxidative	0.7151	0.9963	0.9626	0.6045
87	oxidative,	root	not applic	not applic	612	oxidative	0.5 h	Oxidative	0.9412	1.0289	0.7215	1.2254
88	oxidative,	root	not applic	not applic	622	oxidative	1 h	Oxidative	0.9366	0.8436	0.7161	1.2283
89	oxidative,	root	not applic	not applic	632	oxidative	3 h	Oxidative	0.9848	1.3065	0.8399	1.7536
90	oxidative,	root	not applic	not applic	642	oxidative	6 h	Oxidative	0.9944	1.2983	0.6797	1.5574
91	oxidative,	root	not applic	not applic	652	oxidative	12 h	Oxidative	1.0655	1.118	0.7284	1.2264
92	oxidative,	root	not applic	not applic	662	oxidative	24 h	Oxidative	0.9702	1.2563	0.7963	1.4949
93	UV-B, aer	seedling,	not applic	not applic	771	UV-B	0.25 h	UV-B stre	0.9415	0.521	1.0949	0.4179
94	UV-B, aer	seedling,	not applic	not applic	711	UV-B	0.5 h	UV-B stre	0.8974	0.6728	0.9397	0.4185
95	UV-B, aer	seedling,	not applic	not applic	721	UV-B	1 h	UV-B stre	1.1395	0.7484	1.2341	0.5901
96	UV-B, aer	seedling,	not applic	not applic	731	UV-B	3 h	UV-B stre	1.0044	2.6923	1.0363	0.7901
97	UV-B, aer	seedling,	not applic	not applic	741	UV-B	6 h	UV-B stre	0.9267	1.8987	0.9109	0.7914
98	UV-B, aer	seedling,	not applic	not applic	751	UV-B	12 h	UV-B stre	0.9417	0.943	0.9435	0.7341
99	UV-B, aer	seedling,	not applic	not applic	761	UV-B	24 h	UV-B stre	0.9629	0.6294	1.0804	0.5231
100	UV-B, roo	root	not applic	not applic	772	UV-B	0.25 h	UV-B stre	0.9356	0.8375	0.7339	1.3516

101	UV-B, roo	root	not applic	not applic	712	UV-B	0.5 h	UV-B stre	0.9063	0.728	0.7708	1.1278
102	UV-B, roo	root	not applic	not applic	722	UV-B	1 h	UV-B stre	0.989	0.8192	0.7445	1.4012
103	UV-B, roo	root	not applic	not applic	732	UV-B	3 h	UV-B stre	1.1348	1.2468	0.6929	1.5943
104	UV-B, roo	root	not applic	not applic	742	UV-B	6 h	UV-B stre	1.2263	1.1037	0.7044	1.8248
105	UV-B, roo	root	not applic	not applic	752	UV-B	12 h	UV-B stre	1.1413	1.0412	0.7218	1.5063
106	UV-B, roo	root	not applic	not applic	762	UV-B	24 h	UV-B stre	1.0394	1.0928	0.8888	1.4839
107	wounding	seedling,	not applic	not applic	871	wounding	0.25 h	Wound sti	1.6652	0.5257	1.1078	0.4488
108	wounding	seedling,	not applic	not applic	811	wounding	0.5 h	Wound sti	3.0894	0.675	1.124	0.4903
109	wounding	seedling,	not applic	not applic	821	wounding	1 h	Wound sti	2.9794	0.767	1.1707	0.5544
110	wounding	seedling,	not applic	not applic	831	wounding	3 h	Wound sti	0.9326	0.9363	1.0286	0.7694
111	wounding	seedling,	not applic	not applic	841	wounding	6 h	Wound sti	0.8785	1.3528	1.0257	0.7713
112	wounding	seedling,	not applic	not applic	851	wounding	12 h	Wound sti	0.7696	1.1622	1.0153	0.6943
113	wounding	seedling,	not applic	not applic	861	wounding	24 h	Wound sti	1.1641	0.5782	1.2627	0.5574
114	wounding	root	not applic	not applic	872	wounding	0.25 h	Wound sti	0.8323	1.1014	0.7127	1.6017
115	wounding	root	not applic	not applic	812	wounding	0.5 h	Wound sti	0.9407	0.7739	0.7688	1.3888
116	wounding	root	not applic	not applic	822	wounding	1 h	Wound sti	0.9546	0.7624	0.7772	1.3233
117	wounding	root	not applic	not applic	832	wounding	3 h	Wound sti	0.9354	0.9255	0.7207	1.732
118	wounding	root	not applic	not applic	842	wounding	6 h	Wound sti	0.9286	1.22	0.7593	1.7508
119	wounding	root	not applic	not applic	852	wounding	12 h	Wound sti	0.8188	1.3453	0.7428	1.6219
120	wounding	root	not applic	not applic	862	wounding	24 h	Wound sti	0.9383	1.0032	0.7775	1.2464
121	heat, aeri	seedling,	not applic	not applic	971	heat	0.25 h	0.25h of 3	0.9944	0.4375	1.1889	0.5718
122	heat, aeri	seedling,	not applic	not applic	911	heat	0.5 h	0.5h of 3E	0.9104	0.7396	1.0783	0.6607
123	heat, aeri	seedling,	not applic	not applic	921	heat	1 h	1.0h of 3E	0.8499	0.5667	1.3347	0.4473
124	heat, aeri	seedling,	not applic	not applic	931	heat	3 h	3.0h of 3E	0.6435	0.7657	1.2155	0.6885
125	heat, aeri	seedling,	not applic	not applic	981	heat	4 (3hr + :	3.0h of 3E	0.7562	1.2964	1.0559	1.1899
126	heat, aeri	seedling,	not applic	not applic	941	heat	6 (3hr + :	3.0h of 3E	0.8344	1.3954	0.9624	0.9139
127	heat, aeri	seedling,	not applic	not applic	951	heat	12 (3hr +	3.0h of 3E	0.9578	1.1156	0.9637	0.807
128	heat, aeri	seedling,	not applic	not applic	961	heat	24 (3hr +	3.0h of 3E	0.8563	0.3984	1.1578	0.6037
129	heat, root	root	not applic	not applic	972	heat	0.25 h	0.25h of 3	0.8913	0.945	0.7258	1.7841
130	heat, root	root	not applic	not applic	912	heat	0.5 h	0.5h of 3E	0.9833	0.9923	0.7444	1.1672
131	heat, root	root	not applic	not applic	922	heat	1 h	1.0h of 3E	0.8054	0.8639	0.7805	1.1287
132	heat, root	root	not applic	not applic	932	heat	3 h	3.0h of 3E	0.4612	1.6284	0.8596	1.799
133	heat, root	root	not applic	not applic	982	heat	4 (3hr + :	3.0h of 3E	0.8892	1.9283	0.8426	3.0001
134	heat, root	root	not applic	not applic	942	heat	6 (3hr + :	3.0h of 3E	1.2133	1.0825	0.7542	1.355

135	heat, root	root	not applic	not applic	952	heat	12 (3hr +	3.0h of 38	1.0204	0.8128	0.6978	1.5683
136	heat, root	root	not applic	not applic	962	heat	24 (3hr +	3.0h of 38	0.9268	0.7399	0.8181	1.6062
137	ctrl, cell c	cell cultur	not applic	not applic	C0	control	0 h	incubated	1.2132	1.1904	1.5103	0.9419
138	ctrl, cell c	cell cultur	not applic	not applic	C1	control	3 h	3.0h at 25	1.0735	1.144	1.1761	0.9231
139	ctrl, cell c	cell cultur	not applic	not applic	C2	control	6 h	6.0h at 25	1.2649	1.1643	1.4924	0.9018
140	ctrl, cell c	cell cultur	not applic	not applic	C3	control	12 h	12.0h at 2	1.0695	1.3065	1.3502	1.0004
141	ctrl, cell c	cell cultur	not applic	not applic	C4	control	24 h	24.0h at 2	1.117	1.3131	1.1496	1.0495
142	heat, cell	cell cultur	not applic	not applic	C5	heat	0.25 h	0.25h of 3	1.1057	1.2071	1.8947	0.8236
143	heat, cell	cell cultur	not applic	not applic	C6	heat	0.5 h	0.5h of 38	0.91473	1.19817	1.17776	0.6794
144	heat, cell	cell cultur	not applic	not applic	C7	heat	1 h	1.0h of 38	0.91	1.20788	1.18705	0.7297
145	heat, cell	cell cultur	not applic	not applic	C8	heat	3 h	3.0h of 38	0.90528	1.21758	1.19634	0.7681
146	heat, cell	cell cultur	not applic	not applic	C9	heat	4 (3hr +	3.0h of 38	0.90056	1.22729	1.20563	0.8992
147	heat, cell	cell cultur	not applic	not applic	C10	heat	6 (3hr +	3.0h of 38	0.89584	1.237	1.21492	0.827
148	heat, cell	cell cultur	not applic	not applic	C11	heat	12 (3hr +	3.0h of 38	0.89112	1.24671	1.22421	0.9986
149	heat, cell	cell cultur	not applic	not applic	C12	heat	24 (3hr +	3.0h of 38	0.88639	1.25641	1.2335	0.8667

Microarray sample key - chemical and hormonal treatment

sample #	tissue cl	tissue	genotype	age	NASC sa	stress	time	treatmer	PAPS1	PAPS2	PAPS3	PAPS4
1	baseline v	seedling	Col	not applic	80	not applic	0.5 h	H2O	0.8833	0.6182	0.8989	0.7593
2	baseline v	seedling	Col	not applic	78	not applic	1 h	H2O	0.7026	0.6906	1.2425	0.8548
3	baseline v	seedling	Col	not applic	79	not applic	3 h	H2O	0.7702	0.5437	1.0272	0.6965
4	baseline v	seedling	Col	not applic	50	not applic	0.5 h	10µM ABA	0.7271	0.5234	0.9716	0.6991
5	baseline v	seedling	Col	not applic	48	not applic	1 h	10µM ABA	0.6	0.6101	0.9898	0.6227
6	baseline v	seedling	Col	not applic	49	not applic	3 h	10µM ABA	0.7859	0.852	1.0709	0.7317
7	baseline v	seedling	Col	not applic	53	not applic	0.5 h	10µM ACC	0.7016	0.4862	1.004	0.6105
8	baseline v	seedling	Col	not applic	51	not applic	1 h	10µM ACC	0.7375	0.5185	1.0231	0.635
9	baseline v	seedling	Col	not applic	52	not applic	3 h	10µM ACC	0.845	0.7699	1.0268	0.8358
10	baseline v	seedling	Col	not applic	56	not applic	0.5 h	10nM BL	0.6112	0.305	1.0706	0.6083
11	baseline v	seedling	Col	not applic	54	not applic	1 h	10nM BL	0.665	0.393	0.9487	0.7215
12	baseline v	seedling	Col	not applic	55	not applic	3 h	10nM BL	0.6564	0.3004	1.0413	0.5933
13	baseline v	seedling	Col	not applic	65	not applic	0.5 h	1µM GA	0.8498	0.4335	0.9506	0.7589
14	baseline v	seedling	Col	not applic	57	not applic	1 h	1µM GA	0.7615	0.7328	0.9152	0.7951
15	baseline v	seedling	Col	not applic	64	not applic	3 h	1µM GA	0.6134	0.5797	1.0054	0.6066
16	baseline v	seedling	Col	not applic	68	not applic	0.5 h	1µM IAA	0.7482	0.5404	0.92	0.6589
17	baseline v	seedling	Col	not applic	66	not applic	1 h	1µM IAA	0.8035	0.6143	0.9897	0.786
18	baseline v	seedling	Col	not applic	67	not applic	3 h	1µM IAA	0.8028	0.5722	1.0637	0.7513
19	baseline v	seedling	Col	not applic	71	not applic	0.5 h	1µM MJ	0.6037	0.4773	0.9993	0.6833
20	baseline v	seedling	Col	not applic	69	not applic	1 h	1µM MJ	0.4988	0.4852	0.9772	0.6475
21	baseline v	seedling	Col	not applic	70	not applic	3 h	1µM MJ	0.5116	0.6441	1.0214	0.6763
22	baseline v	seedling	Col	not applic	83	not applic	0.5 h	1µM zeati	0.8507	0.7333	0.921	0.7233
23	baseline v	seedling	Col	not applic	81	not applic	1 h	1µM zeati	0.8186	0.632	0.9284	0.8091
24	baseline v	seedling	Col	not applic	82	not applic	3 h	1µM zeati	0.5839	0.3382	1.1397	0.5639
25	substance	seedling	Col	not applic	20	not applic	3 h	H2O	0.6934	0.7559	0.9031	0.8264
26	substance	seedling	Col	not applic	19	not applic	12 h	H2O	0.5713	1.1991	0.8815	0.885
27	substance	seedling	Col	not applic	5	not applic	3 h	10µM Brz	0.7529	1.1233	0.8701	1.017
28	substance	seedling	Col	not applic	4	not applic	12 h	10µM Brz	0.8175	0.9933	0.8921	1.0568
29	substance	seedling	Col	not applic	2	not applic	3 h	10µM Brz	0.8561	0.8329	0.8657	1.0243
30	substance	seedling	Col	not applic	1	not applic	12 h	10µM Brz	0.7516	0.8392	0.8891	0.979
31	substance	seedling	Col	not applic	22	not applic	3 h	10µM pac	0.8452	0.9076	0.8717	1.0967
32	substance	seedling	Col	not applic	21	not applic	12 h	10µM pac	0.8667	1.2799	0.9111	1.0896

