

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Role of the RNA-Binding Protein UBP1C in Translational Repression During Hypoxia in *Arabidopsis thaliana*

Permalink

<https://escholarship.org/uc/item/6db908pg>

Author

Sorenson, Reed Stephen

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Role of the RNA-Binding Protein UBP1C in Translational Repression
During Hypoxia in *Arabidopsis thaliana*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

by

Reed Stephen Sorenson

December 2012

Dissertation Committee:

Dr. Julia Bailey-Serres, Chairperson

Dr. Zhenbiao Yang

Dr. Xuemei Chen

Copyright by
Reed Stephen Sorenson
2012

The Dissertation of Reed Stephen Sorenson is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

I thank Julia Bailey-Serres, an excellent mentor. If I had to do it again, I would do it with Julia. She has been patient and supportive, professional and caring – from successfully writing funded grants that pay my salary to making a triple layer stress granule cake. She has allowed me sufficient autonomy to test my ideas and develop as an independent scientist and forced me to focus my time and attention towards achieving specific goals. I recommend her mentorship to any graduate school candidate.

I thank the members of the Bailey-Serres Group that have shared in my successes and frustrations in the lab, and been a continuous source of ideas and camaraderie; Piyada Juntawong, Seung Cho Lee, Takeshi Fukao, Charles Jang, Julian Peña-Castro, Kayla Kaiser, Greg Barding, Ruth Chang, Angelika Mustroph, Cris Branco-Price, Huijun Yang, Erin Brinton, Maureen Hummel, Rejbana Alam, Daniel Chen, Elaine Yeung, Matthew Collin, and Teruko Oosumi.

I thank the people of the institutions that made possible my training and education. At UC, Riverside, I thank the faculty who served on my advisory committees Dr. Xeumei Chen, Dr. Zhenbiao Yang, Dr. Thomas Eulgem, Dr. Carol Lovatt, and Dr. Ryan Julian. I thank the Department of Botany and Plant Sciences for taking a chance on me, and accepting me into the Plant Biology program, and for financial support. I thank the Graduate Division for financial support. I thank the National Science Foundation as the source of funding that provided the majority of my financial support. California State University, Chico was the first stepping stone towards this dissertation. It was there that I began to enjoy doing experiments and deepened my understanding of biology. It was there that I first even considered the idea of a PhD. I thank Dr. Larry Hanne and Dr. Jim Pushnik for their guidance and mentorship.

I thank my family, the source of my happiness. Carly, my wife, has been the absolute best spouse ever and made this achievement possible. I thank my children Levi, Phaedra,

Emery, and Caleb for the sacrifices that they have, for the most part, unknowingly made. I thank my parents, obviously, for giving and sacrificing so much for me and my siblings, and contributing so much to shaping my character. And, I acknowledge the hand of Providence in my life.

Publications

Bailey-Serres J., Sorenson R., Juntawong P. 2009. Getting the message across: cytoplasmic ribonucleoprotein complexes. *Trends in Plant Science*, **14**: 443-53.

ABSTRACT OF THE DISSERTATION

Role of the RNA-Binding Protein UBP1C in Translational Repression
During Hypoxia in *Arabidopsis thaliana*

by

Reed Stephen Sorenson

Doctor of Philosophy, Graduate Program in Plant Biology
University of California, Riverside, December 2012
Dr. Julia Bailey-Serres, Chairperson

Oxygen deficient plants suffer an energy limitation that can suppress growth and development and lead to premature cell death. Underlying adaptation, plants have evolved integrated genetic and cellular programs to respond to low oxygen stress (hypoxia/anoxia). Mechanisms of diverse adaptation share common themes, such as enhanced oxygen-independent energy supply, regulated growth and energy consumption, and prioritization of energy use. Tailored protein synthesis conserves energy by reduction in synthesis of non-stress-relevant proteins (90% of cellular mRNA) involved in normal growth and development, and funnels resources to translation of hypoxia-induced mRNAs encoding proteins that drive carbohydrate mobility and substrate-level phosphorylation. It was hypothesized that the mRNAs that are translationally repressed during hypoxia associate with RNA binding proteins, which protect them or allow their degradation, and facilitate their rapid return of the sequestered mRNAs to translational complexes within minutes of reoxygenation. Stress granules (SGs) and processing bodies (PBs) are two structures of up to 4 μ M in diameter, characterized in non-plant eukaryotic cells as sites of mRNA sequestration and degradation during stress. Conservation of genes encoding orthologs of these proteins in plants suggests a corresponding system for management of cytosolic non-translated mRNA. We initiated a survey study of predicted RNA binding proteins that are orthologous to non-plant SG and PB proteins encoded by *Arabidopsis thaliana*. To determine their roles in normal development,

T-DNA insertion mutant alleles were identified and surveyed for abnormal growth and development. In addition, several were fused to fluorescent proteins to determine their subcellular localization and make observations of the effect of overexpression on normal growth. From this survey, plant *OLIGOURIDYLATE BINDING PROTEIN (UBP)1C* was further characterized for its role during hypoxia stress. UBPA and UBPA-GFP fusion proteins relocated from diffuse to granular cytoplasmic localization upon hypoxia treatment, which was rapidly reversed by reoxygenation, concomitant with restoration of protein synthesis. These granules required completion of translation for their assembly. Mutation of these genes led to altered survival of oxygen deprivation. Global quantitative profiling of the mRNAs associated with UBPA revealed that transcripts with U-rich 3'-untranslated regions are highly UBPA-associated during non-stress conditions. Upon treatment with hypoxia most cellular mRNAs increased association with UBPA with the exception of the mRNAs that were highly induced by hypoxia and associated with polyribosomes under stress and those already highly associated. Upon reoxygenation this association was rapidly reversed. In conclusion, UBPA association with mRNA during hypoxia is linked to the dynamic sequestration and turnover of transcripts observed during hypoxia in the model plant *Arabidopsis thaliana*.

Contents	Page
List of Figures	xiii
List of Tables	xvi
Chapter 1. Molecular controls of plant hypoxia response	
1.1 Abstract	1
1.2 Introduction to the dissertation	1
1.2.1 Tailoring to the low oxygen energy crisis	1
1.2.2 Low oxygen acclimation: time is of the essence	4
1.2.3 Targeted translation of anaerobic polypeptides facilitates metabolic adjustment	6
1.2.4 Early discovery that low oxygen stress limits protein synthesis	10
1.2.5 Clues of mRNA translational selection: ribosomal protein modifications	13
1.2.6 Clues of mRNA translational selection: limited mRNA-polysome association	15
1.2.7 Clues of mRNA translational selection: mRNA sequestration	17
1.3 Statement of Dissertation Objective	20
1.4 References	21
Chapter 2. Processing body and stress granule protein genes: Cellular localization and mutant phenotypes	
2.1 Abstract	29
2.2 Introduction	30
2.2.1 A repertoire of RNA binding proteins orchestrate post-transcriptional gene regulation	30
2.2.1 A repertoire of RNA binding proteins orchestrate post-transcriptional gene regulation	31

2.2.3 Decapping component Dhh1 regulates gene expression in non-plant eukaryotes	33
2.2.4 Plant <i>UBP1</i> , <i>RBP45</i> , and <i>RBP47</i> proteins are related to the animal TIA-1 family	34
2.2.5 G3BP1 aids assembly of SGs	35
2.3 Results	36
2.3.1 Gene expression meta-analysis of selected genes	36
2.3.2 Mutation in <i>Arabidopsis thaliana</i> RNA helicases of the yeast Dhh1 family have developmental defects	37
2.3.3 Distinct phenotypes were observed in plants with mutant <i>UBP1A</i> and <i>UBP1C</i> alleles	39
2.3.4 Cellular localization of fluorescent protein fusions of putative stress granule and processing body proteins	40
2.4 Discussion	41
2.4.1 Variable gene expression implicates putative SG and PB protein in tissue specific roles	42
2.4.2 Dhh1-family RNA helicases of <i>Arabidopsis thaliana</i> play a role in development	42
2.4.3 Developmental abnormalities among other putative stress granule and processing body component gene mutants	44
2.4.4 Putative stress granule and processing body proteins localize to granules in conditions of low oxygen	44
2.5 Materials and Methods	45
2.5.1 Plant growth and genetic material	45
2.5.2 Allele-specific PCR	46

2.5.3 Generation of transgenic plants	47
2.5.4 Confocal microscopy	49
2.6 References	50
Chapter 3. Role of the RNA-Binding Protein UBP1C in Translational Repression	
During Hypoxia in <i>Arabidopsis thaliana</i>	
3.1 Abstract	88
3.2 Introduction	89
3.2.1 Plant <i>UBP1</i> is a homolog of mammalian TIA-1	89
3.2.2 TIA-1 and TIAR proteins regulate post-transcriptional processes	90
3.2.3 TIA1/R proteins regulate alternative splicing, nucleocytoplasmic transport, and mRNA stability	91
3.2.4 Stress granule aggregation and repression of ribosomal protein synthesis	93
3.2.5 <i>UBP1</i> s of plants	94
3.2.6 <i>UBP1</i> s and translation	96
3.3 Results	98
3.3.1 Plant <i>UBP1</i> s are triple RRM-domain proteins that are related to animal TIA1/Rs	98
3.3.2 Regulated expression of <i>UBP1</i> genes in Arabidopsis	99
3.3.3 <i>ubp1c-1</i> seedling phenotypes were recapitulated by artificial miRNA targeted silencing	100
3.3.4 <i>ubp1c-1</i> seedlings display other early seedling phenotypes	103
3.3.5 <i>ubp1c-1</i> morphological phenotypes were not complemented by <i>UBP1C</i> overexpression or genomic constructs	103
3.3.6 Overexpression of UBP1C-GFP causes unique phenotypes	104
3.3.7 <i>ubp1c-1</i> is hypersensitive to hypoxic stress	105

3.3.8 <i>ubp1a</i> mutants confer morphological phenotypes that are distinct from those of <i>ubp1c</i> mutants	106
3.3.9 <i>ubp1a</i> mutants show hyposensitivity to hypoxic stress	107
3.3.10 UBP1C-GFP and UBP1A-GFP reversibly accumulate in cytoplasmic granules in response to hypoxia	107
3.3.11 Polyadenylated mRNA associates with UBP1C-GFP granules	109
3.3.12 UBP1C-bound mRNA is efficiently immunoprecipitated	110
3.3.13 UBP1C co-immunopurifies with mRNAs	112
3.3.14 Comparative analysis of total mRNA, polysomal mRNA and UBP1C-associated mRNA	113
3.3.15 Dynamic regulation of UBP1C association with mRNAs	115
3.3.16 Analysis of cluster mRNA features	118
3.4 Discussion	120
3.4.1 UBP1 paralogs function in multiple aspects of plant physiology	120
3.4.2 Relatives of TIA-1/R in plants	120
3.4.3 UBP1s role during stress	122
3.4.4 UBP1C-associated RNAs	125
3.4.5 Phenotypic effects of <i>ubp1c</i> mutations	128
3.4.6 Working model of UBP1 function and hypoxia-triggered SGs	129
3.5 Materials and Methods	131
3.5.1 Genetic material and growth conditions	131
3.5.2 Generation of transgenic plants	132
3.5.3 Gene alignment and analysis of evolutionary relationships	135
3.5.4 Stress treatments	136
3.5.5 Immunoprecipitation of UBP1C-RNA complexes	137

3.5.6 RNA purification and quantitative real-time PCR	139
3.5.7 Microarray hybridization and data analysis	140
3.5.8 Protein separation and immunoblot detection	143
3.5.9 Detection of starch	143
3.5.10 Confocal microscopy	143
3.5.11 Detection of poly(A) mRNA in situ hybridization with oligo(dT)	144
3.5.12 Chlorophyll measurements	145
3.6 References	145
Chapter 4. Dynamics of polysome levels during hypoxia and reoxygenation	
4.1 Abstract	220
4.2 Introduction	220
4.3 Results	223
4.3.1 Polysome dynamics during hypoxia stress are altered in <i>mpk6-3</i> mutant	223
4.4 Discussion	224
4.5 Materials and Methods	225
4.5.1 Genetic material and growth conditions	225
4.5.2 Polysome profiling	225
4.6 References	226
Chapter 5. General Conclusions	232
5.1 References	235
Appendix A	238

List of Figures	Page
Figure 2.1 RBPs mediate post-transcriptional gene regulation	55
Figure 2.2 Domain architecture of selected RNA binding proteins of <i>Arabidopsis thaliana</i>	57
Figure 2.3 Distinct tissue expression by putative SG/PB protein genes of <i>Arabidopsis thaliana</i>	59
Figure 2.4 Altered phylotaxy of <i>rh6-1</i> and <i>rh8-1</i> alleles	61
Figure 2.5 Ectopic overexpression of RH6-CFP results in a dwarf phenotype	63
Figure 2.6 Ectopic overexpression of DCP5 results in a dwarf phenotype	65
Figure 2.7 <i>rh12-2</i> exhibits distinct developmental phenotypes	67
Figure 2.8 Morphological variations in mutants of <i>UBP1A</i>	69
Figure 2.9 <i>ubp1c-1</i> seedlings exhibit prostrate growth	71
Figure 2.10 Granular localization of four putative processing body proteins	73
Figure 2.11 Hypoxia-triggered subcellular relocalization of DCP2-YFP	75
Figure 2.12 Co-localization of YFP-DCP1 and RH12-CFP	77
Figure 2.13 <i>UBP1C</i> -GFP localizes to the nucleus, cytoplasm, and cytoplasmic granules	79
Figure 2.14 Localization of PAB2-mRFP (Col-0)	81
Figure 3.1 Phylogeny of plant <i>UBP1</i> and related proteins based on conserved RNA recognition motifs	153
Figure 3.2 Conservation and divergence of plant <i>UBP1</i> and RBP45/47 proteins	155
Figure 3.3 <i>UBP1</i> gene family members are generally co-expressed	163
Figure 3.4 Relative cell-type specific expression in seedlings is distinct between <i>Arabidopsis thaliana UBPs</i>	165

Figure 3.5	<i>UBP1</i> , <i>RBP45</i> , <i>RBP47</i> gene family members show distinct tissue expression	167
Figure 3.6	<i>UBP1</i> gene loci and expression in mutant lines	169
Figure 3.7	Diagrammatic representation of TDNA expression constructs used for transformation of <i>UBP1A</i> and <i>UBP1C</i>	171
Figure 3.8	<i>ubp1c-1</i> and <i>ubp1a-3</i> confer unique seedling growth phenotypes	173
Figure 3.9	The <i>ubp1c-1</i> null allele confers pleiotropic phenotypes	175
Figure 3.10	Expression of an artificial microRNA gene targeting <i>UBP1C</i> (<i>amiR-ubp1c</i>) in Col-0 recapitulates <i>ubp1c-1</i> phenotypes with varying severity	177
Figure 3.11	Near constitutive ectopic overexpression of <i>UBP1C</i> -GFP results in altered growth	179
Figure 3.12	<i>ubp1c-1</i> seedlings are hypersensitive to oxygen deprivation	181
Figure 3.13	Altered phylotaxy of <i>ubp1a-2</i> and <i>ubp1a-3</i> are similar	183
Figure 3.14	<i>UBP1A</i> mutant alleles are not sensitive to hypoxia and <i>ubp1a-2</i> rescues the hypersensitivity to hypoxia of <i>ubp1c-1</i>	185
Figure 3.15	Cytoplasmic <i>UBP1C</i> -GFP dynamically and reversibly aggregates into granular structures in response to hypoxia treatment	187
Figure 3.16	<i>UBP1C</i> granule formation is inhibited by cycloheximide in hypoxia and heat	189
Figure 3.17	<i>UBP1A</i> may have unique and redundant function with <i>UBP1C</i>	191
Figure 3.18	At <i>UBP1C</i> -GFP granules contain polyA ⁺ RNA	193
Figure 3.19	Immunoprecipitation of <i>UBP1C</i> mRNA complexes	195
Figure 3.20	Microarray detection of total and <i>UBP1C</i> -associated RNAs is reproducible	197

Figure 3.21	Distributions of microarray probeset signal log intensities from total and IP samples are not similar	199
Figure 3.22	Fraction and treatment comparison mRNAs detected by microarrays	201
Figure 3.23	Transcripts poorly associated with FLAG-UBP1C reversibly increase association during hypoxia	203
Figure 3.24	Heatmap comparisons of immunoprecipitated RNAs	205
Figure 3.25	Heatmap comparisons of non-UBP1C-associated RNAs	207
Figure 3.26	Percent nucleotide composition by cluster	209
Figure 3.27	3'UTR local nucleotide composition maxima	211
Figure 3.28	Working model for UBP1 function in <i>Arabidopsis thaliana</i>	213
Figure 4.1	Polysomal dynamic of reoxygenation after 2 h hypoxia treatment	228
Figure 4.2	<i>mpk6-3</i> loss-of-function mutants have altered polysomal dynamics after 2 h hypoxia treatment	230

List of Tables	Page
Table 2.1 Overview of mRNP proteins and predicted RNA Binding Proteins of <i>Arabidopsis</i>	83
Table 2.2 TDNA mutant lines	84
Table 2.3 Primers used for genotyping TDNA mutant lines	85
Table 2.4 Genomic DNA segments cloned for generation of native, FLAG-tagged gene constructs	86
Table 2.5 List of fluorescent protein-tagged overexpression lines of putative stress granule or processing body proteins	87
Table 3.1 Summary of strategies and data for complementation or recapitulation of <i>ubp1c-1</i> phenotypes	215
Table 3.2 Gene ontology categories enriched in cluster 7	216
Table 3.3 Gene ontology categories enriched in clusters 5, 8, 12, and 17.	217
Table 3.4 Gene ontology categories enriched in clusters 1 and 2	218
Table 3.5 Gene ontology categories enriched in cluster 9	219

Chapter 1

Plant Molecular Regulation of Hypoxia Response

1.1 Abstract

In plant cells, severe oxygen deprivation (hypoxia/anoxia) dramatically curtails the efficient production of ATP via mitochondrial oxidative phosphorylation. Survival is mediated by limiting energy consumption and maximizing the substrate level production of ATP. This entails regulation of gene expression at transcriptional, post-transcriptional and translational levels. Based on studies on *Arabidopsis thaliana*, low oxygen stress promotes an energy conservation program that tailors cytosolic protein synthesis to the proteins required for sufficient anaerobic metabolism to maintain cell viability. For example, mRNAs that encode enzymes associated with energy-efficient sucrose catabolism and anaerobic metabolism are efficiently translated. Despite maintained abundance during low oxygen stress, mRNAs encoding proteins associated with growth, such as components of the ribosome and translational machinery, are poorly translated. This chapter summarizes the robust evidence of translational regulation during hypoxia and proposes the hypothesis that an important component of this process is sequestration of mRNAs from the translational machinery until reoxygenation. Non-organelle cytosolic domains known as stress granules and processing bodies, which in animals contain translationally repressed RNA and RNA binding proteins, are found in plants. Their possible role in mRNA sequestration and regulation of polyribosome activity is discussed.

1.2 Introduction to the dissertation

1.2.1 Tailoring to the low oxygen energy crisis

In plants, insufficient oxygen limits efficient production of ATP by mitochondria. As the terminal electron acceptor in the mitochondrial electron transport chain, oxygen

completes the circuit of energetic electrons, which originate in carbon-carbon bonds of organic molecular fuel, the bulk of which is lipid or carbohydrate. Energy of the electron is harnessed and transferred to phosphoanhydride bonds of adenosine triphosphate (ATP), the cellular energy currency, requiring an electrochemical gradient across the inner mitochondrial membrane driving ATP synthesis by the F1-ATPase. Lack of oxygen for cytochrome *c* oxidase ($K_m = 0.14 - 1.7 \mu\text{M}$; Drew, 1997) breaks this circuit. Consequentially, levels of NADH reductant and semiubiquinone build up, and oxygenic phosphorylation slows. ATP generation from carbohydrate becomes limited to the substrate level (i.e. ATP yielding reactions of glycolysis and the tricarboxylic acid cycle). It has been proposed that plant survival of severe oxygen deprivation (i.e. <1% oxygen) is largely controlled by the management of the cellular energy budget. Tolerance mechanisms enhance the ability to maintain the cellular energy supply and reduce consumption (Bailey-Serres and Voeselek, 2008, 2010).

Protein synthesis consumes a large portion of the plant cell energy currency. This process involves three distinct phases, initiation, elongation and termination (Kawaguchi and Bailey-Serres, 2002). Initiation involves the selection of the mRNA by initiation factors and the scanning of the 5'-untranslated region (UTR) of the mRNA by the pre-initiation complex composed of the 40S ribosomal subunit and initiation factors. Elongation follows recognition of the initiator methionine and the joining of the 60S ribosomal subunit. It involves the peptidyl transferase reaction catalyzed by the ribosome and translocation facilitated by an elongation factor. Termination occurs when the first stop codon in the mRNA enters the peptidyl (P) site of the ribosome and involves a Release factor. The process of translation requires ATP or GTP reactions of initiation and elongation. Initiation of an mRNA requires GTP to form the pre-initiation complex and ATP in the scanning process dependent upon the degree of secondary structure of the 5'-UTR of the transcript. In the elongation phase, a

minimum of four energetic equivalents of ATP is consumed in the addition of each amino acid to the nascent peptide. The average coding sequence (CDS) length of the model plant *A. thaliana*, (~1350 bp or about 450 codons) (Alexandrov et al., 2006). Therefore, greater than 1800 ATP equivalents are required per transcript per round of protein production. This is roughly equivalent to as few as 25 glucose molecules catabolized, when ATP is generated in the presence of oxygen (30-36 mol ATP mol⁻¹ glc yield) or as many as 450 when ATP is generated in the absence of oxygen (estimated 2-4 mol ATP mol⁻¹ glc anaerobic yield).

During periods of growth, there are high demands for synthesis of core components of the cell; ribosomes, other organelles, the plasma membrane and cell wall. After maturation of the cell, energy is invested in production of secondary metabolites and to maintain homeostasis in a dynamic environment. For years, it has been known that oxygen deprivation causes a strong global decrease in translating polyribosomes (polysomes: multiple-ribosomes loaded mRNAs). This was observed for roots of soybean (*Glycine max*) (Lin and Key, 1967), coleoptiles and roots of maize (*Zea mays* L.) (Bailey-Serres and Freeling, 1990), and whole seedlings of *A. thaliana* (Branco-Price et al., 2005). Mustroph et al. (2009) demonstrated that the reduction in protein synthesis was 2-fold greater in the root than the shoot of Arabidopsis seedlings grown on medium containing sucrose. Regulation of energy expended on protein synthesis benefits the stressed plant in two ways: (1) decreased energy consumption for biosynthetic activities allows reallocation to processes required for cell homeostasis, such as maintenance of the proton gradient across the plasma and vacuolar membranes; and (2) restriction of protein synthesis to the newly synthesized transcripts encoding proteins and enzymes important for energy efficient anaerobic metabolism tailors energy investment to survival. This process involves the selective translation of a subset of cellular mRNAs, despite the maintained abundance of mRNAs associated with biosynthetic processes. This strategy prioritizes energy investment in

proteins that optimize cellular metabolism for substrate level phosphorylation. The mRNA discriminatory mechanism of translation initiation during low oxygen has been an outstanding question in the scientific literature for more than thirty years (Sachs et al., 1980).

This chapter reviews the dynamic regulation of protein synthesis by oxygen availability within the context of the fundamental shift in basal metabolism and the regulatory programs that facilitate this shift. In the last section the objectives of this dissertation are outlined.

1.2.2 Low oxygen acclimation: time is of the essence

Coordination of gene regulation in response to changes in oxygen availability is complex. The rate the velocity and severity of oxygen deprivation can limit acclimation at the level of gene regulation or metabolism. The temporal dynamics of low oxygen stress imposed on an organism or organ may depend upon a variety of environmental factors, such as the rate or depth of a flood, temperature or floodwaters, amount of light and carbon dioxide for photosynthesis, presence of other aerobic organisms. Selection pressure for flooding survival differs in riparian environments where rains slowly raise the water table over a season as compared to flood plains where plants experience rapid but transient partial to complete submergence. Indeed antithetical genetic controls of submergence responses in rice and other species have been elucidated (Voeselek et al., 2006; Bailey-Serres and Voeselek, 2008, 2010). These include the low-oxygen escape strategy that enables growth out of a slow flood and the low-oxygen quiescence strategy that enables energy conservation during a short term deep flood. Particularly in the case of low oxygen quiescence strategy, cells of aerial and root tissue must simultaneously balance efficient usage of diminishing oxygen and/or ATP levels as well as prepare for its complete absence and a severe energy crisis.

Plants require energy to transition from aerobic to anaerobic metabolism because transcription and translation of key enzymes for fermentation is required. Because of the requirement for energy expenditure, it can be hypothesized that reductions in oxygen levels may initiate the changes in gene expression that enable anaerobic metabolism prior to the onset of full anoxia. Although the notion that plants respond proactively at the level of respiration to declining levels of oxygen through changes in metabolism is controversial (Armstrong and Beckett, 2011), there is evidence that genes associated with anaerobic metabolism are upregulated under conditions of modest oxygen deprivation (8% O₂) prior to reaching oxygen values below the K_m of cytochrome c oxidase (van Dongen et al., 2009; Licausi et al., 2011). Resolution of this debate might be achieved by measurement of cellular concentrations of oxygen by non-invasive technologies at the same time as monitoring gene expression. Nonetheless, evidence of activation of genes encoding proteins relevant to metabolic adaptation at intermediate external oxygen conditions was demonstrated in older literature. For example, the activation of expression of the two *Alcohol Dehydrogenase (Adh)* genes of maize was observed in cells cultured under 15% to 0% oxygen. Cells incubated under 15% oxygen showed a four-fold induction in *Adh1* and *Adh2* transcripts, with higher levels of accumulation at lower oxygen concentrations (Paul and Ferl, 1991). The pre-treatment of maize seedling roots or root tips with low oxygen (2-4 kPa) prior to anoxia dramatically enhanced glucose consumption, maintenance of ATP charge, ethanol biosynthesis, and cellular survival (Saglio et al., 1988; Johnson et al., 1989, 1994; Hole et al., 1992). The enhanced metabolic acclimation was partially explained by higher accumulation of transcripts encoding ADH and other key enzymes of anaerobic metabolism (Andrews et al., 1994). Of the two maize genes encoding this enzyme, *Adh1* was constitutively expressed at low levels in the root tip and strongly inducible by low oxygen stress, whereas *Adh2* transcript accumulation was only induced by low oxygen (Andrews et al., 1994).

The importance of ADH for anaerobic acclimation was demonstrated in root tips using an *adh1* loss-of-function mutant (*adh1*⁻) (Andrews et al., 1994). In the *adh1*⁻ mutant, the benefit of hypoxia pretreatment was magnified. Survival of severe, abrupt anoxia, in which ADH activity was not induced in the mutant or wildtype, was reduced to less than 12 h, compared survival of 24 h by *Adh1*⁺ root tips (Roberts et al., 1984). In the mutant, pretreatment induced *Adh2*, resulting in an increase in ADH specific activity to about half the level in untreated *Adh*⁺ root tips. This was sufficient for *Adh1*⁻ mutant viability, demonstrating that pre-induction of modest levels of ADH are necessary and sufficient for low oxygen survival of maize seedling roots (Johnson et al., 1994). Of note, hypoxia pretreatment also enhanced viability of *Adh1*⁺ seedlings, suggesting that other genes induced or proteins synthesized during hypoxia pre-treatment facilitate acclimation and survival of anoxia (Andrews et al., 1994).

1.2.3 Targeted translation of anaerobic polypeptides facilitates metabolic adjustment

By monitoring *in vivo* protein synthesis and polysome levels in maize roots deprived of oxygen it was recognized that altered protein synthesis contributed to a specific reallocation of cellular resources aimed at stress acclimation. Sachs et al. (1980) identified a group of ~ 20 soluble proteins that were synthesized in maize root tips during oxygen deprivation. These were designated the anaerobic polypeptides (ANPs) of maize. In this study, roots of seedlings incubated in buffer sparged with N₂(g) or Ar(g) were pulse-labeled with ³H-leucine or ³⁵S-methionine. To resolve the newly synthesized proteins, a cell extract of soluble proteins was prepared and separated by 2-dimensional gel electrophoresis, using native and denaturing dimensions, and the radiolabeled proteins were detected by autoradiography. Only minor changes in total protein accumulation were observed. However, the aerobic roots synthesized *de novo* a significantly more complex pattern of proteins than

the oxygen deprived roots. The 20 ANPs were observed after 6 hours and for at least 70 hours of the stress.

The ANPs produced in the oxygen starved seedling roots were estimated to represent 70% of all protein synthesis (Sachs et al., 1980). The discovery of the ANPs preceded the development of high-throughput technologies for analysis of gene expression and protein identification. Consequentially, a combination of genetics and biochemistry was required to identify these proteins. The major ANPs were recognized as the products of the two *Adh* genes, encoding an enzyme required for regeneration of NAD⁺ from NADH to sustain anaerobic metabolism (Sachs et al., 1980). Other ANPs that were identified included enzymes involved in carbohydrate catabolism, glycolysis and fermentation: sucrose synthase 1 and 2 (SUS1/2) (Bailey-Serres et al., 1988; Manjunath et al., 1998), phosphoglucomutase (Manjunath et al., 1998), glyceraldehyde-3-phosphate dehydrogenase (Russell and Sachs, 1991), fructose-1,6-diphosphate aldolase (Kelley and Tolan, 1986), glucose-6-phosphate isomerase (Kelley and Freeling, 1984a), enolase (Lal et al., 1998), and pyruvate decarboxylase (Laszlo and St. Lawrence, 1983). In many cases the increase in the synthesis of the ANPs was concomitant with elevation of their transcripts (Sachs and Ho, 1986).

Despite the very clear distinctions in *de novo* protein synthesis from non-stressed and anaerobically treated roots, Sachs et al. (Sachs et al., 1980) found that a generally similar spectrum of mRNAs were present under the two conditions by *in vitro* translation of the poly(A)⁺ mRNAs and the evaluation of the proteins produced by 1-dimensional gel electrophoresis. This provided the first hint of a mechanism of selective synthesis of proteins during low oxygen stress. Based on the characterization of the maize ANPs, the translated mRNAs were primarily proteins that contribute to anaerobic metabolism, and hence low-oxygen survival. The evaluation of the low oxygen responses in soybean roots, in a system similar to that used to study maize, also showed translational regulation based on selective

synthesis *in planta* relative to the total cellular mRNAs translatable *in vitro*. In this species, only four major proteins were synthesized under anoxia, one of which was ADH (Sachs et al., 1980; Russell et al., 1990). Later studies on rice coleoptiles (*Oryza sativa* L.) provided evidence that the metabolic adjustments to low oxygen stress tend to maximize the output of substrate level ATP production, or minimize its consumption through use of pyrophosphate or UTP (Huang et al., 2008; Bailey-Serres and Voesecek, 2008).

Key to anaerobic metabolism is the consumption of stored or soluble carbohydrates. In some species, starch catabolism to hydrolysable sugars occurs through the activity of α -amylase. As for the enzymes responsible for anaerobic fermentation, increases in α -amylase activity in oxygen-starved organs require protein synthesis. Intriguingly, the regulation of changes in metabolism in response to low oxygen stress was shown to be distinct between flooding sensitive tubers of *Solanum tuberosum* and rhizomes of the flooding tolerant species, *Acorus calamus*. Although each expressed α -AMYLASE mRNA within their respective storage organs, only *Acorus* displayed a high level of α -AMYLASE mRNA translation and enzymatic activity under anoxia (Arpagaus and Braendle, 2000). By contrast, *S. tuberosum* tubers had little α -AMYLASE mRNA translation and activity. This suggested that selective translation of mRNAs may underlie differential mobilization of stored carbohydrate and metabolic differences in submerged organs of these species.

Sucrose released from storage can be consumed or translocated through the phloem to sink tissue where it is imported and metabolized. The hydrolysis of sucrose into two hexose-phosphates can be accomplished by either of two enzymes, invertase or sucrose synthase. Two sucrose synthase (SUS) genes of maize (*Sh1* and *Sus1*) were recognized as an ANP and shown to be upregulated at the mRNA (Springer et al., 1986) and protein level (Bailey-Serres et al., 1988). Upon oxygen deprivation, multiple species increase levels of sucrose synthase and repress the activity of invertase. Invertase requires two equivalents of

ATP to phosphorylate the two hexoses produced from sucrose. Alternatively, SUS requires ATP, UDP and pyrophosphate but yields UTP, the energetic equivalent of the ATP input. Thus, the net requirement for sucrose catabolism by SUS is pyrophosphate and dinucleotide. Pyrophosphate is the byproduct of many biochemical reactions such as the charging of tRNA by aminoacyl-tRNA synthetases and, therefore, the shift to use of SUS under low oxygen stress is believed to be a mechanism of energy conservation (Zeng et al., 1999). This enzymatic shift occurs in many plants including potato in which the K_m for invertase is 7-15 mM and the K_m for SUS is 40-200 mM (Bologa et al., 2003) indicating the need for regulation of enzyme levels for an effectual shift in bulk activity.

Some low oxygen scenarios promote an increase in carbohydrate flux through glycolysis (the Pasteur Effect). This may occur in order to meet energy demands because the energetic yield from metabolized sugar is reduced by ~18 fold following a shift to fermentation, involving pyruvate catabolism to lactate by lactate dehydrogenase (LDH) or to alcohol by pyruvate decarboxylase (PDC) production of acetaldehyde which is converted to ethanol by alcohol dehydrogenase (ADH) enzymes. Both of these reactions regenerate NAD^+ needed to sustain glycolysis. ADH and PDC were identified as ANPs in root tips of maize (Sachs et al., 1980) followed by glycolytic enzymes glucose phosphate isomerase and fructose-1,6-diphosphate aldolase (Kelley and Freeling, 1984a, 1984b). Pyruvate can also be metabolized to other end-products during low oxygen stress, including Alanine, GABA and succinate (Limami et al., 2008; Rocha et al., 2010; Bailey-Serres et al., 2012). The production of these end-products suggests that additional ATP can be generated and NAD^+ regenerated by conversion of succinyl CoA to succinate from 2-oxoglutarate via reactions of the tricarboxylic acid cycle. This pathway requires glutamine and aspartate inputs. The hypoxic accumulation of *ALANINE AMINOTRANSFERASE (AlaAT)* in several species, including barley (*Hordeum vulgare*), maize (*Zea mays*), soybean (*Glycine max*), and

Medicago truncatula (Ricoult et al., 2006; see Miyashita et al., 2007), led to the discovery in *Arabidopsis* that AlaAT1 activity contributes to the rapid reconversion of accumulated Alanine to pyruvate upon reoxygenation (Miyashita et al., 2007). Hence, the production of Alanine may result in less carbon loss than production of ethanol, which escapes by diffusion out of cells.

1.2.4 Early discovery that low oxygen stress limits protein synthesis

In the early 1960s, there was a basic understanding that protein synthesis involved the ribosome that catalyzed the process, mRNA that determined the sequence of the polypeptide, and tRNA that functioned as an interpreter between the ribosome and mRNA. A major breakthrough in the understanding of this process came when the interaction between the mRNA and ribosomes was deciphered. Multiple ribosome-loaded mRNAs, or polyribosomes (polysomes for short), were discovered as the dominant structure of protein synthesis in a rabbit reticulocyte lysate by Warner et al. (1962, 1963). These stable complexes could be detected when cells were osmotically lysed and cell contents sedimented through a 15-30% sucrose density gradient. The differential centrifugation separated peaks of an optical density typical of nucleic acids (A_{260}), representing one to many ribosomes connected by a single RNA molecule. Incorporation of ^3H -leucine into protein occurred in the more dense complexes, which became the size of the less active, and less dense complex by a short ribonuclease treatment. Electron micrographs of the predominant complexes showed ribosome aggregates with the number of ribosomes corresponding to the sequential order of peak density. Shortly following these studies, conditions regulating polysome abundance were identified, including inhibition of mitochondrial oxidative phosphorylation and anoxia.

The first experiments characterizing plant polysomes from soybean seedling roots contributed to the general understanding of polysome function and included use of low

oxygen stress to study polysome disassociation and association. mRNA, known as DNA-like RNA (D-RNA) at the time, was shown to associate with polysomes (Lin et al., 1966). It was hypothesized that protein synthesis was energy dependent therefore anaerobiosis was used as a treatment to study mRNA-ribosome association. Corroborating this data, experiments with chemical treatments were used to dissect the dependencies and dynamics of translation. Lin and Key (1967) used dinitrophenol (DNP) to deplete proton gradients and thereby ATP synthesis. DNP treatment causes a dramatic depletion in polysome levels within minutes of application to roots. Oxygen deprivation was also known to limit aerobic ATP production. These studies of Lin and Key (1967) revealed a number of key aspects of polysomal protein synthesis: (1) both DNP and anaerobiosis caused polysome dissociation; (2) polysome depletion was not caused by a decrease in mRNA abundance as depleted polysomes recovered even when 90% of RNA synthesis was blocked by the presence of Actinomycin D, an inhibitor of RNA polymerase II; (3) the translational elongation inhibitor cycloheximide prevented polysome:mRNA dissociation, which occurred following completion of polypeptide synthesis, indicating that initiation (the recruitment of mRNA to the ribosome) was likely the control point in protein synthesis; and (4) thirty minutes of re-aeration of oxygen deprived seedling roots was sufficient for recovery of large polysomes. Of note, the oxygen deprivation was applied by bubbling $N_2(g)$ through media in a stationary flask and aeration was achieved by shaking the submerged seedlings in the flask without bubbling $N_2(g)$.

The next steps toward understanding the nature of translational control were made using maize seedling roots. These roots were treated as described by Sachs et al., (1980) and pulse labeled with either ^{35}S -methionine or ^{32}P -inorganic phosphate (Bailey-Serres and Freeling, 1990). Ribosomes were isolated by differential centrifugation and fractionated on one-dimensional SDS-polyacrylamide gels to compare newly synthesized ribosome-associated proteins produced under control or oxygen deprivation conditions. The stress

dramatically reduced the synthesis of most of the ribosomal proteins, with only four proteins of the complex labeled during the stress. In addition, the phosphorylation of a 31 kDa ribosomal protein was dramatically reduced by the stress. This protein was identified as ribosomal protein S6 (RPS6) (Williams et al., 2003). In addition, this work included a quantitative comparison of polysome distribution between free 40S and 60S subunits, free ribosomes (80S), small polysomes (2-5 ribosomes), and large polysomes (≥ 5 ribosomes). The levels of large polysomes decrease following hypoxia, whereas small polysome levels changed little and free subunits dramatically increased. After two hours re-aeration of the seedlings, polysome levels partially recovered (Bailey-Serres and Freeling, 1990).

The finding that few proteins were synthesized and polysome levels declined during oxygen deprivation led to experiments that investigated dynamics of translation of individual mRNAs in roots of maize seedlings. The first systematic study evaluated the expression of two proteins, the low oxygen-induced *Adh1* and the essential mitochondrial adenine nucleotide translocator (ANT) (Fennoy and Bailey-Serres, 1995). *Adh1* was chosen as a representative ANP and *Ant* was chosen as a representative of genes expressed during normal growth. A comparison was made of the regulation of production of *Adh1* and *Ant* at multiple levels of gene expression, including evaluation of run-on nuclear transcription, steady-state transcript abundance, sedimentation density of ribosome complexes associated with the transcripts, and *de novo* synthesis of the two proteins. The results confirmed that hypoxia activates *Adh1* transcription, producing a stable transcript that is highly translated. By contrast, *Ant* was transcribed and its transcript was initially stable, but slowly degraded over 10-24 h of hypoxia. Under control conditions, *Ant* mRNA sedimented with large polysomes, indicating it was actively translated. By contrast, after 12 h of oxygen deprivation, *Adh1* and *Ant* transcripts showed a different distribution across polysome gradient samples, with *Adh1* mRNA predominantly associated with large polysomes and *ANT* mRNA

associated with small polysomes. The reduced association of *Ant* mRNA with polysomes corresponded to reduced *de novo* synthesis of ANT, as measured by pulse-labeling followed by immunoprecipitation with a specific antiserum (Fennoy and Bailey-Serres, 1995). A subsequent study that evaluated the run-on transcription, steady state mRNA levels and polysome association of *Adh1*, *Adh2*, *Aldolase1*, *Enolase1*, *Sh1* and *Sus1*, and an unidentified ANP (all ANPs) along with *Ant*, *Actin*, *Catalase 2*, *eukaryotic initiation factor 4A*, *Phosphogluconate dehydrogenase 1*, *ribosomal protein P2a* and *ubiquitin polyprotein 1* (all housekeeping mRNAs) provided additional evidence of selective translation of ANP mRNAs during low oxygen stress in this system (Fennoy et al., 1998). These directed gene studies further the suggestion from the results of Sachs et al. (1980) that ANP mRNAs were selectively translated during low oxygen stress.

1.2.5 Clues of mRNA translational selection: ribosomal protein modifications

How *ADH1* and its companion mRNAs are selected for translation when thousands of other mRNAs are repressed is still unknown. However, this discrimination must include both a unique characteristic of the mRNAs selected for translation as well as changes in the ribosome and translation factor machinery, which initiate translation and recognize these mRNAs. Incremental progress on this front was made by Bailey-Serres and Dawe (1996). They used a luciferase (LUC) and β -glucuronidase (GUS) reporter systems to gauge sequences that enable preferential translation under low oxygen stress by coupling a series of truncated 5'-UTR and N-terminal coding regions of the maize *Adh1* gene to the GUS reporter. The contribution of the *Adh1* 3'-UTR was also evaluated. mRNAs were synthesized in vitro that included the 5'-m⁷Gppp cap and 3'-poly(A)⁺ tail required for efficient translation (Kawaguchi and Bailey-Serres, 2002). The mRNAs were electroporated into protoplasts generated from cultured maize cells, along with a LUC with a 5'-UTR containing the viral 5'-UTR of tobacco mosaic virus that is routinely used as a translational enhancer. For the

control condition, protoplasts were gently perfused with 40% oxygen, ensuring far lower ADH enzyme activity than under 20% oxygen. The presence of the 5'-UTR and first 18 codons and 3'-UTRs enhanced reporter activity by 57-fold in transiently transformed protoplasts cultured under 5% oxygen. Removal of only the 3'-UTR reduced the accumulation by one third and sequential truncation from the 3'-end of the 5'-UTR reduced the enhancement in a step-wise manner. By contrast, the LUC control mRNA was translated at 10-fold lower levels under the stress condition. Importantly, the different culture conditions had no significant effect on the stability of the mRNAs tested. These data demonstrate that *Adh1* mRNA sequences are sufficient to enable a chimeric RNA with a reporter gene CDS to circumvent repression of translation.

Alterations in abundance and post-translational modification of ribosomal proteins have also been observed in response to hypoxia and may reflect changes in ribosome activity. A reduction in a group of small, acidic ribosomal proteins was detected among the ribosomal proteins in polysome complexes (Bailey-Serres and Freeling, 1990). Later these were confirmed to be the ribosomal P proteins, including one that is novel to plants (Szick et al., 1998). The P proteins form a stalk that protrudes from the large ribosomal subunit and assemble onto the ribosome in the cytoplasm and not the nucleus, as do other ribosomal proteins. In maize, four proteins make up this acidic stock, P0, a 36 kDa protein, and P1, P2 (a and b), and P3, which are 12 – 15 kDa (Szick et al., 1998). The complex that is formed by these proteins can be salt-washed from polysomes (Bailey-Serres et al., 1997). Gel mobility shift of P1 and P3 were clearly shown in response to hypoxia driving the hypothesis that they are post-translationally modified as a means of polysome regulation during hypoxia (Szick-Miranda and Bailey-Serres, 2001), however the ramification of this heterogeneity in ribosomal protein phosphorylation or composition was not investigated in greater depth because of lack of tractable mutant resources in maize.

In addition to the ribosomal P proteins, another ribosomal protein and several initiation factors are post-translationally modified in response to hypoxia in seedling roots or root tips of maize. Five phosphorylation sites of the carboxyl terminus of ribosomal protein (RPS6) were characterized, along with the demonstration that the protein is dephosphorylated in response to anoxia and heat (Williams et al., 2003). The reduction in phosphorylation was dependent on okadaic acid-inhibited phosphatase activity and the increase in phosphorylation was blocked by LY-294002, a compound that inhibits phosphatidylinositol-3 kinases (Williams et al., 2003). Other studies on maize seedling roots showed that phosphorylation of eukaryotic Initiation Factor (eIF) 4A and Eukaryotic Elongation Factor (eEF) 2 rapidly increased and that of eIF4B decreased in response to low oxygen stress (Webster et al., 1991; Szick-Miranda et al., 2003). Also modified is the m⁷-Gppp cap-binding protein, eIF4E. Two phosphorylated isoforms of eIF4E accumulated following 1 or 6 h hypoxia treatment in seedling roots (Manjunath et al., 1999). While the direct effects of altered translation factor phosphorylation in oxygen starved seedlings are not known, these changes largely coincide with the global decrease in protein synthesis and enhanced selective mRNA translation. An increasing number of examples of ribosome heterogeneity and activity of distinct or modified ribosomes affecting specific aspects of development or environmental responses (Kim et al., 2004, 2007; Rosado et al., 2010; Szakonyi and Byrne, 2011; Chen et al., 2012) elevates the question of whether these modifications are functionally significant, individually and/or collectively, during hypoxia.

1.2.6 Clues of mRNA translational selection: limited mRNA-polysome association

Hypoxia acclimation entails complex changes in gene expression in plants. These have been cataloged in numerous studies globally profiling total cellular mRNA transcripts using microarray hybridization (Branco-Price et al., 2005; Mohanty et al., 2005; Liu et al., 2005; Agarwal and Grover, 2005; Klok et al., 2002; Loreti et al., 2005; Lasanthi-Kudahettige

et al., 2007). Selective translation of stress-relevant gene transcripts was effectively demonstrated on a global scale in *A. thaliana* employing a technology pioneered for this purpose. Polysome immunoprecipitation was developed as a means to compare total cellular mRNA steady state abundance with the that of polysome-associated mRNAs (Zanetti et al., 2005). This method utilizes transgenic Arabidopsis overexpressing HIS₆-FLAG-epitope tagged ribosomal protein L18 (FLAG-RPL18), which is incorporated into fully functional ribosomes. Total and immunopurified polysomal mRNAs were compared by microarray hybridization in non-stressed and hypoxia stressed tissue following 2 or 9 h, as well as 9 h hypoxia plus 1 h recovery (Branco-Price et al., 2008). This study illuminated the gene expression landscape of hypoxic whole seedlings. The results showed trends in mRNA abundance and polysome association that were predicted based on polysome sedimentation experiments and work on ANPs. Polysome association decreased for a majority of genes. Within this group, the total abundance of most of the individual mRNAs was maintained, demonstrating transcript stability for a large pool (65%) of the total cellular content of mRNAs. On the other hand, a smaller pool (21%) of translationally repressed mRNA was unstable and decreased in total abundance during the period of translational repression. Selective translation was further validated by the striking induction and polysome association (presumably undergoing translation) of a relatively large number of mRNAs, including numerous ANP orthologs (106 genes compared to the 20 ANPs of maize root tip or four ANPs of soybean).

The polysome immunopurification was further employed to probe polysomal mRNAs of specific cell-types by limiting expression of FLAG-RPL18 by promoter fusion (Mustroph et al., 2009). This analysis revealed a group of 49 core hypoxia-response genes induced by all cell types. Of these more than 50% were genes encoding proteins of no known biological function, termed HYPOXIA RESPONSIVE UNKNOWN PROTEINS (HUPs). The

investigation of the low oxygen and submergence survival of lines overexpressing a number of these proteins demonstrated that many have a functional role in low oxygen responses (Mustroph et al., 2010; Lee et al., 2011). Of the 49 core response genes, nine corresponded to genes associated with anaerobic metabolism, including *ADH1*, *AlaAT1/2*, *PDC1*, *PDC2*, and *SUS4* (Mustroph et al., 2009, 2010). This finding supports the conclusion that cells are energetically autonomous and predisposed to generate their own energy currency in response to oxygen deprivation.

1.2.7 Clues of mRNA translational selection: mRNA sequestration

Outside of regulation of the initiation, elongation and termination phases of translation by initiation and elongation factors as well as the ribosome, cellular domains referred to stress granules (SGs) and processing bodies (PBs) have been identified as sites of stress-triggered, translationally-repressed mRNA accumulation in both animals and plants. It has been proposed that at these sites, particularly in SGs, mRNAs are actively sequestered during stress manifesting their translational repression (Kedersha et al., 2005; Parker and Sheth, 2007; Anderson and Kedersha, 2008). These mRNA protein complexes (mRNPs) can be visualized by use of fluorescent protein tagging of individual components and range in size from 1-5 μm in diameter. Evidence suggests they are active sites of mRNA sorting, storage, turnover, and possibly silencing (Anderson and Kedersha, 2008; Kedersha et al., 2005). Condition-specific translational repression, whether at the genome or gene level, can result a failure of re-initiation of translation and subsequent sequestration into a SG, which may then facilitate storage or transfer to a PB where mRNA degradation occurs (Anderson and Kedersha, 2008; Decker et al., 2007). SG and PB mRNAs may be heterogeneous, therefore it may be most informative to consider specific proteins and their co-localization and dynamics.

PBs contain the enzymes that facilitate removal of the 5'-m⁷Gppp cap (DCPs) and several 5'→3' exonucleases that digest decapped mRNA, scaffold proteins, and RNA helicases. Mutational analyses indicate that the components of plant PBs are required for proper development and response to dehydration stress. For example, the protein VARICOSE (VSC) of *Arabidopsis*, a partner of DCP1 is required for post-embryonic development (Xu et al., 2006) as are DCP1 and DCP2 (Xu et al., 2006; Goeres et al., 2007; Iwasaki et al., 2007). Homozygous null mutations in genes encoding any of these three proteins lead to an arrest of post-embryonic development (Xu et al., 2006), confirming that they function in indispensable processes. Several confocal laser microscopy studies that utilize fluorescent proteins for visualization have demonstrated co-localization of other proteins in cytoplasmic foci that contain DCP1. Other evidence suggests that DCP1, DCP2 and VCS interact in *bona fide* PBs. *Arabidopsis* DCP2, DCP1 and VCS co-immunoprecipitate, colocalize in cytoplasmic foci, and interact in the yeast two-hybrid system. Also, DCP2 colocalizes in protoplasts with yeast PB protein DExD/H-box helicase 1 (DHH1p) and its plant equivalent RNA helicase12 (RH12) in transient expression assays. An interaction between *Arabidopsis* XRN4, the cytosolic ortholog of human XRN1, and both DCP2 and DCP1 was confirmed by bimolecular fluorescence complementation in tobacco (*Nicotiana plumbaginifolia*) protoplasts (Xu et al., 2006; Goeres et al., 2007; Iwasaki et al., 2007; Bailey-Serres et al., 2009; Xu and Chua, 2011). A compensatory mechanism for loss of decapping was found in the ecotype Landsberg *erecta*, giving it a phenotype of suppression in *vcs* mutants (Zhang et al., 2010). mRNAs that were not degraded in decapping mutants were targeted for degradation by the SOV protein. This ribonuclease was localized to cytoplasmic granules, but coassociation with other PB proteins were not reported (Zhang et al., 2010). It will be interesting to determine if SOV is also a component of PBs.

There is evidence that PBs form in response to heat shock and hypoxia, since these stresses promoted fluorescently tagged DCP1 to accumulate in granular cytoplasmic foci in leaf protoplasts of *Arabidopsis* (Weber et al., 2008). This change in cytoplasmic distribution of DCP1 was not observed if the translational inhibitor cycloheximide was applied along with the stress, indicating that their formation requires completion of protein synthesis. The apparentness of DCP1-containing granules was exacerbated in a mutant with reduced cytosolic 5'→3' exonuclease activity (*xrn4-5*), consistent with the hypothesis that these foci are the site of mRNA turnover.

SGs function in mRNA sequestration, sorting and storage in animals and possibly in plants (Anderson and Kedersha, 2008). In animals, these dynamic granular structures are reportedly composed of mRNAs and proteins including T-cell intracellular antigen 1 (TIA1) and TIA1-related (TIAR), eIF4E, eIF4G, eIF4A, eIF3, PABP, G3BP1, and 40S ribosomal subunits (Anderson and Kedersha, 2008). TIA1 contains a prion-related domain (PRD) that is critical for SG aggregation following stress-triggered eIF2 α phosphorylation, which blocks initiation of translation under conditions of nutrient starvation (Gilks et al., 2004). Besides its role in SG, TIA1 has also been characterized in inhibition of translation, regulation of alternative splicing, mRNA decay (Del Gatto-Konczak et al., 2000; see refs. in Yamasaki et al., 2007) and in nucleocytoplasmic transport (Zhang et al., 2005; Eisinger-Mathason et al., 2008).

In plants, the gene families of oligouridylate-binding protein 1 (UBP1), RNA-binding P protein (RBP)45 and RBP47 of *A. thaliana* share amino acid sequence conservation with animal Tia-1. These proteins have demonstrated influence in pre-mRNA processing and stability. UBP1 was localized to the nucleus and cytoplasm, shown to enhance splicing of low-efficiency U-rich introns and maintain abundance of transcripts driven by specific promoters (Lorković et al., 2000; Lambermon et al., 2000, 2002). UBP1 and RBP47

colocalized with one another, eIF4E, and mRNAs in cytoplasmic granules in transfected protoplasts following heat shock. This recruitment of UBP1 to granules upon heat was dependent on conservation of RRM domain sequences and was decreased by truncation of the glutamine-rich region reminiscent of the Tia-1 PRD (Weber et al., 2008). The similarity of these characteristics with animal SG points to their near conservation across eukaryotes (Anderson and Kedersha, 2008). It is hypothesized that plants cells enduring hypoxia likely employ these structures to manage translationally inactive mRNAs, as observed in mammals (Gottschald et al., 2010). Maintenance of mRNA sequestration from polysomes could add an additional layer of translational control. Plant orthologs of these mammalian RBPs that localize to PB or SG should be targeted for study. The studies could include monitoring the dynamics in cytosolic location of RBPs and their co-location with other proteins and mRNA. By tagging specific RBPs or production of specific antisera, the mRNP complexes could be isolated for evaluation of associated proteins and mRNAs. Finally, mutant analysis might reveal if SG/PB contribute to survival during oxygen deprivation.

1.3 Statement of Dissertation Objective

In this dissertation our objective is to identify putative SG and PB proteins that may function in sequestration of mRNAs during oxygen deprivation.

In chapter 2, we complete a reverse genetic survey of *A. thaliana* genes with similarity to those of non-plants that encode proteins that aggregate in SG and PB. The results indicate that a number of the selected genes localize into cytoplasmic granules in response to hypoxia treatment when overexpressed as a fluorescent protein fusion in stably transformed plants. In addition, homozygous plants with Transfer-DNA mutation alleles presented with developmental phenotypes consistent with a role for these genes in

developmental control. Based on these results we chose a member of the *UBP1* gene family for in depth analyses.

In chapter 3, the role of UBPC1 in response to hypoxia was targeted for study by analysis of mutant alleles, subcellular localization, and global transcript profiling of UBPC1-associated RNAs. The results suggest that *UBP1A* and *UBP1C*, which have high sequence similarity and overlapping transcript accumulation, have unique roles in development. Both proteins dynamically aggregate in microscopic granules resembling SG during hypoxia stress. Furthermore, global transcript analysis of UBPC1-associated mRNAs revealed that during normal growth conditions, UBPC1 associates with mRNAs that have U-rich 3'-UTRs and are enriched in genes targeted for export to the plasma membranes or cell wall. Upon hypoxia, a majority of cellular RNAs reversibly associate with UBPC1, including those that are translationally repressed and stably stored, such as ribosomal protein gene transcripts. Transcripts that are translationally repressed and unstable also show increased UBPC1 association, leading to the suggestion that their routing to PBs is via SGs.

In chapter 4, we evaluate the response of polysome dynamics to low oxygen and a brief recovery phase in wildtype plants and *mpk6-3* mutants. The results demonstrate that during the same period of time that UBPC1 granules are formed and reverse polysome levels inversely decrease and recover. Comparison of wildtype polysome dynamics with those of *mpk6-3* plants supports a role for *MPK6* in regulation of mRNA sequestration.

1.4 References

- Agarwal, S. and Grover, A.** (2005). Isolation and transcription profiling of low-O₂ stress-associated cDNA clones from the flooding-stress-tolerant FR13A rice genotype. *Ann Bot* **96**: 831–844.
- Alexandrov, N., Troukhan, M., Brover, V., Tatarinova, T., Flavell, R., and Feldmann, K.** (2006). Features of Arabidopsis genes and genome discovered using full-length cDNAs. *Plant Mol. Biol.* **60**: 69–85.

- Anderson, P. and Kedersha, N.** (2008). Stress granules: the Tao of RNA triage. *Trends Biochem. Sci.* **33**: 141–150.
- Andrews, D.L., MacAlpine, D.M., Johnson, J.R., Kelley, P.M., Cobb, B.G., and Drew, M.C.** (1994). Differential induction of mRNAs for the glycolytic and ethanolic fermentative pathways by hypoxia and anoxia in maize seedlings. *Plant Physiol* **106**: 1575–1582.
- Armstrong, W. and Beckett, P.M.** (2011). The respiratory down-regulation debate. *New Phytologist* **190**: 276–278.
- Arpagaus, S. and Braendle, R.** (2000). The significance of {alpha}-amylase under anoxia stress in tolerant rhizomes (*Acorus calamus* L.) and non-tolerant tubers (*Solanum tuberosum* L., var. Desiree). *J. Exp. Bot.* **51**: 1475–1477.
- Bailey-Serres, J. and Dawe, R.K.** (1996). Both 5' and 3' sequences of maize adh1 mRNA are required for enhanced translation under low-oxygen conditions. *Plant Physiol.* **112**: 685–695.
- Bailey-Serres, J. and Freeling, M.** (1990). Hypoxic stress-induced changes in ribosomes of maize seedling roots. *Plant Physiol.* **94**: 1237–1243.
- Bailey-Serres, J., Fukao, T., Gibbs, D.J., Holdsworth, M.J., Lee, S.C., Licausi, F., Perata, P., Voeselek, L.A.C.J., and Dongen, J.T. van** (2012). Making sense of low oxygen sensing. *Trends in Plant Science* **17**: 129–138.
- Bailey-Serres, J., Kloeckener-Gruissem, B., and Freeling, M.** (1988). Genetic and molecular approaches to the study of the anaerobic response and tissue specific gene expression in maize. *Plant, Cell & Environment* **11**: 351–357.
- Bailey-Serres, J., Sorenson, R., and Juntawong, P.** (2009). Getting the message across: cytoplasmic ribonucleoprotein complexes. *Trends in Plant Science* **14**: 443–453.
- Bailey-Serres, J., Vangala, S., Szick, K., and Lee, C.** (1997). Acidic Phosphoprotein Complex of the 60S Ribosomal Subunit of Maize Seedling Roots (Components and Changes in Response to Flooding). *Plant Physiol.* **114**: 1293–1305.
- Bailey-Serres, J. and Voeselek, L.A.** (2010). Life in the balance: a signaling network controlling survival of flooding. *Current Opinion in Plant Biology* **13**: 489–494.
- Bailey-Serres, J. and Voeselek, L.A.C.** (2008). Flooding stress: acclimations and genetic diversity. *Annual Review of Plant Biology* **59**: 313–339.
- Bologa, K.L., Fernie, A.R., Leisse, A., Ehlers Loureiro, M., and Geigenberger, P.** (2003). A bypass of sucrose synthase leads to low internal oxygen and impaired metabolic performance in growing potato tubers. *Plant Physiol.* **132**: 2058–2072.
- Branco-Price, C., Kaiser, K.A., Jang, C.J.H., Larive, C.K., and Bailey-Serres, J.** (2008). Selective mRNA translation coordinates energetic and metabolic adjustments to

cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. *The Plant Journal* **56**: 743–755.

- Branco-Price, C., Kawaguchi, R., Ferreira, R., and Bailey-Serres, J.** (2005). Genome-wide analysis of transcript abundance and translation in *Arabidopsis* seedlings subjected to oxygen deprivation. *Ann. Bot. (Lond)* **96**: 647–660.
- Chen, M.-Q., Zhang, A.-H., Zhang, Q., Zhang, B.-C., Nan, J., Li, X., Liu, N., Qu, H., Lu, C.-M., Sudmorgen, Zhou, Y.-H., Xu, Z.-H., and Bai, S.-N.** (2012). *Arabidopsis* NMD3 Is Required for Nuclear Export of 60S Ribosomal Subunits and Affects Secondary Cell Wall Thickening. *PLoS One* **7**.
- Decker, C.J., Teixeira, D., and Parker, R.** (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J. Cell Biol.* **179**: 437–449.
- van Dongen, J.T., Fröhlich, A., Ramírez-Aguilar, S.J., Schauer, N., Fernie, A.R., Erban, A., Kopka, J., Clark, J., Langer, A., and Geigenberger, P.** (2009). Transcript and metabolite profiling of the adaptive response to mild decreases in oxygen concentration in the roots of *Arabidopsis* plants. *Ann Bot* **103**: 269–280.
- Drew, M.C.** (1997). Oxygen deficiency and root metabolism: Injury and acclimation under hypoxia and anoxia. *Annual Review of Plant Physiology and Plant Molecular Biology* **48**: 223–250.
- Eisinger-Mathason, T.S.K., Andrade, J., Groehler, A.L., Clark, D.E., Muratore-Schroeder, T.L., Pasic, L., Smith, J.A., Shabanowitz, J., Hunt, D.F., Macara, I.G., and Lannigan, D.A.** (2008). Codependent functions of RSK2 and the apoptosis-promoting factor TIA-1 in stress granule assembly and cell survival. *Mol. Cell* **31**: 722–736.
- Fennoy, S.L. and Bailey-Serres, J.** (1995). Post-transcriptional regulation of gene expression in oxygen-deprived roots of maize. *The Plant Journal* **7**: 287–295.
- Fennoy, S.L., Nong, T., and Bailey-Serres, J.** (1998). Transcriptional and post-transcriptional processes regulate gene expression in oxygen-deprived roots of maize. *Plant J.* **15**: 727–735.
- Del Gatto-Konczak, F., Bourgeois, C.F., Le Guiner, C., Kister, L., Gesnel, M.-C., Stévenin, J., and Breathnach, R.** (2000). The RNA-binding protein TIA-1 is a novel mammalian splicing regulator acting through intron sequences adjacent to a 5' splice site. *Molecular and Cellular Biology* **20**: 6287–6299.
- Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P.** (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell* **15**: 5383–5398.
- Goeres, D.C., Van Norman, J.M., Zhang, W., Fauver, N.A., Spencer, M.L., and Sieburth, L.E.** (2007). Components of the *Arabidopsis* mRNA decapping complex are required for early seedling development. *Plant Cell* **19**: 1549–1564.

- Gottschald, O.R., Malec, V., Krasteva, G., Hasan, D., Kamlah, F., Herold, S., Rose, F., Seeger, W., and Hänze, J.** (2010). TIAR and TIA-1 mRNA-binding proteins co-aggregate under conditions of rapid oxygen decline and extreme hypoxia and suppress the HIF-1 α pathway. *J Mol Cell Biol* **2**: 345–356.
- Hole, D.J., Cobb, B.G., Hole, P.S., and Drew, M.C.** (1992). Enhancement of Anaerobic Respiration in Root Tips of *Zea mays* following Low-Oxygen (Hypoxic) Acclimation. *Plant Physiol* **99**: 213–218.
- Huang, S., Colmer, T.D., and Millar, A.H.** (2008). Does anoxia tolerance involve altering the energy currency towards PPI? *Trends in Plant Science* **13**: 221–227.
- Iwasaki, S., Takeda, A., Motose, H., and Watanabe, Y.** (2007). Characterization of Arabidopsis decapping proteins AtDCP1 and AtDCP2, which are essential for post-embryonic development. *FEBS Lett.* **581**: 2455–2459.
- Johnson, J., Cobb, B.G., and Drew, M.C.** (1989). Hypoxic Induction of Anoxia Tolerance in Root Tips of *Zea mays*1. *Plant Physiol* **91**: 837–841.
- Johnson, J.R., Cobb, B.G., and Drew, M.C.** (1994). Hypoxic Induction of Anoxia Tolerance in Roots of Adh1 Null *Zea mays* L. *Plant Physiol* **105**: 61–67.
- Kawaguchi, R. and Bailey-Serres, J.** (2002). Regulation of translational initiation in plants. *Curr. Opin. Plant Biol.* **5**: 460–465.
- Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P.** (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* **169**: 871–884.
- Kelley, P. and Freeling, M.** (1984a). Anaerobic expression of maize glucose phosphate isomerase I. *J. Biol. Chem.* **259**: 673–677.
- Kelley, P.M. and Freeling, M.** (1984b). Anaerobic expression of maize fructose-1,6-diphosphate aldolase. *J. Biol. Chem.* **259**: 14180–14183.
- Kelley, P.M. and Tolan, D.R.** (1986). The complete amino acid sequence for the anaerobically induced aldolase from maize derived from cDNA clones. *Plant Physiol* **82**: 1076–1080.
- Kim, B.-H., Cai, X., Vaughn, J.N., and von Arnim, A.G.** (2007). On the functions of the h subunit of eukaryotic initiation factor 3 in late stages of translation initiation. *Genome Biol.* **8**: R60.
- Kim, T.-H., Kim, B.-H., Yahalom, A., Chamovitz, D.A., and von Arnim, A.G.** (2004). Translational regulation via 5' mRNA leader sequences revealed by mutational analysis of the Arabidopsis translation initiation factor subunit eIF3h. *Plant Cell* **16**: 3341–3356.

- Klok, E.J., Wilson, I.W., Wilson, D., Chapman, S.C., Ewing, R.M., Somerville, S.C., Peacock, W.J., Dolferus, R., and Dennis, E.S.** (2002). Expression profile analysis of the low-oxygen response in *Arabidopsis* root cultures. *Plant Cell* **14**: 2481–2494.
- Lal, S.K., Lee, C., and Sachs, M.M.** (1998). Differential regulation of enolase during anaerobiosis in maize. *Plant Physiol* **118**: 1285–1293.
- Lambermon, M.H.L., Fu, Y., Wiczorek Kirk, D.A., Dupasquier, M., Filipowicz, W., and Lorković, Z.J.** (2002). UBA1 and UBA2, two proteins that interact with UBP1, a multifunctional effector of pre-mRNA maturation in plants. *Mol. Cell. Biol.* **22**: 4346–4357.
- Lambermon, M.H.L., Simpson, G.G., Kirk, D.A.W., Hemmings-Mieszczak, M., Klahre, U., and Filipowicz, W.** (2000). UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J.* **19**: 1638–1649.
- Lasanthi-Kudahettige, R., Magneschi, L., Loreti, E., Gonzali, S., Licausi, F., Novi, G., Beretta, O., Vitulli, F., Alpi, A., and Perata, P.** (2007). Transcript profiling of the anoxic rice coleoptile. *Plant Physiol.* **144**: 218–231.
- Laszlo, A. and St. Lawrence, P.** (1983). Parallel induction and synthesis of PDC and ADH in anoxic maize roots. *Molecular and General Genetics MGG* **192**: 110–117.
- Lee, S.C., Mustroph, A., Sasidharan, R., Vashisht, D., Pedersen, O., Oosumi, T., Voesenek, L.A.C., and Bailey - Serres, J.** (2011). Molecular characterization of the submergence response of the *Arabidopsis thaliana* ecotype Columbia. *New Phytol.* **190**: 457–471.
- Licausi, F., Weits, D.A., Pant, B.D., Scheible, W.-R., Geigenberger, P., and van Dongen, J.T.** (2011). Hypoxia responsive gene expression is mediated by various subsets of transcription factors and miRNAs that are determined by the actual oxygen availability. *New Phytologist* **190**: 442–456.
- Limami, A.M., Glévarec, G., Ricoult, C., Cliquet, J.-B., and Planchet, E.** (2008). Concerted modulation of alanine and glutamate metabolism in young *Medicago truncatula* seedlings under hypoxic stress. *J Exp Bot* **59**: 2325–2335.
- Lin, C.Y. and Key, J.L.** (1967). Dissociation and reassembly of polyribosomes in relation to protein synthesis in the soybean root. *J. Mol. Biol.* **26**: 237–247.
- Lin, C.Y., Key, J.L., and Bracker, C.E.** (1966). Association of D-RNA with polyribosomes in the soybean root. *Plant Physiol.* **41**: 976–982.
- Liu, F., VanToai, T., Moy, L.P., Bock, G., Linford, L.D., and Quackenbush, J.** (2005). Global transcription profiling reveals comprehensive insights into hypoxic response in *Arabidopsis*. *Plant Physiol.* **137**: 1115–1129.

