University of Kentucky
UKnowledge

Theses and Dissertations--Plant and Soil Sciences

Plant and Soil Sciences

2011

FUNCTIONAL CHARACTERIZATION OF WD REPEAT PROTEINS, AtCstF50 AND AtFY IN CLEAVAGE AND POLYADENYLATION

Lavanya Dampanaboina
University of Kentucky, ldamp2@uky.edu

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation

Dampanaboina, Lavanya, "FUNCTIONAL CHARACTERIZATION OF WD REPEAT PROTEINS, AtCstF50 AND AtFY IN CLEAVAGE AND POLYADENYLATION" (2011). *Theses and Dissertations--Plant and Soil Sciences*. 2.
https://uknowledge.uky.edu/pss_etds/2

This Doctoral Dissertation is brought to you for free and open access by the Plant and Soil Sciences at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Plant and Soil Sciences by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Lavanya Dampanaboina, Student

Dr. Arthur G. Hunt, Major Professor

Dr. Arthur G. Hunt, Director of Graduate Studies

FUNCTIONAL CHARACTERIZATION OF WD REPEAT PROTEINS, AtCstF50 AND AtFY IN CLEAVAGE AND POLYADENYLATION

ABSTRACT OF DISSERTATION

A dissertation submitted for partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the College of Agriculture at the University of Kentucky

By

Lavanya Dampanaboina

Lexington, Kentucky

Director: Dr. Arthur G. Hunt, Professor

Department of Plant and Soil Sciences

Lexington, Kentucky

2011

Copyright © Lavanya Dampanaboina 2011

ABSTRACT OF DISSERTATION

FUNCTIONAL CHARACTERIZATION OF WD REPEAT PROTEINS, AtCstF50 AND AtFY IN CLEAVAGE AND POLYADENYLATION

Polyadenylation is an essential post-transcriptional modification resulting in a mature mRNA in eukaryotes. Three cis-elements the Far Upstream Element (FUE), Near Upstream Element (NUE), and Cleavage Site (CS) - guide the process of cleavage and polyadenylation with the help of multi-subunit protein complexes cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF) along with cleavage factors and poly(A) polymerase. Protein-protein interactions play an important role in the cleavage and polyadenylation process. WD repeat proteins play an important role in protein-protein interactions and have diverse functions in plant system. In the present study WD repeat proteins AtCstF50 and AtFY were studied for their role in polyadenylation process.

Mammalian CstF50 is a WD repeat protein that is one of the subunit of CstF that aids in the cleavage step by associating with CPSF and cleavage factors. AtCstF50 was functionally characterized using T-DNA knock-out lines and by identifying the proteins that interacts with it in the process. Results shows that AtCstF50 is essential and was identified as part of CPSF complex, which is different from its mammalian counter part. CPSF was known to interact with Fip (factor interacting with PAP), Poly(A) polymerase and Poly(A) binding protein and AtCstF50 also interacts with these complexes.

AtFY is a 3' end processing factor which contains WD repeats is one of the subunits of the CPSF complex in Arabidopsis polyadenylation machinery. The AtFY interacts with FCA and promotes the alternative polyadenylation and also plays a role in polyadenylation site choice of FCA mRNA. We characterized the FY expression and localization of FY in the cell by fusing with RFP reporter. Results show that FY accumulates in the nucleus while FY with deleted calmodulin binding domain localizes both to the nucleus and outside the nucleus. The individual N-terminal and C-terminal

domains also localized in the nucleus suggesting that they are multiple nuclear localization signals in FY and calmodulin might play a direct or indirect role in FY localization. Using a tethering assay we proved that AtFY is able to recruit the 3' end processing complex in the proximal polyadenylation site choice of the reporter mRNA.

Key words: Polyadenylation, WD repeat proteins, AtCstF50, AtFY, tethering assay

Lavanya Dampanaboina

December, 2011

FUNCTIONAL CHARACTERIZATION OF WD REPEAT PROTEINS, AtCstF50 AND AtFY IN CLEAVAGE AND POLYADENYLATION

By

Lavanya Dampanaboina

Dr. Arthur G. Hunt

Director of Dissertation

Arthur G. Hunt

Director of Graduate Studies

December, 2011

ACKNOWLEDGEMENTS

My dissertation work was carried out under the guidance of Dr. Arthur G. Hunt at the University of Kentucky, Lexington, KY. First, I would like to thank Dr. Arthur G. Hunt for giving me an opportunity to work in his laboratory and I would like to thank my committee members, Dr. Bruce Downie, Dr. George Wagner and Dr. Ling Yuan for their guidance and valuable suggestions in successfully completing my thesis objectives. Enlightening discussions during my committee meetings led to novel ideas and methods to carryout different experiments. I would like to acknowledge all the people who are instrumental in my academic and research accomplishments particularly, my teachers at school, colleges and Universities.

I would like to extend my heartfelt gratitude to faculty, Chairs (Drs. Todd Pfeiffer and Michael Barrett), DGS (Dr. Arthur G. Hunt) and staff of the Department of Plant & soil sciences for their help and encouragement during my graduate study at UK. I am grateful to the graduate school and Department of Plant and Soil Sciences for their financial support during my graduate study at the UK. I would like to extend my sincere gratitude to the present and former lab members of Dr. Hunt, Carol von Lanken, Liuyin Ma, Bobby Gaffney, Jason Briones, Balasubrahmanyam Addepalli and Amrita Bandyopadhyay. My special thanks to Dr. Sharyn Perry for her help with microscopy, Dr. Joe Chappell for sharing the equipment and supplies and Dr. Michael Goodin for providing the reporter vectors for my research. I really appreciate the help and support by Amy Crume in providing me the plant material for my experiments.

I would like to thank my father, Satyanarayana and mother Vijaya and brother for their support. I would like to thank my husband Venugopal Mendu for his constant support and encouragement to pursue my higher studies at UK. It was a great pleasure to have my baby Harini with me who made my life joyful. I really appreciate my friends Umei, Xun, Jeanne Marie Hartman. Finally, I express my sincere thanks to all who helped me directly and indirectly during my research work and thesis preparation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS....................iii

TABLE OF CONTENTS....................iv

LIST OF TABLES....................vii

LIST OF FIGURES....................viii

LIST OF FILES....................x

CHAPTER ONE: Review of literature....................1

Introduction....................1

Importance of poly(A) tail in eukaryotic mRNA....................1

Mechanism of polyadenylation in mammals....................7

The Cleavage and Polyadenylation Specificity Factor (CPSF) complex....................8

The Cleavage Stimulation Factor (CstF) Complex....................10

Poly(A) polymerase (PAP)....................12

Mechanism of polyadenylation in Yeast....................13

Mechanism of polyadenylation in plants....................16

CHAPTER TWO: Functional characterization of the Arabidopsis orthologue of Mammalian CstF50 (Cleavage stimulation Factor subunit-50)....................25

I. Introduction....................25

II. Materials and Methods....................27

PCR genotyping of T-DNA insertions of *CstF50.1.1* and *CstF50.1.2*....................27

Southern blot analysis of *CstF50.1.1*....................27

Random primer labeling of DNA....................28

Hybridization....................29

Reverse transcription....................29

Yeast two-hybrid assay....................29

CstF-50 localization....................30

III. Results..33

Characterization of T-DNA insertion mutants *CstF501.1 and CstF501.2* in the gene encoding AtCstF-50……………...33

The AtCstF-50 protein-protein interaction network……………...................................34

Localization of AtCstF-50 to the Endoplasmic Reticulum…………….........................35

IV. Discussion……………..36

AtCstF-50 also interacts with one of the three isoforms of the nuclear poly(A) binding protein……………………………………………………………………………37

CHAPTER THREE: Characterization of FY, an orthologue of yeast Pfs2p.....................54

I. Introduction..54

II. Materials and Methods..56

Vector construction 56

Agro-infiltration experiments for transient expression studies...57

Localization studies using microscopy...58

RNA extraction, cDNA synthesis and PCR amplification..58

III. Results...59

Full length FY localizes to the nucleus 59

Fusion proteins containing the FY-NTD domain and FY-CTD domains localize to the nucleus……………………………………………………………………………...59

Mutation in the calmodulin binding domain alters the localization of FY.......................60

IV. Discussion……………..61

CHAPTER FOUR: Functional Characterization of the polyadenylation factor FY in 3'end processing using tethering assay……………………..…….........................75

I. Introduction..75

II. Materials and Methods..78

Design and plasmid vector construction...78

Agro-infiltration experiments for transient expression studies...79

Fluorescence Microscopy..80

III. Results……..81

IV. Discussion……………...83

CHAPTER FIVE: Summary……..101

REFERENCES...104

VITA...113

LIST OF TABLES

Table 2.1 List of primers used for genotyping of CstF50 T-DNA insertion lines.…40

Table 2.2 Summary of pair-wise interactions of AtCstF50 with plant polyadenylation factors.……………………….……………….……………41

Table 3.1 Primers used for the localization experiments.…………………..………63

Table 4.1 Primers used for the tethering assay.…………………...………..………85

LIST OF FIGURES

Figure 1.1 Cis-elements in Mammals, Yeast and Plants that guide the process of cleavage and polyadenylation ……………………………………..………21

Figure 1.2 Mammalian 3' end processing machinery highlighting CPSF, CstF, CF-Im CF-IIm, PAP and symplekin association with RNA and RNA pol II…...22

Figure 1.3 Yeast 3' end processing machinery showing complexes CPF, CFIA, CFIB and CFII …………………………………………………………………23

Figure 1.4 Plant 3' end processing machinery showing complexes CPSF, CstF, Poly(A) polymerase along with other complexes involved in plant cleavage and polyadenylation…………..………………………………..24

Figure 2.1 At*CstF-50* gene made up of 13 exons and 12 introns……..…………….…..42

Figure 2.2 T-DNA insertion in 3rd intron and 4th exon representing both SALK lines 075594 (*cstf-50-1-1)* and 075595 (*cstf-50-1-2)*..……….……………43

Figure 2.3 PCR confirmation of T-DNA insertion lines.…………………………...44

Figure 2.4 Genotyping of SALK_075594 (*cstf-50-1-1*) line….…………………45

Figure 2.5 Restriction analysis of T-DNA and genomic region of *CSTF50*...………46

Figure 2.6 *Eco*RI digestion of the genomic DNA Columbia-0 (C) and digested genomic DNA from mutant (SALK_075594) plants…………………47

Figure 2.7 Southern blots of WT plants (lane designated as "C") or plants putatively containing the SALK_075594 T-DNA insertion..……….…………..48

Figure 2.8 Reverse Transcription analysis of intron insertion line…………………49

Figure 2.9 PCR analysis of SALK_075595 line..…………………………….….50

Figure 2.10 Pairwise protein-protein interaction studies using yeast two hybrid system…………………………………………………………...51

Figure 2.11 Localization studies of AtCstF50.………………………………..……53

Figure 3.1 Fusion constructs of FY with RFP for transient expression studies..….64

Figure 3.2 Different domains of FY used for localization studies..………………....65

Figure 3.3 Localization of GFP and RFP nuclear markers……………………..66

Figure 3.4 Localization of FY fused RFP……………………………….........67

Figure 3.5 Localization of FY-NTD domain fused to RFP………………….........68

Figure 3.6 Localization of FY-CTD domain fused to RFP…………………………......69
Figure 3.7 Localization of cFY fused to RFP……………………………………..........70
Figure 3.8 Localization of cFY-CTD fused to RFP...………………….…………..........71
Figure 3.9 A possible model for FY localization to the nucleus..….…………………......71
Figure 3.10 Domains of FY responsible for interaction with DDB1 Ubiquitin Ligase……………………………………………………………………73
Figure 3.11 Possible roles of calmodulin binding to FY in flowering pathway..……74
Figure 4.1 Tethering constructs of FY for testing the Tethering assay…………….87
Figure 4.2 Constructs made to test the tethering assay………………….………..88
Figure 4.3 Principle involved in tethering assay…………………………….……..89
Figure 4.4 Model for FY function in recruiting 3' end processing complex………90
Figure 4.5 Two constructs reporter construct and tethering construct made to test the hypothesis. ……………………………………………………………91
Figure 4.6 Poly(A) signal is necessary for the GFP reporter expression…...………92
Figure 4.7 Reporter construct (GFP reporter, wild type MS2 binding sites, NLS and no poly(A) signal) for the tethering assay does not express GFP..………93
Figure 4.8 Tethering assay with reporter construct and tethering construct with internal RFP nuclear control……………………………………………..94
Figure 4.9 Control constructs made for tethering assay……………….……….95
Figure 4.10 Co-infiltration of reporter construct with RFP nuclear marker and tethered control construct (with out FY)…………………………………96
Figure 4.11 Co-infiltration of tethered construct with reporter control construct and RFP nuclear marker…………………………………………………97
Figure 4. 12 Tethering assay model to test whether absence of calmodulin binding domain affects FY function…………………………………………98
Figure 4.13 Co-infiltration of tethered construct cFY with reporter construct and RFP nuclear marker………………………………………………………99
Figure 4.14 Summary of tethering assay…………………………………………100

LIST OF FILES

Lavanya_dissertation.pdf……………………..……………………..………….32 MB

CHAPTER ONE
Review of Literature

Introduction
Importance of poly(A) tail in eukaryotic mRNA

Transcription is the first step of gene expression during which pre-messenger RNA (pre-mRNA) is synthesized and processed into a mature mRNA that acts as a template for protein synthesis (Keene 2010). The mRNA is the primary product of transcription and is synthesized by RNA polymerase II in all eukaryotes (Wahle and Ruegsegger 1999). The pre-mRNA has to undergo many post-transcriptional modifications such as 5'capping, splicing and polyadenylation before translation into protein. All these modifications occur in the nuclear compartment (Bai and Tolias 1998). Once the pre-mRNA undergoes these chemical modifications, it is now called mature mRNA and is ready to be transported to the cytosol for translation.

Three prime end processing is a universal modification observed in all eukaryotes. Polyadenylation (addition of the poly(A) tail at the 3'OH end of RNA in a non-template dependent manner) is an essential post-transcriptional modification also known as 3'end processing. The poly(A) tail is necessary for nuclear export, cytoplasmic localization, mRNA stability, translocation and translation of mRNA (Andreassi and Riccio 2009). Nuclear mRNA's that are not processed at their 3'ends are either subjected to degradation or not transported to cytosol for translation (Millevoi and Vagner 2010). Defects in mRNA 3' end processing leads to many deficiencies while accumulation of such unprocessed products results in diseases (Danckwardt, Hentze et al. 2008). These defects show how important 3' end processing is for maintaining cell viability, growth and development (Mandel, Bai et al. 2008).

Polyadenylation factors interact with transcription factors, elongation factors, splicing factors and transcription termination factors, all proteins that connect multiple processes of transcription, splicing, capping and those involved in transport of mRNA (Millevoi and Vagner 2010). The phosphorylated C-terminal domain (CTD) of RNA polymerase II

is an essential component of the 3' end RNA processing complex and helps to assemble the polyadenylation factors for carrying out 3' end processing (Ryan, Murthy et al. 2002). The CTD thus plays a major role in the network of post-transcriptional modifications (Hirose and Manley 1998) (Millevoi and Vagner 2010).

For all transcripts produced by RNA polII, the poly(A) tail is added by canonical nuclear poly(A) polymerases. Once the poly(A) tail is synthesized it is bound by a nuclear poly (A) binding protein that regulates the tail length, release of RNA from the transcription site, and export from nucleus to the cytosol. In the cytoplasm, the poly(A) tail is bound by a different, cytoplasmic poly(A) binding protein that facilitates a sort of circularization of the mRNA by interaction with 5'cap binding proteins. The cytoplasmic poly(A) binding protein also protects the 3'end from 3' to 5' RNA exonucleases. In contrast, the non-canonical poly(A) polymerases polyadenylate and target mRNA for degradation. These poly(A) tails act as receptors for degradation and is the common mechanism of regulation in the nucleus, mitochondrion and choloroplast. Polyadenylation-mediated degradation was initially thought to be operating only in prokaryotes but these pathways are also observed in yeast (*Sachharomyces cerevisiae*), Drosophila (*Drosophila melanogaster*), trypanosomes (*Trypanosoma brucei*), human beings (*Homo sapiens*) and plants (*Arabidopsis thaliana*) (Lange, Sement et al. 2009).

The poly(A) tail at the 3'end of mRNA mediates both cap-dependent and cap-independent translation initiation. Cap-dependent translation initiation is carried out by interactions between poly(A) binding protein, eIF4G (eukaryotic initiation factor G/ a large scaffolding protein) and eIF4E (eukaryotic initiation factor E, a cap binding protein). These protein interactions brings both 5' and 3'ends of mRNA together resulting in a circular mRNP that plays a critical role in mRNA metabolism and turnover. In yeast, mRNA turnover is brought about by deadenylation whereby a small poly(A) tail is a signal for decapping, followed by a XRN1 (exoribonuclease) dependent degradation pathway. The presence of nonsense codons in other than regular positions also activates deadenylation-independent decapping in yeast. Two major pathways driving mRNA degradation pathways were identified in mammals and budding yeast (Garneau, Wilusz et

al. 2007). Deadenylation is catalyzed by 3' exonucleases specific to the poly(A) tail (Shock, Fischer et al. 2007). In mammals, a highly conserved eukaryotic RNase D homolog PARN (poly(A) specific ribonuclease) is a major deadenylase. In budding yeast two complexes, pan2/pan3p (pab1p-dependent poly(A) nuclease) (Brown, Tarun et al. 1996) and Ccr4p/Pop2p, drive the process of deadenylation. The second complex Ccr4p/Pop2p is part of bigger complex, the Ccr4/Not complex, and is the major deadenylase in yeast (Martine A 2003). After deadenylation, the degradation of transcripts occurs via decapping by Dcp1p (decapping protein) and Dcp2p for decay in the 5'- to 3'-direction. Other factors such as Edc1p-3p (enhancer of decapping proteins), Lsm (like Sm complex), and the Dhh1p (DEAD box helicase) regulate decapping. The degradation of decapped transcripts is mediated by Xrn1p (exoribonuclease), a 5'-3' exonuclease, and in the 3'to 5' direction by the exosome (Meyer, Temme et al. 2004). The exosome complex is a multisubunit complex with 10 necessary proteins nine of which have 3'to 5' exonuclease domains. Rrp6p (ribosomal RNA processing 6) (Briggs, Burkard et al. 1998) is a 11th non-essential protein that associates with the exosome only in the nucleus. Two helicases, Skip2p and Mtr4p (mRNA transport 4), also interact with the exosome. The free cap generated in 3'- to 5'-decay is hydrolyzed by the Dcs1p (scavenger decapping enzyme) protein (Shock, Fischer et al. 2007).

Three prime end processing is defined as a coupled process of cleavage and polyadenylation carried out in three steps, namely: 1) site choice for polyadenylation; 2) cleavage of pre-mRNA at that site and; 3) poly(A) tail addition. Three prime end processing is guided by the cis-elements present on the RNA. These cis-elements are identified by a series of trans-factors that identify these elements and carry out the processing event. These trans-factors are multi-subunit proteins and each subunit has a specific function. The process is complex as it involves many factors and their interactions among each other and with RNA. The transcription elongation complex is connected to the polyadenylation of pre-mRNA by acting as a platform for the assembly of polyadenylation factors that coordinates with transcription (Zorio and Bentley 2004) via associations with the RNA polymerase II phosphorylated C-terminal domain (Colgan and Manley 1997; Hirose and Manley 1998; Proudfoot and O'Sullivan 2002; Proudfoot,

Furger et al. 2002). TFIID (Transcription factor II D) helps in loading Cleavage and Polyadenylation Specificity Factor (CPSF) onto the RNA polymerase II complex which associates it with the poly(A) signal sequences at the 3'end of the RNA (Hirose and Manley 1998). The CTD of RNA polymerase II with 52 hepta peptide repeats (YSPTSPS) plays an important role in cleavage and polyadenylation (Hirose and Manley 2000). These heptad repeats of RNA pol II CTD are conserved from yeast to humans (Zorio and Bentley 2004). The lack of the carboxyl terminal domain of RNA polymerase II large subunit drastically reduces the cleavage reaction (Adamson, Shutt et al. 2005). The presence of defective poly(A) signals also affects transcription termination and can lead to transcription read-through (Proudfoot 2004).

Multiple cis-elements (Fig.1.1) guide the process of cleavage and polyadenylation by employing protein machinery. In eukaryotes there are usually three cis-elements, the cleavage site, an A-rich element and one or two U-rich elements, listed from most proximal to most distal (further 5') from the site of eventual polyadenylation (Millevoi and Vagner 2010). The involvement of many RNA elements and RNA binding proteins in 3'end processing suggests multiple steps of regulation for pre-mRNA processing. Cleavage and polyadenylation factors involved in 3'end processing determine the quality and specificity of this process. The first step in this process, polyadenylation site choice, is guided by the assembly of the protein factors that further drive the cleavage step and polyadenylation of pre-mRNA. In mammals, the upstream A-rich element is recognized by Cleavage and Polyadenylation Specificity Factor (CPSF) and the downstream GU rich element is recognized by Cleavage stimulation Factor (CstF). Other cis-elements that are upstream and downstream to the poly(A) signal act as auxiliary elements to increase the specificity of 3' end processing.

As mentioned above, there are two well-studied polyadenylation complexes, which are involved in the recognition of the cis-elements. They are the Cleavage and Polyadenylation Specificity Factor (CPSF) and Cleavage stimulation Factor (CstF). These proteins interact with RNA and other proteins to cleave the pre-mRNA. Cleavage and Polyadenylation Specificity Factor (CPSF) also plays an important role in

polyadenylation by interacting with poly(A) polymerase. The poly (A) polymerase adds the poly(A) tail whereas the nuclear poly(A) binding protein regulates the length of the poly(A) tail. The length of poly (A) tail differs and the length also depends on the kind of organism (Edmonds 2002). For example in yeast the poly(A) tail has a maximum length of 75 residues while in mammals it may be as long as 200 residues (Manley 1995). *In vitro* assays for poly(A) tail addition identified CPSF, PAP (poly(A) polymerase) and poly(A) binding protein as the proteins necessary and sufficient for the polyadenylation of mRNA. Recently two new proteins, CPSF associated factor WDR33 (a WD repeat protein) and serine threonine protein phosphatase I (PPI) have been identified as essential proteins in 3' end processing. Although the core components of the polyadenylation complexes exist in all cells throughout the life of an organism, there are certain differences on the expression of certain polyadenylation factors based on their functions. For example, mammalian testis-specific isoforms of CstF64 (Cleavage stimulation Factor 64) play a critical role in RNA turnover during male gametogenesis (Wallace, Dass et al. 1999; Dass, McMahon et al. 2001), implying the presence of different isoforms of polyadenylation factors in different tissues

The polyadenylation of mRNA also takes place in the cytoplasm. Cytoplasmic elongation of the poly(A) tail is a phenomenon observed in the oocytes and embryos of clams (*Meretrix meretrix*), worms (*Urechis caupo*), frogs (*Xenopus laevis*) and mice (*Mus musculus*) (Rosenthal, Tansey et al. 1983; Rosenthal and Wilt 1986; Paris, Osborne et al. 1988; Fox, Sheets et al. 1989; McGrew, Dworkin-Rastl et al. 1989; Vassalli, Huarte et al. 1989), and is a mechanism used for the regulation of mRNA translation during oogenesis and spermatogenesis. Changes in poly(A) tail length also impact the mRNA localization. Maternal mRNAs of some species are regulated by this process. In Drosophila, specific mRNA expression in specific compartment temporally regulates embryonic patterning and in *Caenorhabditis elegans* this process determines sex. In Xenopus, cytoplasmic mRNAs are stored with a short poly(A) tail of 20-40 nucleotides in the growing oocyte. After fertilization, the poly(A) tail is elongated to 80 to 250 nucleotides and the long poly(A) tail (Gorgoni and Gray 2004) activates translation by its interaction with cytoplasmic poly(A) binding proteins that assemble initiation factors by their interaction

with eukaryotic initiation factor eIF4G. During maturation not all mRNA poly(A) tails are elongated at the same time. Instead, some mRNAs are extended earlier and some later in development. Some mRNAs that have long poly(A) tails are also deadenylated to repress translation. At least four cis-elements function as cytoplasmic polyadenylation elements of Xenopus ooctytes. The C-Rich Element, Cytoplasmic Polyadenylation Element, and U Rich Embryonic Cytoplasmic Polyadenylation Element (CPE), mediate cytoplasmic polyadenylation in the Xenopus zygote and embryo (Simon, Tassan et al. 1992; Paillard, Maniey et al. 2000). The A rich element (AAUAAA) and a specific cytoplasmic polyadenylation element (CPE) drive the process of cytoplasmic polyadenylation in general. The expression of mRNA or polyadenylation of the cytoplasmic mRNA might be dependent on the distance between AAUAAA and CPE elements. Three trans-factors are necessary for cytoplasmic polyadenylation, CPSF that binds AAUAAA, cytoplasmic polyadenylation element binding protein (CPEB) that binds the CPE, and poly(A) polymerase that adds the poly(A) tail (Radford, Meijer et al. 2008).

The CPE is recognized by, an RNA binding protein CPEB (Hake and Richter 1994; Stebbins-Boaz, Hake et al. 1996; Hake, Mendez et al. 1998). Structurally CPEB has two RNA recognition motifs and a zinc finger region necessary for its interaction with CPE elements (Radford, Meijer et al. 2008). The cytoplasmic CPSF complex holds, in common with the nuclear CPSF complex, CPSF160, CPSF100, CPSF30 but lacks CPSF73 which is present in the nuclear complex (Dickson, Bilger et al. 1999). CPE also activates the binding of CPSF to the poly(A) signal sequence even in the absence of CPEB. Symplekin, a protein known to regulate polyadenylation and promote gene expression in the nucleus (Hofmann, Schnölzer et al. 2002), is also thought to associate with the CPSF complex in the cytoplasm along with poly(A) polymerase (Radford, Meijer et al. 2008) .

Cytoplasmic mRNA that has CPE elements is translationally repressed and remains in a complex bound by CPEB, eIF4E (Eukaryotic initiation factor 4E) and maskin. At maturation CPEB gets phosphorylated by kinases and assembles into the CPSF complex.

Maskin dissociates from the complex at the time of poly(A) tail elongation. An association between eIF4G and PABP might be responsible for the dissociation of maskin (Mendez and Richter 2001).

Mechanism of polyadenylation in mammals

Polyadenylation is well studied in mammals. At least four cis-elements are required for polyadenylation (Fig.1.1). They are the upstream element 10-30 nucleotides upstream of the cleavage site, a highly conserved AAUAAA sequence, a downstream, less conserved U rich or GU-rich element, and the cleavage site, a CA dinucleotide, which is also the site of polyadenylation (Zhao, Hyman et al. 1999). The AAUAAA sequence element was the first identified element to play a role in polyadenylation (Proudfoot and Brownlee 1976). This element is observed in 80-90 percent of the sequenced mRNA population. The upstream and downstream elements other than the cleavage site regulate the poly(A) site choice and increase the efficiency of cleavage and polyadenylation by coordinating the interaction between different polyadenylation factors while the downstream element also increases the processivity and efficiency of 3' end processing by stabilizing the polyadenylation complexes (Gilmartin, Fleming et al. 1992).

Two multi-subunit complexes recognize the cis-elements present on the pre-mRNA. These are the **Cleavage and Polyadenylation Specificity Factor (CPSF)** that binds to the poly(A) signal sequence AAUAAA and the **Cleavage Stimulation Factor (CstF)** that binds the downstream element (Proudfoot and O'Sullivan 2002). They act with other cleavage factors CF I(m) (mammalian Cleavage factor I) and CF II(m) (mammalian Cleavage factor II). The RNA polymerase II C-terminal domain is also involved in this process along with other multi-subunit complexes that bind to pre-mRNA (Fig.1.2). The nuclear poly(A) binding protein 1 (PabN1) and poly(A) polymerase (PAP) are necessary for the polyadenylation reaction. CPSF160 and CstF64 are the subunits of two different protein complexes that bind to the RNA at specific sequences that signal cleavage (Colgan and Manley 1997).

The Cleavage and Polyadenylation Specificity Factor (CPSF) complex: CPSF is made up of five subunits, i) CPSF160, ii) Fip1 (Factor interacting with poly(A) polymerase), iii) CPSF73, iv) CPSF100, and v) CPSF30 . In the following paragraphs, I provide some of the information available for each of these subunits and what is known about their interactions.

i). CPSF160 binds to the AAUAAA element directly whereas the other CPSF subunits increase the specificity and strength of this binding. CPSF160 has bipartite nuclear localization signal with RNP1 and RNP2 motifs near the N-terminus of the protein. It is not clear whether these domains are responsible for RNA binding or not. CPSF 160 is essential for both cleavage and polyadenylation steps and highly conserved in all eukaryotes. It also interacts with transcription initiation factors (TFIID), the elongation complex of RNA polII and is involved in transcription termination (Dantonel, Murthy et al. 1997). Other studies also suggested that other subunits and complexes such as CstF improve the specificity and binding of CPSF160 to RNA. CPSF160 interacts with CPSF100, hFip1 (human factor interacting with poly(A) polymerase) and PAP.

ii). Human Fip1 is a 66KDa protein that possesses an acidic region, followed by a conserved region of 70 residues and a C-terminal region with a proline-rich domain with alternate arginines and aspartates followed by an arginine-rich region with a bipartite nuclear localization signal (Kaufmann, Martin et al. 2004). The acidic domain (1-111) is responsible for its interaction with PAP and an evolutionarily conserved region (137-243) of hFip1 interacts with CPSF30. The N-terminal region of hFip1 also interacts with CstF77 and CPSF160. CPSF160 also interacts with the C-terminal domain of hFip1. The hFip1-PAP interaction is facilitated by the presence of cleavage factor I(m) (Mandel, Bai et al. 2008). Human Fip1 has an arginine-rich RNA binding motif that binds to the upstream U-rich elements on the pre-mRNA and tethers the RNA to PAP. Human Fip1 forms a ternary complex with CPSF160 and PAP, suggesting a significant role of this subunit in poly(A) signal recognition along with CPSF160 (Kaufmann, Martin et al. 2004) .

iii). CPSF73 and iv) CPSF100 are similar proteins signifying their origin from a common ancestor (Manley and Takagaki 1996). These proteins belong to the metallo-beta-lactamase superfamily (Callebaut, Moshous et al. 2002), other members of which have known nuclease activities. Absence of critical residues in CPSF100 suggests that this subunit is not a nuclease. However, CPSF73(I) has an intact metallo-beta-lactamase domain in its N-terminus. CPSF73-I also has a beta CASP domain. CPSF73 is the endonuclease (Ryan, Calvo et al. 2004; Dominski, Yang et al. 2005) involved in the cleavage reaction and is metal dependent as beta CASP proteins have conserved residues which bind metal ions (Nedea, He et al. 2003; Mandel, Kaneko et al. 2006). The cleavage reaction is zinc dependent and any defects in these binding domains leads to polyadenylation defects and cell death. Mammals possess a second CPSF73 isoform, called RC-68 or Int9. This second isoform might be involved in the processing of small RNAs and interacts with an isoform of CPSF100, RC-72 or Int-11.

v). CPSF30 has a zinc knuckle at its C-terminus and five CCCH zinc finger repeats. It is 40% similar to the yeast Yth1 (a subunit of Yeast Cleavage Polyadenylation factor CPF) that is necessary for its viability in yeast (Barabino, Hubner et al. 1997). Mammals possess a related isoform of unknown function with 54 percent identity to CPSF30. The CCHC zinc knuckle has a consensus sequence CX2CX4HX4C and the zinc fingers have the consensus sequence CX8CX5CX3H. *In vitro*, CPSF30 binds RNA. The loss of the zinc knuckle does not change the specificity of binding but decreases its efficiency. Of the five zinc fingers of CPSF30, the second is the most conserved in eukaryotes. Mutations in the conserved cysteine residues of the second zinc finger are lethal whereas mutations in other residues affect the cleavage reaction and binding to the pre-mRNA (Zarudnaya, Kolomiets et al. 2002). These zinc finger motifs are also involved in protein-protein interactions. The Drososphila homolog of CPSF30, Clipper (CLP), is an endonuclease capable of cleaving RNA hairpins. CLP also has five zinc finger motifs and two zinc knuckle motifs. The zinc finger motifs are highly conserved and are responsible for the endoribonuclease activity.