33	substance	seedling	Col	not applic	24	not applic 3 h	1µM PNOE	0.8725	0.8739	0.9156	0.9607
34	substance	seedling	Col	not applic	23	not applic 12 h	1µM PNOE	0.773	0.8844	0.9262	1.0626
35	substance	seedling	Col	not applic	26	not applic 3 h	10µM pro	0.7549	0.7662	0.8959	1.0531
36	substance	seedling	Col	not applic	25	not applic 12 h	10µM pro	0.8564	1.0352	0.8514	0.8484
37	substance	seedling	Col	not applic	28	not applic 3 h	10µM prol	0.6647	0.843	0.9942	1.0677
38	substance	seedling	Col	not applic	27	not applic 12 h	10µM prol	0.718	0.7194	0.8788	0.9062
39	substance	seedling	Col	not applic	30	not applic 3 h	10µM unic	0.9557	0.9976	1.1218	1.1507
40	substance	seedling	Col	not applic	29	not applic 12 h	10µM unic	0.8729	1.1023	0.8846	1.1157
41	substance	seedling	Col	not applic	12	not applic 3 h	H2O	0.8149	1.1151	0.9488	1.0808
42	substance	seedling	Col	not applic	3	not applic 3 h	10µM 2,4,	0.7284	0.8466	0.9477	0.9199
43	substance	seedling	Col	not applic	6	not applic 3 h	10µM AVC	0.7876	1.0513	0.9128	0.9511
44	substance	seedling	Col	not applic	7	not applic 3 h	10µM AGI	0.9467	2.7522	0.8779	1.1426
45	substance	seedling	Col	not applic	8	not applic 3 h	10µM B9 (	0.8424	0.8494	0.9249	0.8948
46	substance	seedling	Col	not applic	10	not applic 3 h	3µM Brz2:	0.9893	1.1698	0.9574	0.9453
47	substance	seedling	Col	not applic	9	not applic 3 h	10µM cycl	5.538	1.7686	3.5523	0.4389
48	substance	seedling	Col	not applic	18	not applic 3 h	1µM ibupr	0.7971	0.6259	1.4319	0.768
49	substance	seedling	Col	not applic	13	not applic 3 h	10µM MG:	0.831	1.1355	0.9114	0.8763
50	substance	seedling	Col	not applic	14	not applic 3 h	10µM NPA	0.8787	0.8277	0.9211	0.8868
51	substance	seedling	Col	not applic	15	not applic 3 h	10µM PCI	0.702	0.8832	1.3746	0.7845
52	substance	seedling	Col	not applic	11	not applic 3 h	10µM PNC	0.8291	1.7858	1.067	0.9066
53	substance	seedling	Col	not applic	16	not applic 3 h	10µM sali	0.708	0.9182	0.9148	0.9739
54	substance	seedling	Col	not applic	17	not applic 3 h	10µM TIB.	0.8501	1.2174	0.9313	0.9368
55	det2_subs	seedling	det2-1	not applic	43	not applic 3 h	H2O	1.0165	1.4537	0.8854	1.225
56	det2_subs	seedling	det2-1	not applic	31	not applic 3 h	1µM 3-del	0.9778	1.2904	0.8762	1.2566
57	det2_subs	seedling	det2-1	not applic	32	not applic 3 h	1µM 3-del	0.8291	1.2169	0.9525	1.1304
58	det2_subs	seedling	det2-1	not applic	33	not applic 3 h	1µM 6-de	1.0009	1.1218	0.877	1.0006
59	det2_subs	seedling	det2-1	not applic	34	not applic 3 h	1µM 6-de	1.0701	1.6329	0.9226	1.2925
60	det2_subs	seedling	det2-1	not applic	35	not applic 3 h	1µM 6-de	0.9965	1.2947	0.8512	1.2566
61	det2_subs	seedling	det2-1	not applic	36	not applic 3 h	1µM 6-de	0.993	1.6377	0.8466	1.1666
62	det2_subs	seedling	det2-1	not applic	37	not applic 3 h	10nM bra:	0.9493	1.0237	0.8263	1.1116
63	det2_subs	seedling	det2-1	not applic	38	not applic 3 h	10µM cam	1.1279	1.8168	1.2057	1.3673
64	det2_subs	seedling	det2-1	not applic	39	not applic 3 h	100nM ca:	0.8552	1.2774	0.8478	1.0105
65	det2_subs	seedling	det2-1	not applic	40	not applic 3 h	1µM catha	0.8643	1.5351	0.9006	1.265
66	det2_subs	seedling	det2-1	not applic	41	not applic 3 h	1µM teast	1.0173	1.2341	0.8934	1.1863

67	det2_subs	seedling	det2-1	not applic	42	not applic	3 h	1µM typh	0.9642	1.1147	0.8918	1.069
68	det2-1 / E	seedling	det2-1	not applic	77	not applic	0.5 h	H2O	0.6624	0.6629	1.0298	0.7938
69	det2-1 / E	seedling	det2-1	not applic	75	not applic	1 h	H2O	0.7362	0.679	0.9298	0.7373
70	det2-1 / E	seedling	det2-1	not applic	76	not applic	3 h	H2O	0.6936	0.731	1.0372	0.6224
71	det2-1 / E	seedling	det2-1	not applic	74	not applic	0.5 h	10nM BL	0.7571	0.4927	1.0832	0.6766
72	det2-1 / E	seedling	det2-1	not applic	72	not applic	1 h	10nM BL	0.7532	0.8485	0.9458	0.656
73	det2-1 / E	seedling	det2-1	not applic	73	not applic	3 h	10nM BL	0.7274	0.455	1.0624	0.5074
74	ga1-5 / G	seedling	ga1-5	not applic	63	not applic	0.5 h	H2O	1.0404	0.5779	0.9437	0.7683
75	ga1-5 / G	seedling	ga1-5	not applic	61	not applic	1 h	H2O	0.8233	0.3283	0.9601	0.7994
76	ga1-5 / G	seedling	ga1-5	not applic	62	not applic	3 h	H2O	0.9973	0.6553	0.8657	0.8229
77	ga1-5 / G	seedling	ga1-5	not applic	60	not applic	0.5 h	1µM GA	1.0089	0.5543	0.933	0.8821
78	ga1-5 / G	seedling	ga1-5	not applic	58	not applic	1 h	1µM GA	0.7714	0.3009	1.0033	0.7783
79	ga1-5 / G	seedling	ga1-5	not applic	59	not applic	3 h	1µM GA	1.0802	0.4686	0.9395	0.8381
80	ga1-3 / G	seeds	ga1-3	not applic	90	not applic	3 h	control	1.6608	2.2231	1.0331	1.5665
81	ga1-3 / G	seeds	ga1-3	not applic	91	not applic	6 h	control	1.6688	1.5321	0.9988	1.5371
82	ga1-3 / G	seeds	ga1-3	not applic	92	not applic	9 h	control	1.5429	1.9346	1.1961	1.5402
83	ga1-3 / G	seeds	ga1-3	not applic	93	not applic	3 h	5µm GA	1.5398	2.1343	0.9381	1.4016
84	ga1-3 / G	seeds	ga1-3	not applic	94	not applic	6 h	5µm GA	1.4138	1.7008	0.9505	1.5645
85	ga1-3 / G	seeds	ga1-3	not applic	95	not applic	9 h	5µm GA	1.7783	1.3274	0.9823	1.0684
86	hormone	seedling	Ler	not applic	45	not applic	No h	No	0.8432	0.3857	0.9054	0.8146
87	hormone	seedling	ga1-5 (Le	not applic	44	not applic	No h	No	0.7734	0.3881	0.9666	0.8255
88	hormone	seedling	Ws	not applic	46	not applic	No h	No	0.8867	0.3429	1.0312	0.797
89	hormone	seedling	bri1	not applic	47	not applic	No h	No	0.8338	0.4122	0.9746	1.0325
90	zeatin	seedling	Col	not applic	85	not applic	3 h	control	1.0458	0.6786	0.9741	0.592
91	zeatin	seedling	ARR21Co	not applic	87	not applic	3 h	control	0.5862	0.4574	0.9886	0.5696
92	zeatin	seedling	Col	not applic	84	not applic	3 h	control	1.1687	0.7455	0.9846	0.5565
93	zeatin	seedling	Col	not applic	86	not applic	3 h	20µm zea	0.8041	0.3666	0.8981	0.4789
94	zeatin	seedling	ARR22ox	not applic	88	not applic	3 h	control	0.3834	1.1962	1.0406	0.8184
95	zeatin	seedling	ARR22ox	not applic	89	not applic	3 h	20µm zea	0.7059	1.1528	0.9494	0.7428
96	sulfate sta	seedling	Col	not applic	98	not applic	2 h	0µm sulfa	1.0544	1.358	0.9063	1.0424
97	sulfate sta	seedling	Col	not applic	99	not applic	4 h	0µm sulfa	0.9994	2.0616	0.8782	1.3508
98	sulfate sta	seedling	Col	not applic	100	not applic	8 h	0µm sulfa	1.0092	1.3523	0.8782	1.2309
99	sulfate sta	seedling	Col	not applic	96	not applic	12 h	0µm sulfa	0.9369	1.16	0.8684	1.0393
100	sulfate sta	seedling	Col	not applic	97	not applic	24 h	0µm sulfa	0.9093	1.1349	0.9376	1.0635

101	sulfate sta	seedling	Col	not applic	101	not applic	0 h	1500µm s	0.866	1.3705	0.9478	1.3758
102	sulfate sta	seedling	Col	not applic	104	not applic	2 h	1500µm s	1.0967	1.2633	0.8916	1.2753
103	sulfate sta	seedling	Col	not applic	105	not applic	4 h	1500µm s	0.984	1.4631	0.9739	1.1289
104	sulfate sta	seedling	Col	not applic	106	not applic	8 h	1500µm s	1.0015	1.2318	0.9869	1.1575
105	sulfate sta	seedling	Col	not applic	102	not applic	12 h	1500µm s	0.9616	1.0876	0.9107	0.9917
106	sulfate sta	seedling	Col	not applic	103	not applic	24 h	1500µm s	0.951	0.9873	0.8558	1.0754
107	ABA	seeds	Col	not applic	109	not applic	0 h	No	2.2783	0.7697	1.0794	1.8083
108	ABA	seeds	Col	not applic	110	not applic	24 h	No	1.2859	0.7449	0.9009	0.9502
109	ABA	seeds	Col	not applic	107	not applic	24 h	3µm ABA	1.3282	0.8839	0.9545	1.2313
110	ABA	seeds	Col	not applic	108	not applic	24 h	30µm ABA	1.5451	1.0462	0.9284	1.2919
111	seed imbi	seeds	Col	not applic	111	not applic	0 h	imbibition	2.5976	0.9348	1.4038	2.2661
112	seed imbi	seeds	Col	not applic	112	not applic	1 h	imbibition	2.1339	0.9012	1.3226	1.5889
113	seed imbi	seeds	Col	not applic	113	not applic	3 h	imbibition	2.7469	1.9829	1.4858	2.2663
114	temperatu	seeds	Col	not applic	115	not applic	96 h	4°C	1.5523	1.6435	1.01	1.6626
115	temperatu	seeds	Col	not applic	114	not applic	96 h	22°C	2.42	4.7519	1.1887	3.607

Microarray sample key - biotic stress and light conditions

sample #	tissue cl	tissue	genotype	age	NASC sar	stress	time	treatmer	PAPS1	PAPS2	PAPS3	PAPS4
1	Pseudomc	leaf	not applic	not applic	17	not applic	0 h	untreated	0.7077	0.7549	1.0512	1.1225
2	Pseudomc	leaf	not applic	not applic	33	not applic	2 h	infiltrated	1.4036	1.1404	1.0245	1.2379
3	Pseudomc	leaf	not applic	not applic	34	not applic	6 h	infiltrated	0.5762	1.2325	0.9043	1.63
4	Pseudomc	leaf	not applic	not applic	35	not applic	24 h	infiltrated	0.7175	0.6535	1.1187	1.3606
5	Pseudomc	leaf	not applic	not applic	21	not applic	2 h	infiltrated	0.7675	1.2934	0.9939	1.1502
6	Pseudomc	leaf	not applic	not applic	22	not applic	6 h	infiltrated	0.4737	0.7248	0.908	0.9205
7	Pseudomc	leaf	not applic	not applic	23	not applic	24 h	infiltrated	0.49	0.9321	0.9728	1.159
8	Pseudomc	leaf	not applic	not applic	24	not applic	2 h	infiltrated	0.6854	1.449	0.9898	1.1957
9	Pseudomc	leaf	not applic	not applic	25	not applic	6 h	infiltrated	0.504	1.6078	0.9461	1.2477
10	Pseudomc	leaf	not applic	not applic	26	not applic	24 h	infiltrated	0.5328	0.7638	0.9327	1.1069
11	Pseudomc	leaf	not applic	not applic	27	not applic	2 h	infiltrated	1.0282	1.7144	1.0138	1.2719
12	Pseudomc	leaf	not applic	not applic	28	not applic	6 h	infiltrated	0.682	1.411	0.9158	1.5618
13	Pseudomc	leaf	not applic	not applic	29	not applic	24 h	infiltrated	0.7824	0.9328	0.9628	1.4368
14	Pseudomc	leaf	not applic	not applic	30	not applic	2 h	infiltrated	0.9727	1.7301	0.9875	1.1235
15	Pseudomc	leaf	not applic	not applic	31	not applic	6 h	infiltrated	0.5417	1.2895	0.9228	1.2174
16	Pseudomc	leaf	not applic	not applic	32	not applic	24 h	infiltrated	0.6372	0.7127	0.9473	1.264
17	Phytophth	leaf	not applic	not applic	C.6	not applic	6 h	for the tre	0.7485	0.7667	1.0872	0.956
18	Phytophth	leaf	not applic	not applic	C.12	not applic	12 h	for the tre	0.5963	0.7912	0.9647	0.8997
19	Phytophth	leaf	not applic	not applic	C.24	not applic	24 h	for the tre	0.5282	0.997	0.9644	1.0164
20	Phytophth	leaf	not applic	not applic	Pi.6	not applic	6 h	for the tre	0.4777	0.9064	1.0241	0.9279
21	Phytophth	leaf	not applic	not applic	Pi.12	not applic	12 h	for the tre	0.4665	0.9073	0.9532	0.9964
22	Phytophth	leaf	not applic	not applic	Pi.24	not applic	24 h	for the tre	0.5237	1.0161	0.894	1.0225
23	elicitors	leaf	not applic	not applic	1,1	not applic	1 h	infiltrated	2.491	0.58	1.1074	0.6077
24	elicitors	leaf	not applic	not applic	1,4	not applic	4 h	infiltrated	0.4543	0.5531	1.0151	0.76
25	elicitors	leaf	not applic	not applic	2,1	not applic	1 h	infiltrated	2.3342	0.6553	1.0583	0.6883
26	elicitors	leaf	not applic	not applic	2,4	not applic	4 h	infiltrated	0.4197	0.6652	0.9781	0.7345
27	elicitors	leaf	not applic	not applic	3,1	not applic	1 h	infiltrated	2.4739	0.7777	1.059	0.6938
28	elicitors	leaf	not applic	not applic	3,4	not applic	4 h	infiltrated	0.4819	0.8284	0.9607	0.8469
29	elicitors	leaf	not applic	not applic	4,1	not applic	1 h	infiltrated	2.9112	1.1902	1.2114	0.6935
30	elicitors	leaf	not applic	not applic	4,4	not applic	4 h	infiltrated	0.3799	1.2025	0.9475	0.684
31	elicitors	leaf	not applic	not applic	5,1	not applic	1 h	infiltrated	2.7843	1.1125	1.0866	0.6793
32	elicitors	leaf	not applic	not applic	5,4	not applic	4 h	infiltrated	0.3718	1.2073	0.9336	0.8413