- Loreti, E., Poggi, A., Novi, G., Alpi, A., and Perata, P.** (2005). A genome-wide analysis of the effects of sucrose on gene expression in *Arabidopsis* seedlings under anoxia. *Plant Physiol.* **137**: 1130–1138.
- Lorković, Z.J., Wieczorek Kirk, D.A., Klahre, U., Hemmings-Mieszczak, M., and Filipowicz, W.** (2000). RBP45 and RBP47, two oligouridylate-specific hnRNP-like proteins interacting with poly(A)⁺ RNA in nuclei of plant cells. *RNA* **6**: 1610–1624.
- Manjunath, S., Lee, C.-H.K., Winkle, P.V., and Bailey-Serres, J.** (1998). Molecular and biochemical characterization of cytosolic phosphoglucomutase in maize expression during development and in response to oxygen deprivation. *Plant Physiol.* **117**: 997–1006.
- Manjunath, S., Williams, A.J., and Bailey-Serres, J.** (1999). Oxygen deprivation stimulates Ca²⁺-mediated phosphorylation of mRNA cap-binding protein eIF4E in maize roots. *The Plant Journal* **19**: 21–30.
- Miyashita, Y., Dolferus, R., Ismond, K.P., and Good, A.G.** (2007). Alanine aminotransferase catalyses the breakdown of alanine after hypoxia in *Arabidopsis thaliana*. *Plant J.* **49**: 1108–1121.
- Mohanty, B., Krishnan, S.P.T., Swarup, S., and Bajic, V.B.** (2005). Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. *Ann Bot* **96**: 669–681.
- Mustroph, A., Lee, S.C., Oosumi, T., Zanetti, M.E., Yang, H., Ma, K., Masihi, A., Fukao, T., and Bailey-Serres, J.** (2010). Cross-kingdom comparison of transcriptomic adjustments to low oxygen stress highlights conserved and plant-specific responses. *Plant Physiol.*: Provisional acceptance.
- Mustroph, A., Zanetti, M.E., Jang, C.J.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J.** (2009). Profiling translomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A* **106**: 18843–18848.
- Parker, R. and Sheth, U.** (2007). P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**: 635–646.
- Paul, A.-L. and Ferl, R.J.** (1991). Adh1 and Adh2 regulation. *Maydica* **36**: 129–134.
- Ricoult, C., Echeverria, L.O., Cliquet, J.-B., and Limami, A.M.** (2006). Characterization of alanine aminotransferase (AlaAT) multigene family and hypoxic response in young seedlings of the model legume *Medicago truncatula*. *J. Exp. Bot.* **57**: 3079–3089.
- Roberts, J.K.M., Callis, J., Jardetzky, O., Walbot, V., and Freeling, M.** (1984). Cytoplasmic acidosis as a determinant of flooding intolerance in plants. *Proc. Natl. Acad. Sci. U.S.A.* **81**: 6029–6033.
- Rocha, M., Licausi, F., Araújo, W.L., Nunes-Nesi, A., Sodek, L., Fernie, A.R., and Dongen, J.T. van** (2010). Glycolysis and the tricarboxylic acid cycle are linked by

alanine aminotransferase during hypoxia induced by waterlogging of *Lotus japonicus*. *Plant Physiol.* **152**: 1501–1513.

- Rosado, A., Sohn, E.J., Drakakaki, G., Pan, S., Swidergal, A., Xiong, Y., Kang, B.-H., Bressan, R.A., and Raikhel, N.V.** (2010). Auxin-Mediated Ribosomal Biogenesis Regulates Vacuolar Trafficking in *Arabidopsis*. *Plant Cell* **22**: 143–158.
- Russell, D.A. and Sachs, M.M.** (1991). The maize cytosolic glyceraldehyde-3-phosphate dehydrogenase gene family: organ-specific expression and genetic analysis. *Mol. Gen. Genet.* **229**: 219–228.
- Russell, D.A., Wong, D.M.-L., and Sachs, M.M.** (1990). The anaerobic response of soybean. *Plant Physiol* **92**: 401–407.
- Sachs, M.M., Freeling, M., and Okimoto, R.** (1980). The anaerobic proteins of maize. *Cell* **20**: 761–767.
- Sachs, M.M. and Ho, T.H.D.** (1986). Alteration of Gene Expression During Environmental Stress in Plants. *Annual Review of Plant Physiology* **37**: 363–376.
- Saglio, P.H., Drew, M.C., and Pradet, A.** (1988). Metabolic acclimation to anoxia induced by low (2-4 kPa partial pressure) oxygen pretreatment (hypoxia) in root tips of *Zea mays*. *Plant Physiol* **86**: 61–66.
- Springer, B., Werr, W., Starlinger, P., Bennett, D.C., Zokolica, M., and Freeling, M.** (1986). The Shrunken gene on chromosome 9 of *Zea mays* L is expressed in various plant tissues and encodes an anaerobic protein. *Molecular and General Genetics MGG* **205**: 461–468.
- Szakonyi, D. and Byrne, M.E.** (2011). Ribosomal protein L27a is required for growth and patterning in *Arabidopsis thaliana*. *Plant J.* **65**: 269–281.
- Szick, K., Springer, M., and Bailey-Serres, J.** (1998). Evolutionary analyses of the 12-kDa acidic ribosomal P-proteins reveal a distinct protein of higher plant ribosomes. *PNAS* **95**: 2378–2383.
- Szick-Miranda, K. and Bailey-Serres, J.** (2001). Regulated heterogeneity in 12-kDa P-protein phosphorylation and composition of ribosomes in maize (*Zea mays* L.). *J. Biol. Chem.* **276**: 10921–10928.
- Szick-Miranda, K., Jayachandran, S., Tam, A., Werner-Fraczek, J., Williams, A.J., and Bailey-Serres, J.** (2003). Evaluation of Translational Control Mechanisms in Response to Oxygen Deprivation in Maize. *Russian Journal of Plant Physiology* **50**: 774–786.
- Voesenek, L.A.C.J., Colmer, T.D., Pierik, R., Millenaar, F.F., and Peeters, A.J.M.** (2006). How plants cope with complete submergence. *New Phytol.* **170**: 213–226.
- Warner, J.R., Knopf, P.M., and Rich, A.** (1963). A multiple ribosomal structure in protein synthesis. *Proc Natl Acad Sci U S A* **49**: 122–129.

- Warner, J.R., Rich, A., and Hall, C.E.** (1962). Electron microscope studies of ribosomal clusters synthesizing hemoglobin. *Science* **138**: 1399–1403.
- Weber, C., Nover, L., and Fauth, M.** (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.* **56**: 517–530.
- Webster, C., Gaut, R.L., Browning, K.S., Ravel, J.M., and Roberts, J.K.** (1991). Hypoxia enhances phosphorylation of eukaryotic initiation factor 4A in maize root tips. *J. Biol. Chem.* **266**: 23341–23346.
- Williams, A.J., Werner-Fraczek, J., Chang, I.-F., and Bailey-Serres, J.** (2003). Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. *Plant Physiol.* **132**: 2086–2097.
- Xu, J. and Chua, N.H.** (2011). Processing bodies and plant development. *Curr Opin Plant Biol* **14**: 88–93.
- Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H.** (2006). Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell* **18**: 3386–3398.
- Yamasaki, S., Stoecklin, G., Kedersha, N., Simarro, M., and Anderson, P.** (2007). T-cell intracellular antigen-1 (TIA-1)-induced translational silencing promotes the decay of selected mRNAs. *J. Biol. Chem.* **282**: 30070–30077.
- Zanetti, M.E., Chang, I.-F., Gong, F., Galbraith, D.W., and Bailey-Serres, J.** (2005). Immunopurification of polyribosomal complexes of Arabidopsis for global analysis of gene expression. *Plant Physiol.* **138**: 624–635.
- Zeng, Y., Wu, Y., Avigne, W.T., and Koch, K.E.** (1999). Rapid Repression of Maize Invertases by Low Oxygen. Invertase/Sucrose Synthase Balance, Sugar Signaling Potential, and Seedling Survival. *Plant Physiol.* **121**: 599–608.
- Zhang, T., Delestienne, N., Huez, G., Krays, V., and Gueydan, C.** (2005). Identification of the sequence determinants mediating the nucleo-cytoplasmic shuttling of TIAR and TIA-1 RNA-binding proteins. *J. Cell. Sci.* **118**: 5453–5463.
- Zhang, W., Murphy, C., and Sieburth, L.E.** (2010). Conserved RNasell domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. *Proc Natl Acad Sci U S A* **107**: 15981–15985.

Chapter 2

Processing body and stress granule protein genes: Cellular localization and mutant phenotypes

2.1 Abstract

Plant genomes encode a large variety of RNA binding proteins that can be employed in managing RNA biosynthesis, processing, transport, stability, sequestration, and translation. During oxygen deprivation a large shift in mRNA out of polyribosomes can lead to formation of non-translated mRNA:ribonucleoprotein complexes. The nature and composition of these complexes is not well characterized in plants. In non-plant eukaryotes, large cytoplasmic domains with translationally repressed mRNAs have been described in some detail. These are termed stress granules (SG) and processing bodies (PB), being distinguished by the presence of enzymes that mediate 5'm⁷GpppN decapping and endonucleolytic decay in the latter. To the end of further identifying plant RNA binding proteins that may function in plant SG and/or PB, a survey of *Arabidopsis thaliana* RNA binding protein with similarity to animal and yeast SG/PB proteins was completed. Mutant alleles were identified and screened for developmental phenotypes. Florescent protein fusion constructs were overexpressed to identify putative gain-of-function phenotypes and make observations of the effect of submergence on cellular localization. We report a number of developmental defects of mutant lines. Furthermore, we observed cytoplasmic granular localization for a number of putative PB and SG proteins, which was responsive to submergence. Based on these results, the UBP1 family was selected for more detailed study.

2.2 Introduction

2.2.1 A repertoire of RNA binding proteins orchestrate post-transcriptional gene regulation

Gene expression is the continuum from the gene to the mRNA to the functional protein. This involves regulation at the levels of chromatin structure, DNA methylation, initiation of transcription, pre-mRNA capping, splicing, and polyadenylation, followed by a quality-control round of mRNA translation at the nuclear pore, which triggers non-sense mediated decay (NMD) or export to the cytoplasm where localization, translation, storage, and ultimately mRNA degradation occur. Many of these processes involve RNA binding proteins (RBPs) that are often components of mRNA-ribonucleoprotein (mRNP) complexes. The recognition of differentially expressed genes, based on modulation of abundance of individual mRNAs, is a keystone of reverse genetic analyses in plants. But, unfortunately, the isolation of the mRNA from homogenized organs dissociates all RNPs, so transcripts that are stored or targeted for degradation are indistinguishable from those that are undergoing translation. Consequently, steady-state transcript abundance merely provides a snapshot that reflects the balance between mRNA production and turnover. A more precise view of gene regulation requires the quantitative evaluation of mRNAs associated with specific RBPs or mRNP complexes. Additional knowledge of plant RBPs and mRNPs is needed to achieve this goal.

Plant genomes encode hundreds of RBPs that possess one or more motifs that bind RNA (Lorković and Barta, 2002) (*e.g.* RNA recognition motif (RRM; Cléry et al., 2008), K homology domain (KH; Valverde et al., 2008), Pumilio or Puf domain (PUM; Lu et al., 2009), and pentatricopeptide repeat (PPR; Schmitz-Linneweber and Small, 2008)) (Table 2.1). The majority of the ~1100 *Arabidopsis thaliana* proteins with putative RNA binding or modification activity are uncharacterized; many appear to be unique to plants and therefore might participate in plant-specific processes (Lorković and Barta, 2002). Nuclear RBPs function in

the addition of the 5'-m⁷GpppN-cap to the pre-mRNA, the co- and post-transcriptional removal of introns (Lorković and Barta, 2002; Lorković, 2009; Valverde et al., 2008; Reddy, 2007), cleavage and 3' polyadenylation (Hunt et al., 2008), mRNA export (Chekanova et al., 2007), NMD (Pendle et al., 2005), and siRNA and miRNA biogenesis (Voinnet, 2009) (Figure 2.1). Some nuclear RBPs participate in ribosome biogenesis (Pendle et al., 2005), chromatin modification (Pontes and Pikaard, 2008), and telomere maintenance (Kannan et al., 2008). Genetic screens have uncovered a number of nuclear RBPs that are critical to development and stress responses (Lorković, 2009; Chinnusamy et al., 2007). Importantly, some RBPs associate with transcripts in the nucleus and remain bound after export to the cytoplasm. The history of nuclear processing events these RBPs report provides a continuum between nuclear and cytoplasmic processes. In the cytoplasm, RBPs are constituents of a number of mRNPs including the mRNA-40S ribosome pre-initiation complex, ribosome, signal recognition particle (SRP), processing bodies (PBs), stress granules (SG), and mRNA localization machinery (Figure 2.1). Importantly, some cytosolic RBPs bind to cohorts of transcripts via conserved motifs and thereby facilitate the coordinated localization and activity of cadres of mRNAs (Okita and Choi, 2002). The cytoplasmic domains referred to as SGs and PBs have been characterized as sites of non-translated mRNA storage, sorting, and degradation in animal cells. In plant cells, SG or SG-associated proteins have received little attention. Although PB function has not been studied, the decapping complex components, which mark PBs have a demonstrated role in seedling development.

2.2.2 The mRNA decapping complex in plants regulates post-embryonic seedling development.

The main decapping complex components have been identified in *A. thaliana*. *In vitro* evidence of their interaction and biochemical activity are consistent with a conserved functional role in mRNA decapping (Xu et al., 2006). Mutations in genes encoding the three

main decapping complex components, *DCP1*, *DCP2*, and *VARICOSE (VCS)*, confer similar seedling lethal phenotypes: seedling meristem arrest, abnormal cotyledon vein patterning, embryonic traits retained in seedlings, and high abundance of embryonic expressed mRNAs in seedlings (Xu et al., 2006; Goeres et al., 2007; Iwasaki et al., 2007; Xu and Chua, 2011). Another plant decapping component, *DCP5*, was also identified. It accumulates with *DCP1* during seed maturation and peaks 1-2 days after germination and then levels decline to undetectable levels by 6 days following germination. *DCP2* levels increase upon germination and decrease by day 6 (Xu and Chua, 2009). It has been proposed that mRNA accumulation and storage during seed development contributes to seed maturation but following germination the stored mRNAs are targeted for decay. It was further proposed that these steps are regulated by this uncoupled *DCP1/DCP5* and *DCP2* expression during seed maturation and germination with maximal decapping activity when all three are well expressed just after germination. (Xu and Chua, 2009). In *dcp1-1*, *dcp2-1*, *vcs-6*, and *dcp5-1* mutants, higher transcript abundance was maintained in 6-day-old seedlings of numerous seed storage proteins (SSPs), oleosins (OLEOs), and late embryogenesis abundant (LEA) genes. There were large discrepancies in magnitude between mutants, however, which was not well explained (Xu and Chua, 2009). Besides changes in mRNA, protein levels of embryonic expressed genes *OLEO1* and *OLEO2* remained high in these mutant, except *vcs-6*, whereas in wildtype plants these mRNAs and proteins decline following germination (Xu and Chua, 2009).

While the role of decapping in plant development is well established, few studies have evaluated decapping components for stress regulation and PB formation (Chapter 1). There is some evidence that the decapping complex responds to hypoxia and heat by aggregation into granules believed to represent plant PBs (Weber et al., 2008). It was shown that *DCP1* is phosphorylated at serine 237 by MITOGEN ACTIVATED PROTEIN KINASE 6

(MPK6) during dehydration stress (Xu and Chua, 2012). MPK6 is also strongly activated by hypoxia and further activated by reoxygenation (for further discussion see Chapter 4). These data suggest that PB and mRNA decapping might be targets of regulation and contribute to plant stress acclimation.

2.2.3 Decapping component Dhh1 regulates gene expression in non-plant eukaryotes

RH6, *RH8*, and *RH12* are putative *A. thaliana* orthologs of yeast Dhh1. Dhh1 is a decapping activator (Coller et al., 2001) and translational repressor (see citations in Drummond et al., 2011). Characterization of the molecular functions of Dhh1 indicates that it can be localized in polysomes or PBs and plays a role in decapping dependent destabilization of mRNAs.

During a diauxic shift, Dhh1 associates with large polysomes and enters into PBs. Tethering Dhh1 to mRNAs causes a decrease in abundance of the bound transcripts that is dependent on decapping (Carroll et al., 2011). Functional domains for the protein include an ATPase domain that is not required for decapping and RNA degradation (Figure 2.2), but functions in RNA binding (Carroll et al., 2011) (Cheng et al., 2005)(Dutta et al., 2011). There is some evidence that Dhh1 regulates specific mRNAs. In yeast, Dhh1 has function to targeting porin mRNA for decay by a region not conserved in other eukaryotes (Chang and Lee, 2012). *Xenopus* ortholog Xp54 represses specific transcripts during development (cited in Drummond et al., 2011) and stores translationally inactive maternal mRNAs in ribonucleoprotein complexes. The human ortholog of Dhh1, DDX6/Rck/p54, has also been shown to participate in mRNA regulation during maturation of reticulocytes into erythrocytes. In this instance, translational control of gene expression is paramount following nuclear extrusion from the cell. The mRNA of a gene expressed late in reticulocyte differentiation was shown to be associated with DDX6 in a translationally inactive state in PB-like granules until maturation required translation of the mRNA (Naarmann et al., 2010). DDX6 also plays

a tissue-specific role in post-transcriptional regulation during hypoxia in animal cells. Under oxygen replete conditions Hypoxia Inducible Factor (HIF1) α mRNA is translationally repressed and targeted for decapping-mediated decay via an interaction between DDX6 and the HIF1 α IRES. However, during hypoxia *DDX6* mRNA is targeted for decay by hypoxia-induced miR-130. This effects derepression of the transcription factor synthesis and leads to expression of downstream hypoxia-induced genes (Saito et al., 2011).

The family of DDX6 and Dhh1 proteins regulates numerous cellular activities by regulation of protein:mRNA interaction and mRNA translation, storage, and decapping. Ernoul-Lange et al. (2012) proposed a model in which DDX6 is recruited to mRNA via interaction with an mRNA-specific RNA binding protein or by non-specific mRNA binding along the length of an mRNA. This disrupts the 5'-3' bridge made by eIF4E, eIF4G, and PABP. mRNAs are then sequestered to PBs by DDX6 and the decapping complex can be recruited to initiate decay. However, there is no consensus in the model of translational repression. In contrast to the Ernoul-Lange model, Sweet et al. (2012) provide evidence that yeast Dhh1 represses translation by slowing ribosome translocation. Despite the lack of consensus on the mechanism of Dhh1 function, this family of proteins has been identified in the crossroads of mRNA turnover and translation. *A. thaliana RH12* has been shown to colocalize with decapping complex protein DCP2 (Xu et al., 2006), however the details of its specific developmental and/or stress roles remain to be addressed.

2.2.4 Plant UBP1, RBP45, and RBP47 proteins are related to the animal TIA-1 family

SG component proteins have only begun to be identified in plants. To date, these include members of the UBP1 and RBP47 protein families (Weber et al., 2008). The TIA-1 and TIA-related (TIA-1/R) proteins of mammals are well characterized. These RRM proteins (Figure 2.2) play roles in a number of steps of mRNA metabolism including regulation of alternative splicing, mRNA nucleocytoplasmic transport, translational repression, and stress

granule formation. In plant cells three families of genes have similarity to TIA-1/R including UBP1, RBP45, and RBP47 (Figure 2.2). UBP1 genes are the most similar to TIA-1/R (Lorković and Barta, 2002) (Chapter 3). The proteins encoded have been characterized as heteronuclear RNA-binding proteins that bind to U-rich RNA. We identified mutants of UBP1s and determined their cellular localization (this chapter), and associated mutant phenotypes (Chapter 3).

2.2.5 G3BP1 aids assembly of SGs

Ras-GTPase activating protein SH3 domain binding protein (G3BP) is involved in numerous biological processes in eukaryotes. These include myelodysplastic syndrome, cancer progression, development, and neural cell maintenance. At the cellular level, G3BP functions in regulation of protein stability, protein nuclear or cytoplasmic localization, signal transduction, cellular structure, and RNA metabolism. Conserved protein domain structure from N- to C-terminus include the nuclear transport factor 2 domain (NTF2), an acid rich region, a proline rich (PxxP) region, an RNA recognition motif (RRM), and an RGG/glycine rich region (Figure 2.2) (for review see Irvine et al., 2004). Among the numerous activities that have been ascribed to this protein is assembly of SGs (Tourrière et al., 2003). The overexpression of the protein led to SG formation; whereas, overexpression of only the central domain of the protein inhibited arsenite-triggered SGs. G3BP participation in SG formation was regulated by (de)phosphorylation of serine 149. Phosphomimetic S149E failed to oligomerize and form SGs, but S149A, which was nonphosphorylatable could have diffuse cellular distribution or in SG complexes (Tourrière et al., 2003). A number of genes in *A. thaliana* are similar to G3BP and may also participate in SG function.

As introduced in Chapter 1, low oxygen stress causes global decrease in protein synthesis. Many mRNAs that move off of polysomes in response to the stress are stable during transient hypoxia. It was hypothesized that these mRNA may be sequestered in

translationally inactive complexes such as SGs. To begin to address this hypothesis, we sought putative orthologous genes of *A. thaliana* to genes encoding yeast and animal SG and PB components. Our intent was to identify proteins, which localize in stress-induced granules that might associate with non-translating mRNAs during oxygen deprivation. This chapter contains a survey of mutant and overexpression line phenotypes and of fluorescent protein subcellular localization. The group of genes selected includes putative orthologs for decapping complex components: *DCP2*, *DCP1*, *VARICOSE*, *DCP5*, *RH6*, *RH8*, and *RH12*. It also includes putative orthologs encoding SG components: *UBP1A*, *UBP1B*, *UBP1C*, *RBP47A*, *NARR1*, *NARR2*, *NARR3*, *NARR4*, and *PABP2*. Available Transfer-DNA (T-DNA) insertion mutation alleles were identified and their growth observed for developmental defects. Of the putative orthologs, several were selected for cellular localization studies based on overexpression of fluorescent protein fusions or complementation analysis by generation and transformation of genomic gene constructs.

2.3 Results

2.3.1 Gene expression meta-analysis of selected genes

The selected putative SG and PB proteins displayed specific patterns of gene transcript accumulation across developmental stages, organs, tissues and regions based on public transcriptome data. Genevestigator (www.genevestigator.com) was used to generate anatomically specific expression profiles from 6,458 microarrays. Uniquely high levels of transcript accumulation were observed in specific tissues and/or cell types, some of which are critical in development (Figure 2.3).

Of the probsets observed, five had high signal in the sperm cell: *RBP45B*, *RBP47C*, *RH8*, and *DCP1*. Transcripts of *UBP1A*, *RBP45A*, *RBP45*, and *RBP47B* were well expressed in the embryo and suspensor. Consistent with a previous report *DCP5* but not *DCP2* mRNA

was abundant in endosperm and testa along with *UBP1A* and *NARR2*. *NARR2* transcript accumulation was similar in these tissues. A number of other genes monitored had uniquely higher expression in endosperm or testa (Figure 2.3).

In rosette tissue, only *UBP1C* and *RBP45C* had high expression in the hypocotyl and only *RH6* was well expressed in the stem, buds, apex and leaf primordia. *NARR1* and *RH12* were also well expressed in the shoot apex and axillary shoot. However, in rosette leaves, although many of the selected genes had moderate transcript accumulation, none were maximally expressed either in juvenile or adult leaves with the exception that upon leaf senescence *UBP1C* and *DCP1* transcripts increased (Figure 2.3).

The root tip meristematic zone was particularly notable. This region of active division had highly levels of transcripts encoding *RBP45A*, *RBP47A*, *RBP47B*, *RBP47C/C'*, *NARR1*, *NARR3*, *RH6*, *RH8*, *RH12*, *VCS*, and *DCP1*. These transcripts were also generally higher in root tissue. *RH8* was the only gene within the group evaluated that had increased transcript accumulation in lateral roots (Figure 2.3).

Floral tissue had many putative SG and PB protein transcripts moderately expressed, but few that were highly expressed. *RBP45A* transcripts were high in the stamen abscission zone. *NARR4* and *UBP1B* transcripts were abundant in the ovule along with a handful of moderately expressed genes. *RBP45B* and *RBP47C* transcripts were high and all others very low in pollen and anthers. *RH6* has the highest expression in the pedicel and floral stem tissue. *RBP47A* and *NARR1* are high in imbibed seed (Figure 2.3).

2.3.2 Mutation in *Arabidopsis thaliana* RNA helicases of the yeast *Dhh1* family have developmental defects

RNA HELICASE (RH) 6, *RH8*, and *RH12* (At2g45810, At4g00660, and At3g61240, respectively) encode putative RNA helicases. All three have protein sequence similarity to human (*DDX6/RCK*), frog (*Xp54*) and yeast (*Dhh1*) RNA helicases. *RH6* and *RH12* are 90%

identical and RH12 is 75% identical to RH8. All three are 71% identical to human DDX6. DExD/H box helicases include eukaryotic initiation factor 4A (eIF4A) and contain highly conserved sequence domains with an aspartate-glutamate-x/alanine-aspartate/histidine (DExD/H)-box (Figure 2.2). A number of T-DNA mutant alleles were identified in available collections: *rh6-1*, *rh6-2*, *rh8-1*, *rh8-2*, *rh12-1*, *rh12-2*, *rh12-3*, and *rh12-4* (Table 2.4A, 2.6A). Of these, developmental phenotypes were observed in homozygous plants of *rh6-1*, *rh8-1*, and *rh12-2*.

rh6-1 contains a T-DNA insertion in the first exon of *RH6* within the coding sequence (CDS) (Figure 2.4A). *rh6-1* plants exhibited occasional defects in phylotaxy including twin flowers/silques, cauline leaves with axillary bud apical to a flower or silique, and leafy pedicels (Figure 2.4A). *rh6-2*, with a T-DNA insertion in the fourth intron, had no observable phenotypes. When RH6 was fused to cyan fluorescent protein (CFP) and overexpressed under the control of the CaMV 35S promoter, rosettes became dwarfed with round leaf blades and short petioles. Inflorescences were of normal size and morphology but displayed decreased and variable internode length in several independent insertion lines (Figure 2.5). This overexpression phenotype was similar to that produced by overexpression of the DCP5-CFP fusion in Columbia-0 (Figure 2.6)

rh8-1 plants contain a T-DNA insertion within the 5'-UTR of the *RH8* gene. It was previously shown that in this genotype RH8 transcript was greatly reduced (Huang et al., 2010). Phenotypes similar to *rh6-1* were observed in phylotaxy including twin flowers/silques, premature primary meristem termination, accessory flowers to an axillary bud, and cauline leaves apical to flowers (Figure 2.4B). Similar phenotypes were not observed in *rh8-2*, but this insertion was in an intron and therefore may not have disrupted transcript accumulation.

Of the four *rh12* mutants alleles (Figure 2.7A), two exhibited unrelated phenotypes. *rh12-1*, which contains a T-DNA in the 5'-UTR, had a recessive seedling lethal,

morphologically malformed phenotype that did not segregate with the *rh12-1* allele indicating that it was likely due to a second site of T-DNA insertion (Table 2.2). *rh12-2* seedlings were slightly smaller than wildtype plants (Figure 2.7B) and rosette leaves were also slightly small and dark green in appearance (Figure 2.7B). *rh12-1* and *rh12-2* mutants were transformed with an epitope tagged genomic *RH12* construct, which included 585 bp of upstream intergenic sequence as the promoter and the annotated transcript of 2969 bp (Table 2.4). The transgene failed to complement the observed phenotypes. It was subsequently confirmed that the mutant phenotype segregated independently from *rh12-1*. We thereby, generated an *rh12-1 gRH12-FH* plant line for future experimental use. In the *rh12-2* mutant, only two of 72 T1 plants identified by basta selection lacked the dark green leaf phenotypes. These two plants had larger light green leaves in a bushy rosette that was distinct in morphology from Col-0 (Figure 2.7C). These two plants accumulated high levels of anthocyanin causing a dark purple/black coloring of the stem when mature and senescent (not shown).

2.3.3 Distinct phenotypes were observed in plants with mutant *UBP1A* and *UBP1C* alleles

Distinct morphological phenotypes were observed in mutant genotypes of *UBP1A* and *UBP1C*. T-DNA insertion alleles were identified within *UBP1* genes (Table 2.2). Of three alleles of *UBP1A*, *ubp1a-2* and *ubp1a-3* exhibited similar defects in leaf phylotaxy (see Chapter 3). *ubp1a-1* plants, with a T-DNA insertion in the promoter did not have defects in phylotaxy (Figure 2.8). No *UBP1A* insertions were identified within the CDS. A single T-DNA mutant of *UBP1C*, *ubp1c-1*, had an obvious developmental phenotype. The site of insertion was within the CDS of the first exon. *ubp1c-2* and *ubp1c-3* insertion sites are both in the gene upstream sequence ~300 and ~800 bp, respectively. *ubp1c-1* grew prostrate on soil in low light and some seedlings arrested growth early in seedling development. On MS agar

plates with 1% sucrose, seedlings grew similar to wildtype plants but had slightly smaller cotyledons (Figure 2.9).

2.3.4 Cellular localization of fluorescent protein fusions of putative stress granule and processing body proteins

PB proteins of plants were shown to aggregate into granular foci visible by confocal microscopy in response to oxygen deprivation of transfected protoplasts (Weber et al., 2008). In the same study, SG marker proteins aggregated into similar granules in response to heat. To test if any of the targeted proteins showed similar redistribution following seedling submergence, a collection of transgenic lines was made or acquired in which putative SG or PB genes were fused to a fluorescent protein (Table 2.5). The CaMV 35S promoter was used to drive the transgene in all of these genetic constructs. Granule formation was visualized for a number of the proteins evaluated.

Putative markers of PBs were all localized in granular cytoplasmic domains. YFP-RH8 constitutively localized to large cytoplasmic granular domains. However, following submergence, it increased in smaller cytoplasmic granules (Figure 2.10A). A similar pattern of localization was observed for RH6-CFP, except that very large (5 μm) granules were not present in non-stressed cells (Figure 2.10B). RH12-CFP also localized to cytoplasmic granules following submergence (Figure 2.10C). DCP5-CFP did not express well in most tissues, however, it was expressed in root tips where it was localized to cytoplasmic granules (Figure 2.10D). YFP-DCP1 and DCP2-YFP localization were also determined. YFP-DCP1 was constitutively localized to large cytoplasmic granules with little diffuse fluorescence in the cytoplasm and little change following submergence (Figure 2.12). By contrast, DCP2-YFP appeared to be both cytoplasmic and nuclear localized in a diffuse distribution without stress. After 90 min of submergence, intense foci with YFP fluorescence formed (Figure 2.11A, 2.11B). In seedlings treated similarly and desubmerged for 20 min, the intense

punctate appearance reversed to that of diffuse distribution (Figure 2.11A). This phenomenon was quantified by detection of granules on the same field of cells throughout a time course. We found that the number of granules detected increased following extended submergence (70 min) and reversed by 20 min reoxygenation (Figure 2.11C). To further support the hypothesis that these granules represent processing bodies, YFP-DCP1 and RH12-CFP expressing plants were crossed and YFP and CFP localization were compared in F1 lines (Figure 2.12). These two proteins had very good colocalization in cytoplasmic granules.

Putative markers of SGs, UBP1A-GFP and UBP1C-GFP, localized diffusely in the cytoplasm, in cytoplasmic granules, and in the nucleus (Figure 2.13; Chapter 3). Nuclear localization was confirmed by correlated localization of 4',6-diamidino-2-phenylindole (DAPI) staining with UBP1C-GFP. PolyA binding protein (PAB2) mRFP fusion protein did not localize to granules, but localized to lobe tips in epidermal pavement cells on the margins of cotyledons. Older cells in the center of the cotyledon had low expression of the protein, with the exception of the guard cells in which the protein accumulated to higher levels (Figure 2.14).

2.4 Discussion

In this survey of selected putative SG and PB proteins, observations were made of mutant phenotypes (Table 2.2), overexpression phenotypes, and subcellular dynamics of fluorescently tagged proteins. Based on these results, the UBP1 family was selected for further study (Figure 2.9, 2.10, 2.17, Chapter 3). Following the initiation of this survey, one reports was published confirming these results for DCP5 (Xu and Chua, 2009). A handful of intriguing observations are worth further investigation.

2.4.1 Variable gene expression implicates putative SG and PB protein in tissue specific roles

Elevated levels of a gene transcript within a particular tissue or cell type can hint at specific function. Based on the variable tissue expression profiles of our group of selected genes, it seems that a number of them might have roles at specific developmental stages (Figure 2.2). However, moderate expression of genes might also be important in tissues development. Many of these genes were highly expressed during embryo and seed development and in the root meristem. For example, decapping complex genes *DCP1*, *DCP2*, and *VCS* mRNAs were highly expressed in few tissue types, and none accumulated to high levels within flowers, despite the high levels of uncapped mRNAs reported in inflorescence (Jiao et al., 2008). In addition, moderate levels of gene expression may be sufficient for one functional activity, but higher concentration may be required in a second function. Overall, the diversity of tissue-specific expression profiles is consistent with the hypothesis of widespread and specific RNA binding protein regulation of development.

2.4.2 Dhh1-family RNA helicases of *Arabidopsis thaliana* play a role in development

Mutants of *RH6*, *RH8*, and *RH12* were identified in available T-DNA collections. *rh6-1*, *rh8-1*, and *rh12-2*, each with a T-DNA inserted into exonic sequence, exhibited developmental phenotypes (Figure 2.4, 2.7). *rh8-1* has been previously reported to have resistance to potyvirus, but no developmental defects were reported (Huang et al., 2010). Since RHs are proposed to be a component of PBs, disruption of these genes might be anticipated to share phenotypes with *dcp1*, *dcp2*, or *vsc-7* (e.g. seedling lethality, incomplete cotyledon vasculature, embryonic traits). Further experiments will be required to confirm the effect of these insertions on transcript abundance and segregation of phenotype with the mutant allele. It has also been suggested that analysis has been initiated of an *rh12* mutant, named *dhh1-1*, although no data has been provided (Xu and Chua, 2011). The fact that there are three highly similar DDX6/Dhh1 homologs in plants prompts consideration of

possible functional redundancy and/or genetic compensation. We provide evidence that these three proteins all localization into granules that form during hypoxia (Figure 2.10). Dhh1 in yeast has been implicated in multiple steps of RNA processing including relaxing of RNA secondary structure, PB aggregation, and decapping. Single gene disruption has revealed subtle developmental phenotypes (Figure 2.4, 2.7). These phenotypes may be indicative of differences in target gene specificity. Overlap in gene function might be revealed by crossing *RH* mutants to generate the double and triple mutant combinations. If the decapping complex depends on RH function then a phenotype similar to that of the *dcp2*, *dcp1*, and *vcs* would be expected in its absence.

Lack of clear complementation of dark green rosette color in *rh12-2 gRH12-FH* plants might indicate that this phenotype is not caused by the *rh12-2* insertion event (Figure 2.8). This was the case in the phenotype found among *rh12-1* progeny. T-DNA insertion into another gene has not been ruled out. On the other hand, *rh12-2* could lead to a dominant negative truncated gene product that cannot be rescued by the wildtype gene. *rh12-1* might still cause gene disruption and loss-of-function with no accompanying phenotype and requires further investigation.

Overexpression of both RH6-CFP and DCP5-CFP, but not YFP-RH8 or RH12-CFP, caused clear and similar differences in rosette development (Figure 2.5, 2.6). Dwarfed rosettes with round leaf blades and short petioles but normal sized inflorescence with variable node length supports the hypothesis that these proteins function in a similar pathway. This pronounced phenotype was not observed in the overexpression lines of RH8 and RH12, raising the question of whether this represents non-redundant gene function. This hypothesis is consistent with the fact that we did not observe this phenotype in overexpression lines of YFP-DCP1 or DCP2-YFP.

2.4.3 Developmental abnormalities among other putative stress granule and processing body component gene mutants

Of the list of mutant genotypes, obvious developmental defects were observed in *narr1-1* and *rbp47a-1* (Table 2.2). However, the confirmation of causality of genotype to phenotype is required. If validated, the unique role of these RNA binding proteins is implicated in specific aspects of development. Whether or not the proteins participate in RNA sequestration or whether there is high specificity in target mRNA binding that leads to regulation of gene expression will be interesting questions to explore.

2.4.4 Putative stress granule and processing body proteins localize to granules in conditions of low oxygen

Consistent with the conclusion that SG and PB are conserved structures of a eukaryotic cell, we observed the localization of putative marker proteins in cytoplasmic foci with characteristics of these RNA granules. *In planta*, fluorescent RH6, RH8, RH12, and DCP2 localized into cytoplasmic granules when triggered by submergence incubation under a coverslip (Figure 2.10, 2.11). Coverslip incubation limits the oxygen diffusion to two dimensions through water and becomes hypoxic as oxygen is consumed by cells. DCP2 granule formation was reversed by reoxygenation, consistent with the hypothesis that PBs form in response to hypoxia (Figure 2.11A, 2.11C). This work corroborates experiments by Weber et al. (2008). They observed DCP1 granules form in protoplasts within 2 min of coverslip application. We did not see YFP-DCP1 relocalize during hypoxia *in planta* as it was constitutively localized in granules. In this work, RH12 colocalized with DCP1 in constitutive DCP1 granules, which suggests that DCP1 overexpression may drive granule formation (Figure 2.12). This corroborates a previous report that RH12 colocalizes with PB marker DCP2 (Xu et al., 2006). DCP5 also localized to granules (Figure 2.10), consistent with the report showing DCP5 colocalization with DCP1 (Xu and Chua, 2009). Collectively, these

results indicate that the targeted proteins can be stably or transiently localized into granular foci similar to characterized SGs or PBs, where they may function. Future experimentation on these proteins is likely to enhance our understanding of RNA metabolism in plants.

2.5 Materials and Methods

2.5.1 Plant growth and genetic material

A. thaliana ecotype Columbia-0 (Col-0) was used as wildtype. T-DNA insertion lines were obtained from the Arabidopsis Biological Resource Center, Ohio State University, Columbus, OH (SALK or SAIL lines); or from the Institut Jean-Pierre Bourgin, Institut National de la Recherche Agronomique, Versailles, France (FLAG T-DNA lines).

35S:RPL18-GFP was generated previously (Mustroph et al., 2009). ET5803 was a gift from Dr. Rob Martienssen, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. *35S:YFP-DCP1* in Col-0 was a gift from Martin Crespi, Institut des Sciences du Végétal, Centre National de la Recherche Scientifique, Gif-sur-Yvette Cedex, France. Selected gene loci were UBP1A, AT1G54080; UBP1B, AT1G17370; UBP1C, AT3G14100; RBP45C, AT4G27000; RBP45A, AT5G54900; RBP45B, AT1G11650; RBP45, AT5G19350; RBP47A, AT1G49600; RBP47B, AT3G19130; RBP47C/C' AT1G47490, AT1G47500; NARR1, AT5G43960; NARR2, AT3G25150; NARR3, AT5G60980; NARR4, AT5G48650; RH6, AT2G45810; RH8, AT4G00660; RH12, AT3G61240; VCS, AT3G13290; DCP1, AT1G08370; DCP2, AT5G13570; DCP5, AT1G26110. Associated mutant alleles can be seen in Table 2.2.

Seeds were sterilized by incubation in 70% (v/v) ethanol, followed by 1% (v/v) sodium hypochlorite and rinsing; followed by stratification in water at 4°C in the dark for 3-5 days. Plants were grown on vertically oriented plates with 1% (w/v) Phytigel (Sigma, St Louis, MO), 0.43% (w/v) Murashige and Skoog Salts (MS; Caisson Labs, North Ogden, Utah), potassium hydroxide to adjust pH to 7.8, and with or without 1% (w/v) sucrose in a

16h light ($125 \mu\text{E m}^{-2} \text{s}^{-1}$), 8h dark cycle at constant 23°C . Alternatively, following sterilization and stratification plants were sown in Sunshine LC1 soil mix (JM McConkey, Sumner, Washington) with 1.87 g L^{-1} Marathon insecticide (Crop Production Services, Riverside, California) and 1.4 g L^{-1} osmocote 14-14-14 fertilizer, and watered every other day under $\sim 100 \mu\text{E m}^{-2} \text{s}^{-1}$ light. Plants were imaged using a Rebel T1i camera. Images were processed using ImageJ software (<http://rsb.info.nih.gov/ij/>).

2.5.2 Allele-specific PCR

Homozygous plants were identified by allele-specific polymerase chain reaction of genomic DNA and by marker gene selection (Table 2.3). Genotyping by PCR was accomplished using two primer pairs specific for the wildtype or insertion mutant allele. Wildtype-specific primers anneal to 5' and 3' flanking regions of the T-DNA insertion position and generate product only from the wildtype allele. A mutant-allele-specific primer pair includes a primer complementary to the TDNA left border and the respective flanking primer (see Table 2.3). DNA was extracted from leaves or seedlings tissue and was ground in $400 \mu\text{L}$ DNA extraction buffer (100 mM Tris [$\text{pH } 8.0$], 50 mM ethylene-diamine-tetra-acetic acid (EDTA), 500 mM NaCl, 0.2% SDS) and incubated at 50°C for 30 min . To this slurry was added $100 \mu\text{L}$ 24:1 chloroform:isoamyl alcohol, and the mixture was vortexed 30 s . The tubes were centrifuged 2 min at $18,000g$. $200 \mu\text{L}$ of the upper aqueous phase was mixed with $200 \mu\text{L}$ 100% (v/v) isopropanol and incubated 10 min at room temperature. The tubes were again centrifuged at $18,000g$ for 15 min . Pelleted DNA was washed in 70% (v/v) ethanol and dried. DNA was resuspended in $50 \mu\text{L}$ sterile water. PCR reaction mixture ($50 \mu\text{M}$ each primer, 10 mM dNTP, $5 \text{ U } \mu\text{L}^{-1}$ Taq polymerase, 10 mM Tris [$\text{pH } 8.3$], 50 mM KCl, 1.0 mM MgCl_2 , DNA diluted $1 \mu\text{L}$ per $20 \mu\text{L}$ reaction mix) was thermocycled using a standard protocol: 94°C , 3 min ; and 94°C , 30 s ; variable annealing temp, 30 s ; products $> 1.5 \text{ kbp}$ 68°C or products $< 1.5 \text{ kbp}$ 72°C for 1 min per 1 kbp repeated 34 times.

2.5.3 Generation of transgenic plants

Putative orthologous proteins of mammalian stress granule and processing bodies were identified by BLAST search. Gene constructs were made by cloning the *A. thaliana* PCR amplified open reading frame from full-length cDNA prepared from total RNA extracted from one-week old, whole wild type seedlings using the RNeasy Plant Mini Kit (Qiagen, Valencia, California). Total RNA was reverse transcribed by Superscript II reverse transcriptase (Invitrogen, Carlsbad, California) and primed with oligo-deoxy-thymine following the manufacturer's directions. Coding sequence was amplified from cDNA using Hercules II Fusion polymerase (Agilent Technologies, Santa Clara, California) according to the manufacturer's recommendation using the following primer pairs; UBP1C: 5'-CAC CAT GCA GAA TCC GAG ACT GAA GCA ACA TC-3', 5'-CAT GAG TTG CTG TGC GG-3'; UBP1A 5'-CACC ATG CAG AAT CAA AGG CTT ATT AAG CAG CAA CA-3', 5'-CTG ATA GTA CAT GAG CTG CTG ATG GGC-3'; PABP2: 5'-CACC ATG GCG CAG GTT CAA CTT-3', 5'-AGA GAG GTT CAA GGA AGC GA-3'; DCP2: 5'-CACC ATG TCG GGC CTC CAT CGA TCA T-3', 5'-AGC TGA ATT ACC AGA TTC CAA CGC CT-3'; DCP5: 5'-CACC ATG GCG GCT GAT AAT ACG-3', 5'-GGT AGT ACG ATT TGA TAC GCC T-3'; RH6: 5'-CACC ATG AAT AAT AAT AAT AA-3', 5'-CTG ACA GTA GAT TGC CTT GT-3'; RH8: 5'-CACC ATG AAC AAT CGA GGA AGG TAC C-3', 5'-TTA TTG GCA ATA AAT TGC CTG AT-3'; RH12: 5'-CACC ATG AAT ACT AAC AGA GGA AGA T-3', 5'-ATC GAT CAA GCA ATC TAC TGT CAG-3'. For genomic clones, C-terminal FLAG-HIS₆ fusion constructs, gene products were amplified from genomic DNA using the following primers: RH6: 5'-CACC TTT CTC TCT TTC TTT CGG ATG TTA-3', 5'-CTG ACA GTA GAT TGC CTT GT-3'; RH12: 5'-CACC TCA AGG TTT GTT TTG CCA TCA-3', 5'-ATC GAT CAA GCA ATC TAC TGT CAG-3'; DCP2: 5'-CACC AAG TTT GAC GAT TTG CAG GC-3', 5'-AGC TGA ATT ACC AGA TTC CAA CGC CT-3'; VCS: 5'-CACC ACT TTT GAG TTC AGT TGA TCA GTT, 5'-TTT GCA ACC CAT AAG CAT

GG-3'; UBP1C: 5'-CACC CAA AGG CAA AGC TGT CTG TT-3', 5'-CAT GAG TTG CTG TGC GG-3'.

PCR products were separated on 1.0% (w/v) agarose Tris-acetate- EDTA (TAE) gel (40 mM Tris-HCl [pH 8.4], 20 mM sodium acetate, 1 mM EDTA) containing 0.01% (w/v) ethidium bromide by electrophoresis in TAE buffer at 110 V constant voltage for 40-60 min. Following electrophoresis the amplified DNA was cut from the gel and extracted using the QIAquick Gel Extraction Kit (Qiagen) following the prescribed steps. The amplicon was cloned into the pENTR/D-Topo vector (Invitrogen) and transformed into Oneshot Top10 chemically competent cells (Invitrogen) according to the manufacturer's protocols. Plasmid from positive clones was extracted from 5 mL overnight cultures using the Qiagen Spin Miniprep Kit (Qiagen). Sequencing of inserts was performed by the Genomics Core Instrumentation Facility (Gencore) of the Institute for Integrative Genome Biology (IIGB) at the University of California, Riverside. Only sequences matching published, full-length cDNA sequence 100% (www.arabidopsis.org) were used for gene construction. Clones were transferred to target binary vectors (for specific target vectors see Table 2.4, 2.5) using the LR recombinase reaction (Invitrogen) and again transformed into Oneshot Top10 cells for binary vector amplification (Earley et al., 2006; Nakagawa et al., 2007). *gateway-FH-ocsT (bar)* was constructed from the *gateway-FH-ocsT (nptII)* (Mustroph et al., 2009) XhoI/HindIII 9486 bp fragment and the XhoI/HindIII-digested PCR product (1226 bp) containing the *bar* gene amplified from pEarlygate 102 (Earley et al., 2006) using the following primers: bar+HindIII_F, 5'-CCG AAG CTT CTC TGA GAG GAG CAA CTG TG-3'; bar+XhoI_R, 5'-CGA CTC GAG TCG AGC TCG GAT CTG ATA ATT-3'.

Purified plasmid was electroporated into *Agrobacterium tumefaciens* GV3101 cells according to standard procedure. *Agrobacterium*-mediated transformation of *A. thaliana* ecotype Columbia-0 was performed according the floral dip method (Clough and Bent,

1998). Briefly, binary vector-transformed *Agrobacterium tumefaciens* cells were grown to stationary phase in 500 mL Luria broth (1% (w/v) peptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride) with 50 $\mu\text{g mL}^{-1}$ of kanamycin or chloramphenicol, and 50 $\mu\text{g mL}^{-1}$ gentamycin for plasmid selection. Cells were sedimented by centrifugation at 5,000 *g* for 15 min and resuspended in 250 mL 5% (w/v) sucrose and 0.05% (v/v) Silwet-L77® (Crompton, Greenwich, Connecticut, USA). Inflorescences were dipped 15-20 s into *Agrobacterium* and covered for 24h. Seed was harvested 1 month later and transgenic T1 seed was selected on the appropriate antibiotic (50 $\mu\text{g mL}^{-1}$ kanamycin, 25 $\mu\text{g mL}^{-1}$ ammonium glufosinate [basta], or 50 $\mu\text{g mL}^{-1}$ hygromycin).

2.5.4 Confocal microscopy

Plants grown 3-5 d on MS agar with 1% sucrose were used for microscopy experiments. Hypoxia was imposed by submergence in water under a coverslip (24 x 50 mm) on a standard microscope slide (3 x 1 in.) for the specified amount of time. *35S:YFP-RH8-7*, *35S:RH6-CFP-28*, *35S:RH12-CFP-2*, *35S:DCP5-CFP-1*, and *35S:DCP2-YFP-1* transgenic lines were used for microscopy experiments.

FP fusion proteins within plant organs were imaged with a Leica SP2 laser scanning confocal microscope (Wetzlar, Germany) using the 488 nm wavelength Ar laser line at 40% power. Emission wavelengths between 500 and 600 nm were collected as GFP fluorescence. For YFP fluorescence detection, excitation light was provided by a 514 nm laser line at 50% power. 525-600 nm fluorescent light was collected as YFP signal. For CFP detection, a 442 nm laser at 50% power was used for excitation. CFP emission was captured 450-500 nm. Images were processed and analyzed using Imagej software, including granule size and quantity analysis (<http://rsb.info.nih.gov/ij/>).

2.6 References

- Arciga-Reyes, L., Wootton, L., Kieffer, M., and Davies, B.** (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. *Plant J.* **47**: 480–489.
- Bailey-Serres, J.** (1999). Selective translation of cytoplasmic mRNAs in plants. *Trends Plant Sci.* **4**: 142–148.
- Carroll, J.S., Munchel, S.E., and Weis, K.** (2011). The DExD/H box ATPase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics. *J. Cell Biol.* **194**: 527–537.
- Chaikam, V. and Karlson, D.** (2008). Functional characterization of two cold shock domain proteins from *Oryza sativa*. *Plant Cell Environ.* **31**: 995–1006.
- Chang, I.-F., Szick-Miranda, K., Pan, S., and Bailey-Serres, J.** (2005). Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiol.* **137**: 848–862.
- Chang, L.-C. and Lee, F.-J.S.** (2012). The RNA helicase Dhh1p cooperates with Rbp1p to promote porin mRNA decay via its non-conserved C-terminal domain. *Nucl. Acids Res.* **40**: 1331–1344.
- Chekanova, J.A., Gregory, B.D., Reverdatto, S.V., Chen, H., Kumar, R., Hooker, T., Yazaki, J., Li, P., Skiba, N., Peng, Q., Alonso, J., Brukhin, V., Grossniklaus, U., Ecker, J.R., and Belostotsky, D.A.** (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell* **131**: 1340–1353.
- Cheng, Z., Collier, J., Parker, R., and Song, H.** (2005). Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* **11**: 1258–1270.
- Chinnusamy, V., Zhu, J., and Zhu, J.-K.** (2007). Cold stress regulation of gene expression in plants. *Trends Plant Sci.* **12**: 444–451.
- Cléry, A., Blatter, M., and Allain, F.H.-T.** (2008). RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* **18**: 290–298.
- Clough, S.J. and Bent, A.F.** (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**: 735–743.
- Collier, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A., and Parker, R.** (2001). The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* **7**: 1717–1727.
- Drummond, S.P., Hildyard, J., Firczuk, H., Reamtong, O., Li, N., Kannambath, S., Claydon, A.J., Beynon, R.J., Evers, C.E., and McCarthy, J.E.G.** (2011). Diauxic shift-dependent relocalization of decapping activators Dhh1 and Pat1 to polysomal complexes. *Nuc. Acids Res.*

- Dutta, A., Zheng, S., Jain, D., Cameron, C.E., and Reese, J.C.** (2011). Intermolecular interactions within the abundant dead-box protein dhh1 regulate its activity in vivo. *J. Biol. Chem.*
- Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S.** (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**: 616–629.
- Ernoul-Lange, M., Baconnais, S., Harper, M., Minshall, N., Souquere, S., Boudier, T., Bénard, M., Andrey, P., Pierron, G., Kress, M., Standart, N., Cam, E. le, and Weil, D.** (2012). Multiple binding of repressed mRNAs by the P-body protein Rck/p54. *RNA* **18**: 1702–1715.
- Fang, Y. and Spector, D.L.** (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living *Arabidopsis* plants. *Curr. Biol.* **17**: 818–823.
- Giavalisco, P., Wilson, D., Kreitler, T., Lehrach, H., Klose, J., Gobom, J., and Fucini, P.** (2005). High heterogeneity within the ribosomal proteins of the *Arabidopsis thaliana* 80S ribosome. *Plant Mol. Biol.* **57**: 577–591.
- Goeres, D.C., Van Norman, J.M., Zhang, W., Fauver, N.A., Spencer, M.L., and Sieburth, L.E.** (2007). Components of the *Arabidopsis* mRNA decapping complex are required for early seedling development. *Plant Cell* **19**: 1549–1564.
- Huang, T.-S., Wei, T., Laliberté, J.-F., and Wang, A.** (2010). A host RNA helicase-like protein, AtRH8, interacts with the Potyviral Genome-Linked Protein, VPg, associates with the virus accumulation complex, and is essential for infection. *Plant Physiol* **152**: 255–266.
- Hunt, A.G., Xu, R., Addepalli, B., Rao, S., Forbes, K.P., Meeks, L.R., Xing, D., Mo, M., Zhao, H., Bandyopadhyay, A., Dampanaboina, L., Marion, A., Von Lanken, C., and Li, Q.Q.** (2008). *Arabidopsis* mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. *BMC Genom.* **9**: 220.
- Irvine, K., Stirling, R., Hume, D., and Kennedy, D.** (2004). Rasputin, more promiscuous than ever: a review of G3BP. *Int. J. Dev. Biol.* **48**: 1065–1077.
- Iwasaki, S., Takeda, A., Motose, H., and Watanabe, Y.** (2007). Characterization of *Arabidopsis* decapping proteins AtDCP1 and AtDCP2, which are essential for post-embryonic development. *FEBS Lett.* **581**: 2455–2459.
- Jiao, Y., Riechmann, J.L., and Meyerowitz, E.M.** (2008). Transcriptome-wide analysis of uncapped mRNAs in *Arabidopsis* reveals regulation of mRNA degradation. *Plant Cell* **20**: 2571–2585.
- Kannan, K., Nelson, A.D.L., and Shippen, D.E.** (2008). Dyskerin is a component of the *Arabidopsis* telomerase RNP required for telomere maintenance. *Mol. Cell. Biol.* **28**: 2332–2341.

- Kant, P., Kant, S., Gordon, M., Shaked, R., and Barak, S.** (2007). STRESS RESPONSE SUPPRESSOR1 and STRESS RESPONSE SUPPRESSOR2, two DEAD-box RNA helicases that attenuate Arabidopsis responses to multiple abiotic stresses. *Plant Physiol.* **145**: 814–830.
- Kawaguchi, R. and Bailey-Serres, J.** (2002). Regulation of translational initiation in plants. *Curr. Opin. Plant Biol.* **5**: 460–465.
- Kim, B.-H., Cai, X., Vaughn, J.N., and von Arnim, A.G.** (2007). On the functions of the h subunit of eukaryotic initiation factor 3 in late stages of translation initiation. *Genome Biol.* **8**: R60.
- Kim, J.S., Jung, H.J., Lee, H.J., Kim, K.A., Goh, C.-H., Woo, Y., Oh, S.H., Han, Y.S., and Kang, H.** (2008). Glycine-rich RNA-binding protein 7 affects abiotic stress responses by regulating stomata opening and closing in *Arabidopsis thaliana*. *Plant J* **55**: 455–66.
- Kim, Y.-O., Kim, J.S., and Kang, H.** (2005). Cold-inducible zinc finger-containing glycine-rich RNA-binding protein contributes to the enhancement of freezing tolerance in *Arabidopsis thaliana*. *Plant J.* **42**: 890–900.
- Kobayashi, K., Otegui, M.S., Krishnakumar, S., Mindrinos, M., and Zambryski, P.** (2007). INCREASED SIZE EXCLUSION LIMIT 2 encodes a putative DEVH box RNA helicase involved in plasmodesmata function during Arabidopsis embryogenesis. *Plant Cell* **19**: 1885–1897.
- Lambermon, M.H.L., Simpson, G.G., Kirk, D.A.W., Hemmings-Mieszczak, M., Klahre, U., and Filipowicz, W.** (2000). UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J.* **19**: 1638–1649.
- Lorković, Z.J.** (2009). Role of plant RNA-binding proteins in development, stress response and genome organization. *Trends Plant Sci.* **14**: 229–236.
- Lorković, Z.J. and Barta, A.** (2002). Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nuc. Acids Res.* **30**: 623–635.
- Lorković, Z.J., Wieczorek Kirk, D.A., Klahre, U., Hemmings-Mieszczak, M., and Filipowicz, W.** (2000). RBP45 and RBP47, two oligouridylate-specific hnRNP-like proteins interacting with poly(A)⁺ RNA in nuclei of plant cells. *RNA* **6**: 1610–1624.
- Lu, G., Dolgner, S.J., and Hall, T.M.T.** (2009). Understanding and engineering RNA sequence specificity of PUF proteins. *Curr. Opin. Struct. Biol.* **19**: 110–115.
- Ma, L., Xie, B., Hong, Z., Verma, D.P.S., and Zhang, Z.** (2008). A novel RNA-binding protein associated with cell plate formation. *Plant Physiol.* **148**: 223–234.
- Manuell, A.L., Yamaguchi, K., Haynes, P.A., Milligan, R.A., and Mayfield, S.P.** (2005). Composition and structure of the 80S ribosome from the green alga *Chlamydomonas*

reinhardtii: 80S ribosomes are conserved in plants and animals. *J. Mol. Biol.* **351**: 266–279.

- Moore, M.J. and Proudfoot, N.J.** (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**: 688–700.
- Mustroph, A., Zanetti, M.E., Jang, C.J.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J.** (2009). Profiling translomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A* **106**: 18843–18848.
- Naarmann, I.S., Harnisch, C., Müller-Newen, G., Urlaub, H., Ostareck-Lederer, A., and Ostareck, D.H.** (2010). DDX6 recruits translational silenced human reticulocyte 15-lipoxygenase mRNA to RNP granules. *RNA* **16**: 2189–2204.
- Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., Watanabe, Y., Nakamura, K., Kimura, T., and Ishiguro, S.** (2007). Improved gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Bioscience, Biotechnology, and Biochemistry* **71**: 2095–2100.
- Nakaminami, K., Karlson, D.T., and Imai, R.** (2006). Functional conservation of cold shock domains in bacteria and higher plants. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 10122–10127.
- Nishimura, T., Wada, T., Yamamoto, K.T., and Okada, K.** (2005). The *Arabidopsis* STV1 protein, responsible for translation reinitiation, is required for auxin-mediated gynoecium patterning. *Plant Cell* **17**: 2940–2953.
- Okita, T.W. and Choi, S.B.** (2002). mRNA localization in plants: targeting to the cell's cortical region and beyond. *Curr. Opin. Plant Biol.* **5**: 553–559.
- Pendle, A.F., Clark, G.P., Boon, R., Lewandowska, D., Lam, Y.W., Andersen, J., Mann, M., Lamond, A.I., Brown, J.W.S., and Shaw, P.J.** (2005). Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Mol. Biol. Cell* **16**: 260–269.
- Pontes, O. and Pikaard, C.S.** (2008). siRNA and miRNA processing: new functions for Cajal bodies. *Curr. Opin. Genet. Dev.* **18**: 197–203.
- Reddy, A.S.N.** (2007). Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant Biol.* **58**: 267–294.
- Saito, K., Kondo, E., and Matsushita, M.** (2011). MicroRNA 130 family regulates the hypoxia response signal through the P-body protein DDX6. *Nuc. Acids Res.*
- Schmitz-Linneweber, C. and Small, I.** (2008). Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13**: 663–670.

- Streitner, C., Danisman, S., Wehrle, F., Schöning, J.C., Alfano, J.R., and Staiger, D.** (2008). The small glycine-rich RNA binding protein AtGRP7 promotes floral transition in *Arabidopsis thaliana*. *Plant J.* **56**: 239–250.
- Sweet, T., Kovalak, C., and Collier, J.** (2012). The DEAD-Box Protein Dhh1 Promotes Decapping by Slowing Ribosome Movement. *PLoS Biol* **10**.
- Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., and Tazi, J.** (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. *The Journal of Cell Biology* **160**: 823–831.
- Valverde, R., Edwards, L., and Regan, L.** (2008). Structure and function of KH domains. *FEBS J.* **275**: 2712–2726.
- Voinnet, O.** (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* **136**: 669–687.
- Wang, C., Washida, H., Crofts, A.J., Hamada, S., Katsube-Tanaka, T., Kim, D., Choi, S.-B., Modi, M., Singh, S., and Okita, T.W.** (2008). The cytoplasmic-localized, cytoskeletal-associated RNA binding protein *OsTudor-SN*: evidence for an essential role in storage protein RNA transport and localization. *Plant J.* **55**: 443–454.
- Wang, S. and Okamoto, T.** (2009). Involvement of polypyrimidine tract-binding protein (PTB)-related proteins in pollen germination in *Arabidopsis*. *Plant Cell Physiol.* **50**: 179–190.
- Weber, C., Nover, L., and Fauth, M.** (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.* **56**: 517–530.
- Xu, J. and Chua, N.H.** (2011). Processing bodies and plant development. *Curr Opin Plant Biol* **14**: 88–93.
- Xu, J. and Chua, N.-H.** (2009). *Arabidopsis* decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. *Plant Cell* **21**: 3270–3279.
- Xu, J. and Chua, N.-H.** (2012). Dehydration stress activates *Arabidopsis* MPK6 to signal DCP1 phosphorylation. *EMBO J.* **31**: 1975–1984.
- Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H.** (2006). *Arabidopsis* DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell* **18**: 3386–3398.
- Yoine, M., Ohto, M., Onai, K., Mita, S., and Nakamura, K.** (2006). The lba1 mutation of UPF1 RNA helicase involved in nonsense-mediated mRNA decay causes pleiotropic phenotypic changes and altered sugar signaling in *Arabidopsis*. *Plant J.* **47**: 49–62.

Figure 2.1. RBPs mediate post-transcriptional gene regulation. Transcript biogenesis is tightly coupled with pre-mRNA processing events, including 5'-capping, intron splicing and 3'-polyadenylation. RBPs bind to the transcript to form an mRNP. The exon junction (EJC) complex marks the sites of intron removal, whereas other RBPs bind the 5'-cap, the 3'-poly(A) tail and along the mRNA in sequence specific or non-specific manner. Nuclear mRNPs are exported to the cytoplasm and must pass a quality control checkpoint or be subjected to nonsense-mediated decay (NMD). The cytoplasmic routing and activity of mRNAs is determined by sequence features, binding of specific RBPs or characteristics of the nascent peptide, such as a signal peptide. mRNAs and polysomes might be targeted within the cell along microtubules (MTs) or actin filaments (AFs) or to neighboring cells by passage through plasmodesmata. The initiation of translation is a regulated process that results in the formation of a polysome. Regulation of initiation can shift the balance between actively translating polysomes and large mRNP aggregates, such as stress granules and processing bodies, which are associated with mRNA sequestration or degradation, respectively. Signaling initiated by reduced energy status is hypothesized to inhibit the target of rapamycin (TOR) pathway and repress translation. miRNA processing and export involves participation of multiple RNPs. Plant miRNA functions in mRNA slicing and degradation and might also contribute to translational repression. Abbreviations: Ago1, argonaute1; XRN, exoribonuclease; eIF4E, eukaryotic initiation factor 4E; ISE2, increased size exclusion limit 2; nCBPs, nuclear cap-binding proteins; UBP1, oligouridylate-binding protein 1; PABP, poly(A)-binding protein; RBP47, RNA-binding protein 47; RH12, RNA helicase12; VCS, Varicose.

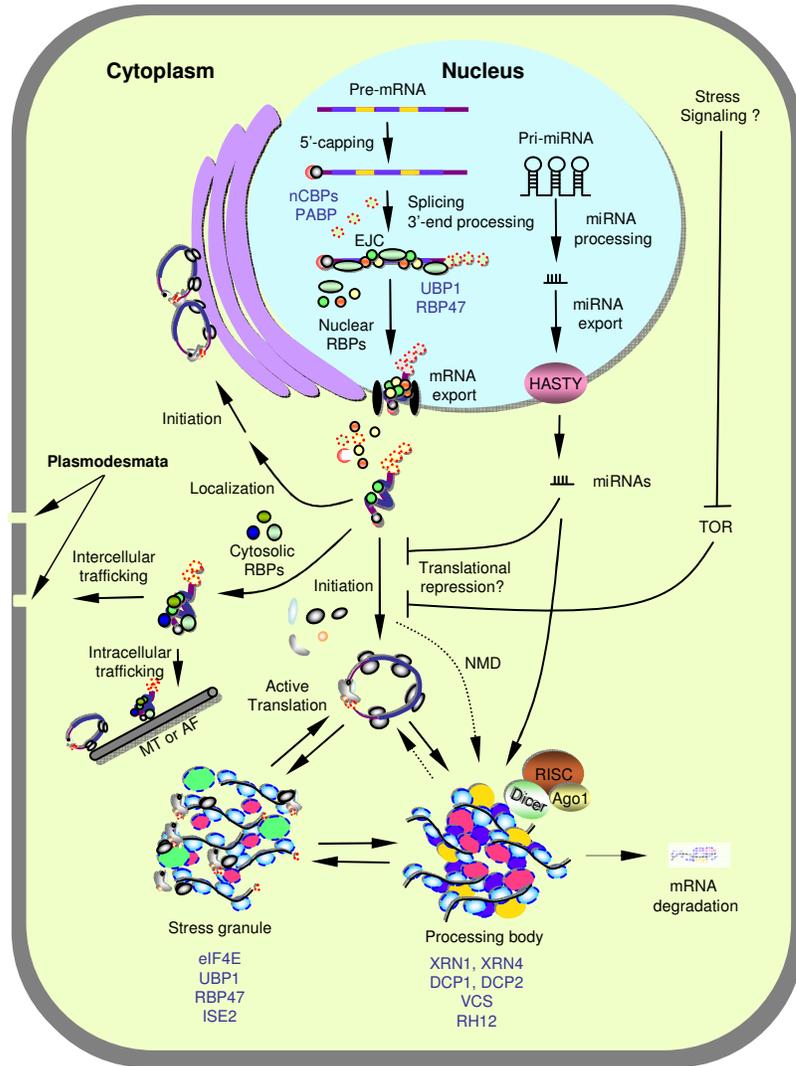


Figure 2.2. Domain architecture of selected RNA binding proteins of *Arabidopsis thaliana*. Representative domain architecture of each family of protein is indicated with symbols. Lines (amino acid sequence with no recognized domain conservation) are scaled to amino acid number of the shortest (black) and longest (gray) protein family members in *A. thaliana*. Abbreviations: a.a., amino acids; LEM, LAP2, emerin, MAN1; NTF2, nuclear transport factor 2; RGG, arginine, glycine, glycine; RRM, RNA recognition motif; Q/N, glutamine, asparagine; DEAD, aspartic, glutamic, alanine, aspartic; DUF, domain of unknown function. Domain symbol size is not to scale.