The Cleavage Stimulation Factor (CstF) Complex: CstF binds to the downstream GU rich element and interacts with the proteins of the CPSF complex that is bound to the conserved poly(A) signal sequence (Zhao, Hyman et al. 1999). The CstF complex is made up of 3 subunits, namely i) CstF64, ii) CstF77 and iii) CstF50. This multi-subunit complex is necessary for the cleavage step of polyadenylation. In the following paragraphs I will furnish a brief expose of what is known about each of these subunits and their interactions.

CstF64 binds the downstream GU rich element. The N-terminus of CstF64 has an RNA binding domain followed by a so-called hinge domain. The RNA binding domain in this protein is responsible for binding the downstream GU-rich element. NMR studies of CstF64 revealed the presence of a U dinucleotide specific binding site and a highly mobile RNA-protein interface that allows the protein to bind to a broad range of GU-rich downstream elements. CstF64 is also essential for cell viability and changes in the levels of this protein affect cell growth and gene expression. The C-terminus of CstF64 has 12 consecutive MEARA/G repeats that form a long alpha helix. This region is bound by proline and glycine residues and the proportion varies in different organisms. A second isoform of CstF64, known as TCstF64 (Tau CstF64), is expressed specifically in male germ cells. TCstF64 functions in male gametogenesis and spermatogenesis (Mandel, Bai et al. 2008). The hinge region in CstF64 is highly conserved and is responsible for its interaction with CstF77 and symplekin (Takagaki and Manley 2000; Hatton, Eloranta et al. 2002). Symplekin is an assembly/scaffolding protein that might be involved in the assembly of CstF and CPSF complexes through its interaction with CstF64(Takagaki and Manley 2000).

CstF77 bridges between the two other subunits of the CstF complex and is similar to the Drosophila suppressor of forked protein. CstF77 has a proline-rich domain required for interacting with two other subunits of CstF and also for self association (Takagaki and Manley 2000). The N-terminus of the CstF77 protein has 12 repeats called the HAT motif (half of a tetratricopeptide repeat (TPR)) that mediate protein-protein interactions (Preker and Keller 1998). The HAT domain has two subdomains, HAT-N (1-240) with

HAT motifs 1-5 and HAT-C (241-549) with HAT motifs 6-12. Dimer formation is mediated by the HAT-C subdomain (Bai, Auperin et al. 2007). The HAT domain is followed by a proline-rich domain that binds to the hinge region of CstF64 and the WD-40 domain of CstF50. CstF77 also interacts with the CTD of RNA polymerase II but with less affinity than CstF50 (McCracken, Fong et al. 1997). The other protein that interacts with CstF77 is CPSF160; this interaction involves the HAT-C domain and is important for CPSF 160 interaction with the poly(A) signal sequence AAUAAA.

CstF50 is a protein with seven WD-40 repeats. The deletion of the last WD-repeat inhibits the interaction of CstF50 with CstF77. CstF50 associates with itself and the N-terminal region is sufficient for this interaction. CstF50 also interacts with the C-terminal domain of RNA polymerase II, thus connecting transcription with polyadenylation. CstF50 also interacts with the BRCA1 (breast cancer type I susceptibility protein) associated protein BARD1 (BRCA1 associated RING domain protein) and with the DNA replication repair factor PCNA. This complex association inhibits nuclear polyadenylation of mRNA when there is DNA damage and may play a role in tumor suppression (Kleiman and Manley 2001).

The interactions of the proteins of the CPSF and CstF complexes with two recognition elements on the RNA promote the positioning of a) Cleavage Factor I(m) (CF I(m) and b) Cleavage Factor II(m) (CF IIm) at the pre-mRNA cleavage site.

CF I(m): CF I(m) binds to RNA sequences similar to UGUAA that may lie upstream of the AAUAAA hexamer. CFIm is a heterotrimeric complex, comprised of 3 subunits that are 68 or 72 kDa, 59kDa, and 25 kDa. De-phosphorylation of a protein in CF I(m) or CF II(m) by phosphatases abolishes the cleavage step in mammals suggesting its role in the cleavage step (Ryan 2007). At the N-terminus of the 68kDa protein, there is an RNA binding domain that interacts with the 25kDa subunit. The C-terminal domain of the 68kDa subunit has RS (arginine, serine)-, RD (arginine, aspartic acid)- and RE (arginine, glutamic acid)- amino-acid repeats akin to SR proteins involved in pre-mRNA splicing.

CF I(m) was identified as a component of the purified spliceosome (Rappsilber, Ryder et al. 2002; Zhou, Sim et al. 2002).

CF II(m): CF II(m) has 2 subunits, Pcf11 and Clp1 (de Vries, Ruegsegger et al. 2000). Pcf11 has a conserved domain that interacts with RNA polII and mutations in this region reduce the efficiency of binding to polII, resulting in transcription termination defects. The N-terminal domain of Pcf11p carries out this interaction with RNA polymerase II ser2 CTD (Barillà, Lee et al. 2001). The function of the C-terminal domain of Pcf11p is not known. Clp1 interacts with CF I(m) and CPSF. Clp1 has a Walker motif which is responsible for ATPase activity (Walker, Saraste et al. 1982). However, biochemical studies indicate that Clp1 lacks such a function, but instead has 5' kinase activity that is involved in tRNA splicing (Walker, Saraste et al. 1982) and in the activation of artificial siRNAs. Clp1 and Pcf11p also interact with each other (Gross and Moore 2001) and Pcf11 has conserved arginine and tryptophan residues at 480 and 489 positions mostly responsible for the binding interface between the two proteins. Immuno-depletion of Clp1 affects cleavage but not polyadenylation (Mandel, Bai et al. 2008).

Poly(A) polymerase (PAP): PAP adds the poly(A) tail to the processed pre-mRNA. PAP binds RNA weakly and non-specifically in the absence of the CPSF complex. *In vitro* studies reveal that PAP does not need any other factor for adding poly(A), but that association with other factors influences enzyme activity and the tail length. CPSF increases the specificity and strength of PAP binding to the RNA substrate. PAP is template-independent and has high specificity for its ATP substrate. PAPs are present in multiple isoforms in cells and tissues (Kyriakopoulou, Nordvarg et al. 2001). There are 6 isoforms of PAP that are produced from a alternatively spliced transcripts from a single gene. PAPI, PAPII and PAP IV are the full-length forms that are functionally active. PAP has three domains namely the N-terminal, middle and C-terminal domain. The active site lies in the middle domain and catalysis requires binding to two metal ions Mn/Mg in the N-terminal domain. These divalent metal ion binding sites are highly conserved. A triad of aspartate residues in the catalytic site is necessary for function. The catalytic region is followed by a nuclear localization signal (NLS). The C-terminus plays a role in protein-

protein interactions and interacts with hFip1 and CPSF160. PAP is substrate specific and undergoes a conformational change to close the active site when bound to a ternary complex of Mg-ATP and pre-mRNA substrates; this change is not brought about by Mg-GTP. CPSF not only specifies the region for cleavage but also promotes polyadenylation by interacting with poly(A) polymerase.

The initial addition of adenine residues is slow and reaches a minimum length of 10. Then PabN1 binds to the poly(A) emerging once it has reached a length of 11-14 residues. The next phase is the elongation phase where PabN1 promotes the processive addition of adenosine monophosphate until the poly(A) tail reaches a length of 200-300 AMP residues (Wahle 1991). PABPN1 regulates mRNA poly(A) tail length and associates with proteins involved in mRNA synthesis and sorting (Liu, Quesada et al. 2007).

Mechanism of polyadenylation in Yeast

There are four cis-elements necessary for 3'end processing in yeast. These are the efficiency element (AU-rich), the positioning element (A-rich), and two U-rich elements present upstream and downstream from the cleavage site (Fig.1.1). The efficiency element is not required for cleavage but improves the efficiency of cleavage. The positioning element is the poly(A) signal sequence similar to the mammalian A rich element (AAUAAA). It is located 10-30 nucleotides upstream of the cleavage site. The positioning element sequence and its distance from the cleavage site is critical for cleavage. The other elements present on either side of the cleavage site are generally U-rich. The cleavage site contains a pyrimidine followed by many adenosine residues. The cleavage reaction is similar to the mammalian cleavage reaction. Mutations in these elements do not eliminate cleavage but reduces its efficiency. The downstream and upstream U-rich elements are present in yeast and plants where as the mammalian system lacks these elements. The yeast system is more like the plant system rather than the animal system (Graber, Cantor et al. 1999).

The protein complexes that carry out the cleavage and polyadenylation in yeast are: 1) the Cleavage and Polyadenylation Factor (CPF); 2) the Cleavage Factor IA (CF IA); 3) the Cleavage Factor IB (CF IB) and; 4) the Poly(A) Binding Protein (Pab1p) (Fig.1.3). These will be described below and what is known about the protein subunits comprising them divulged.

1). Cleavage and Polyadenylation Factor is similar to mammalian CPSF but can be resolved into two sub-complexes, Cleavage Factor II (CFII) and Polyadenylation Factor I (PFI). CPF has additional factors named Pfs2, Ssu72, Mpe1, Glc7 and Ref2 that are required for 3' end processing. CFII is similar to the mammalian CPSF except the subunits homologous to CPSF30 and hFip1 belong to PFI. Fip1p was first identified in a yeast system by its interaction with Pap1. The yeast Pta1, a homolog of symplekin, is a component of CFII and is involved in both the cleavage and the polyadenylation reaction.

2). The CF IA possesses orthologues of both the mammalian CF II(m) and CstF complex but lacks an ortholog of mammalian CstF50. The mammalian CstF complex is vital for cleavage whereas yeast CF IA is necessary for both cleavage and polyadenylation. CF IA, CF IB and CFII are involved in cleavage while the polyadenylation reaction is driven by CF IA, CF IB, PFI, PAP1 and Pab1p. CF IB has a subunit, Hrp1, whose ortholog is not found in mammals. CPSF160 recognizes the A-rich hexamer sequence in mammals whereas its homolog in yeast, Yhh1p, binds near the A-rich element close to the cleavage site and functions in cleavage. Similarly CstF64 of mammals recognizes the downstream GU-rich element where as the yeast homolog Rna15 with the help of Rna14 recognizes the A-rich positioning element. So these homologs bind different elements though they have sequence homology. RNA14 also interacts with Fip1p but not strongly.

CF IA is made up of 4 subunits; Rna14, Rna15, Pcf11 and Clp1. Rna14 and Rna15 are the yeast homologues of CstF77 and CstF64, respectively. Rna15 has an RNA binding domain type RRM (RNA recognition motif) in its amino terminal region and involved in the recognition of the A-rich positioning element. Mutations in the conserved amino acids of Rna15 disrupt binding to the positioning element. Deletion of the C-terminus of

Rna15 protein slows down growth and leads to cell death. The C-terminus of Rna15 interacts with CF IA and Pcf11 and also with some transcription factors. Therefore, Rna15 can also regulate transcription. It also has glutamine and asparagine residues similar to the transcriptional regulators. Rna14 has a HAT repeat and binds to Rna15. Electron microscopy studies show that Rna14 and Rna15 can also form heterotetramers of the type 2:2. Heterologously expressed Rna14 exists as dimers. Pcf11 has leucine-zipper and interacts with the C-terminal domain of RNA polymerase II. Pcf11 has a stretch of glutamines that helps in the interaction with Rna14 and Rna15. Clp1 has a P-loop motif suggesting a role in ATP and GTP binding. CFIB has a single polypeptide called Hrp1/Nab4p. It has 2 RNA recognition motifs RNP1 and RNP2 along with a stretch rich in arginine and glycine. It interacts with Rna14 and Rna15 but not with Pap1. Arginine methylation facilitates the movement of Hrp1 from the nucleus to cytosol and vice-versa. Hrp1 participates in the usage of cleavage site but not in the cleavage reaction.

3). CFII has 4 polypeptides: Cft1/Yhh1, Cft2/Ydh1, Brr5/Ysh1, and Yth1. Cft1 and Cft2 are the homologues of mammalian CPSF160 and CPSF100. Cft2 is involved in the recognition of the efficiency element and poly(A) site. Brr5 is a homologue of mammalian CPSF73 and Yth1 is the yeast homologue of CPSF30. Yth1 has five zinc finger motifs and the second zinc finger is essential. The second zinc finger is highly conserved and mutations in the conserved cysteine residues are lethal. Mutations other than the conserved residues abolish cleavage activity and also affect RNA binding. The fourth and fifth zinc fingers are necessary for its interaction with Fip1p and Brr5/Ysh1p.

The polyadenylation factor PFI was initially identified as poly(A) addition factor. It contains Fip1p, Pap1, Yth1, Pfs1, and Pfs2. Fip1 contains acidic residues in the amino-terminal region and prolines in the carboxyl-terminus. It plays a major role in the assembly of polyadenylation factors during polyadenylation. Pap1 is a 64 KDa protein and has a nucleotidyl transferase activity.

4). The yeast poly(A) binding protein Pab1p is present in both the nucleus and cytosol. Nuclear poly(A) binding protein is encoded by *NAB2* gene and is essential for survival (Anderson, Wilson et al. 1993). Pab1 plays a key role in the deadenylation dependent mRNA turnover and also has a role in translation initiation. *In vitro* studies reveal that CF IA, CF IB and CFII are required for cleavage where as CPF, CF IA, CF IB and Pap1 are necessary for polyadenylation reaction. RNA polII increases the efficiency of both cleavage and polyadenylation and is not essential but important for 3' end processing.

Mechanism of polyadenylation in plants

Polyadenylation in plants requires at least three cis-elements. They are: 1) the far upstream element (FUE); 2) the near upstream element (NUE) and; 3) the cleavage/ polyadenylation site. The far upstream element (FUE) is a GU-rich element 50 or more nucleotides upstream from the cleavage or polyadenylation site (Fig.1.1). The near upstream element (NUE) is an A-rich element that may be analogous to the mammalian poly(A) signal AAUAAA; it is 10 to 30 nucleotides upstream to the cleavage/polyadenylation site. The cleavage/polyadenylation site is a U-rich region surrounding the YA cleavage site. The sequences of these elements are not highly conserved in plants at the nucleotide level.

The plant polyadenylation machinery consists of orthologues of the mammalian CPSF and CstF subunits along with CFI, CFII, poly(A) polymerases and the mammalian nuclear poly(A) binding protein. The CPSF complex in Arabidopsis includes: i) AtCPSF100; ii) AtCPSF73-I; iii) AtCPSF73-II; iv) AtCPSF30; v) AtFIPS5; vi) AtCPSF160; and; vii) AtFY (Fig.1.4). All these factors are in the nucleus suggesting their association in a complex (Elliott, Dattaroy et al. 2003) (Xu, Ye et al. 2004).

i). The core subunit of the CPSF complex is CPSF100. It is an essential gene and mutations that affect *CPSF100* are embryo lethal. AtCPSF100 interacts with the N-terminal domain of poly(A) polymerase (220 amino acids), an interaction not seen in other systems; this suggests that it is different from the mammalian and yeast orthologues.

ii and iii). Plants have two isoforms of CPSF73 (Hunt 2008). Both CPSF73-I and CPSF73-II are essential genes that reside on chromosomes I and II respectively, and mutations in either CPSF 73 gene cause developmental defects. In addition, CPSF73(I) over-expression affects female gametogenesis and plants with one copy of the CPSF73(II) gene show defects in male gametogenesis.

iv). The AtCPSF30 possesses three zinc finger motifs (Addepalli and Hunt 2007). AtCPSF30 is encoded by a gene, whose transcript is alternatively processed in Arabidopsis to produce a larger protein or a smaller protein. The function of the larger protein is not known but it might play a role in 3'end processing. The smaller protein, AtCPSF30, has an RNA binding activity. RNA binding is inhibited by calmodulin, suggesting a role of calcium signaling in 3'end processing. The small protein is homologous to Yeast, Yth1 and mammalian, CPSF30. Yeast CPSF30 is essential but plant CPSF30 is not, suggesting that there may be other proteins that carry out similar functions in plants.

v). AtFIPS5 interacts with several Arabidopsis PAPs and AtCPSF30 and is also an RNA binding protein. AtFIPS5 is also a component of the plant CPSF complex through its interactions mediated by AtCPSF30. This is similar to the mammalian counterparts, hFip1 and CPSF30, belonging to the CPSF complex and to the yeast Yth1p and Fip1p of the PF-I complex (Zhao, Hyman et al. 1999) (Mandel, Bai et al. 2008).

vi). AtCPSF160 is one of the conserved and a large subunit of CPSF and binds RNA in the process of 3' end processing. The subunits of CPSF complex interact with each other similar to mammalian and yeast counter parts. CPSF subunits also interact with Arabidopsis poly(A) polymerase (Xu, Zhao et al. 2006). AtCPSF160 forms stable complexes with FY and forms different sized complexes along with CPSF100 suggesting the role of CPSF subunits in RNA mediated chromatin modifications (Manzano, Marquardt et al. 2009). Interaction of AtCPSF160 with AtCPSF30 also alters CPSF30 localization from cytosol to nucleus (Rao, Dinkins et al. 2009). AtCPSF160 is also part of

the polyadenylation complex in pollen carrying out polyadenylation and is different from the canonical polyadenylation involving binding to the AU rich element (Hunt, Xu et al. 2008).

vii). AtFY; There are two orthologues of CstF50 in Arabidopsis, AtCstF50 and FY. FY is a WD 40 protein like CstF50 but plays a key role in the transition of vegetative- to the reproductive-phase. FY is a component of autonomous pathway and plays an important role in flowering. It is a 3'-end processing factor that regulates the levels of FCA, an RNA binding protein. FY in concert with FCA promotes flowering by repressing expression of FLC, a central repressor of flowering (Simpson, Dijkwel et al. 2003). FY represses FLC by chromatin modifications through FLD (a demethylase) (Liu, Quesada et al. 2007). FY also interacts with ubiquitin ligase, a component of the 26S proteasome degradation machinery however, the involvement of FY in the protein degradation pathway is not clear. AtCstF50 does not interact with CstF77 and shows a different interaction from that of yeast and mammalian CstF50. The presence of FY is a unique feature of the plant cleavage and polyadenylation complex and FY stably associates with CPSF160 and CPSF100. Lack of a functional FCA, or lack of the FCA interaction domain, disrupts formation of these complexes. FY is also involved in the chromatin modification by its association with FCA in the repression of FLC through FLD (Manzano, Marquardt et al. 2009).

AtCstF77 in plants also possess a HAT motif that is conserved in all eukaryotes. AtCstF77 interacts with CstF64, CPSF160 and CPSF30 and also binds RNA (Hunt, Xu et al. 2008; Bell and Hunt 2010). The Arabidopsis homologue of CstF64 has an RNA binding motif and binds RNA. CstF64-like homolog ESP1 associates with AtCPSF100 and AtSYMS5 (Herr, Molnar et al. 2006). In plants there might be two CstF complexes and one of these is a canonical AtCstF64 that binds the GU-rich downstream element through an RNA recognition motif and the other is similar in its role to ESP1 that, in association with an RNA binding protein, is involved in the identification of APA (alternative polyadenylation) sites (Zhao, Hyman et al. 1999; Zhao, Xing et al. 2009).

AtCLPS3 is an ortholog of mammalian hClp1 and AtCLPS3 might be involved in plant development-related pre-mRNA processing. It is considered a polyadenylation factor which, although it interacts peripherally with the CPSF complex through AtCPSF30, is not a core component of CPSF *per se* (Xing, Zhao et al. 2008). AtPCFS4 associates with AtCLPS3 in the same CSPF complex. AtPCFS4 regulates flowering through FCA where as its human (hPcf11) and yeast (Pcf11p) orthologues are known to function in transcription termination (Sadowski, Dichtl et al. 2003; Zhang, Klatt et al. 2007). The ortholog of symplekin, AtSYM5, might have a role in gene silencing (Herr, Molnar et al. 2006).

There are four isoforms of poly(A) polymerase in Arabidopsis. One isoform, (At3g06560, PAPS3) is shorter (482 amino acids) while other 3 isoforms PAPS1 [(AT1g17980)], PAPS2 [(At2g25850)] and PAPS4 [(At4g32850)]) produce proteins between 700 and 800 amino acids. PAPS3 is expressed in the cytosol while other three are expressed in the nucleus. All four isoforms are alternatively spliced, single gene products expressed in a tissue specific manner (Meeks, Addepalli et al. 2009).

There are PABP's in nucleus as well as in the cytoplasm. Multiple isoforms of cytoplasmic poly(A) binding proteins (PABPC) are common in plants and animals. There are eight PABP genes in *Arabidopsis thaliana.* Four classes of PABPC's were identified in Arabidopsis by expression analysis and phylogenetic comparisons. Class I includes PAB3 and PAB5 and their expression is restricted to the reproductive tissue. Class II has members PAB2, PAB4 and PAB8 that are expressed more. Class III composes of PAB6 and PAB7 that have low expression. Class IV contains only PAB1 with weak tissue specific expression (Mangus, Evans et al. 2003). Duplication events were thought to be responsible for the generation of classes I, II and III from comparison studies of rice and Arabidopsis that might have occurred before 200 million years. Loss of introns and conserved nature of PAB's suggest that PABP gave origin to class II, III and IV while class I is derived from class II. The main function of PABP is to bind the poly(A) tail of mRNA as a translation initiation factor. Association of cap binding proteins and PABP's bring the two ends of an mRNA together and enhance translation (Dufresne, Ubalijoro et

al. 2008). Cytosolic PABP's play an important role in the stability of mRNA, translation initiation, translation termination and ribosome recycling.

Nuclear poly(A) binding protein (PABN) is encoded by single gene in animals but there are three isoforms of PABN in *Arabidopsis thaliana* (Hunt, Xu et al. 2008). PABN1 interacts with Fip1 and this interaction is specific to the plant system with no reports of such interaction in other systems. Association of all the three forms of PABN along with four isoforms of PAP with AtFIPS5 complex was interesting and shows similarities to the functions of Fip1 in mammals in binding RNA and stimulating PAP activity but not in its association with poly(A) binding protein (Forbes, Addepalli et al. 2006). Nuclear PABP's are essential for poly(A) tail synthesis, regulation of poly(A) tail length, mRNA maturation and also facilitates mRNA export from nucleus to the cytosol. Both nuclear and cytosolic PABP's aid in mRNA function (Mangus, Evans et al. 2003).

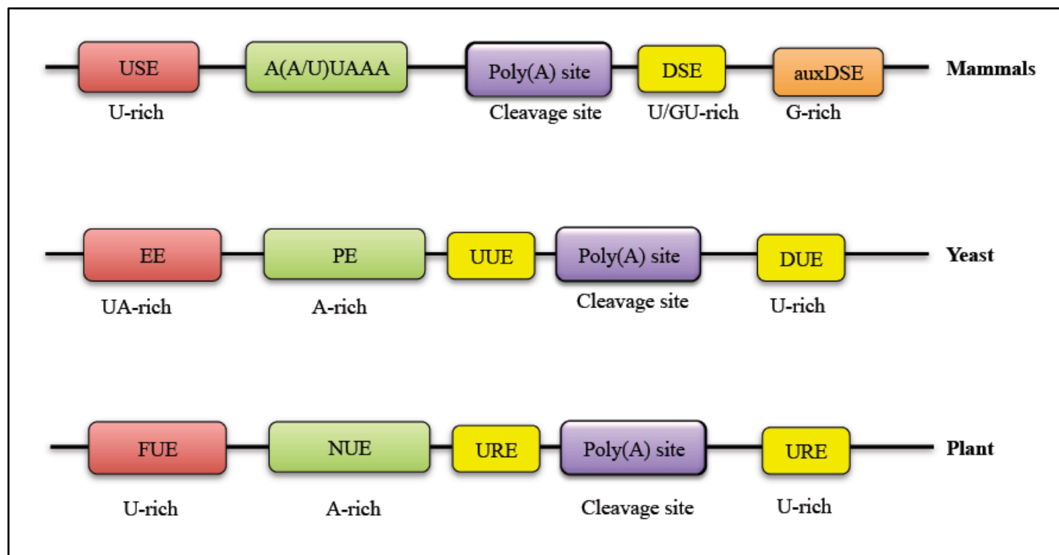


Figure 1.1 Cis-elements in Mammals, Yeast and Plants that guide the process of cleavage and polyadenylation. Elements in the boxes represent cis-elements and sequence below the box represent composition of the cis-element and similar color code denotes similar elements among these systems.

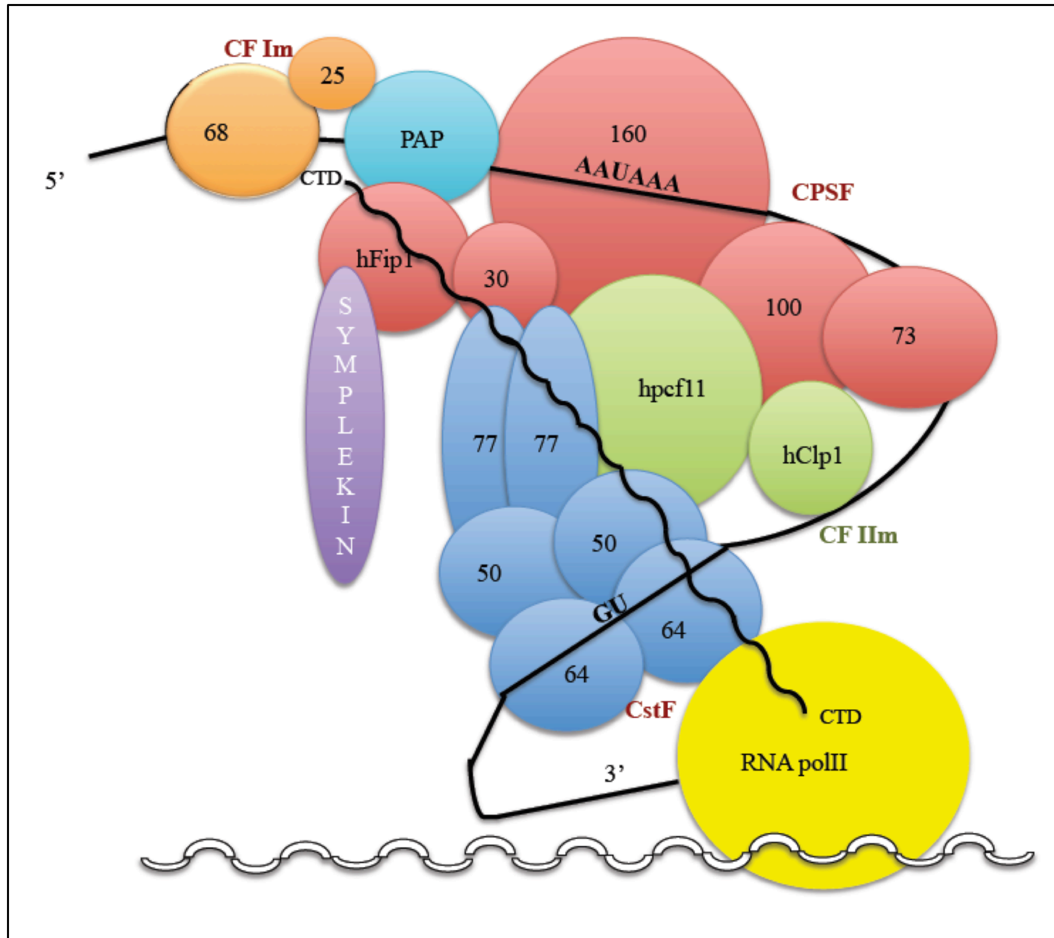


Figure 1.2 Mammalian 3' end processing machinery highlighting CPSF, CstF and its interactions with RNA at A rich element and GU rich element, poly(A) polymerase, CF Im and CF IIm along with symplekin and association with RNA polII (For more details on each subunit see description provided in the chapter).

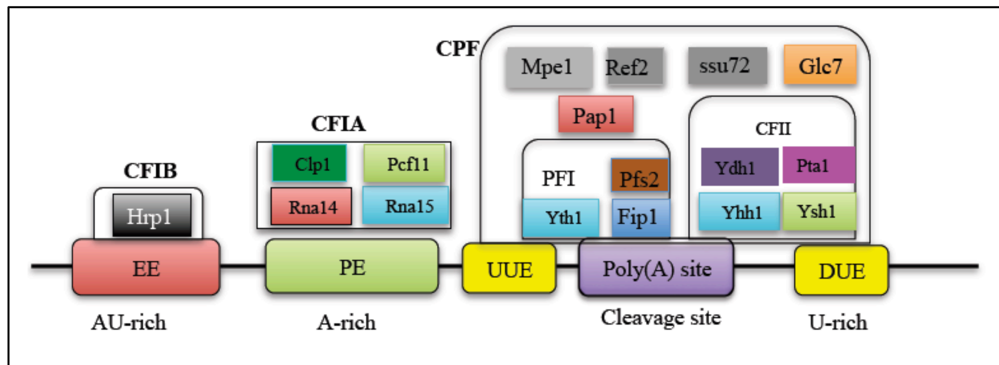


Figure 1.3 Yeast 3' end processing machinery showing complexes CPF, CF IA, CF IB and CFII. Each complex has multiple subunits and interactions among the multi-subunit complexes and RNA with the help of cis-elements guide cleavage and polyadenylation (For more details on each subunit see description provided in the chapter).

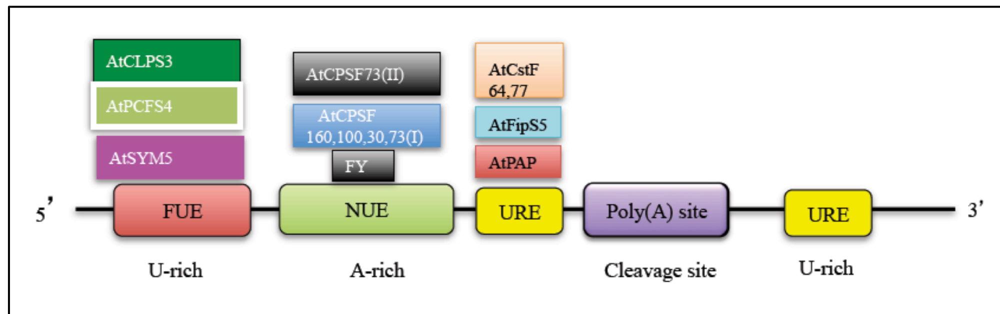


Figure 1.4 Plant 3' end processing machinery showing complexes CPSF, CstF, Poly(A) polymerase along with other complexes involved in plant cleavage and polyadenylation. Each complex has multiple subunits and interactions among the multi-subunit complexes and RNA with the help of cis-elements guide 3' end processing in *Arabidopsis thaliana* (For more details on each subunit see description provided in the chapter).