33	elicitors	leaf	not applic	not applic	6,1	not applic 1 h	infiltrated	3.0932	1.2369	1.2082	0.704
34	elicitors	leaf	not applic	not applic	6,4	not applic 4 h	infiltrated	0.3665	0.6453	0.9454	0.6988
35	elicitors	leaf	not applic	not applic	7,1	not applic 1 h	infiltrated	2.1991	0.7797	1.0729	0.7306
36	elicitors	leaf	not applic	not applic	7,4	not applic 4 h	infiltrated	0.3972	0.8398	0.9378	0.8138
37	darkness	hypocotyl	not applic	not applic	DS	not applic 45 min	light treat	1.0223	1.0505	1.0308	1.0405
38	darkness	hypocotyl	not applic	not applic	DL	not applic 240 min	light treat	1.0564	1.2993	0.9926	1.1585
39	UV-A	hypocotyl	not applic	not applic	AS	not applic 45 min	light treat	1.0958	0.8919	1.0231	0.9579
40	UV-A	hypocotyl	not applic	not applic	AL	not applic 240 min	light treat	1.0962	0.962	0.974	0.8905
41	UV-A/B	hypocotyl	not applic	not applic	US	not applic 45 min	light treat	1.0664	1.0316	1.0009	1.0626
42	UV-A/B	hypocotyl	not applic	not applic	UL	not applic 240 min	light treat	0.9531	1.0455	0.9699	1.0474
43	blue light	hypocotyl	not applic	not applic	BS	not applic 45 min	light treat	0.958	0.7787	1.0529	0.8931
44	blue light	hypocotyl	not applic	not applic	BL	not applic 240 min	light treat	1.0263	0.9334	0.9582	0.9557
45	red (pulse	hypocotyl	not applic	not applic	PS	not applic 45 min	light treat	0.8657	0.9789	1.046	0.9355
46	red (pulse	hypocotyl	not applic	not applic	PL	not applic 240 min	light treat	1.027	1.0997	0.9731	1.0758
47	red light	hypocotyl	not applic	not applic	RS	not applic 45 min	light treat	0.9553	1.0408	1.0154	1.0703
48	red light	hypocotyl	not applic	not applic	RL	not applic 240 min	light treat	1.0032	0.9821	0.9942	1.0455
49	far-red lig	hypocotyl	not applic	not applic	FS	not applic 45 min	light treat	1.0092	0.9706	1.0071	0.9756
50	far-red lig	hypocotyl	not applic	not applic	FL	not applic 240 min	light treat	0.8673	1.0283	0.9652	0.9241
51	white	hypocotyl	not applic	not applic	WS	not applic 45 min	light treat	0.9885	0.8263	1.0211	0.9694
52	white	hypocotyl	not applic	not applic	WL	not applic 240 min	light treat	1.0099	1.0812	0.9764	0.9983

A2: NASC keys and data for figures 3.6 – 3.9 in Chapter 3

Microarray sample key - developmental stage

sample #	tissue cl	tissue	genotype	age	NASC sa	CstF50	CstF64	CstF77
1	root	roots	Wt	7 days	ATGE_3	0.8993	1.1765	0.9714
2	root	roots	Wt	17 days	ATGE_9	0.6731	1.0237	1.0936
3	root	root	Wt	15 days	ATGE_93	0.4976	1.0076	0.7764
4	root	root	Wt	8 days	ATGE_94	0.5575	1.157	1.0379
5	root	root	Wt	8 days	ATGE_95	0.5917	1.1817	1.1568
6	root	root	Wt	21 days	ATGE_98	0.5415	1.0948	1.0861
7	root	root	Wt	21 days	ATGE_99	0.609	1.1313	1.0812
8	stem	hypocotyl	Wt	7 days	ATGE_2	0.9329	0.8486	0.7728
9	stem	1st node	Wt	21+ days	ATGE_28	1.1898	0.8737	0.7282
10	stem	stem, 2nd	Wt	21+ days	ATGE_27	0.9771	0.8044	0.8414
11	leaf	cotyledon:	Wt	7 days	ATGE_1	1.0814	0.9145	0.541
12	leaf	leaves 1 +	Wt	7 days	ATGE_5	1.2744	0.7416	0.4626
13	leaf	rosette le:	Wt	10 days	ATGE_10	1.4319	0.9113	0.6017
14	leaf	rosette le:	gl1-T	10 days	ATGE_11	1.1719	0.8526	0.59
15	leaf	rosette le:	Wt	17 days	ATGE_12	0.7882	0.8654	0.8604
16	leaf	rosette le:	Wt	17 days	ATGE_13	0.8042	0.8361	0.754
17	leaf	rosette le:	Wt	17 days	ATGE_14	0.8581	0.8688	0.7032
18	leaf	rosette le:	Wt	17 days	ATGE_15	1.0383	0.7801	0.5703
19	leaf	rosette le:	Wt	17 days	ATGE_16	1.0376	0.8569	0.6489
20	leaf	rosette le:	Wt	17 days	ATGE_17	1.2672	0.8018	0.6627
21	leaf	rosette le:	gl1-T	17 days	ATGE_18	1.3815	0.9536	0.7046
22	leaf	leaf 7, pet	Wt	17 days	ATGE_19	0.9003	0.7884	0.6345
23	leaf	leaf 7, prc	Wt	17 days	ATGE_20	0.8691	0.8832	0.6086
24	leaf	leaf 7, dis	Wt	17 days	ATGE_21	0.7869	0.8771	0.7222
25	leaf	leaf	Wt	15 days	ATGE_91	1.0125	0.8574	0.6375
26	leaf	senescing	Wt	35 days	ATGE_25	0.7112	1.4874	1.5215
27	leaf	cauline le:	Wt	21+ days	ATGE_26	1.0424	1.1178	1.0003
28	whole pla	seedling,	Wt	7 days	ATGE_7	1.4302	0.9401	0.856
29	whole pla	seedling,	Wt	8 days	ATGE_96	0.7365	0.8485	0.6579
30	whole pla	seedling,	Wt	8 days	ATGE_97	0.6997	0.8278	0.7168
31	whole pla	seedling,	Wt	21 days	ATGE_10(	0.6254	0.735	0.7323
32	whole pla	seedling,	Wt	21 days	ATGE_101	0.4859	0.7738	0.738

33	whole plai	developm	Wt	21 days	ATGE_22	1.026	0.8033	0.5824
34	whole plai	as above	Wt	22 days	ATGE_23	1.1137	0.7886	0.7136
35	whole plai	as above	Wt	23 days	ATGE_24	1.0414	0.8574	0.7791
36	whole plai	vegetative	Wt	7 days	ATGE_87	1.099	0.8983	0.7002
37	whole plai	vegetative	Wt	14 days	ATGE_89	1.5962	0.9301	0.8216
38	whole plai	vegetative	Wt	21 days	ATGE_90	1.5193	0.965	0.8655
39	apex	shoot ape	Wt	7 days	ATGE_4	1.3371	0.8939	0.774
40	apex	shoot ape	Wt	7 days	ATGE_6	1.1875	0.9477	1.0447
41	apex	shoot ape	Wt	14 days	ATGE_8	1.5846	0.959	1.2767
42	apex	shoot ape	Wt	21 days	ATGE_29	1.4727	1.1432	1.4285
43	apex	shoot ape	clv3-7	21+ days	ATGE_46	1.6204	1.0621	1.3512
44	apex	shoot ape	lfy-12	21+ days	ATGE_47	1.4306	0.9402	1.3462
45	apex	shoot ape	ap1-15	21+ days	ATGE_48	1.4835	1.1146	1.5567
46	apex	shoot ape	ap2-6	21+ days	ATGE_49	1.6246	1.013	1.4237
47	apex	shoot ape	ufo-1	21+ days	ATGE_52	1.4743	1.0385	1.3776
48	apex	shoot ape	ap3-6	21+ days	ATGE_50	1.5808	1.0469	1.5866
49	apex	shoot ape	ag-12	21+ days	ATGE_51	1.3571	1.0691	1.5629
50	flowers	flowers st	Wt	21+ days	ATGE_31	1.2399	0.9653	1.2489
51	flowers	flowers st	Wt	21+ days	ATGE_32	1.3314	0.8668	1.0737
52	flowers	flowers st	Wt	21+ days	ATGE_33	0.8343	0.8223	0.8616
53	flowers	flower sta	clv3-7	21+ days	ATGE_53	0.8371	0.8881	0.7412
54	flowers	flower sta	lfy-12	21+ days	ATGE_54	1.2921	0.9508	1.1832
55	flowers	flower sta	ap1-15	21+ days	ATGE_55	1.3438	0.9792	1.5274
56	flowers	flower sta	ap2-6	21+ days	ATGE_56	1.1837	0.9336	1.1181
57	flowers	flower sta	ufo-1	21+ days	ATGE_59	1.1919	0.9233	1.2283
58	flowers	flower sta	ap3-6	21+ days	ATGE_57	1.3641	1.0671	1.1386
59	flowers	flower sta	ag-12	21+ days	ATGE_58	1.1167	0.9525	1.1455
60	flowers	flowers st	Wt	21+ days	ATGE_39	0.6473	0.8186	0.7084
61	flowers	flower	Wt	28 days	ATGE_92	1.1127	0.8516	0.8937
62	floral orga	flowers st	Wt	21+ days	ATGE_40	1.2214	0.8211	0.651
63	floral orga	flowers st	Wt	21+ days	ATGE_34	0.6529	0.8368	0.8108
64	floral orga	flowers st	Wt	21+ days	ATGE_41	0.4473	1.3388	1.0331
65	floral orga	flowers st	Wt	21+ days	ATGE_35	1.0428	1.1626	0.8219
66	floral orga	flowers st	Wt	21+ days	ATGE_42	0.5194	1.2222	0.9

67	floral orga	flowers st	Wt	21+ days	ATGE_36	0.3651	1.7759	0.5715
68	floral orga	flowers st	Wt	21+ days	ATGE_43	0.3368	1.5575	0.5873
69	floral orga	mature pc	Wt	6 wk	ATGE_73	0.1642	4.086	0.0725
70	floral orga	flowers st	Wt	21+ days	ATGE_37	1.2907	0.9729	1.3597
71	floral orga	flowers st	Wt	21+ days	ATGE_45	1.1464	0.8765	1.0858
72	seeds	siliques, w	Wt	8 wk	ATGE_76	1.094	0.7741	0.8015
73	seeds	siliques, w	Wt	8 wk	ATGE_77	0.921	0.6939	1.0779
74	seeds	siliques, w	Wt	8 wk	ATGE_78	0.8188	0.7339	0.9349
75	seeds	seeds, sta	Wt	8 wk	ATGE_79	1.0553	0.752	1.083
76	seeds	seeds, sta	Wt	8 wk	ATGE_81	1.1087	0.9779	1.2786
77	seeds	seeds, sta	Wt	8 wk	ATGE_82	0.4113	0.8231	2.3896
78	seeds	seeds, sta	Wt	8 wk	ATGE_83	0.2455	1.2399	3.2155
79	seeds	seeds, sta	Wt	8 wk	ATGE_84	0.2146	1.0501	2.799

Microarray sample key - abiotic stress

sample #	tissue cl	tissue	genotype	age	NASC sa	stress	time	treatmer	CstF50	CstF64	CstF77
1	control, a	seedling,	not applic	not applic	1	untreated	0 h	Control: M	1.0412	0.5265	0.7065
2	control, a	seedling,	not applic	not applic	71	mock	0.25 h	Control: T	1.0918	0.5703	0.7641
3	control, a	seedling,	not applic	not applic	11	mock	0.5 h	Control: T	0.9766	0.4547	0.6794
4	control, a	seedling,	not applic	not applic	21	mock	1 h	Control: T	1.1596	0.6206	0.8079
5	control, a	seedling,	not applic	not applic	31	mock	3 h	Control: T	1.3427	0.7452	0.7918
6	control, a	seedling,	not applic	not applic	81	mock	4 h	Control: T	1.0127	0.9091	0.7414
7	control, a	seedling,	not applic	not applic	41	mock	6 h	Control: T	1.3249	0.9351	0.6328
8	control, a	seedling,	not applic	not applic	51	mock	12 h	Control: T	1.2354	0.8973	0.7068
9	control, a	seedling,	not applic	not applic	61	mock	24 h	Control: T	1.1876	0.4554	1.0759
10	control, rc	root	not applic	not applic	2	untreated	0 h	Control: M	0.7913	1.1319	1.1893
11	control, rc	root	not applic	not applic	72	mock	0.25 h	Control: T	0.7915	1.2652	1.2349
12	control, rc	root	not applic	not applic	12	mock	0.5 h	Control: T	0.8626	0.9968	1.1042
13	control, rc	root	not applic	not applic	22	mock	1 h	Control: T	0.855	1.1572	1.0706
14	control, rc	root	not applic	not applic	32	mock	3 h	Control: T	0.847	1.1692	1.2145
15	control, rc	root	not applic	not applic	82	mock	4 h	Control: T	0.8228	1.3688	1.363
16	control, rc	root	not applic	not applic	42	mock	6 h	Control: T	0.7606	1.2138	1.1983
17	control, rc	root	not applic	not applic	52	mock	12 h	Control: T	0.9217	1.3725	1.2802
18	control, rc	root	not applic	not applic	62	mock	24 h	Control: T	0.8359	1.1118	1.2043
19	cold, aeria	seedling,	not applic	not applic	111	cold	0.5 h	Cold stres	1.0285	0.3802	0.7001
20	cold, aeria	seedling,	not applic	not applic	121	cold	1 h	Cold stres	1.1466	0.6876	0.6842
21	cold, aeria	seedling,	not applic	not applic	131	cold	3 h	Cold stres	0.9743	0.5902	0.7184
22	cold, aeria	seedling,	not applic	not applic	141	cold	6 h	Cold stres	0.9538	1.3695	0.7414
23	cold, aeria	seedling,	not applic	not applic	151	cold	12 h	Cold stres	0.7627	3.0141	0.8279
24	cold, aeria	seedling,	not applic	not applic	161	cold	24 h	Cold stres	0.6026	1.3448	0.7643
25	cold, root	root	not applic	not applic	112	cold	0.5 h	Cold stres	0.8183	1.2798	1.2218
26	cold, root	root	not applic	not applic	122	cold	1 h	Cold stres	0.7121	1.1334	0.9941
27	cold, root	root	not applic	not applic	132	cold	3 h	Cold stres	0.7454	1.2145	1.2478
28	cold, root	root	not applic	not applic	142	cold	6 h	Cold stres	0.6084	1.2049	1.0607
29	cold, root	root	not applic	not applic	152	cold	12 h	Cold stres	0.4672	1.6681	1.184
30	cold, root	root	not applic	not applic	162	cold	24 h	Cold stres	0.3479	2.0635	0.9526
31	osmotic, a	seedling,	not applic	not applic	211	osmotic	0.5 h	Osmotic s	1.0212	0.5911	0.6032
32	osmotic, a	seedling,	not applic	not applic	221	osmotic	1 h	Osmotic s	1.0547	0.7166	0.7897