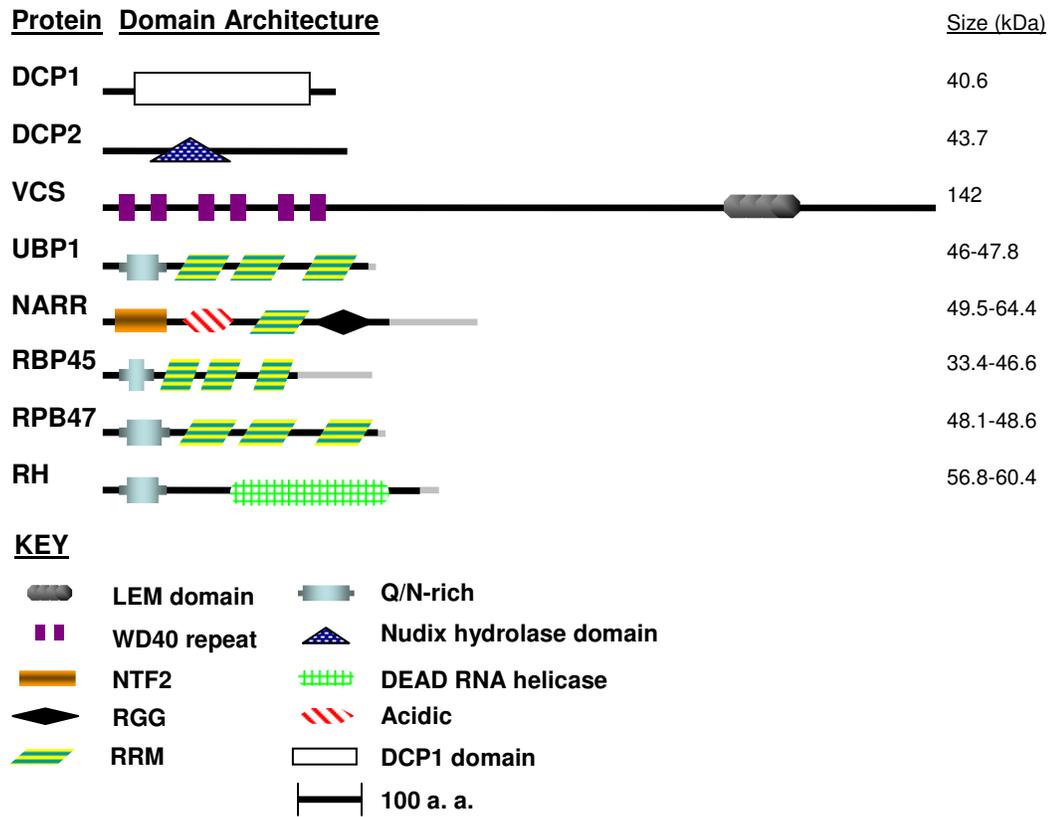


Figure 2.3. Distinct tissue expression by putative SG/PB protein genes of *Arabidopsis thaliana*. A heat map comparing organ-specific gene expression shows examples of unique, organ, zone and tissue-specific expression profiles. The Anatomy tool in Genevestigator 4 (www.genevestigator.com) was used to extract expression values (probe set intensity) from publically available Affymetrix At22k GeneChip microarray data after robust multi-chip adjustment (RMA) normalization followed by adjustments for inter-experiment comparisons. Percent of the expression potential (maximum \log_2 expression value of all 6,458 studies included) is indicated by shading intensity. Gene name and identifier are as follows: UBP1A, AT1G54080; UBP1B, AT1G17370; UBP1C, AT3G14100; RBP45C, AT4G27000; RBP45A, AT5G54900; RBP45B, AT1G11650; RBP45, AT5G19350; RBP47A, AT1G49600; RBP47B, AT3G19130; RBP47C/C' AT1G47490, AT1G47500; NARR1, AT5G43960; NARR2, AT3G25150; NARR3, AT5G60980; NARR4, AT5G48650; RH6, AT2G45810; RH8, AT4G00660; RH12, AT3G61240; VCS, AT3G13290; DCP1, AT1G08370; DCP2, AT5G13570; DCP5, AT1G26110.

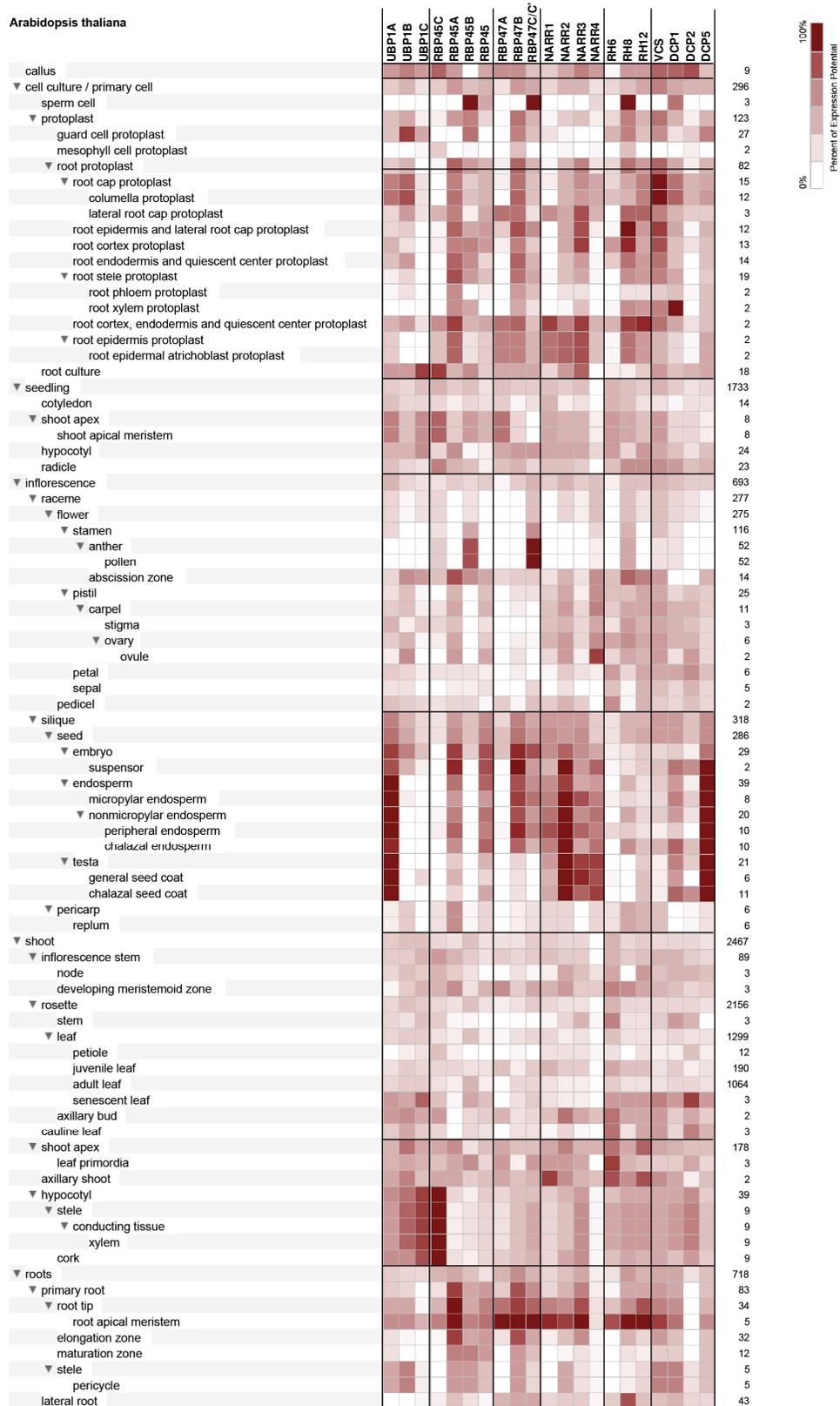


Figure 2.4 Altered phylotaxy of *rh6-1* and *rh8-1* alleles. Line diagram (upper panels) indicating sites of T-DNA insertion in mutant alleles of *RH6* (A) and *RH8* (B). *rh6-1* and *rh8-1* alleles exhibit rare developmental defects (lower panels). Defects included (a) twin flower/silique, (b) twin cauline leaf and axillary buds, (c) kinked inflorescence stem, (d) cauline leaf and adjacent flower, (e) SAM termination in silique, (f) pedicel with leaf and axillary flower, and (g) cauline leaf apical to flower/silique. Not all plants exhibited these phenotypes. Phylotaxy defects in *rh6-2* and *rh8-2* alleles were not determined.

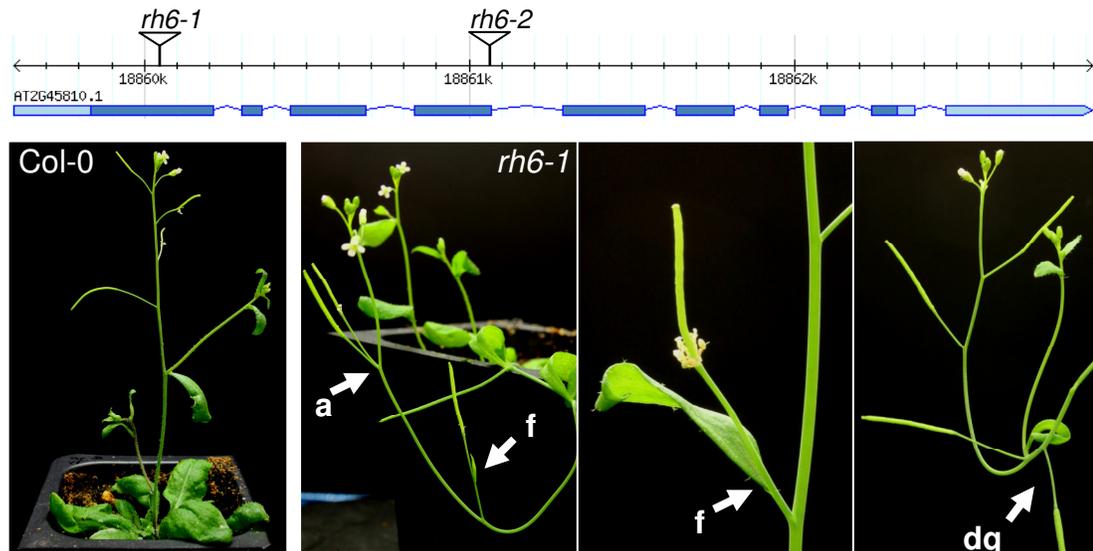
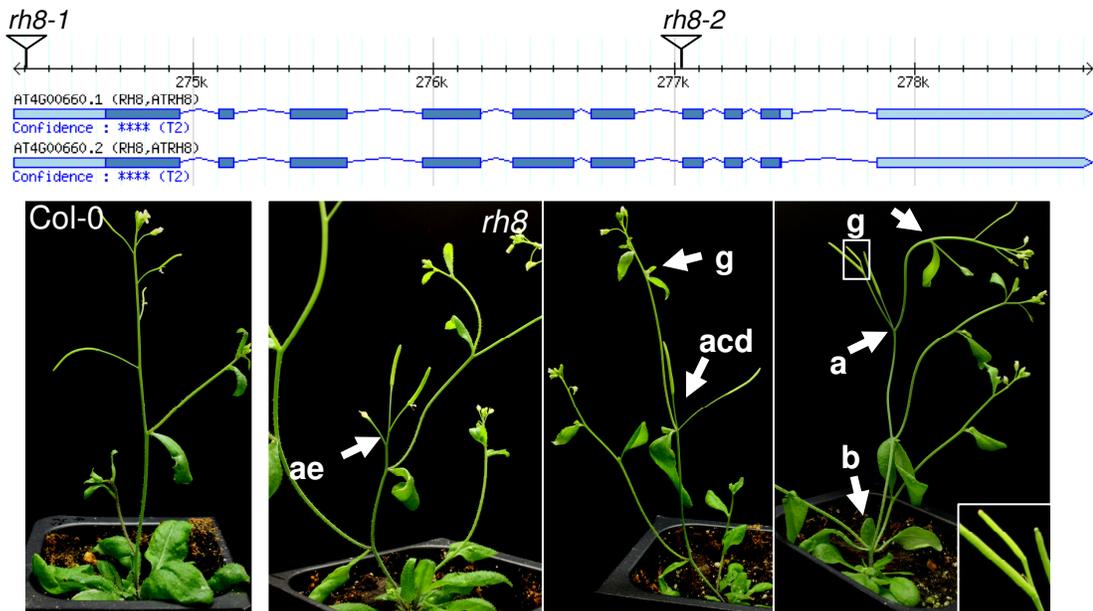
A**B**

Figure 2.5. Ectopic overexpression of RH6-CFP results in a dwarf phenotype. Rosette leaves of independent transgenics *35S:RH6-CFP-28* or *35S:RH6-CFP-17* were round with short petioles compared to Col-0. Flowers and siliques were not smaller but had shorter internodes.

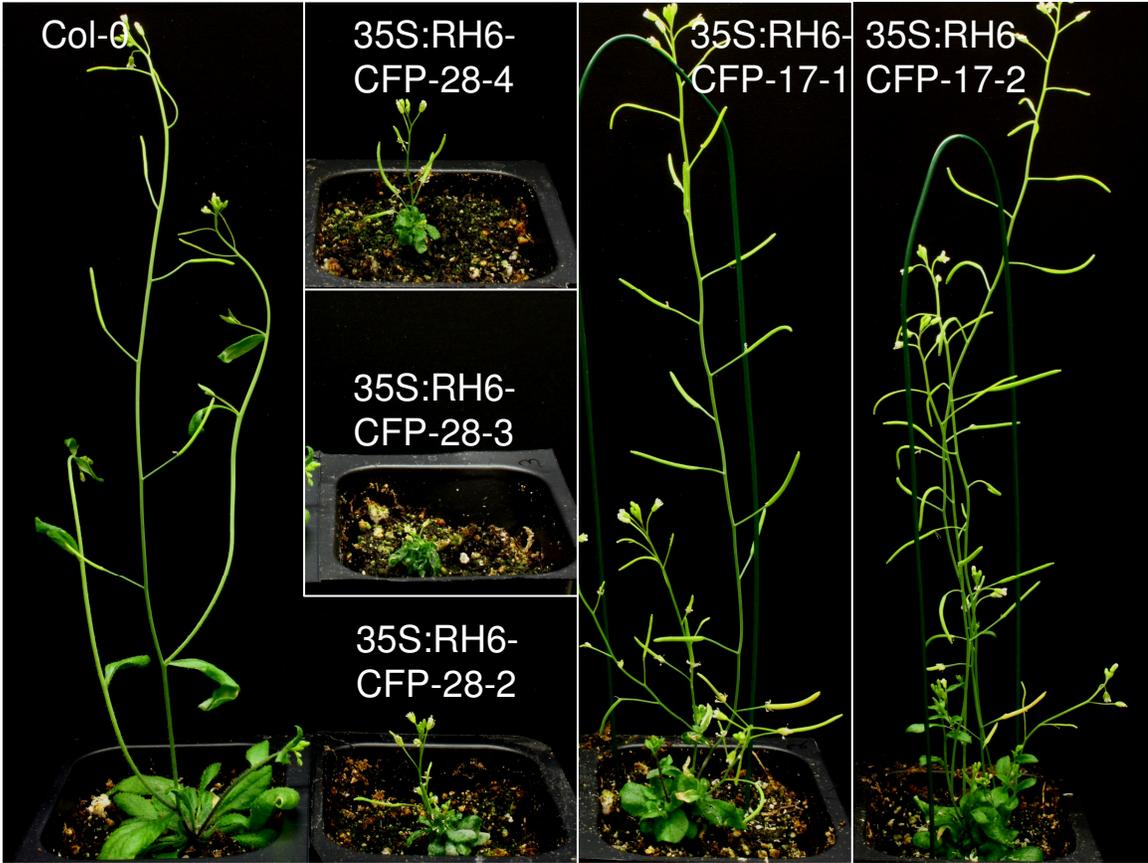


Figure 2.6. Ectopic overexpression of DCP5 results in a dwarf phenotype. (A) Rosettes of five independent T1 lines never *35S:DCP5-CFP-HA* never exceeded a maximum diameter of 2 cm. T1 plants were late flowering and did not bolt until after 5 months, T2 plants (shown for two independent transgenics) did not display delayed flowering. Occasional oversized siliques produced oversized seed and had variable internode length. (B). Juvenile leaves of T2 plants grew twisted with conspicuous serrated edges. Bars = 5 mm.

A



B



Figure 2.7. *rh12-2* exhibits distinct developmental phenotypes. (A) Line diagram of T-DNA insertion mutant alleles of *RH12*. (B) Comparison of Col-0, and *rh12-2*. 7-day-old seedlings grown on MS agar supplemented with 1% (w/v) sucrose and flowering plants. (C) Representative *rh12-2 gRH12-FH* T1 plants. Col-0 (column 1) and *rh12-2* (column 2) at the reproductive stage, are compared to basta-resistant *rh12-2 gRH12-FLAG* T1 plants (column 3-4). Of 71 T1 plants, 69 appeared similar to *rh12-2* (column 3) and two were larger and less affected but distinct from wildtype plants (column 4).

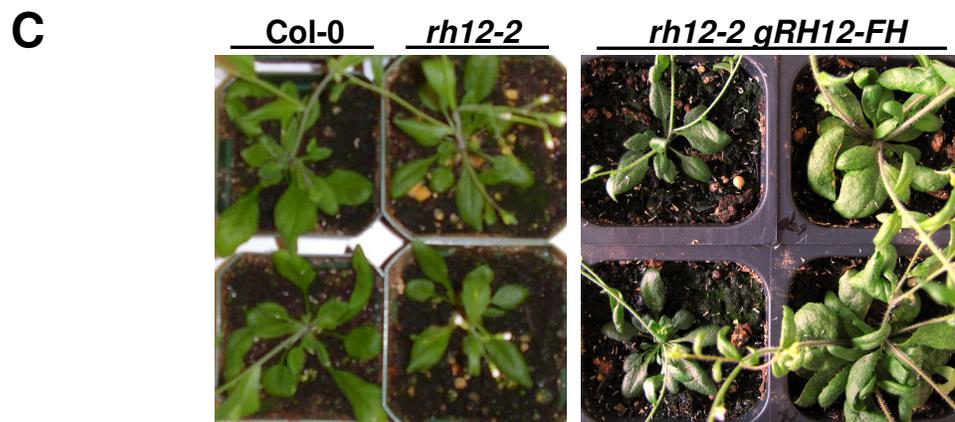
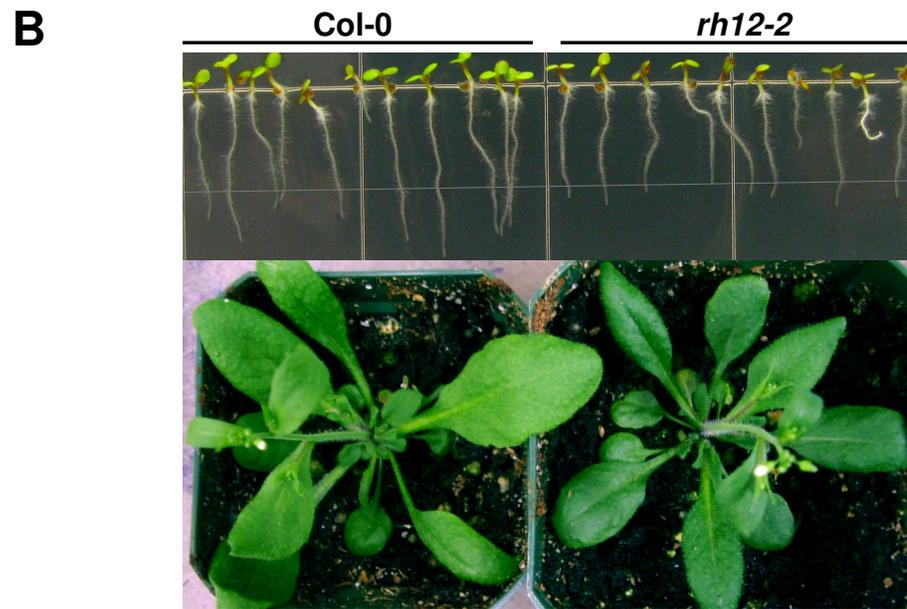
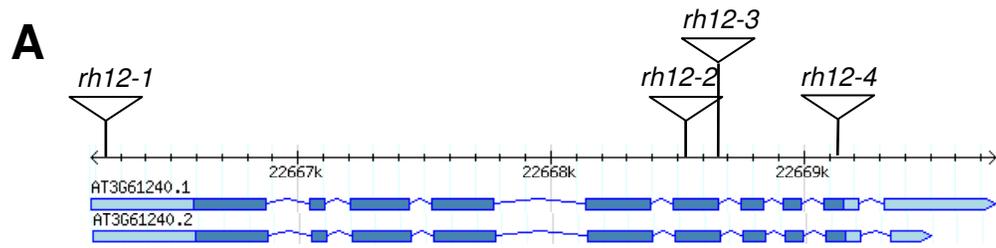


Figure 2.8 Morphological variations in mutants of *UBP1A*. (A) Line diagram (upper panels) indicating sites of T-DNA insertion in mutant alleles of *UBP1A*. Defects in (B) *ubp1a-2*, and (C) *ubp1a-3* include (a) twin flower/silique, (b) twin cauline leaf with floral buds, (c) kinked inflorescence stem, (d) twin axillary bud, (e) pedicel terminated leaf and axillary bud, (f) aborted cauline leaf, (g) cauline leaf w/adjacent flower, and (h) SAM terminates in silique.

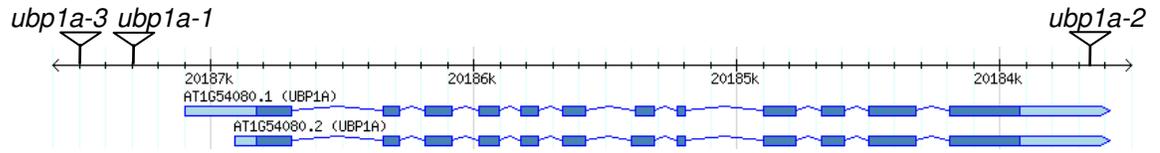
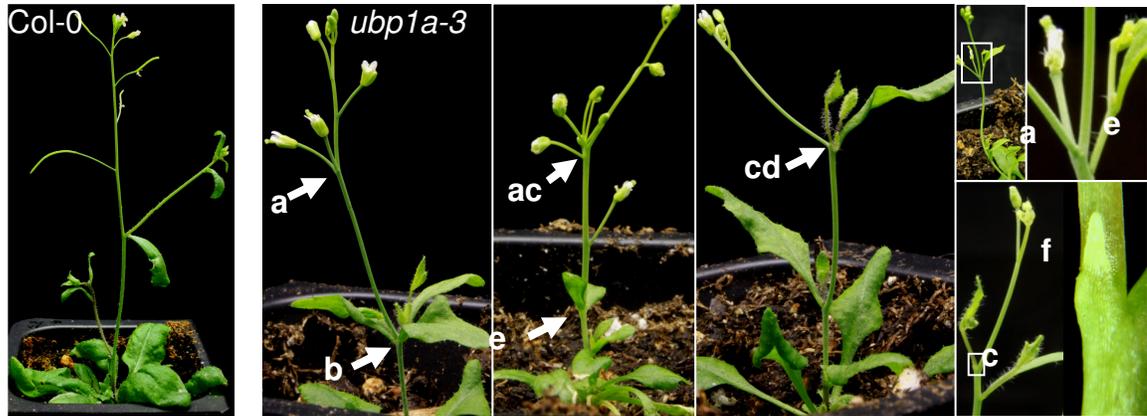
A**B****C**

Figure 2.9. *ubp1c-1* seedlings exhibit prostrate growth. (A) Line diagram (upper panels) indicating sites of T-DNA insertion in mutant alleles of *UBP1C*. When seedlings were grown (B) in low light ($<8 \mu\text{E m}^{-2} \text{s}^{-1}$) on soil but not (C) on MS agar plates containing 1% (w/v) sucrose, *ubp1c-1* seedlings grew prostrate and arrested growth. (D) The white box in (C) is enlarged. *ubp1c-1* seedling cotyledons were slightly smaller than those of Col-0.

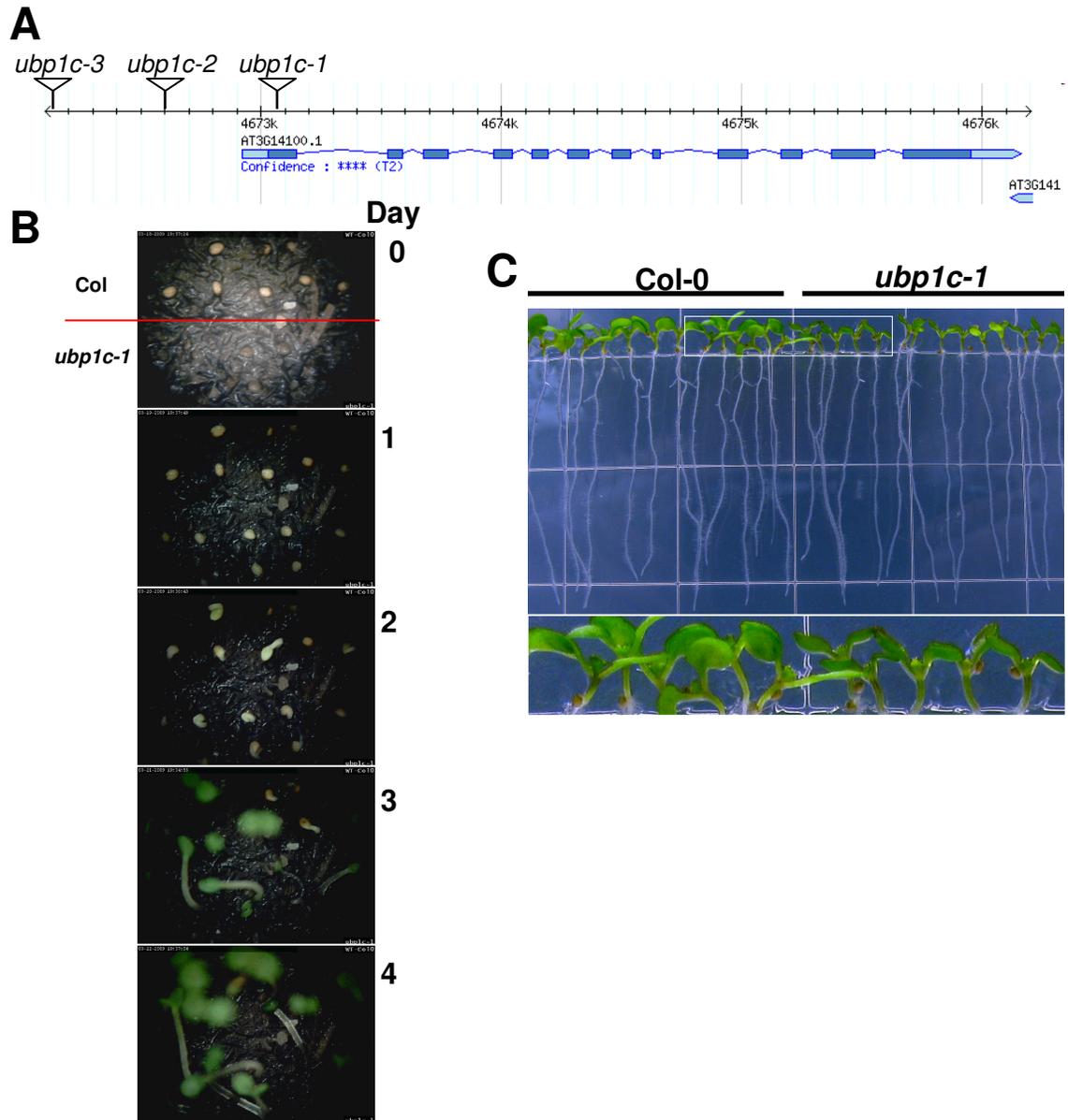


Figure 2.10. Granular localization of four putative processing body proteins. (A) *35S:YFP-RH8* fluorescent protein localization in hypocotyl epidermal cells after 90 min aeration (NS, uncovered bathed in 200 μ L 0.1% MS salt solution on a standard microscope slide) or hypoxia stress (HS) imposed by submergence under coverslip. (B) *35S:RH6-CFP* fluorescent protein localization in hypocotyl epidermal cells after 60 min NS or HS. (C) *35S:RH12-CFP* fluorescent protein localization in cotyledon epidermal cells. (D) Maximum projection image of 8-image Z-series DCP5-CFP fluorescence observed in root tips of 3-d-old seedlings. Bars = 10 μ m.

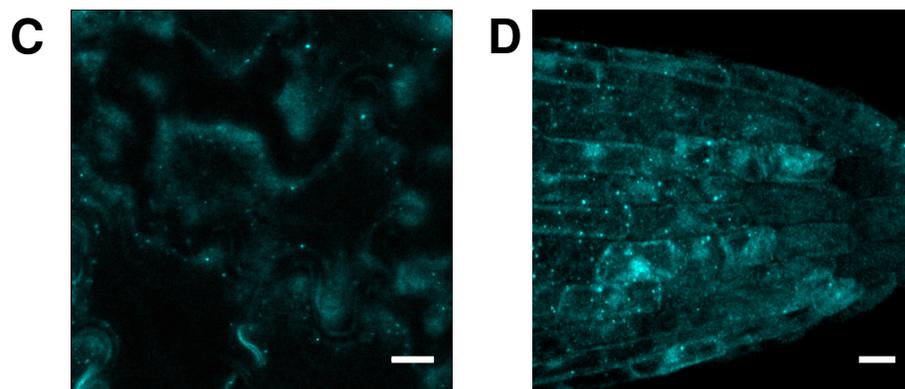
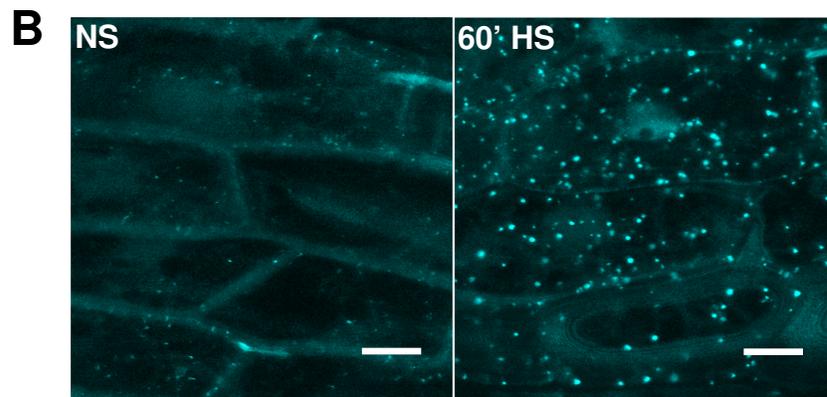
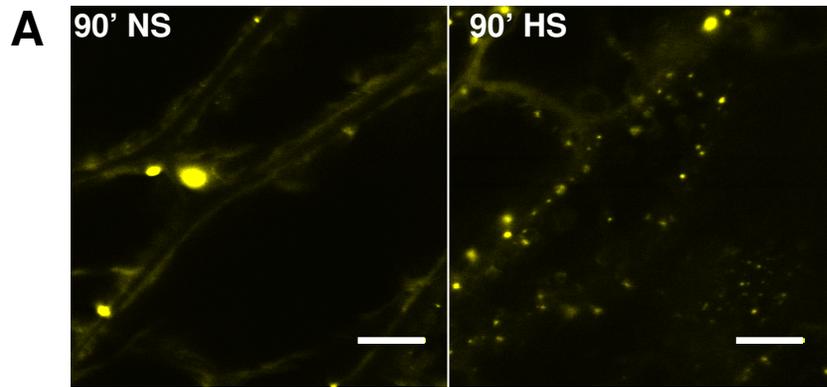
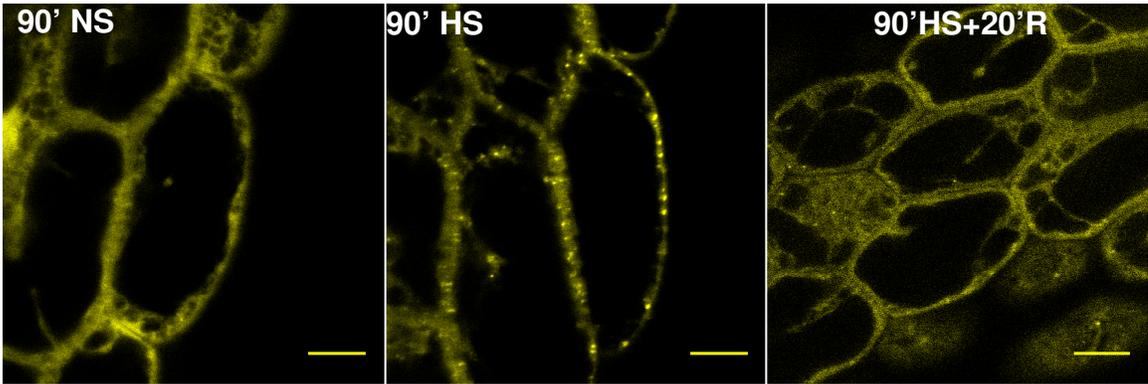
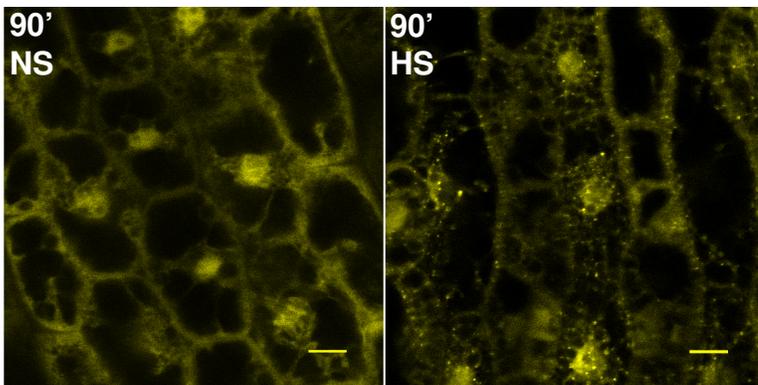


Figure 2.11. Hypoxia-triggered subcellular relocalization of DCP2-YFP. (A, B) Confocal micrographs of hypocotyl epidermal cells of 3-day-old *35S:DCP2-YFP* seedlings treated with 90 min non-stress (NS, uncovered in 200 μ L 0.1% MS salt solution on a standard microscope slide), submergence under a coverslip (HS, in 0.1% MS salt solution on a standard microscope slide), or submergence plus 20 min recovery by aeration (+20'R). (C) Number of granules detected in cells (evaluated from 10 images evenly distributed along the z-axis of a single cell layer) after 10 min (10' S), 70 min (70' S), and 70' S followed by 20 min recovery of the same field of cells. Bars = 10 μ m.

A



B



C

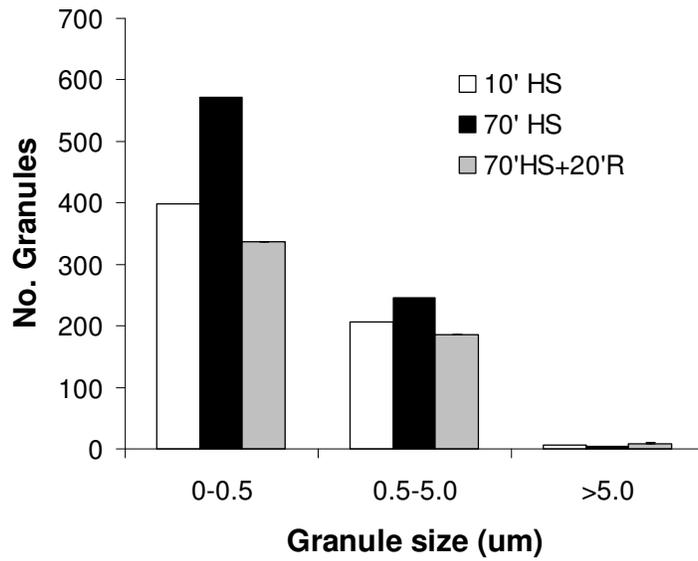


Figure 2.12. Co-localization of YFP-DCP1 and RH12-CFP. Plant lines overexpressing YFP-DCP1 and RH12-CFP were crossed. F1 seedlings were used to visualize localization. CFP and YFP fluorescence colocalized in constitutively formed cytoplasmic granules in hypocotyl epidermal cells. Bar = 10 μ m.

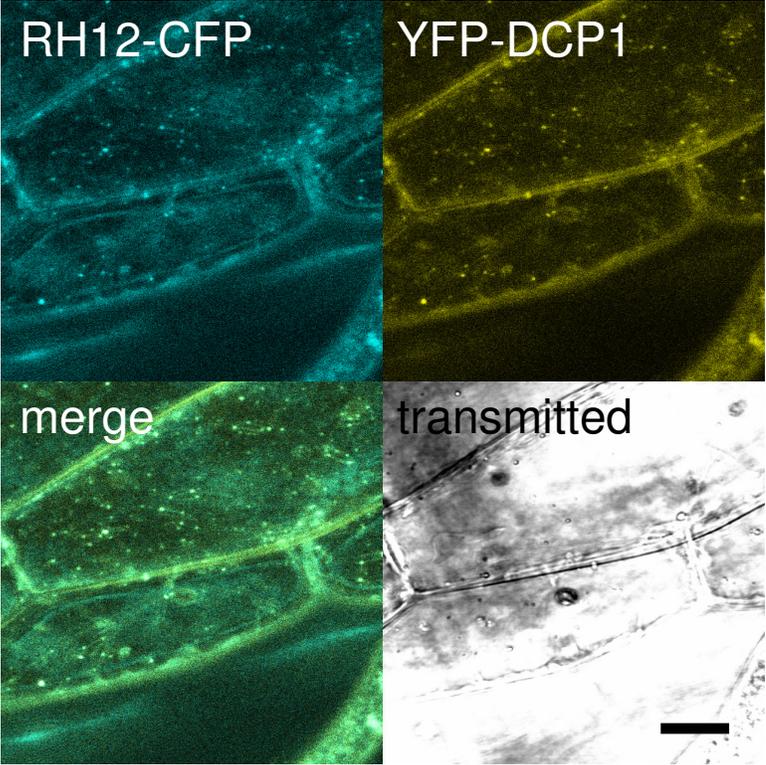


Figure 2.13. UBP1C-GFP localizes to the nucleus, cytoplasm, and cytoplasmic granules. UBP1C localization in a cotyledon epidermal cell colocalized with the nuclear stain, DAPI and also was observed in the cell perimeter and in cytoplasmic foci.

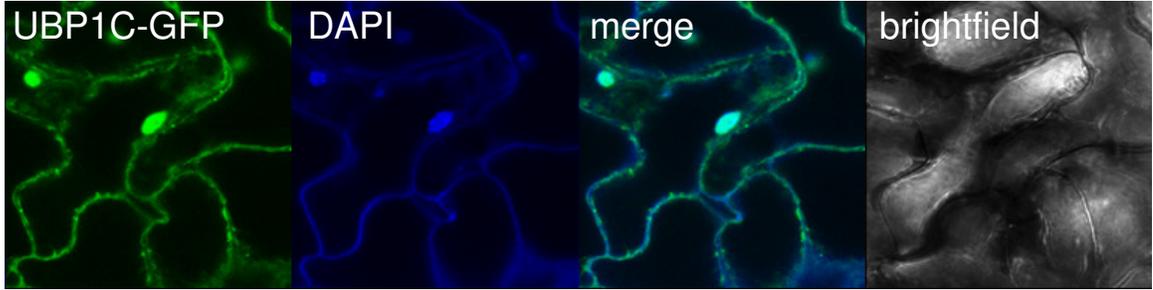
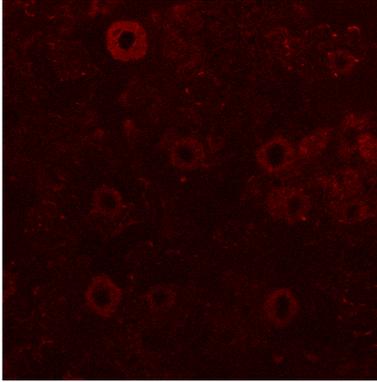
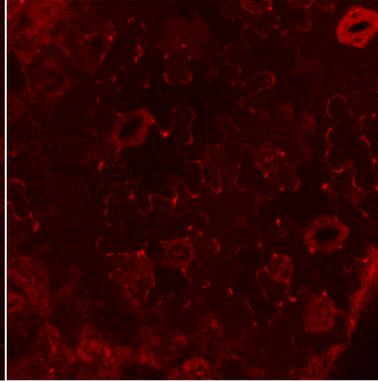


Figure 2.14. Localization of PAB2-mRFP (Col-0). RFP from overexpression lines localized to cytoplasmic lobe tips (white arrows) in epidermal pavement cells of 3-day old cotyledon margins. RFP also accumulated to high levels in guard cells.

cotyledon center



cotyledon margin



Margin pavement cell

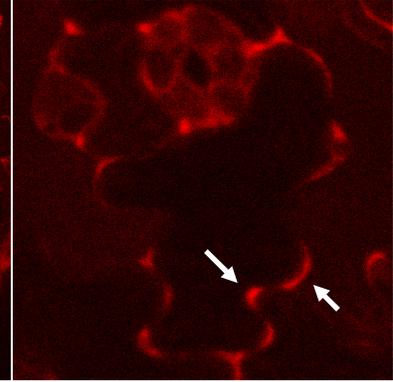


Table 2.1. Overview of mRNP proteins and predicted RNA Binding Proteins of *Arabidopsis*

Protein Group (Gene no. in <i>A. thaliana</i>)	Proteins (no. in <i>A. thaliana</i> and specific names¹)	References
Cytoplasmic Ribosome RNP		
RPL(127)	Large (60S) ribosomal subunit proteins (48 families)	(Chang et al., 2005; Giavalisco et al., 2005; Manuell et al., 2005; Nishimura et al., 2005)
RPS(102)	Small (40S) ribosomal subunit proteins (32 families); RACK1(3)	(Chang et al., 2005; Giavalisco et al., 2005)
RPP(15)	60S ribosomal P proteins (6 families)	(Chang et al., 2005)
Cytoplasmic mRNA management		
Processing Bodies (PB) ² (putative, 8)	AtCAR1, AtDCP2, AtDCP1, AtRH12, AtRH6, AtRH8, VARICOSE, AtXRN4	(Xu et al., 2006; Iwasaki et al., 2007; Goeres et al., 2007; Weber et al., 2008)
Stress Granules (SG) ² (putative, 51)	AtNUC-L1, NTF-RRM(9), RBP45(4), RBP47(4), RHAU-like(9), PUM(26), UBP1(3)	(Lu et al., 2009; Weber et al., 2008; Lorković et al., 2000; Lambermon et al., 2000; Chaikam and Karlson, 2008)
mi/siRNA processes(28)	AGO(10), DCL1(4), DDL, DRB(5), SDE(5), SE, SGS3, RDR(6)	(Voinnet, 2009; Fang and Spector, 2007)
Tudor/Agenet(47)	subcellular trafficking of mRNAs (<i>i.e.</i> similar to OsTudor-SN)	(Wang et al., 2008)
Cytoplasmic translation factors		
eIF(67)	eukaryotic initiation factors (eIF4E, eIF(iso)4E, eIF4A(3) and eIF4B bind RNA), PABP(9)	(Bailey-Serres, 1999; Kawaguchi and Bailey-Serres, 2002; Kim et al., 2007; Weber et al., 2008)
eEF(12); eRF(4)	eukaryotic elongation factors; eukaryotic release factors	
RNPs involved in nuclear processes		
Pre-mRNA processing(24)	CBP20, CBP80, CC1-like(3), AtRBPA/B(6), AtRBPH/F(2), AtRBPI/P(3), AtRSP31, RSP41, RSZp22, RSZ33, SCL30a, AtSF1/BBP, AtSRp30, SR45, AtSRp34b, SR-RRM, U1-70K, U1A, AtU2AF35a, AtU2AF65a, AtU2AF65b, U2B", U2SF3b53b	(Reddy, 2007; Moore and Proudfoot, 2009)
Exon Junction Complex(8)	AtALY-4, AtEIF4A-III, AtRNPS1, AtUAP56-2, UPF1, UPF2, UPF3, AtY14	(Yoine et al., 2006; Arciga-Reyes et al., 2006)
Other RBPs		
Glycine-rich-RBP(17)	RNA chaperones involved in stress acclimation, precise functions unknown: AtAZ1(4), GR-RBP(8), ATE1 AtCSP(4)	(Nakaminami et al., 2006; Kim et al., 2005; Streitner et al., 2008; Kim et al., 2008)
Other RRM(133)	Diverse nuclear and cytoplasmic processes or function unknown: AtAML(5), AtCSTF-64, AtCUG-BP(2), AtCYP59, FCA(2), FLK, FPA, AtHB54, AtKINESIN13A, AtLA1, MCT2, nucleolin-like, PHIP1, PTB(3), AtRANGAP1, RBP37(6), AtREF(4), S-RBP(16), TAF15(2), TBP-BP, TEL(2), 30K-RRM(8), RRM+KH+ZnFinger[CCHC](1), RRM+ZnFinger(16), RRM+Lupis La(4), RRM+other(7), RRM+enz. domain(20), RRM[only recognized domain](26)	(Lorković and Barta, 2002; Ma et al., 2008; Wang and Okamoto, 2009)
K-homology (KH)(25)	Diverse processes or functions unknown: FLOWERING LOCUS K, HEN4(2), PEPPER, KH+Zn Finger[CCCH](2), KH+other(4), KH[only recognized domain](15)	(Lorković and Barta, 2002)
DEAD-Box RNA Helicase(73) excludes 36 proteins with chromatin remodeling domains	Diverse processes or function unknown: BIRH1, DRH1, HEN2(3), LOS4, PRH75, RecQ-like(2), STRS1, STRS2, AtSUV3-like, RH+Zn finger(5), RH+DUF1605(16), RH+Q-motif(36), DEAD2(4)	(Xu et al., 2006; Kobayashi et al., 2007; Kant et al., 2007)
Pentatricopeptide repeat proteins (PPR)(485)	Plastid and mitochondrial RNA processing, splicing, editing, translation, degradation	(Schmitz-Linneweber and Small, 2008)

Abbreviations: At, *Arabidopsis thaliana*; Os, *Oryza sativa*. ¹*Arabidopsis* gene names based on The *Arabidopsis* Information Resource. ²Association with these complexes might be transient.

Table 2.2. TDNA mutant lines. Mutants of RNA binding proteins with a putative role in RNA granules regulating translation and mRNA sequestration were obtained and screened for developmental phenotypes.

Mutant	Locus	Stock Accession	Ecotype	Insert Position	ORF Position*	Marker Resistance	Observations of Developmental Phenotypes
<i>dcp2-1</i>	At5g13570	Salk_000519	Columbia	exon 3	244	NPTII	as reported by Coeres et al., 2007
<i>narr1-1</i>	At5g43960	Salk_023921	Columbia	exon 9	1288	NPTII	only stem and roots greens; slow growth; requires sugar; flowers produced, but no seed
<i>narr2-1</i>	At3g25150	Salk_144662	Columbia	intron 5		NPTII	none observed
<i>narr3-1</i>	At5g00980	Salk_87_G09	Columbia	exon 5	940	BAR	none observed
<i>narr4-1</i>	At5g48650	Salk_027468	Columbia	exon 9	1274	NPTII	none observed
<i>rbp47a-1</i>	At1g49600	Salk_072433	Columbia	exon 1	149	NPTII	aberrant root hair patterning; high frequency of branched root hairs
<i>rht2-1</i>	At3g61240	Salk_024905C	Columbia	5'UTR	-368	NPTII	seedling growth arrest on MS +1% suc; multiple swollen spherical cells; disordered SAM (seg > 12/80)
<i>rht2-2</i>	At3g61240	Salk_016921C	Columbia	exon 6	1125	NPTII	shrunken cotyledons; all adult leaves have dark green appearance; smaller rosette
<i>rht2-3</i>	At3g61240	Salk_028633	Columbia	exon 6	1218	NPTII	none observed
<i>rht2-4</i>	At3g61240	Salk_148563C	Columbia	exon 9	1448	NPTII	none observed
<i>rht6-1</i>	At2g45810	Sail_111_H08	Columbia	exon 1	173	BAR	occasional aberrant phytotaxy
<i>rht6-2</i>	At2g45810	Sail_849_H06	Columbia	intron 4		BAR	patches of white aborted tissue in primary root, cotyledon
<i>rht8</i>	At4g00660	Salk_016830C	Columbia	5'UTR	-335	NPTII	occasional aberrant phytotaxy
<i>rht8-2</i>	At4g00660	Salk_029698	Columbia	intron 6		NPTII	none observed
<i>ubp1a-1</i>	At1g54080	Salk_116327	Columbia	promoter		NPTII	none observed
<i>ubp1a-2</i>	At1g54080	Sail_1185_C11	Columbia	3'UTR		BAR	occasional aberrant phytotaxy
<i>ubp1a-3</i>	At1g54080	Salk_068668	Columbia	-417 of TSS		NPTII	phytotaxy of stem abnormal/variable; twin-multiple flowers/siliques; poor fecundity; pollen
<i>ubp1b-1</i>	At1g17370	Salk_040505	Columbia	3'UTR lip		NPTII	none observed
<i>ubp1b-2</i>	At1g17370	Flag_071F09	Wassilewskija	intron 3		NPTII/BAR	none observed
<i>ubp1c-1</i>	At3g14100	Sail_363_G02	Columbia	exon 1	100	BAR	sucrose-dependant growth; dark-colored seed; agravitropic seedling shoot; weak cotyledon expansion
<i>ubp1c-2</i>	At3g14100	ET5803	Landsberg erecta	-314 of TSS		NPTII	none observed
<i>ubp1c-3</i>	At3g14100	Flag_384B08	Wassilewskija	-781 of TSS		NPTII/BAR	none observed
<i>vcs-7</i>	At3g13300	Salk_032031	Columbia			NPTII	as reported by Coeres et al., 2007

Table 2.3. Primers used for genotyping TDNA mutant lines. Genotypes were identified by PCR of genomic DNA using these primers and marker segregation.

Mutant	Genotype Primer 1	5' Primer Sequence**	Genotype primer 2	3' Primer Sequence**	Left Border Primer***	Left Border	Prod. Size* (wt)	Prod. Size** (T-DNA)	Homozygous (plant number; genomic PCR)	Plant (No.) with Progeny 100% Kan/Basta Resistance
<i>dcp2-1</i>	A15g13570F	caccATGTGGGCTCCATCGATCAT	DCP2gR	CCCCACAGGGTTTACCCT	LB_6313R	5'	688	849	NA	silenced
<i>narr1-1</i>	A15g43960F	caccATGGGACTCCTTATCCTGGAGCGA	A15g43960R	CGACACCCACCGCGGTAGTAA	LB_6313R	3'	2497	362	8,11,15,17	none
<i>narr2-1</i>	A13g25150F	caccATGGGACTTGTAGGTGCACAGCAA	A13g25150R	CGCAGCAAGACACACACGGG	LB_6313R	3'	2937	1163	2,10,11	nd
<i>narr3-1</i>	A15g60980F	caccATGGCAGCAGCAGGAAGCTAGTC	A15g60980R	AGATGAACCCACACCTCGAGCT	LB2	3'	2037	1022	1,4,5,7,8,9,11,12	nd
<i>narr4-1</i>	A15g48650F	caccATGGATTCTACTGCTGCAACCAAC	A15g48650R	GTACGAGTTGATGCTGGCGA	LB_6313R	5'	2251	2451	7,9,14,15,16	nd
<i>rbp47a-1</i>	A11g49600F	caccATGCAGACACCAACAACACGGTTC	A11g49600R	TCAAGAAGCTCCCGGACTGC	LB_6313R	3'	2915	449	5,6,10	none
<i>rh12-1</i>	proA1RH12_F	caccTCAAAGTTTGTGGCCATCA	RH12gR	GAAACTGATCGTGGAACTCAGG	LB_6313R	5'	847	924	1	none
<i>rh12-2</i>	RH12gIF	CGCCACATTCCTGTCACTG	RH12gt-2R	CACGTTTCATGCACAGTTACAAG	LB_6313R	3'	950	947	1-5	1 (weak)
<i>rh12-3</i>	RH12gIF	CGCCACATTCCTGTCACTG	RH12gt-2R	CACGTTTCATGCACAGTTACAAG	LB_6313R	3'	950	854	nd	2,6
<i>rh12-4</i>	RH12gIF	CGCCACATTCCTGTCACTG	RH12gt-2R	CACGTTTCATGCACAGTTACAAG	LB_6313R	5'	950	1177	nd	4 (weak)
<i>rh6-1</i>	RH6_SEQ_F	CGGAATCGAATCCACCTC	RH6_SEQ_R	GCTTCTTGAGATACCGATCCTTG	LB2	5'	1515	584	4,14	22
<i>rh6-2</i>	RP-A12g45810	AGTTGATCAGTGGAACTTCCAG	RH6_SEQ_R	GCTTCTTGAGATACCGATCCTTG	LB2	5'	686	623	nd	nd
<i>rh8</i>	gRH8_F3	CTTGCAATAGCGTTAATCTTTGAGC	RH8_Start_R	GCCAAATCCAGGAGGGT	LB_6313R	5'	691	590	4	1,3
<i>rh8-2</i>	gRH8_F5	GCCTCTAAACCATTGCTTC	A14G00660R	TTGGCAATAAATTCCTGATCG	LB_6313R	5'	1582	1473	3,18	nd
<i>ubp1a-1</i>	UBP1A-1_F	CAACAACATCACCTTGAAGCA	UBP1A-1_R	GCCCCAAGCTGATATAT	LB_6313R	5'	586	808	nd	1,12
<i>ubp1a-2</i>	UBP1A-2_F	GCCCCATCAGCAGCTCAT	UBP1A-2_R	TTCCACACAAGCTGTGGG	LB2	3'	1284	1277	5	5
<i>ubp1a-3</i>	UBP1A-1_F	CAACAACATCACCTTGAAGCA	UBP1A-1_R	GCCCCAAGCTGATATAT	LB_6313R	5'	586	600	6	6
<i>ubp1b-1</i>	nd				3'					none
<i>ubp1b-3</i>	A11g17370F	caccATGCAGAGTTGAAGCAGCAGCA	A11g17370R	TACTGGTAGTACATGAGCTGCTGGG	LB_6313R	5'	2984	1427	1,2,3,8	1,2,3,8
<i>ubp1c-1</i>	A13g14100F	caccATGCAGAAATCCGAGACTGAAGCAACATC	UBP1C-2_R	GAGGAGCCAAAACACCCAG	LB2	3	121	321	7,9,10	7,11
<i>ubp1c-2</i>	proUBP1C_F	caccCAAAGGCAAAGCTGTCTGTT	UBP1C-2_R	GAGGAGCCAAAACACCCAG	DS3-3	3'	1566	837	3,6	3,9,10
<i>ubp1c-3</i>	proUBP1C_F	caccCAAAGGCAAAGCTGTCTGTT	UBP1C-2_R	GAGGAGCCAAAACACCCAG	nd	5'	1566	862	7	1,7

** T-DNA left border orientation with respect to the gene sense strand. ** 5' "cacc" sequences were used in cloning and are not gene specific; wt, wildtype allele specific product size (5' & 3' Primers); T-DNA, mutant allele-specific product size (Left border primer & respective opposing primer gene primer). *** Left border primer sequences: LB_6313, 5'-TCA AAC AGG ATT TTC GCC TGC T-3'; LB2, 5'-GCT TCC TAT TAT ATC TTC CCA AAT TAC CAA TAC A-3'; DS3-3, 5'-GAA ATT GAA AAC GGT AGA GGT-3'.

Table 2.4. Genomic DNA segments cloned for generation of native, FLAG-tagged gene constructs. Genomic DNA segments, including upstream intergenic (promoter) and gene transcript sequence, were cloned and fused to a carboxy-terminal Flag epitope tag in the *gateway-FH-OCST (nptII)* T-DNA vector (Mustroph et al. 2009) or *gateway-FH-OCST (bar)* using the Gateway™ system for transformation into loss-of-function mutant lines.

Transgenic Line name	Gene	Promoter Length	Transcript Length	Line for Transformation	Binary Destination Vector
<i>ubp1c-1 gUBP1C-FH</i>	UBP1C	1343	3029	<i>ubp1c-1</i>	gateway-FH-ocsT (nptII)
<i>rh12-1 gRH12-FH</i>	AtRH12	585	2969	<i>rh12-1</i>	gateway-FH-ocsT (bar)
<i>rh12-2 gRH12-FH</i>	AtRH12	585	2969	<i>rh12-2</i>	gateway-FH-ocsT (bar)

Table 2.5. List of fluorescent protein-tagged overexpression lines of putative stress granule or processing body proteins.
 Proteins were tagged with a tag by translational fusion and stably transformed into *Arabidopsis thaliana* plants via *Agrobacterium tumefaciens* mediated transformation and screened.

flicDNA ORF Locus	Gene Construct	Putative RNA Granule	Selection Marker	Genetic Background	Line No.	Binary Vector*
At5G13570	35S:DCP2-YFP-HA	PB	BAR	Columbia-0	1,5	pEarleygate 101
At1G08370	35S-YFP-DCP1	PB	BAR	Columbia-0	(gift)	pEarleygate 104
At3G61240	35S:RH12-CFP-HA	PB	BAR	Columbia-0	2	pEarleygate 102
At2G45810	35S:RH6-CFP-HA	PB	BAR	Columbia-0	1,28	pEarleygate 102
At4G00660	35S:YFP-RH8	PB	BAR	Columbia-0	1,3,7,12	pEarleygate 104
At1G26110	35S:DCP5-CFP-HA	PB	BAR	Columbia-0	1	pEarleygate 102
At1G54080	35S:UBP1A-GFP-His	SG	BAR	Columbia-0	4,16	pEarleygate 103
At3G14100	35S:UBP1C-GFP-His	SG	BAR	Columbia-0	1,21,31	pEarleygate 103
At3G14100	35S:UBP1C-GFP-His	SG	BAR	<i>ubp1c-1</i>	7,11,16,17	pEarleygate 103
AT4g34110	35S:PAB2-mRFP	SG	HPT	Columbia-0	6,8,9,27	pGWB554
AT5G15200	35S:RPS4-mRFP	Ribos, SG	HPT	Columbia-0	2,3	pGWB554

* pEarleygate vectors – see Earley et al. 2006, pGWB vectors – see Nakagawa et al. 2007. Abbreviations: GY/CFP, green/yellow/cyan fluorescent protein; SG, stress granule; PB, processing body; Ribos, ribosome; NA, not applicable; BAR, basta resistance; NPTII, neophosphate transferase II; HPT, hygromycin phosphotransferase.

Chapter 3

Role of the RNA-Binding Protein UBP1C in Translational Repression During Hypoxia in *Arabidopsis thaliana*

3.1 Abstract

Plants subjected to oxygen deprivation experience an energy crisis. As a result, key energy consuming processes are curtailed. Protein synthesis is globally decreased and limited to the selective translation of hypoxia-responsive gene transcripts. Most of the translationally repressed mRNA is stable and rapidly return to polysome complexes upon reoxygenation. In animals, translationally repressed mRNAs enter large cytoplasmic complexes that form via aggregation of Tia-1. *Arabidopsis thaliana* encodes three oligouridylate binding proteins (UBP1s) identified as homologs of Tia-1 that have been used as stress granule markers. Strong conservation of RNA Recognition Motif (RRM) domains in UBP1s indicates a conserved, ancient cellular role of these proteins. Here we characterized the biological function of two UBP1s expressed at high levels during vegetative development, UBP1A and UBP1C. Correlated transcript accumulation profiles suggest similar regulation of *OLIGOURIDYLATE BINDING PROTEIN 1A (UBP1A)* and *OLIGOURIDYLATE BINDING PROTEIN 1C (UBP1C)*, with some notable exceptions. Phenotypic analysis of two mutant alleles of *UBP1C* demonstrated non-redundant roles in seed color, seedling meristem growth, cotyledon expansion, starch accumulation, and cytokinin response. The overexpression of *UBP1C* led to dwarf plants and premature death of cotyledons. Analysis of mutant alleles of *UBP1A* also demonstrated altered phylotaxy of flowering plants, decreased organ size, and low seed yield. *ubp1a* and *ubp1c* mutant seedlings were altered in survival of hypoxia. Whereas *ubp1c-1* was hypersensitive to hypoxia, FLAG-UBP1C overexpressors were less sensitive. *ubp1a-3* had enhanced survival as compared to wildtype. GFP-fusion to both

UBP1C and UB1A permitted visualization of the dynamic relocation of these proteins from diffuse cytosolic localization to granular forms, which quickly reversed by reoxygenation. The formation of UB1 granules was inhibited by cycloheximide, indicating the requirement of completion of translational elongation. UB1C granules contained polyadenylated RNAs. To assess the mRNAs that associate with UB1C, an immunopurification scheme was developed to isolate UB1C-mRNA complexes from seedlings under control, low-light, oxygen deprivation and reoxygenation conditions. Microarray analysis of RNA from these complexes identified mRNAs enriched in UB1C complexes under control and short term hypoxia. UB1C-associated mRNAs had U-rich 3'- untranslated regions and were enriched for membrane-associated and cell wall targeted proteins including a large number of AGPs as well as a large number of transcription factors and hormone regulated genes. The UB1C-bound mRNAs maintained association under hypoxia. However, in addition most other genes increased association and included those that were stable and translationally repressed and those that were unstable; reoxygenation rapidly reversed the hypoxia-triggered UB1C association. This study demonstrates selective and dynamic binding of UB1C to transcripts, particularly the subset that is translationally repressed during oxygen deprivation.

3.2 Introduction

3.2.1 Plant UB1 is a homolog of mammalian TIA-1

In pursuit of mechanisms that regulate mRNA stability and sequestration during hypoxia stress in *Arabidopsis thaliana*, we identified a number of putative RNA binding proteins as candidates for cytosolic translational regulation of mRNAs (Chapter 2). The selection was based on amino acid sequence homology to non-plant proteins involved in

processing bodies (PBs), and/or stress granules (SGs). These large cellular structures have been characterized as locations of mRNA storage, sorting, triage in times of stress. Among this group was a family of proteins with homology to mammalian TIA-1 (T-cell intracellular antigen 1) and TIAR (TIA-1 related) proteins (TIA1/R). These two proteins possess three RNA Recognition Motif (RRM) domains, bind RNAs and are involved in numerous cellular activities including SG formation. Because TIA1/R have some degree of preferential mRNA binding and mediate mRNA aggregation into SGs, we chose to characterize the roles of the closest plant TIA1/R orthologs, *OLIGOURIDYLATE BINDING PROTEIN 1 (UBP1)*, which is encoded by three genes in *Arabidopsis thaliana*.

3.2.2 TIA-1 and TIAR proteins regulate post-transcriptional processes

Numerous studies of TIA1/R have shown that they regulate multiple post-transcriptional processes. In animals, TIA-1 and TIAR function as translational regulators of the potent pro-inflammatory signal Tumor Necrosis Factor- α (TNF- α) and of apoptosis (Gueydan et al., 1999; Piecyk et al., 2000; Förch and Valcárcel, 2001; Eisinger-Mathason et al., 2008). In mouse, TIA1/R genes are predominantly expressed in brain, spleen, and testis (Beck et al., 1996). TIA-1 and TIAR genes each encode two alternatively spliced isoforms characterized by the presence of three RNA recognition motifs (RRM1-3). The most closely related RNA-binding proteins are the poly(A) binding proteins (PABPs) that function in the stabilization and translation of mRNAs in the cytoplasm (Coller et al., 1998).

Although animal TIA-1 and TIAR have similar tissue expression, a high degree of homology, and similar activity *in vitro*, genetic analyses have demonstrated distinct functions *in vivo*. TIAR knock-out mice display partial embryonic lethality and defects in maturation of germ cells, whereas TIA-1 knock out mice are viable (Förch and Valcárcel, 2001), suggesting these proteins have unique roles during development. The individual TIA1/R

proteins display antagonistic, overlapping, and distinct nuclear, nucleocytoplasmic, and cytoplasmic roles in regulating the splicing, accumulation and translation of mRNAs.

3.2.3 TIA1/R proteins regulate alternative splicing, nucleocytoplasmic transport, and mRNA stability

TIA1/R proteins have nuclear functions. TIAR was shown to enhance the splicing of weak 5'-splice sites by binding an AU-rich element adjacent to the splice site and stabilizing U1 snRNP splicing factor binding to pre-mRNAs (Del Gatto-Konczak et al., 2000) such as Fibroblast Growth Factor Receptor 2 (FGFR2) and the Fas receptor. TIAR is a pro-apoptotic nuclear protein, which relocalizes within 30 min to the cytoplasm upon Fas stimulation, a signal of apoptosis, prior to nuclear disintegration and DNA fragmentation (Taupin et al., 1995). Fas-activated serine threonine kinase (FASTK) phosphorylates TIA1/R and leads to enhanced Fas exon 6 retention and increased pro-apoptotic Fas message accumulation (Izquierdo and Valcárcel, 2007). This forms a positive feedback loop regulated by activity and localization of TIA1/R. The role of TIA-1/R can be antagonized by overlapping competitive RNA binding by Hu antigen R (HuR), another triple RRM-domain protein known to bind within the 3'-UTR of several oncogenes (Izquierdo, 2010; Kim et al., 2011).