Copyright © Lavanya Dampanaboina 2011

CHAPTER TWO

Functional characterization of the Arabidopsis orthologue of Mammalian CstF-50 (Cleavage stimulation Factor subunit-50)

Introduction

All eukaryotic mRNA's are modified at their 3' ends for efficient translation of mRNA into protein (Dichtl and Keller 2001). Polyadenylation is a post-transcriptional modification initiated by endo-nucleolytic cleavage, followed by addition of poly(A) tail (Colgan and Manley 1997). This process results in mature mRNA from all RNA polymerase II transcription products. The process of polyadenylation is mediated by numerous protein factors that interact with each other and with pre-mRNA for cleavage and poly(A) tail addition (Keller 1995). The pre-mRNA is endo-nucleolytically cleaved at the cleavage site i.e. 10-30 nucleotides downstream from the poly(A)site, followed by addition of 200-300 (A) residues at the 3'end (Lutz 2008). Polyadenylation occurs in the nuclear compartment along with other post-transcriptional modifications such as capping and splicing. In mammals, many protein factors were identified with specific functions in the polyadenylation machinery. For example, the multi-subunit protein complexes, Cleavage and Polyadenylation Specificity factor (CPSF) and Cleavage stimulation Factor (CstF), associate with each other to identify the poly(A) site using a recognition sequence (AAUAAA) and a downstream GU/U-rich sequence for efficient cleavage and polyadenylation. These complexes in turn associate with Cleavage Factor I, Cleavage Factor II, and Poly(A) Polymerase (PAP) to catalyze 3' end processing.

In mammals, the Cleavage stimulation Factor is a hetero-trimeric protein made up of three sub-units: i) CstF-64, ii) CstF-77 and, iii) CstF-50 (Takagaki and Manley 2000). The CstF complex is necessary for the cleavage step and also determines the efficiency of polyadenylation (MacDonald, Wilusz et al. 1994). This complex also stabilizes the CPSF complex bound to the poly(A) signal sequence for efficient cleavage of the 3' end of pre-mRNAs before adding the poly(A) tail. One of the subunits of CstF is a WD repeat-containing protein of about 50 kDa (CstF50). CstF50 interacts with CstF77 and WD motif is necessary for its interaction. CstF50 also interacts with RNA polymerase II CTD

(C-terminal domain) and is critical for mediating 3' end processing (McCracken, Fong et al. 1997) (Fong and Bentley 2001). CstF-50 also interacts with BARD1 (BRCA1 associated ring domain protein) which associates with the breast cancer susceptibility gene product BRCA1 (Breast Cancer type I Susceptibility protein) (Tsuzuki, Wu et al. 2006). *In vitro* studies show that the interaction of CstF-50 with BARD1 represses the polyadenylation machinery (Kleiman and Manley 2001). Therefore, 3' end processing also rescues from DNA damage and tumor formation by inhibiting polyadenylation by the association of CstF50 with DNA repair proteins (Kleiman and Manley 2001).

The Arabidopsis orthologue of mammalian CstF-50 is encoded by, At5g60940. Like its mammalian orthologue AtCstF-50 is also a WD repeat protein but contains six WD repeats and shows 55% similarity and 37% identity to human CstF-50. A reverse genetics approach was employed in the present study to address the question whether AtCstF-50 is essential or not in plant cleavage and polyadenylation. This study uses two T-DNA insertion mutants obtained from the Arabidopsis Biological Resource Center (ABRC). These mutants were analyzed by genotyping the lines using PCR. In addition, in order to test the role of AtCstF-50 in the network of protein complexes involved in cleavage and polyadenylation, the yeast two-hybrid assay was used. Interactions of AtCstF-50 with other polyadenylation factors help us understand the importance of AtCstF-50 and to see how different or similar it is to its yeast and mammalian counterparts. Finally, the subcellular distribution of AtCstF-50 was also tested.

MATERIALS AND METHODS

PCR genotyping of T-DNA insertions of *CstF50.1.1* and *CstF50.1.2*

Two T-DNA insertion lines (SALK_075594 and SALK_075595) were identified in the SALK Institute Genomic Analysis Laboratory (SIGnAL) database (Alonso, Stepanova et al. 2003) and obtained from The Arabidopsis Biological Resource Center (ABRC, The Ohio State University, Ohio, USA). The first line, SALK_075594, has a T-DNA insertion in the third intron; in this chapter, the mutant is called *CstF50.1.1*. The other line, SALK_075595, has the T-DNA insertion in the fourth exon; in this chapter, the mutant is called *CstF50.1.2*. The seed that was obtained from TAIR was grown in the seed house at 72^0F under short day conditions (8 hr light, 16hr dark) and seed collected from all F_1 plants was mixed together. The bulked seed was germinated and each seedling was transplanted into individual pots. Leaf material was collected from 35 to 50 individual plants and frozen in liquid nitrogen. Genomic DNA was made using DNazol (Invitrogen Inc., Carlsbad, CA, USA). Genotyping was done using the Polymerase Chain Reaction (PCR). Both gene specific and T-DNA specific primers were used along with T-DNA border primers. To determine whether T-DNA is present or not, both TL (T-DNA Left border) and TR (T-DNA Right border) primers were used in combination with CstF-50 gene specific forward and reverse primers as the orientation of the T-DNA insertion was initially unknown. Similarly PCR reactions were also performed using gene specific primers and T-DNA-specific internal gene primers. Once the genotypes of the plants were determined in each line, T-DNA homozygotes were further confirmed by Southern blotting.

Southern blot analysis of *CstF50.1.1*

Four individual plants (putative T-DNA insertion plants 23, 33, 43 and 45, obtained from Amanda Marion) along with wild type Columbia 0 were selected from the CstF-50 T-DNA insertion line SALK_075594. Leaf material was collected from individual plants and genomic DNA was extracted using DNazol (Invitrogen Inc., Carlsbad, CA, USA). Approximately 5-10 μg of DNA were digested with *Eco*RI and the digested genomic

DNA was separated on 0.8 percent agarose gel. The separated DNA fragments were photographed using a UV trans-illuminator before transferring to the nylon membrane.

After electrophoresis, the gel was cutoff at the bottom left corner. Then the agarose gel containing DNA was depurinated, denatured, neutralized and transferred to a nylon membrane using the capillary transfer method (as described in Molecular cloning second Edition by Sambrook, Fritsch and Maniatis pg 9.34 section 1). Transfer buffer 20X SSC (3M NaCl, 0.3M sodium citrate, pH 7) was used for capillary transfer. After overnight transfer of DNA from the agarose gel to the membrane, the membrane was washed in 6X SSC for 5 minutes to remove any traces of agarose and air-dried for 30 minutes on Whatmann paper (Whatmann Inc., Maidstone, Kent, UK).

Random primer labeling of DNA

The gene specific and T-DNA probes were synthesized using the Prime-It II Random Primer Labeling Kit (Stratagene Inc., La Jolla, California, USA). Both T-DNA and gene specific products were amplified using PCR. The PCR product was gel purified using Qiagen DNA purification columns (Qiagen Inc., Valencia, CA, USA). At least 100ng of DNA was used for probe synthesis. The initial reaction was set using 6µl of probe DNA and 10 µl of random primer made up to a final volume of 34 µl with water. The sample was labeled and heated at 95^0C for five minutes followed by brief chilling on ice. Then 10 µl of dCTP buffer, 1µl Klenow (Promega, Madison, WI, USA) and 5 µl of γ^{32}P-dCTP (New England Nuclear, Boston, MA, USA) was added to the above reaction mix, vortexed and incubated at 37^0C for 30 minutes. After incubation the reaction was stopped by adding 2µl of stop mix (0.5 M EDTA, pH 8.0). The above reaction mix (50µl) was passed through a Sephadex G-50 spin column to separate the radiolabeled probe from unincorporated dNTPs (as described in Appendix E, Molecular cloning second Edition by Sambrook, Fritsch and Maniatis). "Fractions" of four drops of eluate were collected and the radioactivity was determined by scintillation counting. Tubes with higher radioactivity were pooled together and boiled for five minutes and cooled on ice for 3 minutes. Both gene specific and T-DNA probe DNA's were made using the above procedure.

Hybridization

The nylon membrane was separated from the gel and air-dried at room temperature. Then it was UV cross-linked using a Stratagene UV crosslinker and hybridized with DNA probe. Before hybridization the membrane was put in the pre-hybridization buffer (0.12M $NaHPO_4$ pH 7.2, 0.25M NaCl, 7% SDS and 1mM EDTA) and incubated at 65^0C for one to two hours. Freshly prepared probe was added to the freshly made hybridization buffer, added to the filter, and incubated overnight at 65^0C with agitation. After hybridization, the membrane was washed with 2XSSC for 10 minutes at 30^0C or at room temperature and, subsequently, with 0.1%SSC, 0.1%SDS 2 times, for 20 minutes each at 65^0C. The membrane was wrapped in saran wrap and exposed to the phosphor-imager screen for three to four days. A Typhoon 9400 scanner was used to measure the signal using Image Quant software.

Reverse transcription

RNA samples were treated with RNase free DNaseI (Fermentas, Glen Burnie, MD, U.S.A.), at 37^0C for one hour. DNase treated samples were cleaned using Qiagen RNA columns (Qiagen Inc., Valencia, CA, USA) and used for making cDNA. An oligo-dT primer and M-MLV reverse transcriptase (Fermentas, USA) was used according to the manufacturer's instructions. The RNA in the RNA-DNA hybrid was digested with RNAaseH (New England Biolabs, Ipswich, MA, USA) at 37^0C for 1 hour after heat killing RT at 70^0C for 5 minutes. PCR was done using cDNA template with gene specific primers and checked on an agarose gel.

Yeast two-hybrid assay

Yeast two hybrid assay was performed to identify interactions of CstF-50 with other Arabidopsis polyadenylation factors (Table 2). The vectors used in this study are pGAD-C(1) with the activation domain and pGBD-C(1) with the binding domain. The plasmid containing the activation domain and the plasmid containing the binding domain was transformed into the yeast strain PJ69-4A using a transformation protocol (James, Halladay et al. 1996; Hunt, Xu et al. 2008). The PJ69-4A yeast strain is the Gal 4 reporter based system employed in these studies. The genotype of PJ69-4A is: MATa trp1-901

leu2-3, 112 ura3-52 his3-200 gal4del gal80del LYS2::GAL1-HIS3 GAL2-Ade2 met2::GAL7-lacZ. Fresh competent cells of PJ69-4A were made using the polyethylene glycol and lithium acetate method (Gietz et al. 1992). Protein coding sequences of all polyadenylation factors listed in Table 2 were cloned in both AD and BD vectors as described (Xu, Zhao et al. 2006). For reciprocally testing interactions with all polyadenylation factors (listed in Table 2), first the AD test construct containing AtCstF-50 (AD50) was co-transformed with a BD plasmid containing each Arabidopsis polyadenylation factor. Similarly the BD test construct containing AtCstF-50 (BD50) was also co-transformed with each of the Arabidopsis polyadenylation factors in the activation domain plasmid. Negative controls were made for all the polyadenylation factors that were tested for interaction with the complementary empty plasmid (ADX with BD and BD50 with AD respectively). AD64 and BD77 plasmids co-transformation is used as a general positive control for all interactions and other interactions were identified in comparison with the positive control. The empty vectors AD and BD plasmids are co-transformed and used as negative control for all interactions.

Yeast tranformants were selected on synthetic media without leucine and tryptophan (-LW) amino acids. The yeast strain transformed was plated on -LW selection media and incubated at 30^0C for 3-5 days until yeast colonies appeared. For high stringency, the yeast transformants were selected again on synthetic media lacking histidine, leucine and tryptophan (-HLW) along with 3-amino-1, 2, 4-triazole (3-AT, a competitive inhibitor of histidine production) (Forbes, Addepalli et al. 2006) (Hunt, Xu et al. 2008). The transformation mixture was plated on –LW double selection media. The double tranformants (growing on double selection media) were picked and restreaked on –HLW media to identify the interacting proteins. Only colonies with interacting proteins grow on the triple selection media.

CstF-50 localization

The CstF-50 coding region was fused, in frame, to the Red Fluorescent Protein (RFP) coding region carried in the pGDR binary vector (Goodin, Dietzgen et al. 2002) and the localization pattern of CstF-50 was studied by introducing pGDR into *Agrobacterium*

tumefaciens strain LBA4404. Upon establishment in Agrobacterium, the chimeric gene encoding CstF-50::RFP was introduced transiently into *Nicotiana benthamiana* leaves using Agro-infiltrations for expression studies. Nuclear localization markers for positive controls were made by cloning a seven amino acid long (PKKKRKV) monopartite, nuclear localization signal of SV-40 large T-antigen into pGDG and pGDR vectors (Kalderon, Roberts et al. 1984). Oligonucleotides containing SV40 NLS with *Bam*HI over hangs were annealed and ligated with *Bgl*II digested pGDG and pGDR vectors. Other markers including chloroplast and GFP markers were kindly provided by Dr. Michael Goodin, Department of Plant Pathology, University of Kentucky, USA.

All the constructs along with control vectors were transformed into Agrobacterium strain LBA4404 using the freeze thaw method (An, Ebert et al. 1988). Agrobacterium was freshly grown on LB agar plates with respective antibiotics for 2 days at 28^0C. The cultures plates were scraped with a spatula and suspended in the freshly made MES buffer (10mM $MgCl_2$ and 10mM MES). The suspensions were measured for the cell density in the spectrophotometer at 600nm. Using the optical density (OD) readings, the cell density was set to 0.8 OD. Acetosyringone was added to all samples at 150μM concentrations and incubated at room temperature for 2-3 hours. For co-infiltrations equal volumes of different Agrobacterium samples were mixed before infiltrations. The GFP reporter with the NLS marker was used as an internal control in all co-infiltrations as all fusion constructs are made with RFP.

Three weeks old, *Nicotiana benthamiana* plants were selected and leaves infiltrated with the Agrobacterium suspensions using a 1ml syringe without a needle. Healthy looking and fully expanded lower leaves were selected for infiltration experiments. Gentle pressure was applied with a finger on the adaxial surface of the leaf while injecting the suspension buffer from the abaxial surface of the leaf. The desired amount of the buffer with bacterial suspension was infiltrated in the intravenous regions. More than two infiltration sites were made with the tip of the syringe to complete the infiltration in a single leaf. At least two leaves were infiltrated with the same suspension on a single

plant. The plants were incubated for 48 to 72 hours under 16hr/8hr light and dark photoperiods.

Infiltrated leaves (after 48-72hr of incubation) were made into small sections near the infiltration site. The section, with the abaxial side facing the cover slip was placed on a glass slide in a drop of water and mounted gently with a cover slip. The slide was observed under UV light with a Zeiss Axioplan2 HB100. There is an additional filter in the microscope called DIC (Differential Interference Contrast) where the unstained tissue can be visualized and this can be used to visualize both GFP (excites at 488nm) and DsRed (excites at 543nm) markers. The GFP expressing epidermal cells were visualized using the green filter (ex: D470/40; D535/40; beam splitter 500 DCLP) and DsRed is observed under the red filter (ex: HQ545/30X; em: HQ610/75M, Q570LP) filter sets. Images were captured using the Zeiss AxioCam MRc5 attached to the microscope. AxioVision was used to adjust the exposure for each filter and also get the merged picture for the co-infiltration samples along with the pictures in green and red filters.

RESULTS

Characterization of T-DNA insertion mutants *CstF501.1 and CstF501.2* in the gene encoding AtCstF-50

A reverse genetics approach was used to address the question whether *AtCstF-50* (At5g60940) is essential to any stage of plant development or not. The gene has 13 exons and 12 introns (Fig.2.1). Two T-DNA insertions in CstF-50 were genotyped using PCR. Some putative T-DNA mutants (plants 23, 33, and 45) were identified previously (Amanda Marion, unpublished data) from a kanamycin resistance screen of individual plants from SALK line 075594 that has an insertion in the third intron of the gene (Fig.2.1 and Fig.2.2). These results were confirmed by PCR using new sets of primers (Fig.2.3). Amplification using border primers TR or TL, each in combination with a gene specific forward primer, show that the T-DNA was inserted in the third intron in the reverse orientation as the amplification was observed with the TR primer (Fig.2.4B). Plants 23 and 33 were homozygous for the T-DNA insertion and plant 45 was heterozygous for the T-DNA insertion. Plant 43 lacked the T-DNA and was thus concluded to be a kanamycin-resistant "escape."

To confirm the PCR results, genomic DNA of individual plants 23, 33, 43, and 45 along with that from the wild type was analyzed by Southern blotting. The T-DNA insertion is between third and fourth exon. In the T-DNA there are two *Eco*RI sites while in the genomic DNA, flanking the T-DNA, there are three3 more *Eco*RI sites (Fig.2.5b). In preparation for the Southern blot, genomic DNA was digested with *Eco*RI (Fig.2.6). The blots were hybridized to the T-DNA specific probe to confirm the presence of the T-DNA in both T-DNA homozygotes (plants 23 and 33) and the heterozygote (45) (Fig.2.7A). Blots hybridized to the gene-specific probe show two bands in the T-DNA homozygotes and one band in wild-type sample (Fig.2.7B); the sizes of the bands (4.4kb and 1.4kb) are consistent with the restriction map and confirm the PCR-based genotyping. Reverse transcription-PCR reactions using RNA isolated from above plants showed that the *CstF-50* gene was expressed in all four plants (Fig.2.8). The results also indicate that the intron

containing the T-DNA in the homozygous lines was spliced out, allowing the production of a functional mRNA.

When the second insertion line in the fourth exon of *CSTF-50* (*CstF50 1.2*) was studied, only wild-type plants and individuals heterozygous for the T-DNA insertion were obtained in a 1:2 ratio. Fourty five plants are screened in this process (Fig.2.9). This result thus suggests that a functional *CstF-50* gene product is essential at some stage of the Arabidopsis life cycle.

In order to test whether AtCstF50 has any role in the embryo development, seeds were collected from individual siliques of heterozygous plants of *CstF50 1.2.* Seeds were made transparent and observed under microscope for embryo defects. Severe embryo defects were observed in some seed but this result was not consistent with all the siliques containing seed. These results suggest that it might be possible that AtCstF50 might have a role in the post germination effects rather than embryo development itself. Future work needs to be focused in these lines to identify the exact role of CstF50 in seed development.

The AtCstF-50 protein-protein interaction network

A previous report indicated that the Arabidopsis CstF-50 protein did not interact with the Arabidopsis CstF-77 or CstF-64 proteins (Yao, Song et al. 2002), suggesting that the plant protein may differ from its mammalian counterpart. To further explore this, yeast two hybrid assays were employed to study the interactions of CstF-50 with other polyadenylation factor subunits (Table 2). Most of the polyadenylation factors cloned in AD and BD vectors were full length except FIPS5 where partial protein coding clones (Hunt, Xu et al. 2008) with the first 137 amino acids, designated as the N-terminal domain (FIPS5NTD), and last 500 amino acids as the C-terminal domain (FIPS5CTD) were used (Forbes, Addepalli et al. 2006). The strong, positive interaction observed between AtCstF-64 and AtCstF-77 was used as a positive control. Similarly, co-transformants of empty AD and BD plasmids were used as negative controls. Interactions with different polyadenylation factors were tested in both combinations (reciprocal

testing). For example, the AD::CSTF-50 (AD50) was tested with BD::PAPIV and AD::PAPIV was tested with BD::CstF-50. When reciprocal testing was performed, it became evident that some polyadenylation factors have self-activation domains so they could only be used in the AD plasmid as their introduction into the BD plasmid resulted in self-activation, which confounded interpretation of interactions. The other combination, where these clones with activation domains reside in the AD plasmid, can still help in understanding their interactions with proteins in the BD plasmid. The negative control for each plasmid used in the interaction was also performed (e.g. AD::PAPIV with BD and BD::CstF50 with AD).

The results of this study showed that AtCstF-50 interacts with subunits of Cleavage and Polyadenylation Specificity Factor, CPSF100 and CPSF30 (Fig.2.10A). AtCstF-50 also interacts with two poly(A) polymerase isoform (PAPIV and PAPIII) (Fig.2.10B and Fig.2.10C), with one nuclear poly(A) binding protein isoform (Fig.2.10D) and with both N-terminal and C-terminal domains of FIPS5 (Fig.2.10E and Fig2.10F). In contrast, no interactions were seen with CstF-77 or CstF-64. Negative results were also obtained with the other combinations, also suggesting a lack of interactions.

Localization of AtCstF-50 to the Endoplasmic Reticulum

The yeast two hybrid interaction study show that AtCstF-50 is different from the mammalian CstF-50 (Hunt, Xu et al. 2008). To further explore this, the coding region of *AtCstF-50* was translationally fused to the *RFP* coding sequence and its subcellular localization was tested along with a GFP nuclear marker. Distinct spots appear that do not colocalize with the GFP nuclear marker as seen in Fig. 2.11 panel 1 A, B, and C. To investigate further where AtCstF-50 localizes, other markers were employed in the study. A chloroplast-targeted GFP marker was co-infiltrated with RFP translationally fused to AtCstF-50. Results show that AtCstF-50 did not colocalize with the chloroplast GFP marker (Fig. 2.11, panel 2 D, E, F). However, when the AtCstF-50-RFP construct was co-infiltrated with an ER-targeted GFP marker, the RFP colocalized with the endoplasmic reticulum marker (Fig. 2.11, panel 3 G, H, I).

DISCUSSION

To test whether AtCstF-50 is essential or not in the life cycle of Arabidopsis and also how important it is in the cleavage and polyadenylation process, two approaches were employed. One was a reverse genetics approach using T-DNA insertions obtained from Arabidopsis Biological Resource Center (ABRC). The other approach was a series of yeast two hybrid assays to test how AtCstF-50 interacts with other factors. The present study with T-DNA insertion mutants indicates that AtCstF-50 is essential. One mutant with a T-DNA insertion in the third intron of the gene encoding CstF-50 intron yielded homozygous plants that accumulated properly spliced mRNAs. Therefore, the intron possessing T-DNA sequences might be spliced out and these plants are able to produce functional AtCstF-50. A mutant with a T-DNA insertion in the fourth exon yielded no homozygous plants in an F2 population. This suggests that insertion of T-DNA within the exon might cause some severe defects, and is consistent with the conclusion that the plant cannot survive without functional AtCstF-50 gene products. These results suggest that CstF-50 is required for a functional 3' end processing machinery.

The yeast two-hybrid results show that AtCstF-50 does not interact with either AtCstF-64 or AtCstF-77. This confirms an earlier report (Yao, Song et al. 2002) and reveals a substantial difference between plants and animals. This raises the possibility that the plant CstF-50 might be part of different complex and not part of the CstF complex.

The interaction of AtCstF-50 with at least one PAP isoform suggests a different link that has not been reported in any other systems. Using computational tools poly(A) polymerases were identified in eight plant genomes ranging from one to six forms. In Arabidopsis there are four isoforms of poly(A) polymerases that are expressed in a tissue specific manner (Addepalli, Meeks et al. 2004). In other organisms, different poly(A) polymerases function differently from the regular poly(A) polymerases; one such example is the testis specific poly(A) polymerase in the mouse (Kashiwabara, Noguchi et al. 2002) (Kashiwabara, Zhuang et al. 2000) (Lee, Lee et al. 2000). The interaction of AtCstF-50 with PAPIV supports the proposal that CstF-50 might be part of CPSF

complex instead of a plant CstF complex, and that it might have a role in cleavage and polyadenylation as opposed to just a cleavage role as is the case in Mammals.

In mammals, PAP interacts with Fip1, CPSF, CFI-25 and PabN (Thuresson, Astrom et al. 1994; Murthy and Manley 1995) (Kaufmann, Martin et al. 2004) (Kim and Lee 2001; Kerwitz, Kuhn et al. 2003). In yeast, PAP also interacts with Fip1 and CFI (Kessler, Zhelkovsky et al. 1995) (Preker, Lingner et al. 1995). In Arabidopsis, all four PAP isoforms interact with CPSF100, CPSF30, FipS5, PabN and CFIS (Elliott, Dattaroy et al. 2003 ; Forbes, Addepalli et al. 2006; Hunt, Xu et al. 2008). The AtCstF-50-PAP interaction suggests that AtCstF might be involved in the poly(A) polymerase reaction itself, something that is not the case in mammals where it plays a role in cleavage.

AtCstF-50 also interacts both with FIPS5NTD and CTD domains. FIPS5 is an RNA binding protein like mammalian orthologue and interacts with RNA and bridges different polyadenylation factors with poly(A) polymerase. The FIPS5 N-terminal domain interacts with CstF-77, CPSF30, CFIm-25, poly(A) polymerase and the nuclear poly(A) binding protein where as the C-terminal domain has RNA binding domain and interacts with CstF-64 (Forbes, Addepalli et al. 2006). It is not clear how these interactions of AtCstF-50 with FIPS5 might affect the process of polyadenylation.

AtCstF-50 also interacts with one of the three isoforms of the nuclear poly(A) binding protein

AtCstF-50 also interacts with two of the subunits of CPSF, AtCPSF100 and AtCPSF30. The interaction of CstF-50 with CPSF100 has been confirmed by pull down assay (Dr. Hunt, unpublished results). The Arabidopsis CPSF complex is composed of AtCPSF160, AtCPSF100, AtCPSF73-I, AtCPSF73-II, AtCPSF30, AtPAPS2 and FY. In mammals this complex consists of CPSF160, CPSF100, CPSF73 and CPSF30. Immunopurification studies in plant nuclear extracts show the presence of CPSF subunits (CPSF160, CPSF100, CPSF73(I), CPSF73(II), CPSF30) in a complex (Baillat, Hakimi et al. 2005; Herr, Molnar et al. 2006); however, these experiments failed to identify AtCstF-50 in the same complex.

With these interactions summary, a model can be developed where AtCstF-50 is a subunit of CPSF complex interacting with AtCPSF30 and AtCPSF100. AtCstF-50 being a WD repeat protein might be interacting with other subunits of CPSF rather than having direct interaction with the mRNA being processed for cleavage and polyadenylation. The interactions with FIPS5 and poly(A) polymerase might influence the interaction of CPSF and PAP that is required for the polyadenylation of mRNA.

The polyadenylation machinery is localized to the nucleus (Meeks, Addepalli et al. 2009; Rao, Dinkins et al. 2009) and subunits that do not have a nuclear localization signal are presumably targeted to the nucleus through protein-protein interactions (Rao, Dinkins et al. 2009). The localization studies presented here suggest that interactions with other polyadenylation factors are probably important for the nuclear localization of AtCstF-50.

Mammalian CstF-50 has six WD repeats and WD motif is responsible for its interaction with CstF-77 and its localization to the nucleus. The interaction summary of AtCstF-50 shows that it is very different from its mammalian counterpart. It interacts with AtCPSF30, AtCPSF100, PAPS3 and PAPS4. AtCPSF30, an RNA binding endonuclease forms the central factor in the network of interaction bridging cleavage and polyadenylation complexes (Hunt, Xu et al. 2008). A recent report on AtCPSF30 transient studies suggest that its interaction with AtCPSF160 and AtCPSF73(I) (Xu, Zhao et al. 2006) changes its localization from cytosol to nucleus. SimilarlyAtCPSF30 can also change the localization of AtCPSF100 to the cytosol (Rao, Dinkins et al. 2009).

From these studies, it would seem that when a particular factor interacts with multiple factors or part of a multi-subunit complex, this association might change the pattern of localization of the group. It would be interesting to examine the biological roles of cleavage and polyadenylation factors outside of the nucleus. Additionally, it would be interesting to co-express positive interacting partners of AtCstF-50 with AtCstF-50 to see if AtCstF-50 then localized to the nucleus or still to the ER. All polyadenylation factors are necessary for nuclear polyadenylation and localization of these polyadenylation factors to other compartments suggests that it might be a mechanism for the dissociation

of the polyadenylation complexes or necessity of the same factors for cytoplasmic polyadenylation or transport of the mRNA by the association of some of these polyadenylation factors to the target site or a mechanism for down-regulating levels of AtCstF50 in plants to maintain a stoichiometric ratio of CstF50 protein.

It might be possible that there might be some kind of mechanism by which AtCstF50 association with a particular polyadenylation factor is favored. The signaling mechanism might be mediated by calmodulin as observed in case of CPSF subunit AtCPSF30 which is a known calmodulin binding protein, or perhaps a signaling pathway mediated by kinases or phosphatases. These mechanisms might affect the accessibility or inaccessibility of interacting domains, which in turn would impact its localization, expression, or degradation.

To summarize, these studies show that AtCstF-50 might have different roles in the plant, and other proteins that interact with this protein might affect its functioning in polyadenylation in the nucleus and elsewhere.

Table 2.1 List of primers used for genotyping of CstF50 T-DNA insertion lines.

Name of the primer	Primer sequence	Purpose
T-DNA specific primer	Forward primer (7104): CTT ATA TAG AGG AAG GGT CTT GCG AA	T-DNA mutant screening
	Reverse primer (7920): GCA GGT CCC CAG ATT AGC CTT TTC AA	
T-DNA border primers	Left border primer: AAA CTG GAA CAA CAC TCA ACC CTA TCT CGG	Screening *CstF-50* mutants
	Right border primer: TTC TCC GCT CAT GAT CAG ATT GTC GTT TCC	
Gene specific primers for *CSTF-50*	Forward primer: GGT CTT GCA GCG GAG AAC AAT GGG A;	Screening *CstF-50-1-1* mutants
	Reverse primer: CTT TGA TGA GCC CTT CGC ATG ATT TAA CA	
NPT2 primers	Forward Primer: CCT GTC CGG TGC CCT GAA TG	Confirm T-DNA presence
	Reverse primer: CCA CAG TCG ATG AAT CCA GAA AAG	
CSTF-50 RT primers	Forward primer: GGA GGA GGA AGA AGA CGA TGG TTG A	For RT reaction
	Reverse primer: CTT TGC AAC AAG CTC CAA AAG ACG	
T-DNA probe DNA	Forward primer: CTT ATA TAG AGG AAG GGT CTT GCG AA	T-DNA PCR product
	Reverse primer- GCA GGT CCC CAG ATT AGC CTT TTC AA	
CSTF-50 (gene specific) probe DNA	Forward primer (601): GGA GAA CAA TGG GAC ATT GAG A	CstF 50 PCR product
	Reverse primer(1601): CCT GCA GAA TTT GGA GTT AAA G	
CSTF-50 gene specific primers (for Exon insertion)	Forward primer: TGT TAA TCA TGC GAA GGG CTC ATC AAA G	Screening *CstF-50-1-2* mutants
	Reverse primer: GGACTAAATCTTGCACACCTAACAACG	

Table 2.2 Summary of pair-wise interactions of AtCstF50 with plant polyadenylation factors.