33	osmotic, a	seedling,	not applic	not applic	231	osmotic	3 h	Osmotic s	0.9655	0.8151	0.8032
34	osmotic, a	seedling,	not applic	not applic	241	osmotic	6 h	Osmotic s	0.7737	0.9811	0.7414
35	osmotic, a	seedling,	not applic	not applic	251	osmotic	12 h	Osmotic s	0.7725	0.8608	0.8854
36	osmotic, a	seedling,	not applic	not applic	261	osmotic	24 h	Osmotic s	0.523	0.4603	1.1775
37	osmotic, r	root	not applic	not applic	212	osmotic	0.5 h	Osmotic s	0.5998	1.1593	1.2448
38	osmotic, r	root	not applic	not applic	222	osmotic	1 h	Osmotic s	0.5568	1.094	1.2897
39	osmotic, r	root	not applic	not applic	232	osmotic	3 h	Osmotic s	0.5475	1.0976	1.214
40	osmotic, r	root	not applic	not applic	242	osmotic	6 h	Osmotic s	0.5914	1.0985	1.2582
41	osmotic, r	root	not applic	not applic	252	osmotic	12 h	Osmotic s	0.5018	1.0856	1.1456
42	osmotic, r	root	not applic	not applic	262	osmotic	24 h	Osmotic s	0.6917	1.1026	1.177
43	salt, aeria	seedling,	not applic	not applic	311	salt	0.5 h	Salt stres:	1.0438	0.4999	0.6972
44	salt, aeria	seedling,	not applic	not applic	321	salt	1 h	Salt stres:	1.2325	0.6834	0.8716
45	salt, aeria	seedling,	not applic	not applic	331	salt	3 h	Salt stres:	1.1055	0.6259	0.9027
46	salt, aeria	seedling,	not applic	not applic	341	salt	6 h	Salt stres:	1.1027	0.8682	0.6835
47	salt, aeria	seedling,	not applic	not applic	351	salt	12 h	Salt stres:	1.2689	0.9061	0.7125
48	salt, aeria	seedling,	not applic	not applic	361	salt	24 h	Salt stres:	0.8653	0.5416	0.6932
49	root, aeria	root	not applic	not applic	312	salt	0.5 h	Salt stres:	0.6717	1.0678	1.116
50	root, aeria	root	not applic	not applic	322	salt	1 h	Salt stres:	0.5968	1.1419	1.1568
51	root, aeria	root	not applic	not applic	332	salt	3 h	Salt stres:	0.5679	1.1724	0.9543
52	root, aeria	root	not applic	not applic	342	salt	6 h	Salt stres:	0.5068	1.2676	0.9484
53	root, aeria	root	not applic	not applic	352	salt	12 h	Salt stres:	0.5526	1.1835	1.0937
54	root, aeria	root	not applic	not applic	362	salt	24 h	Salt stres:	0.5424	1.0745	0.7924
55	drought, a	seedling,	not applic	not applic	471	drought	0.25 h	Drought s	1.086	0.4215	0.7872
56	drought, a	seedling,	not applic	not applic	411	drought	0.5 h	Drought s	1.0702	0.5146	0.7765
57	drought, a	seedling,	not applic	not applic	421	drought	1 h	Drought s	0.8773	0.502	0.776
58	drought, a	seedling,	not applic	not applic	431	drought	3 h	Drought s	1.1992	0.7343	0.7388
59	drought, a	seedling,	not applic	not applic	441	drought	6 h	Drought s	1.2805	0.9588	0.6721
60	drought, a	seedling,	not applic	not applic	451	drought	12 h	Drought s	1.2858	0.9251	0.7282
61	drought, a	seedling,	not applic	not applic	461	drought	24 h	Drought s	1.0856	0.4726	0.9368
62	drought, r	root	not applic	not applic	472	drought	0.25 h	Drought s	0.7183	1.0207	1.253
63	drought, r	root	not applic	not applic	412	drought	0.5 h	Drought s	0.8515	0.9446	0.9477
64	drought, r	root	not applic	not applic	422	drought	1 h	Drought s	0.7293	1.0449	1.1447
65	drought, r	root	not applic	not applic	432	drought	3 h	Drought s	0.724	1.0086	0.9968
66	drought, r	root	not applic	not applic	442	drought	6 h	Drought s	0.905	1.2712	1.2581

67	drought, r	root	not applic	not applic	452	drought	12 h	Drought s	0.9226	1.1889	1.1958
68	drought, r	root	not applic	not applic	462	drought	24 h	Drought s	0.8758	1.0446	1.2182
69	genotoxic	seedling,	not applic	not applic	511	genotoxic	0.5 h	Genotoxic	1.146	0.4219	0.7982
70	genotoxic	seedling,	not applic	not applic	521	genotoxic	1 h	Genotoxic	1.1845	0.5733	0.7263
71	genotoxic	seedling,	not applic	not applic	531	genotoxic	3 h	Genotoxic	1.4914	0.6848	0.7724
72	genotoxic	seedling,	not applic	not applic	541	genotoxic	6 h	Genotoxic	1.433	0.8405	0.7632
73	genotoxic	seedling,	not applic	not applic	551	genotoxic	12 h	Genotoxic	1.1176	0.8044	0.7893
74	genotoxic	seedling,	not applic	not applic	561	genotoxic	24 h	Genotoxic	0.9665	0.4365	0.9487
75	genotoxic	root	not applic	not applic	512	genotoxic	0.5 h	Genotoxic	0.8873	1.1364	1.1742
76	genotoxic	root	not applic	not applic	522	genotoxic	1 h	Genotoxic	0.8494	1.1256	1.2307
77	genotoxic	root	not applic	not applic	532	genotoxic	3 h	Genotoxic	0.7915	1.1115	1.1231
78	genotoxic	root	not applic	not applic	542	genotoxic	6 h	Genotoxic	0.8238	1.135	1.2772
79	genotoxic	root	not applic	not applic	552	genotoxic	12 h	Genotoxic	0.7892	1.1394	1.3177
80	genotoxic	root	not applic	not applic	562	genotoxic	24 h	Genotoxic	0.8729	1.0233	1.4966
81	oxidative,	seedling,	not applic	not applic	611	oxidative	0.5 h	Oxidative	1.163	0.558	0.7798
82	oxidative,	seedling,	not applic	not applic	621	oxidative	1 h	Oxidative	1.2337	0.6275	0.8896
83	oxidative,	seedling,	not applic	not applic	631	oxidative	3 h	Oxidative	1.3959	0.6378	0.7026
84	oxidative,	seedling,	not applic	not applic	641	oxidative	6 h	Oxidative	1.3642	0.8034	0.6905
85	oxidative,	seedling,	not applic	not applic	651	oxidative	12 h	Oxidative	1.0854	0.7557	0.6834
86	oxidative,	seedling,	not applic	not applic	661	oxidative	24 h	Oxidative	0.8987	0.4033	0.6378
87	oxidative,	root	not applic	not applic	612	oxidative	0.5 h	Oxidative	0.8033	1.0168	1.2307
88	oxidative,	root	not applic	not applic	622	oxidative	1 h	Oxidative	0.9185	1.1621	1.2708
89	oxidative,	root	not applic	not applic	632	oxidative	3 h	Oxidative	0.7186	1.2571	1.2717
90	oxidative,	root	not applic	not applic	642	oxidative	6 h	Oxidative	0.7677	1.1895	1.2393
91	oxidative,	root	not applic	not applic	652	oxidative	12 h	Oxidative	0.8304	1.3481	1.1204
92	oxidative,	root	not applic	not applic	662	oxidative	24 h	Oxidative	0.6576	1.1029	1.2511
93	UV-B, aer	seedling,	not applic	not applic	771	UV-B	0.25 h	UV-B stre	1.0719	0.6019	0.9622
94	UV-B, aer	seedling,	not applic	not applic	711	UV-B	0.5 h	UV-B stre	0.9492	0.4592	0.8087
95	UV-B, aer	seedling,	not applic	not applic	721	UV-B	1 h	UV-B stre	0.9481	0.5919	0.7584
96	UV-B, aer	seedling,	not applic	not applic	731	UV-B	3 h	UV-B stre	0.8566	0.8654	0.9279
97	UV-B, aer	seedling,	not applic	not applic	741	UV-B	6 h	UV-B stre	1.0034	1.3733	1.1265
98	UV-B, aer	seedling,	not applic	not applic	751	UV-B	12 h	UV-B stre	1.4234	1.0137	0.9275
99	UV-B, aer	seedling,	not applic	not applic	761	UV-B	24 h	UV-B stre	1.0487	0.6242	0.8941
100	UV-B, roo	root	not applic	not applic	772	UV-B	0.25 h	UV-B stre	0.7662	1.2883	1.2288

101	UV-B, roo	root	not applic	not applic	712	UV-B	0.5 h	UV-B stre	0.8081	1.0471	1.2115
102	UV-B, roo	root	not applic	not applic	722	UV-B	1 h	UV-B stre	0.9024	1.1035	1.225
103	UV-B, roo	root	not applic	not applic	732	UV-B	3 h	UV-B stre	0.8627	1.3542	1.3527
104	UV-B, roo	root	not applic	not applic	742	UV-B	6 h	UV-B stre	0.8083	1.2254	1.4781
105	UV-B, roo	root	not applic	not applic	752	UV-B	12 h	UV-B stre	0.8796	1.2771	1.4394
106	UV-B, roo	root	not applic	not applic	762	UV-B	24 h	UV-B stre	0.6006	1.1309	1.4288
107	wounding	seedling,	not applic	not applic	871	wounding	0.25 h	Wound sti	0.9278	0.4462	0.8318
108	wounding	seedling,	not applic	not applic	811	wounding	0.5 h	Wound sti	0.9631	0.6347	0.8467
109	wounding	seedling,	not applic	not applic	821	wounding	1 h	Wound sti	0.7291	0.6894	1.0107
110	wounding	seedling,	not applic	not applic	831	wounding	3 h	Wound sti	1.2153	0.8408	0.6547
111	wounding	seedling,	not applic	not applic	841	wounding	6 h	Wound sti	1.6056	1.0395	0.6727
112	wounding	seedling,	not applic	not applic	851	wounding	12 h	Wound sti	1.0925	0.9336	0.7142
113	wounding	seedling,	not applic	not applic	861	wounding	24 h	Wound sti	1.0713	0.5143	0.9633
114	wounding	root	not applic	not applic	872	wounding	0.25 h	Wound sti	0.7621	1.0171	1.2125
115	wounding	root	not applic	not applic	812	wounding	0.5 h	Wound sti	0.8038	1.1168	1.1681
116	wounding	root	not applic	not applic	822	wounding	1 h	Wound sti	0.8446	1.1867	1.308
117	wounding	root	not applic	not applic	832	wounding	3 h	Wound sti	0.8684	1.2454	1.3745
118	wounding	root	not applic	not applic	842	wounding	6 h	Wound sti	0.7973	1.0984	1.2314
119	wounding	root	not applic	not applic	852	wounding	12 h	Wound sti	0.7941	1.11	1.1917
120	wounding	root	not applic	not applic	862	wounding	24 h	Wound sti	0.7504	1.0389	1.2241
121	heat, aeri	seedling,	not applic	not applic	971	heat	0.25 h	0.25h of 3	1.4136	0.489	0.9397
122	heat, aeri	seedling,	not applic	not applic	911	heat	0.5 h	0.5h of 38	1.1835	0.5248	0.7891
123	heat, aeri	seedling,	not applic	not applic	921	heat	1 h	1.0h of 38	1.1382	0.4908	0.6194
124	heat, aeri	seedling,	not applic	not applic	931	heat	3 h	3.0h of 38	0.6246	0.7543	0.4125
125	heat, aeri	seedling,	not applic	not applic	981	heat	4 (3hr + :	3.0h of 38	1.7355	0.987	1.113
126	heat, aeri	seedling,	not applic	not applic	941	heat	6 (3hr + :	3.0h of 38	1.451	0.8037	0.8007
127	heat, aeri	seedling,	not applic	not applic	951	heat	12 (3hr +	3.0h of 38	1.4255	1.0087	0.7926
128	heat, aeri	seedling,	not applic	not applic	961	heat	24 (3hr +	3.0h of 38	1.273	0.4225	0.897
129	heat, root	root	not applic	not applic	972	heat	0.25 h	0.25h of 3	0.8808	1.1788	1.112
130	heat, root	root	not applic	not applic	912	heat	0.5 h	0.5h of 38	0.86	1.2032	1.3651
131	heat, root	root	not applic	not applic	922	heat	1 h	1.0h of 38	0.666	1.1113	1.3163
132	heat, root	root	not applic	not applic	932	heat	3 h	3.0h of 38	0.4891	1.0967	1.2902
133	heat, root	root	not applic	not applic	982	heat	4 (3hr + :	3.0h of 38	0.6585	1.8486	1.6702
134	heat, root	root	not applic	not applic	942	heat	6 (3hr + :	3.0h of 38	1.0516	1.5219	1.548