Both TIA-1 and TIAR bind RNA. The *in vitro* SELEX binding assay was used to evaluate the binding activities of the individual RRMs of these proteins. This determined that RRM2 of human TIA-1 and TIAR binds to uridylylate-rich elements (Dember et al., 1996). Whereas RRM1 did not bind RNA in the assay, RRM 3 bound non-specifically to RNA. Although RRM2 was necessary and sufficient to bind a poly-uridylylate element, the presence of RRM1 and RRM3 increased the binding affinity. Independent studies reported that TIA1/R binds RNA in the nanomolar range (Kim et al., 2011). Distinct RRM domain functionality was illustrated for RRM2 and RRM3 in nuclear import/export by mutation analysis of TIA-1 (Zhang et al., 2005). The RNA binding activity of the RRM2 domain is required for nuclear

import and accumulation and is dependant on Ran-GTPase activity. By contrast, nuclear export was dependant on RNA binding of the RRM3 domain and is independent of the Ran-GTPase and CRM1 pathways (Zhang et al., 2005).

There is also increasing evidence that TIA1/R proteins function in the regulation of stability and translation of specific mRNAs in the cytoplasm. It was found that the AU-rich element (ARE) of the TNF- α 3'-UTR is bound by either TIAR (Gueydan et al., 1999) and TIA-1 (Piecnyk et al., 2000). Mice lacking TIA-1 are unaffected in TNF- α mRNA abundance and stability but are hypersensitive to the upstream signal lipopolysaccharide, also known as endotoxin, a chemical component of Gram-negative bacteria outer membranes and potent inducer of inflammation in animals. These mice show increased association of TNF- α with polysomes, suggesting a role for TIA1/R in the repression of translation of TNF- α . It was also found that TIA1/R association with the ARE of the β -subunit of mitochondrial H⁺ ATPase transcript correlated with increased steady state accumulation of the target transcript (Izquierdo, 2006). To further elucidate TIA-1 targets in human cells, RNA immunoprecipitation (RIP) was used to enrich mRNA (López de Silanes et al., 2005). The enriched mRNAs were identified by microarray hybridization and evaluated for over-represented motifs. A 37 bp bipartite motif with a U-rich segment adjacent to an AU-rich segment was identified in the 3'-UTRs of mRNAs that bound TIA-1. The accumulation of proteins encoded by targeted mRNAs was shown to increase following TIA-1 siRNA-mediated knock down in control and heat stress conditions, supporting the hypothesis that TIA-1 limits the translation of the mRNAs it binds.

To date, the role of TIA1/R in splicing has not been linked to its role in translation. However, nuclear binding of these proteins could predispose mRNAs associated with TIA1/R to be co-regulated in translation or sequestration in the cytoplasm. The fact that RRM2 is responsible for nuclear accumulation and U-rich element binding and the fact that TIA1/R

also binds cytosolic mRNAs via a U-rich motif predominantly located within 3'-UTR sequences conceptually links nuclear activity, RNA binding, export and cytosolic control of translation in non-growth-limiting conditions.

3.2.4 Stress granule aggregation and repression of ribosomal protein synthesis

TIA1/R play a major role in cellular response to stress by aggregating mRNAs and proteins into stress granules. Stress granules (SGs) are large aggregates (up to 4-5 μm diameter) of ribonucleoprotein (RNP) complexes which form in response to stresses such as oxidative stress, UV, nutrient limitation, heat, and hypoxia (Gottschald et al., 2010). The mRNAs and proteins they contain are in dynamic exchange with the cytoplasm as treatment with the translation elongation inhibitor cycloheximide causes the stabilization of polysomes and their depletion; and fluorescence recovery after photobleaching (FRAP) experiments using GFP-fused TIA-1 or PABP-I have shown dynamic protein exchange with the cytoplasm (Kedersha et al., 2000). The role of TIA1/R in SG aggregation is facilitated via both RNA binding activity and regulation of a prion-related, glutamine-rich domain (Gilks et al., 2004).

The complete protein composition of SG is not known but a number of proteins have been recognized using immunohistochemistry and numerous reviews have been published (Ivanov and Nadezhdina, 2006; Anderson and Kedersha, 2006, 2008, 2009; Anderson, 2010). Along with mRNAs and TIA1/R, many other proteins are present in SG complexes of mammalian cells. These include translation initiation factors eIF4E, eIF4G, eIF4A, eIF3, and PABP, ternary complex-deficient 48S preinitiation complex, G3BP1, RHAU, FMRP, and others (see reviews above).

In human cells, TIA1/R proteins are implicated in the regulation of mRNAs containing 5'-terminal oligopyrimidine tracts (5'TOPs) in response to nutrient limitation. Transcripts with this element include many that encode components of cytosolic ribosomes and translation

factors. It has long been purported that the translation of mRNAs with a 5'TOP is curtailed under growth and energy limiting conditions as a means to limit energy expenditure on cell growth (e.g. ribosome biogenesis). It was shown by Damgaard et al. (2011) that upon amino acid starvation 5'TOP mRNAs increase their association with TIA1/R, as shown by use of an RNA immunoprecipitation (RIP) assay. The siRNA-mediated knock-down of TIA1/R genes, derepressed 5'TOP ribosomal protein translation during amino acid starvation, as shown by ³⁵S-methionine incorporation. The derepression was correlated with 5'TOP mRNA retention in dense polysome complexes during amino acid starvation only when TIA1/R was knocked down. Notably, this phenomenon was reversed when siRNA-resistant TIA-1 or TIAR genes were co-transfected with siRNA targeting endogenous TIA-1/R (Damgaard and Lykke-Andersen, 2011).

3.2.5 UBP1s of plants

In plants, UBP1 was initially identified as a heteronuclear RNA binding protein involved in pre-mRNA splicing in tobacco (Lambermon et al. (2000). It was later recognized as sharing amino acid sequence similarity to TIA1/R. Plant introns differ from that of other eukaryotes because of higher bias in U and A-content. U-rich elements in introns increase the efficiency but not the accuracy of splicing of introns in transiently transformed tobacco cells (Goodall and Filipowicz, 1989; Ko et al., 1998). UBP1 was among a number of RBPs recognized as U-rich element binding proteins in Arabidopsis. The others included RBP45, RBP47 (Lorković et al., 2000) and the two UBP1 interacting proteins UBA1 and UBA2 (Lambermon et al., 2002). It was shown that UBP1 associates with U-rich regions within introns and 3'-UTRs *in vitro* and enhances reporter gene abundance in an intron- and promoter-specific manner in transiently transfected protoplast of tobacco (Lambermon et al., 2000). It was also suggested that UBP1 enhances transcript abundance of intronless genes in a promoter-specific manner. Later, the unrelated UBP1-associated 1 (UBA1) and UBP1-

associated 2 (UBA2) proteins were identified by yeast two-hybrid screen of UBP1 against an *A. thaliana* cDNA library. UBA1 was isolated by screening the library with either *Nicotiana plumbaginifolia* (*Np*)*UBP1* or *AtUBP1b*. Reciprocal screens using UBA1a or UBA2a as bait confirmed interaction of these proteins with all three *AtUBP1* family members (A, B, and C) as well as each other. UBA1a was also confirmed to interact with itself. *In vitro* pull down assays confirmed that the interaction of these proteins is independent of RNA. UBA1 and UBA2 both localize to the nucleus and bind U-rich RNA. However, neither they nor the UBP1-related three-RRM U-rich binding proteins RPB45 and RBP47 stimulated splicing of weak introns (Lorković et al., 2000).

These conclusions were based on an artificial system using synthetic genes and introns, with viral promoters and 3'-UTRs, with *A. thaliana* proteins in *N. plumbaginifolia* protoplasts transformed with high levels of plasmid (Lambermon et al., 2000, 2002). The biological significance of these observations in the nuclear pre-mRNA splicing process has yet to be determined *in planta*. Also, to date, no studies have addressed the possible role of UBP1 in nucleocytoplasmic transport of RNAs.

UBP1s also appear to function in the cytoplasm. Weber and colleagues (Weber et al., 2008) established that similar to TIA1/R, UBP1B and RBP47 are markers of heat-triggered plant SGs, which are distinct from previously reported heat shock granules (Nover et al., 1989). They also demonstrated that a UBP1 with mutated RRM3 was unable to bind RNA and has a dominant negative phenotype preventing cytosolic granule formation in heat-stress. More recently, UBP1B was used as a marker of plant SGs in studies of so-called siRNA bodies, which contain proteins associated with siRNA mediated gene silencing: Argonaute7 (AGO7), SGS3 and RDR6 (Jouannet et al., 2012). It was found that upon heat stress, GFP-tagged AGO7 and SGS3 were recruited to UBP1B-containing SGs.

The extent to which the functions of animal TIA1/R proteins have been conserved in plant UBP1s is unknown. The reported similarities in the cellular activities of UBP1s and TIA1/R suggests that core characteristics may have been retained and tuned by evolution to control completely different physiology. The question of whether the three isoforms of UBP1 in Arabidopsis have redundant or specialized function as that of the four isoforms of TIA1/R of animals is also unanswered. However, distinct mechanisms of their regulation have been shown.

McCue et al. (2012) demonstrated a unique pathway which regulates *UBP1B* via siRNA854 targeting *UBP1B* but not *UBP1A* or *UBP1C*. siRNA854 was predicted to target four elements within *UBP1B* 3'-UTR. The siRNA854 family is reportedly conserved from animals to plants and recognizes TIA-1 and *UBP1B* orthologs, respectively (Arteaga-Vázquez et al., 2006). The target sites are absent from the 3'-UTRs of *UBP1A* or *UBP1C*. This causes a *UBP1B* expression pattern that is suppressed in the presence of siRNA854. siRNA854 is generated from transcriptionally activated *Atilia6* retrotransposon, an event which occurs in the pollen vegetative nucleus or in mutants with altered epigenetic silencing as in the *DECREASE IN DNA METHYLATION 1 (ddm1)* mutant. They also show that a *ubp1b* loss-of-function mutant has decreased survival to salt and osmotic stress (McCue et al., 2012) possibly indicating a unique role for UBP1B. Whether this is an isoform-specific stress phenotype or it is partially redundant and shared among *UBP1* mutants has not been tested. Logically, stress phenotypes are likely related to the functional role these proteins play in SG.

3.2.6 UBP1s and translation

It is hypothesized that the sequestration of mRNAs in SGs limits them from translation. Numerous stimuli cause regulated changes in translational status of cohorts of transcripts. These changes have been monitored at the genomic scale in response to

dehydration, hypoxia, sucrose starvation, (reviewed in Bailey-Serres et al., 2009), ozone (Puckette et al., 2012), gibberellic acid (Ribeiro et al., 2012), and photomorphogenic light (Liu et al., 2012), salinity, and heat (Matsuura et al., 2010; Ueda et al., 2012) . Of these treatments, low oxygen stress causes severe energy limitation and a dramatic decrease in protein synthesis. Branco-Price et al. (Branco-Price et al., 2008) measured changes in mRNA-polysome association by means of polyribosome (polysome) RNA immunoprecipitation and microarray hybridization on the Affymetrix ATH1 Genechip® (RIP-Chip; Zanetti et al., 2005). By this means of translating ribosome affinity profiling (TRAP) they identified over 65% of the total cellular mRNAs detected in the assay as translationally repressed but stable during the stress. These transcripts are reloaded onto polysomes following 1 h reoxygenation. We hypothesized that this group of mRNAs, which includes many ribosomal protein genes, are sequestered into SG and stored until reoxygenation, in a sub-lethal stress treatment. The mRNAs that bind SG marker proteins *in planta* during normal growth have not been identified, nor is it known if the mRNA clients of these proteins are altered by environmental factors.

In this study we focus on the biological function of Arabidopsis *UBP1A* and *UBP1C*. T-DNA insertion mutants of these genes cause distinct developmental and hypoxia stress phenotypes. However, both *AtUBP1C-GFP* and *AtUBP1A-GFP* also form cytosolic granules in response to hypoxia *in planta*. A focused investigation on *UBP1C* cell biology and RNA binding activity was carried out. *UBP1C* granule formation is dependent on translational elongation, a characteristic of SGs. *UBP1C* mutants have poor hypoxia survival, whereas FLAG-*UBP1C* overexpression in the mutant improved hypoxia survival. By use of FLAG-*UBP1C* transgenics and RIP-Chip, we measured the levels of *UBP1C*-associated RNAs under normal growth conditions in the light and dark in response to hypoxia and upon reoxygenation. We found that under normal growth conditions *UBP1C* preferentially binds to

a subset of RNAs with U-rich 3'-UTRs. Associated mRNAs only modestly shift in response to anticipated diurnal darkness, but, when treated with hypoxia, mRNAs that are poorly or not associated under non-stress and do not have high U composition increase association. This association is reversed by 20 min of reoxygenation. The transcripts associated with UBP1C under hypoxia are highly enriched in those that are stable but translationally repressed during hypoxia.

3.3 Results

3.3.1 Plant UBP1s are triple RRM-domain proteins that are related to animal TIA1/Rs

Plant UBP1 proteins were recognized as sharing RRM domain similarity with animal TIA1/R (Weber et al., 2008). The animal protein family most similar to the triple RRM TIA1/Rs is the quadruple RRM-containing polyadenylate binding proteins (PABPs). In plants, an expansion of similar RRM-domain proteins early in land plant evolution resulted in two additional groups, RBP45s and RBP47s, which retained greater similarity to UBP1s than PABPs based on queries with the Protein Basic Local Alignment Search Tool (BLASTP). To better elucidate the evolutionary relatedness of these proteins, a phylogeny was generated by amino acid sequence alignment of representative animal TIA1/R, plant-specific UBP1, RBP45 and RBP47, and animal and plant PABPs. This phylogeny confirms that UBP1s are more closely related to TIA-1/R than to RBP45 and RBP47 (Figures 3.1, 3.2). The *Arabidopsis thaliana* genome encodes three *UBP1* genes: *UBP1A* (At1g54080), *UBP1B* (At1g17370) and *UBP1C* (At3g14100). The encoded proteins are 83% identical to one another at the amino acid level and are well conserved in close relatives (i.e. *Arabidopsis lyrata*, *Thellungiella parvula*, and *Brassica rapa*) (Figure 3.1).

A comparison of the N- and C-terminal regions of UBPs and RBP45/47s revealed divergence in length and sequence between these protein groups. However, both possess unstructured, Q- and QN-rich stretches characteristic of prion-like domains (Figure 3.2).

3.3.2 Regulated expression of *UBP1* genes in *Arabidopsis*

The strong amino acid sequence similarity of *Arabidopsis* *UBP1A-C* suggests that *UBP1* genes may be genetically redundant. To investigate overlap in expression profiles, publicly available expression data were evaluated using Genevestigator (Zimmermann et al., 2004) and other tools. Indeed, *UBP1A*, *UBP1B*, and *UBP1C* display highly correlated mRNA accumulation across hundreds of global gene expression studies. The co-regulation was evident in a co-expression network (Figure 3.3A). *UBP1C* is connected to a hub centered on a FRIGIDA-like hydroxyproline-rich glycoprotein (At3g22440) shown to associate with indole-3-glycerol phosphate synthase (IGPS) a key enzyme in tryptophan and auxin biosynthesis (Bürkle et al., 2005). *UBP1A* and *UBP1C* correlated expression matched well overall, co-correlating with most other *Arabidopsis* genes (Figure 3.3B). Nonetheless, distinctions in the regulation of *UBP1* genes were apparent in experiments assessing condition-, tissue-, and cell-type specific expression. Among examples of condition specific regulation, enhanced accumulation of *UBP1C* mRNA in rosette leaves of plants exposed to prolonged darkness and intensified in dark plus submergence (Lee et al., 2011) suggests that this gene may play a role in regulation of carbon or energy usage (Figure 3.3D).

mRNA presence in polysome complexes (Mustroph, Juntawong, et al., 2009) were evaluated as a measure for cell-type specific expression in seedlings under both control conditions and after 2 h of hypoxia (Figure 3.4). Of the three genes, the absolute level of *UBP1A* mRNA in polysomes was the highest and showed little difference between the 19 cell-types/ regions assayed. *UBP1B* mRNA levels in polysomes was also consistent throughout cell types with slightly higher abundance in the root endodermis. *UBP1C* mRNA,

however, was more abundant in the polysomes of shoots. Epidermal cells expressed *UBP1C* the highest followed by vascular bundle sheath cells and mesophyll cells (Figure 3.4). A survey of publically available DNA microarray datasets that monitor total mRNA abundance, identified high *UBP1A* mRNA accumulation in seed tissues including embryo, endosperm, testa, and suspensor cells as well as meristems (seedling shoot apical meristem and root meristem) (Figure 3.5). *UBP1B* mRNA was enriched in ovules, mature shoot apices, root pericycle, and stamen abscission zones (Figure 3.5) whereas that of *UBP1C* was high in hypocotyls and in senescent leaves (Figure 3.5).

Evaluation of the abundance of mRNAs encoded by *RBP45* and *RBP47* gene family members also highlighted variation (Figure 3.5). Expression profiles suggest that *RBP45B* and *RBP47C* are uniquely expressed at high levels within the sperm cell. *RBP47B* mRNA is abundant in the endosperm and along with *RBP45A* and *RBP45* mRNA accumulates in embryos and the suspensor. *RBP47A* transcripts increased in imbibed seed and meristematic zones. *RBP45A* mRNA was high in stamen abscission zones, and the root tip and meristem. *RBP47A* and *RBP47B* are both also highly expressed in the root tip and meristem. Taken together, these data demonstrate that members of the *UBP1*, *RBP45* and *RBP47* were independently expressed across tissue-types of Arabidopsis.

3.3.3 *ubp1c-1* seedling phenotypes were recapitulated by artificial miRNA targeted silencing

Three publicly available transfer-DNA (T-DNA) insertion alleles for *UBP1C* were identified but only one was due to an insertion into an exon. *ubp1c-1* (SAIL_363_G02) contains a T-DNA within exon 1 at position +99 from the translation start site (Figure 3.6A; data not shown). Semi-quantitative RT-PCR with a primer 5' to the insertion site failed to produce an amplicon of the exonic region (Figure 3.6 B), indicating that *ubp1c-1* does not produce full-length *UBP1C* mRNA. However, primers that bind to the T-DNA (left) border and span the intron 6 splice site produced an amplicon at similar levels (primers f & h) as primers

that bind just 3' of the insertion site and the intron 6 splice site in wild-type Col-0 (primers e & h) (Figure 3.6C). This suggests that transcription initiating within the T-DNA insertion may lead to production of an N-terminally truncated form of UBP1C in *ubp1c-1*.

Because a second insertion allele within exonic or intron sequence of *UBP1C* was unavailable, we overexpressed an artificial miRNA designed to uniquely target *UBP1C* (*amiR-ubp1c*) using the CaMV 35S promoter (Figure 3.7G, 3.6A). *amiR-ubp1c-32* and *amiR-ubp1c-48* transgenic seedlings accumulated reduced levels of *UBP1C* mRNA, to as low as 3 % of wildtype (Figure 3.6).

We found that *ubp1c-1* plants exhibit multiple phenotypic traits in seeds and seedlings. A number of these phenotypes were recapitulated in the *amiR-ubp1c* transgenics (see below). First, *ubp1c-1* seedlings exhibited sucrose-rescued growth arrest with incomplete penetrance. When grown on MS agar plates, growth was arrested in both the root and shoot within 3 days of germination (radicle emergence from the seed coat) (Figure 3.8A). Seedling development was fully rescued by transfer of the arrested seedlings to medium containing 29 mM (1%) sucrose. When quantitatively assayed using seed lots produced at the same time, seedling arrest was observed in <25% of Col-0 seedlings but in >95% of *ubp1c-1* plants (Figure 3.8B). Incremental increases in concentration of sucrose in the growth medium proportionally rescued the growth of the shoot (Figure 3.8A, 3.8C) and root (Figure 3.8A, 3.8D). Most of the *amiR-ubp1c* lines displayed a higher percentage of seedling growth arrest than the wildtype, emphasizing that penetrance was variable.

In addition to the dependency on sucrose for seedling establishment, seeds of *ubp1c-1* were darker in color compared to wild-type. When crossed to Col-0 and the F₁ selfed, the F₂ phenotypic segregation ratio of seed color was 3:1 light to dark (Chi-square test p-value=0.585, n=228; Figure 3.9A). To investigate linkage of these two phenotypes, the F₂ seed were separated by seed color and grown in the presence or absence of sucrose. We

found that 70% of the dark seed (n=41) were arrested in development on plates lacking sucrose, whereas, only 11% of the light colored seed (n=97) were sucrose dependent, supporting linkage (Chi-square p-value<0.005). From the same population, 76 F₂ seedlings were genotyped using PCR and allele specific primer pairs. All of the growth-arrested seedlings were homozygous *ubp1c-1* (data not shown), demonstrating genetic linkage of <5 map units (Chi-square test p-value=0.0455).

To further validate recapitulation of *ubp1c-1* phenotypes in the *amiR-ubp1c* lines, a more in depth analysis was undertaken. *amiR-ubp1c* was introduced into Col-0 plants via *Agrobacterium*-mediated floral dip transformation and T₁ transgenic plants were selected using kanamycin in presence of sucrose plus kanamycin or by growth arrest on media lacking sucrose and kanamycin. Each method identified transgenic plants. These plants were grown and self-crossed. The T₂ populations were scored for percent growth arrest, confirming varying penetrance from 0-50% in segregating T₂ populations (Figure 3.10A). Five lines with approximately 10, 20, 30, 40, and 50% growth arrest in the segregating T₂ populations were selected for further characterization. The lines with the highest percent growth arrest had subtly segregating seed color similar to *ubp1c-1*. Therefore, seed from these segregating T₂ populations was separated by eye into two groups - one with light colored seed and the other with dark colored seed. These seed were scored for growth arrest following seven days of growth on medium lacking sucrose. The darker seed had >60% arrested growth, whereas the lighter seed had <40% growth arrest (Figure 3.10B, 3.10C). These seedlings were then transferred to MS plates with sucrose and their root length was measured following one and seven additional days of growth. Growth was rescued in both *ubp1c-1* and *amiR-ubp1c* lines (Figure 3.10C, 3.10D); confirming that both *ubp1c-1* and *amiR-ubp1c* seedlings display both a darker seed color and sucrose-rescued

growth arrest. There were no phenotypic differences that were noted between wildtype, *ubp1c-1*, and *amiR-ubp1c* lines throughout vegetative and floral development (Figure 3.10F).

3.3.4 *ubp1c-1* seedlings display other early seedling phenotypes

Additional seedling phenotypes were observed for the *ubp1c-1* mutant but were not evaluated in the *amiR-ubp1c* lines. *ubp1c-1* seedlings germinated in low light (<15 μ E) in soil exhibited prostrate growth and appear agravitropic (Figure 3.9B). Because hypocotyl gravitropism is dependant on starch granule accumulation in the columella of the root cap and endodermis, and, therefore on sucrose biosynthesis and transport, we evaluated chlorophyll and starch accumulation in the mutant plants. Based on histochemical staining with iodine, starch granules were less evident in mutant plants except a very small amount in the root tip of the stunted plant after seven days of growth without sucrose (Figure 3.9C). However, when growth was rescued with exogenous sucrose, starch accumulation in *ubp1c-1* seedlings was similar to wild type plants grown without sucrose (Figure 3.9C). Chlorophyll content of two-week old plants was measurably reduced as well (Figure 3.9D). Cell division is driven by the plant hormone cytokinin. To test the *ubp1c-1* response to cytokinin, the growth response to the cytokinin analog kinetin was assayed. *ubp1c-1* had a distinct response to kinetin. Whereas Col-0 plants developed concentration-dependant anthocyanin accumulation, *ubp1c-1* did so to a lesser degree (Figure 3.9E); even though under other conditions anthocyanin did accumulate to a high degree (data not shown).

3.3.5 *ubp1c-1* morphological phenotypes were not complemented by UBP1C overexpression or genomic constructs

Frequently mutant phenotypes can be complemented by expression of the wildtype cDNA or gene in the mutant background. This is accomplished by use of a gene construct driven by a constitutive or the endogenous promoter or by transformation with a DNA

fragment that includes the deduced genic region. Multiple attempts were made to complement the seedling phenotypes of *ubp1c-1*. Two overexpression fusion proteins representing both N- and C-terminal fusion constructs failed to complement the sucrose dependence phenotype. Neither *ubp1c-1 35S:UBP1C-GFP* (Figure 3.7B) nor *ubp1c-1 35S:FLAG-UBP1C* (Figure 3.7C) seedlings would grow for longer than 3 days on medium without sucrose (Figure 3.8A, 3.8B) or had the lighter seed color phenotype. Multiple native gene constructs also failed to complement the mutant. The promoter sequence of these constructs included the upstream intergenic sequence (1343 bp) as well as the 5'-UTR, and full gene sequence including introns. A C-terminal FLAG-His epitope tag was encoded without the native 3'-UTR but with the terminator of the octopine synthase gene. FLAG-tagging the N-terminus and inclusion of the 3'-UTR also did not alter the result. The T₂ generation of plants transformed with the various native gene constructs included plants that displayed slight dwarfism and lower seed yield (Figure 3.11A, 3.11B; Table 3.1).

3.3.6 Overexpression of UB1C-GFP causes unique phenotypes

The full-length UB1C cDNA fused to GFP and driven by the CaMV 35S promoter was stably transformed into the *A. thaliana* ecotype Col-0 and *ubp1c-1* genetic backgrounds. The transgenics displayed normal to dwarf stature in independent lines in the *ubp1c-1* background. In the dwarfed plants, the flowers, stems, and leaves all appeared smaller compared to wild-type (Figure 3.12A). Mature plants with severe growth reduction also presented terminal flower defects similar to the terminal flower mutants *tfl1* (Hanano and Goto, 2011) and *tfl2* (Kotake et al., 2003). Defects include multiple fused or open carpels, twisted and malformed gynoecia, and petaloid stamen (Figure 3.12B). Seedlings expressing 35S:UB1C-GFP in populations segregating 3:1 (GFP+:GFP-) had premature chlorotic cotyledons and seedling death. At three weeks of age on medium lacking sucrose, arrested *ubp1c-1* seedlings appear green and healthy compared to those expressing UB1C-GFP,

which were yellow and shriveled (Figure 3.12C). This yellow cotyledon phenomenon was observed in seedlings grown for the equivalent amount of time with exogenous sucrose as well. However, emerging rosette leaves of these transgenics appeared green and of normal size (data not shown). The yellow cotyledon phenotype was also apparent in lines of *ubp1c-1 35S:FLAG-UBP1C*, but again the emerging rosette leaves appeared normal (data not shown).

3.3.7 *ubp1c-1* is hypersensitive to hypoxic stress

To determine if UBP1C contributes to plant survival during hypoxia stress, survival of Col-0 and mutant plants were compared over a time course of hypoxia. Ten-day-old plants of Col-0, *ubp1c-1*, *amiR-ubp1c-32*, and two independent lines of *ubp1c-1 35S:FLAG-UBP1C* (20 and 24) were subjected to 10, 11.5, or 13 h of oxygen deprivation in Ar(g) purged Plexiglass chambers in near darkness. Plant survival was determined following 5 days. A duration-dependant decrease in survival was observed. Significantly higher survival was seen in wildtype plants at each time point. After 13 h stress, less than 40% of Col-0 plants died, whereas more than 80% of the mutant plants died (Figure 3.13A). Tissue of surviving *ubp1c-1* was visibly more damaged than surviving Col-0 tissue (Figure 3.13B). The hypersensitivity to hypoxia was also seen in 7 d plants as well (data not shown). *amiR-ubp1c* plants also had more reduced survival, but were not significantly different than Col-0 plants. In contrast to the lack of effect on morphologic phenotypes of *ubp1c-1* by over expression of FLAG-UBP1C, hypersensitivity to hypoxia by *ubp1c-1* was partially rescued in these lines, as at 13 h hypoxia both overexpression lines had significantly higher survival (Figure 3.13A, 3.13B).

3.3.8 *ubp1a* mutants confer morphological phenotypes that are distinct from those of *ubp1c* mutants

Three mutant alleles of *UBP1A* were identified within public collections of T-DNA mutants (Figure 3.6). Two alleles, named *ubp1a-2* and *ubpa-3*, displayed similar phenotypes. The sites of T-DNA insertion were located in the 3'-UTR and the promoter region of the gene, respectively (Figure 3.6). A third allele, *ubp1a-1*, with the insertion in the promoter, exhibited no obvious visible phenotypes. qPCR revealed that the level of *UBP1A* gene transcript in whole seedlings was not altered in any of these alleles (data not shown). Nonetheless, the detection of similar phenotypes in *ubp1a-2* and *ubp1a-3* suggests that the abnormalities are due to alteration of *UBP1a* expression, presumably transcription based on the sites of insertion. Each has altered phylotaxy of the mature reproductive stem, with a more severe phenology observed in *ubp1a-3* (Figure 3.13A). *ubp1a-3* exhibited twin flowers/siliques at most floral positions. Occasionally, a large cluster of flowers formed at a single position (Figure 3.13B). Twin and triplet cauline leaves were also frequent, or a cauline leaf and axillary bud formed with a partner flower/silique at the same position along the stem. At these sites of multiple organs the stem had a kinked or 'zigzag' morphology (Figure 3.13A, 3.13B). Additionally, the apical meristem was sometimes prematurely terminated in a twin or triplet silique. Less often, a flower was formed basally to a cauline leaf, an axillary bud formed with a leaf stump, or cauline leaf and axillary bud formed at the end of tissue that had the appearance of a pedicel. Similar phenotypes were observed among *ubp1a-2* plants (Figure 3.13B) but at a much lower frequency. More than half of the *ubp1a-2* mutant plants observed had at least a single defect in stem phylotaxy, whereas 100% of *ubp1a-3* plants had multiple phylotaxy alterations as well as wavy or shriveled rosette leaves (Figure 3.13A), thin spindly stems (Figure 3.14) and lower seed yields (not shown). When grown on MS agar plates with 1% sucrose, *ubp1a-3* shoots and roots grew significantly more slowly than wild type (Figure

3.8F) with a high amount of root curvature (Figure 3.8E). Root length and rosette leaf morphology was unaltered in *ubp1a-2*. Seedling growth arrest in the absence of sucrose did not occur in more than 25% of the *ubp1a* mutants, the same frequency observed in the wild type (Figure 3.8A). No genetic interaction was apparent the morphologic phenotypes in *ubp1a-2 ubp1c-1* double mutant plants. Although these double mutants exhibited sucrose-dependant seedling growth and dark seed characteristic of *ubp1c-1* and altered phylotaxy of the flowering stem characteristic of *ubp1a-2* (data not shown).

3.3.9 UBP1A mutants show hyposensitivity to hypoxic stress

The decreased survival of hypoxia by *ubp1c-1* led us to assay survival of *UBP1A* mutants under imposed oxygen deprivation. Ten-day-old Col-0, *ubp1c-1*, *ubp1a-1*, *ubp1a-2*, *ubp1a-3*, and *ubp1c-1 ubp1a-2* double mutant genotypes were compared. Oxygen deprivation was imposed as described and survivors were counted following 6 days of recovery. *ubp1a-3* survived significantly better than Col-0 and appeared more resistant (Figure 3.14A, 3.14B). *ubp1c-1 ubp1a-2* double mutant plant survival was intermediate to Col-0 and *ubp1c-1* suggesting a genetic interaction between these two genes during hypoxia (Figure 3.14A).

3.3.10 UBP1C-GFP and UBP1A-GFP reversibly accumulate in cytoplasmic granules in response to hypoxia

Transgenic plants overexpressing UBP1C-GFP fusion protein (Figure 3.7B) were used to visualize UBP1C subcellular localization and dynamics. In rapidly imaged seedlings, UBP1C-GFP was visible in the nucleus, with one or two regions of high concentration and throughout the cytoplasm. When cells of a cauline leaf (Figure 3.15A) or seedling (Figure 3.15E) were subjected to coverslip-imposed hypoxia by submergence in buffer over a time course under a coverslip, UBP1C-GFP aggregated into distinct granular cytoplasmic foci,

whereas the nuclear distribution did was not noticeably altered. The cytoplasmic granule size and number increased gradually over 45 min to a diameter of approximately 4-5 μm (Figure 3.15B). In the same assay, quantitative and temporal monitoring revealed that the granule number was maintained for ~10-15 min then increased over the following ~15-20 min and then leveled off (Figure 3.15C). This phenomenon was more rapid when oxygen was quickly removed by inclusion of Oxyrase in the solution used to bathe the seedling sample, a commercial antioxidant that rapidly consumes oxygen in the conversion of lactate to pyruvate and formate. Within 4 min of bathing in solution containing Oxyrase, cytoplasmic UBP1C-GFP granules formed in a manner qualitatively similar to those induced by 45 min of submergence. After 22 min of Oxyrase treatment, diffuse UBP1C fluorescence completely disappeared and very large granules formed (Figure 3.15D).

To determine if UBP1C-GFP granule formation was reversible, seedlings submerged for 60 min under a coverslip were imaged and then desubmerged for 20 min, maintaining hydration, before reimaging the same cells (Figure 3.15E). We found that granule size was dramatically reduced after the reoxygenation. Granule formation was also dependent on continued translational elongation. The formation of SGs is known to be blocked by cycloheximide, which is a chemical that halts ribosome translocation, maintaining mRNA in polysome complexes and, thereby, preventing completion of the termination phase and formation of mRNP granules. We confirmed that hypoxia-induced UBP1C-GFP aggregation was prevented by simultaneous treatment with cycloheximide. Incubation of seedlings under a coverslip in a dark, Ar purged chamber for 1 h caused formation of large UBP1C-GFP granules. In seedlings treated identically in the presence of $200 \mu\text{g mL}^{-1}$ cycloheximide, even small granules did not form (Figure 3.16A). We also tested whether UBP1C-GFP granules formed in cells of seedlings that were heat stressed for 1 hr at 37°C with or without cycloheximide treatment. Large granules were observed in heat-treated cells, but

cycloheximide- and heat-treated cells had little to no granule formation (Figure 3.16B). These results demonstrate that UBP1C-GFP granule formation in response to hypoxia caused by submergence or heat stress requires release of mRNA from translating ribosomes.

An UBP1A-GFP fusion protein was constructed (Figure 3.17A) similar to UBP1C-GFP and was stably transformed into the Col-0 ecotype genetic background. T₃ homozygous plants were used to assay the cellular localization response of UBP1A-GFP during submergence, as was UBP1C. To examine the change in size of UBP1A-GFP granules over time in response to hypoxia, an analysis was performed with maximum projection images made from z-stacks. This analysis confirmed that, similar to UBP1C-GFP the diffuse cytoplasmic UBP1A-GFP signal faded as granules formed over a 70 min time course of submergence (Figure 3.17A). The GFP signal intensity was reflective of protein abundance. The sum intensity within small granules (1-2 μ m diameter) peaked after 30 min, whereas, larger granules formed progressively over later time points (Figure 3.17B). Additionally, 20 min of reoxygenation dispersed UBP1A-GFP granules and by 30 min localization had returned to a nearly pre-stressed state (Figure 3.17A, 3.17B). The result was consistent with that of UBP1C-GFP.

3.3.11 Polyadenylated mRNA associates with UBP1C-GFP granules

To further characterize UBP1C-GFP granules, 2 h hypoxia-treated 3-day-old seedlings were rapidly fixed with formaldehyde and hybridized *in situ* with Cy3-oligo(d)T. GFP and Cy3 fluorescence were co-localized in UBP1C-GFP granules (Figure 3.18A). In an independent assay, UBP1C-GFP and *Arabidopsis* POLY A BINDING-PROTEIN 2 (PAB2)-mRFP fusion proteins were transiently transfected into *Nicotiana benthamiana* leaves by *Agrobacterium* infiltration. After 2 days, the fluorescent proteins were visualized under coverslip submergence. Granules were observed in which both PAB2 and UBP1C-GFP were

co-localized (Figure 3.18B). Together these data indicate that polyadenylated mRNA is co-localized in UBP1C-GFP stress granules that form in response to hypoxia.

3.3.12 UBP1C-bound mRNA is efficiently immunoprecipitated

A method was established for immunopurification of N-terminally FLAG epitope-tagged UBP1C in association with mRNA from 7-d-old seedlings of the genotype *ubp1c-1 35S:HF-UBP1C-20* (Figure 3.19A). This transgenic did not display evident phenotypic distinction from wildtype Col-0 when grown on medium containing 1% sucrose for 7-days under normal growth conditions. The goal was to test the hypothesis that mRNAs that are stable but poorly translated under hypoxia are sequestered in UBP1C granules. To address this, four samples of seedlings were prepared in three biological replicates. Seedlings grown under normal conditions (control light) a shift to dim light (mock dark), short-term and sub-lethal exposure to hypoxia (2 h hypoxia [2HS]), and short term reoxygenation (2HS followed by 20 min reoxygenation [+20'R]).

Pilot experiments using our standard protocol for polysome isolation by differential centrifugation (Mustroph, Juntawong, et al., 2009) indicated that a large proportion of cellular UBP1C pelleted when the crude fraction was centrifuged at 16,000 *g* (data not shown). To retain most of the UBP1C in the crude supernatant, the cell homogenate was filtered and centrifuged at a low speed (1,500 *g*) to remove the cell debris (Figure 3.19B, compare S-1.5k and P-1.5k fractions). To affinity purify FLAG-UBP1C, the 1,500 *g* supernatant was incubated with anti-FLAG conjugated magnetic beads. Beads were extensively washed prior to elution by competitive displacement with the FLAG₃ peptide. The calculated molecular mass of the FLAG-UBP1C fusion protein is 52.2 kDa. Two proteins with an apparent molecular mass of 52 kDa were detected in the 1,500 *g* crude extract (S-1.5k) and UBP1C IP fractions, but were not detectable in the unbound fraction (UNB). The lower molecular mass band was less intense than the upper band and barely detectible in the S-1.5k fraction, but

following IP was obvious. The specificity of the IP with anti-FLAG conjugated magnetic beads was tested by comparison of the results with anti-HA conjugated magnetic beads (M), which did not IP the 52 kDa protein (Figure 3.19B).

Because large cytosolic mRNP complexes can contain the small (40S) ribosomal subunit (see refs. in Ivanov and Nadezhdina, 2006), we tested whether UBP1C co-purified with ribosomal protein S6 (RPS6), a core protein of the 40S subunit. RPS6 was detectable in the unbound fraction but absent from the IP eluate, indicating that the 40S subunit is not co-immunopurified with UBP1C by this method. Heating the magnetic beads after release of UBP1C with FLAG₃ peptide in SDS loading buffer eluted additional UBP1C, indicating that release of the complex from the IgGs was improved by the presence of an ionic detergent. Neither UBP1C nor other proteins were clearly detectable in the IP eluate by silver-staining, although several non-specific protein bands were evident (Figure 3.19B).

The UBP1C IP fraction contained mRNA. The UBP1C IP and mock IP samples were evaluated for the abundance of mRNAs by qRT-PCR. *ADH1*, *UBP1A*, and *RPL37B* were selected to represent three classes of transcripts. *ADH1* mRNA is quickly induced by hypoxia and well-loaded onto polysomes, whereas *RPL37B* mRNA is stable but poorly translated until reoxygenation (Branco-Price et al., 2008). *UBP1A* contains U-rich sequences in its 3'-untranslated region predicted to bind UBP1 proteins (Lambermon et al., 2000). As a control, the non-specific association of these mRNAs with the magnetic beads was determined to be very low based on a mock IP with anti-HA conjugated beads (Figure 3.19C). When compared with total mRNA abundance, the levels of these transcripts in the UBP1C IP varied by 2 orders of magnitude, with levels of *ADH1* the lowest and *UBP1A* the highest. Transfer of seedlings to darkness or hypoxia reduced *UBP1A* association with UBP1C by approximately two-fold and increased *RPL37B* association by ~2.5-fold. UBP1C association with *UBP1A* and *RPL37B* mRNAs was reversed by 20 min of reoxygenation. *ADH1* levels in

the UBP1C IP were much lower than the other two transcripts and increased in response to hypoxia by 2-fold as compared to the 35-fold increase in the steady state abundance of this message (Figure 3.19D). These results demonstrate that UBP1C association with three representative mRNAs is selective and dynamically regulated by the environment in seedlings.

3.3.13 UBP1C co-immunopurifies with mRNAs

We hypothesized that mRNAs may co-immunopurify with UBP1C based on its known binding to RNA (Lambermon et al., 2000). To determine whether the immunopurification method yielded mRNA, UBP1C IP and mock IP samples were processed to isolate RNA and three gene transcripts were evaluated by qRT-PCR. *ADH1*, *UBP1A*, and *RPL37B* were selected as *ADH1* mRNA is quickly induced by hypoxia and well-loaded onto polysomes, *RPL37B* mRNA is stable but poorly translated under hypoxia but translated upon reoxygenation (Branco-Price et al., 2008), and *UBP1A* contains U-rich sequences in its 3'-untranslated region predicted to bind UBP1 proteins (Lambermon et al., 2000). All three mRNAs were detected in IP samples and mock IP. First, the level of each transcript in UBP1C IP fractions was compared with total mRNA abundance. In this comparison, mRNA background signal from mock anti-HA conjugated beads was similar for all three mRNAs but 10-1000 fold lower than UBP1 IP signals (Figure 3.19C). The levels of these transcripts in the UBP1C IP varied by 2 orders of magnitude, with levels of *ADH1* the lowest and *UBP1A* the highest. This indicates that *UBP1A* mRNA levels were highly enriched in the UBP1C IP, whereas *ADH1* mRNA was not. Transfer of seedlings to darkness or hypoxia reduced the relative *UBP1A* association with UBP1C by approximately two-fold and increased *RPL37B* association by ~2.5-fold. UBP1C association with *UBP1A* and *RPL37B* mRNAs was reversed by 20 min of reoxygenation, indicating that binding was conditionally regulated. *ADH1* levels in the UBP1C IP were much lower than the other two transcripts, although there

was 2-fold increase in *ADH1* mRNA in the UBP1C IP following 2 h of hypoxia there was a 35-fold increase in the steady state abundance of this message (Figure 3.19D). These results demonstrate that UBP1C association with three representative mRNAs is selective and dynamically regulated by environmental cues.

3.3.14 Comparative analysis of total mRNA, polysomal mRNA and UBP1C-associated mRNA

To evaluate the selectivity and dynamics in UBP1C-association with mRNAs at the global scale, UBP1C IP and total cellular mRNA samples from the four treatment samples were hybridized to the Affymetrix ATH1 microarray using independently generated biological replicate samples. This microarray platform assays 22,810 probe pair sets, uniquely representing 20,922 gene models. The signal values obtained with bioreplicate RNA samples were highly correlated (Pearson correlation coefficients >0.97) (Figure 3.20) indicating that the treatment and purification procedures were very reproducible. Importantly, the signal values of total mRNA samples of FLAG-UBP1C seedlings grown under control conditions in this experiment were highly correlated to *FLAG-RPL18 total* mRNA of seedlings grown under the same conditions by Branco-Price et al. (2008) (Pearson correlation coefficient 0.98), indicating that the overexpression of UBP1C had limited overall effect on the transcriptome of seedlings under these growth conditions.

To compare the data generated with the total mRNA and UBP1C IP samples, we first plotted the probe-pair set signal intensity distributions by the \log_2 expression values. This revealed a difference in shape between total RNA and UBP1C-associated mRNA populations. The UBP1C IP sample peaked at a signal \log_2 value of 4, with little skew towards higher values, except in the 2HS sample. By contrast the values for total RNA had a major peak just above a signal \log_2 value of 3 and a second broad peak between values of 6 and 8. The extensive global scale analysis of translational regulation by hypoxia and

reoxygenation performed on the *35S:FLAG-RPL18* line (Branco-Price et al., 2008) was carried out under similar conditions, providing a dataset for comparison to the UBP1C IP sample. In those samples, the intensity distribution of total and polysomal mRNA populations was slightly different, with the observed global decline in mRNAs associated with polysomes under hypoxia evident (Figure 3.21). Because of these differences in signal value distribution in these datasets, we decided that quantile normalization of signal intensities generated by hybridization of Affymetrix microarrays employed by the robust multi-array average (RMA) algorithm were not optimal. Instead, the signal intensities were normalized by use of linear scaling so that median feature values were equal between all samples.

To determine the number of mRNAs that were found in association with UBP1C, we first utilized the MAS5 “present” and “absent” calls algorithm to compare genes and samples. We applied a filter to require at least one “present” or “marginal” call out of two bioreplicates. Of the 22,810 probe pair sets, over 13,000 were “present” in at least one treatment in both total and UBP1C-IP mRNA populations (Figure 3.22A). This indicates that while a majority (8,047) was “present” in total and IP fractions of all treatments, ~3000 genes were only present in both total and IP following hypoxia treatment or hypoxia and reoxygenation. Only 3,335 mRNAs of the 17,299 transcripts detected in the total mRNA pool were never above background in any UBP1C IP sample. Interestingly, some genes were only “present” in the IP fractions (439) and most of these were only present in a single treatment (Figure 3.22A). Of the mRNAs that were observed only in total RNA (average \log_2 intensity 4.1) or only in the UBP1C IP mRNA (average \log_2 intensity 3.44), most had a low signal value compared with mRNAs which were “present” in both fractions (average \log_2 intensity 6.21) indicating that many mRNAs not “present” in both fractions are low abundance messages closer to the limit of detection.

Differential gene expression of total RNAs in response to 2 h dark, 2 h hypoxia, and 2 h hypoxia plus 20 min reoxygenation were compared by Venn diagram. We found the mock dark treatment resulted in only 24 differentially expressed genes (DEGs) (adj P-value < 0.05, |SLR|>1.0). There were 666 hypoxia-responsive DEGs and an additional 480 reoxygenation-responsive DEGs (Figure 3.22B). Meta-comparison of the DEGs identified by Branco-Price et al. (2008) (2 h hypoxia versus 2 h mock dim light control) with those of this study (2 h hypoxia versus 2 h light control) identified an overlap of 324 genes. Together with the scatter plot of the SLR comparisons genes from these two datasets (Figure 3.22D), we conclude that the two treatments produced highly similar results but were not identical, either due to distinctions in the growth and treatment conditions and/or distinctions in response of the Col-0 and FLAG-UBP1C genotypes.

3.3.15 Dynamic regulation of UBP1C association with mRNAs

To obtain an initial perspective on the effect of darkness, hypoxia and reoxygenation on the association of mRNAs with UBP1C, all transcripts that were “present” in both total and UBP1C IP populations were evaluated by comparison of UBP1C-association (UBP1C IP / Total RNA SLR) between the light sample versus the other treatments using a scatter plot of the 13,964 probe pair sets (Figure 3.23). The change in slope of the best-fit line demonstrated a global dynamic in mRNA-UBP1C association. The analysis revealed that UBP1C-association was very similar in light-grown and dark-treated seedlings (slope = 0.98, $r^2 = 0.955$). The association with UBP1C increased for many mRNAs following 2 h hypoxia (slope = 0.57, $r^2 = 0.741$), concomitant with a decrease in some highly associated mRNAs. Twenty minutes of reoxygenation was sufficient to nearly completely reverse the hypoxia-increased UBP1C association (slope = 0.8743, $r^2 = 0.877$).

We reasoned that dynamics in UBP1C association with mRNAs could parallel the change in translation status of mRNAs in response to hypoxia stress and reoxygenation

reported by Branco-Price et al. (2008) . To assess this, a meta-analysis was performed using the linear normalized data from the UBP1C IP and total mRNA described above and the immunopurified polysomal and total mRNAs quantified by Branco-Price et al. (2008). For this, we generated SLR comparisons that systematically evaluated the change in mRNA association with immunopurified UBP1C and polysome complexes relative to total mRNA levels in the same sample. Comparisons were made between UBP1C IP mRNA and total mRNA samples, separately (e.g. IP dark / IP light) and together (e.g. [IP dark / total dark] / [IP light / total light]). Similar comparisons were made for the SLR values that depict the change in total and polysomal mRNAs in the Branco-Price et al. dataset. We also assessed the similarity in the appropriate control and 2 h hypoxia mRNA signal values produced in the two datasets using Pearson correlation on the signal values (FLAG-RPL18 Total 2N vs. FLAG-UBP1C Total mRNA L = 0.976; FLAG-RPL18 Total 2H vs. FLAG-UBP1C Total mRNA H = 0.976) and SLR comparisons (0.727) (Figure 3.24; UBP1C/RPL18 total).

These 31 SLR comparisons were used in fuzzy *k*-means clustering to identify genes with similar regulation and dominant patterns of distinct regulation (Figure 3.24). The clusters were ordered based on the median value of UBP1C-association in light-grown seedlings. Therefore, the cluster number is a rough metric of UBP1C association, with cluster 1 mRNAs having the highest and cluster 20 mRNAs the lowest UBP1C association in light grown seedlings. The reversible association of UBP1C with mRNAs in response to hypoxia that is illustrated in Figure 3.23 can be seen in the cluster analysis. Hypoxia-induced and reoxygenation-reversed UBP1C association is illustrated in the IP/T and [IP/T]/[IP/T] columns. The increase in UBP1C association seen in 2 h Hypoxia / Light treatment comparison ([IP/T]/[IP/T]) by the more poorly-associated genes of clusters 15-20 is greater than that of cluster 6 for example. This trend was consistent throughout the data set, with the exception of clusters 7 and 10. These had many mRNAs which were induced and well translated

during hypoxia and increased in the UBP1C IP during hypoxia, however, as found for *ADH1* mRNA by qRT-PCR (Figure 3.19C-D) this increase was much lower relative to the increase in the total mRNA. On the other hand, clusters 3, 11, and 19 contained mRNAs that were induced by hypoxia but were not well translated until reoxygenation. For these transcripts the level of UBP1C association did increase, suggesting that their limited association with polysomes could be a cause or consequence of increased UBP1C association. Seven clusters included genes, which were relatively well translated under non-stress conditions (FLAG-RPL18 IP/T 2, 9 h mock) but were rapidly translationally repressed during hypoxia (clusters 1, 5, 8, 12, 15, 17, and 20). Low-oxygen unstable transcripts were parsed into 5 clusters (2, 9, 13, 14, and 18). Interestingly, there was no clear regulatory trend between UBP1-association and either hypoxia-mRNA stabilization or destabilization, as both stable and unstable mRNAs showed increased UBP1C association in response to the stress.

We performed a similar analysis using the SLR value comparisons generated for the 3,335 transcripts which were below the threshold of detection (absent) in the UBP1C-IP samples. Similar patterns of enrichment were observed. These genes were clustered into 5 clusters that are labeled non-associated (NA) 1-5. Although these genes were identified as “absent” in UBP1C IP samples, the heatmap (Figure 3.25) shows similar patterns as seen in the heatmap of UBP1C-associated mRNAs.

Because co-regulated genes are frequently associated with similar biological processes or molecular function, we used Gene Ontology analysis to evaluate the individual clusters. As expected, the hypoxia-responsive clusters, particularly cluster 7, were enriched in genes associated with anaerobic metabolism and stress response (Table 3.2). The clusters of genes that increased UBP1C-association during low oxygen stress included several clusters (5, 8, 12, and 17) enriched in proteins which function in translation and or ribosomal biogenesis (Table 3.3). As noted, genes in these clusters are well translated,

however, during hypoxia stress their messages are stable but dramatically decrease association with polysomes in the cells, until reoxygenation causes rapid reloading.

Transcripts of clusters 1 & 2 with high levels of UBP1C association had a number of gene ontology groups enriched (Table 3.4). These categories include extracellular proteins, plasma membrane-associated proteins, as well as endoplasmic reticulum to Golgi vesicle-mediated transport. They were also enriched in transcripts that encode nuclear-targeted proteins and/or have transcription factor activity. Others were also identified which regulate RNA metabolic processes.

3.3.16 Analysis of cluster mRNA features

As shown in the directed gene study of figure 3.19C-D, *UBP1A* transcript was highly associated with UBP1C; this gene was assigned to cluster 1 (Figure 3.23). *UBP1A* contains a U-rich 3'-UTR sequence. To determine if U-rich sequences are enriched in 5'- and/or 3'-UTR sequences of UBP1C-bound transcripts, each cluster was evaluated for nucleotide composition within the 5'-UTR, coding sequence (CDS), introns, and 3'-UTR of each mRNA. U-richness was generally higher in both the 3'-UTR and introns as compared to other regions. Introns had consistent nucleotide composition across the clusters. However, a trend of decreasing 3'-UTR U-richness was correlated with decreasing median value of UBP1C-association in each cluster (Pearson correlation = 0.924). This trend was complemented by increasing A-richness (Pearson correlation = -0.898), whereas G and C 3'-UTR content was consistent across all clusters (Figure 3.26). By contrast, the most strongly hypoxia-induced genes, enriched in cluster 7, had A-rich 5'-UTRs. This pattern was not apparent in clusters containing other hypoxia-induced genes (Figure 3.26). This analysis indicates that a bias for U in the 3'-UTR predominates in mRNAs that were enriched in the UBP1C IP fraction.

We considered that the high U bias in the 3'-UTRs of UBP1C-associated mRNAs may be more evident if a scanning window was used to determine the maximum U-richness of each 3'-UTR. By use of a window of 25 nt, the distribution of maximum richness of each cluster was compared (Figure 3.27A). This revealed that maximum U-richness was well correlated with UBP1C association (Pearson correlation = 0.929). Cluster 1 peak density was ~3.5 nt higher than that of cluster 20, with the maximum A richness declining as cluster number decreased. Thus, total 3'-UTR U-richness and maximum U-richness in a 25 nt window was strongly correlated with constitutive UBP1C binding. C and G maximum content in a 25 nt sliding window was not different between clusters. Because UBP1C association was only correlated with U-richness in the 3'-UTR and not the 5'-UTR, we investigated whether maximum U-richness is localized in a specific position along the sequences relative to the stop codon and site of polyadenylation. For most clusters, the maximum U-rich sequence windows were located from 10-60% along the length of the 3'-UTR (Figure 3.27B). This prevalence was the most notable for clusters 1-3, which had the highest constitutive levels of UBP1C association (Figure 3.27B). The analysis also revealed that the maximum A-richness peaked very close to the 3' end, in all clusters. By contrast, the maximum C-richness peaked just following the translation stop codon in all clusters (Figure 3.27B).

In an attempt to identify sequence motifs in the 3'-UTRs of the constitutive UBP1C-associated mRNAs (clusters 1 and 2), the Multiple Em for Motif Elucidation algorithm was used, with the 3'-UTRs of clusters 19 and 20 genes as reference. The analysis identified a 30 nt stretch of U's, further indicating the presence of U-rich motif segments in the most highly UBP1C-associated mRNAs. C was identified as the predominant secondary residue within the motif (Figure 3.27C). However, there was no specific motif identified nor were there residues that were invariant. These data provide the first global level demonstration of poly(U)-rich mRNAs as the target of UBP1C.

3.4 Discussion

3.4.1 *UBP1* paralogs function in multiple aspects of plant physiology

In summary, mutant and expression analysis of *A. thaliana* *UBP1* paralogs suggests unique roles for these proteins in development and potential redundant function during hypoxia response in the cytoplasm. These genes have uniquely high levels of mRNA accumulation in specific tissues and cell-types but are generally well expressed in most cells. *UBP1A* and *UBP1C* are localized to both nucleus and cytoplasm. Mutants of *UBP1A* and *UBP1C* show altered growth and development. During conditions favoring growth, *UBP1C* associates with a subset of cellular mRNAs with U-rich elements in their 3'-UTR. These transcripts encode proteins involved in multiple pathways including regulation of cell wall biosynthesis, hormone-regulated genes, and transcription factors. Under hypoxic stress, *UBP1C* and likely other RBPs, including *UBP1A*, *UBP1B*, and possibly *RBP47/45s* form cytosolic granules consistent with characterized SGs. mRNAs poorly associated with *UBP1C* under non-stress conditions show increased *UBP1C* association. These mRNAs have less U-rich 3'-UTR elements and include mRNAs encoding a wide variety of proteins. mRNAs that are bound to *UBP1C* under non-stress conditions are poised for quick sequestration into SGs upon stress. This may lead to rapid decrease in key energy consuming processes such as cell growth.

3.4.2 Relatives of *TIA-1/R* in plants

Plant RRM domain-containing protein genes include the *RBP45*, *RBP47*, *UBP1*, and *PABP* families ((Lorković and Barta, 2002; Gomez-Porrás et al., 2011). Phylogenetic analysis confirmed that the RRM of the *UBP1* proteins are more similar to those of the animal *TIA1/R* family than to the plant *RBP45* or *RBP47* families, which do not have orthologous groups in animals (Figure 3.1). The expansion and divergence of these families

in plants suggests evolution of the triple RRM proteins as a mechanism of adaptation, as strictly conserved paralogous groups within the UBP1 family are not apparent. The conservation of the RRM in UBP1 relative to TIA1/R proteins, contrasted with divergence of the domains in the RBP45s and RBP47s, suggesting conservation of specific cellular function of animal TIA/R and plant UBP1s (Figure 3.1). The maintenance of unstructured, glutamine-rich regions within these proteins that are known to function in protein aggregation further supports the hypothesis that UBP1s and TIA1/Rs are functionally conserved.

The high sequence similarity and co-expression (Figure 3.2) between UBP1 family members is consistent with the hypothesis that these protein may be functional redundant. Indeed, we provide evidence of similarity in the cytoplasmic localization dynamics during hypoxia of both UBP1C-GFP (Figure 3.16) and UBP1A-GFP (Figure 3.18). However, within UBP1s, variation of DNA and RNA regulatory sequences and in coding sequence outside of the RRM regions may have facilitated functional divergence. This is exemplified by the presence of functional siRNA854-target sites that are also semi-conserved in animal TIA-1 in *UBP1B* mRNA that is absent in *UBP1A* and *UBP1C* mRNA (McCue et al., 2012). We provide data supporting this hypothesis as well. Regulatory divergence in settings in which gene expression is not correlated but exhibits distinct responses may facilitate functional divergence (Figure 3.4, 3.5). For example, UBP1C mRNA accumulation is induced in the shoot more than 3-fold in response to 24 h darkness and this is enhanced an additional ~50% in dark plus submergence (Figure 3.3C). In addition, alleles of *UBP1C* (*ubp1c-1* and *amiR-ubp1c*), and alleles of *UBP1A* (*ubp1a-2* and *ubp1a-3*) exhibit distinct phenology plausibly indicative of unique protein function or distinct spatiotemporal regulation (Figure 3.4, 3.5, 3.14). Altogether, these data suggest that UBP1A and UBP1C are distinctly regulated and may be incompletely functionally redundant. The analysis of additional alleles and

mutant combinations for these loci is needed to better understand their overlapping versus unique functions.

Multi-purpose TIA-1/R proteins also have redundant and independent functions, and at times are antagonistic to one another (Förch and Valcárcel, 2001; Izquierdo and Valcarcel, 2007). Our data support the conclusion that UBP1C may play multiple cellular roles. The fact that overexpression failed to complement the developmental phenotypes of *ubp1c-1* (Table 3.1, Figure 3.8), but did rescue the hypersensitivity to hypoxia (Figure 3.13) in two independent insertion alleles points to rescue of UBP1C function responsible for hypoxia tolerance but not its role in development.

3.4.3 UBP1s role during stress

What are the functions of UBP1A, UBP1B, and UBP1C? TIA1/R function in multiple roles in mammalian cells, one of which is SG aggregation. Arabidopsis UBP1B and RBP47 were shown to aggregate into granules in response to heat stress (Weber et al., 2008). UBP1s, RBP45s, and RBP47s all have conserved, unstructured, and glutamine (Q)-rich domains at their N- and C-termini (Figure 3.2), reminiscent of the prion-related domain. Domains like these are responsible for protein aggregation *in vivo* in multiple examples (see Michelitsch and Weissman, 2000; Gilks et al., 2004; Decker et al., 2007). Deletion of the Q-rich N-terminus of RBP47 caused a qualitative decrease in heat-triggered granule accumulation in protoplasts (Weber et al., 2008). We have shown both UBP1A (Figure 3.18) and UBP1C aggregate during hypoxia, and at least UBP1C aggregates during heat stress (Figure 3.16, 3.17). Hypoxia stress imposes severe energetic restriction, which causes a global reduction in translational initiation and overall reduction of translating ribosomes. The altered survival of hypoxia by *ubp1c-1* and *ubp1a-3* mutants supports the conclusion that UBP1s play a role in plant survival of oxygen deprivation and reoxygenation (Figure 3.13, 3.15).

Polyadenylated RNAs are present in UBP1C granules formed under hypoxia. Lack of RNA-binding capacity in mutated UBP1B RRM domains (Weber et al., 2008), inhibited heat-triggered granule formation in a dominant negative fashion in protoplasts. Inhibition of translation elongation by cycloheximide of hypoxia- or heat-triggered SGs also prevents SG formation (Figure 3.16A, 3.16B). All of these results point to granule aggregation dependent on multiple protein-mRNA interactions, which require mRNA release from polysomes. Several studies have identified RNA binding proteins in SGs of plant cells (Weber et al., 2008). One possibility is that there is indiscriminate recruitment of RNA binding proteins and their clients to SGs. This would effectively sequester mRNA from other processes and may explain the presence of siRNA bodies containing AGO7 in SGs following heat stress (Jouannet et al., 2012). This conclusion is further supported by the non-specific increase in association of UBP1C with a large number of cellular mRNAs in response to hypoxia (Figure 3.23, 3.24). Out of all mRNAs associated with UBP1C, those with the most increased interaction were the ones which were least associated in control conditions.

The exception to this trend was the hypoxia-induced genes, grouped in cluster 7 in Figure 3.24. The median \log_2 fold increase in cluster 7 mRNA abundance in response to hypoxia was 2.18 in the FLAG-UBP1C genotype. These hypoxia-induced genes also increased in UBP1C association during the stress but to a far lesser extent relative to the induction of steady-state transcript, leading to a decrease or no change in UBP1C association of a majority of the transcripts in this cluster (median SLR -0.08) (Figure 3.24). These mRNAs were also well translated during the stress treatment suggesting that they escaped stress granule sequestration and had some feature allowing them to avoid the global translational inhibition. Cluster 7 had notably high A composition within their 5'-UTRs (Figure 3.26). Kawaguchi and Bailey-Serres (2005) reported that A-richness upstream of the start codon was typical of mRNAs well translated during drought. In contrast, the steady

state abundance of genes of cluster 3 increased by hypoxia (median SLR 0.22) but did not associate well with polysomes until reoxygenation. Cluster 3 5'-UTRs were also not A-rich. This group maintains or increases association with UBP1C during hypoxia (median SLR 0.24, Figure 3.24).

The cytoplasmic relocation of both UBP1C and UBP1A in response to hypoxia was dynamic. The velocity and severity of oxygen deprivation directly correlated with granule size and number (Figure 3.16B, 3.16C, 3.16D). The transient, dynamic nature of these large complexes is also illustrated by their rapid dissolution following 20 min reoxygenation (Figure 3.16E, 3.18) and the reversal of increases in UBP1C mRNA-binding triggered by hypoxia. The rapid regulation of UBP1C granule aggregation leads to the hypothesis that it is regulated by signal transduction directly linked to sensing of oxygen concentration or energy status. The evolution of such a dynamic system demonstrates the selective pressure for efficient cellular management of cytoplasmic mRNA populations in reacting to environmental change and energy or growth limitation. A rapid, dynamic response to oxygen deprivation is essential to conserve sufficient cellular energy in order to alter programs of gene expression necessary to curtail growth and shift metabolism driven by oxidative phosphorylation to one sustained by substrate level phosphorylation and fermentation. If UBP1C facilitates this metabolic transition, it might explain the decreased survival of the *ubp1c-1* mutant (Figure 3.13, 3.15).

The biological requirement of fine temporal control of translation and SG aggregation in hypoxia is similar to that required during mammalian inflammatory response in which post-transcriptional regulons are quickly activated and deactivated as discussed by Anderson (2010). In the case of inflammation, fine control of gene expression is required to prevent excessive tissue damage. Deregulated inflammatory response leads to multiple disorders and disease in mammals.

Intriguingly, *ubp1c-1* hypersensitivity is suppressed by the *ubp1a-2* allele. More work is needed to understand this, as *ubp1a-2* alone was also hypersensitive. A second allele, *ubp1a-3*, with more severe phenologic alterations survived mild hypoxia with no growth reduction in a treatment that reduced growth of wild-type plants by 40% (data not shown). Preliminary qRT-PCR data indicate that the abundance of *UBP1A* mRNA was not reduced in 7-day-old seedlings of *ubp1a-2* or *ubp1a-3*. Phenotypes of these alleles might result from misexpression or conditional expression of *UBP1A*. In fact, increased survival may be due to higher *UBP1A* mRNA in stressed tissue. This might explain the enhanced survival of double mutant *ubp1c-1 ubp1a-2* and *ubp1a-3* (Figure 3.15). These experiments might be accomplished when seed for these genotypes produced at the same time becomes available.

3.4.4 UBPC-associated RNAs

Immunopurified UBPC was associated with mRNAs, but its co-purification with proteins has not been tested. Based on the pronounced aggregation of UBPC and the absence of co-purifying proteins at similar stoichiometry, UBPC apparently does not interact with other RNPs or ribosomes in a multisubunit complex. In the future, mass spectrometric proteomic analysis of the UBPC IP could be used to determine if the mRNA-UBPC IP includes other proteins. In mammals, the small ribosomal subunit has been reported to colocalize with SG (Kedersha et al., 2002), however López de Silanes et al. (2005) also performed TIA-1 immunopurification and found no evidence of RPS6 co-purification under control or heat stress conditions. In the same study, the overall RIP-Chip transcript signals increased under heat stress in a manner similar to our result.

The considerable variation in number of mRNAs associated with UBPC under control and hypoxia stress conditions leads to the question whether the protein serves distinct functions under the two conditions. The variation in phenotypes in common with the *ubp1c-1* mutants and *amiRNA-ubp1c* lines also hints at distinct function under different

conditions or based on overall abundance of the protein. This is also suggested by the variation in phenotypic effects in the UBP1C overexpression lines. In the absence of stress, UBP1C is associated with mRNAs which are predominately transcription factors or are targeted for vesicle-mediated export and have U-rich stretches within their 3'-UTR (Table 3.4). The fact that 3'-UTR U-richness correlates with UBP1C association and not with that of introns in this data set suggests that UBP1C is predominantly associated with the 3'-UTRs of mature, spliced transcripts during conditions favoring growth (Figure 3.26).

In the presence of stress, the U-richness is less of a determinant of association. It is possible that the increase in number of mRNAs that bind to the protein under low oxygen stress could reflect a change in affinity to RNA in general. This could be due to a post-transcriptional modification or change in oxidation state, which is rapidly reversed upon reoxygenation. This might be investigated by further study of UBP1C protein using mass spectrometry or 2D-isoelectric focusing gel electrophoresis to discern condition specific isoforms due to post-translational modifications such as phosphorylation.

Among highly UBP1C-associated genes of clusters 1 and 2, enriched gene ontology categories that involved proteins targeted to the plasma membrane or exported from the cell included “anchored to membrane”, “plant cell wall”, “cell wall constituent”, “cellulose synthase activity”, “transporter activity”, “apoplast”, “ER to Golgi vesicle-mediated transport”, and “cell-cell signaling” (Table 3.4). These two clusters include many members of the hydroxyproline-rich glycoprotein superfamily. If we count genes from this family from clusters 1-4, we find 16 out of 39 proline-rich extensin or extensin-like genes, 27 from 89 other hydroxyproline-rich glycoprotein genes, and 28 arabinogalactan protein (AGP) genes out a total of 50 AGP-like or FASCICLIN-like AGP proteins in Arabidopsis. In addition, 27 nodulin or early-nodulin-like protein genes out of 85 were also present. Nodulin-like genes are predicted to be membrane glycosylphosphatidylinositol-anchored chimeric AGPs (Mashiguchi et al., 2009). AGPs have

been suggested to be involved in numerous biological processes from cell expansion and division, abiotic stress resistance, control of development, and hormone signaling among others (Tan et al., 2012).