No.	Arabidopsis gene	Polyadenylation factor tested for interaction	Results of Interaction (No/ Yes)
1.	At1g17760	CstF-77	No
2.	At1g71800	CstF-64	No
3.	At5g60940	CstF-50	No
4.	At5g13480	FY	No
5.	At5g51660	CPSF160	No
6.	At5g23880	CPSF100	Yes
7.	At1g61010	CPSF73-I	No
8.	At2g01730	CPSF73-II	No
9.	At1g30460	CPSF30	Yes
10.	AT1g17980	PAPS1	No
11.	At2g25850	PAPS2	No
12.	At3g06560	PAPS3	Yes
13.	At4g32850	PAPS4	Yes
14.	At3g66652	FIPS3	No
15.	At5g58040	FIPS5 NTD and FIPS5 CTD	Yes
16.	At4g29820	CFIS1	No
17.	At4g25550	CFIS2	No
18.	At3g04680	CLPS3	No
19.	At5g39930	CLPS5	No
20.	At1g66500	PCFS1	No
21.	At4g04885	PCFS4	No
22.	At5g43620	PCFS5	No
23.	At5g65260	PABN1	No
24.	At5g51120	PABN2	No
25.	At5g10350	PABN3	Yes

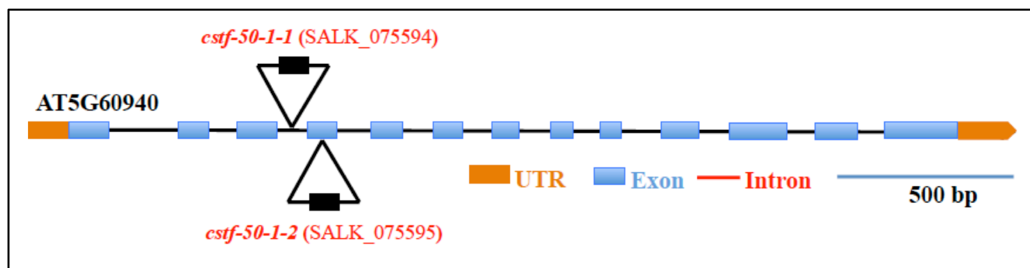


Figure 2.1 At*CstF-50* gene made up of 13 exons and 12 introns. Figure shows the site of T-DNA insertions of SALK_075594 (***cstf-50-1-1)*** and SALK_075595 (***cstf-50-1-2)*** lines.

T DNA

5' and 3' UTR Intron Exon ==== T DNA

Figure.2.2 T-DNA insertion in 3rd intron and 4th exon representing both SALK_075594 (***cstf-50-1-1)*** and SALK_075595 (***cstf-50-1-2)***.

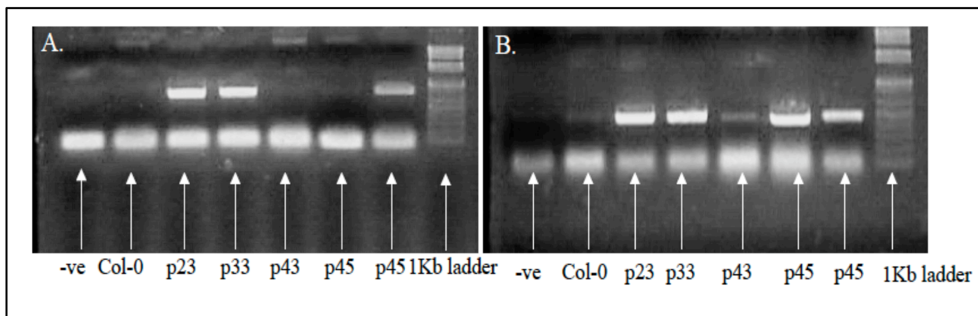


Figure.2.3 PCR confirmation of T-DNA insertion lines. With T-DNA (Panel A) and *NPT2* gene specific (Panel B) primers in SALK_075594 (***cstf-50-1-1***). (-ve: no template; Col-0: Wild type; independent T-DNA insertion plant numbers 23, 33, 43 and 45).

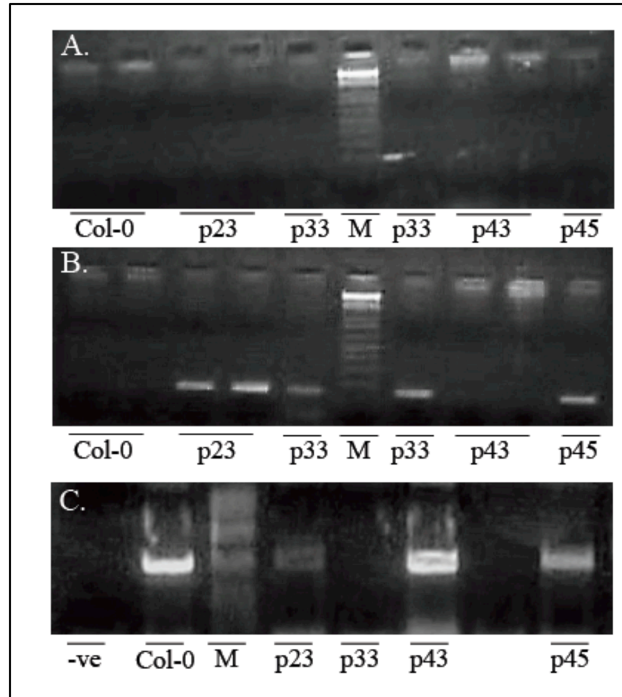


Figure 2.4 Genotyping of SALK_075594 (*cstf-50-1-1*) line.

Panel A: No amplification with T-DNA Left border and gene specific forward primers.

Panel B: Observed amplification in plant numbers 23, 33 and 45 with T-DNA Right border and gene specific forward primers.

Panel C: Observed no amplification in plant numbers 23 and 33 with gene specific forward primers. (–ve: No template; M: 1KB DNA marker).

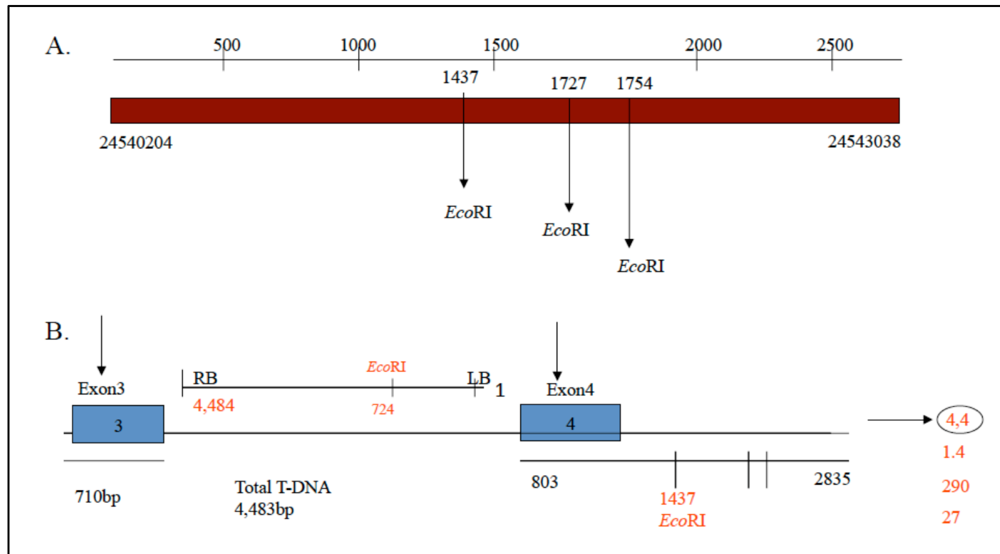


Figure.2.5 Restriction analysis of T-DNA and genomic region of *CSTF50*.

A. Restriction analysis of the genomic region of *CSTF50* shows three *Eco*RI sites.

B. Restriction analysis of the T-DNA shows two *Eco*RI sites and including the three sites in the genomic region release 4 DNA fragments of sizes 4.4 kb, 1.4 kb, 290 base pairs and 27 base pairs.

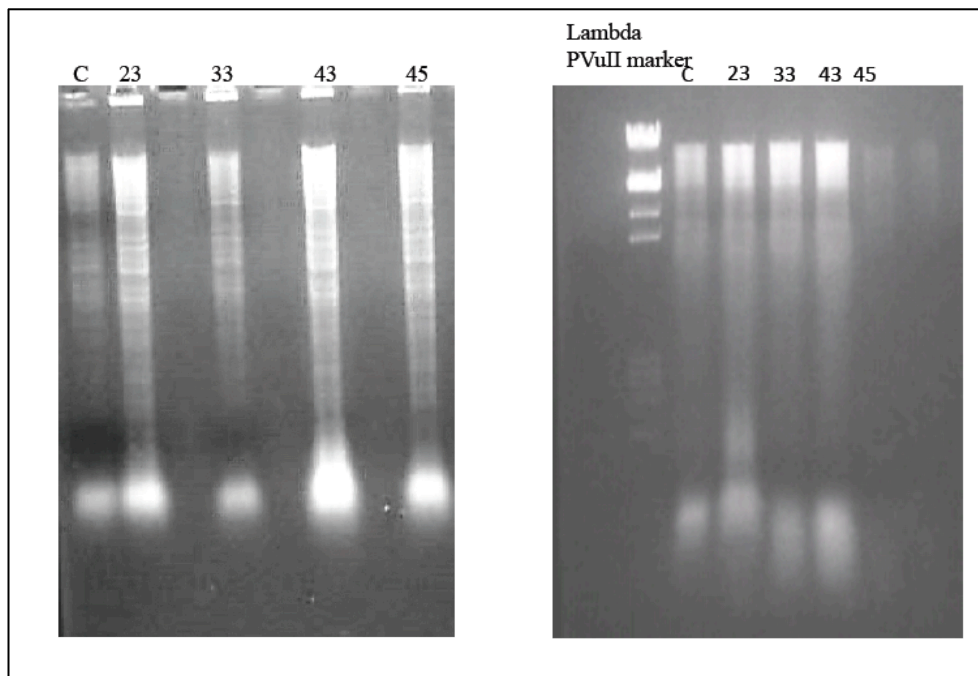


Figure 2.6 *Eco*RI digestion of the genomic DNA Columbia-0 (C) and digested genomic DNA from mutant (SALK_075594) plants (23, 33, 43 and 45). The digested DNA was run on 0.8 percent agarose gel containing Ethidium bromide. Lane designations are printed above the gel image, including C: control DNA (from WT Columbia-0). The gel on the right included a molecular weight marker (Lambda PVuII marker) of known sizes 21 kb, 4.4kb, 4.2kb, 4.1kb, 3.19kb, 3.638kb, 2.296kb, 1.7 kb, 636 bp, 579bp, 532bp, 468 bp, 343bp and 141 bp.

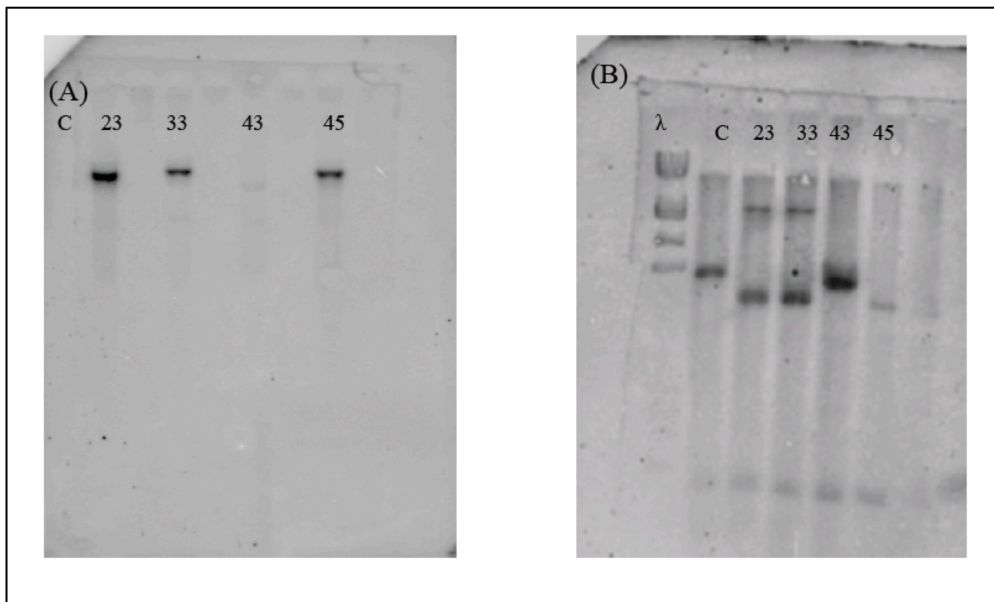


Figure 2.7 Southern blots of WT plants (lane designated as "C") or plants putatively containing the SALK_075594 T-DNA insertion (23, 33, 43, 45). The blot depicted in (A) was hybridized to the T-DNA probe while the blot depicted in (B) was hybridized to the gene specific probe. Lambda refers to the DNA molecular weight marker of known sizes 21 kb, 4.4kb, 4.2kb, 4.1kb, 3.19kb, 3.638kb, 2.296kb, 1.7 kb, 636 bp, 579bp, 532bp, 468 bp, 343bp and 141 bp.

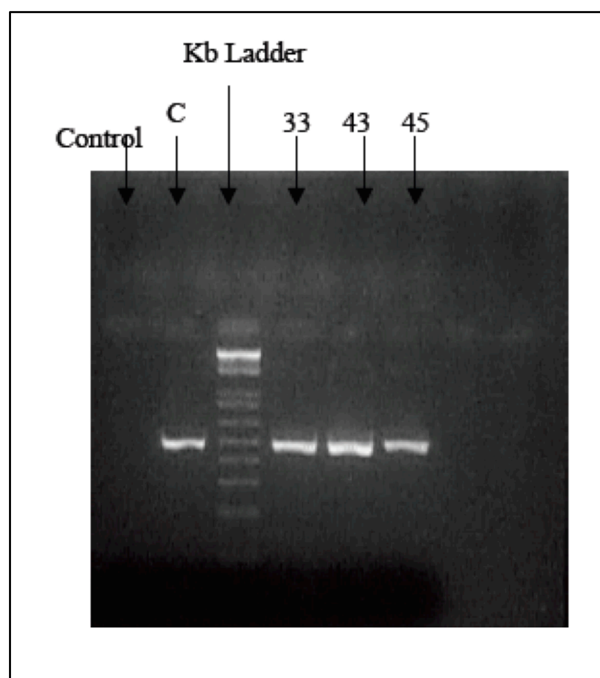


Figure 2.8 Reverse Transcription analysis of intron insertion line. *CSTF50* (SALK_075594) mutant analysis for the presence of gene product in negative control with no template DNA, Columbia-0, plants 33, 43 and 45 using cDNA as a template for the polymerase chain reaction and the expected size is 400 base pairs.

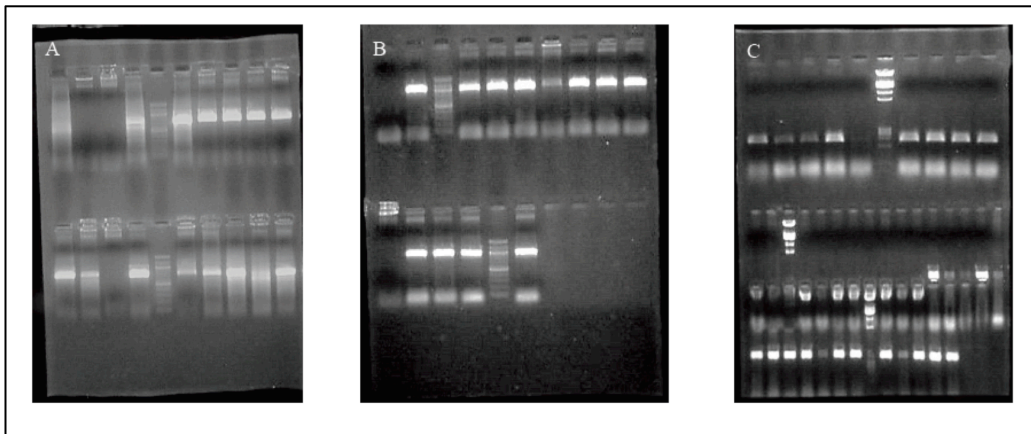


Figure 2.9 PCR analysis of SALK_075595 line.
A. Using T-DNA primers (yields a DNA fragment size of 700 bp,
B. Using gene specific primers (yields 900bp DNA fragment)
C. Using *NPT2* primers (yields 400bp DNA fragment).

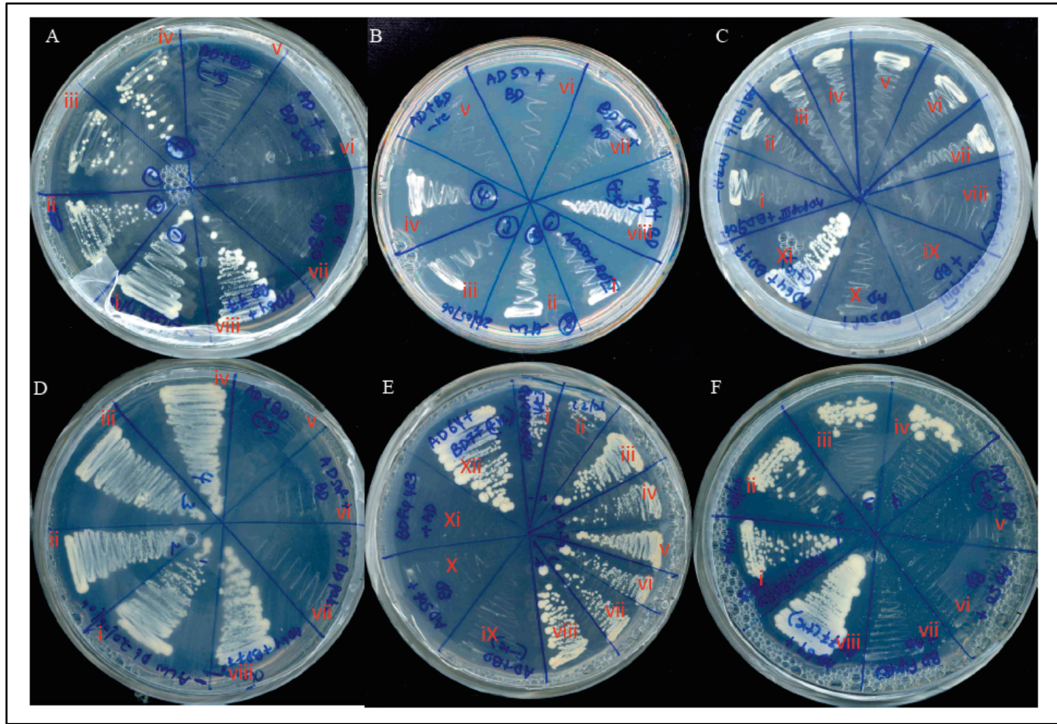


Figure 2.10 Pairwise protein-protein interaction studies using yeast two hybrid system.

Plate-A: Testing interaction of BD-CstF50 with AD-CPSF30: AtCstF50 showed positive interaction with AtCPSF30 (Numbers i, ii, iii, iv are independent colonies; number v, negative control with empty vectors AD and BD; number vi, vii are negative controls with one empty vector AD with BD-CstF50 and BD vector with AD-CPSF30; number viii shows positive control with AD-CstF64 with BD-CstF77).

Figure 2.10 Continued….

Figure 2.10 Continued….

Plate-B: Testing interaction of AD-CstF50 with BD-PAPIV: AtCstF50 showed positive interaction with PAPIV (Numbers i, ii, iii, iv are independent colonies; number v, negative control with empty vectors AD and BD; number vi, vii are negative controls with one empty vector AD-CstF50 with BD and AD vector with BD-PAPIV; number viii shows positive control with AD-CstF64 with BD-CstF77).

Plate-C: Testing interaction of BD-CstF50 with AD-PAPIII: AtCstF50 showed positive interaction with PAPIII (Numbers i, ii, iii, iv,v,vi, vii are independent colonies; number viii, negative control with empty vectors AD and BD; number ix, X are negative controls with one empty vector AD-PAPIII with BD and AD vector with BD-CstF50; number Xi shows positive control with AD-CstF64 with BD-CstF77).

Plate-D: Testing interaction of AD-CstF50 with BD-PABC: AtCstF50 showed positive interaction with PABC (Numbers i, ii, iii, iv are independent colonies; number v, negative control with empty vectors AD and BD; number vi, vii are negative controls with one empty vector AD-CstF50 with BD and AD vector with BD-PABC; number viii shows positive control with AD-CstF64 with BD-CstF77).

Plate-E: Testing interaction of AD-CstF50 with BD-FIPS5NTD: AtCstF50 showed positive interaction with FIPS5NTD (Numbers i, ii, iii, iv,v,vi,vii,viii are independent colonies; number ix, negative control with empty vectors AD and BD; number X, Xi are negative controls with one empty vector AD-CstF50 with BD and AD vector with BD-FIPS5NTD; number Xii shows positive control with AD-CstF64 with BD-CstF77).

Plate-F: Testing interaction of AD-CstF50 with BD-FIPS5CTD: AtCstF50 showed positive interaction with FIPS5CTD (Numbers i, ii, iii, iv are independent colonies; number v, negative control with empty vectors AD and BD; number vi, vii are negative controls with AD-CstF50 with empty vector BD and AD vector with BD-FIPS5CTD; number viii shows positive control with AD-CstF64 with BD-CstF77).

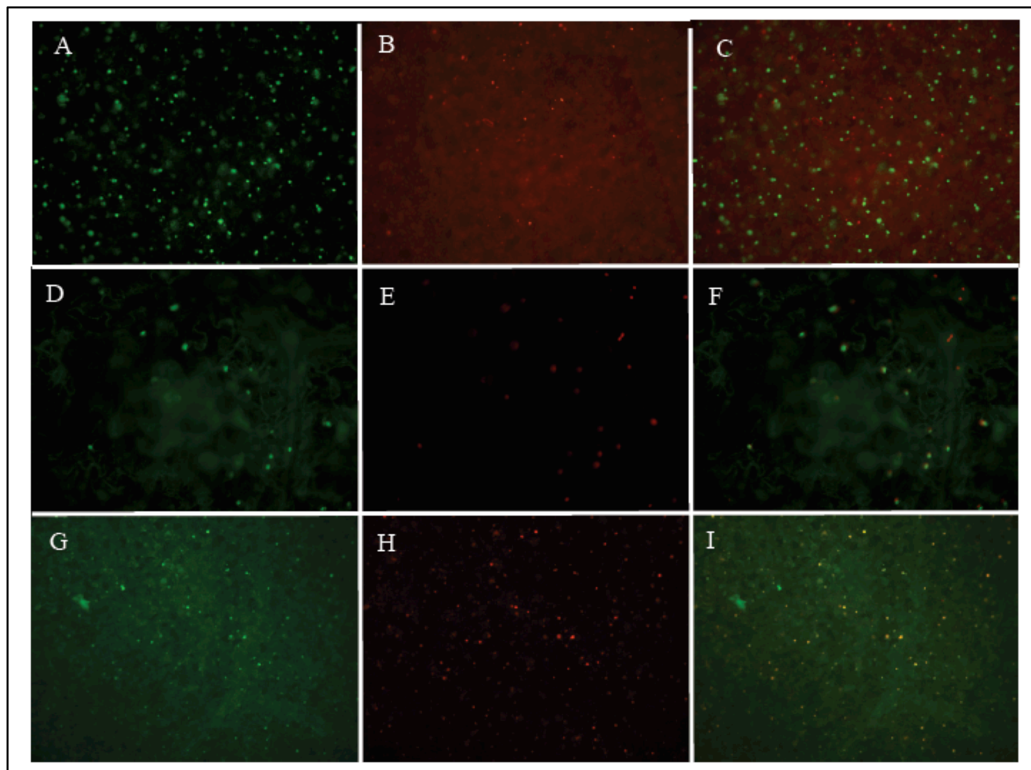


Figure 2.11 Localization studies of AtCstF50.

Panel1: Co-infiltration of RFP-AtCstF50 using GFP nuclear marker
A. Nuclear GFP in green filter
B. AtCstF50 fused to RFP in red filter
C. Merged picture of RFP-AtCstF50 and nuclear GFP

Panel2: Co-infiltration of RFP-AtCstF50 using GFP ER marker
D. ER-GFP in green filter
E. AtCstF50 fused to RFP in red filter
F. Merged picture of RFP-AtCstF50 and ER-GFP.

Panel3: Co-infiltration of GFP and RFP nuclear markers
G. GFP nuclear marker in green filter
H. RFP nuclear marker in red filter
I. Merged picture of GFP and RFP nuclear markers.

Copyright © Lavanya Dampanaboina 2011

CHAPTER THREE

Characterization of FY, an orthologue of yeast Pfs2p

Introduction

In eukaryotes, the 3′ ends of RNA polymerase II-generated transcripts are cleaved and polyadenylated and this is an essential step for translation, stability, transcription termination, transport of mRNA's to the cytosol and functions associated with regulation through APA and 3'UTR's (Hammell, Gross et al. 2002; Buratowski 2005) (Holec, Lange et al. 2006). Genetic and biochemical approaches in *Saccharomyces cerevisiae* have defined a large number of conserved proteins required for RNA 3′-end processing, including the polyadenylation factor Pfs2p (Zhao, Hyman et al. 1999). Pfs2p contains seven WD repeats and acts as an interaction surface within the cleavage and polyadenylation factor (CPF) 3′-end processing complex (Ohnacker, Barabino et al. 2000). CPF acts with the cleavage factor I (CFI) complex to direct 3′-end processing of pre-mRNA transcripts (Zhao, Hyman et al. 1999).

The Arabidopsis homologue of Pfs2p is FY, a protein that plays a role in RNA 3′-end processing (Simpson, Dijkwel et al. 2003) (Feng, Jacob et al. 2011) and is involved in the regulation of expression of the *FCA* and *FLC* genes (Feng, Jacob et al. 2011) (Quesada, Macknight et al. 2003). FY physically interacts with FCA, negatively auto-regulates FCA protein levels and promotes flowering by targeting FLC, a central repressor of flowering (Simpson, Dijkwel et al. 2003). FY also takes part in the polyadenylation site choice of FLC apart from FCA and acts as a 3' end processing factor (Feng, Jacob et al. 2011). Analysis of *fy* mutants in *Arabidopsis thaliana*, and of *Nicotiana benthamiana* plants silenced for FY expression showed that FY is required for growth and development in plants (Henderson, Liu et al. 2005) and that both the conserved FY WD repeats and the C terminus are required for repression of *FLC* gene expression.

Calmodulin is a highly conserved calcium binding protein in eukaryotes. It is one of the three classes of calcium sensor molecules (Kim, Chung et al. 2009). More than 50 calmodulin binding proteins have been identified in plants (Snedden and Fromm 1998)

(Yang and Poovaiah 2003; Bouché, Yellin et al. 2005). In plants, targets of calmodulin were identified with distinctive cellular functions such as regulation of metabolism, signal transduction pathway, transcriptional regulation, protein folding, post translational modifications, functioning of cytoskeleton and ion transport (Reddy, Ali et al. 2002; Kim, Chung et al. 2009). Calcium and calmodulin complexes might mediate plant responses directly by changing the activity of target proteins or indirectly by changing the expression of genes involved in the production of downstream effectors (Kim, Chung et al. 2009). It has been reported that the Arabidopsis ortholog of CPSF30 (AtCPSF30) has a calmodulin binding domain and that AtCPSF30 binding to calmodulin inhibits RNA binding by AtCPSF30 in a calcium dependent manner. This example illustrates a possible link between signaling pathway of calmodulin and 3'end processing of pre-mRNAs in plants (Delaney, Xu et al. 2006).

A novel domain has been identified in FY at the C-terminal end that binds to calmodulin (S. Rao and A. G. Hunt, unpublished data). We hypothesize that this calmodulin binding domain (CBD) might be playing a role in the functioning of FY. To test this hypothesis, different domains of FY were fused to RFP and the subcellular locations of the fusion proteins was examined. The results show that nuclear localization of FY is a complicated process, and that the calmodulin-binding domain of FY may play a role in nuclear localization. The regulatory implications of these results are discussed.

MATERIALS AND METHODS

Vector construction

For transient studies, in-frame translational fusion constructs were made using pCAMBIA-1301-derived (CAMBIA, George Street, Brisbane, QLD, Australia) pGDG (with Green Fluorescent Protein (GFP) marker) or pGDR (with Red Fluorescent Protein (DsRED) vector (Goodin, Dietzgen et al. 2002). Constructs encoding the full length FY (1-647 amino acids), the N-terminal domain of FY (1-471 amino acids), the C-terminal domain of FY (465-647 amino acids), the full length FY with the calmodulin binding domain deleted (cFY) binding domain (amino acids 508-520), and the C-terminal domain with calmodulin binding domain deleted, were made by cloning the region encoding for each of the above proteins in-frame with the end of the coding sequence for the RFP reporter in the pGDR vector (Fig.3.1 and Fig.3.2). Full length *FY* was PCR amplified with gene specific primers using cDNA from wild type Arabidopsis plants. The PCR product was digested with *Bam*HI and *Xba*I and cloned into the pGDR vector using the same sites. The sequences encoding the N- and C-terminal domains (NTD and CTD, respectively) of FY were amplified using gene specific NTD primers, CTD primers, using the full length *FY* clone as a template (Table 1). The PCR products were restriction digested with *Bam*HI and *Xba*I, gel purified and ligated into the *Bam*HI and *Xba*I sites of the pGDR vector.

Deletion of the calmodulin binding domain (CBD) from the full length *FY* coding sequence (CDS) was accomplished using primers designed to amplify two fragments flanking the CBD with a compatible restriction site between them. The two fragments were amplified in two separate PCR reactions using four primers (*cFY1* forward primer and *cFY1* reverse primer for fragment 1 amplification and *cFY2* forward and *cFY2* reverse primer for fragment 2, Table 1) and the resultant fragments were independently digested with *Sal*I and *Xho*I and the purified fragments were ligated at 16^0C with T4 DNA ligase (New England Biolabs, Beverly, MA, USA). The ligated product was amplified using the full length *FY* primers resulting in the *FY* with CBD deleted (cFY). The PCR product was digested with *Bam*HI and *Xba*I and cloned at the same sites in the pGDR vector. The cloned cFY was used as a template for amplifying the *cFY* C-terminal

domain. *FY CTD* specific primers (Table 1) were used to amplify the *cFY* CTD, the PCR products were digested with *Bam*HI and *Xba*I, gel purified and cloned at *Bam*HI and *Xba*I sites of pGDR vector. The nuclear localized markers for positive controls were made by cloning the nucleotides encoding the seven amino acid long (PKKKRKV) nuclear localization signal of SV-40 (simian virus 40) T antigen into pGDG and pGDR vectors (Kalderon, Roberts et al. 1984) (Fig. 3.1). This sequence was shown to localize GFP and RFP to the nucleus (Fig. 3.3). Oligonucleotides containing the SV40 NLS and *Bam*HI over hangs were annealed and ligated with *Bgl*II-digested pGDG and pGDR vectors. The expression of individual genes and florescent tags for all the constructs were confirmed by extracting RNA from the infiltrated leaves, performing reverse transcription, and PCR amplifying and then sequencing the resulting amplicons.