135	heat, root	root	not applic	not applic	952	heat	12 (3hr +	3.0h of 38	1.2365	1.3967	1.5835
136	heat, root	root	not applic	not applic	962	heat	24 (3hr +	3.0h of 38	1.03	1.1713	1.4032
137	ctrl, cell c	cell cultur	not applic	not applic	C0	control	0 h	incubated	2.3931	1.3896	0.9495
138	ctrl, cell c	cell cultur	not applic	not applic	C1	control	3 h	3.0h at 25	2.129	1.608	0.7999
139	ctrl, cell c	cell cultur	not applic	not applic	C2	control	6 h	6.0h at 25	2.2144	1.5395	0.9192
140	ctrl, cell c	cell cultur	not applic	not applic	C3	control	12 h	12.0h at 2	2.2038	1.4859	0.819
141	ctrl, cell c	cell cultur	not applic	not applic	C4	control	24 h	24.0h at 2	2.2507	1.6043	0.8711
142	heat, cell	cell cultur	not applic	not applic	C5	heat	0.25 h	0.25h of 3	1.6939	1.0934	0.9536
143	heat, cell	cell cultur	not applic	not applic	C6	heat	0.5 h	0.5h of 38	1.4748	0.9265	0.8198
144	heat, cell	cell cultur	not applic	not applic	C7	heat	1 h	1.0h of 38	1.1246	0.8841	0.8431
145	heat, cell	cell cultur	not applic	not applic	C8	heat	3 h	3.0h of 38	0.5253	1.4806	0.9688
146	heat, cell	cell cultur	not applic	not applic	C9	heat	4 (3hr +	3.0h of 38	0.8156	1.5256	0.9367
147	heat, cell	cell cultur	not applic	not applic	C10	heat	6 (3hr +	3.0h of 38	1.3483	1.524	1.016
148	heat, cell	cell cultur	not applic	not applic	C11	heat	12 (3hr +	3.0h of 38	2.1704	1.4327	0.9811
149	heat, cell	cell cultur	not applic	not applic	C12	heat	24 (3hr +	3.0h of 38	1.9051	1.3419	0.8719

Microarray sample key - chemical and hormonal treatment

sample #	tissue cl	tissue	genotype	age	NASC sa	stress	time	treatmer	CstF50	CstF64	CstF77
1	baseline v	seedling	Col	not applic	80	not applic	0.5 h	H2O	0.9218	0.6892	0.5566
2	baseline v	seedling	Col	not applic	78	not applic	1 h	H2O	1.0848	0.4934	0.7503
3	baseline v	seedling	Col	not applic	79	not applic	3 h	H2O	0.7572	0.6048	0.5791
4	baseline v	seedling	Col	not applic	50	not applic	0.5 h	10µM ABA	0.8341	0.4545	0.3621
5	baseline v	seedling	Col	not applic	48	not applic	1 h	10µM ABA	0.6888	0.5685	0.7352
6	baseline v	seedling	Col	not applic	49	not applic	3 h	10µM ABA	0.5619	0.5911	0.5196
7	baseline v	seedling	Col	not applic	53	not applic	0.5 h	10µM ACC	0.7899	0.7449	0.5799
8	baseline v	seedling	Col	not applic	51	not applic	1 h	10µM ACC	1.0086	0.5515	0.2558
9	baseline v	seedling	Col	not applic	52	not applic	3 h	10µM ACC	0.8927	0.4914	0.8304
10	baseline v	seedling	Col	not applic	56	not applic	0.5 h	10nM BL	0.7524	0.6253	0.1595
11	baseline v	seedling	Col	not applic	54	not applic	1 h	10nM BL	0.6644	0.678	0.3982
12	baseline v	seedling	Col	not applic	55	not applic	3 h	10nM BL	0.8624	0.7519	0.4403
13	baseline v	seedling	Col	not applic	65	not applic	0.5 h	1µM GA	1.0288	0.6265	0.3591
14	baseline v	seedling	Col	not applic	57	not applic	1 h	1µM GA	1.1133	0.5607	0.3821
15	baseline v	seedling	Col	not applic	64	not applic	3 h	1µM GA	0.7501	0.6943	0.2939
16	baseline v	seedling	Col	not applic	68	not applic	0.5 h	1µM IAA	0.9842	0.821	0.4371
17	baseline v	seedling	Col	not applic	66	not applic	1 h	1µM IAA	0.8042	0.738	0.469
18	baseline v	seedling	Col	not applic	67	not applic	3 h	1µM IAA	0.9361	0.7068	0.2545
19	baseline v	seedling	Col	not applic	71	not applic	0.5 h	1µM MJ	0.7503	0.5494	0.371
20	baseline v	seedling	Col	not applic	69	not applic	1 h	1µM MJ	0.6523	0.6849	0.5628
21	baseline v	seedling	Col	not applic	70	not applic	3 h	1µM MJ	0.739	0.5114	0.3319
22	baseline v	seedling	Col	not applic	83	not applic	0.5 h	1µM zeati	1.1505	0.7469	0.6778
23	baseline v	seedling	Col	not applic	81	not applic	1 h	1µM zeati	1.1098	0.7375	0.4834
24	baseline v	seedling	Col	not applic	82	not applic	3 h	1µM zeati	0.7876	0.6529	0.5116
25	substance	seedling	Col	not applic	20	not applic	3 h	H2O	0.9119	0.7142	0.6877
26	substance	seedling	Col	not applic	19	not applic	12 h	H2O	1.1735	0.8711	0.5607
27	substance	seedling	Col	not applic	5	not applic	3 h	10µM Brz	0.9287	0.868	0.4629
28	substance	seedling	Col	not applic	4	not applic	12 h	10µM Brz	0.7671	0.7373	0.7487
29	substance	seedling	Col	not applic	2	not applic	3 h	10µM Brz	1.1492	1.0742	0.6255
30	substance	seedling	Col	not applic	1	not applic	12 h	10µM Brz	1.0253	0.7867	0.8728
31	substance	seedling	Col	not applic	22	not applic	3 h	10µM pac	0.8193	0.7775	0.8288
32	substance	seedling	Col	not applic	21	not applic	12 h	10µM pac	1.1642	0.8385	0.9249

33	substance	seedling	Col	not applic	24	not applic	3 h	1µM PNOE	1.1134	0.7814	0.5876
34	substance	seedling	Col	not applic	23	not applic	12 h	1µM PNOE	0.9113	0.6551	0.6079
35	substance	seedling	Col	not applic	26	not applic	3 h	10µM proj	1.1563	0.897	0.6283
36	substance	seedling	Col	not applic	25	not applic	12 h	10µM proj	0.7644	0.6186	0.778
37	substance	seedling	Col	not applic	28	not applic	3 h	10µM prol	0.8396	0.789	0.8875
38	substance	seedling	Col	not applic	27	not applic	12 h	10µM prol	1.0126	1.0616	0.8398
39	substance	seedling	Col	not applic	30	not applic	3 h	10µM unic	0.9759	0.9414	0.8385
40	substance	seedling	Col	not applic	29	not applic	12 h	10µM unic	0.7532	0.9416	0.9965
41	substance	seedling	Col	not applic	12	not applic	3 h	H2O	1.1227	0.7882	0.5243
42	substance	seedling	Col	not applic	3	not applic	3 h	10µM 2,4,	0.9493	0.8166	0.7991
43	substance	seedling	Col	not applic	6	not applic	3 h	10µM AVG	0.9519	0.7446	0.6014
44	substance	seedling	Col	not applic	7	not applic	3 h	10µM AGI	0.4448	0.9138	0.5666
45	substance	seedling	Col	not applic	8	not applic	3 h	10µM B9 (	0.8648	1.0083	0.5851
46	substance	seedling	Col	not applic	10	not applic	3 h	3µM Brz2(	0.7728	0.7341	0.4618
47	substance	seedling	Col	not applic	9	not applic	3 h	10µM cycl	0.1353	0.4823	0.092
48	substance	seedling	Col	not applic	18	not applic	3 h	1µM ibupr	0.6825	0.9715	0.5035
49	substance	seedling	Col	not applic	13	not applic	3 h	10µM MG(	1.2209	0.9514	0.7179
50	substance	seedling	Col	not applic	14	not applic	3 h	10µM NPA	0.8367	0.6827	0.8477
51	substance	seedling	Col	not applic	15	not applic	3 h	10µM PCI	1.0624	0.613	0.7238
52	substance	seedling	Col	not applic	11	not applic	3 h	10µM PNC	0.7317	0.6103	0.1696
53	substance	seedling	Col	not applic	16	not applic	3 h	10µM sali	0.8018	0.9089	0.7867
54	substance	seedling	Col	not applic	17	not applic	3 h	10µM TIBA	0.8007	1.2659	0.8857
55	det2_subs	seedling	det2-1	not applic	43	not applic	3 h	H2O	0.9602	0.9477	1.5993
56	det2_subs	seedling	det2-1	not applic	31	not applic	3 h	1µM 3-del	1.0011	0.9782	0.9645
57	det2_subs	seedling	det2-1	not applic	32	not applic	3 h	1µM 3-del	1.1896	0.9128	1.0027
58	det2_subs	seedling	det2-1	not applic	33	not applic	3 h	1µM 6-de	0.947	0.9289	1.1651
59	det2_subs	seedling	det2-1	not applic	34	not applic	3 h	1µM 6-de	1.3832	0.9107	0.9409
60	det2_subs	seedling	det2-1	not applic	35	not applic	3 h	1µM 6-de	1.4323	0.913	0.8369
61	det2_subs	seedling	det2-1	not applic	36	not applic	3 h	1µM 6-de	1.4478	0.8762	0.9981
62	det2_subs	seedling	det2-1	not applic	37	not applic	3 h	10nM bra	0.7994	0.7834	0.8419
63	det2_subs	seedling	det2-1	not applic	38	not applic	3 h	10µM cam	1.4323	0.8513	1.1588
64	det2_subs	seedling	det2-1	not applic	39	not applic	3 h	100nM ca	0.9711	0.73	0.9278
65	det2_subs	seedling	det2-1	not applic	40	not applic	3 h	1µM catha	1.2764	0.8748	0.8641
66	det2_subs	seedling	det2-1	not applic	41	not applic	3 h	1µM teast	1.1031	0.8595	1.1918

67	det2_subs	seedling	det2-1	not applic	42	not applic	3 h	1µM typh	1.1058	0.9374	0.6218
68	det2-1 / E	seedling	det2-1	not applic	77	not applic	0.5 h	H2O	1.0311	1.0812	0.761
69	det2-1 / E	seedling	det2-1	not applic	75	not applic	1 h	H2O	1.1529	0.9175	0.8846
70	det2-1 / E	seedling	det2-1	not applic	76	not applic	3 h	H2O	0.9616	0.7378	0.7396
71	det2-1 / E	seedling	det2-1	not applic	74	not applic	0.5 h	10nM BL	0.8613	0.7084	0.7744
72	det2-1 / E	seedling	det2-1	not applic	72	not applic	1 h	10nM BL	1.1374	0.7951	0.7992
73	det2-1 / E	seedling	det2-1	not applic	73	not applic	3 h	10nM BL	0.9473	0.7338	0.3324
74	ga1-5 / G	seedling	ga1-5	not applic	63	not applic	0.5 h	H2O	1.282	1.024	0.7945
75	ga1-5 / G	seedling	ga1-5	not applic	61	not applic	1 h	H2O	1.0855	0.7134	0.3074
76	ga1-5 / G	seedling	ga1-5	not applic	62	not applic	3 h	H2O	1.1045	0.9669	0.4189
77	ga1-5 / G	seedling	ga1-5	not applic	60	not applic	0.5 h	1µM GA	1.0483	0.8484	0.541
78	ga1-5 / G	seedling	ga1-5	not applic	58	not applic	1 h	1µM GA	0.8071	1.0816	0.3026
79	ga1-5 / G	seedling	ga1-5	not applic	59	not applic	3 h	1µM GA	1.1277	0.7827	0.3711
80	ga1-3 / G	seeds	ga1-3	not applic	90	not applic	3 h	control	1.9385	1.8073	2.0248
81	ga1-3 / G	seeds	ga1-3	not applic	91	not applic	6 h	control	1.8155	1.9104	1.7946
82	ga1-3 / G	seeds	ga1-3	not applic	92	not applic	9 h	control	2.6404	1.9266	1.8037
83	ga1-3 / G	seeds	ga1-3	not applic	93	not applic	3 h	5µm GA	1.8288	2.2408	1.9227
84	ga1-3 / G	seeds	ga1-3	not applic	94	not applic	6 h	5µm GA	2.4198	1.7943	1.6115
85	ga1-3 / G	seeds	ga1-3	not applic	95	not applic	9 h	5µm GA	1.6367	2.0782	2.4423
86	hormone	seedling	Ler	not applic	45	not applic	No h	No	0.6975	0.8134	0.6143
87	hormone	seedling	ga1-5 (Le	not applic	44	not applic	No h	No	0.799	0.8885	0.2356
88	hormone	seedling	Ws	not applic	46	not applic	No h	No	0.6648	1.096	0.8206
89	hormone	seedling	bri1	not applic	47	not applic	No h	No	0.7735	0.9454	0.3907
90	zeatin	seedling	Col	not applic	85	not applic	3 h	control	0.9597	0.482	0.3243
91	zeatin	seedling	ARR21Co	not applic	87	not applic	3 h	control	0.8755	0.6395	0.2314
92	zeatin	seedling	Col	not applic	84	not applic	3 h	control	0.9553	0.4865	0.2771
93	zeatin	seedling	Col	not applic	86	not applic	3 h	20µm zea	1.0191	0.6037	0.5231
94	zeatin	seedling	ARR22ox	not applic	88	not applic	3 h	control	0.9812	0.3574	0.4708
95	zeatin	seedling	ARR22ox	not applic	89	not applic	3 h	20µm zea	0.7635	0.7453	0.3725
96	sulfate sta	seedling	Col	not applic	98	not applic	2 h	0µm sulfa	0.6075	1.1629	0.9134
97	sulfate sta	seedling	Col	not applic	99	not applic	4 h	0µm sulfa	0.6671	1.3488	1.2026
98	sulfate sta	seedling	Col	not applic	100	not applic	8 h	0µm sulfa	0.7108	1.3613	1.0104
99	sulfate sta	seedling	Col	not applic	96	not applic	12 h	0µm sulfa	0.6913	1.2269	1.0929
100	sulfate sta	seedling	Col	not applic	97	not applic	24 h	0µm sulfa	0.6428	1.2121	1.1873