It is possible that the first plant UBP1 family member investigated was PRP-binding protein (BP) in *Phaseolus vulgaris* by Zhang and Mehdy (1994). PRP-BP is a 50 kD protein that binds a 3'-UTR, U-rich sequence of the bean *Proline-rich protein (PvPRP)* gene, a homolog of *Arabidopsis thaliana* (AGP) 31 (Liu and Mehdy, 2007). However, *AtAGP31* does not have a U-rich 3'-UTR sequence and was located in cluster 18 in our data set. PRP-BP is redox-regulated and Zhang and Mehdy (1994) speculated that PRP-BP may play a role in destabilizing *PvPRP* in response to a fungal elicitor. In light of our results, this seems unlikely. However, redox-mediated post-translational modification has yet to be examined in *AtUBP1* proteins and may be a means to control UBP1 aggregation and/or RNA-binding activity. Indeed, western blot analysis of UBP1C revealed a double band around 52 kDa which may represent a post-translationally modified form of the protein (Figure 3.19A). Other genes which contributed to enrichment of the above mentioned GO categories include 11 of 31 cellulose synthase genes, 2 apyrase genes, and 44 from a total of 214 leucine-rich repeat family or protein kinase family genes.

Furthermore, genes regulated by auxin are enriched in clusters 1 and 2 and are known to be post-transcriptionally regulated. These include *AUXIN RESPONSE FACTOR (ARF) 17*, (others are *ARF12*, *ARF16*, *ARF18*, *AUXIN-INDUCED 13 (AUX/IAA13)*, *AUXIN-INDUCED IN ROOT CULTURES (AIR)*) and 12 and 23 out of 56 *SMALL AUXIN-UP RESPONSE (SAUR)-like* family genes. *ARF17* is believed to be targeted by miR160 through *AGO1* to regulate adventitious root formation in the hypocotyl (Vaucheret et al., 2004; Sorin et al., 2005). *SAUR* genes from tobacco have been reported to have conserved 3'-UTR element distinct from U-rich elements, called the DST for downstream element (McClure et

al., 1989) which is believed to cause the shortest mRNA half lives known in plants - on the order of 10 to 20 min (Gil and Green, 1996). However, the only direct evidence supporting this is from reporter assays using the complete SAUR-AC1 3'-UTR. These examples illustrate that UBP1C targets may require fine control of expression and may be regulated by multiple layers of post-transcriptional regulation to this end.

Besides these GO categories, 58 genes with DNA-binding domains or known transcription factors were also present in clusters 1 and 2. The following ribosomal proteins were present: L10, L13, L14p/L23e, L39, L7Ae, L37A, S11, S15A, S27, S5, S18; as well as 16 other RNA binding proteins. Cyclin A3;4 and B1;4, key cellular regulators of the cell cycle, 6 of 14 Arabidopsis CLAVATA3/ESR-RELATED proteins which may play regulatory roles in developmental were present, and 13% of genes from clusters 1 and 2 are proteins of unknown function.

3.4.5 Phenotypic effects of UBP1C mutations

While the molecular effect of UBP1C-mRNA binding is unclear, we have shown that UBP1C and UBP1A play a role regulating normal development and facilitate acclimation to stress. Disruption of the *UBP1C* gene caused numerous morphologic phenotypes, suggesting that the gene is involved in multiple pathways in normal growth including seed development, cell division, carbon metabolism, cytokinin response, and shoot gravitropism (Figure 3.8, 3.9, 3.10). These may be related phenotypes at the molecular level; however they also may represent altered expression of distinct mRNA targets of regulation. Cytokinin sensing and signal transduction is required for seedling development.

Loss-of-function mutations in the transcription factor *STIMPY/WOX9* leads to remarkably similar phenotypes, including sucrose-rescued growth arrest (Wu et al., 2005). In these mutants, seedling development halts at the same stage as *ubp1c-1*, starch granules

are absent, shoots are agravitropic, and seeds exhibit a dark color similar to *ubp1c-1* (data not shown). This last phenotype was not reported in the literature. Nevertheless, *UBP1C* and *STIMPY* mutants are not allelic (data not shown). Lack of starch granules would lead to loss of shoot gravity sensing. This phenotype has been linked to expression of cytokinin signal transducers. *STIMPY* mutants have an altered response to kinetin similar to *ubp1c-1*. The cytokinin receptor triple mutant *ahk2-2 ahk3-3 cre1-12* also has a similar growth arrest phenotype and these genes are expressed at a low level in the *stip-1* loss of function mutant. The authors of this work also showed that *stip-1* meristems are small and malformed in arrested seedlings and they postulate that sucrose drives cell division by enhancing cyclin D1 expression. This link points toward *UBP1C* as a positive regulator of cytokinin signaling. Sugar biosynthesis must be sufficient to maintain cellular metabolism because the plants survive for almost 4 weeks without exogenous sugar. Why there is insufficient endogenous sucrose to drive cell division or form starch granules is unknown. The mechanism of this regulation could be the focus of future work.

3.4.6 Working model of *UBP1* function and hypoxia-triggered SGs

We propose a model in which *UBP1* controls multiple aspects of gene expression within distinct developmental and environmental contexts. This includes (1) the previously demonstrated role of *UBP1* in splicing efficiency and (2) the sequestration of mRNAs targeted for conservation or degradation during hypoxia (Figure 3.28).

The distinct transcript accumulation patterns and phenotypes of *ubp1a* and *ubp1c* mutants as well as *UBP1C* overexpression lines support the conclusion that these genes non-redundantly control aspects of development, either through their role in splicing or mRNA sequestration. The distinct low-oxygen sensitivities of *ubp1a* and *ubp1c* mutants as well as *UBP1C* overexpression lines, and the dynamic cytoplasmic re-localization in response to hypoxia and reoxygenation, support a role for *UBP1A* and *UBP1C* in plant

hypoxia-triggered SG formation and hypoxia survival. Based on the immunopurification of UBP1C-associated mRNAs, transcripts with high U-rich 3'-UTRs preferentially associate with UBP1C under normal growth conditions.

UBP1C association with other mRNAs increases in response to hypoxia and was reduced by reoxygenation, consistent with the dynamics of cytosolic UBP1C granule formation. Increased mRNA association with UBP1C was observed for both poorly translated and unstable mRNAs. Hypoxia-induced transcripts were associated with this protein, but at a disproportionately low level as compared to translationally repressed and unstable mRNAs. Therefore we propose two cytoplasmic functions of UBP1C-mRNA association, (1) to poise the cell for rapid sequestration of the transcripts bound by UBP1C prior to hypoxia. As these transcripts are highly enriched in cell wall protein components, membrane-bound structural proteins, ribosomal proteins, and hormone-regulated transcription factors, this association could rapidly curtail energy consumption upon transfer to low oxygen; and (2) to promote the sequestration and translational repression or degradation of mRNAs with increased association with UBP1C upon hypoxia. We propose that in response reoxygenation UBP1C RNA-sequestration in granules is rapidly reversed.

Based on this model, we hypothesize a number of things. (1) Upstream hypoxia-initiated signal transduction leads to altered UBP1 protein- and RNA-binding activity by post-translational modification, enhancing non-specific RNA-binding and aggregation of UBP1 proteins and most mRNAs. (2) UBP1A, UBP1B, and UBP1C redundantly bind U-rich 3'UTRs of mature 3'UTR in the cytoplasm. (3) RNAs that encode stress-relevant polypeptides escape general translational repression and SG sequestration, which drives selective translation of these mRNAs. (4) Unstable mRNAs sequestered into SG are transient and are transferred to processing bodies. (5) UBP1C may participate in spatiotemporal control of translation as many of its targets are nuclear, membrane, or cell wall localized.

3.5 Materials and Methods

3.5.1 Genetic material and growth conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) (Arabidopsis Biological Resource Center) was used as the wild type. *ubp1c-1* (SAIL_363_G02; AT3G14100), *ubp1a-2* (SAIL_1185_C11) and *ubp1a-3* (SALK_068668; AT1G54080) were obtained from the Arabidopsis Biological Resource Center, Ohio State University and confirmed to be homozygous. The *Arabidopsis thaliana* Col-0 *35S:HIS₆FLAG-RPL18* 12-2-4 genotype has been described by Zanetti et al. (2005) and is the same as that used by Branco-Price et al. (2008). *stip-1* was a gift from Xuelin Wu, University of Southern California.

Seeds were sterilized by soaking in 70% (v/v) ethanol for 5 min, followed by 1% (v/v) sodium hypochlorite 5 min and rinsing 5 times with sterile distilled, deionized H₂O (ddH₂O). Seeds were stratified in water at 4°C in the dark for 3-5 d. Seeds were transferred in a sterile laminar flow hood to plastic Petri dishes (10 cm diameter) containing autoclaved solid medium composed of 0.4% (w/v) Phytigel (Sigma, St. Louis, MO), 0.43% Murashige and Skoog Salts (MS salts; Caisson Labs, North Ogden, Utah), adjusted with potassium hydroxide to pH 5.75, and with or without 29 mM (1% w/v) sucrose. Plates were positioned vertically in racks of a growth chamber (Percival Scientific Inc., Perry, IA; model CU36L5C8) with a 16 h light (~120 μE s⁻¹ m⁻²), 8 h dark cycle at 23°C. To evaluate growth responses to kinetin (Sigma, Lot 80K4133), seedlings were grown for 5 d on solid MS medium containing 29 mM sucrose and transferred with forceps to plates with the same medium additionally containing kinetin. For selection of kanamycin resistant transgenics, 50 μg mL⁻¹ kanamycin was additionally included in the medium by dilution from a 50 mg mL⁻¹ stock solution (prepared by dissolving kanamycin (Sigma, catalog no. K-4000) in water) to the autoclaved medium after cooling to ~50°C.

Plants were grown in soil (Sunshine LC1 soil mix, JM McConkey, Sumner, Washington) with 1.87 g L⁻¹ Marathon insecticide (Crop Production Services, Riverside, California) and 1.4 g L⁻¹ osmocote 14-14-14 fertilizer (Scotts, Marysville, OH), and watered every other day under a diurnal cycle of 16 h light (50 μE s⁻¹ m⁻²), 8 h dark at 23°C. Alternatively, sterilized seed were stratified at 4 °C in darkness for 3 d. Imbibed seed were transferred to solid MS media [0.43 % (w/v) Murashige Skoog (MS) salts (Caisson, North Logan, UT)], 0.4 % (w/v) Phytigel (Sigma), 29 mM (1 % w/v) sucrose, pH 5.7) in petri plates. Plates were positioned vertically in a growth chamber (Model # CU36L5, Percival Scientific, Perry, IA) under a long day cycle (16 h light at ~125 μE m⁻² s⁻¹, 8 h darkness) at 23 °C for the time indicated.

Plants were photographed using a Rebel T1i camera (Canon U.S.A., Lake Success, New York). Root lengths were measured from photographic images using ImageJ software (rsbweb.nih.gov/ij/).

3.5.2 Generation of transgenic plants

The *35S:UBP1C-GFP*, *35S:UBP1A-GFP*, and *35S:HIS₆FLAG-UBP1C* genes were constructed using coding sequence (CDS) generated from the full-length cDNA from the loci At3g14100 and At1g54080 for *UBP1C* and *UBP1A*, respectively. Genomic constructs were made similarly by amplification from genomic DNA (gDNA). Total RNA was extracted from 1-wk-old wild type seedlings (Col-0) using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). Total RNA was reverse transcribed into complementary DNA (cDNA) by Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) and primed with oligo-(d)T following the manufacturer's directions. The *UBP1C* CDS was amplified from cDNA using Herculase II Fusion polymerase (Agilent Technologies, Santa Clara, CA) according to the manufacturer's recommendation using the following primers for *UBP1C* CDS from cDNA: 5'-CAC CAT GCA GAA TCC GAG ACT GAA GCA ACA TC-3' and 5'-TTA CTG ATA GT ACA TGA GTT GCT

GTG CCG-3'. Primers used to amplify *UBP1A* CDS from cDNA were: 5'-CACC ATG CAG AAT CAA AGG CTT ATT AAG CAG CAA CA-3' and 5'-CTG ATA GTA CAT GAG CTG CTG ATG GGC-3'. Primers used to amplify the genomic region of *UBP1C* to make *proUBP1C::genomicUBP1C* from gDNA were: 5'-CACC CAA AGG CAA AGC TGT CTG TT-3' and 5'-TTA CTG ATA GT ACA TGA GTT GCT GTG CCG-3'. Primers used to make *proUBP1C::genomicUBP1C::UBP1C_{3'UTR}* from gDNA were: 5'-CACC CAA AGG CAA AGC TGT CTG TT-3' and 5'-TGC TCG GTT AGA GAC TGA CTG A-3'. For generation of the *proUBP1C::FLAG-genomicUBP1C::UBP1C_{3'UTR}* from gDNA, the FLAG peptide sequence was inserted by multi-template PCR of two overlapping products generated using the following primer pairs: 5'-CACC CAA AGG CAA AGC TGT CTG TT-3' and 5'-TCC CTT ATC ATC ATC ATC CTT ATA ATC ACC ATT CTG CAT TTT TTT TCC CC-3'; and 5'-GAT ATG TGA AAG GAG TGG GGA GAC GAA CCG GAC GAA CA-3' and 5'-TGC TCG GTT AGA GAC TGA CTG A-3'.

PCR products were separated on a 1.0% (w/v) agarose Tris-acetate-EDTA (TAE) gel (40 mM Tris-HCl [pH 8.4], 20 mM sodium acetate, 1 mM Ethylene-diamine-tetra-acetic acid) containing 0.01% (w/v) ethidium bromide by electrophoresis in TAE buffer at 110 V constant voltage for 40 min. Following electrophoresis, the amplified DNA was cut from the gel and extracted using the QIAquick Gel Extraction Kit (Qiagen) following the prescribed steps. The amplicons were cloned into the pENTR/D-Topo vector (Invitrogen) and transformed into Oneshot Top10 chemically competent cells (Invitrogen) according to the manufacturer's protocols. Plasmid from positive clones was extracted from 5 mL of bacterial culture grown in Luria broth (LB) medium (1% (w/v) peptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride) plus 50 $\mu\text{g mL}^{-1}$ kanamycin overnight by use of the Qiagen Miniprep Spin kit (Qiagen).

The sequencing of clone inserts was performed by the Genomics Core Instrumentation Facility of the Institute for Integrative Genome Biology at the University of California, Riverside. Only clones with sequences that match the public, full-length cDNA or genomic sequences (www.arabidopsis.org) were used for gene construction.

Two binary expression vectors were constructed for cloning of genomic DNA. GATA-OCS_T(nptII) and GATA-FH-OCS_T(nptII) vectors were constructed from p35S:HF-GATA and p35S:GATA-FH vectors (Mustroph et al., 2010). GATA-OCS_T(nptII) was constructed by removal of the promoter and HIS₆-FLAG tag from 35S:HF-GATA via partial digest of EcoRI and BsrGI. The FLAG-HIS₆ tag was added to this vector to make GATA-FH-OCS_T(nptII) by ligating the Sall/XmaI 907 bp fragment from the p35S:GATA-FH vector into the Sall/XmaI restriction enzyme digested GATA-OCS_T(nptII) vector.

Plant gene regions were cloned into T-DNA binary vectors. cDNA were cloned into the pEarleygate 103 vector (Earley et al., 2006) or expression vector (Mustroph et al., 2010). *proUBP1C::genomicUBP1C* was cloned into the gateway-FH-OCS_T (nptII) vector, *proUBP1C::genomicUBP1C::UBP1C_{3'UTR}* and *proUBP1C::FLAG-genomicUBP1C::UBP1C_{3'UTR}* were cloned into the gateway-OCS_T (nptII) vector. Cloning was performed by LR Clonase reaction (Invitrogen) according to the manufacturer's prescribed steps and transformed into Oneshot Top10 chemically competent cells for plasmid amplification. Amplified plasmid was electroporated into *Agrobacterium tumefaciens* GV3101 cells. *Agrobacterium*-mediated transformation of *A. thaliana* was performed according the floral dip method (Clough and Bent, 1998). Briefly, binary vector-transformed *A. tumefaciens* cells were grown to stationary phase in 500 mL LB with 50 µg mL⁻¹ the appropriate antibiotic for plasmid selection. Cells were pelleted by centrifugation at 5,000 *g* for 15 min and resuspended in 250 mL 5% (w/v) sucrose and 0.05% (v/v) Silwet-L77® (Crompton, Greenwich, CT). Inflorescences were dipped into resuspended *Agrobacterium* and covered

for 24 h. Seed was harvested after maturation. Transgenic T₁ seed was selected on the appropriate antibiotic (50 µg mL⁻¹ kanamycin or 10 µg mL⁻¹ ammonium glyphosate) on solid MS medium in plates. Both Col-0 and *ubp1c-1* genotypes were used for transformation, as indicated.

3.5.3 Gene alignment and analysis of evolutionary relationships

TIA-1 and TIAR-related proteins were identified by use of the basic local alignment search tool (BLAST) (Altschul et al., 1990) using representative taxa of mammals, birds, amphibians, insects, monocotyledonous and dicotyledonous plants and the moss *Physcomitrella patens*. Oligouridylate-binding protein 1 (UBP1), RNA binding protein 45 (RBP45), RNA-binding protein 47 (RBP47), and polyadenylate binding protein (PABP) gene family members were selected. PABPs have four RRM domains, the first of which did not align well with any other protein. Sequences were aligned using the Muscle algorithm in the Mega 5.0 software (Edgar, 2004; Tamura et al., 2007). The alignment was used to infer evolutionary relationship by the Neighbor-Joining method (Saitou and Nei, 1987). The degree of evolutionary distance was computed using the JTT matrix-based method to generate the number of amino acid substitutions per site (units) (Felsenstein, 1985). The analysis involved 99 protein coding (amino acid) sequences. All positions containing gaps and missing data were eliminated. There were a total of 176 positions used from the final alignment.

Expression maps and co-expression analysis were generated by use of online tools. The co-expression network was created using the Atted (atted.jp) network drawer and the co-correlation scatter plot was made on the *Arabidopsis* co-expression data mining tool (<http://www.arabidopsis.leeds.ac.uk/act/index.php>) (Manfield et al., 2006). The cell-type specific enrichment maps were generated from the data of Mustroph et al. (Mustroph, Zanetti, et al., 2009) on the EFP browser (efp.ucr.edu) in Absolute mode. The Genevestigator 4.0

(genevestigator.org) (Zimmermann et al., 2004) Anatomy tool was used to generate heatmaps of organ- and tissue-specific expression from 6,458 samples probed with Affymetrix ATH1 Genechips.

3.5.4 Stress treatments

Seedlings grown in vertical orientation for 7 to 14 d on solid MS medium containing 29 mM sucrose were treated with mixed gases in humidified chambers. This was accomplished by placing the plates oriented in the vertical orientation in sealed 13.1 L Plexiglas chambers, which were purged with 99.99% Ar(g), at the specified flow regime. In this case, seedlings were deprived of O₂ as well as CO₂. Alternatively, treatment was with 2% O₂, 370 ppm CO₂, in a balance of N₂. Treatments were performed either immediately following the end of the subjective day or night as indicated. Treatments were performed at room temperature and under 1-3 $\mu\text{E s}^{-1} \text{m}^{-2}$ light or in the dark. Control treatments were performed on seedlings placed in open Plexiglas chambers under the same light, humidity and temperature conditions. Stress survivors were those plants that demonstrated shoot apical meristem growth after 4 d of recovery from the stress in normal growth conditions.

For microscopy experiments, a portion or the entire seedling was subjected to oxygen deprivation by submergence in 0.106% (w/v) MS medium and carefully covering with a glass coverslip avoiding trapping of bubbles. Reoxygenation was facilitated by removal of the coverslip or desubmergence by careful aspiration of liquid surrounding the seedling and leaving sufficient moisture to prevent drying. Hypoxia treatment included submergence as described plus incubation in a humidified, Ar-purged chamber. Chemical treatments included: addition of 200 ng μL^{-1} cycloheximide to the MS medium with 0.4% (v/v) dimethyl sulfoxide (DMSO) to inhibit elongation of translation or addition of 8 U/mL Oxyrase (Lot PEC00722, Oxyrase Inc., Mansfield, OH) to the MS medium to rapidly remove dissolved oxygen. Heat stress was imposed on seedlings by incubation in 250 μL 0.106% (w/v) MS

solution (DMSO) in the presence or absence of 200 ng μL^{-1} cycloheximide on a standard microscope slide placed in a closed, humidified box at 20°C or 41°C for 2 h. Confocal fluorescence images were taken within 5 min of coverslip application unless otherwise indicated.

Seedling treatments for microarray studies were initiated at the end of subjective day (Zeitgeber time 16) after 7 d of growth on solid MS medium plus sucrose (29 mM). Control light plants were maintained in light. Control dark plants were transferred to near darkness ($<0.05 \mu\text{E m}^{-2} \text{s}^{-1}$) for 2 h in a humidified 13.5 L Plexiglas chamber open to ambient air (dark control, D). For hypoxia treatment (H), plants were placed under the same low light and deprived of oxygen for 2 h by sealing the chamber and continuously flushing it with humidified 99.99% Ar(g) at a rate of 50 mL s^{-1} for 2 h at 20°C. For 2 h hypoxia followed by 20 min reoxygenation (R), reoxygenation was facilitated by removing the top of the chamber and fanning with a rigid plastic fan (~25 x 40 cm) immediately and again after 10 min for 1 minute. All seedlings were harvested rapidly into $\text{N}_2(l)$, ground to a powder and stored at -80°C.

3.5.5 Immunoprecipitation of UBPC-RNA complexes

FLAG-tagged proteins were isolated by immunoprecipitation (IP). All steps were carried out on ice unless indicated otherwise. For each sample, two aliquots of approximately 100 mg frozen ground powder tissue were thawed in 1.5 mL ice cold RNP extraction buffer modified from Oeffinger et al. (Oeffinger et al., 2007) for plant extracts (200 mM Tris [pH 9.0], 110 mM potassium acetate, 0.5% (v/v) Triton-X100, 0.1% (v/v) Tween-20, 1.5% (v/v) protease inhibitor cocktail for plant cell and tissues extracts (Sigma P9599, St. Louis, MO), 5 mM dithiothreitol, 0.1% (v/v) RNaseOUT (Invitrogen, Carlsbad, CA)). The sample tube was briefly vortexed and the extract filtered through two layers of sterilized Miracloth (EMD Millipore, Billerica, MA) into a clean 1.5 mL microfuge tube. The filtrate was centrifuged in a

swinging bucket clinical centrifuge at approximately 1500g for 2 min at 4°C. The supernatant (S-1.5K) was removed and a sample aliquoted for analysis by standard SDS-polyacrylamide gel electrophoresis (PAGE) (15 µL), immunoblotting (IB) (15 µL), and RNA extraction (100 µL per replicate sample). The remaining S-1.5K was transferred to a 1.5 mL microfuge tube containing antibody-conjugated beads for the IP. The pellet (P-1.5K) was saved for PAGE and IB analysis.

Protein G-conjugated paramagnetic beads (Dynabeads, Invitrogen, Carlsbad, CA) were prepared according to the manufacturer's directions. Briefly, for each 100 mg of tissue processed, 100 µL of resuspended Dynabeads were magnetically separated from the storage solution and resuspended in 400 µL α-FLAG M2 monoclonal antibody (Sigma F1804) in phosphate buffered saline Tween (PBST; 137 mM sodium chloride, 2.7 mM potassium chloride, 10 mM sodium phosphate dibasic, 2 mM sodium phosphate monobasic [pH 7.4], 0.02% (v/v) Tween-20). The final FLAG M2 monoclonal antibody concentration was 25 ng µL⁻¹ α-FLAG M2. The bead antibody mixture was incubated for 1 h on a rocker at 20-23°C. Following antibody conjugation to the beads, excess antibody was removed by magnetic separation and washing.

IP of FLAG-UBP1C-containing complexes was accomplished by rocking the S-1.5K-Dynabead slurry at 4 °C for 1.5 h. Following incubation, the Dynabeads were magnetically separated, 30 µL supernatant was saved as the unbound fraction, and the beads were washed with 0.75 mL RNP extraction buffer for 5 min with rocking at 4°C and magnetically collected. Bead washing steps were repeated five more times. FLAG-UBP1C was eluted in 100 µL RNP extraction buffer plus 400 ng µL⁻¹ 3X Flag® Peptide (Sigma F4799) at 4°C for 30 min with rocking. Eluates of replicate IPs were combined for a total of 200 µL. A fraction of the eluate (20 µL) was saved for PAGE and IB analysis and 180 µL for RNA extraction. Used beads were saved for PAGE and IB analysis as the post-IP bead fraction. A mock IP

was performed on dark-treated tissue in exactly the same manner as other samples, except an equivalent amount of α -HA monoclonal antibody clone HA-7 (Sigma H9658) was used in the place of anti-FLAG antibody.

3.5.6 RNA purification and quantitative real-time PCR

For direct RNA extraction, three individual seedlings of each genotype were ground in liquid N₂ using a micropestle. RNA was extracted immediately using 300 μ L Trizol (Invitrogen). Extracted RNA was treated with DNase I (RNase-Free DNase Set, Qiagen) to eliminate genomic DNA and cleaned-up using the RNeasy Mini Kit as directed (Qiagen). The RNA concentration was determined by use of a Nanodrop spectrophotometer (Thermo Scientific, Wilmington, DE). cDNA was synthesized as described above from 120 ng of total RNA in a 20 μ L volume.

For RNA extracted from IP eluate (UBP1C IP) or S-1.5K fraction (total RNA), 1 mL Trizol reagent (Invitrogen) was used according to the manufacturer's protocol. Extracted RNA was further processed as above, except that the RNA was eluted with two volumes of 50 μ L RNase-free water and concentrated by precipitation in 20 μ g glycogen, 300 mM sodium acetate pH 5.2, and 2.5 volumes 100% (v/v) ethanol, followed by mixing, incubation at -20°C for 2 h and centrifugation for 30 min at 16,000 g at 4°C. RNA pellets were solubilized in RNase-free water at a concentration of 100 ng μ L⁻¹ total or 5-10 ng μ L⁻¹ IP RNA. Six percent of the RNA extracted from the total (S-1.5k) and IP fractions was reverse transcribed into complementary DNA (cDNA) by Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) primed with oligo-(d)T following the manufacturer's directions in a 20 μ L volume.

To perform quantitative real-time reverse transcriptase PCR (qRT-PCR), the cDNA reaction mixture was diluted five times in distilled water and 2 μ L were used for each reaction.

A reaction mixture was made using iQ™ SYBR® Green Supermix (Biorad) according to the manufacturer's instructions with primer concentrations of 0.5 μM each. Thermocycling and fluorescence measurements were accomplished on a C1000 Thermocycler (Biorad, Hercules, CA) equipped with a CFX-96 Real-Time PCR Detection System (Biorad). Thermocycle parameters were: 3 min at 95°C, 10 s at 95°C, 15 s at 55°C, 30 s at 72°C with 40 cycles; followed by melt curve analysis from 95°C to 65°C in 0.5°C increments at 5 s each. Relative transcript levels were analyzed using the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001) with the *UBIQUITIN CONJUGASE 5* (UBC5) or *RIBOSOMAL PROTEIN L37B* (*RPL37B*, At1g52300) gene as reference for comparison of total RNA levels. IP mRNA levels were compared against total mRNA abundance for each treatment. Primer pairs used for qRT-PCR were *UBC5*: 5'-TGC GTG ATC GTC CTA CCT ATG AAC A-3' and 5'-TCT CCG GTT TAG CCA TCT ATT AAG C-3'; *RPL37B*: 5'-TCC GCT TCA ATG ACT AAG GGA ACG-3' and 5'-GCG AGG CAC ATT GCG AAG ATA C-3'; *UBP1C*: 5'-GAT CCG CTG CTC TGG CTA TAC TGT C and 5'-CGC ATC CGA ACA ACT GGA AAA GA-3'; *UBP1A*: 5'-GAA GAT GCC CCT GAA AAC AAT CCT C-3' and 5'-TCC CCA GGA ACA CCT TAT CTG TCT G -3'; *ADH1*: 5'-TCT ACT GGG TTA GGA GCA AC-3' and 5'-TTG ATT TCC GAG AAT GGC AC-3'. These primer pairs had efficiencies of 100%, 103.6%, 102.3%, and 106%, respectively.

3.5.7 Microarray hybridization and data analysis

RNA samples were submitted to the Genomics Core Facility of the Institute for Integrative Genome Biology, University of California, Riverside, for quality monitoring and GeneChip hybridization. One μL of each sample was evaluated by use of the Bioanalyzer 2100 (Agilent, Technologies, Santa Clara, CA) RNA 6000 Nanochip (for total RNA) or RNA 6000 Picochip (for UBP1C IP RNA). For microarray analysis, 150 ng total and 15 ng UBP1C IP RNA were used to synthesize biotinylated, amplified RNA (aRNA) via 15 rounds of *in vitro* transcription using the 3' IVT Express Kit (Affymetrix, Santa Clara, CA). aRNA was

fragmented and probed using the Arabidopsis ATH1 Genome Array (Lot 4160460; Affymetrix). Hybridizations were performed with 15 µg fragmented aRNA at 45°C for 16 h on a rotating platform.

Following hybridization and scanning, raw data in .CEL files were extracted and pre-processed using the Bioconductor package (Gentleman et al., 2004) of the R software for statistical computing (R. Development Core Team, 2012). The analysis included .CEL files generated in Branco-Price et al. (2008) for total and polysomal mRNA (Gene Expression Omnibus Record: GSE9719). Microarray features of the above data sets were background corrected, normalized, and probe set intensities were summarized using the *affy* (Gautier et al., 2004) and *limma* (Smyth, 2005) packages. Present, marginal, absent (PMA) calls were made using the Wilcoxon signed rank-based gene expression presence/absence detection algorithm as implemented by MAS5. The “expresso” function calculated expression values using “rma” background adjustment, “constant” normalization for microarray scaling, and “median polish” method for probe set summarization of perfect match only features, specified as “PMonly”.

Following normalization, expression values were further adjusted in the polysomal RNA samples based on levels of polysomes in cell extracts determined by sucrose density gradient fractionation exactly as described for this dataset (Branco-Price et al., 2008). This utilized the ratio of the amount of RNA in polysomes of stressed plants versus the amount of RNA in polysomes of non-stressed plants. Treatment comparisons were made using the *limma* package (Smyth, 2005). “lfit”, “contrasts.fit”, and “ebays” functions were used to calculate log₂ fold change (logFC) values and adjusted p-values (adj.P.Val) using the “fdr” method for controlling the false discovery rate. Annotated gene loci and descriptions were merged based on probeset IDs using the TAIR10 annotation file:

“affy_ATH1_array_elements-2010-12-20.txt” (www.arabidopsis.org, April 30, 2012).

The dataset was filtered to include probsets with a present call (1 "P" or "M" call out of 2 duplicate microarrays) in both total and IP samples from at least one of the treatments. The *e1071* package (Dimitriadou et al., 2011) was used for fuzzy *k*-means clustering of specified SLRs (i.e. differentially expressed genes (DEG) recognized by SLR and FDR as indicated in figure legends) using the "cmeans" function for 20 clusters using Euclidean distance and the "m" parameter set to 1.1. Clustered data were sorted and the median of each cluster determined. Twenty clusters were selected as an optimum for resolution of co-regulated genes based on evaluation of results using 10 to 30 clusters. Data was mapped as a heatmap using the "heatmap.2" function from the *gplots* package. Cluster gene ontology enrichment was determined by the "GOCluster_Report" function using the "simplify" method defined at http://faculty.ucr.edu/~tgirke/Documents/R_BioCond/My_R_Scripts/GOHyperGAll.txt (accessed April 30, 2012). GO categories of Arabidopsis genes were obtained from the "gene_association.tair" file from geneontology.org (accessed April 30, 2012). The distribution of signal log intensities were plotted using the "stat_density" statistic and the "path" geom of the *ggplot2* package (Wickham, 2009).

For nucleotide composition analysis, sequences were obtained from TAIR10 updated December 14, 2010. TAIR10 representative gene models, updated January 3, 2011, were used to reduce redundancy. A number of functions were defined (Appendix A) using functionality of the *Biostrings* (Pages et al.), *ggplot2* (Wickham, 2009), and *grid* (Gentleman et al., 2004) packages and were used to analyze cluster nucleotide composition maxima and plot the analysis. Multiple Em for Motif Elucidation (MEME) (Bailey and Elkan, 1994) was implemented in the *cosmo* package (Bembom et al., 2007) and plotted using modified code of the *seqLogo* package (Bembom) so that "U" was plotted in the place of "T" in the sequence logo.

3.5.8 Protein separation and immunoblot detection

Proteins in the different fractions were evaluated by separation and visualization on a 12% (w/v) polyacrylamide gel (Laemmli, 1970) stained with silver. Duplicate gels were electroblotted onto Hybond ECL nitrocellulose membrane (Amersham, Piscataway, NJ) and probed with monoclonal anti-FLAG M2 antibody horseradish peroxidase (HRP)-conjugate (1:1000; Sigma) or a polyclonal antibody against maize (*Zea mays* L.) ribosomal protein S6 (α RPS6; 1:10,000) (Williams et al., 2003) in combination with secondary goat anti-rabbit IgG HRP-conjugate (1:10,000; Bio-Rad, Hercules, CA). HRP detection was performed with ECL Plus Chemiluminescence system (Amersham) on Hyperfilm ECL (Amersham).

3.5.9 Detection of starch

Seedlings grown on solid MS medium in the presence or absence of sucrose were fixed overnight at 20°C in 40 mL 95% (v/v) ethanol immediately following the transition from light to dark. Seedlings were stained in Lugol's solution (1% (w/v) iodine, 2% (w/v) potassium iodide) 5 min, briefly rinsed in ddH₂O water, imaged with an Olympus CK2 inverted microscope (Center Valley, PA) and photographed with a Cannon Rebel T1i.

3.5.10 Confocal microscopy

GFP fusion proteins within plant organs were imaged with a Leica SP2 laser scanning confocal microscope (Wetzlar, Germany) using the 488 nm wavelength Ar laser line at 40% power. Emission wavelengths between 500 and 600 nm were collected as GFP fluorescence, and wavelengths between 680 nm and 720 nm were collected for chlorophyll fluorescence. Cy3 fluorophore and mRFP were excited using the 543 nm wavelength Ar laser line at 100% power. For Cy3 fluorescence of fixed hybridized tissue wavelengths between 560-650 nm were captured and for mRFP1 577-631 nm.

For analysis of aggregates of proteins (granules), reconstructed z-series were rendered as three-dimensional interactive graphics and analyzed using Bitplane software (Imaris, Inc. South Windsor, CT). Rendered graphics were also exported as movies using Bitplane. Alternatively, Imagej software was used to quantify granule size, number, and intensity from maximum projections of z-series.

For time-resolved fast imaging, a Yokogawa CSU-10 spinning disc confocal scan head (Tokyo, Japan) mounted on an Axiovert 100M inverted microscope (Carl Zeiss, Oberkochen, Germany) was used. The microscope was equipped with a PhotonMax 512B EMCCD camera (Princeton Instruments, Trenton, New Jersey) for high sensitivity fluorescence detection. The 488 nm wavelength Ar laser line and RGB filter cube were used for excitation and emission filtering, respectively. A 10-image z-series, spread evenly across a 24 μm depth, was taken every 10 s (0.5 s exposure time for each image) for 45 min and reconstructed into a movie and analyzed for granule number using Bitplane software.

3.5.11 Detection of poly(A) mRNA in situ hybridization with oligo(dT)

Poly(A) mRNA was detected by *in situ* hybridization as described by Kim et al. (2008) (Kim et al., 2008). Briefly, 4-6 seedlings were transferred to a 5 mL glass vial and fixed in 50% fixation buffer (FB; 120 mM NaCl, 7 mM Na_2HPO_4 , 3 mM NaH_2PO_4 [pH 7.5], 2.7 mM KCl, 0.1% (v/v) Tween-20, 80 mM EGTA, 5% (w/v) formaldehyde, 10% (v/v) DMSO), 50% (v/v) heptane), shaken on a rotary shaker for 30 min at 23-25°C, dehydrated in 100% (v/v) methanol twice for 5 min, and in 100% (v/v) ethanol thrice for 5 min. The seedlings were incubated 30 min in 1:1 ethanol:xylene, then twice in 100% (v/v) ethanol and 100% (v/v) methanol each for 5 min. Tissue was post-fixed in 1:1 methanol:FB without formaldehyde 30 min at 23-25°C, and then rinsed in FB without formaldehyde. Prior to hybridization, seedlings were rinsed in 1 mL PerfectHyb Plus hybridization buffer (Sigma-Aldrich; H-7033) and then prehybridized in 1 mL the same buffer for 1 h at 50°C. 10 μL of 0.5 μM 5'-Cy3-labeled

oligo(d)T 45mer (Integrated DNA Technologies, Coralville, IA) were added and the sample was incubated 12 h at 50°C. Following hybridization, the seedlings were washed 60 min in 2x SSC (300 mM NaCl, 30 mM sodium citrate [pH 7.0]), 0.1% (w/v) SDS at 50°C, and 20 min in 0.2x SSC, 0.1% (w/v) SDS at 50°C. Samples remained in this solution until imaging.

3.5.12 Chlorophyll measurements

Chlorophyll content was determined as previously described (Porra et al., 1989). Briefly, chlorophyll was extracted in cold 80% (v/v) aqueous acetone from 2-wk-old seedlings grown on solid MS medium containing 29.2 mM sucrose. Following organic extraction, absorbance at 663.6, 646.6, and 750 nm was measured and extract concentration was calculated based on the equations: $\mu\text{g Chl}_a \text{ mL}^{-1} = 12.25 * (A663 - A750) - 2.55 * (A646 - A750)$, $\mu\text{g Chl}_b \text{ mL}^{-1} = 20.31 * (A646 - A750) - 4.91 * (A663 - A750)$.

3.6 References

- Anderson, P.** (2010). Post-transcriptional regulons coordinate the initiation and resolution of inflammation. *Nature Reviews Immunology* **10**: 24–35.
- Anderson, P. and Kedersha, N.** (2006). RNA granules. *The Journal of Cell Biology* **172**: 803–808.
- Anderson, P. and Kedersha, N.** (2009). RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* **10**: 430–436.
- Anderson, P. and Kedersha, N.** (2008). Stress granules: the Tao of RNA triage. *Trends Biochem. Sci.* **33**: 141–150.
- Arteaga-Vázquez, M., Caballero-Pérez, J., and Vielle-Calzada, J.-P.** (2006). A family of microRNAs present in plants and animals. *Plant Cell* **18**: 3355–3369.
- Bailey, T. and Elkan, C.** (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology* (AAAI Press, Menlo Park, California), pp. 28–36.
- Bailey-Serres, J., Sorenson, R., and Juntawong, P.** (2009). Getting the message across: cytoplasmic ribonucleoprotein complexes. *Trends in Plant Science* **14**: 443–453.

- Beck, A.R., Medley, Q.G., O'Brien, S., Anderson, P., and Streuli, M.** (1996). Structure, tissue distribution and genomic organization of the murine RRM-type RNA binding proteins TIA-1 and TIAR. *Nucleic Acids Res.* **24**: 3829–3835.
- Bembom, O.** seqLogo: Sequence logos for DNA sequence alignments.
- Bembom, O., Keles, S., and van der Laan, M.J.** (2007). Supervised detection of conserved motifs in DNA sequences with cosmo. *Stat Appl Genet Mol Biol* **6**: Article8.
- Branco-Price, C., Kaiser, K.A., Jang, C.J.H., Larive, C.K., and Bailey-Serres, J.** (2008). Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. *The Plant Journal* **56**: 743–755.
- Bürkle, L., Meyer, S., Dortay, H., Lehrach, H., and Heyl, A.** (2005). In vitro recombination cloning of entire cDNA libraries in *Arabidopsis thaliana* and its application to the yeast two-hybrid system. *Functional & Integrative Genomics* **5**: 175–183.
- Clough, S.J. and Bent, A.F.** (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**: 735–743.
- Coller, J.M., Gray, N.K., and Wickens, M.P.** (1998). mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes Dev* **12**: 3226–3235.
- Damgaard, C.K. and Lykke-Andersen, J.** (2011). Translational coregulation of 5'TOP mRNAs by TIA-1 and TIAR. *Genes & Development* **25**: 2057–2068.
- Decker, C.J., Teixeira, D., and Parker, R.** (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J. Cell Biol.* **179**: 437–449.
- Dember, L.M., Kim, N.D., Liu, K.-Q., and Anderson, P.** (1996). Individual RNA recognition motifs of TIA-1 and TIAR have different RNA binding specificities. *Journal of Biological Chemistry* **271**: 2783–2788.
- Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., and Weingessel, A.** (2011). e1071: Misc functions of the Department of Statistics (e1071), TU Wien.
- Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S.** (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**: 616–629.
- Edgar, R.C.** (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* **32**: 1792–1797.
- Eisinger-Mathason, T.S.K., Andrade, J., Groehler, A.L., Clark, D.E., Muratore-Schroeder, T.L., Pasic, L., Smith, J.A., Shabanowitz, J., Hunt, D.F., Macara, I.G., and Lannigan, D.A.** (2008). Codependent functions of RSK2 and the apoptosis-promoting factor TIA-1 in stress granule assembly and cell survival. *Mol. Cell* **31**: 722–736.

- Felsenstein, J.** (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**: 783–791.
- Förch, P. and Valcárcel, J.** (2001). Molecular mechanisms of gene expression regulation by the apoptosis-promoting protein TIA-1. *Apoptosis* **6**: 463–468.
- Del Gatto-Konczak, F., Bourgeois, C.F., Le Guiner, C., Kister, L., Gesnel, M.-C., Stévenin, J., and Breathnach, R.** (2000). The RNA-binding protein TIA-1 is a novel mammalian splicing regulator acting through intron sequences adjacent to a 5' splice site. *Molecular and Cellular Biology* **20**: 6287–6299.
- Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A.** (2004). affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**: 307–315.
- Gentleman, R.C., Carey, V.J., Bates, D.M., and others** (2004). Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology* **5**: R80.
- Gil, P. and Green, P.J.** (1996). Multiple regions of the Arabidopsis SAUR-AC1 gene control transcript abundance: the 3' untranslated region functions as an mRNA instability determinant. *EMBO J* **15**: 1678–1686.
- Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P.** (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell* **15**: 5383–5398.
- Gomez-Porras, J., Lewinski, M., Riabi-Pachon, D.M., and Staiger, D.** (2011). Molecular evolution of RRM-containing proteins and glycine-rich RNA-binding proteins in plants. *Nature Precedings* <<http://hdl.handle.net/10101/npre.2011.5970.1>>.
- Goodall, G.J. and Filipowicz, W.** (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* **58**: 473–483.
- Gottschald, O.R., Malec, V., Krasteva, G., Hasan, D., Kamlah, F., Herold, S., Rose, F., Seeger, W., and Hänze, J.** (2010). TIAR and TIA-1 mRNA-binding proteins co-aggregate under conditions of rapid oxygen decline and extreme hypoxia and suppress the HIF-1 α pathway. *J Mol Cell Biol* **2**: 345–356.
- Gueydan, C., Droogmans, L., Chalon, P., Huez, G., Caput, D., and Krays, V.** (1999). Identification of TIAR as a protein binding to the translational regulatory AU-rich element of tumor necrosis factor alpha mRNA. *J. Biol. Chem.* **274**: 2322–2326.
- Hanano, S. and Goto, K.** (2011). Arabidopsis TERMINAL FLOWER1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *Plant Cell* **23**: 3172–3184.
- Ivanov and Nadezhkina** (2006). Stress granules: RNP-containing cytoplasmic bodies arising in stress: Structure and mechanism of organization. *Molecular Biology* **40**: 844–850.

- Izquierdo, J.M.** (2010). Cell-specific regulation of Fas exon 6 splicing mediated by Hu antigen R. *Biochemical and Biophysical Research Communications* **402**: 324–328.
- Izquierdo, J.M.** (2006). Control of the ATP synthase β subunit expression by RNA-binding proteins TIA-1, TIAR, and HuR. *Biochemical and Biophysical Research Communications* **348**: 703–711.
- Izquierdo, J.M. and Valcarcel, J.** (2007). Two isoforms of the T-cell intracellular antigen 1 (TIA-1) splicing factor display distinct splicing regulation activities: control of TIA-1 isoform ratio by TIA-1 RELATED protein. *J. Biol. Chem* **282**: 19410–19417.
- Izquierdo, J.M. and Valcárcel, J.** (2007). Fas-activated Serine/Threonine Kinase (FAST K) synergizes with TIA-1/TIAR proteins to regulate Fas alternative splicing. *Journal of Biological Chemistry* **282**: 1539–1543.
- Jouannet, V., Moreno, A.B., Elmayan, T., Vaucheret, H., Crespi, M.D., and Maizel, A.** (2012). Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *The EMBO Journal* **31**: 1704–1713.
- Kawaguchi, R. and Bailey-Serres, J.** (2005). mRNA sequence features that contribute to translational regulation in Arabidopsis. *Nuc. Acids Res.* **33**: 955–965.
- Kedersha, N., Chen, S., Gilks, N., Li, W., Miller, I.J., Stahl, J., and Anderson, P.** (2002). Evidence that ternary complex (eIF2-GTP-tRNA^{iMet})-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol. Biol. Cell* **13**: 195–210.
- Kedersha, N., Cho, M.R., Li, W., Yacono, P.W., Chen, S., Gilks, N., Golan, D.E., and Anderson, P.** (2000). Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J. Cell Biol.* **151**: 1257–1268.
- Kim, H.S., Wilce, M.C.J., Yoga, Y.M.K., Pardini, N.R., Gunzburg, M.J., Cowieson, N.P., Wilson, G.M., Williams, B.R.G., Gorospe, M., and Wilce, J.A.** (2011). Different modes of interaction by TIAR and HuR with target RNA and DNA. *Nucleic Acids Research* **39**: 1117–1130.
- Kim, J.S., Jung, H.J., Lee, H.J., Kim, K.A., Goh, C.-H., Woo, Y., Oh, S.H., Han, Y.S., and Kang, H.** (2008). Glycine-rich RNA-binding protein 7 affects abiotic stress responses by regulating stomata opening and closing in *Arabidopsis thaliana*. *Plant J* **55**: 455–66.
- Ko, C.H., Brendel, V., Taylor, R.D., and Walbot, V.** (1998). U-richness is a defining feature of plant introns and may function as an intron recognition signal in maize. *Plant Mol. Biol.* **36**: 573–583.
- Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K.** (2003). Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. *Plant Cell Physiol.* **44**: 555–564.

- Laemmli, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.
- Lambermon, M.H.L., Fu, Y., Wieczorek Kirk, D.A., Dupasquier, M., Filipowicz, W., and Lorković, Z.J.** (2002). UBA1 and UBA2, two proteins that interact with UBP1, a multifunctional effector of pre-mRNA maturation in plants. *Mol. Cell. Biol.* **22**: 4346–4357.
- Lambermon, M.H.L., Simpson, G.G., Kirk, D.A.W., Hemmings-Mieszczak, M., Klahre, U., and Filipowicz, W.** (2000). UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J.* **19**: 1638–1649.
- Lee, S.C., Mustroph, A., Sasidharan, R., Vashisht, D., Pedersen, O., Oosumi, T., Voesenek, L.A.C., and Bailey - Serres, J.** (2011). Molecular characterization of the submergence response of the *Arabidopsis thaliana* ecotype Columbia. *New Phytol.* **190**: 457–471.
- Liu, C. and Mehdy, M.C.** (2007). A nonclassical arabinogalactan protein gene highly expressed in vascular tissues, AGP31, is transcriptionally repressed by methyl jasmonic acid in *Arabidopsis*. *Plant Physiol.* **145**: 863–874.
- Liu, M.-J., Wu, S.-H., Chen, H.-M., and Wu, S.-H.** (2012). Widespread translational control contributes to the regulation of *Arabidopsis* photomorphogenesis. *Mol Syst Biol* **8**: 566.
- Livak, K.J. and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods* **25**: 402–408.
- López de Silanes, I., Galbán, S., Martindale, J.L., Yang, X., Mazan-Mamczarz, K., Indig, F.E., Falco, G., Zhan, M., and Gorospe, M.** (2005). Identification and functional outcome of mRNAs associated with RNA-binding protein TIA-1. *Mol. Cell. Biol.* **25**: 9520–9531.
- Lorković, Z.J. and Barta, A.** (2002). Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nuc. Acids Res.* **30**: 623–635.
- Lorković, Z.J., Wieczorek Kirk, D.A., Klahre, U., Hemmings-Mieszczak, M., and Filipowicz, W.** (2000). RBP45 and RBP47, two oligouridylylate-specific hnRNP-like proteins interacting with poly(A)⁺ RNA in nuclei of plant cells. *RNA* **6**: 1610–1624.
- Manfield, I.W., Jen, C.-H., Pinney, J.W., Michalopoulos, I., Bradford, J.R., Gilmartin, P.M., and Westhead, D.R.** (2006). *Arabidopsis* Co-expression Tool (ACT): web server tools for microarray-based gene expression analysis. *Nucleic Acids Res.* **34**: W504–509.
- Mashiguchi, K., Asami, T., and Suzuki, Y.** (2009). Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes

in *Arabidopsis*. *Biosci. Biotechnol. Biochem.* **73**: 2452–2459.

- Matsuura, H., Ishibashi, Y., Shinmyo, A., Kanaya, S., and Kato, K.** (2010). Genome-wide analyses of early translational responses to elevated temperature and high salinity in *Arabidopsis thaliana*. *Plant Cell Physiol* **51**: 448–462.
- McClure, B.A., Hagen, G., Brown, C.S., Gee, M.A., and Guilfoyle, T.J.** (1989). Transcription, organization, and sequence of an auxin-regulated gene cluster in soybean. *Plant Cell* **1**: 229–239.
- McCue, A.D., Nuthikattu, S., Reeder, S.H., and Slotkin, R.K.** (2012). Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. *PLoS Genet* **8**: e1002474.
- Michelitsch, M.D. and Weissman, J.S.** (2000). A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions. *PNAS* **97**: 11910–11915.
- Mustroph, A., Juntawong, P., and Bailey-Serres, J.** (2009). Isolation of plant polysomal mRNA by differential centrifugation and ribosome immunopurification methods. *Meth. Mol. Biol.* **553**: 109–126.
- Mustroph, A., Lee, S.C., Oosumi, T., Zanetti, M.E., Yang, H., Ma, K., Masihi, A., Fukao, T., and Bailey-Serres, J.** (2010). Cross-kingdom comparison of transcriptomic adjustments to low oxygen stress highlights conserved and plant-specific responses. *Plant Physiol.*: Provisional acceptance.
- Mustroph, A., Zanetti, M.E., Jang, C.J.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J.** (2009). Profiling translomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A* **106**: 18843–18848.
- Nover, L., Scharf, K.D., and Neumann, D.** (1989). Cytoplasmic heat shock granules are formed from precursor particles and are associated with a specific set of mRNAs. *Mol. Cell. Biol.* **9**: 1298–1308.
- Oeffinger, M., Wei, K.E., Rogers, R., DeGrasse, J.A., Chait, B.T., Aitchison, J.D., and Rout, M.P.** (2007). Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat. Meth.* **4**: 951–956.
- Pages, H., Aboyoun, P., Gentleman, R., and DebRoy, S.** *Biostrings*: String objects representing biological sequences, and matching algorithms.
- Pieczyk, M., Wax, S., Beck, A.R., Kedersha, N., Gupta, M., Maritim, B., Chen, S., Gueydan, C., Kruys, V., Streuli, M., and Anderson, P.** (2000). TIA-1 is a translational silencer that selectively regulates the expression of TNF- α . *EMBO J.* **19**: 4154–4163.
- Porra, R.J., Thompson, W.A., and Kriedemann, P.E.** (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b

extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *BBA Bioenergetics* **975**: 384–394.

- Puckette, M., Iyer, N.J., Tang, Y., Dai, X.-B., Zhao, P., and Mahalingam, R.** (2012). Differential mRNA translation in *Medicago truncatula* accessions with contrasting responses to ozone-induced oxidative stress. *Mol. Plant* **5**: 187–204.
- R. Development Core Team** (2012). R: A language and environment for statistical computing (Vienna, Austria).
- Ribeiro, D.M., Araújo, W.L., Fernie, A.R., Schippers, J.H.M., and Mueller-Roeber, B.** (2012). Translatome and metabolome effects triggered by gibberellins during rosette growth in Arabidopsis. *J Exp Bot* **63**: 2769–2786.
- Saitou, N. and Nei, M.** (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- Smyth, G.K.** (2005). Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, and W. Huber, eds (Springer: New York), pp. 397–420.
- Sorin, C., Bussell, J.D., Camus, I., Ljung, K., Kowalczyk, M., Geiss, G., McKhann, H., Garcion, C., Vaucheret, H., Sandberg, G., and Bellini, C.** (2005). Auxin and light control of adventitious rooting in Arabidopsis require ARGONAUTE1. *Plant Cell* **17**: 1343–1359.
- Tamura, K., Dudley, J., Nei, M., and Kumar, S.** (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**: 1596–1599.
- Tan, L., Showalter, A.M., Egelund, J., Doblin, M.S., and Bacic, A.** (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci* **3**: 140.
- Taupin, J.L., Tian, Q., Kedersha, N., Robertson, M., and Anderson, P.** (1995). The RNA-binding protein TIAR is translocated from the nucleus to the cytoplasm during Fas-mediated apoptotic cell death. *Proceedings of the National Academy of Sciences* **92**: 1629–1633.
- Ueda, K., Matsuura, H., Yamaguchi, M., Demura, T., and Kato, K.** (2012). Genome-wide analyses of changes in translation state caused by elevated temperature in *Oryza sativa*. *Plant Cell Physiol.*
- Vaucheret, H., Vazquez, F., Crété, P., and Bartel, D.P.** (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18**: 1187–1197.
- Weber, C., Nover, L., and Fauth, M.** (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.* **56**: 517–530.
- Wickham, H.** (2009). ggplot2: elegant graphics for data analysis (Springer New York).

- Williams, A.J., Werner-Fraczek, J., Chang, I.-F., and Bailey-Serres, J.** (2003). Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. *Plant Physiol.* **132**: 2086–2097.
- Wu, X., Dabi, T., and Weigel, D.** (2005). Requirement of homeobox gene STIMPY/WOX9 for Arabidopsis meristem growth and maintenance. *Curr. Biol.* **15**: 436–440.
- Zanetti, M.E., Chang, I.-F., Gong, F., Galbraith, D.W., and Bailey-Serres, J.** (2005). Immunopurification of polyribosomal complexes of Arabidopsis for global analysis of gene expression. *Plant Physiol.* **138**: 624–635.
- Zhang, S. and Mehdy, M.C.** (1994). Binding of a 50-kD protein to a U-Rich sequence in an mRNA encoding a proline-rich protein That is destabilized by fungal elicitor. *Plant Cell* **6**: 135–145.
- Zhang, T., Delestienne, N., Huez, G., Kruys, V., and Gueydan, C.** (2005). Identification of the sequence determinants mediating the nucleo-cytoplasmic shuttling of TIAR and TIA-1 RNA-binding proteins. *J. Cell. Sci.* **118**: 5453–5463.
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Grissem, W.** (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiology* **136**: 2621 –2632.

Figure 3.1. Phylogeny of plant UBP1 and related proteins based on conserved RNA recognition motifs. Evolutionary history was inferred from of plant UBP1, RBP45/47, animal TIA1 as well as representative plant and animal poly(A) binding proteins. Amino acid sequence of representative protein isoforms of animal taxa and plants were obtained from public databases. 99 sequences were aligned using the Muscle algorithm in Mega 5.0 [1,2]. The evolutionary distances were computed using the JTT matrix-based method [3] and are in the units of the number of amino acid substitutions per site. A phylogeny was produced using the Neighbor-Joining method [4]. All positions containing gaps and missing data were eliminated. There were a total of 176 positions in the final dataset predominantly from RNA recognition motif (RRM) domains. Bootstrap (1,000 replicates) values are shown above branches [5]. Branch lengths are shown to scale. Species and gene symbol or locus index is used to identify each sequence. Species names were abbreviated as follows: *Arabidopsis lyrata* (Araly), *Arabidopsis thaliana* (Arath), *Brassica rapa* (Brara), *Glycine max* (Glyma), *Medicago truncatula* (Medtr), *Oryza sativa* (Orysa), *Physcomitrella patens* (Phypa), *Populus trichocarpa* (Poptr), *Selaginella moellendorffii* (Selmo), *Solanum tuberosum* (Soltu), *Sorghum bicolor* (Sorbi), *Thellungiella parvula* (Thepa) , *Vitis vinifera* (Vitvi), *Zea mays* (Zeama). RBP45/47 members were only included for *Arabidopsis thaliana*, *Medicago truncatula*, *Oryza sativa*, *Physcomitrella patens*, and *Zea mays*. *Arabidopsis thaliana* UBP1 genes are marked with green diamonds, and *Solanum* gene indices were shortened from PGSC0003DMP400005226, PGSC0003DMP400044537, PGSC0003DMP400008635, and PGSC0003DMP400039690 for convenience.

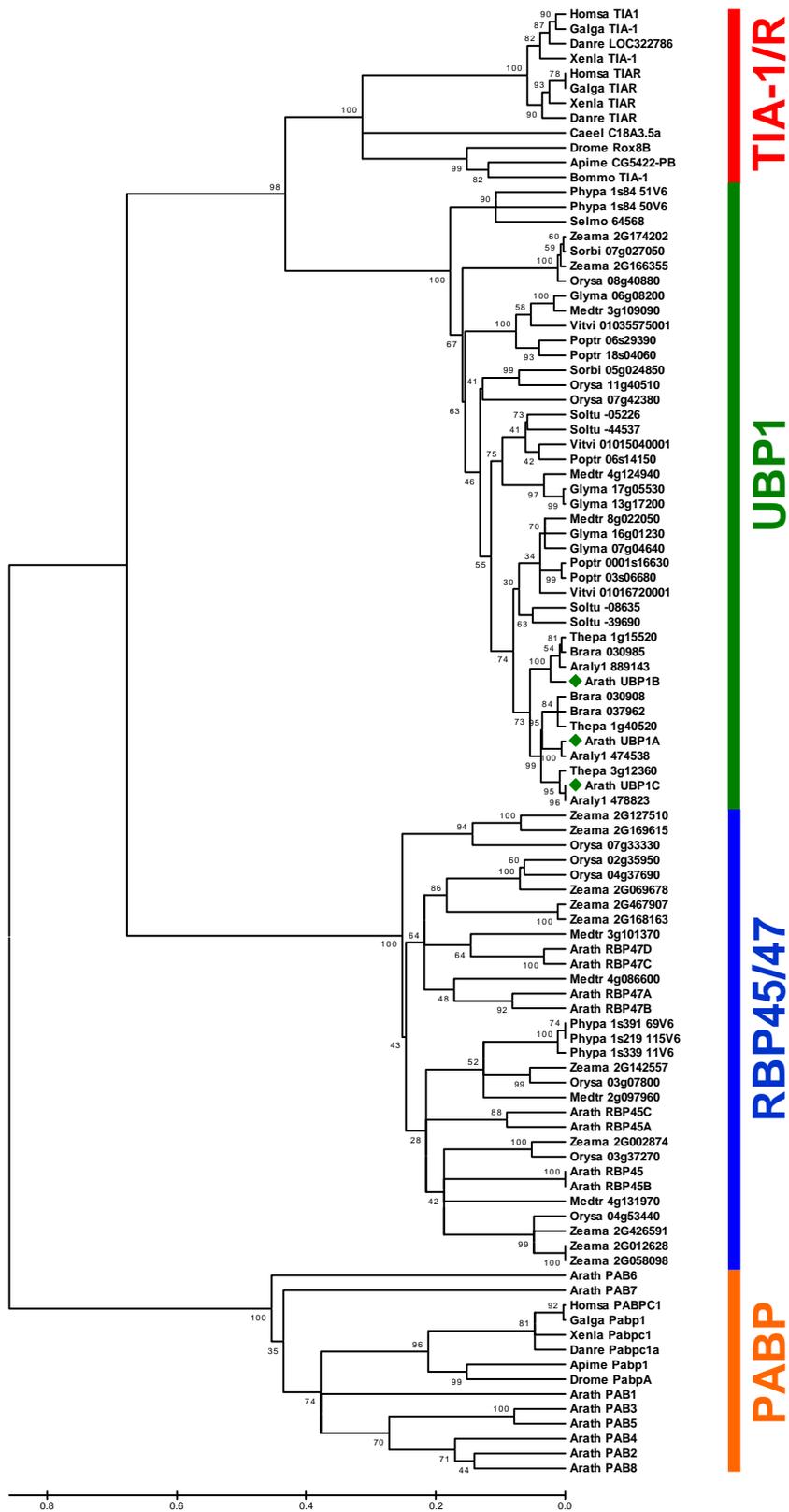
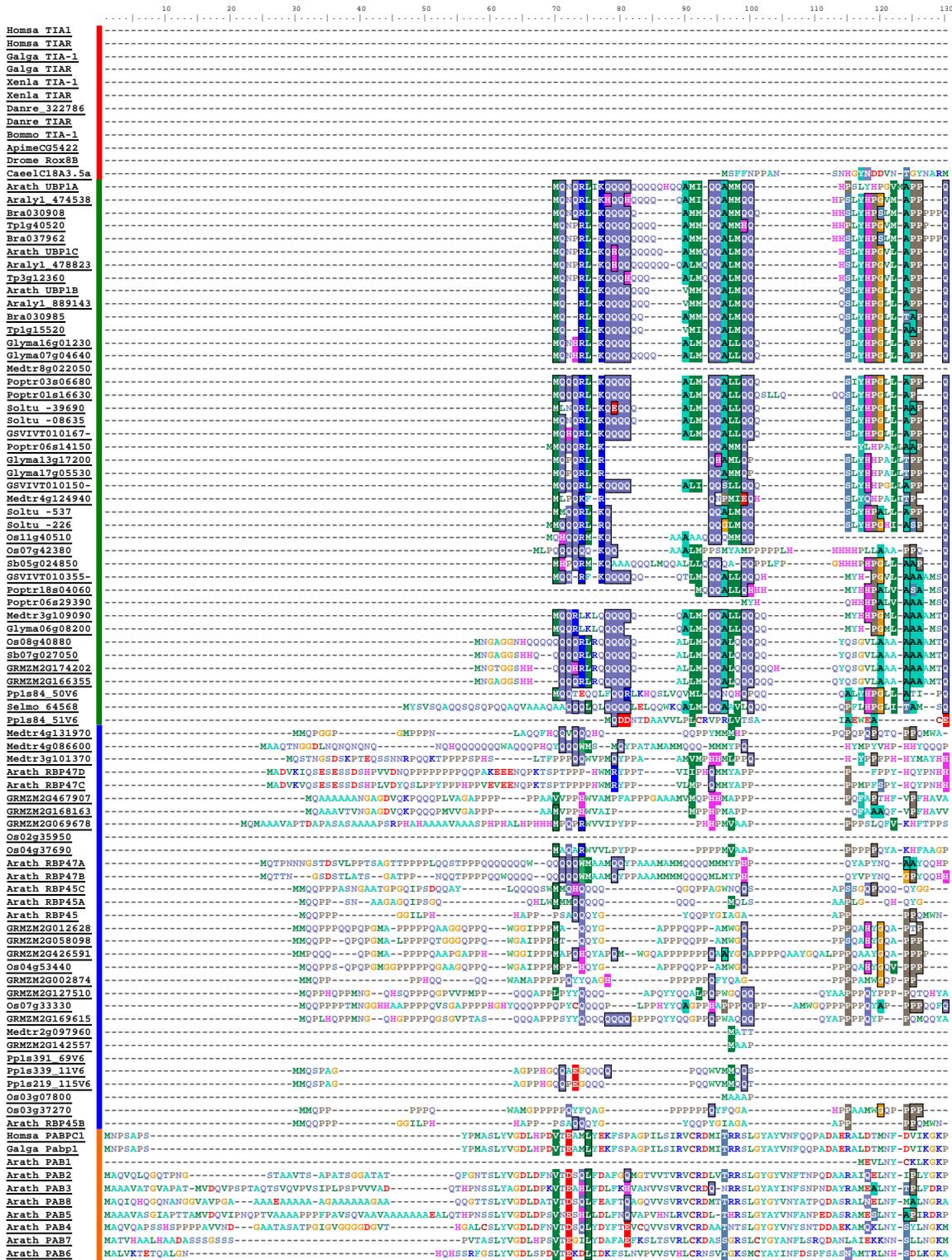
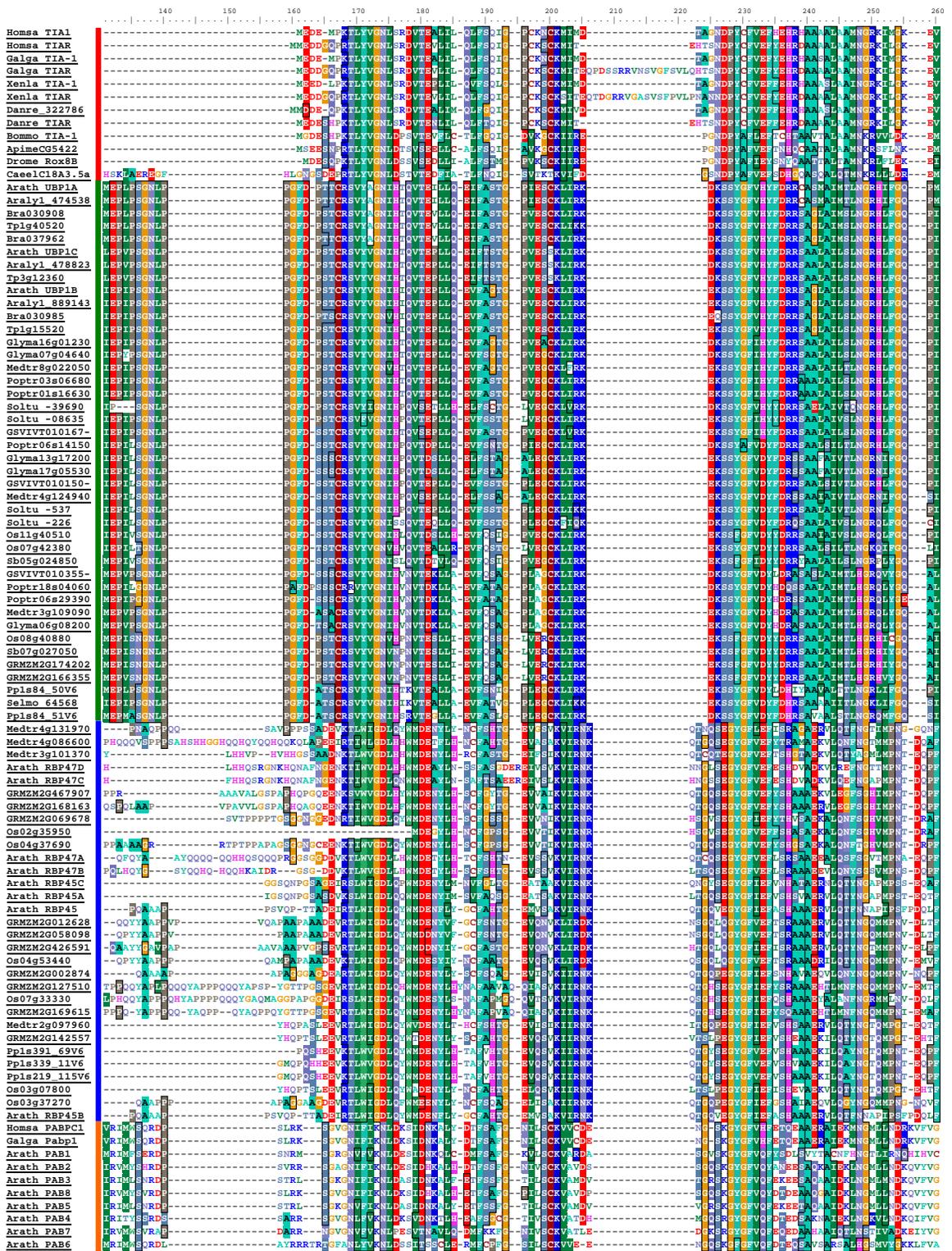


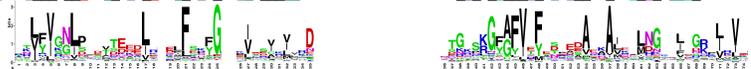
Figure 3.2. Conservation and divergence of plant UBP1 and RBP45/47 proteins. (A) Alignment of amino acid sequences used to generate the tree in Figure 1. PABP family members from frog, fish, and insect were removed following alignment for convenience. The RRM domains are indicated by alignment to the domain sequence logo generated from prosite entry PS50102 based on 2,265 hits in 1,349 different sequences including plant and animal hits. Sequences from the major protein families of Figure 1 are color-coded in similar fashion by a colored bar next to the sequence name. Labeled domains (bottom) are also color coded consistent with the family to which they apply (red: Tia-1; green: UBP1; blue: RBP45/47; orange: PABP). Gene symbols or locus identifiers were used to label the sequences. Abbreviations: PRD, prion-related domain; RRM, RNA recognition motif.