Agro-infiltration experiments for transient expression studies

All the constructs along with control vectors were transformed into *Agrobacterium tumefaciens* strain LBA4404 using a freeze thaw method (An, Ebert et al. 1988). Agrobacterium was freshly grown on LB agar plates with appropriate antibiotics for 2 days at 28^{0}C. The cultures plates were scraped with a spatula and suspended in the freshly made MES buffer (10mM MgCl2 and 10mM MES). The suspensions are measured for the cell density in a spectrophotometer at 600nm. Using the optical density (OD) readings, the cell density was set to 0.8 OD by dilution with fresh 10mM $MgCl_2$ and 10mM MES buffer. Acetosyringone was added to all samples at 150μM concentrations and incubated at room temperature for 2-3 hours. For co-infiltrations, equal volumes of different Agrobacterium samples were mixed before infiltrations. An NLS-GFP reporter was used as an internal control in all co-infiltrations.

Three weeks-old *Nicotiana benthamiana* plants were selected and leaves are infiltrated with the Agrobacterium suspensions using a 1ml syringe without a needle (Goodin, Dietzgen et al. 2002). Healthy looking and fully expanded lower leaves were selected for the infiltration experiments. Gentle pressure was applied with a finger on the adaxial surface of the leaf while injecting the suspension buffer from the abaxial surface of the leaf. The desired amount of the buffer with the bacterial suspension was infiltrated in the

intravenous regions. More than two infiltration sites were made with the tip of the syringe to complete the infiltration in a single leaf. At least two leaves were infiltrated with the same suspension on a single plant. Once the leaves were infiltrated a water soaked appearance is observed. The plants were incubated for 48 to 72 hours under the fluorescent light with 16hr/8hr light and dark photoperiods.

Localization studies using microscopy

Infiltrated leaves (after 48-72hr of incubation) were made into small sections near the infiltration site. The section with the abaxial side up was placed on a glass slide in a drop of water and mounted gently with a cover slip. The slide was observed under UV light using an epifluorescence microscope (Zeiss Axioplan2 HB100). The GFP reporter was visualized using the green filter (ex: D470/40; D535/40; beam splitter 500 DCLP) and RFP was visualized under the red filter (ex: HQ545/30X; em: HQ610/75M, Q570LP). Images were captured using the Zeiss AxioCam MRc5 attached to the microscope. Axio Vision software was used to adjust the exposure for each filter and also to get the merged picture for the co-infiltration samples along with the pictures under green or red filters.

RNA extraction, cDNA synthesis and PCR amplification

The infiltrated leaves that were observed under the microscope and other infiltrated leaf samples were frozen in liquid nitrogen and stored at -80^{0}C in a freezer. The frozen tissue was ground in liquid nitrogen and RNA was isolated using Trizol (Invitrogen). To avoid DNA contamination in the samples, the isolated RNA was treated with DNaseI (Fermentas Inc., Glen Burnie, Maryland, USA) at 37°C for 1hr. After DNase treatment the samples were purified using a spun column (Qiagen RNeasy plant mini kit). Samples were reverse transcribed into cDNA using an oligo-dT primer. Reporter- and gene-specific primers were used to PCR amplify the reporter and coding sequences using cDNA template. For the co-infiltrated samples, the internal control, gene specific primers were also used to test the expression of the transgene through RT-polymerase chain reaction.

RESULTS

Full length FY localizes to the nucleus

Transient expression studies were performed along with controls to determine the localization of FY in the plant cells. These studies revealed that the RFP-FY protein was apparent as a distinct spot in leaf epidermal cells. To determine the site of localization, the RFP-FY construct was co-infiltrated with a positive control encoding *GFP* fused to the SV40 NLS. The results showed that the majority of the RFP-FY co-localized with GFP-NLS indicating the nuclear localization of the RFP-FY (Fig.3.4).

Fusion proteins containing the FY-NTD domain and FY-CTD domains localize to the nucleus

The N-terminal domain of FY has a WD-repeat motif (Tryptophan, Aspartate) necessary for its interaction with E3 ubiquitin ligases, while it lacks two PPLP motifs that are necessary for its interaction with FCA (an RNA binding protein) and Calmodulin Binding Domain (CBD). The C-terminal domain of FY has several PPLP motifs and the calmodulin binding domain, but no WD motif. To test how these domains affect the localization pattern of FY, they were translationally fused to the coding sequence of the GFP marker and, upon transient introduction into tobacco cells, the subcellular distributions of these protein chimeras was determined.

For the *FY*-NTD construct, the nucleotides encoding the *FY* N-terminal domain (1-471 amino acids; FY-NTD) were fused at the C-terminus of cDNA encoding RFP. Transient expression studies in *Nicotiana benthamiana* leaves revealed that the FY-NTD-RFP fusion protein co-localized with GFP-NLS marker, indicating a nuclear localization of FY (Fig.3.5).

The FY-CTD construct (465-647 amino acids; FY-CTD) was fused to the RFP coding region at its 3' end in the pGDR vector. Transient expression studies in *Nicotiana benthamiana* leaves revealed that the FY-CTD-RFP fusion protein co-localized with the GFP-NLS marker, indicating a nuclear localization for the CTD of FY (Fig.3.6).

Consistently, the levels of the FY-NTD fusion protein were low in these studies, whereas the levels of the FY-CTD fusion protein were high. The presence of the WD motifs in the N-terminus of FY might be responsible for lower expression as it also interacts with E3 ubiquitin ligases.

Mutation in the calmodulin binding domain alters the localization of FY

A novel calmodulin binding domain (CBD) was identified and mapped to amino acids 508 to 520 at the C-terminal end of FY (S. Rao, unpublished observations). To characterize this domain, the CBD was deleted and the resulting variant (FY-CBD/cFY) was fused to the C-terminal end of RFP. Transient expression showed that this protein co-localizes with the GFP-NLS marker. However, the cFY-RFP protein was also distributed throughout the cytoplasm (Fig. 3.7).

Since the full length FY localized exclusively to the nucleus (Fig.3.4) while full length cFY also localized to the cytoplasm (Fig.3.7), it seemed possible that the calmodulin-binding domain might overlap a nuclear-localization signal. If this were the case, then deletion of the CBD from the C-terminal domain should cause the resulting protein (cFY-CTD) to be localized in the cytoplasm, in contrast to what is seen with the FY-CTD constructs. To test this, the FY-CTD with mutation in CBD was fused to the C-terminal end of RFP. Transient expression studies showed that the RFP-cFY-CTD co-localizes with the NLS-GFP marker (Fig.3.8). Thus, the calmodulin-binding domain does not appear to overlap with the nuclear localization signal that is responsible for the nuclear accumulation of the FY-CTD fusion protein.

DISCUSSION

Transient expression studies showed that full length FY, the FY N-terminal domain and the FY C-terminal domain all localize to nucleus. Full length FY might be binding to other proteins and then transported into the nucleus, as there is no obvious nuclear localization signal in FY. FY interacts with CPSF subunits 100 and 160 (Hunt, Xu et al. 2008) but smaller FY protein moieties can not form complexes with CPSF (Manzano, Marquardt et al. 2009). So, these interactions may be responsible for the localization of the full length FY to the nucleus but cannot explain the import of fragments of FY. However, the FY-NTD and FY-CTD could also localize to the nucleus; while the interaction with the CPSF must involve both domains (Fig.3.9), it is also possible that FY possesses multiple non-canonical nuclear localization signals that are responsible for the nuclear localization of FY.

Visual comparison suggests that full length FY and the FY-NTD domains have low levels of expression while the FY-CTD accumulated to much higher levels. While this needs to be confirmed, this observation is interesting in light of other reported properties of FY. FY was found to interact with DDB1, an E3 ubiquitin ligase (Lee, Terzaghi et al. 2008). The site of interaction with DDB1 lies in the WD repeat-containing, N-terminal end, suggesting that this domain promotes polyubiquitination and subsequent degradation by the 26S proteasome (Fig.3.10). Therefore, the possible differences in expression may be due to the presence of the WD motif of FY that interacts with an E3 ubiquitin ligase targeting FY or FY-NTD for polyubiquitination and then, to the degradation machinery. The absence of the WD repeat motif in the FY-CTD would be responsible for the greater stability of this portion of FY resulting in apparent higher expression.

Calmodulin (CaM) is one of the calcium sensing proteins, and calcium is a secondary messenger involved in the regulation of many biotic and abiotic stresses in plants affecting their development (Hashimoto and Kudla 2011). Proteins with helix-loop-helix EF hand motifs form a major group of calcium sensing proteins that are responsible for calcium binding and aid in protein conformational changes upon calcium binding. Arabidopsis has 250 putatitive EF hand proteins and 100 known calcium sensor proteins

while *Drosophila melanogaster* and humans have 132 and 83 EF hand proteins respectively (Day, Reddy et al. 2002). In Arabidopsis there are four isoforms of CaM that are encoded by seven genes. There are 50 isoforms of CaM-like proteins (CMLs) in Arabidopsis that show structural differences though they are thought to be evolved from ancestral CaMs (McCormack, Tsai et al. 2005). Possible targets for CaMs and CMLs include transcription factors, metabolic enzymes, kinases, phosphatases, ion transporters and a wide range of other proteins (Reddy, Ali et al. 2011).

The presence of the calmodulin binding domain in the C-terminus of FY, between the PPLP motifs that are necessary for the FY interaction with FCA suggest that when calmodulin binds, FY may not be capable of binding FCA. A similar mechanism has been proposed by Delaney et al in 2006 for AtCPSF30 where binding of calmodulin inhibited the binding of RNA (Delaney, Xu et al. 2006). Assuming the above model also functions in the flowering pathway of Arabidopsis it is thought that in the absence of calmodulin binding, FCA might be easily accessible to FY for the repression of FLC. If calmodulin blocks FCA interaction with FY, then more FCA active protein will be synthesized and hasten FLC repression promoting flowering. The deletion of calmodulin binding would therefore, promote FY binding to FCA, result in less FCA full length protein being produced, alleviate FCA repression of FLC, causing late flowering. So, both FCA levels and FY levels are important for promoting flowering. Calmodulin might be affecting one pathway and the absence of calmodulin might result in early flowering as depicted in Fig.3.11.

Reports in tobacco show that flowering is also regulated by calmodulin dependent protein kinase (MCK1). Lack of calmodulin binding domain in the modified MCK1 aborts flower primordia suggesting key role of calmodulin in flowering (Liang, Wang et al. 2001). Other reports in Arabidopsis shows that calmodulin regulates flowering pathway mediated by photoperiod pathway by regulating the gene expression of clock genes (Murphy, Kemp et al. 2009). With FY being a 3'-end processing factor, functions in autonomous pathway of flowering, with a novel CBD suggests a calmodulin mediated signaling pathway in flowering that has to be unraveled in the future.

Table 3.1 Primers used for the localization experiments.

Full length FY primers	FY forward primer: TTT TTT GGA TCC ATG TAC GCC GGC GGC GAT ATG CAC;	Amplification of full length FY
	FY reverse primer: CCC CCC TCT AGA CTA CTG ATG TTG CTG ATT GTT GTT	
FY NTD primers	FY NTD forward primer: TTT TTT GGA TCC ATG TAC GCC GGC GGC GAT ATG CAC	For the amplification of FY-NTD domain
	FY NTD reverse primer: CCC CCC TCT AGA GAG AGG ATG CAT CAA ATG GCA TTG	
FY CTD primers	FY CTD forward primer: TTT TTT GGA TCC ATG CCA TTT GAT GCA TCC TCT CAA GGG	For the amplification of FY-CTD domain
	FY CTD reverse primer: CCC CCC TCT AGA TAC TGA TGT TGC TGA TTG TTG TTT GGT	
cFY (full length FY with deleted calmodulin binding domain)	cFY1 forward primer: TTT TTT GGA TCC ATG TAC GCC GGC GGC GAT ATG CAC	Amplification of Full length cFY by deletion of calmodulin binding domain.
	cFY1 reverse primer: CCC CCC GTC GAC ATA CCC TTG CTG CTG GCC ACT TCC	
	cFY2 forward primer: TTT TTT CTC GAG CTT CCA ATG CCC AAT ATG CCT CAC	
	cFY2 reverse primer: CCC CCC TCT AGA CTA CTG ATG TTG CTG ATT GTT GTT	
SV40 Nuclear localizatio n signal	SV40 NLS forward primer: GATCG CCA AAA AAG AAG AGA AAG GTA GCC TAA G	For the nuclear localization signal
	SV40 NLS reverse primer: GATCC TTA GGC TAC CTT TCT CTT CTT TTT TGG C	

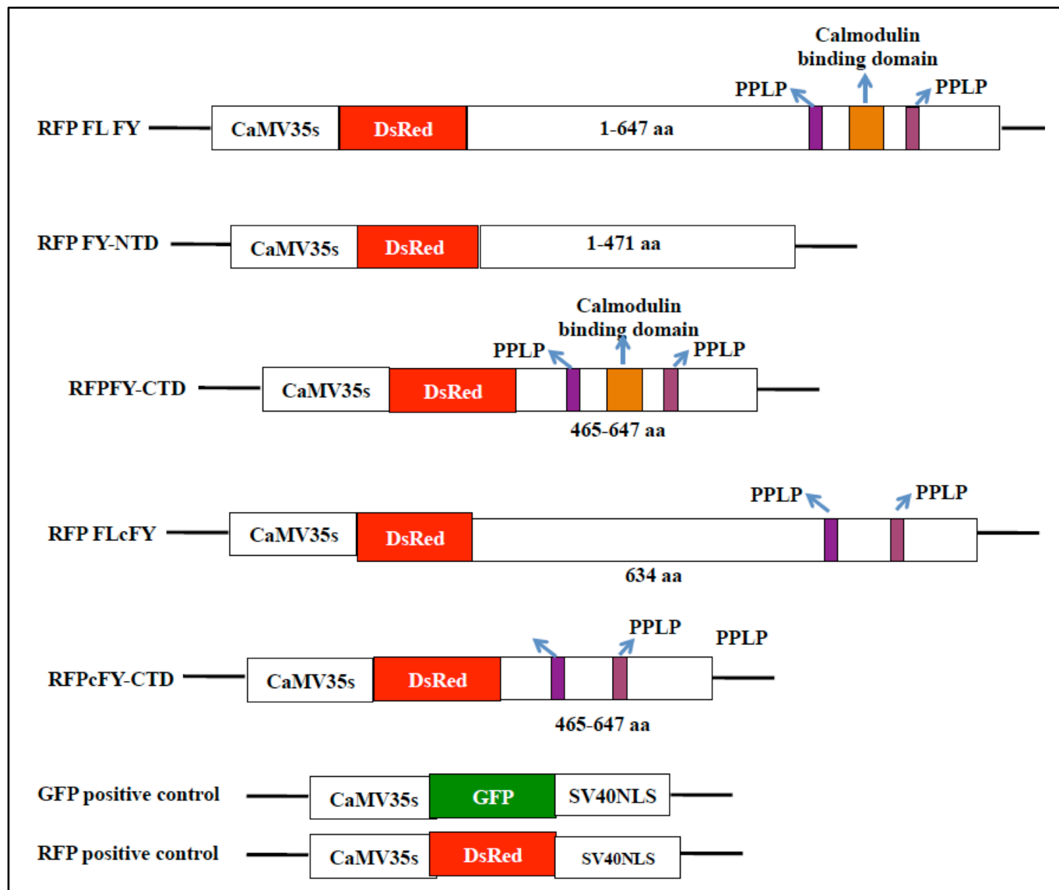


Figure 3.1 Fusion constructs of FY with RFP for transient expression studies

1 MYAGGDMHRG SQMPQPPMMR QSSASSTNIN PDYHHPSGPF DPNVDSFGAK
51 RMRKHTQRRA VDYTSTVVRY IQARTWQRDS RDRTTLQPTP AAAVDMLPTV
101 AYSDNPSTSF AAKFVHASLN KNRCSINRVL WTPSGRRLIT GSQSGEFTLW
151 NGQSFNFEMI LQAHDQPIRS MVWSHNENYM VSGDDGGTLK YWQNNMNNVK
201 ANKTAHKESI RDLSFCKTDL KFCSCSDDTT VKVWDFTKCV DESSLTGHGW
251 DVKSVDWHPT KSLLVSGGKD QLVKLWDTRS GRELCSLHGH KNIVLSVKWN
301 QNGNWLLTAS KDQIIKLYDI RTMKELQSFR GHTKDVTSLA WHPCHEEYFV
351 SGSSDGSICH WIVGHENPQI EIPNAHDNSV WDLAWHPIGY LLCSGSNDHT
401 TKFWCRNRPA DNPRDVLMQN QGYNEQGFGR QPDNFQPSEA SPIPGAFVPG
451 LTRNEGTIPG IGIAMPFDAS SQGDHKQPLP GSMALGAPPL PPGPHPSLLG
501 SGQQQGY QQQ QQHQGHPQQM LPMPNMPHHQ LPPSSHMPLH PHHLPRPMQM
551 PPHGHMPPPS MPMSHQMPGS MGMQGGMNPQ MSQSHFMGAP SGVFQGQPNS
601 GGPQMYPQGR GGFNRPQMIP GYNNPFQQQQ Q PPLPPGPPP NNNQQHQ

- Full length FY – blue + pink (647aa)
- N-terminal domain- blue color (1-471aa)
- C-terminal domain – pink color (465-647aa)
- Full length FY with the deletion of Calmodulin binding domain (underlined pink font region deleted)

Figure 3.2 Different domains of FY used for localization studies.

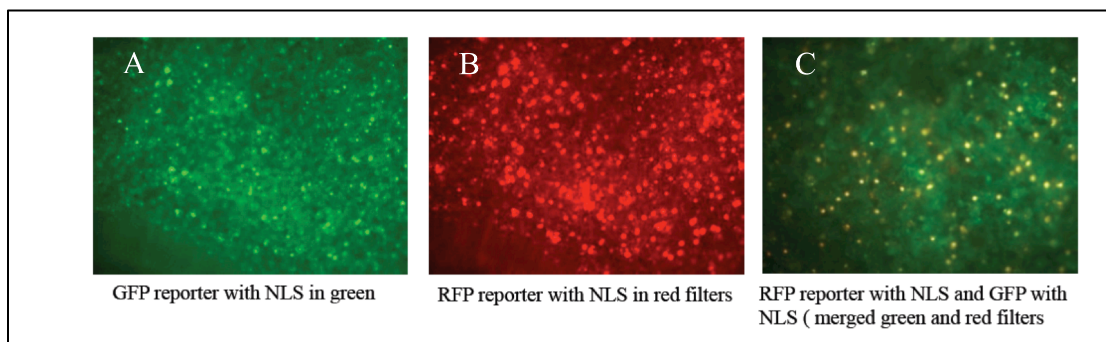


Figure 3.3 Localization of GFP and RFP nuclear markers
Co-infiltration sample of GFP reporter with nuclear localization signal and RFP reporter with nuclear localization signal.

A. Under green filter
B. Under red filter and
C. Merged picture

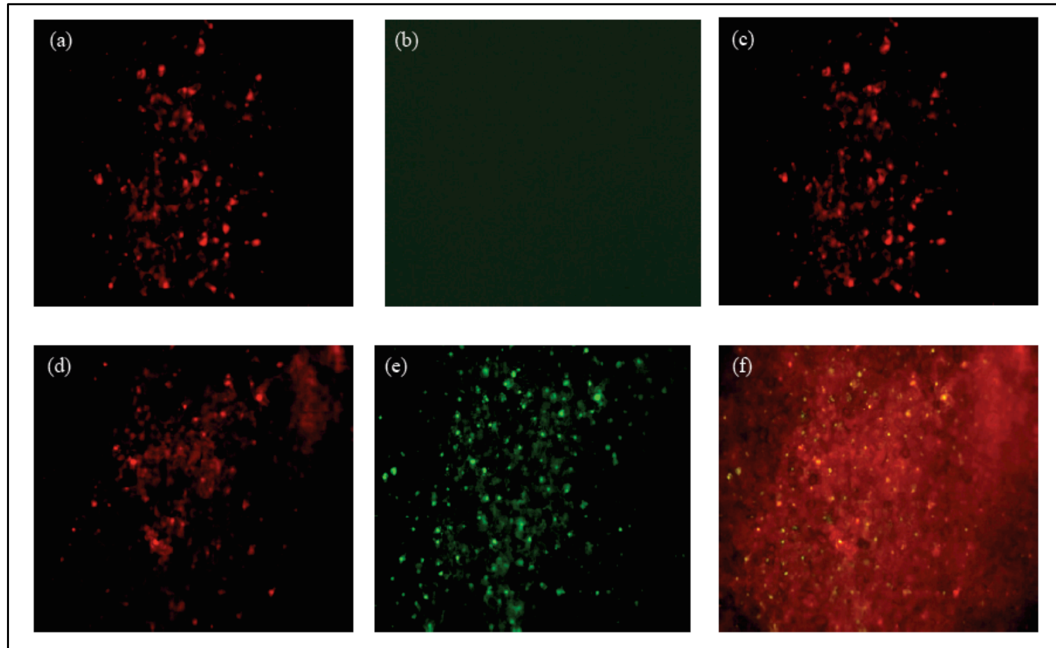


Figure 3.4 Localization of FY fused to RFP.

Upper panel: Infiltration of FY fused to RFP alone in
(a) red filter (b) green filter and (c) merged picture.

Lower Panel: Co-infiltration of FY fused to RFP with GFP nuclear marker
(d) green filter (e) red filter and (f) merged picture

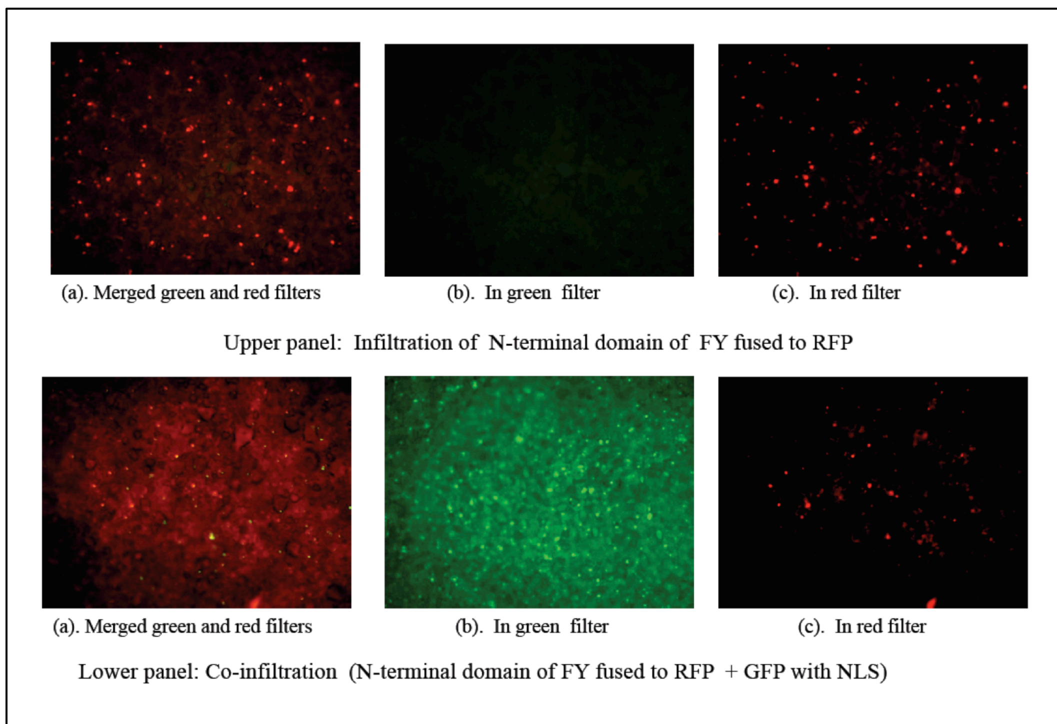


Figure 3.5 Localization of FY-NTD domain fused to RFP. Infiltration of FY-NTD fused to RFP alone and co-infiltration of FYNTD fused to RFP along with GFP nuclear marker

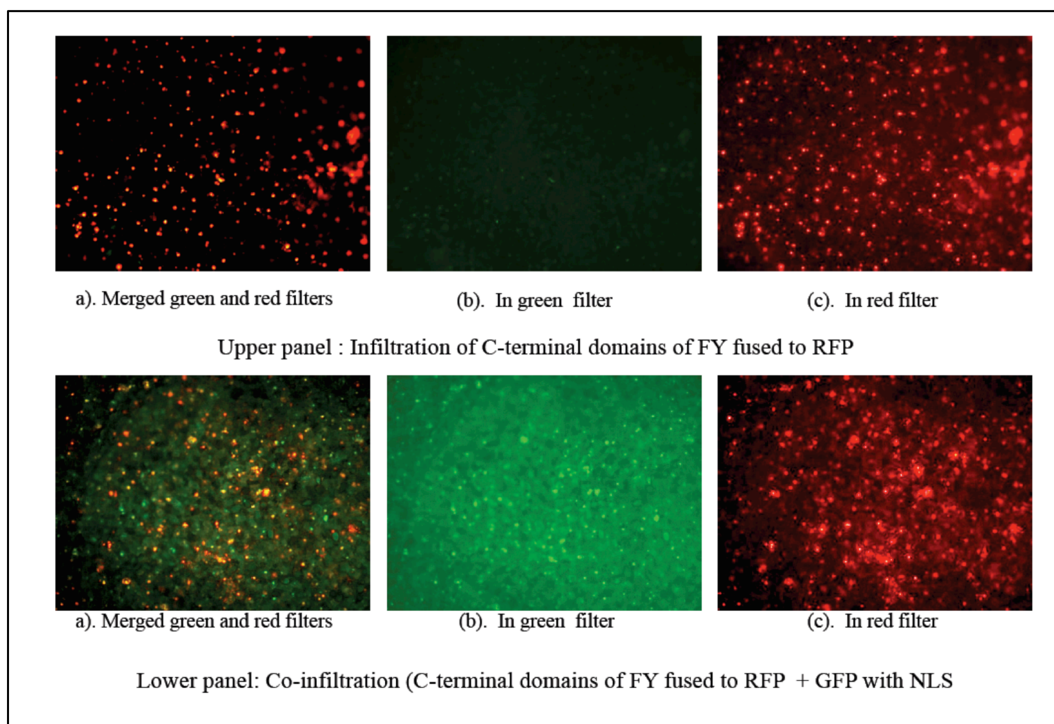


Figure 3.6 Localization of FY-CTD domain fused to RFP. Infiltration of FY-CTD fused to RFP alone and co-infiltration of FY-CTD fused to RFP along with GFP nuclear marker

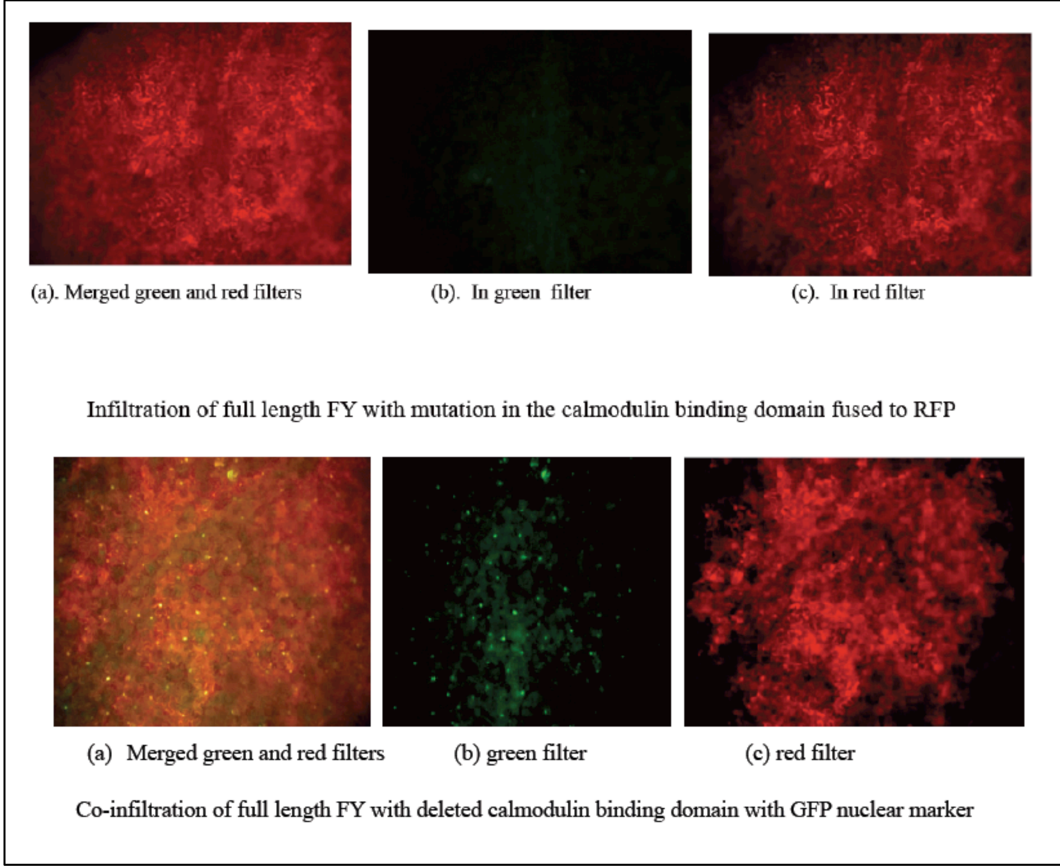


Figure 3.7 Localization of cFY fused to RFP. Infiltration of cFY fused to RFP alone and co-infiltration of cFY fused to RFP along with GFP nuclear marker

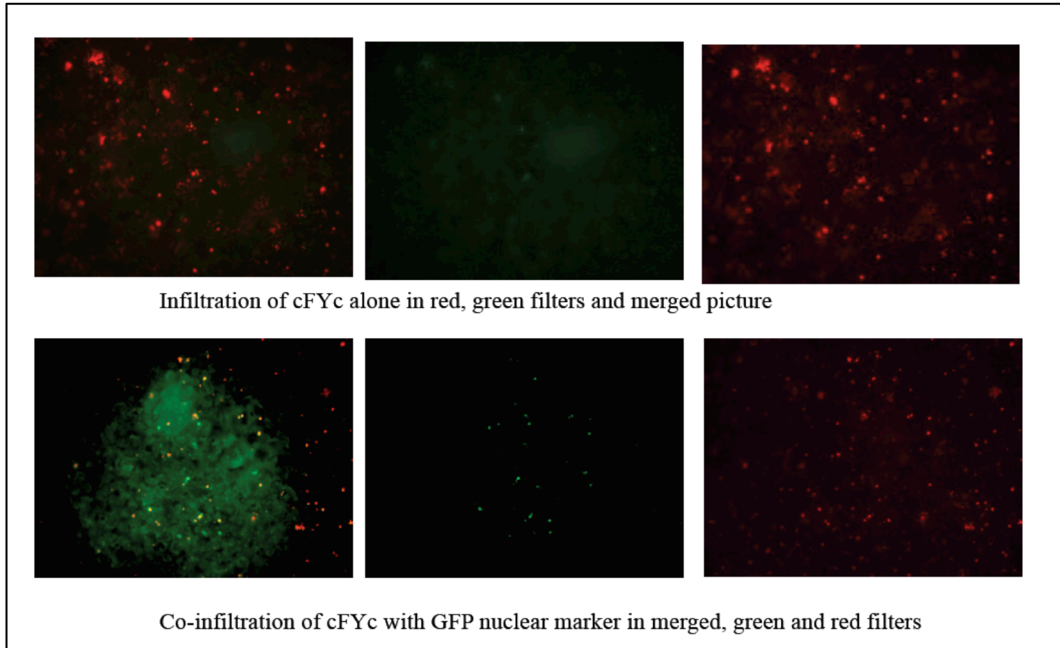


Figure 3.8 Localization of cFY-CTD fused to RFP.
Upper panel is the Infiltration of cFY-CTD fused to RFP alone (a) merged red and green filters (b) green filter and (c) red filter and lower panel shows co-infiltration of cFY fused to RFP along with GFP nuclear marker visualized in (d) merged red and green filters (e) green filter and (f) red filter

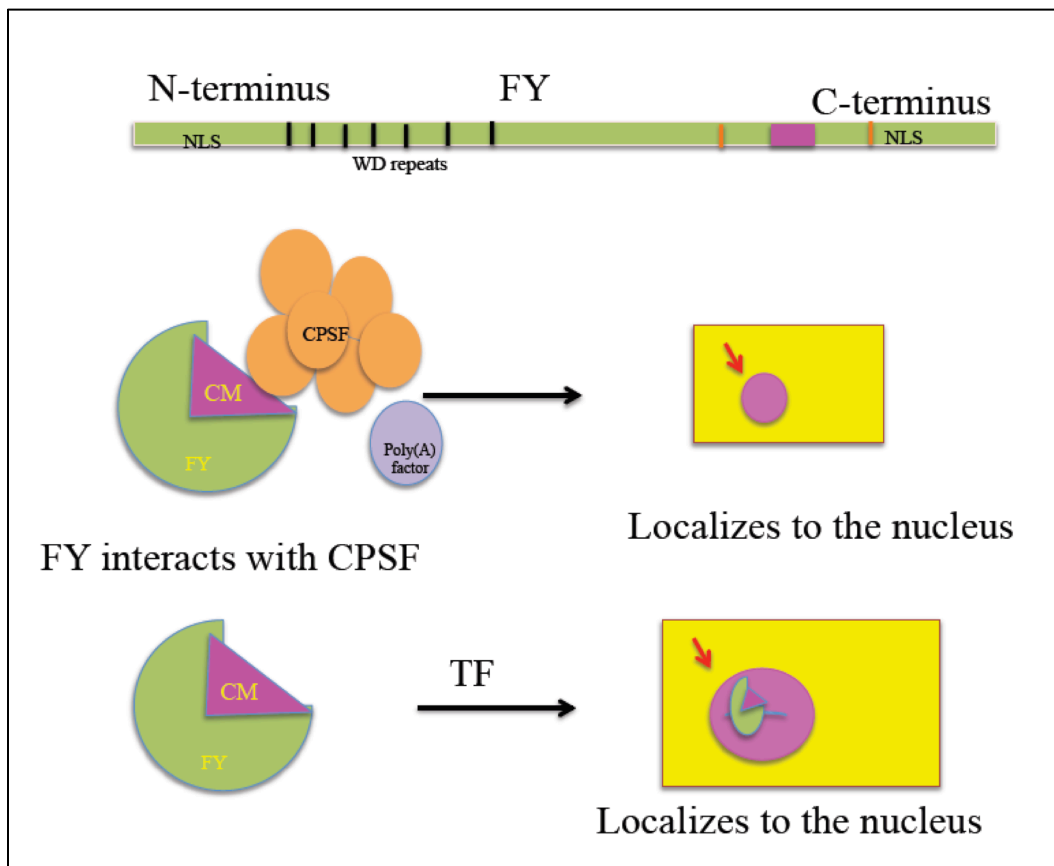


Figure 3.9 A possible model for FY localization to the nucleus. NLS: nuclear localization signal, CPSF: Cleavage and Polyadenylation Specificity Factor, TF: transcription factor.