101	sulfate sta	seedling	Col	not applic	101	not applic	0 h	1500µm s	0.7995	0.9645	1.1389
102	sulfate sta	seedling	Col	not applic	104	not applic	2 h	1500µm s	0.6396	1.2628	1.0016
103	sulfate sta	seedling	Col	not applic	105	not applic	4 h	1500µm s	0.7692	1.1912	1.0168
104	sulfate sta	seedling	Col	not applic	106	not applic	8 h	1500µm s	0.323	1.0615	1.2022
105	sulfate sta	seedling	Col	not applic	102	not applic	12 h	1500µm s	0.5596	1.2267	1.1319
106	sulfate sta	seedling	Col	not applic	103	not applic	24 h	1500µm s	0.543	1.0687	1.2124
107	ABA	seeds	Col	not applic	109	not applic	0 h	No	0.7515	1.8056	7.604
108	ABA	seeds	Col	not applic	110	not applic	24 h	No	2.1257	2.3593	1.6857
109	ABA	seeds	Col	not applic	107	not applic	24 h	3µm ABA	2.206	2.7316	2.0072
110	ABA	seeds	Col	not applic	108	not applic	24 h	30µm ABA	1.7923	2.0832	1.9814
111	seed imbi	seeds	Col	not applic	111	not applic	0 h	imbibition	0.421	1.8705	7.6272
112	seed imbi	seeds	Col	not applic	112	not applic	1 h	imbibition	0.6102	2.0401	6.4705
113	seed imbi	seeds	Col	not applic	113	not applic	3 h	imbibition	0.4937	4.0883	3.7843
114	temperatu	seeds	Col	not applic	115	not applic	96 h	4°C	2.7795	2.9922	1.8741
115	temperatu	seeds	Col	not applic	114	not applic	96 h	22°C	1.0379	1.1183	2.4178

Microarray sample key - biotic stress and light conditions

sample #	tissue cl	tissue	genotype	age	NASC sa	stress	time	treatmer	CstF50	CstF64	CstF77
1	Pseudomc	leaf	not applic	not applic	17	not applic	0 h	untreated	1.3941	0.8814	0.9385
2	Pseudomc	leaf	not applic	not applic	33	not applic	2 h	infiltrated	0.7557	0.8986	1.1208
3	Pseudomc	leaf	not applic	not applic	34	not applic	6 h	infiltrated	1.0942	1.016	0.809
4	Pseudomc	leaf	not applic	not applic	35	not applic	24 h	infiltrated	1.0963	0.9459	1.1476
5	Pseudomc	leaf	not applic	not applic	21	not applic	2 h	infiltrated	0.2905	1.2349	0.8834
6	Pseudomc	leaf	not applic	not applic	22	not applic	6 h	infiltrated	1.0816	1.3411	0.671
7	Pseudomc	leaf	not applic	not applic	23	not applic	24 h	infiltrated	0.2738	1.3261	0.739
8	Pseudomc	leaf	not applic	not applic	24	not applic	2 h	infiltrated	0.4475	1.1038	0.6512
9	Pseudomc	leaf	not applic	not applic	25	not applic	6 h	infiltrated	0.5255	1.2158	0.7847
10	Pseudomc	leaf	not applic	not applic	26	not applic	24 h	infiltrated	0.7672	0.9557	0.8693
11	Pseudomc	leaf	not applic	not applic	27	not applic	2 h	infiltrated	0.6863	1.2183	0.7667
12	Pseudomc	leaf	not applic	not applic	28	not applic	6 h	infiltrated	1.3057	1.2484	0.5995
13	Pseudomc	leaf	not applic	not applic	29	not applic	24 h	infiltrated	1.2928	1.1532	1.1113
14	Pseudomc	leaf	not applic	not applic	30	not applic	2 h	infiltrated	0.4516	1.1999	1.0136
15	Pseudomc	leaf	not applic	not applic	31	not applic	6 h	infiltrated	0.7608	1.4417	0.6715
16	Pseudomc	leaf	not applic	not applic	32	not applic	24 h	infiltrated	0.8363	1.0244	1.3496
17	Phytophth	leaf	not applic	not applic	C.6	not applic	6 h	for the tre	2.3652	0.6655	1.016
18	Phytophth	leaf	not applic	not applic	C.12	not applic	12 h	for the tre	1.7199	0.8438	1.1546
19	Phytophth	leaf	not applic	not applic	C.24	not applic	24 h	for the tre	2.0759	0.6804	0.7056
20	Phytophth	leaf	not applic	not applic	Pi.6	not applic	6 h	for the tre	0.7729	1.0008	0.7322
21	Phytophth	leaf	not applic	not applic	Pi.12	not applic	12 h	for the tre	1.5407	0.8547	1.1354
22	Phytophth	leaf	not applic	not applic	Pi.24	not applic	24 h	for the tre	1.7386	0.7763	1.1465
23	elicitors	leaf	not applic	not applic	1,1	not applic	1 h	infiltrated	0.6197	0.3901	1.3763
24	elicitors	leaf	not applic	not applic	1,4	not applic	4 h	infiltrated	1.5211	1.0784	0.9347
25	elicitors	leaf	not applic	not applic	2,1	not applic	1 h	infiltrated	0.6391	0.489	1.2083
26	elicitors	leaf	not applic	not applic	2,4	not applic	4 h	infiltrated	1.7347	1.1722	1.106
27	elicitors	leaf	not applic	not applic	3,1	not applic	1 h	infiltrated	0.619	0.488	1.3031
28	elicitors	leaf	not applic	not applic	3,4	not applic	4 h	infiltrated	1.415	1.3446	1.0429
29	elicitors	leaf	not applic	not applic	4,1	not applic	1 h	infiltrated	0.6068	0.519	1.3928
30	elicitors	leaf	not applic	not applic	4,4	not applic	4 h	infiltrated	0.836	1.6986	0.9305
31	elicitors	leaf	not applic	not applic	5,1	not applic	1 h	infiltrated	0.5254	0.4544	1.387
32	elicitors	leaf	not applic	not applic	5,4	not applic	4 h	infiltrated	0.7279	1.5986	0.9612

33	elicitors	leaf	not applic	not applic	6,1	not applic	1 h	infiltrated	0.5243	0.5128	1.1685
34	elicitors	leaf	not applic	not applic	6,4	not applic	4 h	infiltrated	0.7932	1.5314	0.9742
35	elicitors	leaf	not applic	not applic	7,1	not applic	1 h	infiltrated	0.5419	0.5533	1.1992
36	elicitors	leaf	not applic	not applic	7,4	not applic	4 h	infiltrated	1.6245	1.1444	0.9998
37	darkness	hypocotyl	not applic	not applic	DS	not applic	45 min	light treat	1.0676	1.1553	1.0322
38	darkness	hypocotyl	not applic	not applic	DL	not applic	240 min	light treat	0.9748	1.0094	1.0099
39	UV-A	hypocotyl	not applic	not applic	AS	not applic	45 min	light treat	0.9861	0.9994	1.0603
40	UV-A	hypocotyl	not applic	not applic	AL	not applic	240 min	light treat	0.9828	0.9943	1.0694
41	UV-A/B	hypocotyl	not applic	not applic	US	not applic	45 min	light treat	0.9777	1.0031	1.008
42	UV-A/B	hypocotyl	not applic	not applic	UL	not applic	240 min	light treat	0.9786	1.0405	1.0295
43	blue light	hypocotyl	not applic	not applic	BS	not applic	45 min	light treat	1.0069	1.0797	0.9949
44	blue light	hypocotyl	not applic	not applic	BL	not applic	240 min	light treat	0.9671	0.9881	1.0127
45	red (pulse	hypocotyl	not applic	not applic	PS	not applic	45 min	light treat	1.0287	1.048	1.0175
46	red (pulse	hypocotyl	not applic	not applic	PL	not applic	240 min	light treat	0.9798	1.043	0.9523
47	red light	hypocotyl	not applic	not applic	RS	not applic	45 min	light treat	1.0003	0.9038	0.9928
48	red light	hypocotyl	not applic	not applic	RL	not applic	240 min	light treat	0.9806	0.8655	0.9266
49	far-red lig	hypocotyl	not applic	not applic	FS	not applic	45 min	light treat	0.9851	0.9341	0.9769
50	far-red lig	hypocotyl	not applic	not applic	FL	not applic	240 min	light treat	0.9931	0.953	0.8484
51	white	hypocotyl	not applic	not applic	WS	not applic	45 min	light treat	0.986	0.9693	1.0481
52	white	hypocotyl	not applic	not applic	WL	not applic	240 min	light treat	1.1056	1.0145	1.0213

References

Addepalli B, Hunt AG (2007) A novel endonuclease activity associated with the Arabidopsis ortholog of the 30-kDa subunit of cleavage and polyadenylation specificity factor. Nucl. Acids Res.**:** gkm457

Addepalli B, Meeks LR, Forbes KP, Hunt AG (2004) Novel alternative splicing of mRNAs encoding poly(A) polymerases in Arabidopsis. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression **1679:** 117-128

An G, Ebert PR, Mitra A, Ha SB (1988) Binary Vectors. *In* SB Gelvin., Schilperoort, R., and Verma, D. P. , ed, **Plant Molecular Biology Manual** Vol A3. Kluwer Academic Pub., Dordrecht, Belgium, pp 1-19

Aranda A, Perez-Ortin JE, Moore C, del Olmo ML (1998) Transcription termination downstream of the Saccharomyces cerevisiae FBP1 [changed from FPB1] poly(A) site does not depend on efficient 3'end processing. RNA **4:** 303-318

Aranda A, Proudfoot N (2001) Transcriptional Termination Factors for RNA Polymerase II in Yeast. Molecular Cell **7:** 1003-1011

Audibert A, Simonelig M (1998) Autoregulation at the level of mRNA 3' end formation of the suppressor of forked gene of *Drosophila melanogaster* is conserved in *Drosophila virilis*. Proceedings of the National Academy of Sciences of the United States of America **95:** 14302-14307

Bai Y, Auperin TC, Chou C-Y, Chang G-G, Manley JL, Tong L (2007) Crystal Structure of Murine CstF-77: Dimeric Association and Implications for Polyadenylation of mRNA Precursors. Molecular Cell **25:** 863-875

Ballantyne S, Bilger A, Astrom J, Virtanen A, Wickens M (1995) Poly (A) polymerases in the nucleus and cytoplasm of frog oocytes: dynamic changes during oocyte maturation and early development. RNA **1:** 64-78

Barabino SM, Hubner W, Jenny A, Minvielle-Sebastia L, Keller W (1997) The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. Genes Dev. **11:** 1703-1716

Beelman CA, Parker R (1995) Degradation of mRNA in eukaryotes. Cell **81:** 179-183

Benoit B, Juge F, Iral F, Audibert A, Simonelig M (2002) Chimeric human CstF-77/Drosophila Suppressor of forked proteins rescue suppressor of forked mutant lethality and mRNA 3' end processing in Drosophila. PNAS **99:** 10593-10598

Berry M, Sachar RC (1981) Hormonal regulation of poly(A) polymerase activity by gibberellic acid in embryo-less half-seeds of wheat (Triticum aestivum). FEBS Letters **132:** 109-113

Berry M, Sachar RC (1982) Expression of conserved message of poly(A) polymerase through hormonal control in wheat aleurone layers. FEBS Letters **141:** 164-168

Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ (1998) Coupling Termination of Transcription to Messenger RNA Maturation in Yeast. Science **280:** 298-301

Bonneaud N, Minvielle-Sebastia L, Cullin C, Lacroute F (1994) Cellular localization of RNA14p and RNA15p, two yeast proteins involved in mRNA stability. J Cell Sci **107:** 913-921

Brodsky AS, Silver PA (2000) Pre-mRNA processing factors are required for nuclear export. RNA **6:** 1737-1749

Burkard G, Keller EB (1974) Poly(A) Polymerase and Poly(G) Polymerase in Wheat Chloroplasts. Proceedings of the National Academy of Sciences of the United States of America **71:** 389-393

Calvo O, Manley JL (2001) Evolutionarily Conserved Interaction between CstF-64 and PC4 Links Transcription, Polyadenylation, and Termination. Molecular Cell **7:** 1013-1023

Cao G-j, Pogliano J, Sarkar N (1996) Identification of the coding region for a second poly(A) polymerase in Escherichia coli. PNAS **93:** 11580-11585

Cao G, Sarkar N (1992) Poly(A) RNA in Escherichia coli: Nucleotide Sequence at the Junction of the lpp Transcript and the Polyadenylate Moiety. PNAS **89:** 7546-7550

Chen J, Moore C (1992) Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. Mol. Cell. Biol. **12:** 3470-3481

Christofori G, Keller W (1988) 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor, and a snRNP. Cell **54:** 875-889

Christofori G, Keller W (1989) Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre-mRNA in vitro. Mol. Cell. Biol. **9:** 193-203

Colgan DF, Murthy KG, Zhao W, Prives C, JL. M (1998) Inhibition of poly(A) polymerase requires p34cdc2/cyclin B phosphorylation of multiple consensus and non-consensus sites. EMBO J. **17:** 1053-1062

Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. Genes & Dev. **11:** 2755-2766

Colgan DF, Murthy KGK, Prives C, Manley JL (1996) Cell-cycle related regulation of poly(A) polymerase by phosphorylation. Nature **384:** 282-285