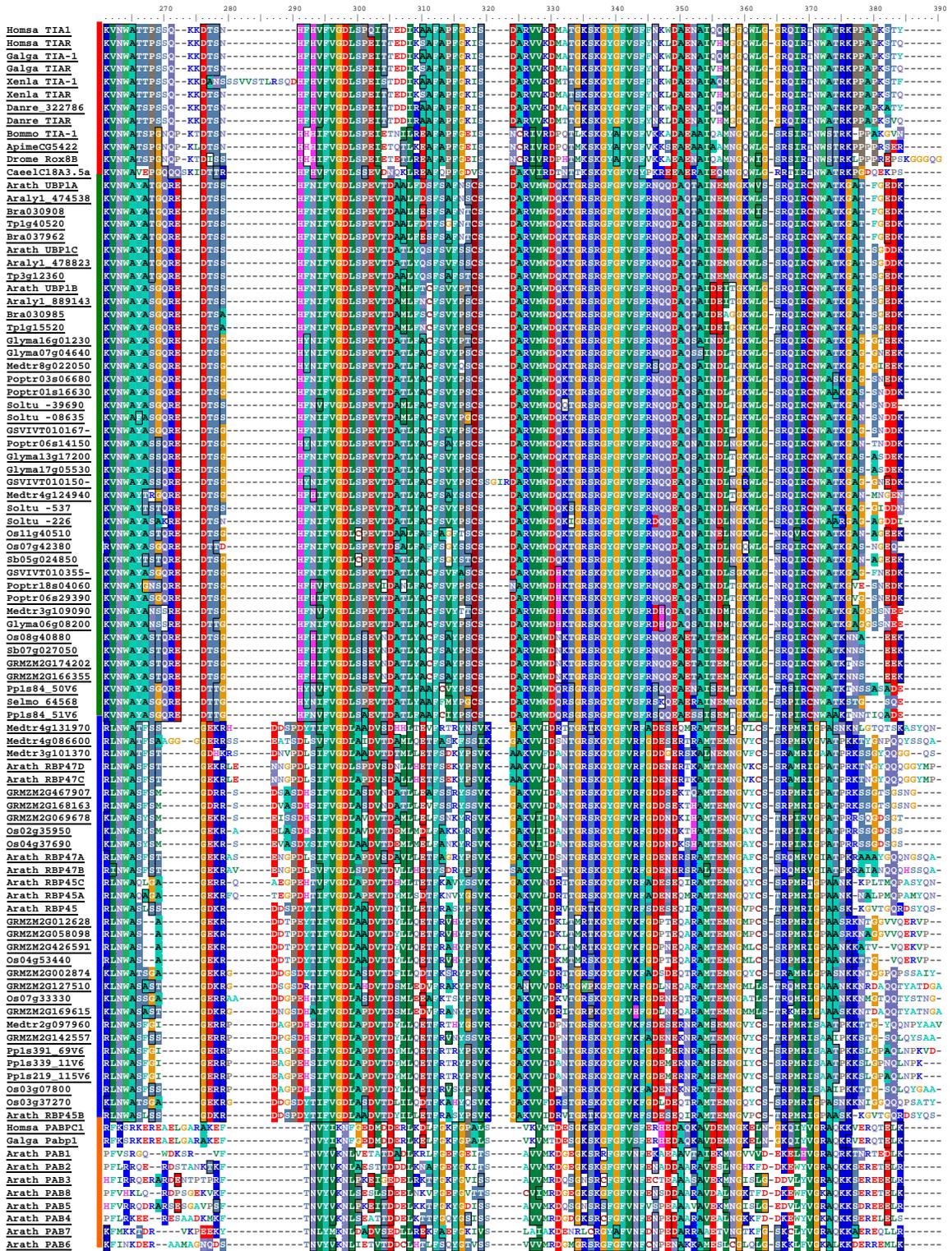


Q-rich RRM Q-rich

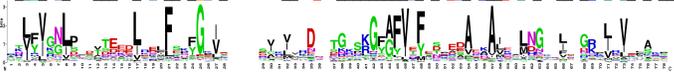


RRM1:





RRM2:



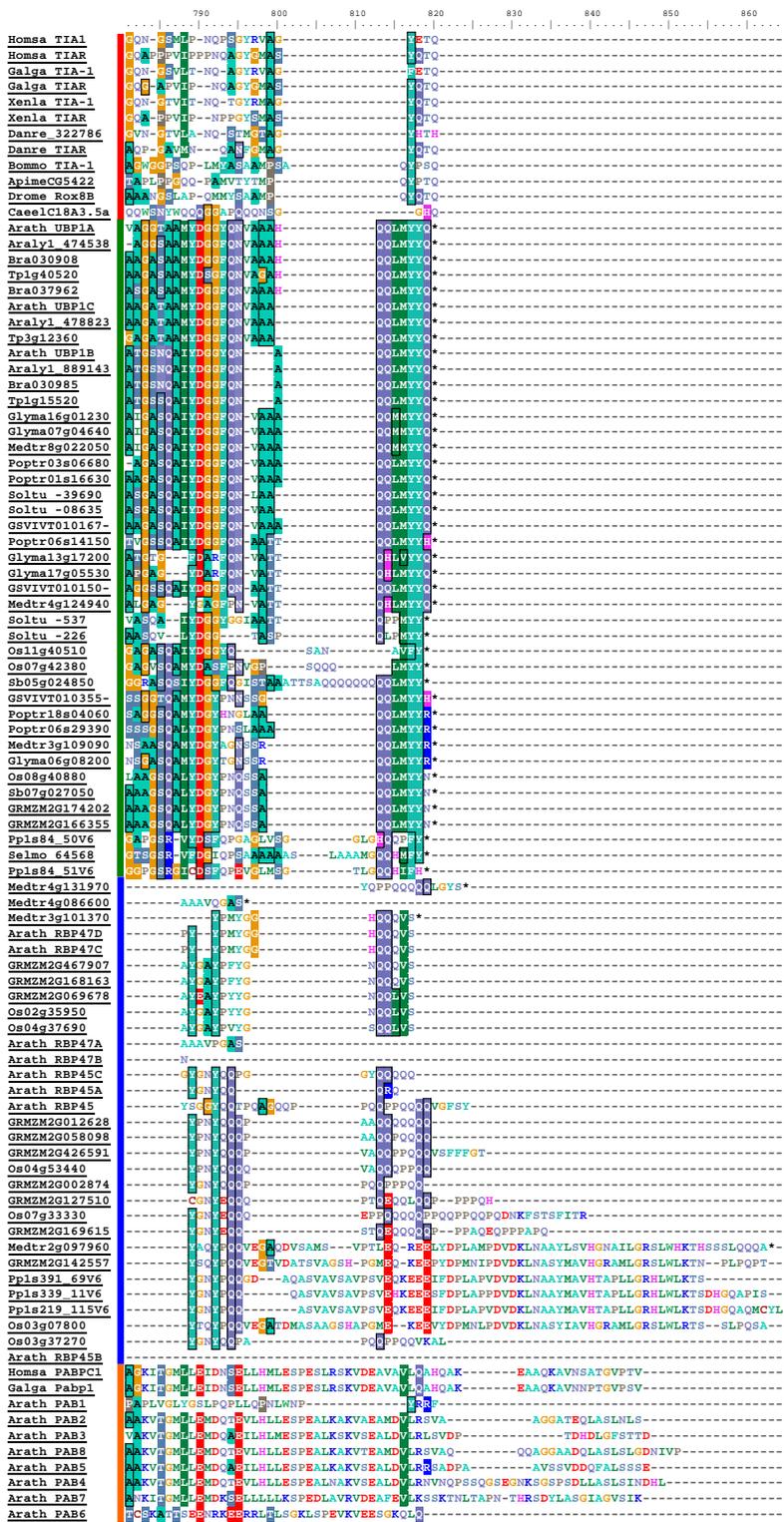


Figure 3.3. *UBP1* gene family members are generally co-expressed. (A) A co-expression network generated by use of the Atted (atted.jp) network drawer demonstrates the co-expression of the *UBP1* gene family members, *UBP1A*, *UBP1B*, and *UBP1C*. (B) Graphical display and Pearson correlation coefficient calculations of co-expression of individual gene probes with *UBP1C* (At22k probe set: 257002_at) on the x-axis and *UBP1A* (At22k GeneChip probe set: 263160_at) on the y-axis were generated by use of the *Arabidopsis* co-expression data mining tool (<http://www.arabidopsis.leeds.ac.uk/act/index.php>) [6]. (C) Distinctions in *UBP1* gene family member co-regulation are found in publicly available microarray data. Shown here, *UBP1C* mRNA accumulation is uniquely induced by prolonged darkness, which is enhanced in combination with submergence [7]. Asterisk indicates a significantly different fold change in both comparisons (adjusted p-value<0.05).

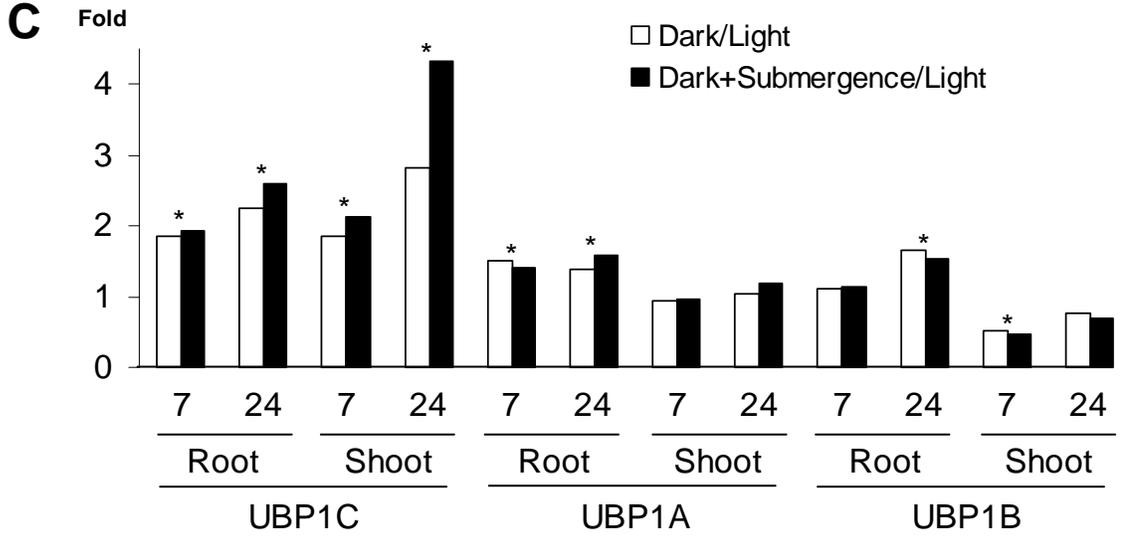
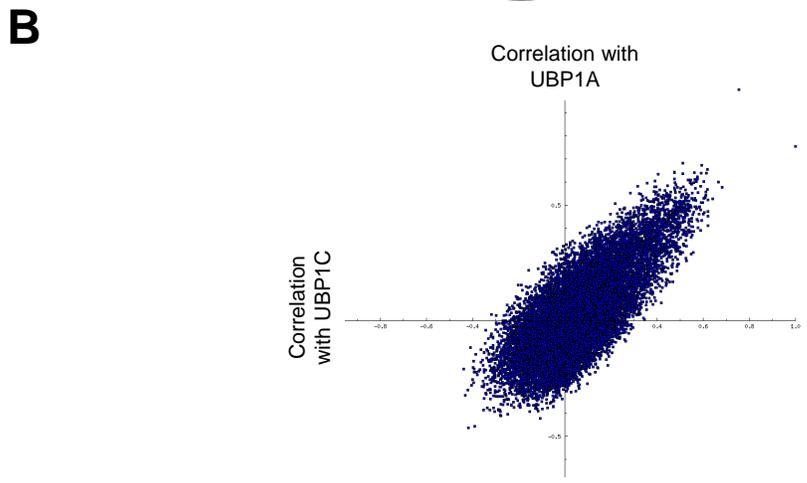
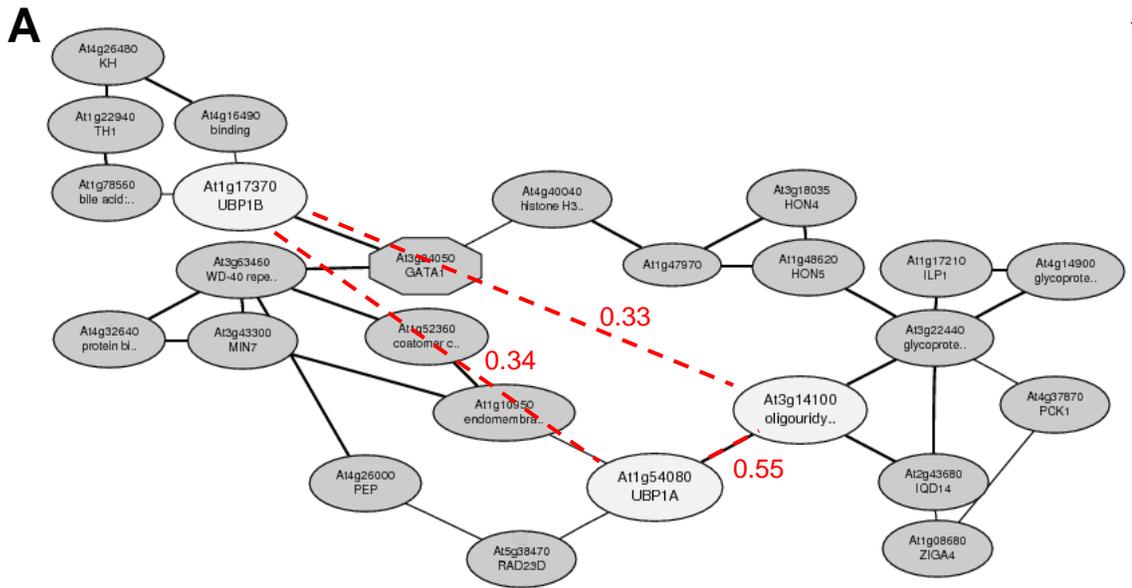


Figure 3.4. Relative cell-type specific expression in seedlings is distinct between *Arabidopsis thaliana* *UBP1s*. (A) *UBP1A*, (B) *UBP1B*, and (C) *UBP1C* expression within distinct translates (mRNA populations isolated in association with cell-type specific expressed, epitope-tagged polyribosomes) isolated from seedlings subjected to mock control treatment or oxygen deprivation are displayed using the eFP browser (efp.ucr.edu) [8].

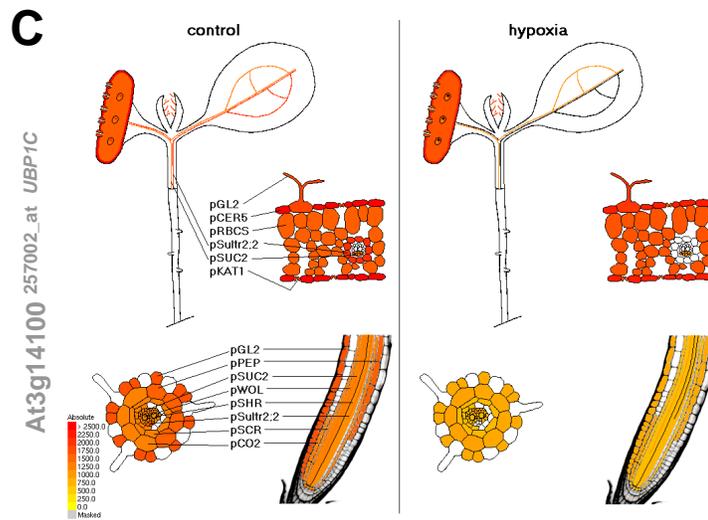
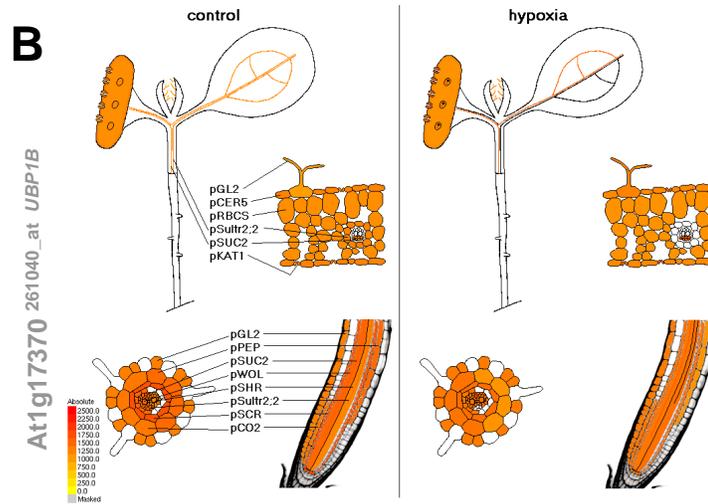
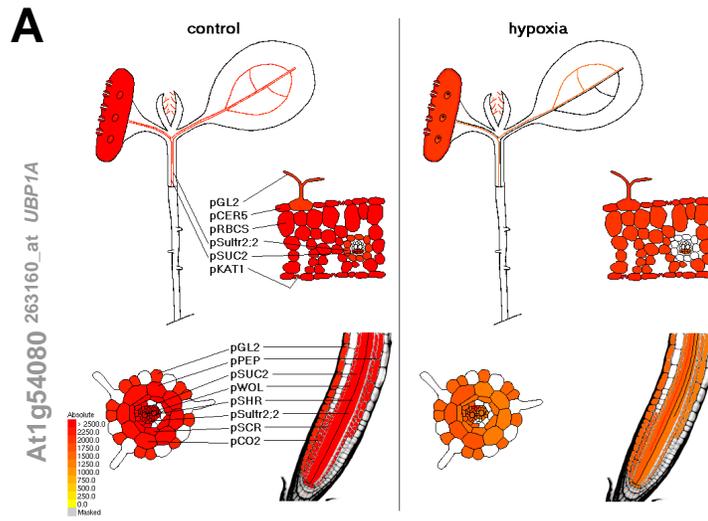


Figure 3.5. *UBP1*, *RBP45*, *RBP47* gene family members show distinct tissue expression. A heat map comparing tissue-specific gene expression of members of *UBP1*, *RBP45*, and *RBP47* gene families shows examples of unique, tissue-specific expression profiles. The Anatomy tool in Genevestigator 4 [9] was used. It extracts expression values from publically available Affymetrix At22k Genechip microarray data and does robust multi-chip adjustment (RMA) normalization followed by a further adjustment for inter-experiment comparison as described on the website (www.genevestigator.com). Percent of the expression potential (maximum \log_2 expression value out of all 6,458 included studies) is indicated by shading intensity. Maximum \log_2 expression values (100%) for each gene were as follows: *UBP1A*: 16.06, *UBP1B*: 14.88, *UBP1C*: 16.04, *RBP45C*: 15.97, *RBP45A*: 15.25, *RBP45*: 14.98, *RBP45B*: 17.98, *RBP47A*: 15.85, *RBP47B*: 15.57, *RBP47C*: 18.60.

Arabidopsis thaliana

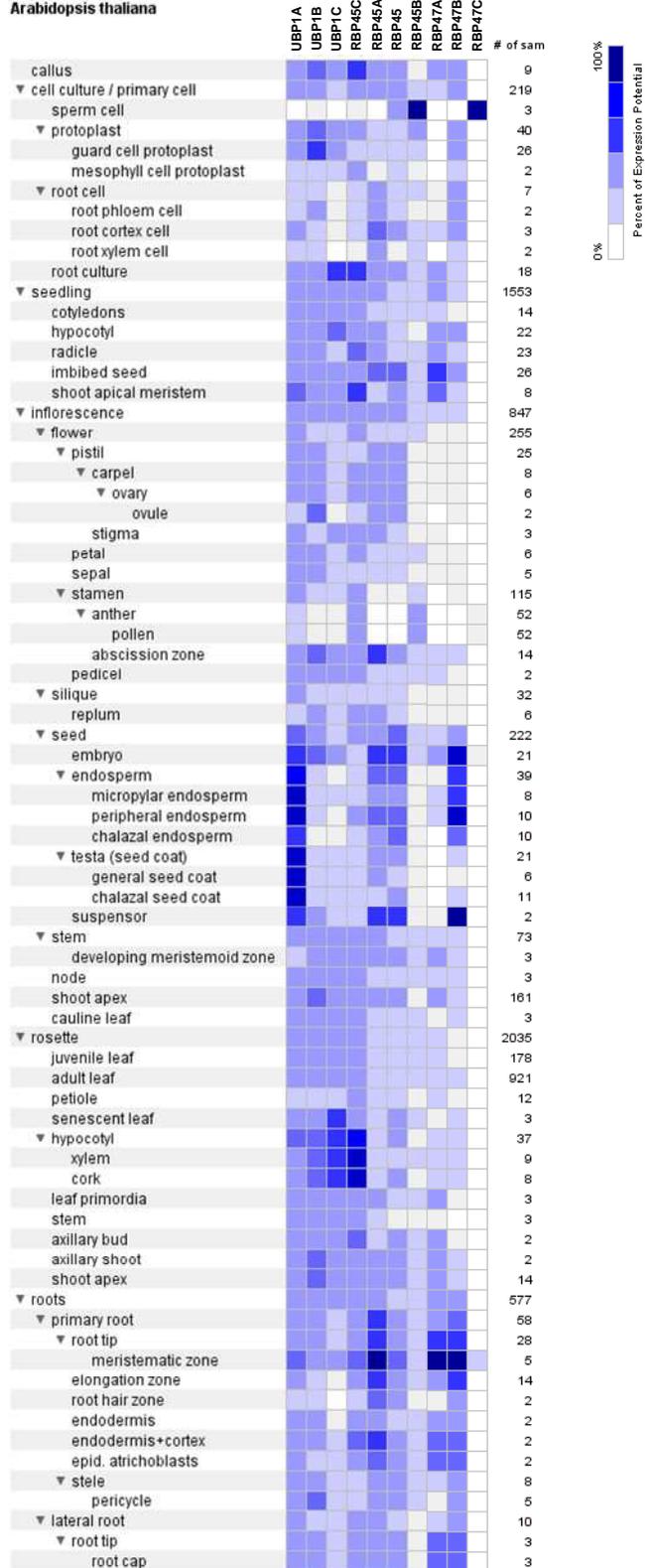


Figure 3.6. UBP1 gene loci and expression in mutant lines. (A) Gene model diagrams of *UBP1A*, *UBP1B*, and *UBP1C* show alternative splice forms, TDNA insertion position of multiple alleles, artificial miRNA-ubp1c target position, and primer binding sites (red arrows; letter labels indicate primers used in B and C) used in gene expression studies. (images source: gbrowse.arabidopsis.org). (B) Semi-quantitative reverse transcription polymerase chain reaction of *UBP1C* shows knock-out of the *UBP1C* gene compared to *Act2* in cDNA from rosette leaves of wild type Col-0 and *ubp1c-1*. (C) Quantitative real-time polymerase chain reaction (qPCR) of multiple products was employed to determine the abundance of wildtype, mutant, and overexpressed *UBP1C* transcripts relative to the *UBC5* gene were present in 1 week old seedlings. Error bars indicate standard error of the mean for three biological replicates.

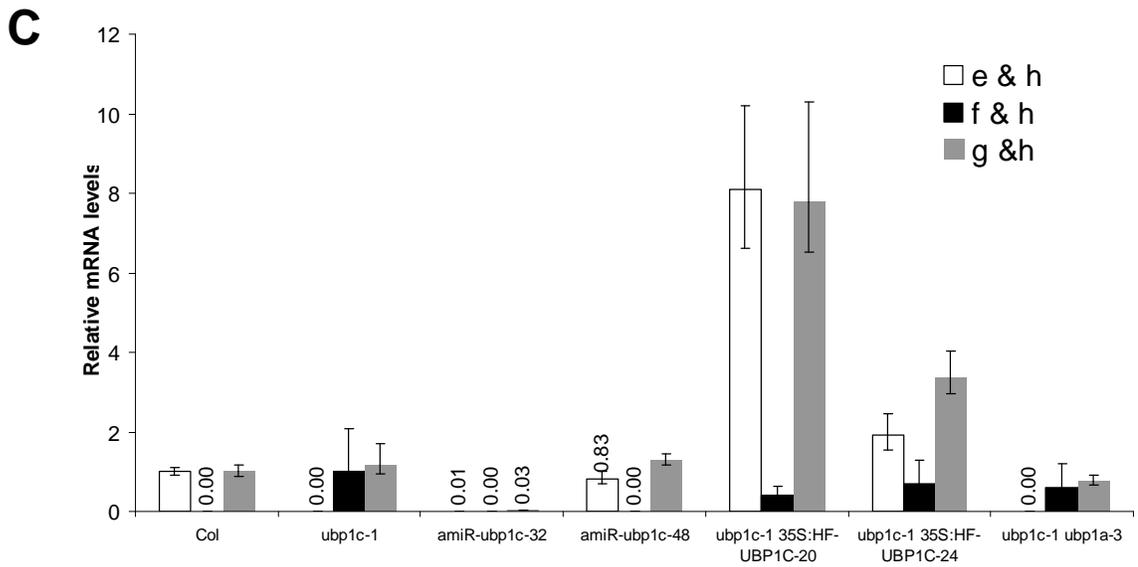
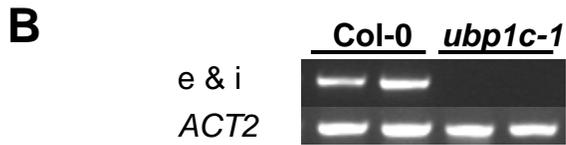
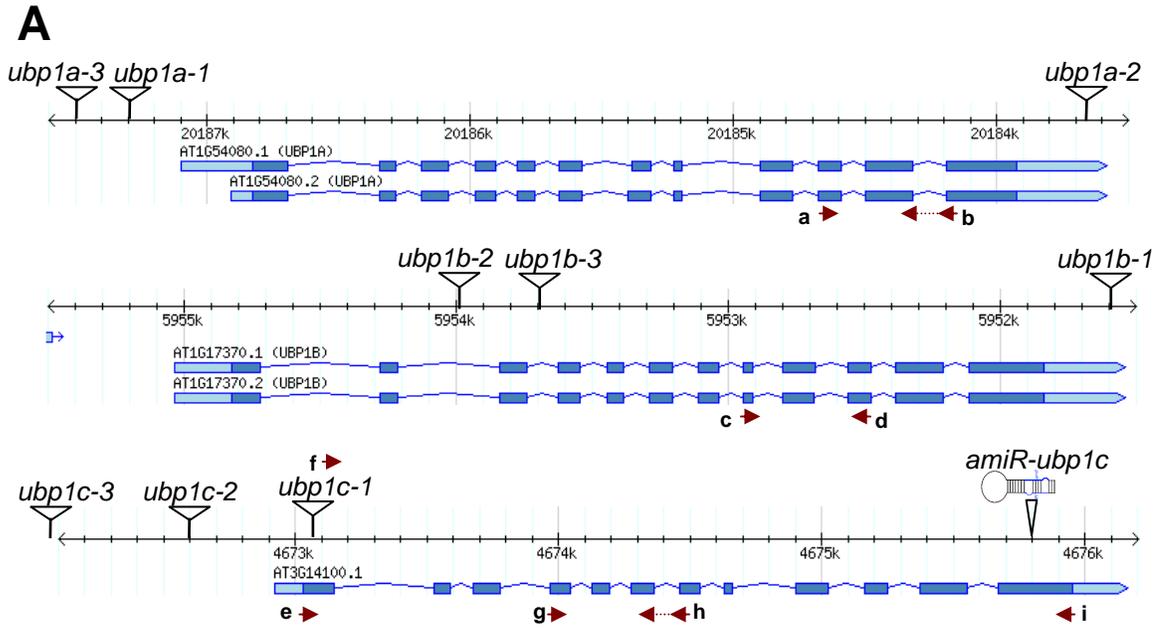
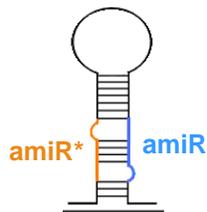
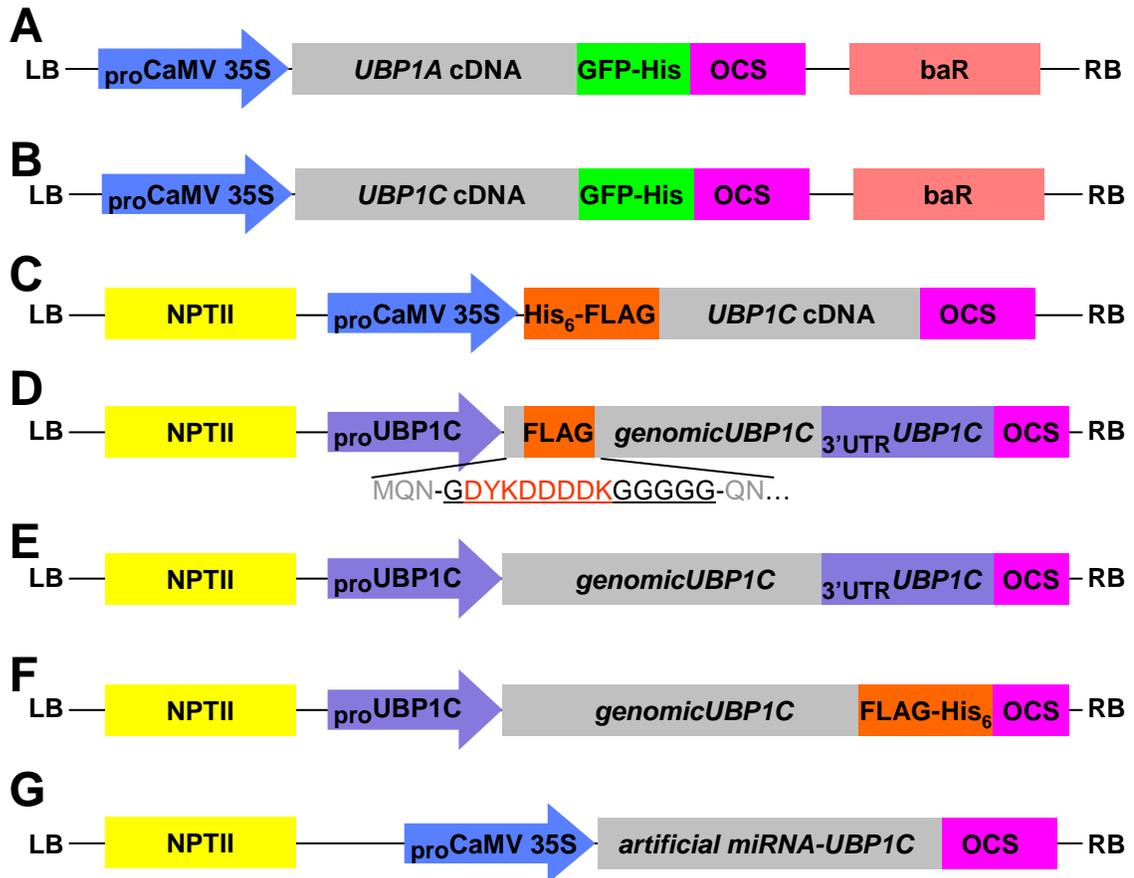


Figure 3.7. Diagrammatic representation of TDNA expression constructs used for transformation of *UBP1A* and *UBP1C*. (A) *35S::UBP1A-GFP*, (B) *35S::UBP1C-GFP*, (C) *35S::HF-UBP1C*, (D) *proUBP1C::FLAG-genomicUBP1C-UBP1C_{3'UTR}* (following the methionine the next two amino acids were retained at the amino terminus before and after the Flag insertion), (E) *proUBP1C::genomicUBP1C-UBP1C_{3'UTR}*, (F) *proUBP1C::genomicUBP1C-FH*, (G) *35S::amiR-ubp1c* (this gene contains a hairpin loop containing the amiR* (blue) and amiR (orange) sequences; match (*), G-U bubble (o), and mismatch [blank] are indicated).



amiR* 5' -AAUGGCACGUGUGAAUGCGUU-3'
 O** *****

amiR 3' -UUAUCGUGCACACUAACGCAA-5'
 O**O*** ***** *****

mRNA 5' -GAUGGCAAGUGUGAAUGCGUU-3'

Figure 3.8. *ubp1c-1* and *ubp1a-3* confer unique seedling growth phenotypes. Stratified seeds were grown on MS medium in vertically oriented plates in a growth chamber with a 16 h light / 8 h dark diurnal cycle. (A) 9 d old seedlings grown in the absence (-suc) or presence (+suc) of 1% (w/v) sucrose. *ubp1c-1* genotypes and *amiR-ubp1c* growth is arrested at 3 days after germination. Among seedlings grown with sucrose, root length is not altered in genotypes except *ubp1a-3*. (B) Penetrance of arrested growth. Seedling growth arrest in the absence of sucrose does not occur in more than 25% of seedlings other than in *ubp1c-1* and *amiR-ubp1c* genotypes. (C) (D) Primary root lengths of seedlings grown in the presence of increasing concentrations of sucrose. Root length was measured after 9 d growth and is significantly different between genotypes at each concentration (Student's *t*-test; $p < 0.001$); error bars represent standard error. (E) *ubp1a-3* plants have slower growing shoots and roots. 7 d old plants are shown. (F) Root length of plants in (E). Col-0 seedlings were significantly different than those of *ubp1a-3* (Student's *t*-test; $p < 0.001$). Error bars represent standard error; ND, not determined.

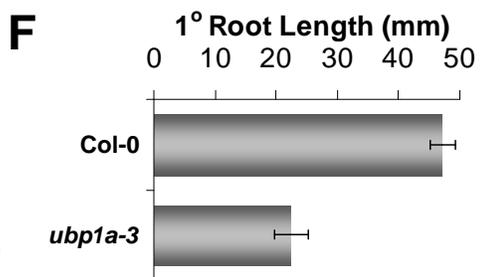
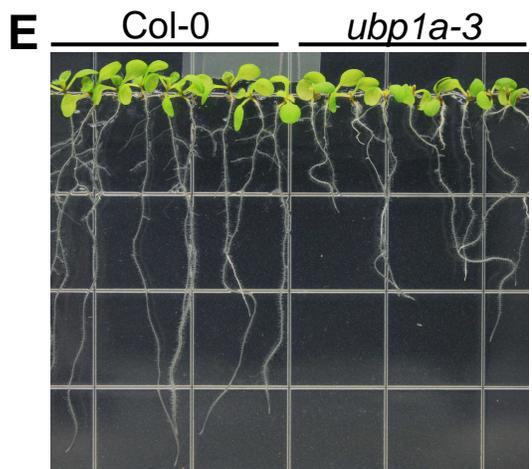
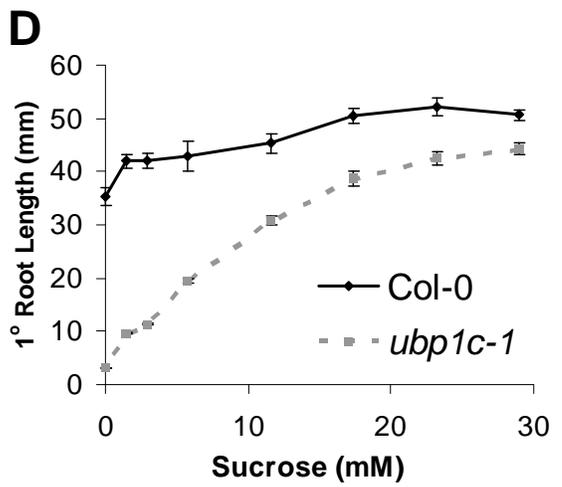
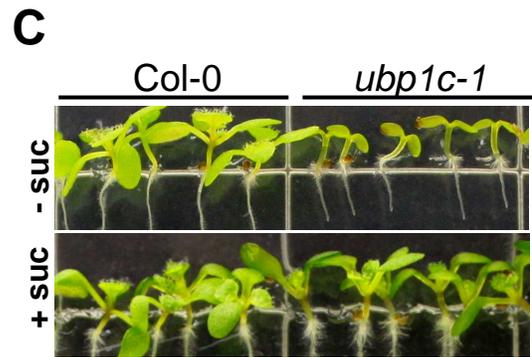
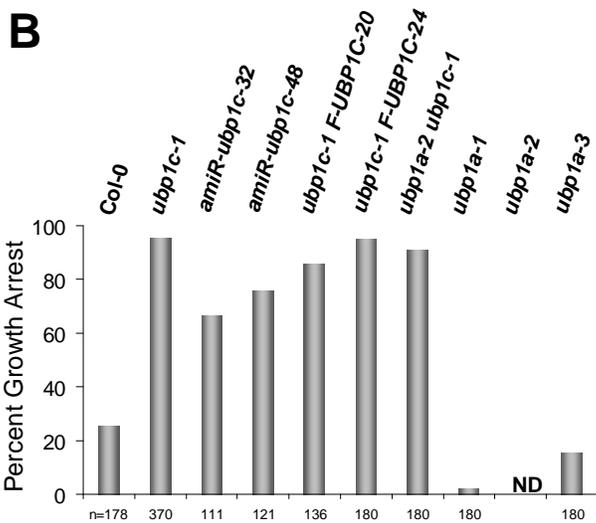
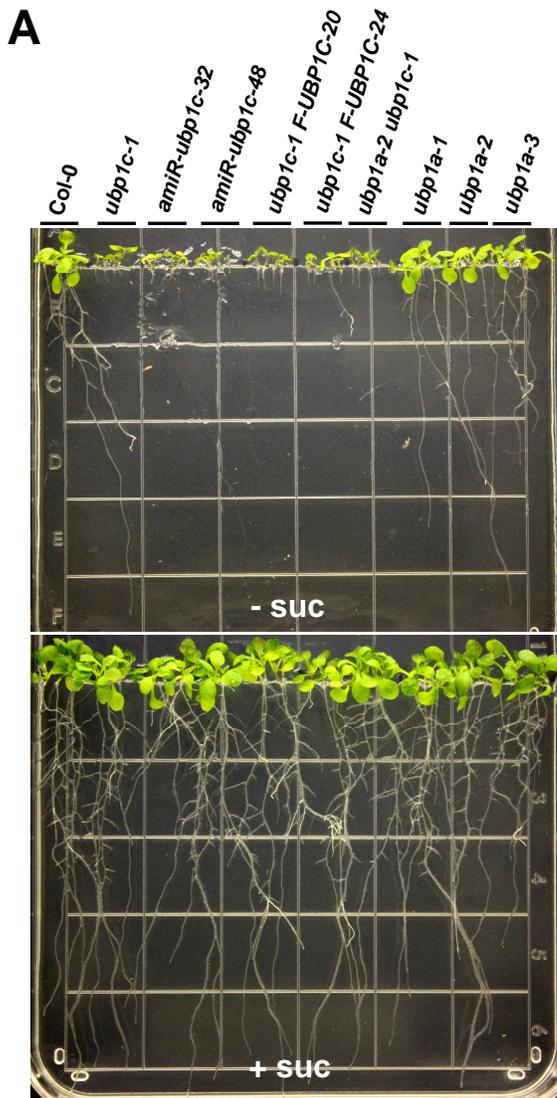


Figure 3.9. The *ubp1c-1* null allele confers pleiotropic phenotypes. (A) *ubp1c-1* seeds are darker than those of wild type. Seed color segregates in the F₂ generation of a Col-0 X *ubp1c-1* F₁ hybrid. Light (black arrow) to dark (white arrow) seed segregation was not different than a 3:1 ratio (Chi-squared p-value = 0.585, n=228). Parental and F₁ plants were simultaneously cultivated, their seed harvested and dried for three weeks. (B) Col-0 (upper) and *ubp1c-1* (lower) seeds were germinated and grown under continuous, low light (~10-15 μ E) for 4 d. (C) Starch granule accumulation is limited in *ubp1c-1* seedlings grown on MS medium without sucrose. Col-0 and *ubp1c-1* seeds were grown on medium in the absence (-suc) or presence of 1% sucrose (+suc) for 12 d. Immediately following the transition from light to dark, seedlings were fixed in 100% (v/v) ethanol overnight at 20°C, stained in Lugol's solution 5 min, and imaged. Cotyledon (1st column), apical hypocotyl (2nd column), and root tip (3rd column) regions are shown. (D) *ubp1c-1* seedlings accumulate less chlorophyll than those of wild type. Chlorophyll levels were determined spectrophotometrically following extraction in cold 80% acetone. Total chlorophyll and a/b ratios were significantly different between genotypes (p<0.05, Student's *t*-test). Error bars represent standard deviation of total chlorophyll. (E) Kinetin treatment reveals altered anthocyanin accumulation in *ubp1c-1* seedlings. Stratified seeds were germinated on solid MS medium containing 30 mM sucrose for 5 d and transferred to identical plates containing up to 40 μ M kinetin. Seedlings were imaged after 8 d of compound treatment.

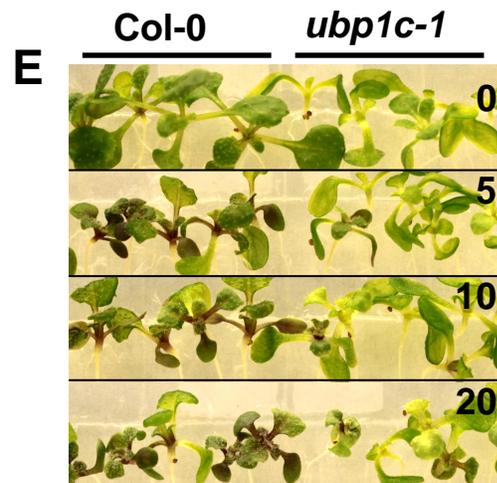
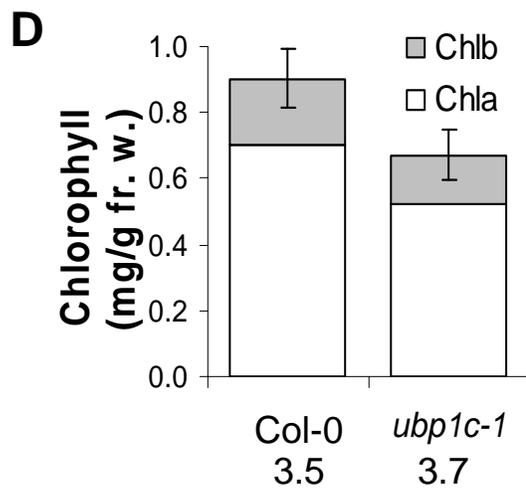
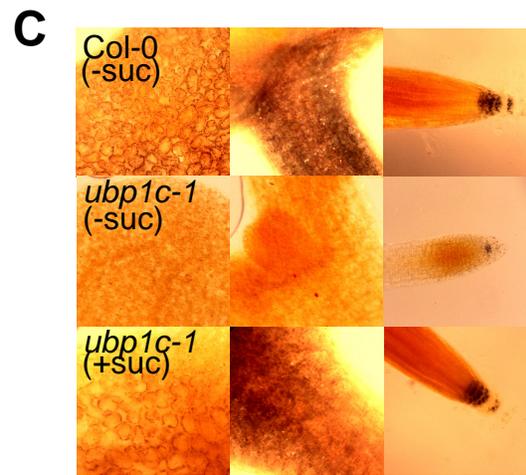
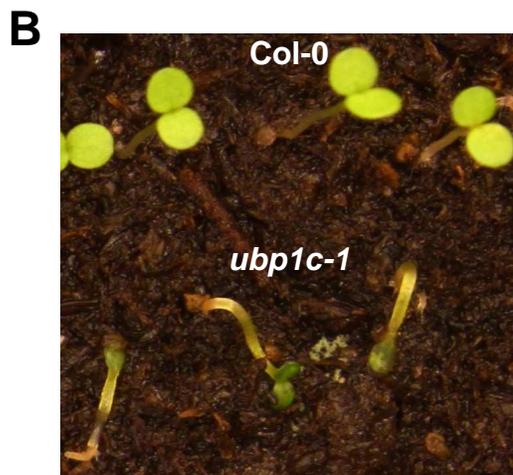
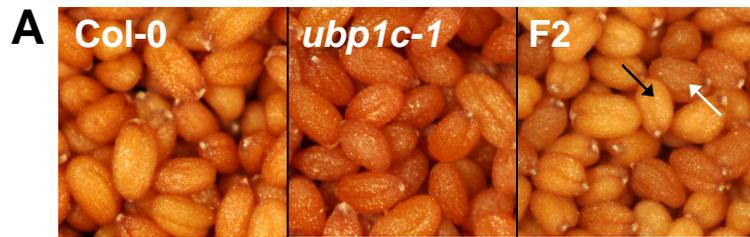


Figure 3.10. Expression of an artificial microRNA gene targeting *UBP1C* (*amiR-ubp1c*) in Col-0 recapitulates *ubp1c-1* phenotypes with varying severity. (A) Segregation of the seedling growth arrest phenotype in the T2 populations is not uniform. (B-C) T2 seed were separated by light and dark color and sucrose-rescued growth arrest was quantified for each group of seed for 5 independent lines. Seedlings from dark seed with arrested growth after 6d were transferred to media with sucrose and root length was measured before (C, upper right) and after (C, lower right) an additional 7d growth (D). (E) T3 *amiR-ubp1c* seeds. (F) 4-week old plants.

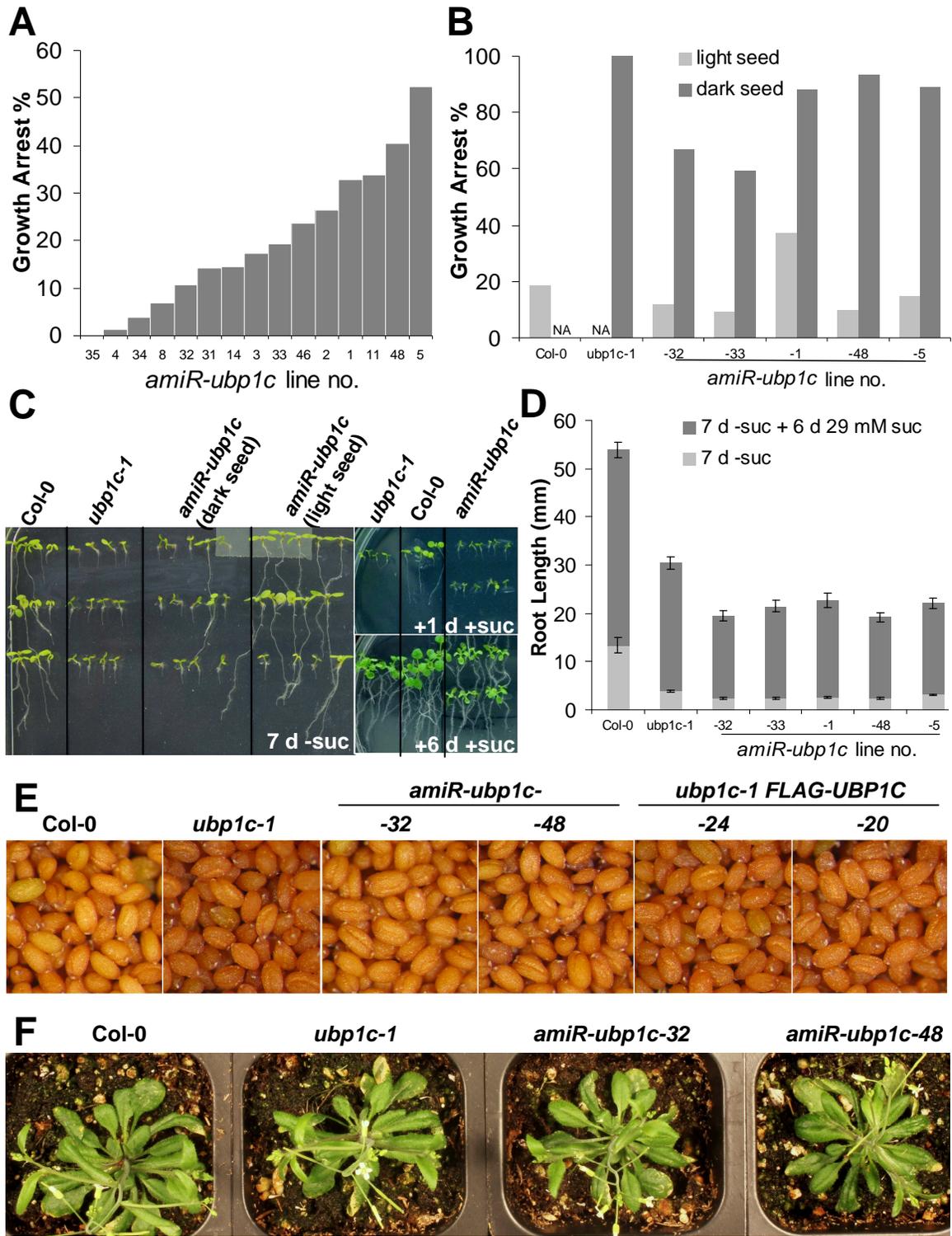


Figure 3.11. Ectopic overexpression or native promoter driven expression of UBP1C fusion proteins results in altered growth. (A) Two four-week old overexpression lines of UBP1C-GFP plants exhibit a dwarf phenotype with varying severity in the T2 population compared to Col-0. (B) Dwarf phenotype is observed among *ubp1c-1 proUBP1C:FLAG-genomicUBP1C-3'UTR_{UBP1C}* plant lines as well. (C) Defects in terminal flowers were apparent in overexpression lines including multiple fused or open carpels (a), twisted and malformed gynonecia (b), and petaloid stamen (arrows). (D) When expressed in the *ubp1c-1* background and grown without sucrose GFP fluorescence co-segregated with a premature yellowing and seedling death in a 3:1 ratio in multiple insertion lines.

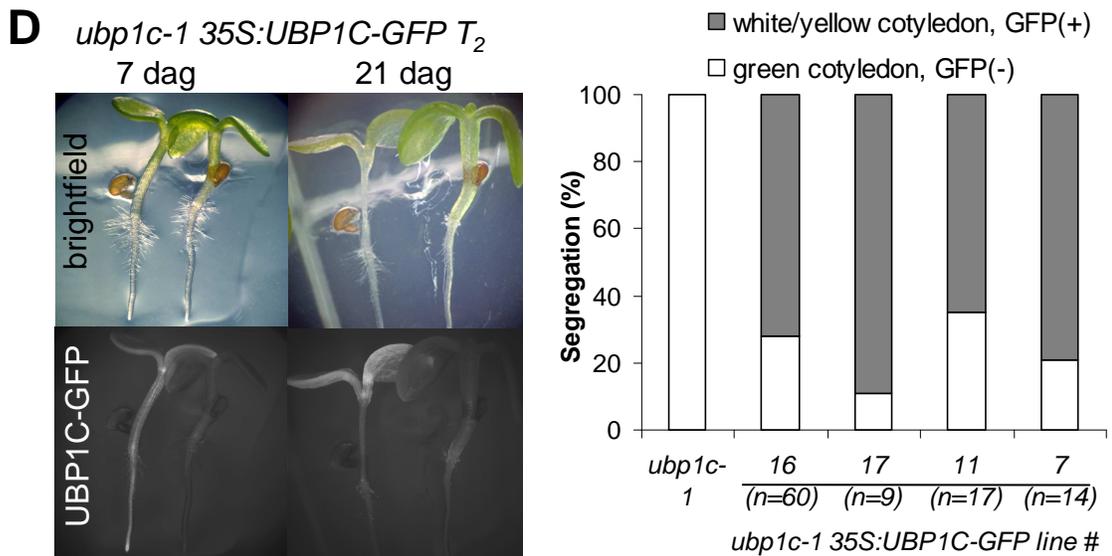
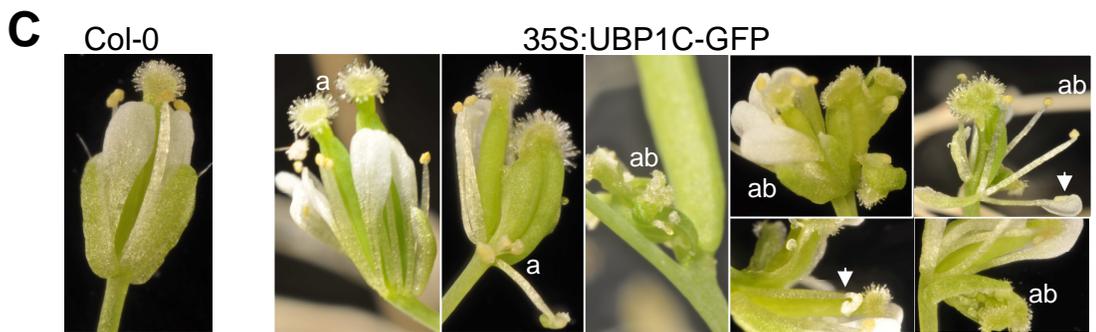
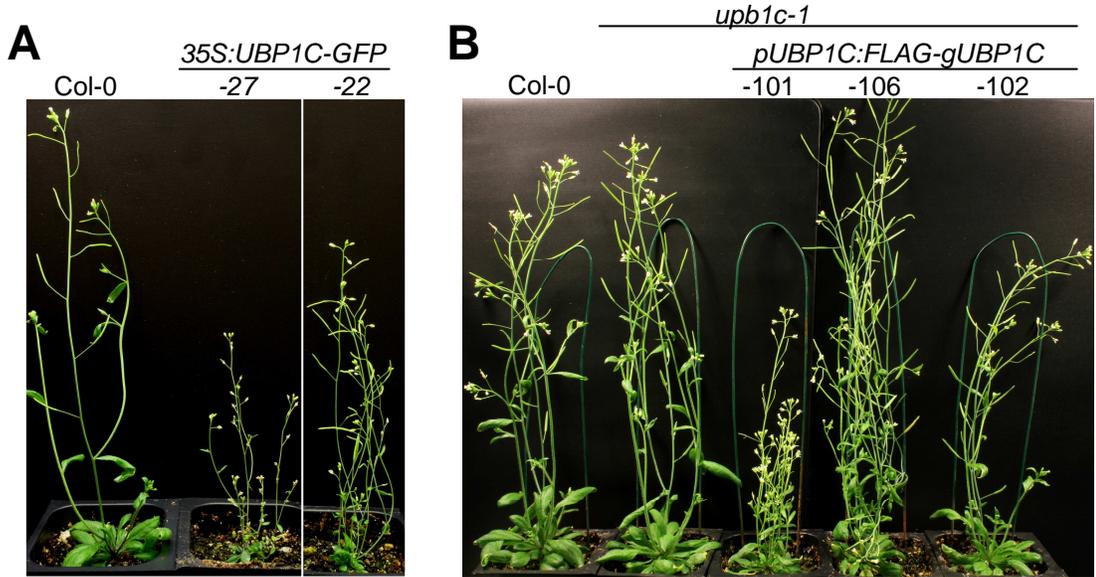


Figure 3.12. *ubp1c-1* seedlings are hypersensitive to oxygen deprivation. Columbia-0 (Col), *ubp1c-1*, *amiR-ubp1c-32*, *ubp1c-1 35S::HF-UBP1C-20*, *ubp1c-1 35S::HF-UBP1C-24* plants were grown on MS agar containing 1% sucrose in vertically oriented plates in a growth chamber with a 16 h light / 8 h dark diurnal cycle for 10.5 days. 2.5 h following the light to dark transition, plates were transferred to 13.5 L plexiglass chambers in very low light ($\sim 1\mu\text{E}$) which were purged with 99.99% Argon gas at $\sim 33\text{ mL s}^{-1}\text{ chamber}^{-1}$ for 4 h followed by Argon flow of $10\text{ mL s}^{-1}\text{ chamber}^{-1}$ until the plants were removed after the 10, 11.5, and 13 h hypoxia stress (HS). Surviving plants were subsequently counted following 6 days recovery at day 17 of growth. Survivors were identified by green tissue growth from the shoot apical meristem. (A) Percent survival of 10 seedlings of each genotype was measured from 12-14 replicate plates. Bars with a similar letter are not significantly different (adjusted p-value < 0.05, one-way ANOVA, TukeyHSD). (B) Representative plants of 0 h (Mock) and 13 h hypoxia treatment following a five day recovery period are shown.

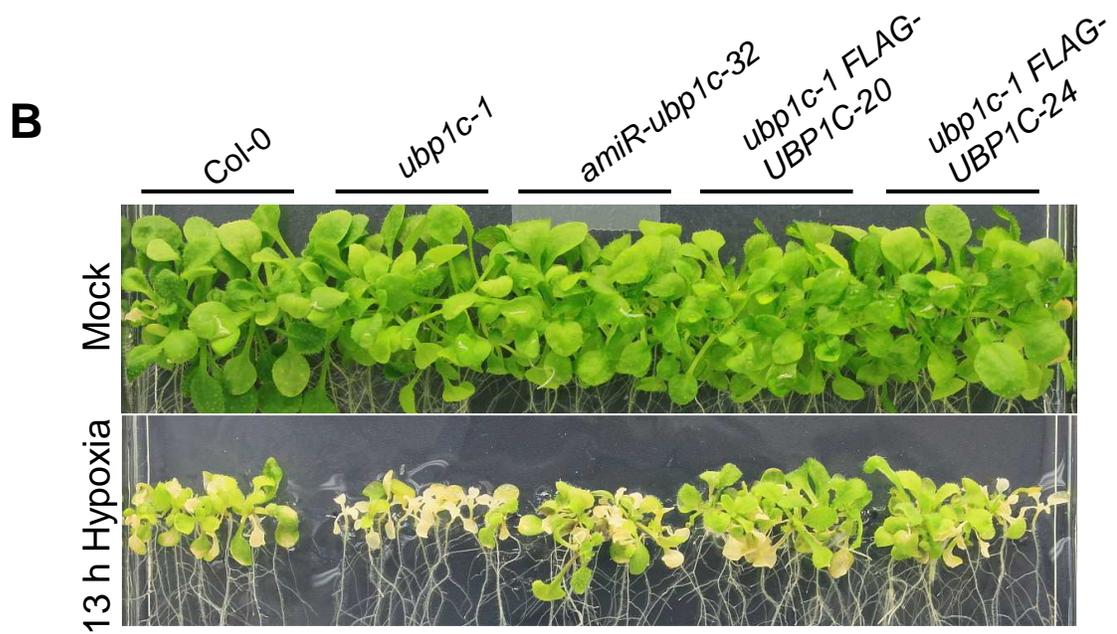
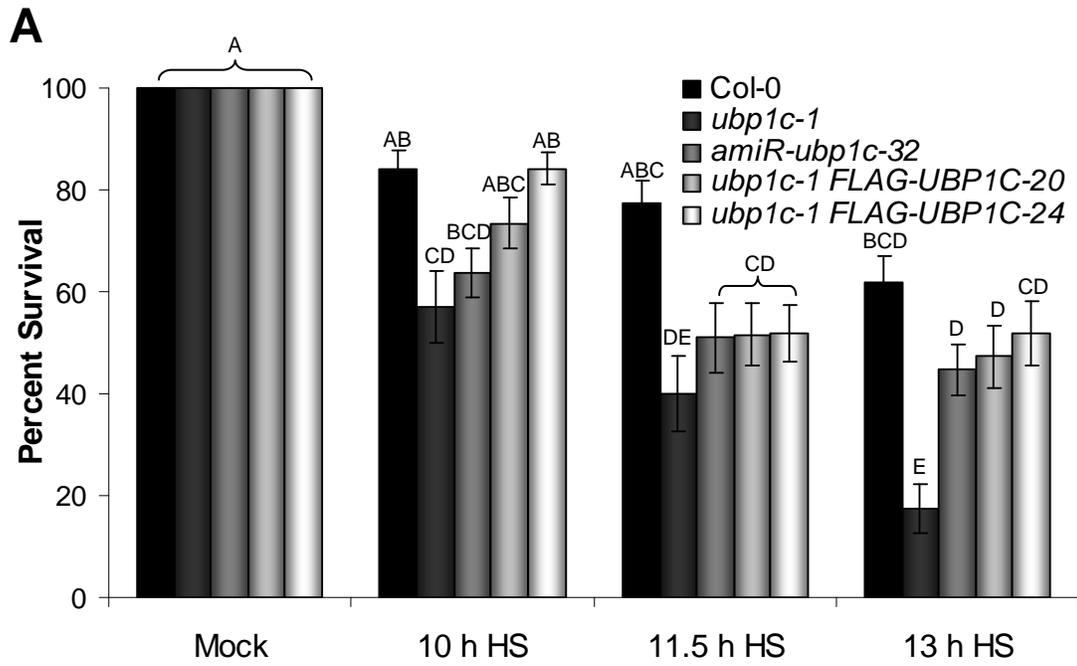


Figure 3.13. Altered phylotaxy of *ubp1a-2* and *ubp1a-3* are similar. (A) 4-week old plants of *UBP1* genotypes. (B) *ubp1a-2* (~50% penetrance) and *ubp1a-3* (100% penetrance) exhibit similar phenology in their phylotaxy including twin flower/silique (a), twin/triplet cauline leaf w/buds (b), kinked inflorescence stem (c), aborted cauline leaf (d), pedicel terminated leaf and axillary bud (e), and cauline leaf with adjacent flower (f).

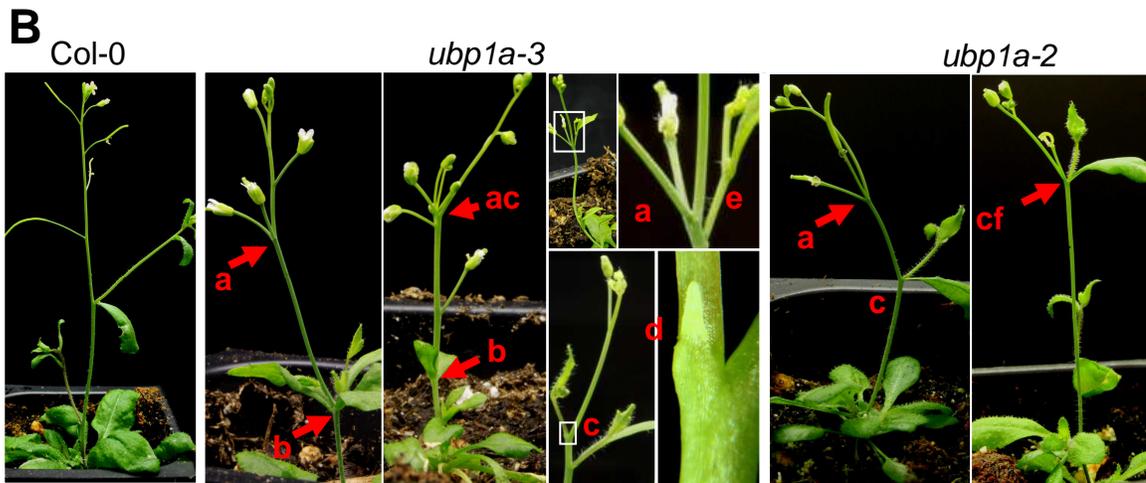
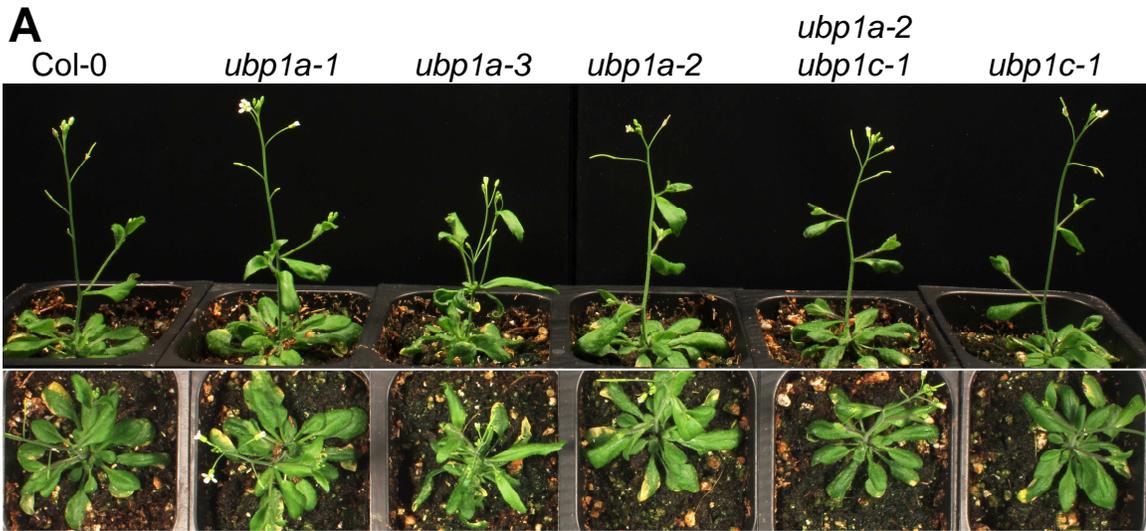


Figure 3.14. *UBP1A* mutant alleles are not sensitive to hypoxia and *ubp1a-2* rescues the hypersensitivity to hypoxia of *ubp1c-1*. (A) *UBP1* genotypes were grown, treated and scored as in Figure 3.12. (A) Percent survival of seedlings of each genotype was measured from 12-14 replicate plates. Bars with a similar letter are not significantly different (adjusted p-value < 0.05, one-way ANOVA, TukeyHSD). (B) Representative plants of 0 h (Mock) and 13 h hypoxia treatment following a five day recovery period are shown. Note: *ubp1a-2* seedlings did not have uniform nor complete germination and were smaller upon treatment.