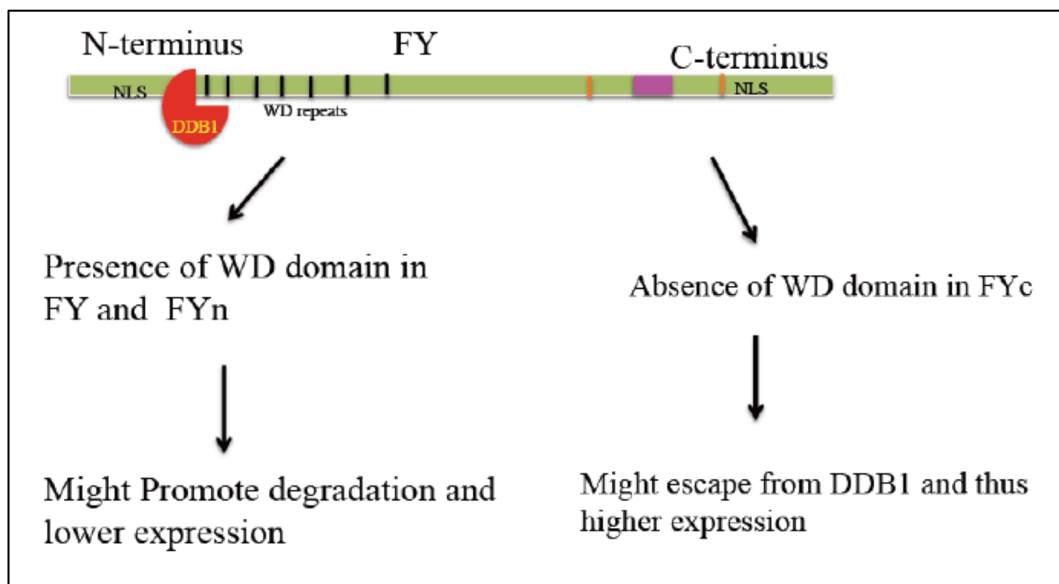


Figure 3.10 Domains of FY responsible for interaction with DDB1 Ubiquitin Ligase.

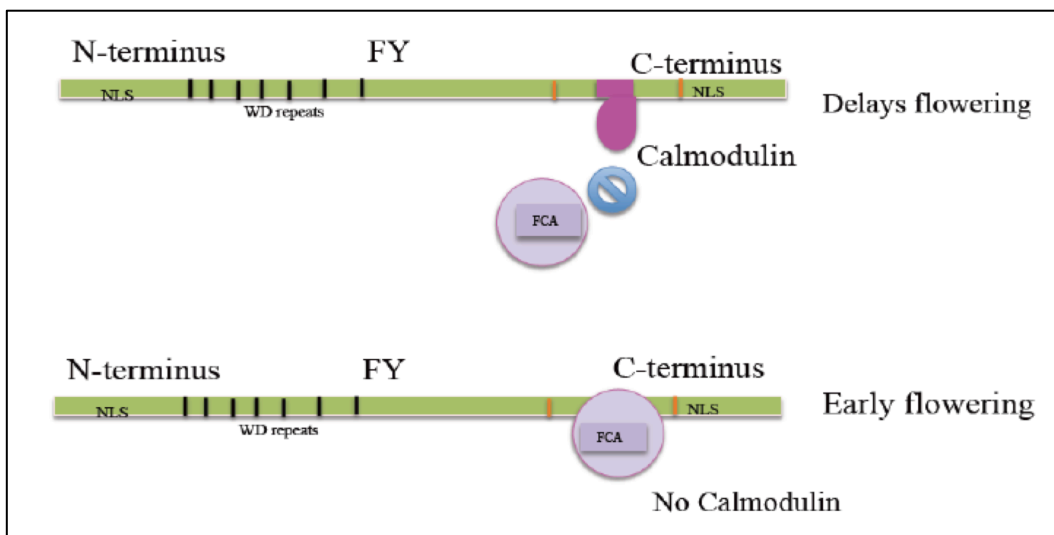


Figure 3.11 Possible roles of calmodulin binding to FY in flowering pathway.

Copyright © Lavanya Dampanaboina 2011

CHAPTER FOUR

Functional Characterization of the polyadenylation factor FY in 3' end processing using a tethering assay

Introduction

FY is an RNA 3' end processing factor and is also part of the autonomous pathway, one of the pathways regulating the transition from vegetative to the reproductive phase in Arabidopsis. The autonomous pathway targets *FLC* gene repression for promoting flowering. FLC is a central repressor of flowering and FY and FCA together promote the low expression of (or an inactive) FLC protein. FY mediates this FLC repression by physical association with FCA, an RNA binding protein (Marquardt, Boss et al. 2006). The repression of FLC gene expression by these proteins (FY and FCA) works at the level of chromatin by affecting histone demethylase activity. CstF64 and CstF77 are also identified to be necessary for histone demethylase activity on nucleosomes by promoting the production of FLC antisense transcripts (Liu, Marquardt et al. 2010). This network of interactions affects chromatin structure at other loci in addition to the *FLC* gene (Liu, Quesada et al. 2007). FY also interacts with AtCPSF100 and AtCPSF160 (Hunt, Xu et al. 2008) (Rao, Dinkins et al. 2009) and forms different complexes that might also play a role in chromatin silencing (Manzano, Marquardt et al. 2009).

The FY and FCA proteins not only repress *FLC* gene expression but also auto-regulate FCA protein levels by controlling polyadenylation site choice in the processing of FCA transcripts. There are no fewer than 4 different mRNA lengths for *FCA* in Arabidopsis due to alternative 3' end processing and demarcated by the greek letters, α, γ, δ and β in order of decreasing size (Marquardt, Boss et al. 2006) with only γ form producing full length active FCA protein. The FY + FCA protein complex binds to FCA pre-mRNA at the proximal polyadenylation site located in the third intron, promoting the formation of a truncated FCA mRNA (β transcript).

FY is an orthologue of the yeast 3'end processing factor Pfs2p. Pfs2p protein is a 53kDa protein with seven WD repeats and is part of the CFII-PFI complex (Ohnacker, Barabino

et al. 2000). Mutations in *Pfs2* cause cleavage and polyadenylation defects. *In vitro* studies show that Pfs2p interacts with Ysh1p, Fip1p and Rna14p. Mutations in *Pfs2* lead to chromosomal segregation defects (Wang, Asakawa et al. 2005). *PFS2* is an essential polyadenylation factor and its absence is lethal in yeast.

FY plays multiple roles in plant growth and development. Null alleles of *FY* are embryo lethal in Arabidopsis and deleterious to growth in *Nicotiana benthamiana* (Henderson, Liu et al. 2005). *fy* mutants are late flowering with more FLC protein accumulation in the mutant plants. Study of hypomorphic *fy* late flowering mutants show the requirement of both WD repeats and C-terminal PPLP motifs in FLC repression (Henderson, Liu et al. 2005). *In vitro* assays show that these PPLP repeats are also necessary for the interaction with FCA (Henderson, Liu et al. 2005). Mutations in *FY* usually cause late flowering phenotypes but a recent report on the hypomorphic *fy5* allele shows an earlier flowering phenotype and depicts FY as an activator or repressor of *FLC,* independent of FCA (Feng, Jacob et al. 2011).

FY is part of the CPSF complex and associates with subunits of CPSF complex other than FCA protein. In contrast, FCA is not associated with polyadenylation factors other than FY (Manzano, Marquardt et al. 2009). FY mutants that lack the motif for FCA interactions do not form these CPSF complexes. This report suggests that FCA, being an RNA binding protein, might initiate the process leading to transient interactions with FY, and in this way play a role in changing interactions of different polyadenylation factors in the pathway of RNA silencing by chromatin modifications (Manzano, Marquardt et al. 2009).

Recent reports show that FY also works independently in FLC repression. Null alleles of *fy* are embryo lethal but hypomorphic alleles of *fy* show different properties in different genetic background. In a winter annual background, hypomorphic *fy* mutants have low FLC levels with no late flowering phenotype. In these studies, it was noticed that the *FLC* gene has two poly(A) sites and both proximal and distal poly(A) sites of *FLC* are expressed at similar levels. In hypomorphic *fy* mutants the proximal poly(A) site of *FLC*

is preferred. Interestingly, in a null *fca* background FY partially represses *FLC* expression and promotes proximal poly(A) site choice (Feng, Jacob et al. 2011).

In the present study we hypothesize that FY, being an RNA 3'-end processing factor, is able to assemble a complete, functional 3' end processing complex if an RNA binding protein like FCA binds to FY and brings it to the site of polyadenylation. This hypothesis is tested using a tethering assay similar to those described by others (Coller, Gray et al. 1998). The results of the study show that FY is able to assemble an RNA 3'-end processing complex.

MATERIAL AND METHODS

Design and plasmid vector construction

Tethering assays using the bacteriophage MS2 coat protein (Coller, Gray et al. 1998; Graveley and Maniatis 1998; De Gregorio, Preiss et al. 1999) were employed to study the role of FY in recruiting 3' end processing complex and carrying out polyadenylation. For transient studies in-frame translational fusion constructs between the MS2 coat protein (MS2CP) and full length FY were made using pCAMBIA-1301 (CAMBIA, George Street, Brisbane, QLD, Australia) derived pGD vectors (Goodin, Dietzgen et al. 2002). MS2 coat protein was PCR amplified using a 5' MS2CP forward primer and a MS2CP reverse primer (Table 1) using the plasmid template kindly provided by Dr. Gravely (Graveley and Maniatis 1998). The PCR products encoding the MS2 coat protein was digested with restriction enzymes *Xho*I and *Hin*d III and cloned in pGD vector at the *Xho*I and *Hin*dIII restriction sites. DNA fragments encoding full length *FY* (1-647 amino acids), the N-terminal domain of FY (1-471 amino acids), the C-terminal domain of FY (465-647 amino acids), the full length FY with the calmodulin binding domain (508-520aa) deleted (cFY) were cloned into *Bam*HI-digested pGD clones that had the MS2 coat protein sequences present in the cassette (Table 1). Full length *FY* was PCR amplified with gene specific primers using cDNA from wild type Arabidopsis (Col-0) plants. The PCR product was digested with *Bam*HI and cloned into the pGD vector using the same site. The N and C-terminal domains of FY were amplified using gene specific primers and full length FY as a template. The PCR products were restriction digested with *Bam*HI, gel purified and cloned into *Bam*HI sites of the pGD vector containing the MS2 coat protein. The structures of these constructs are shown in Figs 4.1 and 4.2.

Deletion of the calmodulin binding domain (S. Rao and A. G. Hunt, unpublished results) from the full length FY was accomplished by designing primers to amplify two fragments flanking the calmodulin binding domain so as to insert a ligation-compatible restriction site between the two fragments. The two fragments were amplified in separate PCR reactions using four primers, the resultant fragments were independently digested with *Sal*I and *Xho*I and the purified fragments were ligated. The ligated products were

amplified using the full length *FY* primers resulting in the *FY* with the calmodulin bynding domain deleted (*cFY*). The PCR product was digested with *Bam*HI and cloned at the same sites in the pGD vector containing the MS2 coat protein.

Fluorescent reporters, for use as positive controls, were made by cloning the nuclear localization signal of the SV40 large T-antigen (PKKKRKV) in the pGDG and pGDR vectors (Fanara, Hodel et al. 2000) (Fig.4.7). Oligonucleotides containing the SV40 NLS coding sequences with *Bam*HI over hangs were annealed and ligated with *Bgl*II digested pGDG or pGDR vectors. To remove the nos terminator from pGDG vector containing the SV40 NLS, the vector was digested with *Bam*HI (last site in the MCS) and *Bss*HII (2 sites in the nos region and removes most of the region). The digested vector was filled in (blunted) using klenow polymerase and blunt end ligated intra-molecularly. The MS2 binding sites were cloned in the resultant vector at the *Hin*dIII site. Another construct was made in which the MS2 binding sites were replaced with a mutated version that does not bind the MS2 coat protein.

Agro-infiltration experiments for transient expression studies

All the constructs, along with control vectors, were transformed into *Agrobacterium tumefaciens* strain *LBA4404* using the freeze thaw method (An, Ebert et al. 1988). Agrobacterium was freshly grown on LB agar plates with appropriate antibiotics for 2 days at 28^{0}C. The culture plates were scraped with a spatula and suspended in freshly made MES buffer (10mM MgCl2 and 10mM MES). The suspensions were measured for the cell density in a spectrophotometer at 600nm. Using the optical density (OD) readings, the cell mass was set to 0.8 OD by dilution with fresh MES buffer. Acetosyringone was added to all samples at 150µM concentrations and incubated at room temperature for 2-3 hours. For co-infiltrations equal volumes of different Agrobacterium samples were mixed before infiltrations.

Three weeks old *Nicotiana benthamiana* plant were selected and leaves were infiltrated with the Agrobacterium suspensions using a 1ml syringe without a needle. Healthy looking and fully expanded lower leaves were selected for the infiltration experiments.

Gentle pressure was applied with a finger on the adaxial surface of the leaf while injecting the suspension buffer from the abaxial surface of the leaf. The desired amount of the buffer with bacterial suspension was infiltrated in the intravenous regions. More than 2 infiltration sites were made with the tip of the syringe to complete the infiltration in a single leaf. At least two leaves were infiltrated with the same suspension on a single plant. Once the leaves are infiltrated, a water soaked appearance was observed. The plants were incubated for 48 to 72 hours at 28˚C under Fluorescent light of 16hr/8hr light and dark photoperiods.

Fluorescence Microscopy

Infiltrated leaves (after 48-72hr of incubation) were made into small sections near the infiltration site. The sections, with the abaxial side up, were placed on a glass slide in a drop of water and mounted gently with a cover slip. The slide was observed under UV light using an epifluorescence microscope (Zeiss Axioplan2 HB100). The GFP expressing epidermal cells were visualized using the green filter (ex: D470/40; D535/40; beam splitter 500 DCLP) and DsRed was observed under the red filter (ex: HQ545/30X; em: HQ610/75M, Q570LP) filter sets. Images were captured using the Zeiss AxioCam MRc5 attached to the microscope. AxioVision software was used to adjust the exposure for each filter and also get the merged picture for the co-infiltration samples along with the pictures in green and red filters.

RESULTS

Proteins that bind RNA and associate with other proteins in the multi-subunit complex mainly regulate gene expression. Tethering assay is the assay where function of protein ‘X’ is tested by separating its RNA binding. Two constructs are required for performing tethering assay. One construct is the tethering construct made by fusing protein ‘X’ to be tested to the MS2 coat protein. The other construct is the reporter construct. Specific binding and affinity of tethered construct to the reporter construct is mediated by the MS2 binding sites in the reporter driven by MS2 coat protein as depicted in Fig.4.3. In the tethering assay employed here, the test protein FY is fused to MS2 coat protein and tethered to genetically engineered reporter mRNA with out poly(A) signal to understand the function of FY. GFP reporter expression in the tethering assay is the indication of FY functioning in recruiting 3’-end processing complex (Fig.4.4).

The reporter construct (Fig.4.5) was modified by deleting the nos terminator, GFP expression was not observed, showing that the presence of the poly(A) tail is necessary for GFP reporter expression (Fig.4.6). In order to test that there was no experimental errors in infiltration, the reporter construct without the nos region was co-infiltrated with a RFP NLS marker. RFP expresses in these samples confirming that GFP expression needs a poly (A) tail (Fig.4.6). In the next step the reporter construct was also tested for GFP expression after introducing wild type MS2 coat protein binding sites to test whether the hairpin structure of MS2 coat protein binding sites alone might affect GFP expression. Similarly, the reporter construct with mutated MS2 coat protein binding sites was also tested for expression of the GFP reporter. No GFP expression was seen with this construct (Fig. 4.7), confirming that poly(A) tail was necessary for GFP reporter expression. These results confirmed that the reporter mRNA with functional MS2 coat protein binding sites can be used in the tethering assay for testing the role of FY in RNA 3’ end processing.

Coinfiltration of the tethering construct (MS2CP-FY) (Fig.4.5) in combination with the reporter construct resulted in GFP reporter expression (Fig.4.8). Coinfiltration of the

MS2CP construct (without FY) i.e. tethered control construct (Fig.4.9) in combination with the reporter construct did not show any GFP expression, confirming the requirement for FY (Fig.4.10). The same experiments were also done using reporter constructs with non-functional MS2 coat protein binding sites i.e reporter control construct (Fig.4.9) in combination with the tethering construct (MS2CPFY) (Fig.4.11) and, as expected, there was no reporter expression suggesting that functional MS2 binding sites are necessary for FY binding at the 3' UTR of the GFP reporter mRNA.

A tethering construct encoding a mutant FY that does not bind calmodulin (cFY) was also tested with the MS2CPGFP reporter construct (Fig.4.12). Interestingly, no GFP expression was seen in these experiments, even though the internal RFP control was expressed (Fig.4.13).

In summation, the GFP reporter expresses in the presence of a tethering construct (MS2CP+FY), and this expression is dependent on both FY and the MS2 coat protein binding sites. These results support the model that FY can directly recruit the 3' processing complex to a pre-mRNA. The results also show that, in the absence of calmodulin binding domain in FY, the expression of GFP goes down. This suggests that the calmodulin-binding domain is important in FY functioning (Fig.4.14).

DISCUSSION

Our hypothesis is that if FY is brought into physical proximity of a pre-mRNA 3' end, in this case using the MS2 coat protein fused to FY and an unprocessed pre-mRNA including the MS2 coat protein binding sites (RNA), FY should be able to assemble other polyadenylation factors, to bring about the cleavage and polyadenylation of the target pre-mRNA. Two vectors were assembled to test the hypothesis that FY is able to assemble the 3' end processing complex when mediated by the MS2 coat protein and that this would result in the polyadenylation of the *GFP* reporter mRNA. A poly (A) tail is necessary for the GFP reporter expression and GFP expression is therefore a measure of FY activity in recruiting the 3'-end processing complex. Two reporter constructs were made, one with wild type MS2 coat protein binding sites and one with a non-functional MS2 coat protein-binding site. The pGDG backbone was also modified to remove the nos terminator. For easy visualization in the assay, nuclear localization signals were added to the GFP and RFP coding regions in the pGDG and GDR vectors, respectively.

The second vector was the tethered construct of the MS2 coat protein and FY. The affinity of the MS2 coat protein to the MS2 coat protein binding sites will bring FY to the 3' UTR of the GFP reporter mRNA. A variation of the tethering construct was made with the MS2 coat protein and cFY (FY-DCBD; see Chapter III) for use in the tethering assay to test the effect of the calmodulin binding domain on FY function. A plasmid encoding just the MS2 coat protein with a nuclear localization signal was used as a negative control for the tethering construct.

The results support this hypothesis in that they show that the MS2 coat protein-FY chimera can promote the polyadenylation of the GFP reporter mRNA. When mutant MS2 binding sites replace wild type MS2 binding sites at the 3' end of the GFP reporter mRNA, the FY-MS2 coat protein chimera is unable to process the GFP reporter pre-mRNA, indicating that it is important for FY to be proximal to the 3'UTR region of the GFP reporter pre-mRNA to function. The tethering assay performed with the MS2 coat protein alone (not fused to FY) resulted in no GFP expression suggesting that FY is

responsible for the assembly of the CPSF complex that results in the cleavage and polyadenylation of the GFP mRNA.

Simpson et al. (2003) proposed a model while studying the functional role of FY in FCA auto-regulation where FY is recruited to the FCA-pre-mRNA and promotes the assembly of the 3' end processing complex and results in the choice of a proximal polyadenylation site in the FCA pre-mRNA. The resulting truncated FCA (β and γ versions) mRNAs produce proteins that are incapable of inhibiting Flowering Locus C (FLC) expression. The results described in this chapter supports this model (Simpson, Dijkwel et al. 2003).

Yeast two hybrid interaction studies of FY reveal that it interacts with CPSF160 and CPSF100 which interact with CPSF30 which, in turn recruits CstF77 and attendant Fip1 proteins, into contact and forms a complex for both cleavage and polyadenylation (Hunt, Xu et al. 2008). The results presented here are consistent with these previous studies, and indicates that FY may promote polyadenylation by assembling the CPSF complex at the appropriate site on a pre-mRNA. The presence of a CPSF complex might also recruit CstF factors and poly(A) polymerase, ensuring the proper RNA 3'-end processing of the pre-mRNA.

Table 4.1 Primers used for the tethering assay

Designation	Sequence (5' -> 3')	Purpose
MS2CP forward and reverse primers	Forward primer: TTT TTT GTC GAC ATG GCT TCT AAC TTT ACT CAG TTC GTT	Amplification of MS2 coat protein
	Reverse primer: CCC CCC AAG CTT TTA TCA GGG CCC ACC ACC ACC ACC GTA	
Wild type MS2 binding sites	Forward primer: AGCT CGT ACA CCA TCA GGG TAC GAAAGGGCGT ACA CCA TCA GGG TAC G	Amplification of Wild type MS2 binding sites
	Reverse primer: AGCT CGT ACC CTG ATGGTGTACG CCCTTT CGTACCC TGATG GTGTACG	
Mutant MS2 binding sites	Forward primer: AGC TCG TTG ATC AGC AGG GTA CGA AAG GGC GTT GAT CAG CAG GGT ACG	Amplification of Mutant MS2 binding sites
	Reverse Primer: AGC TCG TAC CCT GCT GAT CAA CGT TTC CCC GTA CCC TGC TGA TCA ACG	
SV40 Nuclear localization signal	SV40 NLS forward primer: GATCG CCA AAA AAG AAG AGA AAG GTA GCC TAA G;	For the nuclear localization signal
	SV40 NLS reverse primer: GATCC TTA GGC TAC CTT TCT CTT CTT TTT TGG C	
Full length FY primers	FY forward primer: TTT TTT GGA TCC ATG TAC GCC GGC GGC GAT ATG CAC	Amplification of full length FY
	FY reverse primer: CCC CCC GGA TCC CTA CTG ATG TTG CTG ATT GTT GTT	

Table 4.1 continued……….

Table 4.1 continued……..

Designation	Sequence (5' -> 3')	Purpose
FY NTD primers	FY reverse primer: CCC CCC GGA TCC CTA CTG ATG TTG CTG ATT GTT GTT	Amplification of FY-NTD domain
	FY NTD forward primer: TTT TTT GGA TCC ATG TAC GCC GGC GGC GAT ATG CAC	
FY CTD primers	FY NTD reverse primer: CCC CCC GGA TCC GAG AGG ATG CAT CAA ATG GCA TTG	Amplification of FY-CTD domain
	FY CTD forward primer: TTT TTT GGA TCC ATG CCA TTT GAT GCA TCC TCT CAA GGG: FY CTD reverse primer: CCC CCC GGA TTC TGA TGT TGC TGA TTG TTG TTT GGT)	
cFY (full length FY with deleted calmodulin binding domain)	cFY 1 forward primer: TTT TTT GGA TCC ATG TAC GCC GGC GGC GAT ATG CAC	Amplification of Full length cFY by deletion of calmodulin binding domain.
	cFY 1reverse primer: CCC CCC GTC GAC ATA CCC TTG CTG CTG GCC ACT TCC	
	cFY 2 forward primer: TTT TTT CTC GAG CTT CCA ATG CCC AAT ATG CCT CAC	
	cFY 2 reverse primer: CCC CCC GGA TCC CTA CTG ATG TTG CTG ATT GTT GTT)	

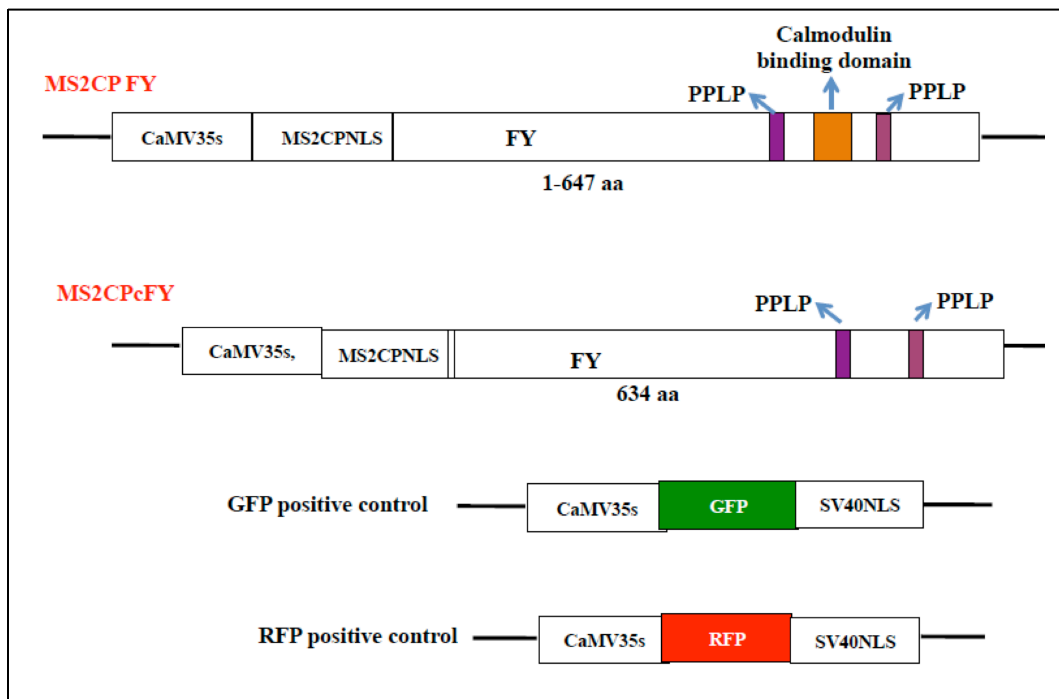


Figure 4.1 Tethering constructs of FY for testing the Tethering assay

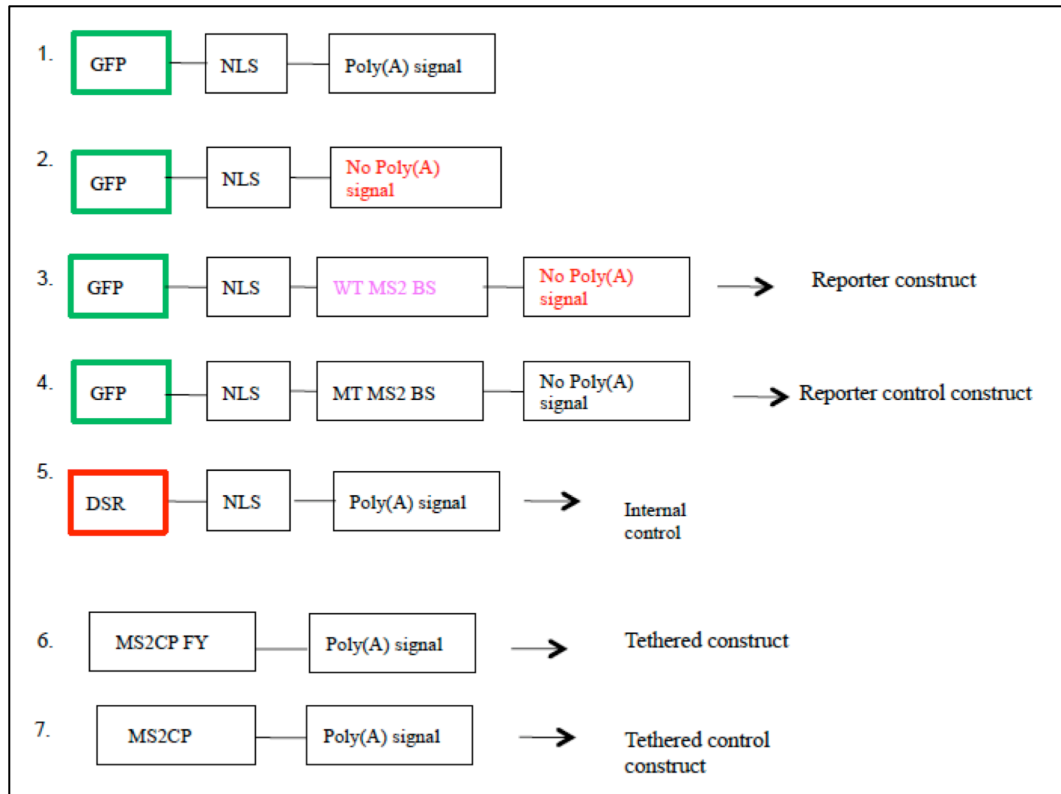


Figure 4.2 Constructs made to test the tethering assay.