Cooke C, Hans H, Alwine JC (1999) Utilization of Splicing Elements and Polyadenylation Signal Elements in the Coupling of Polyadenylation and Last-Intron Removal. Mol Cell Biol **19:** 4971-4979

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. Nucl. Acids Res. **32:** D575-577

D'Alessandro M, Srivastava BIS (1985) Poly(A) polymerase and poly(ADP-ribose) polymerase activities in normal and crown gall tumor tissue cultures of tobacco. FEBS Letters **188:** 239-242

Danckwardt S, Kaufmann I, Gentzel M, Foerstner K U, Gantzert A S, Gehring N H, Neu-Yilik G, Bork P, Keller W, Wilm M, Hentze M W and Kulozik A E (2007) Splicing factors stimulate polyadenylation via USEs at non-canonical 3' end formation signals. EMBO J **26:** 2658–2669

Dasgupta J, Li Q-S, Brian Thomson A, Hunt AG (1998) Characterization of a cDNA encoding a novel plant poly(A) polymerase. Plant Molecular Biology **37:** 729-734

Dasgupta J, Li Q, Thomson AB, Hunt AG (1995) Characterization of a novel plant poly(A) polymerase. Plant Science **110:** 215-226

Dass B, McMahon KW, Jenkins NA, Gilbert DJ, Copeland NG, MacDonald CC (2001) The Gene for a Variant Form of the Polyadenylation Protein CstF-64 Is on Chromosome 19 and Is Expressed in Pachytene Spermatocytes in Mice. J. Biol. Chem. **276:** 8044-8050

Davila Lopez M, Samuelsson T (2008) Early evolution of histone mRNA 3' end processing. RNA **14:** 1-10

del Olmo M, Mizrahi N, Gross S, Moore CL (1997) The Uba2 and Ufd1 proteins of Saccharomyces cerevisiae interact with poly(A) polymerase and affect the polyadenylation activity of cell extracts. Molecular and General Genetics MGG **255:** 209-218

Delaney KJ, Xu R, Zhang J, Li QQ, Yun K-Y, Falcone DL, Hunt AG (2006) Calmodulin Interacts with and Regulates the RNA-Binding Activity of an Arabidopsis Polyadenylation Factor Subunit. Plant Physiol **140:** 1507-1521

Dickson KS, Thompson SR, Gray NK, Wickens M (2001) Poly(A) Polymerase and the Regulation of Cytoplasmic Polyadenylation. J. Biol. Chem. **276:** 41810-41816

Dohmen RJ, Stappen R, McGrath JP, Forrov H, Kolarov J, Goffeau A, Varshavsky A (1995) An Essential Yeast Gene Encoding a Homolog of Ubiquitin-activating Enzyme. J. Biol. Chem. **270:** 18099-18109

Doma MK, Parker R (2007) RNA Quality Control in Eukaryotes. Cell **131:** 660-668

Dreyfus M, Régnier P (2002) The Poly(A) Tail of mRNAs: Bodyguard in Eukaryotes, Scavenger in Bacteria. Cell **111:** 611-613

Duvel K, Valerius O, Mangus DA, Jacobson A, Braus GH (2002) Replacement of the yeast TRP4 3' untranslated region by a hammerhead ribozyme results in a stable and efficiently exported mRNA that lacks a poly(A) tail. RNA **8:** 336-344

Edmonds M (1982) Poly (A) adding enzymes. *In* P Boyer, ed, **The enzymes**, Vol XV. Academic Press, Inc, New York, pp 217-244

Edmonds M, Abrams R (1960) Polynucleotide Biosynthesis: Formation of a Sequence of Adenylate Units from Adenosine Triphosphate by an Enzyme from Thymus Nuclei. J. Biol. Chem **235:** 1142-1149

Edmonds M, Winters MA, Waldo EC (1976) Polyadenylate Polymerases. Progress in Nucleic Acid Research and Molecular Biology **Volume 17:** 149-179

Edwalds-Gilbert G, Milcarek C (1995) Regulation of poly(A) site use during mouse B-cell development involves a change in the binding of a general polyadenylation factor in a B-cell stage-specific manner. Mol. Cell. Biol. **15:** 6420-6429

Elliott BJ, Dattaroy T, Meeks-Midkiff LR, Forbes KP, Hunt AG (2003) An interaction between an Arabidopsis poly(A) polymerase and a homologue of the 100 kDa subunit of CPSF. Plant Molecular Biology **51:** 373-384

Forbes KP (2005) Characterization of plant polyadenylation trans-acting factors-factors that modify poly(A)polymerse activity. PhD Dissertation. University of Kentucky, Lexington

Forbes KP, Addepalli B, Hunt AG (2006) An Arabidopsis Fip1 Homolog Interacts with RNA and Provides Conceptual Links with a Number of Other Polyadenylation Factor Subunits. J Biol Chem **281:** 176-186

Gebauer F, Richter JD (1995) Cloning and characterization of a Xenopus poly(A) polymerase [published erratum appears in Mol Cell Biol 1995 Jun;15(6):3460]. Mol. Cell. Biol. **15:** 1422-1430

Gietz D, Jean AS, Woods RA, Schiestl RH (1992) Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res **20:** 1425

Gilmartin GM, Nevins JR (1991) Molecular analyses of two poly(A) site-processing factors that determine the recognition and efficiency of cleavage of the pre-mRNA. Mol. Cell. Biol. **11:** 2432-2438

Glover-Cutter K, Kim S, Espinosa J, Bentley DL (2008) RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat Struct Mol Biol **15:** 71-78

Goodin M M, Dietzgen R G, Schichnes D, Ruzin S, Jackson A O (2002) pGD vectors: versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves. The Plant Journal **31:** 375-383

Gross S, Moore CL (2001) Rna15 Interaction with the A-Rich Yeast Polyadenylation Signal Is an Essential Step in mRNA 3'-End Formation. Mol. Cell. Biol. **21:** 8045-8055

Gunderson SI, Vagner S, Polycarpou-Schwarz M, Mattaj IW (1997) Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. Genes Dev. **11:** 761-773

Guo Z, Sherman F (1996) 3'-end-forming signals of yeast mRNA. Trends in Biochemical Sciences **21:** 477-481

Hake LE, Richter JD (1994) CPEB is a specificity factor that mediates cytoplasmic polyadenylation during Xenopus oocyte maturation. Cell **79:** 617-627

Hammell CM, Gross S, Zenklusen D, Heath CV, Stutz F, Moore C, Cole CN (2002) Coupling of Termination, 3' Processing, and mRNA Export. Mol Cell Biol **22:** 6441-6457

Hatton LS, Eloranta JJ, Figueiredo LM, Takagaki Y, Manley JL, O'Hare K (2000) The Drosophila homologue of the 64 kDa subunit of cleavage stimulation factor interacts with the 77 kDa subunit encoded by the suppressor of forked gene. Nucl. Acids Res. **28:** 520-526

Helmling S, Zhelkovsky A, Moore CL (2001) Fip1 Regulates the Activity of Poly(A) Polymerase through Multiple Interactions. Mol. Cell. Biol. **21:** 2026-2037

Huang Y, Carmichael GC (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. Mol Cell Biol **16:** 1534-1542

Huber Z, Monarez RR, Dass B, MacDonald CC (2005) The mRNA Encoding {tau}CstF-64 Is Expressed Ubiquitously in Mouse Tissues. Ann NY Acad Sci **1061:** 163-172

Hunt A, Xu R, Addepalli B, Rao S, Forbes K, Meeks L, Xing D, Mo M, Zhao H, Bandyopadhyay A, Dampanaboina L, Marion A, Von Lanken C, Li Q (2008) Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. BMC Genomics **9:** 220

Hunt AG (1994) Messenger RNA3' End Formation in Plants. Annual Review of Plant Physiology and Plant Molecular Biology **45:** 47-60

Hunt AG, Addepalli B (2008) The Interaction Between Two Arabidopsis Polyadenylation Factor Subunits Involves an Evolutionarily-Conserved Motif and Has Implications for the Assembly and Function of the Polyadenylation Complex. Protein and Peptide Letters **15:** 76-88

Hunt AG, Meeks LR, Forbes KP, Das Gupta J, Mogen BD (2000) Nuclear and Chloroplast Poly(A) Polymerases from Plants Share a Novel Biochemical Property. Biochemical and Biophysical Research Communications **272:** 174-181

Jacob S, Janne O, Rose K (1975) Cortico steroid induced changes in RNA poly(A) polymerase and chromatin activity in rat liver. *In* G Talwar, ed, **Regulation of growth and differentiation function in eukaryotic cells**. Raven Press, New York, pp 369-370

James P, Halladay J, Craig EA (1996) Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selection in Yeast. Genetics **144:** 1425-1436

Johnson ES, Ma PCM, Ota IM, Varshavsky A (1995) A Proteolytic Pathway That Recognizes Ubiquitin as a Degradation Signal. J. Biol. Chem. **270:** 17442-17456

Juge F, Audibert A, Benoit B, Simonelig M (2000) Tissue-specific autoregulation of Drosophila suppressor of forked by alternative poly(A) site utilization leads to accumulation of the suppressor of forked protein in mitotically active cells. RNA **6:** 1529-1538

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:** 6603-6613

Kaufmann I, Martin G, Friedlein A, Langen H, Keller W (2004) Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. EMBO **23:** 616-626

Keller W, Minvielle-Sebastia L (1997) A comparison of mammalian and yeast pre-mRNA 3'-end processing. Current Opinion in Cell Biology **9:** 329-336

Kerwitz Y, Kühn U, Lilie H, Knoth A, Scheuermann T, Friedrich H, Schwarz E, Wahle E (2003) Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. EMBO J. **22:** 3705-3714

Kessler M, Henry M, Shen E, Zhao J, Gross S, Silver P, Moore C (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**

Kessler MM, Zhao J, Moore CL (1996) Purification of the Saccharomyces cerevisiae Cleavage/Polyadenylation Factor I. . J Biol Chem **271:** 27167-27175

Kessler MM, Zhellcovsky Alexander M., Skvorak A, Moore Claire L. (1995) Monoclonal Antibodies to Yeast Poly(A) Polymerase (PAP) Provide Evidence for association of PAP with Cleavage Factor I. **Biochemistry 34:** 1750-1759

Kim H, Lee Y (2001) Interaction of Poly(A) Polymerase with the 25-kDa Subunit of Cleavage Factor I. Biochemical and Biophysical Research Communications **289:** 513-518

Kleiman FE, Manley JL (1999) Functional Interaction of BRCA1-Associated BARD1 with Polyadenylation Factor CstF-50. Science **285:** 1576-1579

Kleiman FE, Manley JL (2001) The BARD1-CstF-50 Interaction Links mRNA 3' End Formation to DNA Damage and Tumor Suppression. Cell **104:** 743-753

Lakhani S, Sachar RC (1985) Hormonal regulation of poly(A)+ RNA and poly(A) polymerase activity in wheat embryos. Plant Science **42:** 191-200

Legrand P, Pinaud N, Minvielle-Sebastia L, Fribourg S (2007) The structure of the CstF-77 homodimer provides insights into CstF assembly. Nucl. Acids Res. **35:** 4515-4522

Lei EP, Silver PA (2002) Intron status and 3'-end formation control cotranscriptional export of mRNA. Genes & Development **16:** 2761-2766

Lei EP, Silver PA (2002) Protein and RNA Export from the Nucleus. Developmental Cell **2:** 261-272

Li Q, Hunt AG (1997) The Polyadenylation of RNA in Plants. Plant Physiol **115:** 321-325

Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH (2002) Interactions between mRNA Export Commitment, 3'-End Quality Control, and Nuclear Degradation. Mol. Cell. Biol. **22:** 8254-8266

Lingner J, Kellerman J (1991) Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. Nature **354:** 496

Lutz CS (2008) Alternative Polyadenylation: A Twist on mRNA 3' End Formation. ACS Chemical Biology **3:** 609-617

Lykke-Andersen S, Jensen TH (2006) CUT it out: silencing of noise in the transcriptome. Nat Struct Mol Biol **13:** 860-861

MacDonald CC, 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:** 6647-6654

Manley JL (1995) A complex protein assembly catalyzes polyadenylation of mRNA precursors. Current Opinion in Genetics & Development **5:** 222-228

Mans RJ, Huff NJ (1975) Utilization of ribonucleic acid and deoxyoligomer primers for polyadenylic acid synthesis by adenosine triphosphate: polynucleotidylexotransferase from maize. J. Biol. Chem **250:** 3672-3678

Martin G, Keller, W. and Doublie, S. (2000) Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. EMBO J. **19:** 4193–4203.