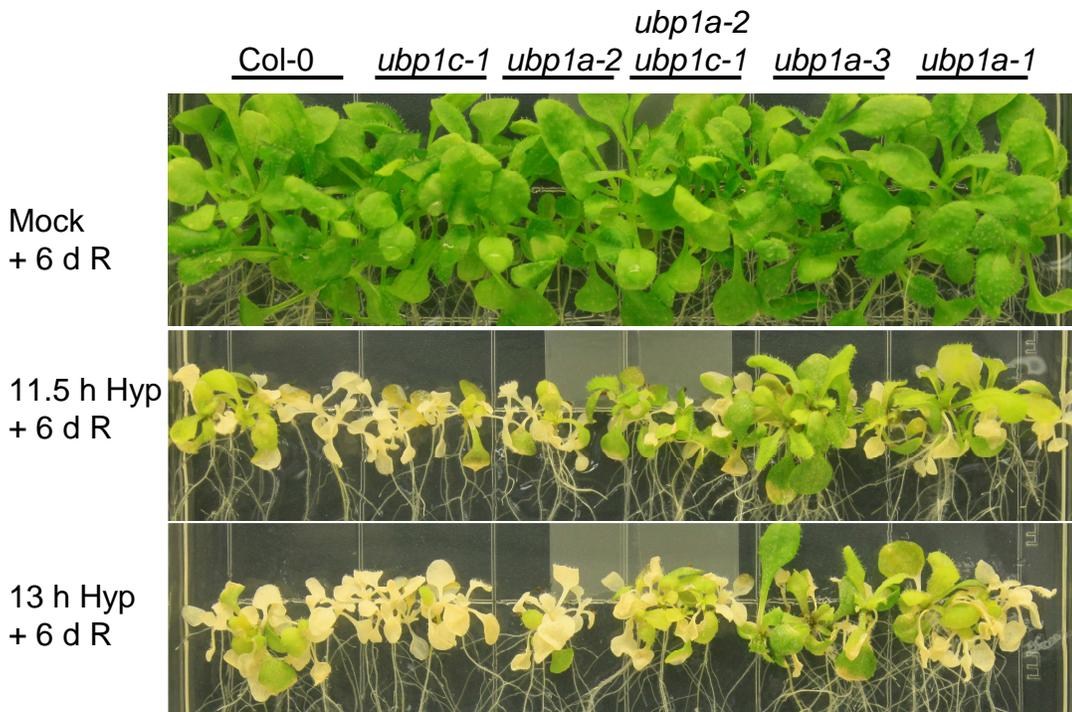
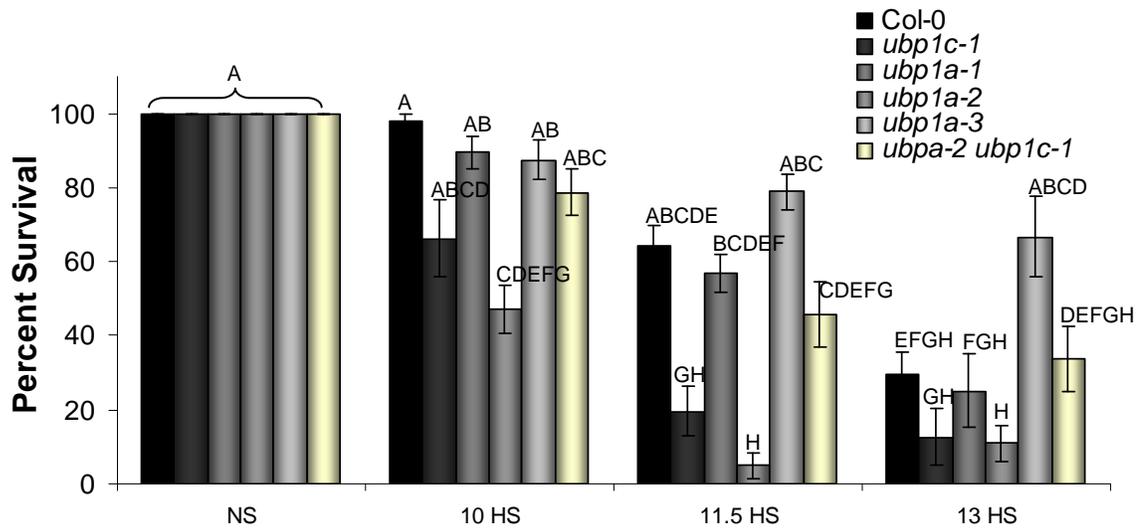


Figure 3.15. Cytoplasmic UBP1C-GFP dynamically and reversibly aggregates into granular structures in response to hypoxia treatment. (A) UBP1C-GFP cytoplasmic granule formation in response to hypoxia. A section of cauline leaf from *ubp1c-1 35S::UBP1C-GFP-7* was submerged (S) under a coverslip and imaged after 8, 25, and 45 min. A 3-dimensional reconstruction from a 10 image z-stack was made and analyzed using Imaris Bitplane Software. GFP and chlorophyll fluorescence channels are false colored green and red, respectively (volume dimensions = 97 μ m x 97 μ m x 9 μ m; tick marks = μ m). (B) Using Bitplane spot-finding algorithm granule number (left) and intensity (right) were determined for rendered volumes in (A). (C) In an experiment similar to (A), three-dimensional reconstruction time-lapse video microscopy shows the granule formation kinetics consistent with a lag phase of oxygen depletion, a response phase, and saturation of the response. Partial initial and final images are shown corresponding with adjacent arrows. (D) Submergence with Oxyrase® induces granule formation much more quickly than does submergence alone. Within 4 min of submergence plus Oxyrase® granules from seedlings of *ubp1c-1 35S::UBP1C-GFP-16* have formed and after 22 min almost all cytoplasmic GFP signal has entered large granules (left 2 images). The guard cells in the white square are enlarged for comparison (right two images) of GFP distribution (volume dimensions=95.3 μ m x 95.3 μ m x 23 μ m; tick marks = μ m). (E) Granule formation reverses upon reoxygenation. 5 day old plants *ubp1c-1 35S::UBP1C-16*. Cells were imaged on a Leica SP2 confocal laser scanning microscope at specified time intervals following coverslip submersion in 0.25X MS salt solution. The plants were desubmerged after 60 min and seedlings were exposed to air for 20 min (60' S +20' R) and same cells were imaged immediately. Fluorophores were excited with an Argon laser at 488nm wavelength at 25% power, emission wavelengths between 500-600 nm were collected. Maximum projection of a 16-image z-stack is shown for each time point (bars = 10 μ m).

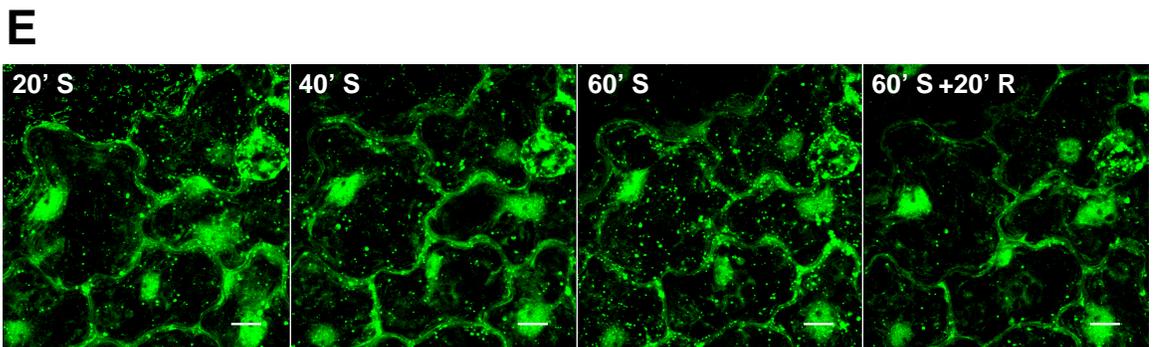
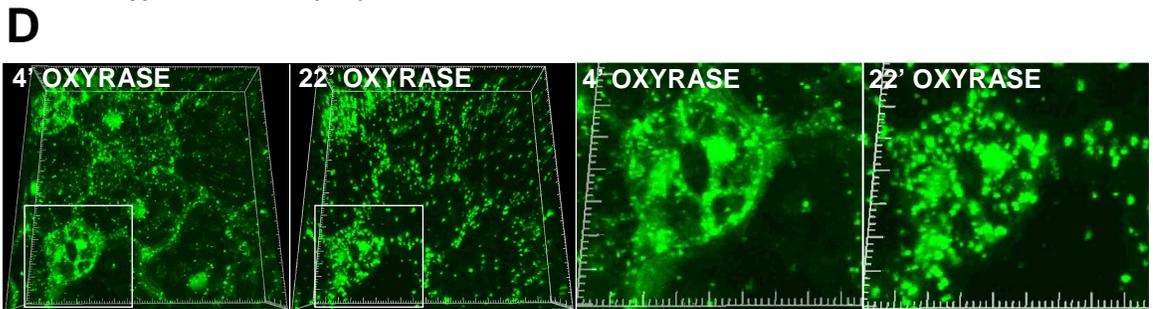
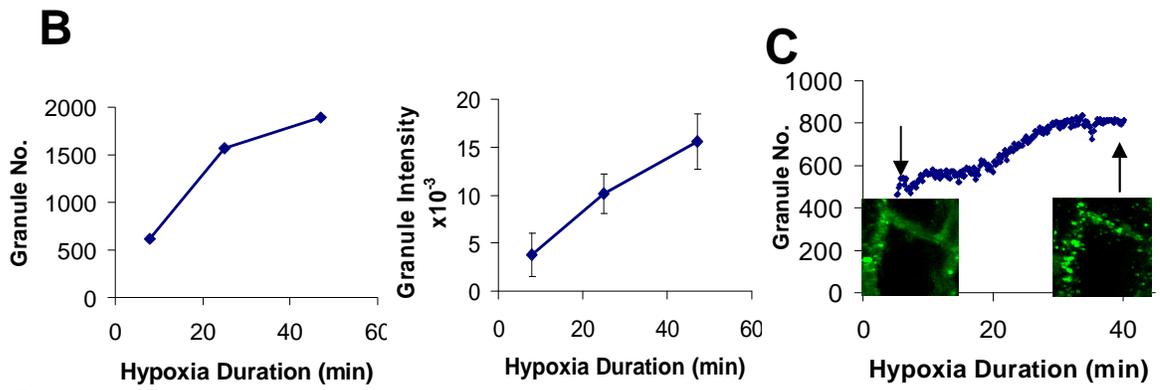
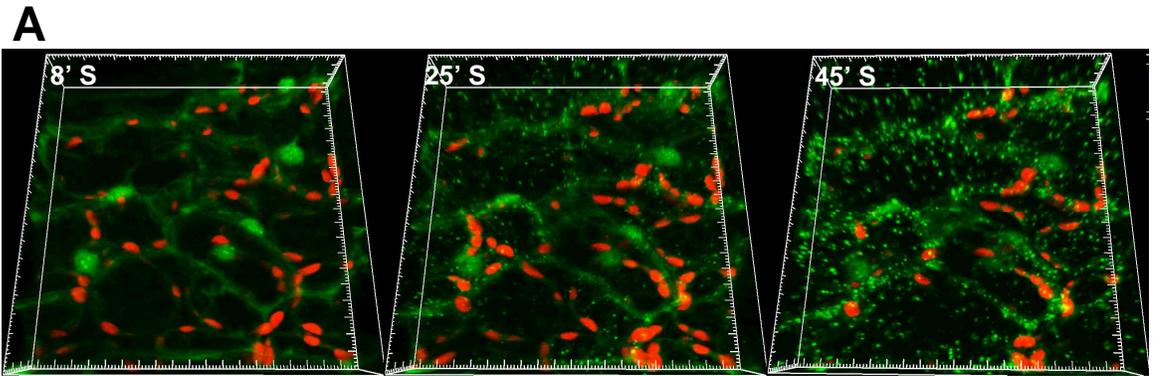


Figure 3.16. UBP1C granule formation is inhibited by cycloheximide in hypoxia and heat. (A) Hypocotyl epidermal cells of *ubp1c-1 35S::UBP1C-GFP-16* were treated with 1 h dark (1 h D), 1 h hypoxia imposed by coverslip application and incubation in an Ar(g) purged chamber 1 h in 0.4% DMSO with (+cyc.) or without 200 ng μl^{-1} cycloheximide in dark (1 h H) or maintained in normal growth conditions in light until imaging (1 h L). Bar=20 μm . (B) Heat-induced UBP1C-GFP granules are also prevented by cycloheximide treatment. *ubp1c-1 35S::UBP1C-GFP-16* seedlings were incubated in dark in 0.25X MS 0.4% DMSO solution with or without cycloheximide on an uncovered slide in a humidified container at 20°C or 41°C for 2 h. Confocal images were taken within 5 min following coverslip application as in Figure 3.15. Bar = 15 μm .

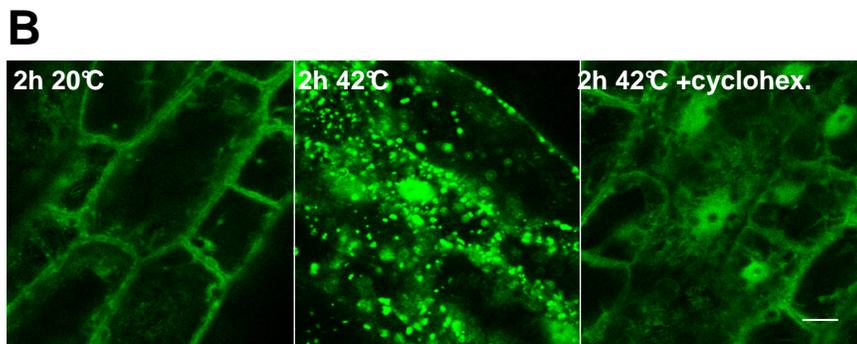
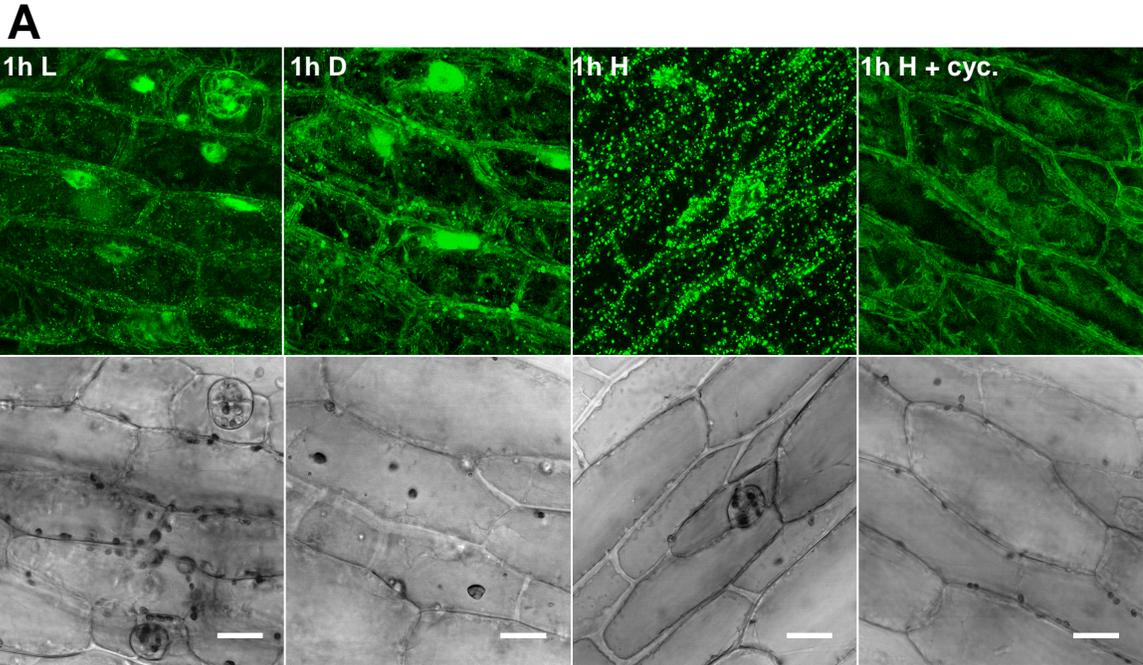


Figure 3.17. UBP1A may have redundant function with UBP1C. (A) Plants over expressing UBP1A cDNA fused to GFP were generated to determine the localization of UBP1A in response to hypoxia. An intact 5d.o. *35S:UBP1A-GFP-33* plant was submerged in water centered under a coverslip. Hypocotyl epidermal cells were imaged following submergence (S) or submergence plus re-aeration (R). Images were taken on Leica SP2 at 8 z-planes evenly spaced over ~25um depth. Fluorophore was excited with an Argon laser at wavelength of 488nm and imaged as in Figure 3.15. Each stack took ~5 min 25 s. Images were processed on ImageJ software: Z-series were combined using the maximum projection function, followed by applying a median filter with a radius of 2 pixels, and adjustment of the gamma to 1.2. Granules were counted using imageJ software.

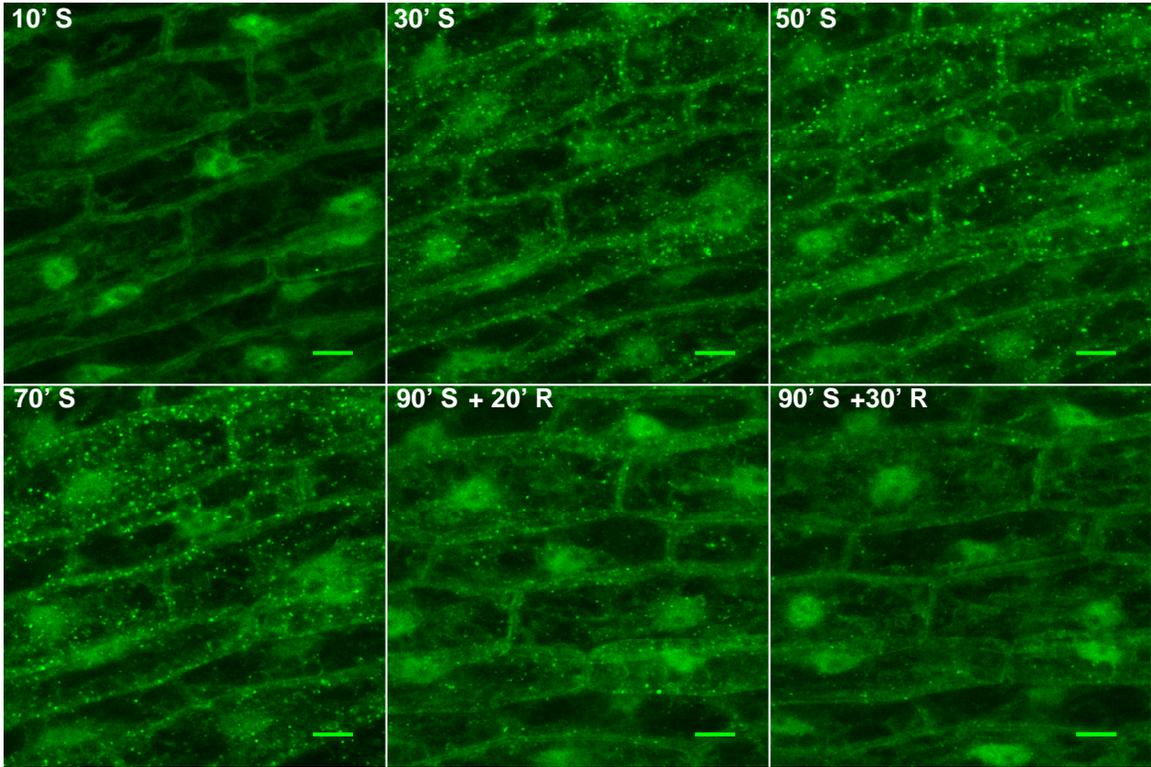
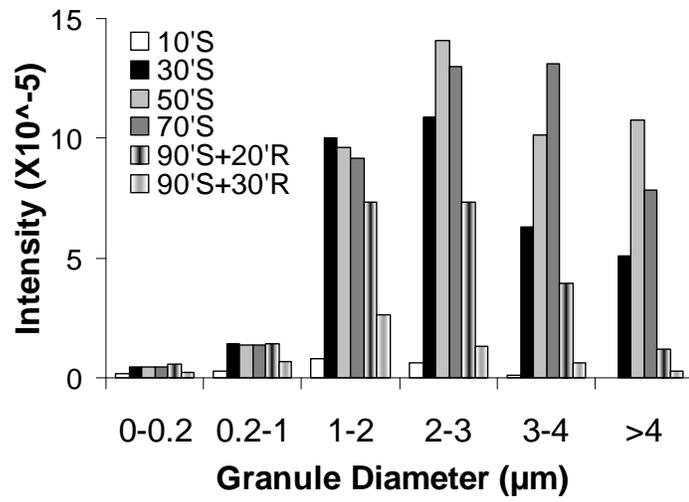
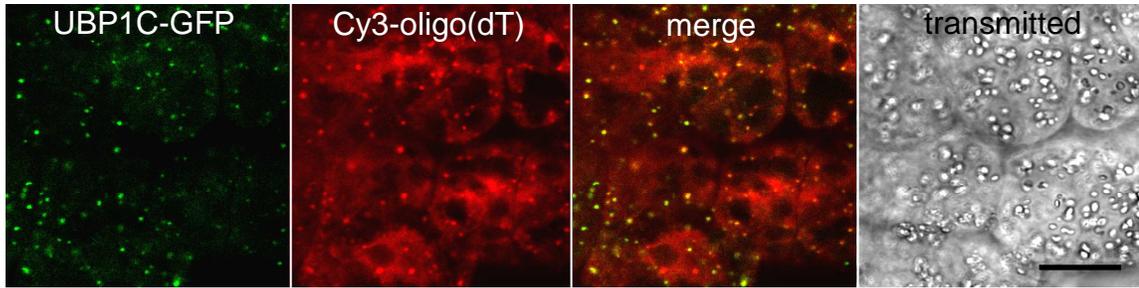
A**B**

Figure 3.18. AtUBP1C-GFP granules contain polyA+ RNA. (A) Mesophyll cells of 3-d-old seedlings over expressing AtUBP1C-GFP-16 were treated 3 h in a plexiglass chamber completely purged with 2% O₂, 0.037% CO₂, 97.963% N₂ and then immediately fixed. Fixed tissue was hybridized with Cy3-oligo(dT). (B) In an *Agrobacterium*-mediated transient transformation assay in leaves of *Nicotiana plumbaginifolia* AtUBP1C-GFP and AtPAB2-mRFP co-localized in granules. Fluorescent images were made on a Leica SP2 confocal microscope and processed using ImageJ (bars = 20 μm).

A



B

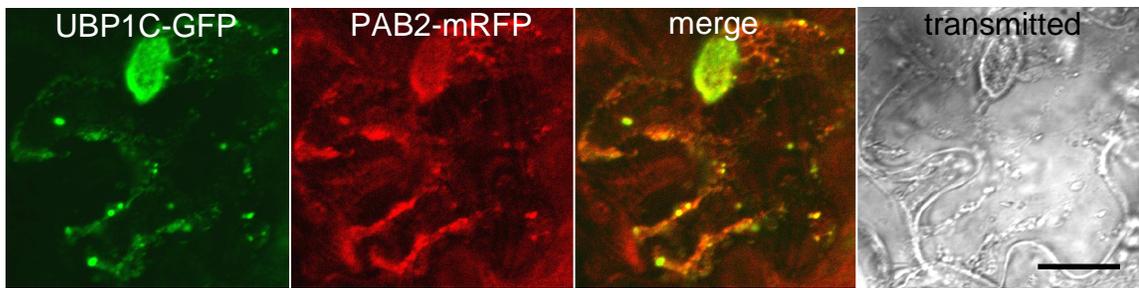


Figure 3.19. Immunoprecipitation of UBP1C mRNA complexes. 7-day-old *ubp1c-1 35S:HF-UBP1C-20* seedlings grown on MS containing 1% (w/v) sucrose were maintained under standard conditions (*i.e.* in the light and in air (L)), or transferred to dim light in air for 2 h (D), deprived of oxygen for 2 h in dim light (H), or deprived of oxygen for 2 h and then recovered for 20 min in air in dim light (R). Seedlings were rapidly harvested and frozen in liquid nitrogen. Immunoprecipitation of UBP1C (IP) was accomplished with the following proportions of tissue and reagents: 200 mg ground tissue was hydrated in 3.0 mL RNP extraction buffer (RNPB) and purified with 200 μ L Dynabeads per 20 μ g α -FLAG M2 antibody. (A) Tissue was resuspended in extraction buffer and filtered through miracloth. Filtrate was centrifuged at $\sim 1500 \times g$ to remove dense, intact organelles and cellular debris (P-1.5K). Supernatant (S-1.5K) was transferred to a new tube for IP and aliquots were saved for analysis by 12% (w/v) SDS-polyacrylamide gel electrophoresis (PAGE) and total RNA extraction. α -FLAG M2 antibody was bound to protein G Dynabeads 1 h in PBST and excess antibody was removed by magnetic separation. Antibody bound beads were combined with S-1.5K and incubated at 4°C with rocking for 1.5 h. Beads were magnetically collected and the IP supernatant was saved as the unbound fraction (UNB). Dynabead-bound UBP1C:RNA complexes were washed 6 times, 5 min each, in RNPB and eluted with 100 μ L 400 ng μ L⁻¹ 3XFLAG peptide. Post-elution, the beads were rinsed 1x in RNPB and saved for PAGE analysis. The eluate (IP) was split for mRNA purification (80%) and (B) immunoblot (IB; 10%) and PAGE analysis (10%). Fractions were diluted into 1X SDS loading buffer and loaded on duplicate 12% polyacrylamide gels. IB analysis was of 0.33% of the total and UNB samples; silver stained gels were 0.11% of the total sample. For the UBP1C-IP, 7.5% of the total IP was loaded on the IB and 2.5% for PAGE. The mobility of molecular weight markers are indicated at the left of each gel in kDa. Blots were probed with α -FLAG and α -RPS6 antibodies. Gels are representative of more than three independent experiments. (C, D) Quantitative real-time reverse-transcription polymerase chain reaction was used to measure relative abundance of mRNA in the UBP1C-IP and S-1.5K (Total) fractions. Data are representative of three independent experiments. UBP1C-IP relative to Total mRNA abundance confirmed preferential association of mRNAs to UBP1C (C) and mRNA abundance relative to *RPL37B* demonstrates anticipated increase in the anaerobic response marker gene *ADH1* in response to the stress treatment (D). Data are the mean \pm SD of three technical replicates.

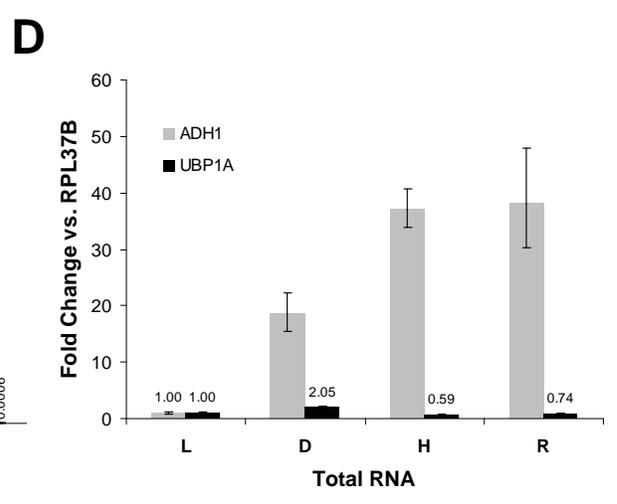
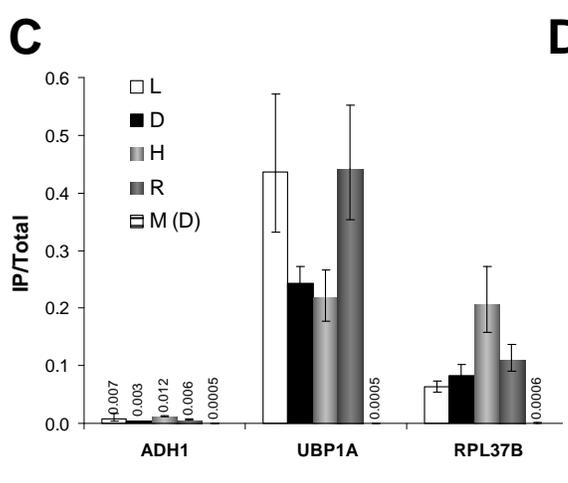
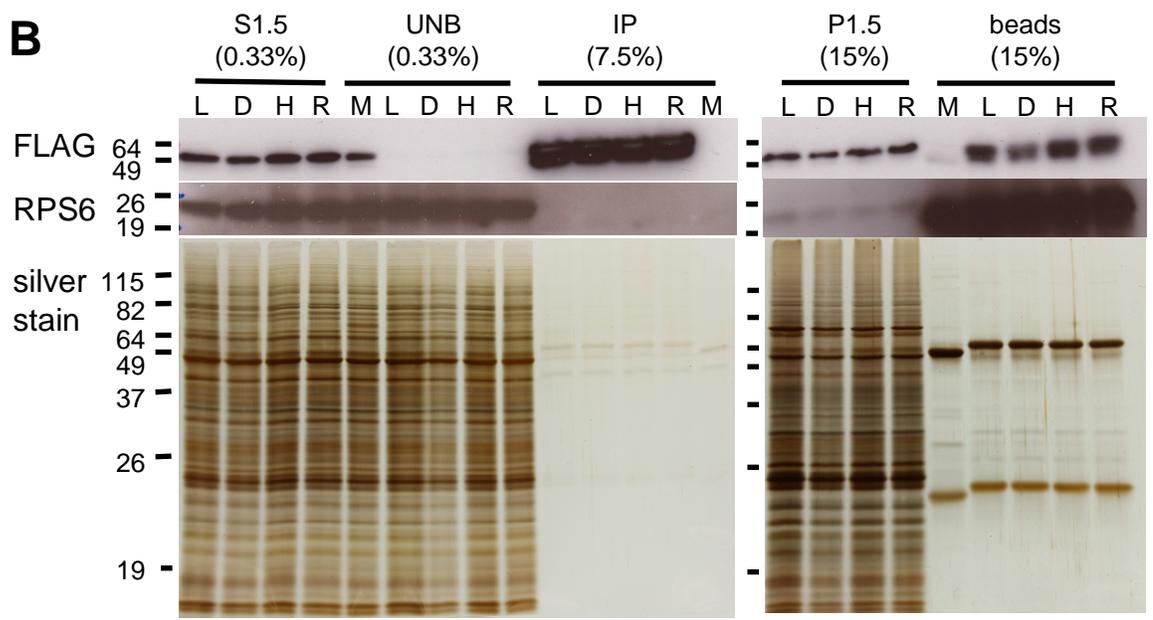
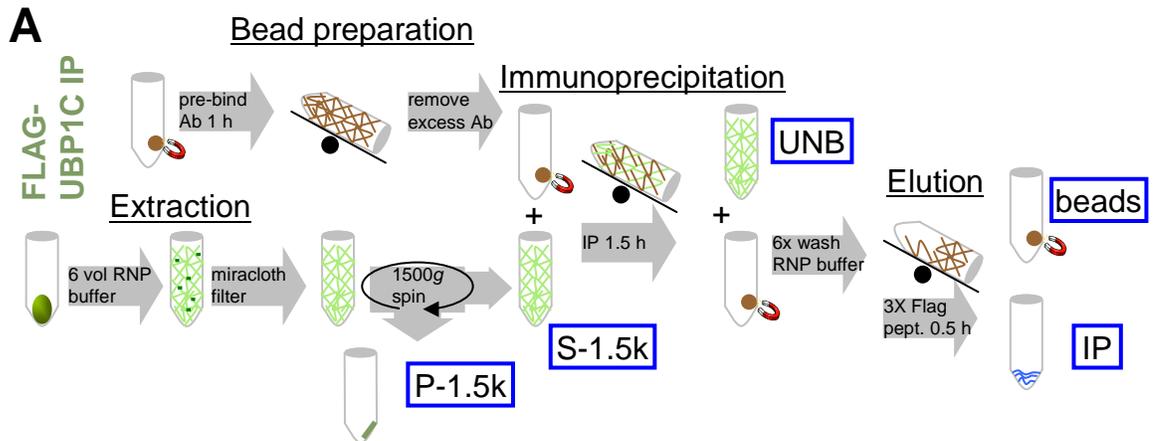


Figure 3.20. Microarray detection of total and UBP1C-associated RNAs is reproducible. RNAs isolated from S-1.5k (Total) and immunoprecipitated FLAG-UBP1C complexes were extracted and probed with Affymetrix ATH1-22k Genechips®. RMA signal log values of 2 bioreplicates are plotted against one another from each treatment. Blue lines represent linear regression fit for each dataset. Pearson correlation coefficients for each comparison label each plot.

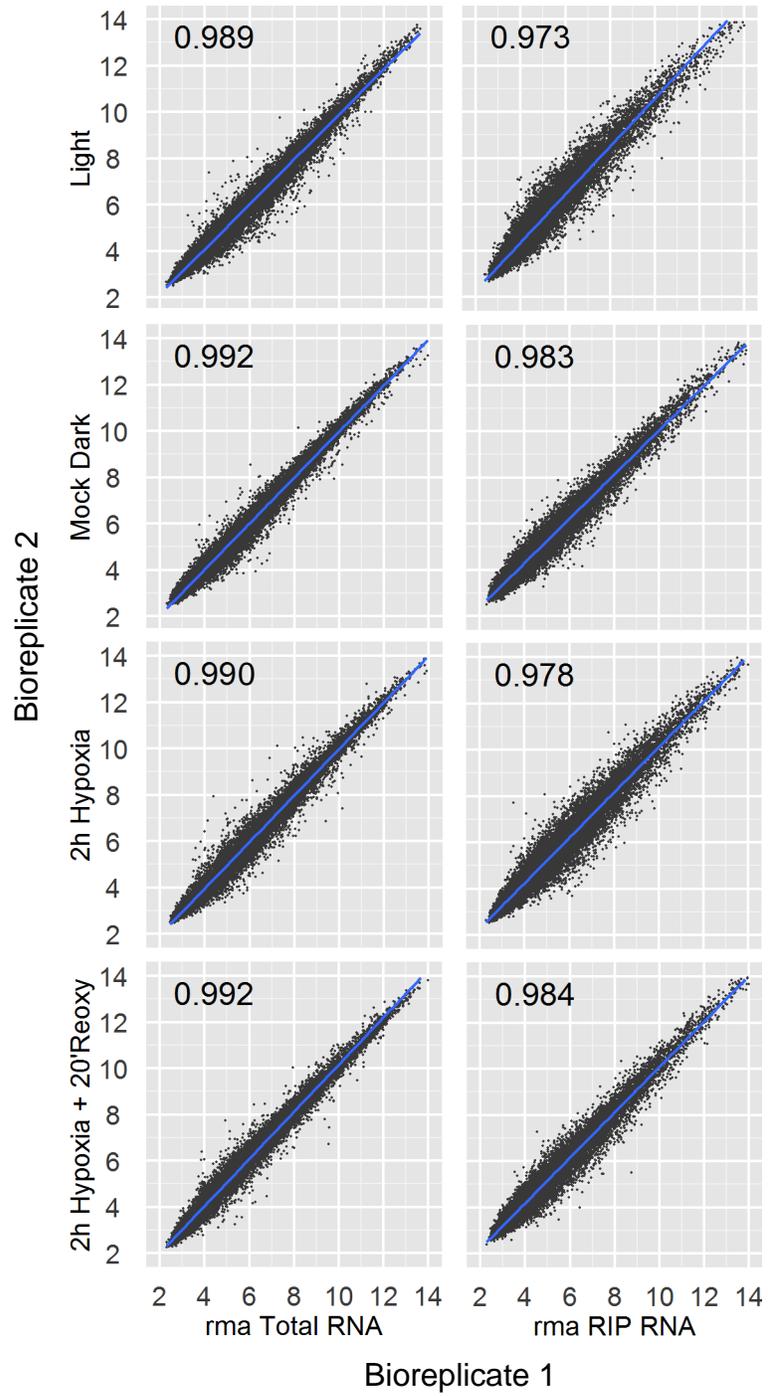


Figure 3.21. Distributions of microarray probeset signal log intensities from total and IP samples are not similar. Following background adjustment, and scaling (including scaling of polysome factors based on polysome density), the density of log base 2 summarized probeset signal values of combined replicate chips were compared across treatments of UBP1C-associated RNAs (A), UBP1C total RNAs (B), and RNAs of total and polysomal RNAs of Branco-Price et. al. (C).

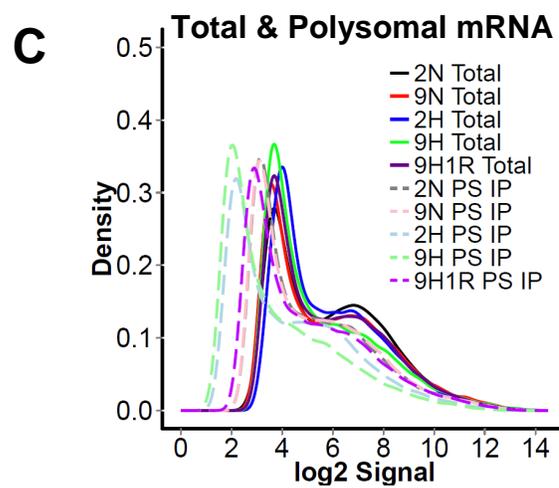
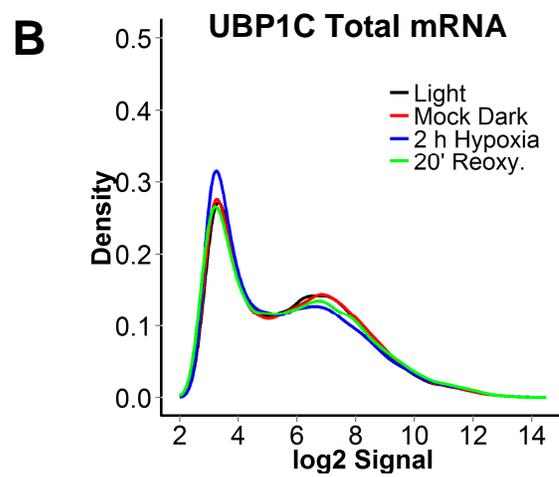
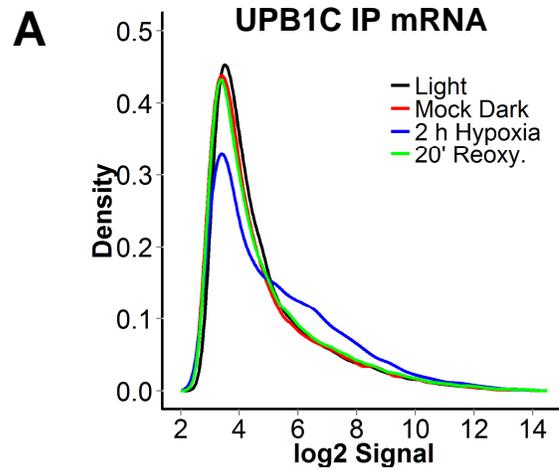


Figure 3.22. Fraction and treatment comparison mRNAs detected by microarrays. (A) Venn diagrams of MAS5 “present” and “absent” calls within total RNA versus UBP1C IP RNA. Distribution of these calls between treatment samples is shown. (B) Venn diagram of differentially expressed genes comparing treatments (D, 2 h dark; H, 2 h hypoxia; R, 2 h hypoxia plus 20 min reoxygenation) versus control (L, light) (cutoff: adjusted p-value<0.05, |SLR|>1.0). Venn diagram (C) and scatter plot (D) comparing differentially expressed genes of total RNA from Branco-Price et al. [10] 2 h hypoxia versus 2 h mock (RPL18 2H/2N) and this study 2 h hypoxia versus control 2 h light (UBP1C H/L).

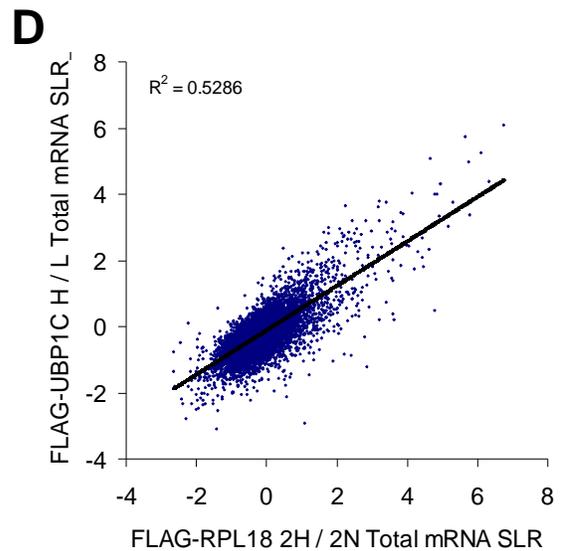
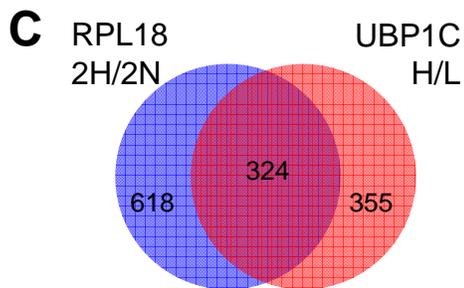
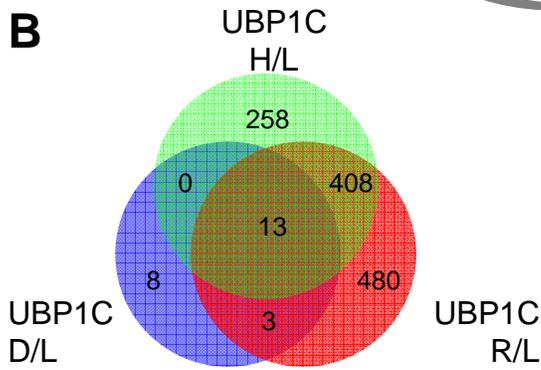
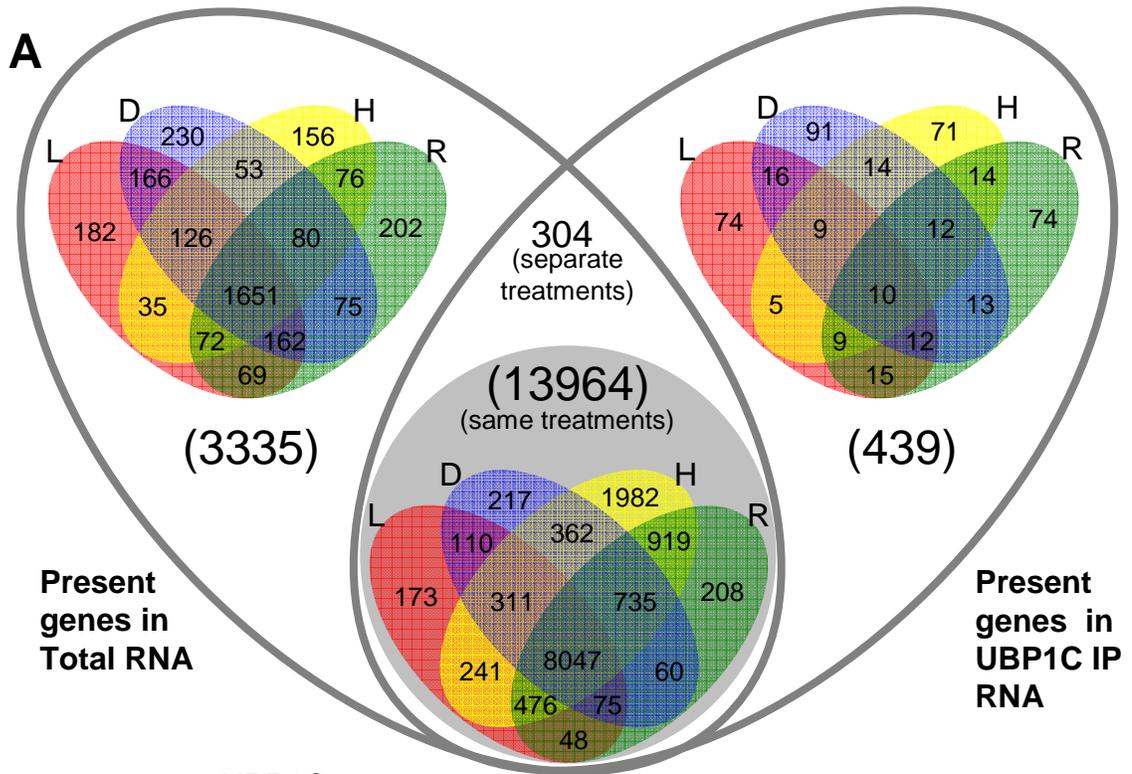


Figure 3.23. Transcripts poorly associated with FLAG-UBP1C reversibly increase association during hypoxia. Immunoprecipitated RNA (IP) abundance was normalized to transcript abundance from whole seedling total RNA (T) extracted from the S-1.5k fraction as a metric of association. Gene transcript levels associated with UBP1C complexes from plants treated with mock darkness, 2 h hypoxia, and 2 h hypoxia plus 20 min reoxygenation in air were compared against those of control plants maintained in light.

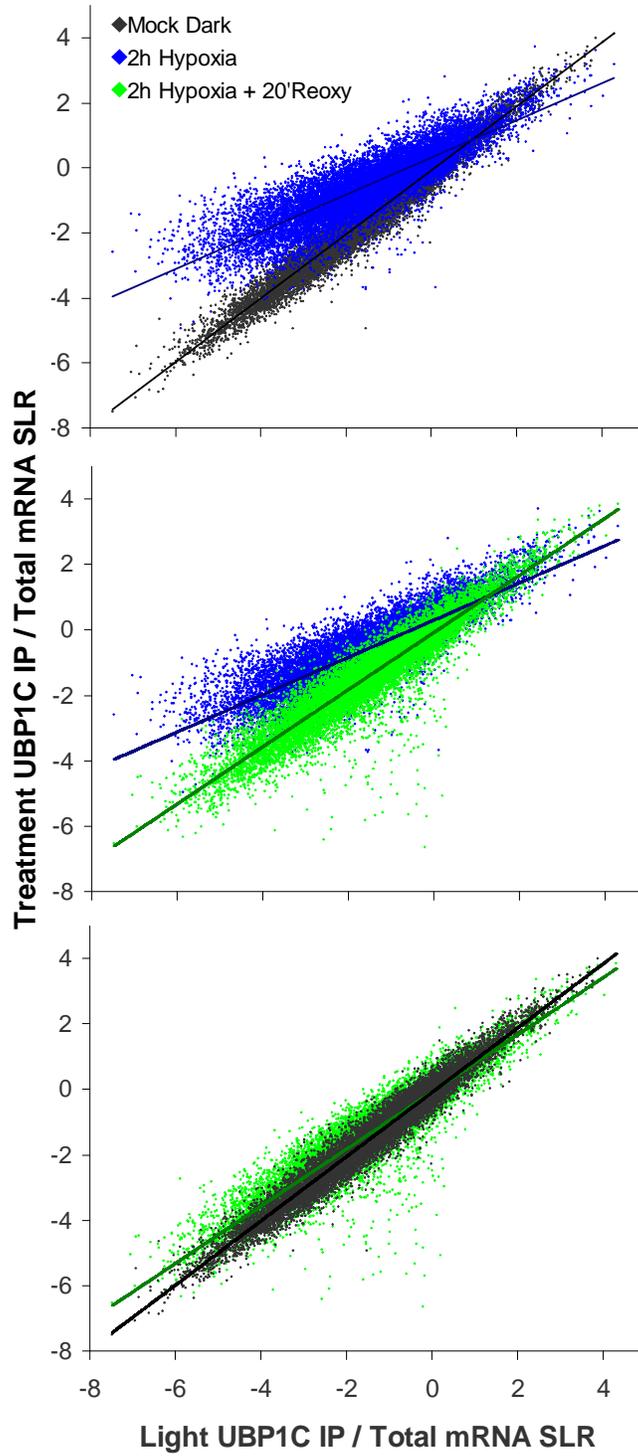


Figure 3.24. Heatmap comparisons of immunoprecipitated RNAs. 13,964 genes with at least one MAS5 “present” or “marginal” call in one of two replicate microarrays, in both total and UBP1C immunoprecipitated samples from the same treatment were included. Affy control probes and mitochondrial and chloroplast genes were excluded. Signal log ratios were used to compare immunoprecipitated RNA against total RNA abundance (IP/T), three types of treatment comparisons were made within each dataset plus a meta-comparison of genotypes between datasets. Treatment comparisons included changes in abundance of IP RNAs (IP), changes in T abundance (T), and change in IP/T. *FLAG-UBP1C* experiments were performed as described in Figure 3.19. Data for *FLAG-RPL18* was reprocessed from Branco-Price et al. [10] and was included for comparison of hypoxia controlled polysome mRNA association. It includes triplicate microarrays of 2 and 9 h mock treatment, 2 and 9 h hypoxia treatment (in *Ar(g)*-purged chambers) in low light, and 9 h hypoxia plus 1 h reoxygenation. All comparisons shown were included in clustering analysis. 20 clusters were generated by fuzzy k-means clustering. Cluster number and size (in genes) are shown. *Clusters were sorted by median SLR of the light control treatment IP/T comparison. **Genes within each cluster were ordered by the [IP/T] 2 h hypoxia / Light comparison.

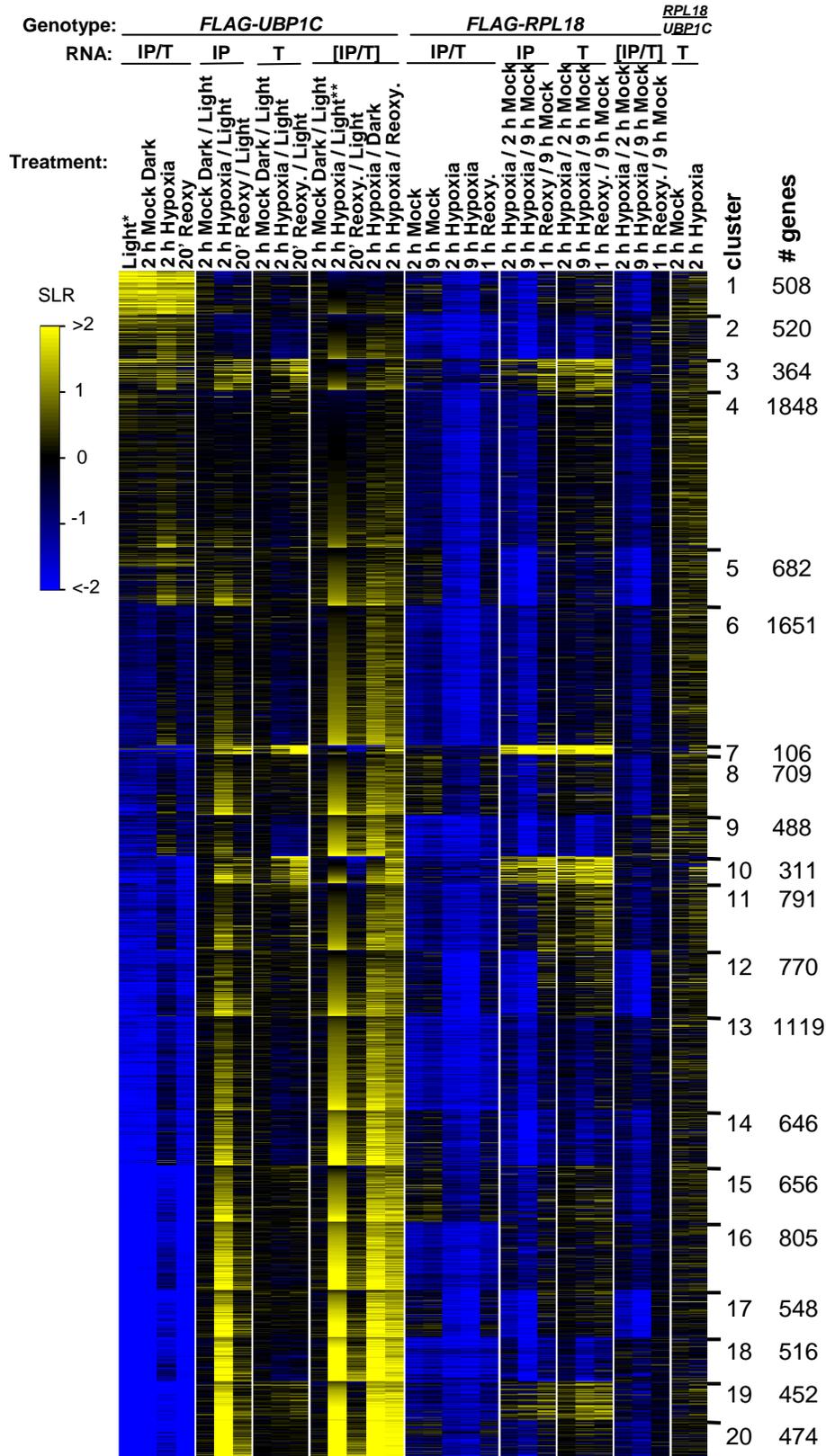


Figure 3.25. Heatmap comparisons of non-UBP1C-associated RNAs. 3,335 genes with at least one MAS5 “present” and all “absent” calls in the UBP1C immunoprecipitated samples were analyzed. Treatment comparisons and k-means clustering for 5 clusters (numbered NA1-NA5) were accomplished as for Figure 3.23.

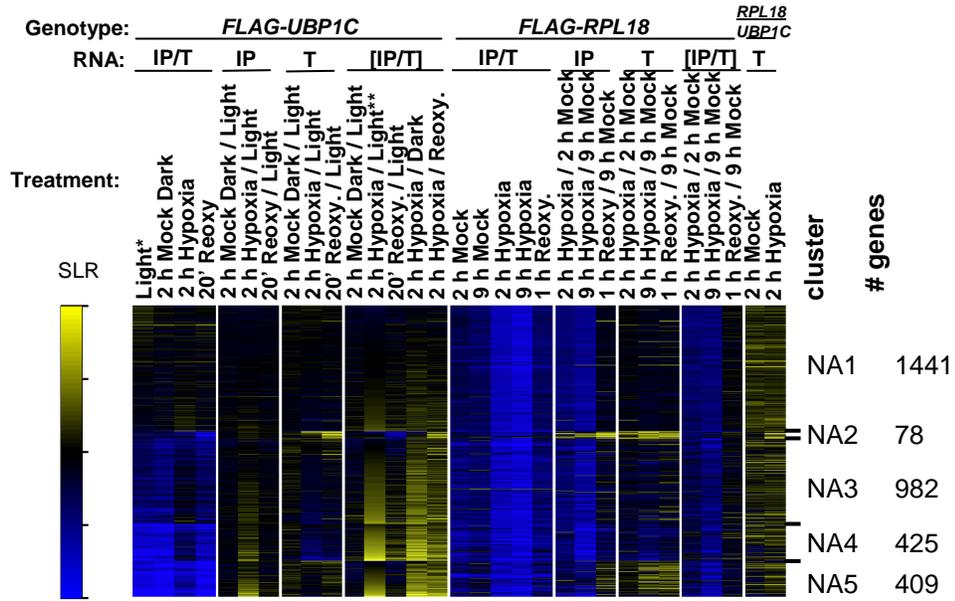


Figure 3.26. Percent nucleotide composition by cluster. Gene sequences including full length transcripts (cDNA), coding sequences (CDS), 5' untranslated regions (5'UTR), introns, and 3' untranslated regions (3'UTR) were obtained from www.arabidopsis.com TAIR10 genome annotation. Sequences of each cluster were analyzed for nucleotide composition. For comparison, cluster medians of Light and 2 h hypoxia (Hypoxia) UBP1C IP/T comparisons are shown for each cluster as a heatmap.

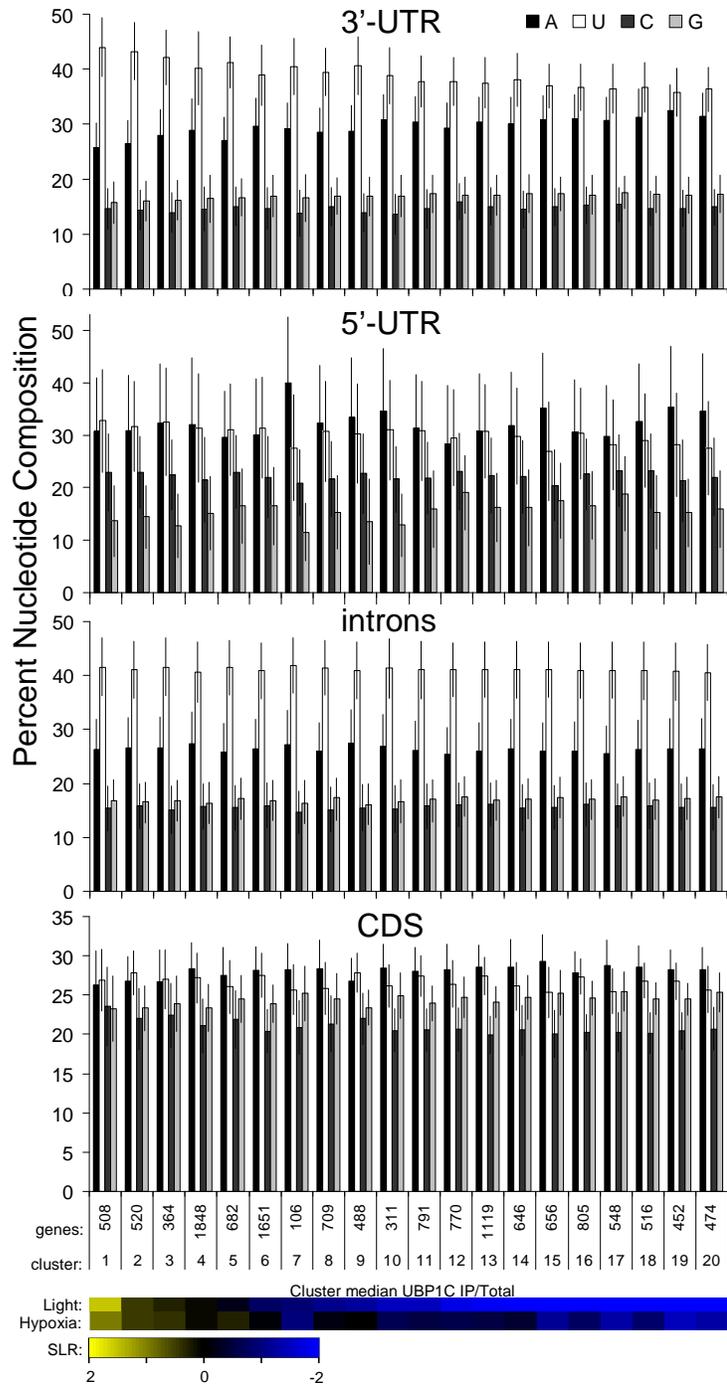
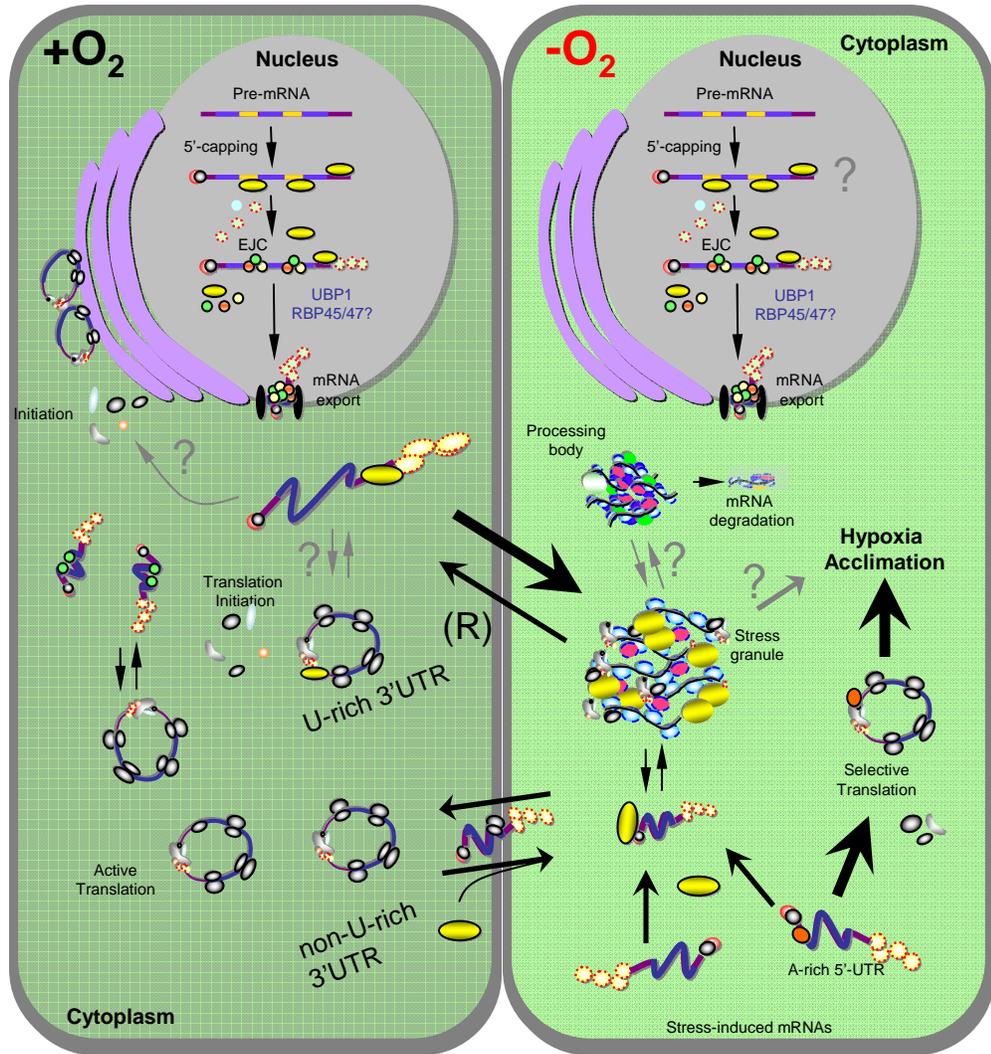


Figure 3.27. 3'UTR local nucleotide composition maxima. The 25 nucleotide sequence window with maximum A, U, C, or G composition was identified within each 3'untranslated region (3'UTR) from genes in each of 20 clusters identified. (A) The density of maximum nucleotide composition. (B) The density of the relative position(s) of each maximum nucleotide composition between each cluster was compared for each nucleotide. (C) Motif identified by "Multile Em for Motif Elicitation" algorithm using 3'UTR sequences of clusters 1 and 2 for motif identification and 3'UTR sequences of clusters 19 and 20 for generation of a background Markov model.

Figure 3.28. Working model for UBP1 function in *Arabidopsis thaliana*. UBP1 paralogs enhance splicing of U-rich introns, associate with U-rich 3'-untranslated regions of mature mRNAs in non-stress conditions, and aggregate in RNA containing stress granules (SGs) in response to hypoxia and other stresses *in vivo*. In hypoxia, escape from SG sequestration and selective translation contributes to production of proteins that facilitate acclimation to hypoxia. Hypothetical pathways need experimental support in plants (gray arrows).



KEY

- | | | | |
|--|---------------------|--|-----------------------|
| | UBP1 | | mRNA |
| | 40S ribosome | | ORF |
| | 60S ribosome | | Initiation Factors |
| | Hypothetical factor | | intron |
| | PAB | | R reoxygenation |
| | cap-binding complex | | exon junction complex |
| | other RNPs | | nuclear pore |

Table 3.1. Summary of strategies and data for complementation or recapitulation of *ubp1c-1* phenotypes. Constructs were made and stably transformed into the respective background for evaluation.

Genotype & Construct	Expression Result	Phenotype
<i>ubp1c-1</i> 35S:UBP1C cDNA-GFP	Construct expressed, Hypoxia induced granule formation	No complementation of sucrose dependence. GFP+ seedlings died 1 week earlier than <i>ubp1c-1</i> without sucrose. Plants are dwarfed and seed yield is low in highly expressing plants.
<i>ubp1c-1</i> <i>genomic</i> :UBP1C-FH-OCS	mRNA expressed and spliced	No complementation of sucrose dependence. Seed yield is lower in highly expressing plants.
<i>ubp1c-1</i> 35S:HF-UBP1C-OCS	UBP1C protein accumulates	No complementation of sucrose dependence; complementation not evident on 0%, 0.2%, or 3% sucrose. Seedlings died 1 week earlier than <i>ubp1c-1</i> without sucrose. Plants are small and seed yield is low in highly expressing plants.
<i>ubp1c-1</i> <i>pro</i> UBP1C::Flag- <i>genomic</i> UBP1C-UBP1C _{3'UTR}	Flag-UBP1C protein accumulates	No complementation of sucrose dependence. Some T1 plants exhibit dwarf phenotype and very low yield.
<i>ubp1c-1</i> <i>genomic</i> :UBP1C-UBP1C _{3'UTR}	Not determined	T1 plants not complemented in sucrose dependence. Some T1 plants exhibit very low yield
Col-0 35S: <i>amiR-ubp1c</i>	Reduced UBP1C transcript	Dark seed coat. Recapitulation of sucrose dependent growth. Phenotype severity different between lines.

Table 3.2. Gene ontology categories enriched in cluster 7.

Ont	GO ID	Padj	Term	Gene Matches	Node Size*
Cluster 7 - 106					
MF	GO:0047800	5.42E-03	cysteamine dioxygenase activity	2	5
MF	GO:0004737	8.10E-03	pyruvate decarboxylase activity	2	6
MF	GO:0016157	8.10E-03	sucrose synthase activity	2	6
MF	GO:0030976	1.93E-02	thiamine pyrophosphate binding	2	9
MF	GO:0008889	4.14E-02	glycerophosphodiesterase	2	13
BP	GO:0006950	7.99E-22	response to stress	49	2438
BP	GO:0009061	8.94E-21	anaerobic respiration	9	9
BP	GO:0009642	9.09E-08	response to light intensity	8	85
BP	GO:0010200	6.09E-04	response to chitin	6	127
BP	GO:0055062	2.01E-03	phosphate ion homeostasis	3	17
CC	GO:0044424	1.60E-02	intracellular part	45	8543

Abbreviations: Ont-ontology, GO-gene ontology, Padj-adjusted P-value, MF-molecular function, BP-biological process, CC-cellular component; *Nodes with fewer than 5 were removed.

Table 3.3. Gene ontology categories enriched in clusters 5, 8, 12, and 17.