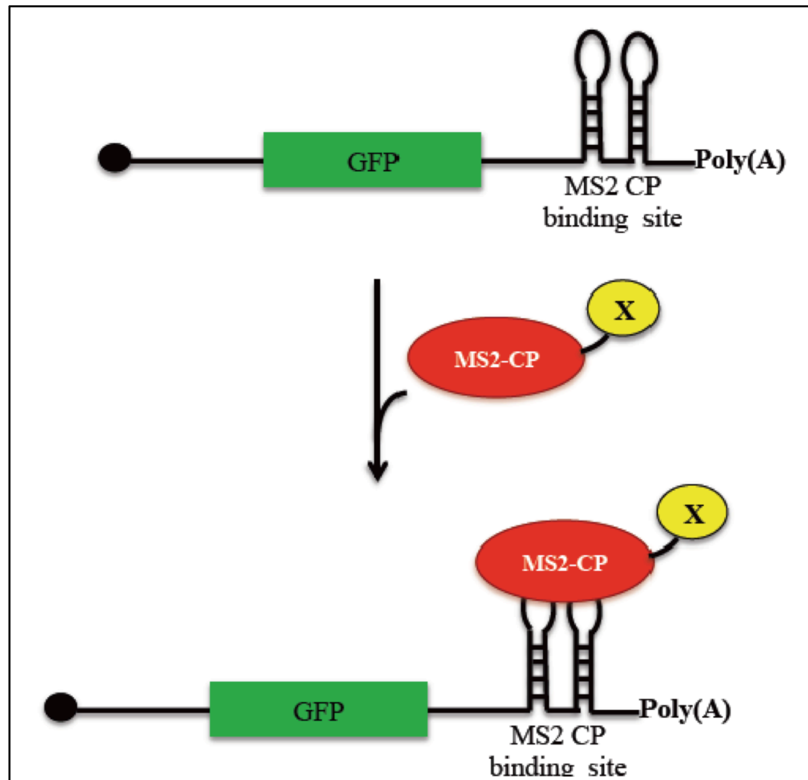


Figure 4.3 Principle involved in tethering assay.

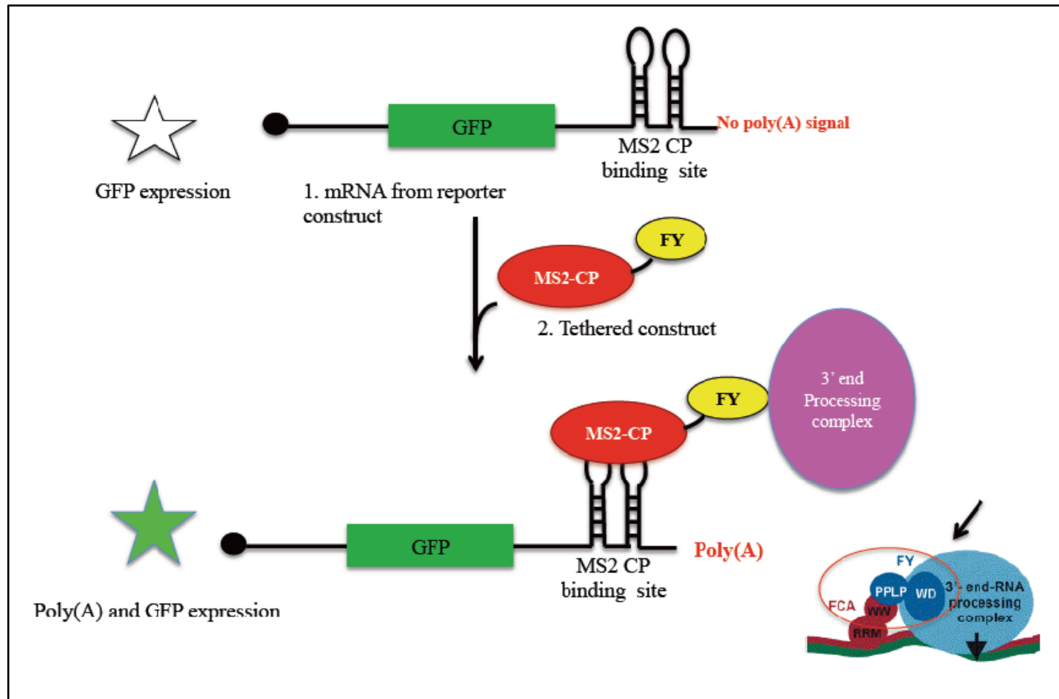


Figure 4.4 Model for FY function in recruiting 3' end processing complex. (FY model in recruiting 3' end processing complex adapted from Simpson et al, 2003)

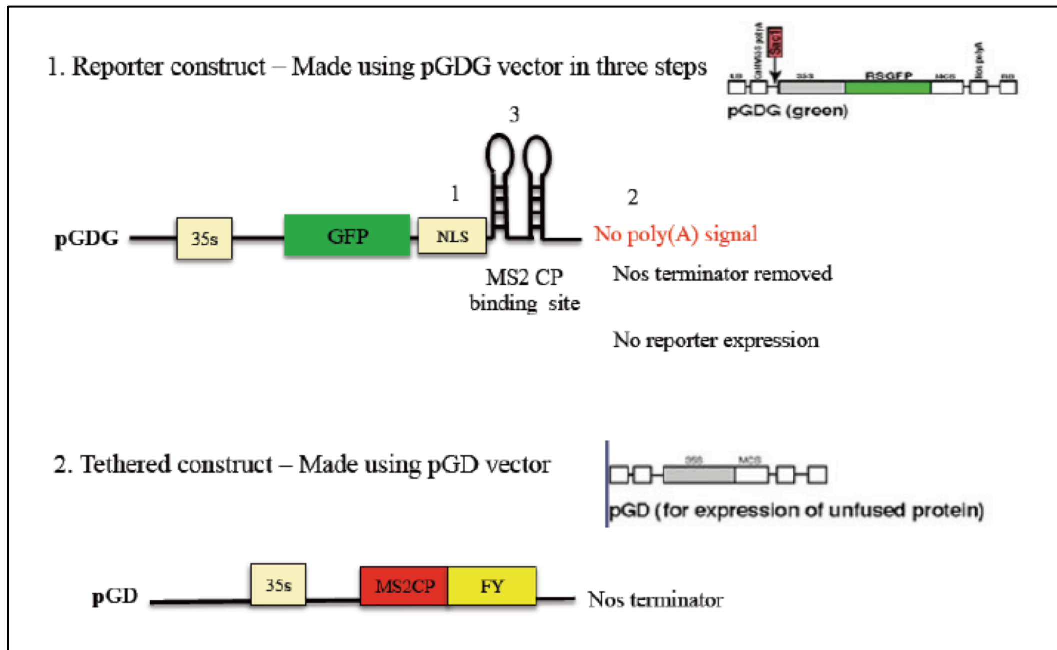


Figure 4.5 Two constructs reporter construct and tethering construct made to test the hypothesis (Goodin et al, 2002).

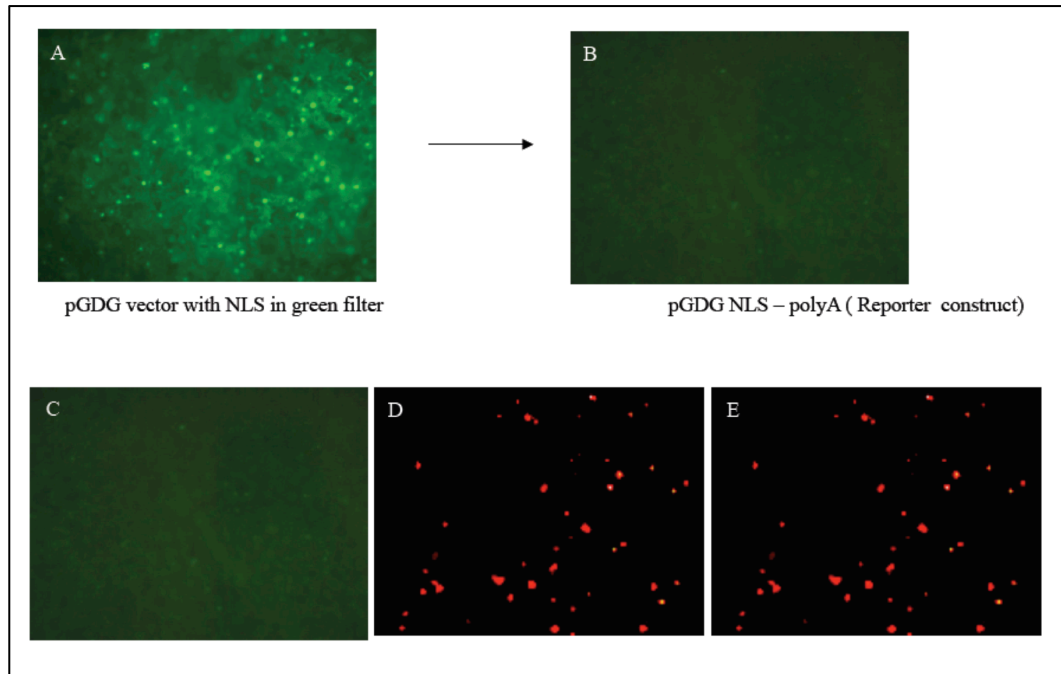


Figure 4.6 Poly(A) signal is necessary for the GFP reporter expression.

Panel 1: A. GFP NLS marker in green filter **B.** GFP NLS marker with no nos terminator in green filter

Panel 2: Co-infiltration of GFP NLS construct with out nos terminator with RFP nuclear marker **C.** GFP NLS with no Nos terminator in green filter **D.** RFP NLS marker in red filter **E.** Merging both green and red filters.

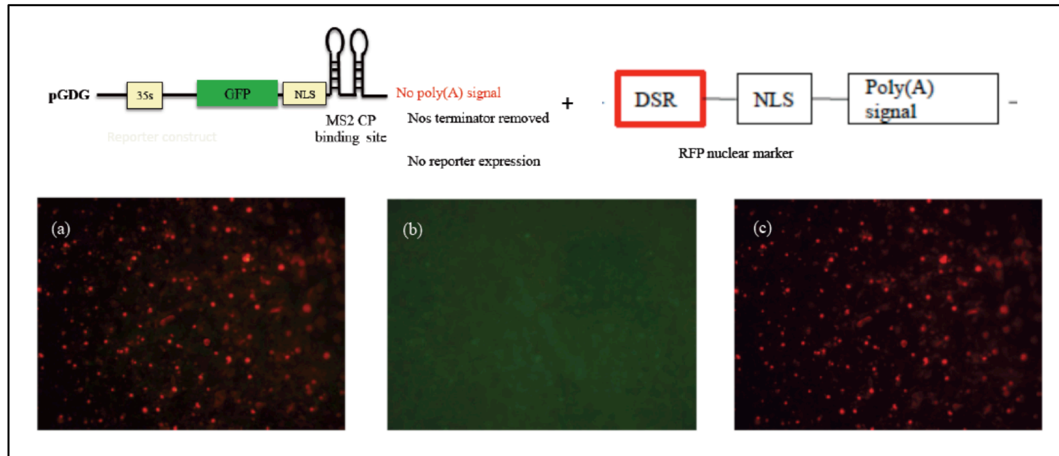


Figure 4.7 Reporter construct (GFP reporter, wild type MS2 binding sites, NLS and no poly(A) signal) for the tethering assay does not express GFP.

Co-infiltration of reporter construct and RFP nuclear marker in (a) merged green and red filter (b) green filter and (c) red filter (Result shows that poly(A) was necessary for GFP reporter expression and introduction of functional MS2 coat protein binding sites do not change the reporter construct expression).

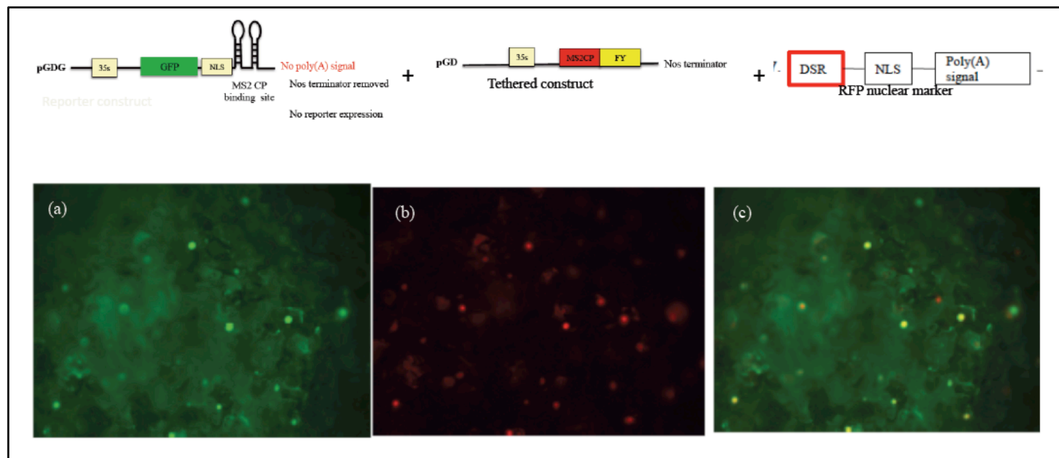
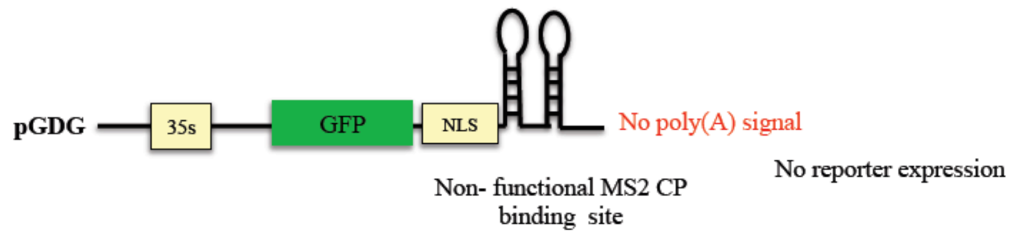


Figure 4.8 Tethering assay with reporter construct and tethering construct with internal RFP nuclear control.
(a) merged green filter (b) green filter and (c) red filter (Result shows that FY was able to recruit 3' end processing complex and was expected from the GFP expression of the reporter construct).

1. Reporter control construct – using pGDG vector



2. Tethered control construct – using pGD Vector

Figure 4.9 Control constructs made for tethering assay.

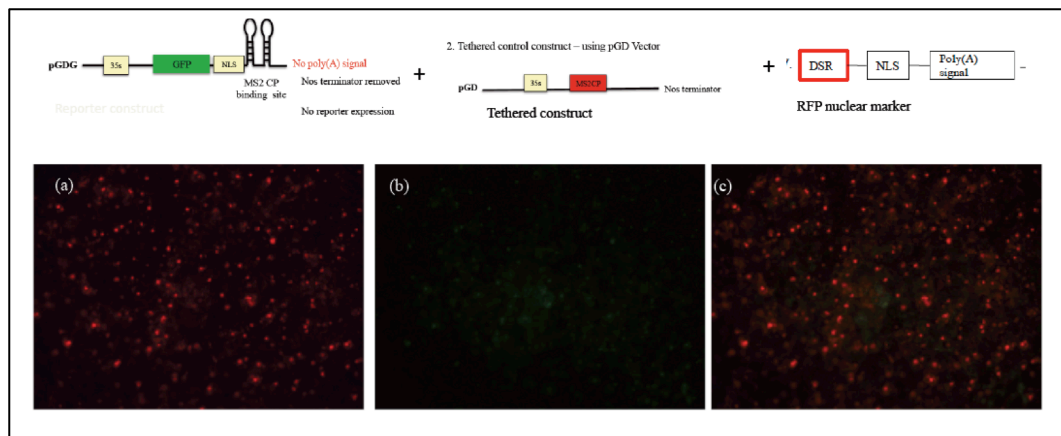


Figure 4.10 Co-infiltration of reporter construct with RFP nuclear marker and tethered control construct (with out FY).
(a) red filter (b) green filter (c) merged red and green filters (Result shows that FY alone drives the GFP expression and absence of FY does not allow poly(A) tail formation on GFP mRNA)

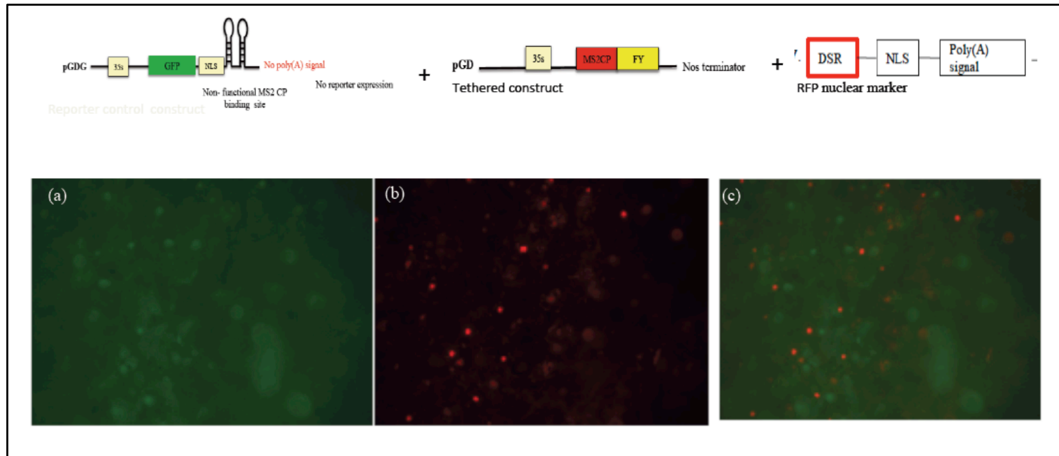


Figure 4.11
Co-infiltration of tethered construct with reporter control construct and RFP nuclear marker.

(a) green filter (b) red filter (c) merged red and green filters (Result shows that Presence of functional MS2 coat protein binding sites is necessary for GFP expression driven by FY)

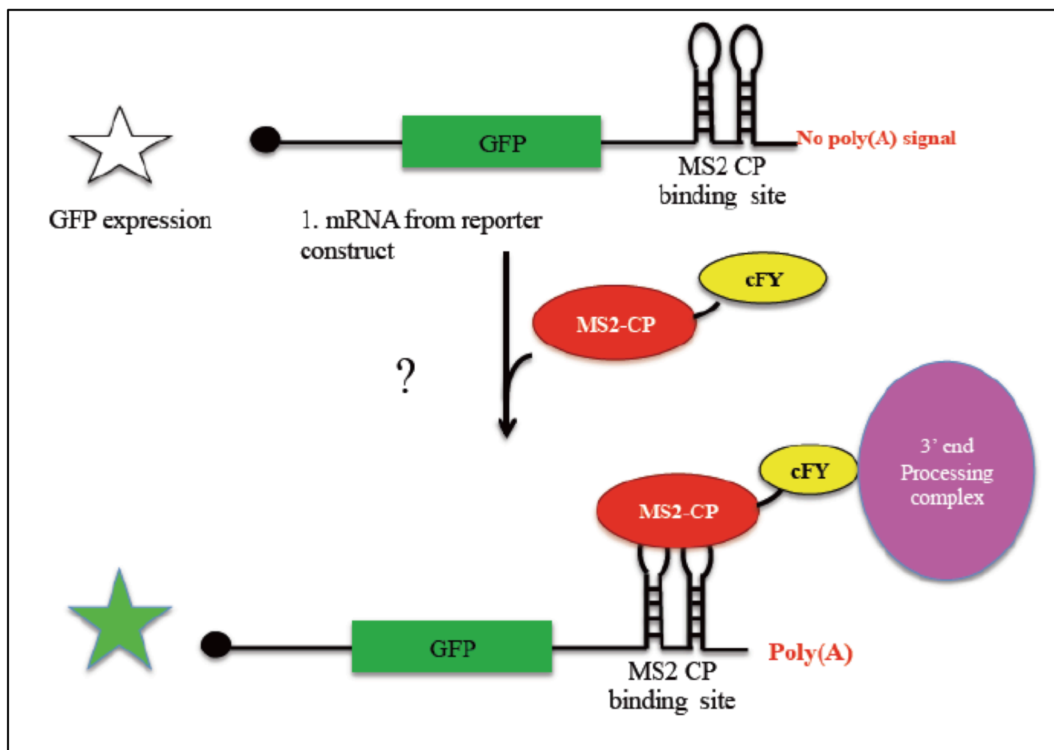


Figure 4.12 Tethering assay model to test whether absence of calmodulin binding domain affects FY function.

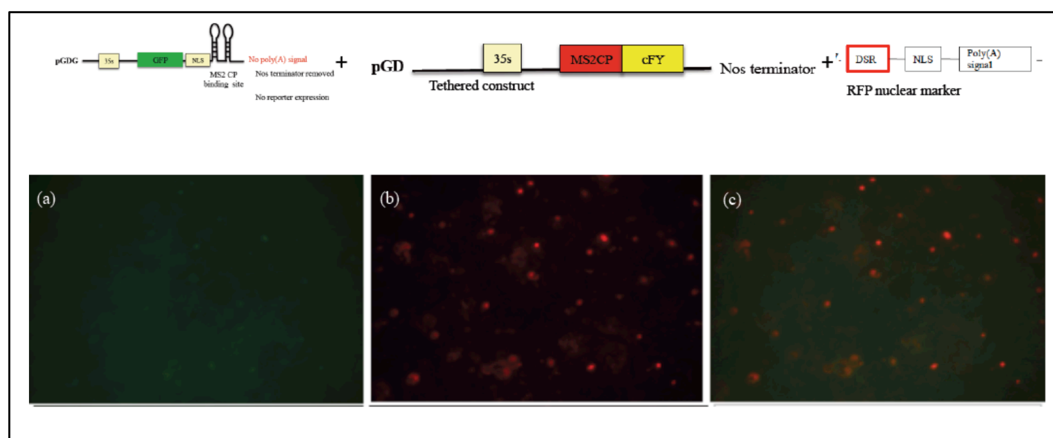


Figure 4.13 Co-infiltration of tethered construct cFY with reporter construct and RFP nuclear marker.
(a) green filter (b) red filter and (c) merged picture of green and red filter. (Results show that Calmodulin binding domain was necessary in FY functioning)

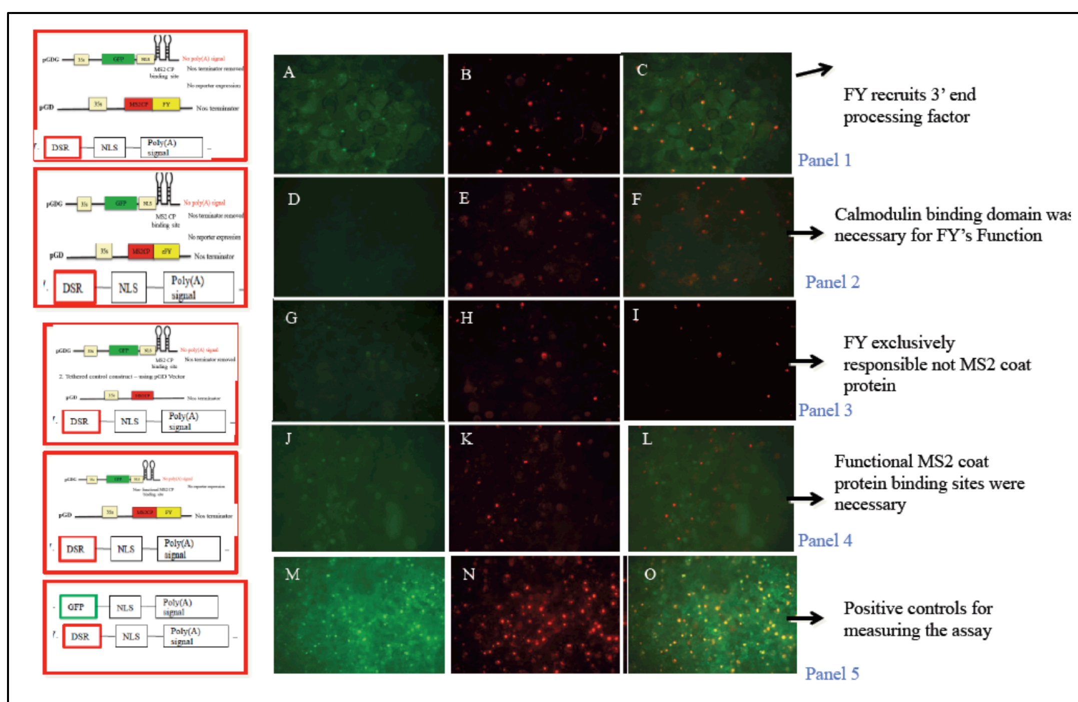


Figure 4.14 Summary of tethering assay.
Panel 1: Tethering assay with reporter construct and tethering construct with internal RFP nuclear control in (A) green filter (B) red filter and (c) merged picture of green and red filter (Result shows that FY was able to recruit 3' end processing complex on the GFP reporter mRNA and polyadenylate)

Panel 2: Co-infiltration of tethered construct cFY with reporter construct and RFP nuclear marker (D) green filter (E) red filter and (F) merged picture of green and red filter. (Results show that CBD was necessary in FY functioning)

Panel 3: Co-infiltration of reporter construct with RFP nuclear marker and tethered control construct (with out FY) (G) green filter (H) red filter and (I) merged picture (Result shows that FY alone drives the GFP expression and absence of FY does not allow poly(A) tail formation on GFP mRNA)

Panel 4: Co-infiltration of reporter construct with RFP nuclear marker and tethered control construct (with out FY) (J) green filter (K) red filter and (L) merged picture (Result shows that FY alone drives the GFP expression and absence of FY does not allow poly(A) tail formation on GFP mRNA)

Panel 5: Co-infiltration of GFP nuclear and RFP nulcear marker (M) green filter (N) red filter and (O) Merged picture of green and red filter.

Copyright © Lavanya Dampanaboina 2011

CHAPTER FIVE

Summary

Polyadenylation is an essential post-transcriptional modification resulting in a mature mRNA from all RNA polII transcripts in eukaryotes. In plants, three cis-elements – the Far Upstream Element (FUE), Near Upstream Element (NUE), and Cleavage Site (CS) - guide the process of cleavage and polyadenylation with the help of multi-subunit protein complexes CPSF (cleavage and polyadenylation specificity factor), cleavage stimulation factor (CstF) along with cleavage factors and poly(A) polymerase. RNA-protein interactions and protein-protein interactions play a vital role in the process of cleavage and polyadenylation. In the present study, WD repeat proteins AtCstF50 and AtFY were studied.

In mammals, Cleavage stimulation Factor 50 (CstF50) is one of the subunits of Cleavage and stimulation Factor complex along with other two subunits CstF64 and CstF77. CstF77 bridges two subunits CstF64 and CstF50. CstF64 and CstF77 strongly interact with each other. CstF50 is a WD repeat protein with seven WD repeats and interacts with RNA polII C-terminal domain with a possible role in transcription termination. The other important role is in tumor suppression by associating with BARD1. This interaction of CstF50 represses polyadenylation and helps in regulating the expression of right processed RNA's. AtCstF50 is also a WD repeat protein but contains six WD repeats. Study of T-DNA knockout lines suggests that AtCstF50 is essential in Arabidopsis polyadenylation machinery. Protein interaction studies showed that AtCstF50 does not interact with any of the CstF subunits neither with CstF64 nor with CstF77 showing that it might be working differently from its mammalian orthologue. AtCstF50 also interacts positively with poly(A) polymerase, Fip (Factor interacting with poly(A) polymerase), poly(A) binding proteins and CPSF subunits, CPSF30 and CPSF100. These results showed that AtCstF50 is part of CPSF complex rather than part of CstF complex. But reports on plant nuclear extracts containing CPSF complexes did not yield CstF50 suggesting that interaction of CstF50 with CPSF is a transient interaction in 3'-end processing.

AtFY is another WD repeat protein involved in 3'-end processing. AtFY contains seven WD repeats at N-terminus and is an orthologue of Yeast Pfs2p and forms part of the CPSF complex. Pfs2p is an essential yeast-processing factor involved in assembling different polyadenylation factors in cleavage and polyadenylation process. AtFY has an extended C-terminal domain with the evolution of PPLP motifs unique to the plant system. Yeast Pfs2p null mutants are lethal while Arabidopsis FY null alleles are embryo lethal and deleterious to growth in *Nicotiana benthamiana.*

FY is also a component of the autonomous pathway, promoting the transition to the reproductive phase in Arabidopsis. In this pathway, FY targets FLC, a central repressor of flowering for repression by associating with another nuclear RNA binding protein, FCA. The FY-FCA interaction promotes the alternative polyadenylation of FCA mRNA itself and results in a truncated FCA protein. It also found to have a role in FLC polyadenylation site choice. FY plays pivotal role in flowering process, alternative polyadenylation of FCA, embryogenesis, chromatin modifications and other unreported functions. FY contains a non-canonical calmodulin binding domain (CBD) in its C-terminal end between the PPLP motifs. It was interesting to note that FY lacking the CBD (cFY) localized mostly to the cytosol while FY localizes to the nucleus. The FY N-terminal domain, C-terminal domain, and C-terminal domain lacking the CBD (cFYc) also localize in the nucleus, indicating the presence of multiple nuclear localization signals in FY. It is also possible that calmodulin might be affecting the interaction of FY with other proteins like CPSF and unknown proteins that might be responsible for nuclear localization. One more observation was the amount of reporter expression was relatively high in the FY-CTD fusion while FY and FY-NTD showed relatively lower reporter expression. It is known that FY WD motif is involved in the interaction of FY with DDB1, an ubiquitin ligase. The expression differences might be attributed to the presence or absence of WD motifs and thus susceptibility to ubiquitin-mediated degradation.