Martincic K, Campbell R, Edwalds-Gilbert G, Souan L, Lotze MT, Milcarek C (1998) Increase in the 64-kDa subunit of the polyadenylation/cleavage stimulatory factor during the G_0 to S phase transition. Proceedings of the National Academy of Sciences of the United States of America **95:** 11095-11100

Martins CRF, Johnson JA, Lawrence DM, Choi T-J, Pisi A-M, Tobin SL, Lapidus D, Wagner JDO, Ruzin S, McDonald K, Jackson AO (1998) Sonchus Yellow Net Rhabdovirus Nuclear Viroplasms Contain Polymerase-Associated Proteins. J. Virol. **72:** 5669-5679

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson S, Wickens M, Bentley D (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature **385:** 357-361

Meeks LR (2005) Isolation and Characterization of the Four Arabidopsis thaliana Poly(A) Polymerase Genes. PhD dissertation. University of Kentucky, Lexington

Meinhart A, Kamenski T, Hoeppner S, Baumli S, Cramer P (2005) A structural perspective of CTD function. Genes & Development **19:** 1401-1415

Mellman DL, Gonzales ML, Song C, Barlow CA, Wang P, Kendziorski C, Anderson RA (2008) A PtdIns4,5P2-regulated nuclear poly(A) polymerase controls expression of select mRNAs. Nature **451:** 1013-1017

Millevoi S, Loulergue C, Dettwiler S, Zeïneb Karaa S, Keller W, Antoniou M and Vagner S (2006) An interaction between U2AF 65 and CF Im links the splicing and 3' end processing machineries. EMBO J. **25:** 4854–4864

Minvielle-Sebastia L, Keller W (1999) mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. Current Opinion in Cell Biology **11:** 352-357

Minvielle-Sebastia L, Preker PJ, Keller W (1994) RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. Science **266:** 1702-1705

Minvielle-Sebastia L, Winsor B, Bonneaud N, Lacroute F (1991) Mutations in the yeast RNA14 and RNA15 genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the RNA15 protein. Mol. Cell. Biol. **11:** 3075-3087

Mitchelson A, Simonelig M, Williams C, O'Hare K (1993) Homology with Saccharomyces cerevisiae RNA14 suggests that phenotypic suppression in Drosophila melanogaster by suppressor of forked occurs at the level of RNA stability. Genes & Development **7:** 241-249

Mizrahi N, Moore C (2000) Posttranslational Phosphorylation and Ubiquitination of the Saccharomyces cerevisiae Poly(A) Polymerase at the S/G2 Stage of the Cell Cycle. Mol. Cell. Biol. **20:** 2794-2802

Monarez RR, MacDonald CC, Dass B (2007) Polyadenylation proteins CstF-64 and tau CstF-64 exhibit differential binding affinities for RNA polymers. Biochem. J **401:** 651-658

Moore CL CJ, Whoriskey J. (1988) Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding. EMBO J. **7:** 3159-3169

Murthy KG, Manley JL (1992) Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. J. Biol. Chem. **267:** 14804-14811

Murthy KG, Manley JL (1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. Genes & Development **9:** 2672-2683

Nag A, Narsinh K, Martinson HG (2007) The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat Struct Mol Biol **14:** 662-669

Nagaike T, Suzuki T, Ueda T (2008) Polyadenylation in mammalian mitochondria: Insights from recent studies. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1779:** 266-269

Niessing J, Sekeris CE (1974) Evidence for the existence of two different poly(A) polymerases and a novel ribohomopolymer polymerase in rat liver nuclei. Biochemical and Biophysical Research Communications **60:** 673-680

Niwa M, MacDonald CC, Berget SM (1992) Are vertebrate exons scanned during splice-site selection? Nature **360:** 277-280

Noble CG, Walker PA, Calder LJ, Taylor IA (2004) Rna14-Rna15 assembly mediates the RNA-binding capability of Saccharomyces cerevisiae cleavage factor IA. Nucl. Acids Res. **32:** 3364-3375

Orava MM, Isomaa VV, Janne OA (1979) Nuclear Poly(A) Polymerase Activities in the Rabbit Uterus. European Journal of Biochemistry **101:** 195-203

Pan Z, Zhang H, Hague LK, Lee JY, Lutz CS, Tian B (2006) An intronic polyadenylation site in human and mouse CstF-77 genes suggests an evolutionarily conserved regulatory mechanism. Gene **366:** 325-334

Parker R, Song H (2004) The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol **11:** 121-127

Pérez Cañadillas JM VG (2003) Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. EMBO J. **22:** 2821-2830

Perumal K, Sinha K, Henning D, Reddy R (2001) Purification, Characterization, and Cloning of the cDNA of Human Signal Recognition Particle RNA 3'-Adenylating Enzyme. J. Biol. Chem. **276:** 21791-21796

Preker PJ, Keller W (1998) The HAT helix, a repetitive motif implicated in RNA processing. Trends in Biochemical Sciences **23:** 15-16

Preker PJ, 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:** 379-389

Proudfoot N (2004) New perspectives on connecting messenger RNA 3' end formation to transcription. Current Opinion in Cell Biology **16:** 272-278

Proudfoot N, O'Sullivan J (2002) Polyadenylation: A tail of two complexes. Current Biology **12:** R855-R857

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA Processing with Transcription. Cell **108:** 501-512

Proweller A, Butler S (1994) Efficient translation of poly(A)-deficient mRNAs in Saccharomyces cerevisiae. Genes & Dev. **8:** 2629-2640

Qu X, Perez-Canadillas J-M, Agrawal S, De Baecke J, Cheng H, Varani G, Moore C (2007) The C-terminal Domains of Vertebrate CstF-64 and Its Yeast Orthologue Rna15 Form a New Structure Critical for mRNA 3'-End Processing. J. Biol. Chem **282:** 2101-2115

Raabe T, Bollum FJ, Manley JL (1991) Primary structure and expression of bovine poly(A) polymerase. Nature **353:** 229-234

Raabe T, Murthy KG, Manley JL (1994) Poly(A) polymerase contains multiple functional domains. Mol. Cell. Biol. **14:** 2946-2957

Raju VS, Reddy PRK (1983) Inhibition of polyadenylation of mRNA by gonadotropin releasing hormone. Biochemical and Biophysical Research Communications **115:** 451-455

Richardson JM, McMahon KW, MacDonald CC, Makhatadze GI (1999) MEARA Sequence Repeat of Human CstF-64 Polyadenylation Factor Is Helical in Solution: A Spectroscopic and Calorimetric Study. Biochemistry **38:** 12869-12875

Richter JD (1999) Cytoplasmic Polyadenylation in Development and Beyond. Microbiol Mol Biol Rev **63:** 446-456

Richter JD (2007) CPEB: a life in translation. Trends in Biochemical Sciences **32:** 279-285

Rothnie HM (1996) Plant mRNA 3′-end formation. Plant Molecular Biology **32:** 43-61

Rothnie HM RJ, 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:** 2200-2210

Rouget C, Papin C, Mandart E (2006) Cytoplasmic CstF-77 Protein Belongs to a Masking Complex with Cytoplasmic Polyadenylation Element-binding Protein in Xenopus Oocytes. J. Biol. Chem. **281:** 28687-28698

Rouillard JM, Brendolise C, Lacroute F (2000) Rna14p, a component of the yeast nuclear Cleavage/Polyadenylation Factor I, is also localised in mitochondria. Molecular and General Genetics MGG **262:** 1103-1112

Ryner LC, Takagaki Y, Manley JL (1989) Multiple forms of poly(A) polymerases purified from HeLa cells function in specific mRNA 3'-end formation. Mol. Cell. Biol. **9:** 4229-4238

Sachs AB, Sarnow P, Hentze MW (1997) Starting at the Beginning, Middle, and End: Translation Initiation in Eukaryotes. Cell **89:** 831-838

Sarkar N (1997) Polyadenylation of mRNA in prokaryotes. Annu. Rev. Biochem. **66:** 173-197

Schul W, Groenhout B, Koberna K, Takagaki Y, Jenny A, Manders E, Raska I, van Driel R, de Jong L (1996) The RNA 3' cleavage factors CstF 64 kDa and CPSF 100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. EMBO J. **15:** 2883-2892

Sharma N, Dey M, Satpathy M, Sachar RC (2002) Evidence of two forms of poly(A) polymerase in germinated wheat embryos and their regulation by a novel protein kinase. Biochemical and Biophysical Research Communications **293:** 403-411

Shatkin AJ, Manley JL (2000) The ends of the affair: Capping and polyadenylation. Nat Struct Mol Biol **7:** 838-842

Shen Y, Ji G, Haas BJ, Wu X, Zheng J, Reese GJ, Li QQ (2008) Genome level analysis of rice mRNA 3'-end processing signals and alternative polyadenylation. Nucl. Acids Res.**:** gkn158

Slomovic S, Portnoy V, Liveanu V, Schuster G (2006) RNA Polyadenylation in Prokaryotes and Organelles; Different Tails Tell Different Tales. Critical Reviews in Plant Sciences **25:** 65-77

Sopory SK, Munshi M (1998) Protein Kinases and Phosphatases and Their Role in Cellular Signaling in Plants. Critical Reviews in Plant Sciences **17:** 245 - 318

Tacahashi Y, Helmling S, Moore CL (2003) Functional dissection of the zinc finger and flanking domains of the Yth1 cleavage/polyadenylation factor. Nucl. Acids Res. **31:** 1744-1752

Takagaki Y, MacDonald CC, Shenk T, Manley JL (1992) The Human 64-kDa Polyadenylylation Factor Contains a Ribonucleoprotein-Type RNA Binding Domain and Unusual Auxiliary Motifs. PNAS **89:** 1403-1407

Takagaki Y, Manley J (1992) A human polyadenylation factor is a G protein beta-subunit homologue. J. Biol. Chem. **267:** 23471-23474

Takagaki Y, Manley JL (1994) A polyadenylation factor subunit is the human homologue of theDrosophila suppressor of forked protein. Nature **372:** 471-474

Takagaki Y, Manley JL (1997) RNA recognition by the human polyadenylation factor CstF. Mol. Cell. Biol. **17:** 3907-3914

Takagaki Y, Manley JL (1998) Levels of Polyadenylation Factor CstF-64 Control IgM Heavy Chain mRNA Accumulation and Other Events Associated with B Cell Differentiation. Molecular Cell **2:** 761-771

Takagaki Y, Manley JL (2000) Complex Protein Interactions within the Human Polyadenylation Machinery Identify a Novel Component. Mol. Cell. Biol. **20:** 1515-1525

Takagaki Y, Manley JL, MacDonald CC, Wilusz J, Shenk T (1990) A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. Genes & Development **4:** 2112-2120

Takagaki Y, Ryner L, Manley J (1989) Four factors are required for 3'-end cleavage of pre-mRNAs. Genes Dev. **3:** 1711-1724

Takagaki Y, Seipelt RL, Peterson ML, Manley JL (1996) The Polyadenylation Factor CstF-64 Regulates Alternative Processing of IgM Heavy Chain Pre-mRNA during B Cell Differentiation. Cell **87:** 941-952

Thuresson A, Astrom J, Astrom A, Gronvik K, Virtanen A (1994) Multiple Forms of Poly(A) Polymerases in Human Cells. PNAS **91:** 979-983

Torchet C, Bousquet-Antonelli C, Milligan L, Thompson E, Kufel J, Tollervey D (2002) Processing of 3'-Extended Read-Through Transcripts by the Exosome Can Generate Functional mRNAs. Molecular Cell **9:** 1285-1296

Tucker M, Parker R (2000) Mechanisms and control of mRNA decapping in Saccharomyces cerevisiae. Annu. Rev. Biochem. **69:** 571-595

Vaňáčová Š, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W (2005) A New Yeast Poly(A) Polymerase Complex Involved in RNA Quality Control. PLoS Biology **3:** e189

Verma N, Sachar RC (1994) Evidence to show the phosphoprotein nature of poly(A) polymerase in germinated wheat embryos: enzyme regulation through phosphorylation. Plant Science **99:** 151-160

Viphakone N, Voisinet-Hakil F, Minvielle-Sebastia L (2008) Molecular dissection of mRNA poly(A) tail length control in yeast. Nucl. Acids Res. **36:** 2418-2433

Wahle E (1991) Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. J. Biol. Chem **266:** 3131-3139

Wahle E, Keller W (1996) The biochemistry of polyadenylation. Trends in Biochemical Sciences **21:** 247-250

Wahle E, Ruegsegger U (1999) 3'-End processing of pre-mRNA in eukaryotes. FEMS Microbiology Reviews **23:** 277-295

Wallace AM, Dass B, Ravnik SE, Tonk V, Jenkins NA, Gilbert DJ, Copeland NG, MacDonald CC (1999) Two distinct forms of the 64,000 Mr protein of the cleavage stimulation factor are expressed in mouse male germ cells. PNAS **96:** 6763-6768

Wallace AM, Denison TL, Attaya EN, MacDonald CC (2004) Developmental Distribution of the Polyadenylation Protein CstF-64 and the Variant {tau}CstF-64 in Mouse and Rat Testis. Biology of Reproduction **70:** 1080-1087

Wilusz J, Shenk T (1988) A 64 kd nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. Cell **52:** 221-228

Wilusz J, Shenk T, Takagaki Y, Manley JL (1990) A multicomponent complex is required for the AAUAAA-dependent cross-linking of a 64-kilodalton protein to polyadenylation substrates. Mol. Cell. Biol. **10:** 1244-1248

Wyers F, Rougemaille M, Badis G, Rousselle J-C, Dufour M-E, Boulay J, Régnault B, Devaux F, Namane A, Séraphin B, Libri D, Jacquier A (2005) Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase. Cell **121:** 725-737

Xu F, Cohen SN (1995) RNA degradation in Escherichia coli regulated by 3' adenylation and 5' phosphorylation. Nature **374:** 180-183

Xu R, Zhao H, Dinkins R, Cheng X, Carberry G, Li Q (2006) The 73 kD Subunit of the cleavage and polyadenylation specificity factor (CPSF) complex affects reproductive development in Arabidopsis. Plant Molecular Biology **61:** 799-815

Xu Y, Hiremath S, Wang T (1983) Androgen regulation of nuclear poly(A) polymerase activities in rat ventral prostate. Biochim. Biophys. acta **760:** 278-282

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:** 2277-2278

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:** 405-445

Zhao W, Manley JL (1996) Complex alternative RNA processing generates an unexpected diversity of poly(A) polymerase isoforms. Mol. Cell. Biol. **16:** 2378-2386

Zhelkovsky AM, Kessler MM, Moore CL (1995) Structure-Function Relationships in the Saccharomyces cerevisiae Poly(A) Polymerase. J. Biol. Chem **270:** 26715-26720

Zorio DAR, Bentley DL (2004) The link between mRNA processing and transcription: communication works both ways. Experimental Cell Research **296:** 91-97

Vita

Amrita Bandyopadhyay was born on March 29, 1981, in Burdwan, India. She received her BS degree in Botany, with honors, from the Calcutta University in 2002. She then attended the Banaras Hindu University for an MS degree in Botany, with first class, in 2005. In the spring of 2006, she joined the Department of Plant and Soil Sciences, University of Kentucky to pursue her MS degree in Plant and Soil Science program. She held the position of a Graduate Research Assistant during her tenure at the department. Her honors and activities include honorary judge at the 2007 Science Fair, *Squires Elementary School*, Lexington, KY and honorary judge at the 2008 Science Fair, *Northern Elementary School*, Lexington, KY. She has published the following paper in peer reviewed journals:

Hunt A, Xu R, Addepalli B, Rao S, Forbes K, Meeks L, Xing D, Mo M, Zhao H, **Bandyopadhyay A**, Dampanaboina L, Marion A, Von Lanken C, Li Q (2008) Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. BMC Genomics **9:** 220

Amrita Bandyopadhyay

April 15, 2009