During FCA auto-regulation FY together with FCA, an RNA binding protein associates with FCA mRNA and recruits 3' end processing complex and controls the poly(A) site choice in intron 3 for the production of truncated FCA transcript. Our hypothesis was FY

being a 3'-end processing factor is able to recruit the 3'-end processing complex and drive the polyadenylation. A tethering assay was employed to test the hypothesis. GFP expression was used as a measure for FY activity in recruiting 3' end processing complex. For the tethering assay a reporter construct and tethering construct were made. The reporter construct contains GFP reporter, nuclear localization signal, functional MS2 coat protein binding site but lacks poly(A) tail and do not show any reporter expression. Tethering construct was made with MS2 coat protein and FY. Co-expression of the reporter and tethering construct should yield GFP expression if FY was able to recruit 3' end processing complex. As expected co-expression of reporter construct and tethered construct resulted in GFP expression proving our hypothesis.

Additional controls used in the tethering assay also showed that poly(A) tail was necessary for GFP expression, functional binding sites of MS2 coat protein drive the process and it is also important for FY to be at the 3'UTR region to act as a 3' end processing factor and FY is exclusively responsible for the reporter expression. To test whether a mutation in CBD in FY affects its function a tethered construct was made with cFY (CBD domain deleted) and used in the tethering assay along with the reporter construct. Lack of GFP expression with cFY showed that this domain is important for the proper function of FY in 3' end processing.

Localization studies showed that absence of CBD changes the localization mostly to the cytosol while tethering assay using cFY resulted in lack of GFP expression. Due to change in the localization of cFY it could not drive the nuclear polyadenylation in the tethering assay, hence there was no GFP expression. A model was proposed based on the tethering results and on FY being part of CPSF complex. FY might be the first factor to be at the 3' UTR of the mRNA and its interactions with other subunits of CPSF, might assemble CstF subunits and able to even pull poly(A) polymerase and poly(A) binding protein necessary for poly(A) tail addition.

Copyright © Lavanya Dampanaboina 2011

REFERENCES

Adamson, T. E., D. C. Shutt, et al. (2005). "Functional Coupling of Cleavage and Polyadenylation with Transcription of mRNA." Journal of Biological Chemistry **280**(37): 32262-32271.

Addepalli, B. and A. G. Hunt (2007). "A novel endonuclease activity associated with the Arabidopsis ortholog of the 30-kDa subunit of cleavage and polyadenylation specificity factor." Nucleic Acids Research **35**(13): 4453-4463.

Addepalli, B., L. R. Meeks, et al. (2004). "Novel alternative splicing of mRNAs encoding poly(A) polymerases in Arabidopsis." Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression **1679**(2): 117-128.

Alonso, J. M., A. N. Stepanova, et al. (2003). "Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana." Science **301**(5633): 653-657.

An, G., P. R. Ebert, et al. (1988). "Binary Vectors, Plant Molecular Biology Manual A3, ." Kluwer Academic Publishers, Dordrecht, Belgium: 1-19.

Anderson, J. T., S. M. Wilson, et al. (1993). "NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability." Molecular and Cellular Biology **13**(5): 2730-2741.

Andreassi, C. and A. Riccio (2009). "To localize or not to localize: mRNA fate is in 3'UTR ends." Trends in Cell Biology **19**(9): 465-474.

Bai, C. and P. P. Tolias (1998). "Drosophila clipper/CPSF 30K is a post-transcriptionally regulated nuclear protein that binds RNA containing GC clusters." Nucleic Acids Research **26**(7): 1597–1604.

Bai, Y., T. C. Auperin, et al. (2007). "Crystal Structure of Murine CstF-77: Dimeric Association and Implications for Polyadenylation of mRNA Precursors." Molecular cell **25**(6): 863-875.

Baillat, D., M.-A. Hakimi, et al. (2005). "Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II." Cell **123**(2): 265-276.

Barabino, S. M., W. Hubner, et al. (1997). "The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins." Genes & Development **11**(13): 1703-1716.

Barillà, D., B. A. Lee, et al. (2001). "Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae." Proceedings of the National Academy of Sciences **98**(2): 445-450.

Bell, S. A. and A. G. Hunt (2010). "The Arabidopsis ortholog of the 77kDa subunit of the cleavage stimulatory factor (AtCstF-77) involved in mRNA polyadenylation is an RNA-binding protein." FEBS Letters **584**(8): 1449-1454.

Bouché, N., A. Yellin, et al. (2005). "Plant-specific calmodulin-binding proteins." Annual Review of Plant Biology **56**(1): 435-466.

Briggs, M. W., K. T. D. Burkard, et al. (1998). "Rrp6p, the Yeast Homologue of the Human PM-Scl 100-kDa Autoantigen, Is Essential for Efficient 5.8 S rRNA 3' End Formation." Journal of Biological Chemistry **273**(21): 13255-13263.

Brown, C. E., S. Z. Tarun, et al. (1996). "PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae." Molecular and Cellular Biology **16**(10): 5744-5753.

Buratowski, S. (2005). "Connections between mRNA 3 ' end processing and transcription termination." Current Opinion in Cell Biology **17**(3): 257-261.
Callebaut, I., D. Moshous, et al. (2002). "Metallo,-β-lactamase fold within nucleic acids processing enzymes: the β-CASP family." Nucleic Acids Research **30**(16): 3592-3601.
Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA polyadenylation." Genes & Development **11**(21): 2755-2766.
Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA, polyadenylation." Genes & Development **11**(21): 2755-2766.
Coller, J. M., N. K. Gray, et al. (1998). "mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation." Genes & Development **12**(20): 3226-3235.
Danckwardt, S., M. W. Hentze, et al. (2008). "3 ' end mRNA processing: molecular mechanisms and implications for health and disease." Embo Journal **27**(3): 482-498.
Dantonel, J.-C., K. G. K. Murthy, et al. (1997). "Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA." Nature **389**: 399-402.
Dass, B., K. W. McMahon, et al. (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." Journal of Biological Chemistry **276**(11): 8044-8050.
Day, I., V. Reddy, et al. (2002). "Analysis of EF-hand-containing proteins in Arabidopsis." Genome Biology **3**(10): research0056.0051 - research0056.0024.
De Gregorio, E., T. Preiss, et al. (1999). "Translation driven by an eIF4G core domain in vivo." EMBO Journal **18**(17): 4865-4874.
de Vries, H., U. Ruegsegger, et al. (2000). "Human pre-mRNA cleavage factor IIm contains homologs of yeast proteins and bridges two other cleavage factors." EMBO Journal **19**(21): 5895-5904.
Delaney, K. J., R. Xu, et al. (2006). "Calmodulin Interacts with and Regulates the RNA-Binding Activity of an Arabidopsis Polyadenylation Factor Subunit." Plant Physiology **140**(4): 1507-1521.
Dichtl, B. and W. Keller (2001). "Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor." EMBO Journal **20**(12): 3197-3209.
Dickson, K. S., A. Bilger, et al. (1999). "The Cleavage and Polyadenylation Specificity Factor in Xenopus laevis Oocytes Is a Cytoplasmic Factor Involved in Regulated Polyadenylation." Molecular and Cellular Biology **19**(8): 5707-5717.
Dominski, Z., X.-c. Yang, et al. (2005). "The Polyadenylation Factor CPSF-73 Is Involved in Histone-Pre-mRNA Processing." Cell **123**(1): 37-48.
Dufresne, P. J., E. Ubalijoro, et al. (2008). "Arabidopsis thaliana class II poly(A)-binding proteins are required for efficient multiplication of turnip mosaic virus." Journal of General Virology **89**(9): 2339-2348.
Edmonds, M. (2002). "A history of poly A sequences: from formation to factors to function." Progress in Nucleic Acid Research & Molecular Biology **71**: 285-389.

Elliott, B. J., T. Dattaroy, et al. (2003). "An interaction between an Arabidopsis poly(A) polymerase and a homologue of the 100 kDa subunit of CPSF." Plant Molecular Biology **51**(3): 373-384.
Elliott, B. J., T. Dattaroy, et al. (2003). "An interaction between an Arabidopsis poly(A) polymerase and a homologue of the 100 kDa subunit of CPSF." Plant Molecular Biology **51**(3): 373-384.
Fanara, P., M. R. Hodel, et al. (2000). "Quantitative Analysis of Nuclear Localization Signal (NLS)-Importin-α Interaction through Fluorescence Depolarization." Journal of Biological Chemistry **275**(28): 21218-21223.

Feng, W., Y. Jacob, et al. (2011). "Hypomorphic Alleles Reveal FCA-Independent Roles for FY in the Regulation of FLOWERING LOCUS C." Plant Physiology **155**(3): 1425-1434.

Fong, N. and D. L. Bentley (2001). "Capping, splicing, and 3′ processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD." Genes & Development **15**(14): 1783-1795.

Forbes, K. P., B. Addepalli, et al. (2006). "An Arabidopsis Fip1 Homolog Interacts with RNA and Provides Conceptual Links with a Number of Other Polyadenylation Factor Subunits." Journal of Biological Chemistry **281**(1): 176-186.

Fox, C. A., M. D. Sheets, et al. (1989). "Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU." Genes & Development **3**(12b): 2151-2162.

Garneau, N. L., J. Wilusz, et al. (2007). "The highways and byways of mRNA decay." Nat Rev Mol Cell Biol **8**(2): 113-126.

Gilmartin, G. M., E. S. Fleming, et al. (1992). "Activation of HIV-1 pre-mRNA 3' processing in vitro requires both an upstream element and TAR." The EMBO Journal **11**(12): 4419-4428.

Goodin, M. M., R. G. Dietzgen, et al. (2002). "pGD vectors: versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves." The Plant Journal **31**(3): 375-383.

Gorgoni, B. and N. K. Gray (2004). "The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: A developmental perspective." Briefings in Functional Genomics & Proteomics **3**(2): 125-141.

Graber, J. H., C. R. Cantor, et al. (1999). "In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species." Proceedings of the National Academy of Sciences **96**(24): 14055-14060.

Graveley, B. R. and T. Maniatis (1998). "Arginine/Serine-Rich Domains of SR Proteins Can Function as Activators of Pre-mRNA Splicing." Molecular cell **1**(5): 765-771.

Gross, S. and C. Moore (2001). "Five subunits are required for reconstitution of the cleavage and polyadenylation activities of Saccharomyces cerevisiae cleavage factor I." Proceedings of the National Academy of Sciences **98**(11): 6080-6085.

Hake, L. E., R. Mendez, et al. (1998). "Specificity of RNA Binding by CPEB: Requirement for RNA Recognition Motifs and a Novel Zinc Finger." Molecular and Cellular Biology **18**(2): 685-693.

Hake, L. E. and J. D. Richter (1994). "CPEB is a specificity factor that mediates cytoplasmic polyadenylation during Xenopus oocyte maturation." Cell **79**(4): 617-627.

Hammell, C. M., S. Gross, et al. (2002). "Coupling of Termination, 3' Processing, and mRNA Export." Molecular and Cellular Biology **22**(18): 6441-6457.

Hashimoto, K. and J. r. Kudla (2011). "Calcium decoding mechanisms in plants." Biochimie **93**(12): 2054-2059.

Hatton, L. S., J. J. Eloranta, et al. (2002). "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." Nucleic Acids Research **28**(2): 520–526.

Henderson, I. R., F. Liu, et al. (2005). "An allelic series reveals essential roles for FY in plant development in addition to flowering-time control." Development **132**(16): 3597-3607.

Herr, A. J., A. Molnar, et al. (2006). "Defective RNA processing enhances RNA silencing and influences flowering of Arabidopsis." Proceedings of the National Academy of Sciences **103**(41): 14994-15001.

Hirose, Y. and J. L. Manley (1998). "RNA polymerase II is an essential mRNA polyadenylation factor." Nature **395**(6697): 93-96.

Hirose, Y. and J. L. Manley (2000). "RNA polymerase II and the integration of nuclear events." Genes & Development **14**(12): 1415-1429.

Hofmann, I., M. Schnölzer, et al. (2002). "Symplekin, a Constitutive Protein of Karyo- and Cytoplasmic Particles Involved in mRNA Biogenesis in Xenopus laevis Oocytes." Molecular Biology of the Cell **13**(5): 1665-1676.

Holec, S., H. Lange, et al. (2006). "Relaxed Transcription in Arabidopsis Mitochondria Is Counterbalanced by RNA Stability Control Mediated by Polyadenylation and Polynucleotide Phosphorylase." Molecular and Cellular Biology **26**(7): 2869-2876.

Hunt, A., R. Xu, et al. (2008). "Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling." BMC Genomics **9**: 220.

Hunt, A., R. Xu, et al. (2008). "Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling." BMC Genomics **9**(1): 220.

Hunt, A. G. (2008). Messenger RNA 3′ End Formation in Plants. Nuclear pre-mRNA Processing in Plants. A. S. N. Reddy and M. Golovkin, Springer Berlin Heidelberg. **326:** 151-177.

James, P., J. Halladay, et al. (1996). "Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selection in Yeast." Genetics **144**(4): 1425-1436.

Kalderon, D., B. L. Roberts, et al. (1984). "A short amino acid sequence able to specify nuclear location." Cell **39**(3, Part 2): 499-509.

Kashiwabara, S.-i., J. Noguchi, et al. (2002). "Regulation of Spermatogenesis by Testis-Specific, Cytoplasmic Poly(A) Polymerase TPAP." Science **298**(5600): 1999-2002.

Kashiwabara, S.-i., T. Zhuang, et al. (2000). "Identification of a Novel Isoform of Poly(A) Polymerase, TPAP, Specifically Present in the Cytoplasm of Spermatogenic Cells." Developmental Biology **228**(1): 106-115.

Kaufmann, I., G. Martin, et al. (2004). "Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase." EMBO Journal **23**(3): 616-626.

Keene, J. D. (2010). "Minireview: Global Regulation and Dynamics of Ribonucleic Acid." Endocrinology **151**(4): 1391–1397.

Keller, W. (1995). "No end yet to messenger-rna 3' processing." Cell **81**(6): 829-832.

Kerwitz, Y., U. Kuhn, et al. (2003). "Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA." EMBO Journal **22**(14): 3705-3714.

Kessler, M. M., A. M. Zhelkovsky, et al. (1995). "Monoclonal antibodies to yeast poly(A) polymerase (PAP) provide evidence for association of PAP with cleavage factor I." Biochemistry **34**(5): 1750-1759.

Kim, H. and Y. Lee (2001). "Interaction of Poly(A) Polymerase with the 25-kDa Subunit of Cleavage Factor I." Biochemical and Biophysical Research Communications **289**(2): 513-518.

Kim, M. C., W. S. Chung, et al. (2009). "Calcium and Calmodulin-Mediated Regulation of Gene Expression in Plants." Molecular Plant **2**(1): 13-21.

Kleiman, F. E. and J. L. Manley (2001). "The BARD1-CstF-50 Interaction Links mRNA 3' End Formation to DNA Damage and Tumor Suppression." Cell **104**(5): 743-753.

Kleiman, F. E. and J. L. Manley (2001). "The BARD1-CstF-50 Interaction Links mRNA 32 End Formation to DNA Damage and Tumor Suppression." Cell **104**(5): 743-753.

Kyriakopoulou, C. B., H. Nordvarg, et al. (2001). "A Novel Nuclear Human Poly(A) Polymerase (PAP), PAP gamma." Journal of Biological Chemistry **276**(36): 33504-33511.

Lange, H., F. M. Sement, et al. (2009). "Polyadenylation-assisted RNA degradation processes in plants." Trends in plant science **14**(9): 497-504.

Lee, J.-H., W. Terzaghi, et al. (2008). "Characterization of Arabidopsis and Rice DWD Proteins and Their Roles as Substrate Receptors for CUL4-RING E3 Ubiquitin Ligases " Plantcell **20**: 152-167.

Lee, Y. J., Y. Lee, et al. (2000). "An intronless gene encoding a poly(A) polymerase is specifically expressed in testis." FEBS Letters **487**(2): 287-292.

Liang, S., X. Wang, et al. (2001). "Mediation of flowering by a calmodulin-dependent protein kinase." Science in China Series C: Life Sciences **44**(5): 506-512.

Liu, F., S. Marquardt, et al. (2010). "Targeted 3' Processing of Antisense Transcripts Triggers Arabidopsis FLC Chromatin Silencing." Science **327**(5961): 94-97.

Liu, F., V. Quesada, et al. (2007). "The Arabidopsis RNA-Binding Protein FCA Requires a Lysine-Specific Demethylase 1 Homolog to Downregulate FLC." Molecular cell **28**(3): 398-407.

Liu, F., V. Quesada, et al. (2007). "The Arabidopsis RNA-Binding Protein FCA Requires a Lysine-Specific Demethylase 1 Homolog to Downregulate FLC." Molecular cell **28**(3): 398-407.

Lutz, C. S. (2008). "Alternative Polyadenylation: A Twist on mRNA 3' End Formation." ACS Chemical Biology **3**(10): 609-617.

MacDonald, C. C., J. Wilusz, et al. (1994). "The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location." Molecular & Cellular Biology **14**(10): 6647-6654.

Mandel, C. R., Y. Bai, et al. (2008). "Protein factors in pre-mRNA 3′-end processing." Cellular and Molecular Life Sciences (CMLS) **65**(7.8): 1099-1122.

Mandel, C. R., S. Kaneko, et al. (2006). "Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease." Nature **444**(7121): 953-956.

Mangus, D., M. Evans, et al. (2003). "Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression." Genome Biology **4**(7): 223.

Manley, J. L. (1995). "Messenger RNA polyadenylylation: A universal modification." Proceedings of the National Academy of Sciences **92**: 1800-1801.

Manley, J. L. and Y. Takagaki (1996). "The End of the Message,Another Link Between Yeast and Mammals." Science **274**(5292): 1481-1482.

Manzano, D., S. Marquardt, et al. (2009). "Altered interactions within FY/AtCPSF complexes required for Arabidopsis FCA-mediated chromatin silencing." Proceedings of the National Academy of Sciences **106**(21): 8772-8777.

Marquardt, S., P. Boss, et al. (2006). "Additional targets of the Arabidopsis autonomous pathway members, FCA and FY." Journal of Experimental Botany **57**(13): 3379-3386.

Martine A, C. (2003). "Global control of gene expression in yeast by the Ccr4-Not complex." Gene **313**(0): 1-16.

McCormack, E., Y.-C. Tsai, et al. (2005). "Handling calcium signaling: Arabidopsis CaMs and CMLs." Trends in plant science **10**(8): 383-389.

McCracken, S., N. Fong, et al. (1997). "The C-terminal domain of RNA polymerase II couples mRNA processing to transcription." Nature **385**(6614): 357-361.

McGrew, L. L., E. Dworkin-Rastl, et al. (1989). "Poly(A) elongation during Xenopus oocyte maturation is required for translational recruitment and is mediated by a short sequence element." Genes & Development **3**(6): 803-815.

Meeks, L. R., B. Addepalli, et al. (2009). "Characterization of Genes Encoding Poly(A) Polymerases in Plants: Evidence for Duplication and Functional Specialization." PLoS one **4**(11): e8082.

Mendez, R. and J. D. Richter (2001). "Translational control by CPEB: a means to the end." Nature Reviews Molecular Cell Biology **2**(7): 521-529.

Meyer, S., C. Temme, et al. (2004). "Messenger RNA turnover in eukaryotes: pathways and enzymes." Critical Reviews in Biochemistry and Molecular Biology **39**(4): 197-216.

Millevoi, S. and S. Vagner (2010). "Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation." Nucleic Acids Res **38**(9): 2757–2774.

Murphy, A. J., F. Kemp, et al. (2009). "Surface Plasmon Resonance Characterization of Calspermin-Calmodulin Binding Kinetics." Analytical Biochemistry **376**: 61-72.

Murthy, K. G. and J. L. Manley (1995). "The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation." Genes & Development **9**(21): 2672-2683.

Nedea, E., X. He, et al. (2003). "Organization and Function of APT, a Subcomplex of the Yeast Cleavage and Polyadenylation Factor Involved in the Formation of mRNA and Small Nucleolar RNA 3' Ends." Journal of Biological Chemistry **278**(35): 33000-33010.

Ohnacker, M., S. M. L. Barabino, et al. (2000). "The WD-repeat protein Pfs2p bridges two essential factors within the yeast pre-mRNA 3 '-end-processing complex." Embo Journal **19**(1): 37-47.

Paillard, L., D. Maniey, et al. (2000). "Identification of a C-rich element as a novel cytoplasmic polyadenylation element in Xenopus embryos." Mechanisms of Development **93**(1-2): 117-125.

Paris, J., H. B. Osborne, et al. (1988). "Changes in the polyadenylation of specific stable RNA during the early development of Xenopus laevis." Gene **72**(1-2): 169-176.

Preker, P. J. and W. Keller (1998). "The HAT helix, a repetitive motif implicated in RNA processing." Trends in Biochemical Sciences **23**(1): 15-16.

Preker, P. J., J. Lingner, et al. (1995). "The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase." Cell **81**(3): 379-389.

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

Proudfoot, N. and G. Brownlee (1976). "3' non-coding region sequences in eukaryotic messenger RNA." Nature **263**: 211-214.

Proudfoot, N. and J. O'Sullivan (2002). "Polyadenylation: A tail of two complexes." Current Biology **12**(24): 855-857.

Proudfoot, N. J., A. Furger, et al. (2002). "Integrating mRNA Processing with Transcription." Cell **108**(4): 501-512.

Quesada, V., R. Macknight, et al. (2003). "Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time." Embo Journal **22**(12): 3142-3152.

Radford, H. E., H. A. Meijer, et al. (2008). "Translational control by cytoplasmic polyadenylation in Xenopus oocytes." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1779**(4): 217-229.

Rao, S., R. Dinkins, et al. (2009). "Distinctive interactions of the Arabidopsis homolog of the 30 kD subunit of the cleavage and polyadenylation specificity factor (AtCPSF30) with other polyadenylation factor subunits." BMC Cell Biology **10**(1): 51.

Rappsilber, J., U. Ryder, et al. (2002). "Large-Scale Proteomic Analysis of the Human Spliceosome." Genome Research **12**(8): 1231-1245.

Reddy, A. S. N., G. S. Ali, et al. (2011). "Coping with Stresses: Roles of Calcium- and Calcium/Calmodulin-Regulated Gene Expression." Plant Cell **23**(6): 2010–2032..

Reddy, V. S., G. S. Ali, et al. (2002). "Genes Encoding Calmodulin-binding Proteins in the Arabidopsis Genome." Journal of Biological Chemistry **277**(12): 9840-9852.

Rosenthal, E. T., T. R. Tansey, et al. (1983). "Sequence-specific adenylations and deadenylations accompany changes in the translation of maternal messenger RNA after fertilization of Spisula oocytes." Journal of Molecular Biology **166**(3): 309-327.

Rosenthal, E. T. and F. H. Wilt (1986). "Patterns of maternal messenger RNA accumulation and adenylation during oogenesis in Urechis caupo." Developmental Biology **117**(1): 55-63.

Ryan, K. (2007). "Pre-mRNA 3' Cleavage is Reversibly Inhibited in Vitro by Cleavage Factor Dephosphorylation." RNA Biology **4**(1): 26-33.

Ryan, K., O. Calvo, et al. (2004). "Evidence that polyadenylation factor CPSF-73 is the mRNA 3,Ä≤ processing endonuclease." RNA **10**(4): 565-573.

Ryan, K., K. G. K. Murthy, et al. (2002). "Requirements of the RNA Polymerase II C-Terminal Domain for Reconstituting Pre-mRNA 3' Cleavage." Molecular & Cellular Biology **22**(6): 1684-1692.

Sadowski, M., B. Dichtl, et al. (2003). "Independent functions of yeast Pcf11p in pre-mRNA 3 ' end processing and in transcription termination." Embo Journal **22**(9): 2167-2177.

Shock, J., K. Fischer, et al. (2007). "Whole-genome analysis of mRNA decay in Plasmodium falciparum reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle." Genome Biology **8**(7): R134.

Simon, R., J. P. Tassan, et al. (1992). "Translational control by poly(A) elongation during Xenopus development: differential repression and enhancement by a novel cytoplasmic polyadenylation element." Genes & Development **6**(12b): 2580-2591.

Simpson, G. G., P. P. Dijkwel, et al. (2003). "FY is an RNA 3 ' end-processing factor that interacts with FCA to control the Arabidopsis floral transition." Cell **113**(6): 777-787.

Simpson, G. G., P. P. Dijkwel, et al. (2003). "FY Is an RNA 32 End-Processing Factor that Interacts with FCA to Control the Arabidopsis Floral Transition." Cell **113**(6): 777-787.

Snedden, W. A. and H. Fromm (1998). "Calmodulin, calmodulin-related proteins and plant responses to the environment." Trends in plant science **3**(8): 299-304.

Stebbins-Boaz, L. E. Hake, et al. (1996). " CPEB controls the cytoplasmic polyadenylation of cyclin, Cdk2 and c-mos mRNAs and is necessary for oocyte maturation in Xenopus. ." EMBO J. **15**(10): 2582-2592.

Takagaki, Y. and J. L. Manley (2000). "Complex Protein Interactions within the Human Polyadenylation Machinery Identify a Novel Component." Molecular & Cellular Biology **20**(5): 1515-1525.

Thuresson, A. C., J. Astrom, et al. (1994). "Multiple forms of poly(A) polymerases in human cells." Proceedings of the National Academy of Sciences **91**(3): 979-983.

Tsuzuki, M., W. Wu, et al. (2006). "A truncated splice variant of human BARD1 that lacks the RING finger and ankyrin repeats." Cancer Letters **233**(1): 108-116.

Vassalli, J. D., J. Huarte, et al. (1989). "Regulated polyadenylation controls mRNA translation during meiotic maturation of mouse oocytes." Genes & Development **3**(12b): 2163-2171.

Wahle, E. (1991). "A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation." Cell **66**(4): 759-768.

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

Walker, J. E., M. Saraste, et al. (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." EMBO Journal **1**(8): 945-951.

Wallace, A. M., B. Dass, et al. (1999). "Two distinct forms of the 64,000 Mr protein of the cleavage stimulation factor are expressed in mouse male germ cells." Proceedings of the National Academy of Sciences **96**(12): 6763-6768.

Wang, S.-W., K. Asakawa, et al. (2005). "Inactivation of the Pre-mRNA Cleavage and Polyadenylation Factor Pfs2 in Fission Yeast Causes Lethal Cell Cycle Defects." Molecular & Cellular Biology **25**(6): 2288-2296.

Xing, D., H. Zhao, et al. (2008). "Arabidopsis PCFS4, a homologue of yeast polyadenylation factor Pcf11p, regulates FCA alternative processing and promotes flowering time." Plant Journal.

Xu, R., X. Ye, et al. (2004). "AtCPSF73-II gene encoding an Arabidopsis homolog of CPSF 73 kDa subunit is critical for early embryo development." Gene **324**(0): 35-45.

Xu, R., H. Zhao, et al. (2006). "The 73 kD Subunit of the cleavage and polyadenylation specificity factor (CPSF) complex affects reproductive development in Arabidopsis." Plant Molecular Biology **61**(4): 799-815.

Yang, T. B. and B. W. Poovaiah (2003). "Calcium/calmodulin-mediated signal network in plants." Trends in plant science **8**(10): 505-512.

Yao, Y., L. Song, et al. (2002). "Cloning and characterization of Arabidopsis homologues of the animal CstF complex that regulates 3' mRNA cleavage and polyadenylation." Journal of Experimental Botany **53**(378): 2277-2278.

Zarudnaya, M. I., I. M. Kolomiets, et al. (2002). "What Nuclease Cleaves Pre-mRNA in the Process of Polyadenylation?" IUBMB Life **54**(1): 27-31.

Zhang, Z., A. Klatt, et al. (2007). "Transcription termination factor Pcf11 limits the processivity of Pol II on an HIV provirus to repress gene expression." Genes & Development **21**(13): 1609-1614.

Zhao, H., D. Xing, et al. (2009). "Unique Features of Plant Cleavage and Polyadenylation Specificity Factor Revealed by Proteomic Studies." Plant Physiology **151**: 1546-1556.

Zhao, J., L. Hyman, et al. (1999). "Formation of mRNA 3' Ends in Eukaryotes: Mechanism, Regulation, and Interrelationships with Other Steps in mRNA Synthesis." Microbiology and Molecular Biology Reviews **63**(2): 405-445.

Zhou, Z., J. Sim, et al. (2002). "Purification and electron microscopic visualization of functional human spliceosomes." Proceedings of the National Academy of Sciences **99**(19): 12203-12207.

Zorio, D. A. R. and D. L. Bentley (2004). "The link between mRNA processing and transcription: communication works both ways." Experimental Cell Research **296**(1): 91-97.

VITA

Name: Lavanya Dampanaboina

Date of Birth: November 30th, 1979

Place of Birth: Bhimavaram, Andhra Pradesh, India

Education:

1996-1999: B.Sc.
Major: Microbiology
Osmania University,
Andhra Pradesh, INDIA.

1999-2001: M.Sc.
Major: Genetics
Osmania University,
Andhra Pradesh, INDIA.

Professional Experience:

2001-2004: Junior Research Fellow
Center for Plant Molecular Biology (CPMB)
Andhra Pradesh, India.

2004 - 2005 Lecturer (Undergraduate level)
Bhavan's New Science College
Osmania University, Hyderabad, India.

Honors and academic activities

- Kentucky Graduate scholarship (KGS) 2003-2008.
- Qualified in GATE (Graduate Aptitude Test in Engineering) India-2001.
- Science fair judge at the 2006 & 2007 Science Fair, Squires Elementary School, Lexington, KY.
- Science fair judge at the 2008 Science Fair, Northern Elementary School, Lexington, KY and Glendover Elementary School, Lexington, KY.
- Served as a Treasurer for Plant and soil science Graduate student association (2009-2011), University of Kentucky.

Research papers

1. Arthur G Hunt, Ruqiang Xu, Balasubrahmanyam Addepalli, Suryadevara Rao, Kevin Forbes, Lisa Meeks, Denghui Xing, Min Mo, Hongwei Zhao, Amrita Bandyopadhyay, **Lavanya Dampanaboina**, Amanda Marion, Carol Von Lanken, and Qingshun Quinn Li (2008) *Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling.* BMC Genomics, 9: 220.

Poster presentations

1. **LavanyaDampanaboina** and Arthur G. Hunt (2011) Characterization of FY, an orthologue of yeast 3' end processing factor Psf2p in plant polyadenylation using tethering assay and 3' RACE. Rustbelt RNA meeting, Crowne Plaza Hotel, Dayton, OH.
2. **LavanyaDampanaboina** and Arthur G. Hunt (2010) Characterization of Pfs2p in 3' end processing using a tethering assay and 3'RACE. Rustbelt RNA meeting, 2010, Cleveland, OH.
3. **LavanyaDampanaboina** and Arthur G. Hunt (2009) Characterization of the plant polyadenylation factor FY, Kentucky Innovation and Entrepreneurship Conference, Louisville, Kentucky.
4. **LavanyaDampanaboina** and Arthur G. Hunt (2008) Characterization of Arabidopsis orthologue of mammalian CstF50 in 3' end processing. Rustbelt RNA meeting, Mt. Sterling, OH.
5. **LavanyaDampanaboina** and Arthur G. Hunt (2006) Role of the 50 KDa subunit of Cleavage stimulation Factor in Polyadenylation in plants. Rustbelt RNA meeting, Mt. Sterling, OH.