Ont	GO ID	Padj	Term	Gene Matches	Node Size*
Cluster 5 - 682 genes					
MF	GO:0003735	1.25E-57	structural constituent of ribosome	91	412
MF	GO:0003700	1.82E-04	seq.-specific DNA binding TF activity	73	1680
MF	GO:0003743	3.85E-02	translation initiation factor activity	9	92
BP	GO:0006412	4.18E-25	translation	103	1320
BP	GO:0042254	3.90E-17	ribosome biogenesis	37	247
BP	GO:0006355	1.63E-04	regulation of transc, DNA-dependent	77	1870
BP	GO:0007264	9.11E-03	small GTPase mediated signal transd.	11	112
CC	GO:0022626	2.81E-61	cytosolic ribosome	84	309
CC	GO:0005794	1.46E-12	Golgi apparatus	60	793
Cluster 8 - 709 genes					
MF	GO:0005515	4.35E-06	protein binding	108	2493
MF	GO:0015035	9.04E-05	protein disulfide oxidoreductase activity	13	95
MF	GO:0003735	1.81E-02	structural constituent of ribosome	24	412
MF	GO:0019787	2.09E-02	small conjugating protein ligase activity	19	292
MF	GO:0016671	2.25E-02	oxidoreductase, sulfur/disulfide	6	35
BP	GO:0006334	2.60E-06	nucleosome assembly	13	70
BP	GO:0045454	3.11E-06	cell redox homeostasis	18	143
BP	GO:0006662	2.26E-03	glycerol ether metabolic process	9	59
BP	GO:0009733	2.56E-02	response to auxin stimulus	20	318
BP	GO:0071901	3.48E-02	neg. regulation of prot. S/T kinase	4	13
CC	GO:0043229	4.92E-19	intracellular organelle	279	7076
Cluster 12 - 770 genes					
MF	GO:0003735	3.28E-14	structural constituent of ribosome	47	412
MF	GO:0005515	9.25E-06	protein binding	114	2493
MF	GO:0008270	2.96E-02	zinc ion binding	67	1551
BP	GO:0006396	2.08E-11	RNA processing	44	440
BP	GO:0080135	7.40E-03	regulation of cellular response to stress	6	25
BP	GO:0006412	3.97E-02	translation	58	1320
CC	GO:0005622	2.47E-30	intracellular	385	9164
Cluster 17 - 548 genes					
MF	GO:0003735	7.52E-18	structural constituent of ribosome	44	412
MF	GO:0016853	8.00E-03	isomerase activity	15	239
MF	GO:0019843	8.12E-03	rRNA binding	5	24
MF	GO:0050897	2.76E-02	cobalt ion binding	6	47
BP	GO:0006412	2.44E-05	translation	54	1320
BP	GO:0015979	3.65E-05	photosynthesis	17	196
BP	GO:0042254	4.84E-05	ribosome biogenesis	19	247
BP	GO:0006457	8.86E-05	protein folding	19	257
BP	GO:0009793	1.48E-02	embryo dev. ending in seed dormancy	20	401
BP	GO:0006367	3.16E-02	transc. initiation RNA pol. II promoter	4	17
BP	GO:0015031	4.67E-02	protein transport	20	438
CC	GO:0005737	3.79E-53	cytoplasm	294	6466

Abbreviations: Ont-ontology, GO-gene ontology, Padj-adjusted P-value, MF-molecular function, BP-biological process, CC-cellular component, TF-transcription factor, ; *Nodes with fewer than 5 were removed.

Table 3.4. Gene ontology categories enriched in clusters 1 and 2.

Ont	GO ID	Padj	Term	Gene Matches	Node Size*
Cluster 1 - 508 genes					
MF	GO:0005199	4.42E-06	structural constituent of cell wall	8	33
MF	GO:0003700	5.38E-05	seq.-specific DNA binding TF activity	57	1680
MF	GO:0003677	3.86E-03	DNA binding	58	1988
MF	GO:0005515	1.15E-02	protein binding	67	2493
BP	GO:0009725	1.05E-09	response to hormone stimulus	52	1081
BP	GO:0051252	1.31E-04	regulation of RNA metabolic process	60	1882
BP	GO:0007267	1.31E-02	cell-cell signaling	6	47
CC	GO:0031225	3.75E-18	anchored to membrane	32	248
CC	GO:0005634	7.83E-03	nucleus	64	2541
CC	GO:0048046	2.09E-02	apoplast	15	346
Cluster 2 - 520 genes					
MF	GO:0016759	3.51E-09	cellulose synthase activity	11	38
MF	GO:0005215	2.03E-06	transporter activity	52	1294
MF	GO:0005524	1.34E-05	ATP binding	69	2068
MF	GO:0004674	1.39E-03	protein serine/threonine kinase activity	36	952
MF	GO:0005515	2.47E-02	protein binding	67	2493
MF	GO:0005199	2.85E-02	structural constituent of cell wall	5	33
MF	GO:0005086	4.69E-02	ARF-GEF activity	3	9
BP	GO:0032502	9.90E-14	developmental process	96	2314
BP	GO:0006888	1.14E-04	ER to Golgi vesicle-mediated transport	7	30
BP	GO:0030244	1.98E-04	cellulose biosynthetic process	8	46
BP	GO:0009846	8.96E-04	pollen germination	7	40
BP	GO:0006886	1.71E-03	intracellular protein transport	18	311
BP	GO:0006468	4.40E-03	protein phosphorylation	40	1146
BP	GO:0009225	6.77E-03	nucleotide-sugar metabolic process	6	37
BP	GO:0032011	2.99E-02	ARF protein signal transduction	3	7
CC	GO:0016020	1.19E-26	membrane	190	4881
CC	GO:0009505	4.30E-05	plant-type cell wall	19	268
CC	GO:0005829	1.37E-04	cytosol	67	2014
CC	GO:0005576	4.46E-02	extracellular region	20	483

Abbreviations: Ont-ontology, GO-gene ontology, Padj-adjusted P-value, MF-molecular function, BP-biological process, CC-cellular component, TF-transcription factor, ARF-GEF-guanine nucleotide exchange factor for the ADP-ribosylation factor GTPases, ER-endoplasmic reticulum; *Nodes with fewer than 5 were removed.

Table 3.5. Gene ontology categories enriched in cluster 9.

Ont	GO ID	Padj	Term	Gene Matches	Node Size*
Cluster 9 - 488					
MF	GO:0003824	1.06E-23	catalytic activity	255	8733
MF	GO:0020037	5.48E-08	heme binding	27	377
MF	GO:0015562	8.08E-04	efflux transmembrane transporter	6	28
MF	GO:0005507	2.68E-03	copper ion binding	15	244
MF	GO:0015140	5.96E-03	malate transmembrane transporter	3	5
BP	GO:0007169	1.01E-15	transmembrane receptor prot. T kinase	24	130
BP	GO:0005975	3.67E-09	carbohydrate metabolic process	51	1025
BP	GO:0006633	1.00E-06	fatty acid biosynthetic process	17	166
BP	GO:0009624	3.09E-06	response to nematode	11	66
BP	GO:0071702	1.66E-05	organic substance transport	21	302
BP	GO:0009827	2.13E-05	plant-type cell wall modification	9	48
BP	GO:0046274	7.05E-05	lignin catabolic process	6	18
BP	GO:0006829	5.40E-03	zinc ion transport	5	22
BP	GO:0048765	6.72E-03	root hair cell differentiation	8	73
BP	GO:0006814	8.47E-03	sodium ion transport	5	24
BP	GO:0016126	1.04E-02	sterol biosynthetic process	5	25
BP	GO:0009807	2.33E-02	lignan biosynthetic process	4	16
BP	GO:0009074	4.56E-02	aromatic AA family catabolic process	3	8
BP	GO:0010345	4.56E-02	suberin biosynthetic process	3	8
CC	GO:0016020	4.24E-31	membrane	192	4881
CC	GO:0005576	6.73E-07	extracellular region	29	483
CC	GO:0009505	7.70E-05	plant-type cell wall	18	268
CC	GO:0042807	4.36E-03	central vacuole	3	6

Abbreviations: Ont-ontology, GO-gene ontology, Padj-adjusted P-value, MF-molecular function, BP-biological process, CC-cellular component, AA-amino acid; *Nodes with fewer than 5 were removed.

Chapter 4

Dynamics of polysome levels during hypoxia and reoxygenation

4.1 Abstract

Hypoxia results in transcriptional and post-transcriptional regulation of gene expression in seedlings of *Arabidopsis thaliana*. Global dynamics in protein synthesis can be monitored by quantitative evaluation of polyribosomes (polysomes). Here, evaluation of polysome levels in oxygen deprived and reoxygenated seedlings determined that the 41% reduction in large polysomes observed in response to 2 hours of stress recovered to 89% of non-stressed levels within 15 minutes of reoxygenation. A number of signaling pathways that function in low oxygen (hypoxia) signal transduction have been identified, including a rapid transient activation of MPK6. Previous comparative evaluation of *mpk6-3* mutants and MPK6 overexpression lines indicated that a primary distinction between genotypes was in the degree of downregulation of transcripts. As MPK6 phosphorylates Decapping Protein 1 (DCP1), which leads to targeted mRNAs degradation in response to dehydration, we tested the hypothesis that MPK6 may regulate changes in polysome levels in response to hypoxia and reoxygenation. By comparison to the polysome loss during hypoxia and recovery upon reoxygenation in wildtype seedlings, loss-of-function *mpk6-3* mutants show reduced loss and slower recovery. This is a novel finding that, along with recently developed genetic tools, opens the ability to address a number of outstanding questions.

4.2 Introduction

Insufficient oxygen levels quickly initiate upstream signal transduction pathways. As previously discussed (Chapter 1), a number of these signal transducers have been identified. Calcium has been identified as a second messenger in plant hypoxia signaling (Subbaiah et al., 1994). It has been proposed that cytosolic Ca^{+2} release is upstream and regulates Rop

signaling controlled by RopGAP4 in *Arabidopsis thaliana* (Baxter-Burrell et al., 2002). An oxidative burst from mitochondria and/or membrane-bound NADPH oxidase is believed to be important in indirect signaling of low oxygen status as well (reviewed by Bailey-Serres and Chang, 2005). In support of this, *rbohD* mutants have reduced induction of *ADH1* and *HSFA2* and have reduced survival during anoxia (Pucciariello et al., 2012). Additionally, low oxygen promotes a strong and pronounced activation of three MITOGEN ACTIVATED PROTEIN KINASEs (MPKs), MPK3, MPK4 and MPK6 (Chang et al., 2012).

Activation of MPK3/4/6 by severe oxygen deprivation of seedlings occurred within 30 minutes of the stress, peaked within 2 h and transiently reoccurred at even higher levels upon reoxygenation (Chang et al., 2012). MPK3/4/6 activation similarly occurred in a dose dependent manner in seedlings treated with mitochondrial electron transport inhibitors potassium cyanide or antimycin A. Of these kinases, the activation of MPK6 was the most pronounced based on in-gel kinase assays. The functional significance of the activation of these kinases was further investigated using loss-of-function mutants for *MPK3* and *MPK6* and overexpression mutants for MPK3/4/6. These studies found that an *mpk6-3* mutant is compensated by the partially redundant gene *MPK3* in a number of contexts. For example, *mpk6-3* mutants display a substantial increase in MPK3 activity in response to oxygen deprivation. In lines overexpressing either MPK3 or MPK6, subtly enhanced survival of oxygen deprivation was observed (Chang et al., 2012). As MPK6 is known to respond to pathogen infection (Li et al., 2012), wound response (Wang et al., 2008), salt and cold (Teige et al., 2004), guard and pavement cell differentiation (Wang et al., 2007), ovule development (Wang et al., 2008), the impact of *MPK6* loss-of-function and overexpression on hypoxia-responsive gene regulation was carefully evaluated in the root and shoot of seedlings (Chang et al., 2012). Surprisingly, only subtle differences were observed in transcriptome analyses that compared mutant, wildtype, and overexpression lines treated with hypoxia.

Most hypoxia-responsive genes did display altered levels in these mutants under either control or stress conditions. However, significantly more genes were down-regulated in the mutant than in WT in response to hypoxia, and conversely fewer genes were down-regulated in the MPK6 overexpression line as compared to hypoxia. A group of genes (~100) was reduced 2-3 fold in both mutant and over expression lines in stress and non-stress in both roots and shoot. This led Chang et al. (2012) to propose that MPK6 may be involved in regulating the stability of mRNAs during oxygen deprivation, plausibly through involvement in the mechanism of sequestration of mRNAs into non-polysomal complexes. We hypothesized that the rapid activation of MPK6 in response to oxygen deprivation and re-activation of MPK6 upon re-oxygenation may be consistent with dynamics in global levels of protein synthesis, as measured by sucrose gradient fractionation and quantitation of polysomes. Moreover, the role of MPK6 activation might be linked to the regulation of transcript stability, rather than transcript synthesis.

A major gateway of RNA degradation is controlled by the decapping complex. There is evidence that the decapping complex responds to hypoxia and heat by aggregation into large cytoplasmic complexes, which have been referred to as Processing Bodies (PBs) (Weber et al., 2008; Chapter 1). In yeast, the active decapping enzyme, DCP2, is regulated by phosphorylation at serine 137 by Ste20, which leads to PB aggregation and stabilization of selected RNAs (Yoon et al., 2010). This serine residue is conserved in plants, suggesting it might also be a target site of regulatory phosphorylation. The Ste20 group kinases are reported activators of MAP kinase cascades (Dan et al., 2001), connecting these pathways. Furthermore, the dehydration of Arabidopsis seedlings was recently shown to result in phosphorylation of DCP1 at serine 237. MPK6 is responsible for this phosphorylation event, which promotes decapping and destabilizes some RNAs (Xu and Chua, 2012). PB formation triggered by dehydration has not been demonstrated.

Based on the knowledge that hypoxia results in reversible down-regulation of translation and sequestration of a subset of mRNAs from translational complexes as a means of energy conservation (Branco-Price et al., 2008), we hypothesized that MPK6 might play a role in the regulation of polysome levels by oxygen availability and therefore oxygen-regulated dynamics in global levels of protein synthesis. As an initial step to address this hypothesis, we determined the dynamics in polysome accumulation in response to hypoxia and reoxygenation in wildtype and *mpk6-3* seedlings. We observed subtle, reproducible differences between wildtype and *mpk6-3* polysome dynamics consistent with our hypothesis.

4.3 Results

4.3.1 Polysome dynamics during hypoxia stress are altered in *mpk6-3* mutant

To assess if MPK6 plays a role in translational dynamics in response to hypoxia or reoxygenation, polysomal recovery dynamics were evaluated in seedlings of *Arabidopsis thaliana* ecotype Columbia (Col-0) and *mpk6-3*. Seedlings that were grown for seven days on MS medium containing 1% (w/v) sucrose were used. First, an analysis was performed with Col-0 to appreciate the timing of polysome recovery. Branco-Price et al. (2008) demonstrated a reduction in polysomal RNA to 46% that of non-stressed plants after 9 h oxygen deprivation which recovered to 88% following an additional 1 h of aeration. Here, 2 hours of hypoxia stress was tested, allowing recovery by 5 or 15 min aeration (Figure 4.1). Quantitation of the levels of polysomes (≥ 2 ribosomes) confirmed reduction to 59% of mock treated after 2 h of hypoxia. Polysome levels recovered rapidly upon reoxygenation, with an increase to 82 and 89% within 5 and 15 min, respectively (Figure 4.2B). A shift in size from smaller to larger polysomes can be seen in this rapid recovery (Figure 4.1).

To determine if MPK6 regulates polysome dynamics, *mpk6-3* mutant seedlings were treated with the same regimen as above with noted differences in the level of reduction and

the speed of recovery (Figure 4.2B). In order to make a direct comparison, wildtype and *mpk6-3* seedlings were treated with hypoxia for 2 h and reoxygenated for 5, 10, or 15 min. This was again repeated in an additional experiment with a single recovery time point after 10 min aeration following 2 h less severe oxygen deprivation. In each of these experiments, the cryopreserved samples were used for quantitation of polysomes. We observed consistently less-responsive polysome regulation in *mpk6-3* seedlings, polysomes decreased less in *mpk6-3* at different severity of oxygen deprivation and recovered more slowly than wildtype plants (Figure 4.2B).

4.4 Discussion

Here we show that polysome levels recover dramatically within minutes of release from hypoxia (Figure 4.1). In an *mpk6-1* loss-of-function mutant, polysome levels were less dramatically reduced by hypoxia and less efficiently recovered by reoxygenation (Figure 2.2). These dynamics in translational activity temporally coincide with the transient activation of MPK3/4/6 in response to oxygen deprivation and reoxygenation (Chang et al. (2012)). Chang et al. (2012) also demonstrated that MPK3 partially compensates for the loss of MPK6 activity. Based on this, we hypothesize that, in a genetic background with MPK6 and MPK3 activity, a further slowing of the polysome hypoxia dynamic may be observed. This could be evaluated using the rescued double *mpk3/mpk6* mutant (Li et al., 2012). Few differences were observed in increases in steady state transcript abundance following a comparable 2 h hypoxia treatment in *mpk6-3* and Col-0, however, many more mRNAs significantly declined in abundance in Col-0 or a *MPK6* overexpression line (Chang et al., 2012). This suggests that MPK6 functions in post-transcriptional regulation.

MPK6 may be involved in the regulation of the process of mRNA sequestration, and/or mRNA degradation. This is supported by the recent finding that phosphorylation of

DCP1 occurs in response to seedling dehydration and promotes mRNA decapping and degradation (Xu and Chua (2012)). During dehydration, MPK6 activation enhanced degradation of RNAs. Here we find a correlation between less translational repression in *mpk6-3* mutants and more destabilization of mRNAs. Possibly, in the absence of MPK6 activation the mechanism of selection of mRNAs for degradation or the sequestration of mRNAs is somehow altered. This might be further studied by careful evaluation of the selective translation of individual mRNAs in the *mpk6-3* mutant and MPK6 overexpression line, as done for Col-0 by Branco-Price et al (2008).

It is hypothesized that translation is repressed in response to reduced levels of oxygen. The rapid activation of MPK3/MPK6 could facilitate this regulation either by targeting translation factors or ribonucleotide binding proteins involved in RNA sequestration.

4.5 Materials and Methods

4.5.1 Genetic material and growth conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) (Arabidopsis Biological Resource Center), and *35S:FLAG-RPL18* (*FLAG-RPL18*) (Zanetti et al., 2005) were used for wildtype comparison. *mpk6-3* (SALK_12507, ABRC) is a T-DNA in exon 4 of locus At2g43790.

Seeds were sterilized, stratified, and grown on MS agar with 1% sucrose as described (Chapter 3). Oxygen deprivation and reoxygenation was also imposed in Argon purged chambers as described (Chapter 3).

4.5.2 Polysome profiling

Ribosome subunits and complexes were isolated from seedling extracts using polysome extraction buffer and detergents by differential centrifugation. Polysomes were concentrated by sedimentation through a sucrose cushion and fractionated over sucrose density gradients to generate absorbance profiles exactly as described by Mustroph et al.

2009 (Mustroph et al., 2009). The proportion of the total absorbance (A_{254}) in polysomes (≥ 2 ribosome per mRNA) was determined as described by Williams et al. (2003).

4.6 References

- Bailey-Serres, J. and Chang, R.** (2005). Sensing and signalling in response to oxygen deprivation in plants and other organisms. *Ann Bot* **96**: 507–518.
- Baxter-Burrell, A., Yang, Z., Springer, P.S., and Bailey-Serres, J.** (2002). RopGAP4-dependent rop GTPase rheostat control of Arabidopsis oxygen deprivation tolerance. *Science* **296**: 2026–2028.
- Branco-Price, C., Kaiser, K.A., Jang, C.J.H., Larive, C.K., and Bailey-Serres, J.** (2008). Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. *The Plant Journal* **56**: 743–755.
- Chang, R., Jang, C., Branco-Price, C., Nghiem, P., and Bailey-Serres, J.** (2012). Transient MPK6 activation in response to oxygen deprivation and reoxygenation is mediated by mitochondria and aids seedling survival in Arabidopsis. *Plant Molecular Biology* **78**: 109–122.
- Dan, I., Watanabe, N.M., and Kusumi, A.** (2001). The Ste20 group kinases as regulators of MAP kinase cascades. *Trends in Cell Biology* **11**: 220–230.
- Li, G., Meng, X., Wang, R., Mao, G., Han, L., Liu, Y., and Zhang, S.** (2012). Dual-level regulation of ACC Synthase activity by MPK3/MPK6 cascade and its downstream WRKY transcription factor during ethylene induction in Arabidopsis. *PLoS Genet* **8**.
- Mustroph, A., Juntawong, P., and Bailey-Serres, J.** (2009). Isolation of plant polysomal mRNA by differential centrifugation and ribosome immunopurification methods. *Meth. Mol. Biol.* **553**: 109–126.
- Pucciariello, C., Parlanti, S., Banti, V., Novi, G., and Perata, P.** (2012). Reactive Oxygen Species-Driven Transcription in Arabidopsis under Oxygen Deprivation. *Plant Physiol.* **159**: 184–196.
- Subbaiah, C.C., Bush, D.S., and Sachs, M.M.** (1994). Elevation of cytosolic calcium precedes anoxic gene expression in maize suspension-cultured cells. *Plant Cell* **6**: 1747–1762.
- Teige, M., Scheikl, E., Eulgem, T., Dóczi, R., Ichimura, K., Shinozaki, K., Dangl, J.L., and Hirt, H.** (2004). The MKK2 Pathway Mediates Cold and Salt Stress Signaling in Arabidopsis. *Molecular Cell* **15**: 141–152.

- Wang, H., Liu, Y., Bruffett, K., Lee, J., Hause, G., Walker, J.C., and Zhang, S.** (2008). Haplo-insufficiency of MPK3 in MPK6 mutant background uncovers a novel function of these two MAPKs in Arabidopsis ovule development. *Plant Cell* **20**: 602–613.
- Wang, H., Ngwenyama, N., Liu, Y., Walker, J.C., and Zhang, S.** (2007). Stomatal development and patterning are regulated by environmentally responsive mitogen-activated protein kinases in Arabidopsis. *Plant Cell* **19**: 63–73.
- Weber, C., Nover, L., and Fauth, M.** (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.* **56**: 517–530.
- Williams, A.J., Werner-Fraczek, J., Chang, I.-F., and Bailey-Serres, J.** (2003). Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. *Plant Physiol.* **132**: 2086–2097.
- Xu, J. and Chua, N.-H.** (2012). Dehydration stress activates Arabidopsis MPK6 to signal DCP1 phosphorylation. *EMBO J.* **31**: 1975–1984.
- Yoon, J.-H., Choi, E.-J., and Parker, R.** (2010). Dcp2 phosphorylation by Ste20 modulates stress granule assembly and mRNA decay in *Saccharomyces cerevisiae*. *J. Cell Biol.* **189**: 813–827.
- Zanetti, M.E., Chang, I.-F., Gong, F., Galbraith, D.W., and Bailey-Serres, J.** (2005). Immunopurification of polyribosomal complexes of Arabidopsis for global analysis of gene expression. *Plant Physiol.* **138**: 624–635.

Figure 4.1. Polysomal dynamic of reoxygenation after 2 h hypoxia treatment. Seedlings were grown seven days on MS agar with 1% (w/v) sucrose and transferred to clear Plexiglas chambers at 22°C with 8-10 $\mu\text{E m}^{-2} \text{s}^{-1}$ room light and incubated 2 h in air (mock) or were sealed and purged with Argon gas 2 h (HS). Following 2 h hypoxia treatment, reoxygenation (R) was facilitated by vigorous fanning of open chambers by hand and harvesting after 5 or 15 min. Polysomes of frozen tissue were extracted and sedimented in a 15-10% sucrose gradient and an absorbance (254 nm) profile along the axis of the gradient was measured.

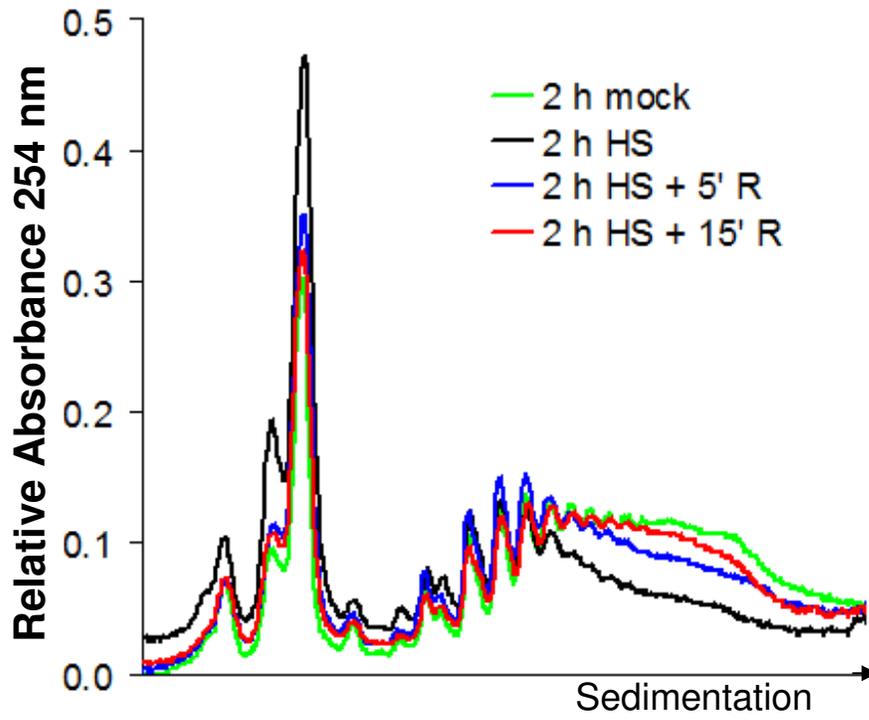
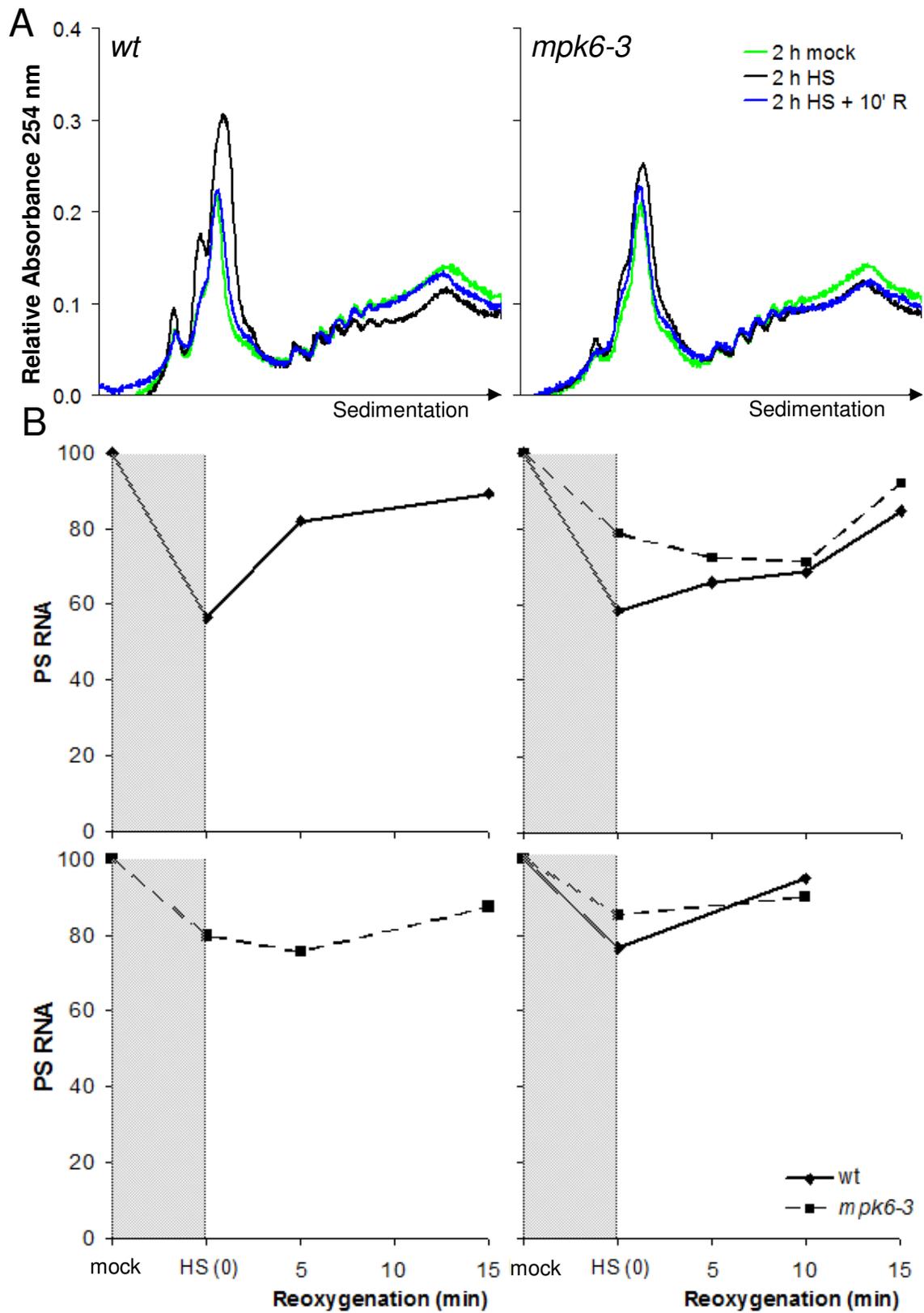


Figure 4.2. *mpk6-3* loss-of-function mutants have altered polysomal dynamics after 2 h hypoxia treatment. *FLAG-RPL18* (wt) and *mpk6-3* seedlings were grown and treated as in Figure 4.1. (A) Absorbance (254 nm) profiles along the sucrose gradient axis. (B) The fraction of RNA in polysomes (PS RNA) of mock treated plants was compared to that of hypoxia and reoxygenation treatments from four independent experiments: *FLAG-RPL18* alone (upper left; Figure 4.1 experiment), *mpk6-3* alone (lower left), Columbia-0 and *mpk6-3* (upper right), *FLAG-RPL18* and *mpk6-3* (lower right, polysomes in A).



Chapter 5

General Conclusions

Over the past 30 years, progress has continued on the question of plant cell selective translation during oxygen deprivation. This dissertation takes another step forward in understanding the cellular response to hypoxia in the context of post-transcriptional gene regulation. After identifying a number of putative proteins that could participate in RNA sequestration during hypoxia, the role of the RNA binding protein UBP1C was investigated. It was demonstrated that dynamic cytosolic relocation of UBP1C was finely regulated by oxygen availability. Nuclear encoded mRNAs that are translationally repressed by hypoxia, independent of their stability, reversibly increase association with UBP1C. It was also demonstrated that MPK6, an upstream signaling protein that is activated by hypoxia and reoxygenation, contributes to regulation of polysomes levels during oxygen deprivation and recovery by reoxygenation. The results allow further refinement of the model of hypoxic mRNA management and post-transcriptional gene regulation.

In non-plant eukaryotes, an elaborate cellular system that manages cytoplasmic mRNA during numerous stresses has been described (Ivanov and Nadezhdina, 2006; Parker and Sheth, 2007; Kedersha and Anderson, 2007; Anderson and Kedersha, 2008). This system is characterized by large cytoplasmic domains comparable in size to a mitochondrion (up to 4-5 μm diameter) and millions of times more voluminous than a ribosome. These domains appear as granules when fluorescent marker proteins are micrographed by confocal microscopy. Dense inclusions appear in transmission electron microscopy and can have ribosome-like structures adjacent to them (Gilks et al., 2004). The internal organization of these stress granules (SGs) and processing bodies (PBs), however, remains a “black box”. Packaging mRNAs into these granules physically decreases their accessibility to ribosomes, driving the hypothesis that mRNA sequestration is a means of controlling protein synthesis.

This suggests that mRNA management is facilitated by a ubiquitous, complex, and organized system of regulated RNA binding proteins that, complementary to regulation of polysome activity, might regulate mRNA accessibility to translational machinery and recruitment to degradative processes.

Data has accumulated supporting the conclusion that the cytosolic mRNA network of SGs and PBs is conserved to some degree in nearly all eukaryotes, RNA granules have been described in a long list of animal context-specific processes (e.g. differential stress composition (Arimoto et al., 2008), in developing neurons (Vessey et al., 2006), for RNA storage in *Xenopus laevis* ovules (Tanaka et al., 2006; Weston and Sommerville, 2006), and in differentiation priming of immune cell (Mazumder et al., 2010). Little can be extrapolated from these examples to mechanisms of plant growth, development, and survival. In plants protein:RNA granules in addition to SGs and PBs have been observed (e.g. siRNA bodies (Jouannet et al., 2012), nuclear dicing bodies (Fang and Spector, 2007), and Tandem Zinc Finger-containing bodies (Pomeranz et al., 2010). A few studies provide evidence that RNA granules form in response to stress in plants (Weber et al., 2008; Pomeranz et al., 2010). The results presented in this dissertation extend this evidence by demonstrating dynamic mRNA association with UBP1C, which forms SG-like granules, during oxygen deprivation.

The genetic resources and tools that have been collected and generated in this study make possible answers to additional questions about SGs and PBs. After an initial survey, we focused our study on UBP1C, developing an efficient method to immunopurify UBP1C:mRNA complexes. The profiling of UBP1C mRNAs, in addition to the contribution to hypoxia gene regulation, provides evidence of a UBP1C mRNA regulon. During uninhibited growth, UBP1C was not localized in granules and was highly associated with a specific subset of cellular mRNAs - the regulon. The level of UBP1C association correlated with 3'-UTR U-richness, consistent with its characterization as an oligouridylate binding protein

(Lambermon et al., 2000). Within this regulon, a number of interesting gene ontology categories were enriched, suggesting that the regulon may play a role in gene expression of transcripts targeted to the cell wall and/or plasma membrane.

The functional significance of UBP1C RNA binding during normal growth will be interesting to determine. Following up this study, there are two pressing questions to be addressed. (1) Demonstrated regulation of UBP1C:mRNA binding leads to the question of whether mRNAs associated with UBP1C under normal growth conditions have altered rates of translation in the *ubp1c-1* mutant plants. This can be addressed experimentally by comparing polysome association of UBP1C-associated mRNAs in wildtype and *ubp1c-1* mutants. (2) Although few proteins were visualized after polyacrylamide electrophoresis separation of the immunoprecipitated UBP1C:mRNA complex, we hypothesize that other RNA binding proteins might be present. To identify these, a scaling-up of the immunopurification method could yield sufficient protein for analysis by liquid chromatography mass spectrometry (LC-MS). Identification of associated proteins would further our understanding of the network of RNA binding proteins participating in RNA sequestration.

Few RNA regulons have been described in plants and there are over 1,100 RNA binding proteins predicted from the *Arabidopsis thaliana* genome (Bailey-Serres et al., 2009). The concept that these proteins may bind a non- or partially-overlapping subsets of cellular mRNAs and influence timing and/or location of translation or decay might represent a rich layer of fine gene regulation. Of course, a challenge will be to tease the target specificity from cell-specific expression of these proteins.

Besides SG-like aggregation during hypoxia, results of our survey are consistent with aggregation of PBs, which among other proteins are composed of *RH6*, *RH8*, and *RH12*. SG and PB can have an intimate relationship (Kedersha et al., 2005). We identified mutants of *RH6*, *RH8* and *RH12*, and confirmed that these proteins aggregate into cytoplasmic granules

which respond to oxygen deprivation. Generation of double and triple mutants of these genes will answer the question of whether these proteins are redundant or essential for decapping in developing seedlings as are DCP1, DCP2, and VCS (Xu et al., 2006; Goeres et al., 2007; Iwasaki et al., 2007). One question is what is the effect of dampening *RH6/8/12* function in turnover, translational repression and/or mRNA sequestration under non-stress or oxygen deprivation conditions. In the triple mutant *rh6 rh8 rh12*, SG dynamics might be altered. Furthermore, more can be done to understand movement between SGs and PBs. mRNAs that were unstable during hypoxia and associated with UBP1C and from there may have passed to PB for degradation. Without a functional decapping complex this transfer of RNA may be affected and alter SG size and dynamics. Lack of XRN4 in PB, caused increase in PB size, a similar result might be found in SG if some mRNAs are transferred to non-functional PBs.

PB/SG dynamics might also be altered in mutants of regulatory proteins. Does MPK6 directly regulate polysomes and/or SG/PB proteins during hypoxia, as it does DCP1 during dehydration (Xu and Chua, 2012)? These questions could lead to more that continue to uncover RNA regulatory mechanisms. This dissertation has made strides towards understanding selective translation in oxygen deprived plants, and by doing so opened the doors to further inquiry of the regulation of gene expression by RNA sequestration.

5.1 References

- Anderson, P. and Kedersha, N.** (2008). Stress granules: the Tao of RNA triage. *Trends Biochem. Sci.* **33**: 141–150.
- Arimoto, K., Fukuda, H., Imajoh-Ohmi, S., Saito, H., and Takekawa, M.** (2008). Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nature Cell Biology* **10**: 1324–1332.
- Bailey-Serres, J., Sorenson, R., and Juntawong, P.** (2009). Getting the message across: cytoplasmic ribonucleoprotein complexes. *Trends in Plant Science* **14**: 443–453.

- Fang, Y. and Spector, D.L.** (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. *Curr. Biol.* **17**: 818–823.
- Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P.** (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell* **15**: 5383–5398.
- Goeres, D.C., Van Norman, J.M., Zhang, W., Fauver, N.A., Spencer, M.L., and Sieburth, L.E.** (2007). Components of the Arabidopsis mRNA decapping complex are required for early seedling development. *Plant Cell* **19**: 1549–1564.
- Ivanov and Nadezhkina** (2006). Stress granules: RNP-containing cytoplasmic bodies arising in stress: Structure and mechanism of organization. *Molecular Biology* **40**: 844–850.
- Iwasaki, S., Takeda, A., Motose, H., and Watanabe, Y.** (2007). Characterization of Arabidopsis decapping proteins AtDCP1 and AtDCP2, which are essential for post-embryonic development. *FEBS Lett.* **581**: 2455–2459.
- Jouannet, V., Moreno, A.B., Elmayer, T., Vaucheret, H., Crespi, M.D., and Maizel, A.** (2012). Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. *The EMBO Journal* **31**: 1704–1713.
- Kedersha, N. and Anderson, P.** (2007). Mammalian stress granules and processing bodies. *Meth. Enzym.* **431**: 61–81.
- Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P.** (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* **169**: 871–884.
- Lambermon, M.H.L., Simpson, G.G., Kirk, D.A.W., Hemmings-Mieszczak, M., Klahre, U., and Filipowicz, W.** (2000). UBP1, a novel hnRNP-like protein that functions at multiple steps of higher plant nuclear pre-mRNA maturation. *EMBO J.* **19**: 1638–1649.
- Mazumder, B., Li, X., and Barik, S.** (2010). Translation control: A multifaceted regulator of inflammatory response. *The Journal of Immunology* **184**: 3311–3319.
- Parker, R. and Sheth, U.** (2007). P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**: 635–646.
- Pomeranz, M.C., Hah, C., Lin, P.-C., Kang, S.G., Finer, J.J., Blackshear, P.J., and Jang, J.-C.** (2010). The Arabidopsis tandem zinc finger protein AtTZF1 traffics between the nucleus and cytoplasmic foci and binds both DNA and RNA. *Plant Physiol.* **152**: 151–165.

- Tanaka, K.J., Ogawa, K., Takagi, M., Imamoto, N., Matsumoto, K., and Tsujimoto, M.** (2006). RAP55, a cytoplasmic mRNP component, represses translation in *Xenopus* oocytes. *J. Biol. Chem* **281**: 40096–40106.
- Vessey, J.P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M.A., and Macchi, P.** (2006). Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. *J Neurosci* **26**: 6496–508.
- Weber, C., Nover, L., and Fauth, M.** (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.* **56**: 517–530.
- Weston, A. and Sommerville, J.** (2006). Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation. *Nucl. Acids Res.* **34**: 3082–3094.
- Xu, J. and Chua, N.-H.** (2012). Dehydration stress activates Arabidopsis MPK6 to signal DCP1 phosphorylation. *EMBO J.* **31**: 1975–1984.
- Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H.** (2006). Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell* **18**: 3386–3398.

Appendix A

R code for analysis of nucleotide composition of microarray expression clusters of *Arabidopsis thaliana*

Code:

```
library(Biostrings)
library(ggplot2)
library(grid)

### Sequence Files (import as DNASTringSets)
At5utr=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_5_utr_20101028.txt", "fasta")
At3utr=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_3_utr_20101028.txt", "fasta")
Atcdna=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_cdna_20101214_updated.txt", "fasta")
Atcds=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_cds_20101214_updated.txt", "fasta")
Atintron=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_intron_20101028.txt", "fasta")
AtcdsREP=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_cds_20110103_representative_gene_model_updated.txt", "fasta")
AtcdnaREP=read.DNASTringSet("C:/Users/Reed/Desktop/Ranalysis/TAIR10_cdna_20110103_representative_gene_model_updated.txt", "fasta")

#for copy to server
setwd("/srv/zpools/tern.ib_bigdata/home/serreslab/rsorenson/WD")

At5utr=read.DNASTringSet("TAIR10_5_utr_20101028.txt", "fasta")
At3utr=read.DNASTringSet("TAIR10_3_utr_20101028.txt", "fasta")
Atcdna=read.DNASTringSet("TAIR10_cdna_20101214_updated.txt", "fasta")
```

```

Atcds=read.DNAStringSet("TAIR10_cds_20101214_updated.txt","fasta")
Atintron=read.DNAStringSet("TAIR10_intron_20101028.txt","fasta")
AtcdsREP=read.DNAStringSet("TAIR10_cds_20110103_representative_gene_model_updated.txt","f
asta")
AtcdnaREP=read.DNAStringSet("TAIR10_cdna_20110103_representative_gene_model_updated.txt",
"fasta")

RepGeneID= substring(names(AtcdnaREP),1,11)

#####
#####
# RepSeqs(SEQSET,Allisos=F,RepGeneIDs=RepGeneID) function definition
# = MAKES NEW DNAStringSet - trims down sequence sets by pulling out only the
representative gene model
# = as determined by RepGeneIDs
# SEQSET - generally DNAStringSet with 3'UTRs, 5'UTRs, introns, etc.
# (read from one of the fasta seq files imported above
# Allisos - if T nothing will happen to SETSEQ
# RepGeneIDs - vector or list of AGIs specifying representative isoforms)
#**timeconsuming (only do this one time per DNAStringSet per R session)
RepSeqs=function(SEQSET, Allisos=F,RepGeneIDs=RepGeneID ) {
  if(Allisos==F) {SeqSetAGIs=substring(names(SEQSET),1,11)
    Repls=intersect(RepGeneIDs,SeqSetAGIs)
    SEQS=SEQSET[which(Repls[1]==SeqSetAGIs)]
    for(i in 2:length(Repls))
SEQS=append(SEQS,SEQSET[which(Repls[i]==SeqSetAGIs)])
    SEQS}
  else SEQSET}

##### Function Definition
#####
# clust.meanATGC(cluster = cluster number, Sequences = DNAStringSet) function definition
# = takes a subset of Sequences which match AGIs from cluster and calculates overall

```

```

# = percent composition of A, T, C, G of each sequence and returns the mean for each
clust.meanATGC=function(cluster,Sequences) {
  clAGI=read.delim(paste("K",No.clust,"_",cluster,"-probes.txt",sep=""))[,2]
  AGIs=substring(names(Sequences),1,9)
  clean=intersect(clAGI,AGIs)
  seqs=Sequences[which(clean[1]==AGIs)]
  for(i in 2:length(clean)) seqs=append(seqs,Sequences[which(clean[i]==AGIs)])
  length(seqs)
  ATGC=data.frame(ID=substring(names(seqs),1,11),
                  Feature=100*alphabetFrequency(seqs)[,c(1,4,2,3)]/width(seqs))
  meanATGC=t(as.matrix(c(signif(cluster,digits=2),length(clean),colMeans(ATGC[,2:5]))))
}
#test=clust.meanATGC(2,At5utr)
#test

#####
# (same for standard deviation ) ####
#####
##### Function Definition
#####
# clust.sdATGC(cluster = cluster number, Sequences = DNASTringSet) function definition
# = takes a subset of Sequences which match AGIs from cluster and calculates overall
# = percent composition of A, T, C, G of each sequence and returns the standard deviation
for each
clust.sdATGC=function(cluster,Sequences) {
  clAGI=read.delim(paste("K",No.clust,"_",cluster,"-probes.txt",sep=""))[,2]
  AGIs=substring(names(Sequences),1,9)
  clean=intersect(clAGI,AGIs)
  seqs=Sequences[which(clean[1]==AGIs)]
  for(i in 2:length(clean)) seqs=append(seqs,Sequences[which(clean[i]==AGIs)])
  length(seqs)
  ATGC=data.frame(ID=substring(names(seqs),1,11),
                  Feature=100*alphabetFrequency(seqs)[,c(1,4,2,3)]/width(seqs))
  sdATGC=t(as.matrix(c(signif(cluster,digits=2),length(clean),apply(ATGC[,2:5],2,sd))))
}

```

```

}

##### FUNCTION DEFINITION
#####
#CompPlotATCG(SEQ,windowSize=30,percent=T) function definition
# = MAKES A TABLE of scanned window percent composition or if percent=F then straight
# = composition
# position = 5' end of window ; Table columns = position, A,T,C,G (%) columns

CompATCG=function(SEQ,windowSize=30,percent=T) {
  compos=matrix(numeric(0),ncol=4)
  for(i in 1:(length(SEQ[[1]])-(windowSize-1))) {
    compos=as.data.frame(rbind(compos,alphabetFrequency(SEQ[[1]][i:(i+windowSize-
1)])[c("A","T","C","G")]))
  }
  if(percent==T) composPercent=as.data.frame(cbind(position=1:(length(SEQ[[1]])-
(windowSize-1)),compos*100/windowSize))
  else if(percent==F) compos=as.data.frame(cbind(position=1:(length(SEQ[[1]])-
(windowSize-1)),compos))
}

#Examples
#test2=CompATCG(CSS,windowSize=25)
#test2[1:5,]

##### FUNCTION DEFINITION
#####
#CompPlot(seq,windosize=30,percent=T) function definition
# = makes a scanning composition plot for a the specified window size of a single sequence
# = e.g.Sequences[1] (Sequences=DNAStringSet object)
# requires: CompATCG() function - defined above
CompPlot=function(seq,windowSize=30,percent=T) {
  if(width(seq)<windowSize) NULL
  else {

```

```

p=ggplot(CompATCG(SEQ=seq,windowSize=windowSize,percent=percent))
p+geom_line(aes(x=position,y=A,color="A"))+
  geom_line(aes(x=position,y=T,color="T"))+
  geom_line(aes(x=position,y=C,color="C"))+
  geom_line(aes(x=position,y=G,color="G"))+
  ylab(if(percent==T) paste("Percent composition of ",windowSize,"nt window",sep="")
      else paste("Composition of ",windowSize,"nt window",sep=""))+
  opts(title=names(seq),
        plot.title=theme_text(size = 10))+

scale_colour_manual("",values=c("A"="green","T"="red","C"="blue","G"="black"),
                    breaks=c("A","T","C","G"))

}

}

#test3=CompPlot(SeqSet[35],32,TRUE)

#print(test3)

##### FUNCTION DEFINITION
#####

#ClusterSeqSet(cluster,Sequences) function definition

# = creates a DNASTringSet of genes in "cluster", from Sequences (DNASTringSet)
# = ignores genes not in the Sequences
# cluster AGIs are supplied from files in wd (name form = "K(No.clust)_(cluster)-
probes.txt"

# **(files are OUTPUT from GO_IN-separate_cluster_genelists.R script)

ClusterSeqSet=function(cluster,Sequences) {
  clAGI=read.delim(paste("K",No.clust,"_",cluster,"-probes.txt",sep=""))[,2]
  AGIs=substring(names(Sequences),1,9)
  clean=intersect(clAGI,AGIs)
  seqs=Sequences[which(clean[1]==AGIs)]
  for(i in 2:length(clean)) seqs=append(seqs,Sequences[which(clean[i]==AGIs)])
  seqs
}

```

```

##### FUNCTION DEFINITION
#####
#RichElem(Sequences=DNAStrngSet) function definition
# = returns a list(maxATCG = table with columns: Gene, maximum A,T,C,G window composition
# = for for each sequence in a DNAStrngSet, and 4 vectors with percent position within
the sequences
# = at which the window composition maxima occur maxApos,maxTpos,maxCpos,maxGpos
# requires: CompATCG() function - defined above

RichElem=function(Sequences,windowSize=30,percentComp=T,Position="percent") {
  richelem=matrix(numeric(0),ncol=5)
  maxApos= vector()
  maxTpos= vector()
  maxCpos= vector()
  maxGpos= vector()
  for(k in 1:length(Sequences)) {
    tt=CompATCG(Sequences[k],windowSize,percent=percentComp) # makes a composition table
for each sequence (k) to find the max below
    richelem=rbind(richelem,
c(names(Sequences[k]),max(tt[, "A"]),max(tt[, "T"]),max(tt[, "C"]),max(tt[, "G"])))
    if(Position=="percent") {
      maxApos=c(maxApos,tt[which(max(tt[, "A"])==tt[, "A"]), "position"]*100/(width(CSS[k])-
(windowSize-1)))
      maxTpos=c(maxTpos,tt[which(max(tt[, "T"])==tt[, "T"]), "position"]*100/(width(CSS[k])-
(windowSize-1)))
      maxCpos=c(maxCpos,tt[which(max(tt[, "C"])==tt[, "C"]), "position"]*100/(width(CSS[k])-
(windowSize-1)))
      maxGpos=c(maxGpos,tt[which(max(tt[, "G"])==tt[, "G"]), "position"]*100/(width(CSS[k])-
(windowSize-1)))
    } else if(Position=="5align") {
      maxApos=c(maxApos,tt[which(max(tt[, "A"])==tt[, "A"]), "position"])
      maxTpos=c(maxTpos,tt[which(max(tt[, "T"])==tt[, "T"]), "position"])
      maxCpos=c(maxCpos,tt[which(max(tt[, "C"])==tt[, "C"]), "position"])
    }
  }
}

```

```

    maxGpos=c(maxGpos,tt[which(max(tt[, "G"])==tt[, "G"]), "position"])
  } else if(Position=="3align") {
    maxApos=c(maxApos,tt[which(max(tt[, "A"])==tt[, "A"]), "position"]-nrow(tt))
    maxTpos=c(maxTpos,tt[which(max(tt[, "T"])==tt[, "T"]), "position"]-nrow(tt))
    maxCpos=c(maxCpos,tt[which(max(tt[, "C"])==tt[, "C"]), "position"]-nrow(tt))
    maxGpos=c(maxGpos,tt[which(max(tt[, "G"])==tt[, "G"]), "position"]-nrow(tt))
  }

  print(paste("max composition - seq", k, "out of", length(Sequences)))}
return(list(maxATCG=data.frame(Gene=richelem[,1],
                                "A"=round(as.numeric(richelem[,2]),1),
                                "T"=round(as.numeric(richelem[,3]),1),
                                "C"=round(as.numeric(richelem[,4]),1),
                                "G"=round(as.numeric(richelem[,5]),1)),

maxApos=round(maxApos,1),maxTpos=round(maxTpos,1),maxCpos=round(maxCpos,1),maxGpos=round(
maxGpos,1)))
}

#test=RichElem(CSS[1:100])
#test

##### FUNCTION DEFINITION
#####

# MaxWinCompPlot(RichElement output list )
# = makes an density plot of maximum comositions windows from a given sequence set
# INPUT : e.g.MaxWindows=RichElem(CSS (cluster sequence set),25,TRUE,"5'align") # list of
names,max ATCG window table, max(A,T,C,G)pos vectors

MaxWinCompPlot=function(maxWindows>window,windowsize, percent=T,cluster) {
  p=ggplot(maxWindows$maxATCG) #p=ggplot(test$maxATCG)
  print(
  p+
  stat_density(aes(A, colour="A"), size=0.1, geom="path")+

```

```

stat_density(aes(T, colour="T"), size=0.1, geom="path")+
stat_density(aes(C, colour="C"), size=0.1, geom="path")+
stat_density(aes(G, colour="G"), size=0.1, geom="path")+
ylab("Density")+
xlab("Composition of Maximum Window")+
scale_colour_manual("", values=c("A"="green", "T"="red", "C"="blue", "G"="black"),
breaks=c("A", "T", "C", "G"))+
  opts(title=paste("K", No.clust, " Cluster ", cluster, " ", Feature),
xlab="Composition of Maximum Window",
panel.background = theme_blank(),
theme_rect(colour = 'white',
fill = 'white',
size = 3,
linetype='solid'),
panel.grid.minor = theme_blank(),
axis.line = theme_segment(colour = 'black', size = 1,
linetype = 'solid'),
plot.margin = unit(c(1,1,1,1), "lines"),
axis.title.x = theme_text(size=16, face="bold"),
axis.text.x = theme_text(colour="black", size=16),
axis.title.y = theme_text(size=18, face="bold", angle=90),
axis.text.y = theme_text(colour="black", size=20),
#legend.background = theme_rect(colour="white", fill =
'white'),
#legend.key=theme_blank(),
#legend.position = c(0.9, 0.95),
#legend.text = theme_text(size = 18)
legend.position = "none"
)+
#scale_y_continuous(limits = c(0,0.1), breaks=seq(from=0, to=0.1, 0.02))+
if(percent == T) scale_x_continuous(limits = c(0,100), breaks=seq(from=0, to=100, by=20))
else {
  if(percent==F) scale_x_continuous(limits =
c(0, window_size), breaks=seq(0, window_size, by=window_size/5))}

```

```

)
}

##### FUNCTION DEFINITION
#####

# MaxWinCompPosPlot(RichElement output list )
# = makes an density plot of position of the maximum comositions windows from a given
sequence set
# INPUT : e.g.MaxWindows=RichElem(CSS (cluster sequence set),25,TRUE,"5'align") # list of
names,max ATCG window table, max(A,T,C,G)pos vectors

MaxWinCompPosPlot=function(maxWindows,Position="percent",cluster) {
  maxWinPos=data.frame()
  for(i in 1:4) for(m in 1:length(maxWindows[[1+i]])) maxWinPos[m,i] =
maxWindows[[1+i]][m]
  colnames(maxWinPos)=c("A", "T", "C", "G")
  p=ggplot(maxWinPos)
  print(
  p+
    stat_density(aes(A, colour="A"), size=0.1, geom="path", na.rm=T)+
    stat_density(aes(T, colour="T"), size=0.1, geom="path", na.rm=T)+
    stat_density(aes(C, colour="C"), size=0.1, geom="path", na.rm=T)+
    stat_density(aes(G, colour="G"), size=0.1, geom="path", na.rm=T)+
    #scale_y_continuous(limits = c(0,0.1),breaks=seq(from=0,to=0.1,0.02))+
    xlab("Position of maximum composition window")+
    ylab("Density")+
    scale_colour_manual("", values=c("A"="green", "T"="red", "C"="blue", "G"="black"),
      breaks=c("A", "T", "C", "G"))+
      opts(title=paste("K",No.clust," Cluster ",cluster," ",Feature),
        panel.background = theme_blank(),
        theme_rect(colour = 'white',
          fill = 'white',
          size = 3,
          linetype='solid'),

```

```

panel.grid.minor = theme_blank(),
axis.line = theme_segment(colour = 'black', size = 1,
linetype = 'solid'),

plot.margin = unit(c(0,0,0,0),"lines"),
axis.title.x = theme_text(size=6,face="bold"),
axis.text.x = theme_text(colour="black",size=6),
axis.title.y = theme_text(size=6,face="bold",angle=90),
axis.text.y = theme_text(colour="black",size=6),
#legend.background = theme_rect(colour="white",fill =
'white'),

#legend.key=theme_blank()
#legend.position = c(0.9, 0.9),
#legend.text = theme_text(size = 18)
legend.position = 'none'
)+
  if(Position == "percent") scale_x_continuous(limits =
c(0,100),breaks=seq(from=0,to=100,by=20)) else {
  if(Position== "5align") scale_x_continuous(limits =
c(0,max(maxWinPos,na.rm=T)),breaks=seq(0,max(maxWinPos,na.rm=T),by=max(maxWinPos,na.rm=T)
/5)) else {
  if(Position== "3align") scale_x_continuous(limits =
c(-max(maxWinPos,na.rm=T),0),breaks=seq(-
max(maxWinPos,na.rm=T),0,by=max(maxWinPos,na.rm=T)/5))
  }}
)
}

```

```
##### ACTIVE CODE
```

```
#####
```

```
# INPUTs
```

```
# **files in working directory for each cluster as: "K(No.clust)_(cluster)-probes.txt"
```

```

# **(these files are OUTPUT from GO_IN-separate_cluster_genelists.R script)

SeqSet=RepSeqs(At3utr)

No.clust = 20

Feature = "3UTR"

OutputFile=paste("K",No.clust,"_",Feature,"_ATGC_ClusterContent.txt",sep="")

OutputFile

#makes a table of clust.meanATGC values for each cluster
# columns = cluster number, number of gene seqs used from cluster, %A mean, %T mean, %C
mean, %G mean

MEANS=matrix(data=numeric(0),ncol=6) #empty matrix of 6 rows
for(i in 1:No.clust) {meanATGC=clust.meanATGC(i,SeqSet) # binds ATGC means from each
cluster into one data.frame

                MEANS=rbind(MEANS,meanATGC)

                print(i)}

colnames(MEANS)[1:2]=c("cluster","size")

MEANS

#makes a table of clust.sdATGC values for each cluster
SD=matrix(data=numeric(0),ncol=6) #empty matrix of 6 rows
for(i in 1:No.clust) {sdATGC=clust.sdATGC(i,SeqSet) # binds ATGC stdev from each cluster
into one data.frame

                SD=rbind(SD,sdATGC)

                print(i)}

#joins MEANS table from above and data from SD table and writes to Output file
content=cbind(MEANS, SD[,3:6])

content

write.table(content,OutputFile,sep="\t",row.names=F)

##### MORE ACTIVE CODE

#####

```

```

# generate PDF FILE a CompPlot for each Seq in Set of DNASTringSets generated from each
cluster.

# therefore 1 pdf file for each cluster with a CompPlot graph for each seq
# this is slow and requires:

# INPUTs

# **files in working directory for each cluster as: "K(No.clust)_(cluster)-probes.txt"
# **(these files are OUTPUT from GO_IN-separate_cluster_genelists.R script)

WinSize=25 #sets window size

for(j in 1:No.clust) {

  CSS=ClusterSeqSet(j,SeqSet) # makes cluster sequence set (CSS)

  pdf(paste("K",No.clust,"_",j,"-",Feature,"_ATGC_SeqComp.pdf",sep=""))
  for(k in 1:length(CSS)) {print(CompPlot(CSS[k],WinSize,TRUE))
                          print(paste("cluster ",j,"-",k))}

  dev.off()
}

pdf(paste("K",No.clust,Feature,"maxcomp",WinSize,"nt_window.pdf",sep=""),width=1.5,
height=1.5)

for(j in 1:No.clust) {

  CSS=ClusterSeqSet(j,SeqSet)

  MaxWindows=RichElem(CSS,windowsize=WinSize,percentComp=F,Position="percent") # list of
names,max ATCG window table, max(A,T,C,G)pos vectors

  print(
    MaxWinCompPlot(maxWindows=MaxWindows,windowsize=WinSize, percent=F,cluster=j)
    MaxWinCompPosPlot(MaxWindows,Position="percent",cluster=j)
  )
}

dev.off()

# for a single cluster

#CSS=ClusterSeqSet(cluster=1,SeqSet)

#pdf(paste("K",No.clust,"_",1,"-",Feature,"_ATGC_SeqComp.pdf",sep=""))

#for(k in 1:length(CSS)) {print(CompPlot(CSS[k],20,TRUE))

```

```

#                                     print(k)}

#dev.off()

##### CODE IN DEVELOPMENT
#####

# makes a table of maximum compositions from the max window for 1 column for each
nucleotide ATCG for each cluster (20 clusters*4nt =80rows)

Adensities=data.frame()
Tdensities=data.frame()
Cdensities=data.frame()
Gdensities=data.frame()

Amaxpos=data.frame()
Umaxpos=data.frame()
Cmaxpos=data.frame()
Gmaxpos=data.frame()

for(j in 1:No.clust) {
  CSS=ClusterSeqSet(j,SeqSet)

  MaxWindows=RichElem(CSS,windowsize=WinSize,percentComp=F,Position="percent")
  for(i in 1:nrow(MaxWindows$maxATCG)) Adensities[i,j]=MaxWindows$maxATCG$A[i]
  for(i in 1:nrow(MaxWindows$maxATCG)) Tdensities[i,j]=MaxWindows$maxATCG$T[i]
  for(i in 1:nrow(MaxWindows$maxATCG)) Cdensities[i,j]=MaxWindows$maxATCG$C[i]
  for(i in 1:nrow(MaxWindows$maxATCG)) Gdensities[i,j]=MaxWindows$maxATCG$G[i]
  colnames(Adensities)[j]=paste("c",j,"_A-",Feature,sep="")
  colnames(Tdensities)[j]=paste("c",j,"_U-",Feature,sep="")
  colnames(Cdensities)[j]=paste("c",j,"_C-",Feature,sep="")
  colnames(Gdensities)[j]=paste("c",j,"_G-",Feature,sep="")

  for(i in 1:length(MaxWindows$maxApos)) Amaxpos[i,j]=MaxWindows$maxApos[i]
  for(i in 1:length(MaxWindows$maxTpos)) Umaxpos[i,j]=MaxWindows$maxTpos[i]
  for(i in 1:length(MaxWindows$maxCpos)) Cmaxpos[i,j]=MaxWindows$maxCpos[i]
  for(i in 1:length(MaxWindows$maxGpos)) Gmaxpos[i,j]=MaxWindows$maxGpos[i]
  colnames(Amaxpos)[j]=paste("c",j,"_A-",Feature," maxpos",sep="")
  colnames(Umaxpos)[j]=paste("c",j,"_U-",Feature," maxpos",sep="")

```

```

colnames(Cmaxpos)[j]=paste("c",j,"_C-",Feature," maxpos",sep="")
colnames(Gmaxpos)[j]=paste("c",j,"_G-",Feature," maxpos",sep="")
}

densmaxATCG=cbind(Adensities,Tdensities,Cdensities,Gdensities)
write.table(densmaxATCG,"densmaxATCG.txt",sep="\t")

write.table(Amaxpos,"Amaxpos.txt",sep="\t")
write.table(Umaxpos,"Umaxpos.txt",sep="\t")
write.table(Cmaxpos,"Cmaxpos.txt",sep="\t")
write.table(Gmaxpos,"Gmaxpos.txt",sep="\t")

#densmaxATCG=read.delim("densmaxATCG.txt")
#ATCGmaxpos=read.delim("ATCGmaxpos.txt")

library("gplots") # Loads the gplots library that contains the heatmap.2() function.

pdf("Maximum 3UTR Composition 25nt.pdf",width=4.0,height=4.5)

##### Pseudo-FUNCTION ##### (making this into a function doesn't work; must use
all of the code each time)
# plots densities of densmaxATCG table into 4 density plots and color codes the lines by
cluster 1 plot / nucleotide

nt="G" # A, U, C, G
linecolors=colorpanel(21,"black","blue","orange")
Feature="3\'UTR"
xdata=densmaxATCG[,grep(paste("_",nt,sep=""),colnames(densmaxATCG))]

p=ggplot(xdata)
p+
  stat_density(aes(x=xdata[,20],colour="20"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,19],colour="19"),size=01,geom="path",na.rm=T)+

```

```

stat_density(aes(x=xdata[,18], colour="18"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,17], colour="17"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,16], colour="16"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,15], colour="15"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,14], colour="14"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,13], colour="13"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,12], colour="12"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,11], colour="11"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,10], colour="10"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,9], colour="9"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,8], colour="8"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,7], colour="7"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,6], colour="6"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,5], colour="5"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,4], colour="4"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,3], colour="3"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,2], colour="2"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,1], colour="1"), size=01, geom="path", na.rm=T) +
scale_y_continuous(limits = c(0,0.3), breaks=seq(from=0, to=0.3, 0.1)) +
xlab(paste("Max ", nt, " Composition (nt)", sep="")) +
ylab("Density") +
scale_colour_manual("Cluster", values=c("1"=linecolors[20],
                                         "2"=linecolors[19],
                                         "3"=linecolors[18],
                                         "4"=linecolors[17],
                                         "5"=linecolors[16],
                                         "6"=linecolors[15],
                                         "7"=linecolors[14],
                                         "8"=linecolors[13],
                                         "9"=linecolors[12],
                                         "10"=linecolors[11],
                                         "11"=linecolors[10],
                                         "12"=linecolors[9],
                                         "13"=linecolors[8],

```

```

"14"=linecolors[7],
"15"=linecolors[6],
"16"=linecolors[5],
"17"=linecolors[4],
"18"=linecolors[3],
"19"=linecolors[2],
"20"=linecolors[1]),
breaks=c("1", "2", "3", "4", "5",
"6", "7", "8", "9", "10",
"11", "12", "13", "14", "15",
"16", "17", "18", "19", "20")+
opts(title=paste(Feature, sep=""),
plot.title.position=c(0.3,0.9),
panel.background = theme_blank(),
theme_rect(colour = 'white',
fill = 'white',
size = 3,
linetype='solid'),
panel.grid.minor = theme_blank(),
axis.line = theme_segment(colour = 'black', size =
1, linetype = 'solid'),
plot.margin = unit(c(1,1,1,1), "lines"),
axis.title.x = theme_text(size=16, face="bold"),
axis.text.x = theme_text(colour="black", size=16),
axis.title.y =
theme_text(size=16, face="bold", angle=90),
axis.text.y = theme_text(colour="black", size=16),
legend.background = theme_rect(colour="white", fill
= 'white'),
legend.key=theme_blank(),
#legend.position = c(0.9, 0.9),
legend.text = theme_text(size = 12),
legend.position = 'right',
legend.key.height = unit(3.25/20, "in")

```

```

)+
scale_x_continuous(limits =
c(0,25),breaks=seq(from=0,to=25,by=5))

##### repeat above for each nt and then turn off device #####
dev.off()

##### for max position plotting
#####

pdf("Posistion of Maximum 3UTR Composition 25nt.pdf",width=4.3,height=4.5)

##### Pseudo-FUNCTION ##### (making this into a function doesn't work; must use
all of the code each time)
# plots densities of densmaxATCG table into 4 density plots and color codes the lines by
cluster 1 plot / nucleotide

nt="G" # A, U, C, G
xdata=Gmaxpos

linecolors=colorpanel(21,"black","blue","orange")
Feature="3\'UTR"

p=ggplot(xdata)
p+
  stat_density(aes(x=xdata[,20],colour="20"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,19],colour="19"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,18],colour="18"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,17],colour="17"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,16],colour="16"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,15],colour="15"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,14],colour="14"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,13],colour="13"),size=01,geom="path",na.rm=T)+
  stat_density(aes(x=xdata[,12],colour="12"),size=01,geom="path",na.rm=T)+

```

```

stat_density(aes(x=xdata[,11], colour="11"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,10], colour="10"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,9], colour="9"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,8], colour="8"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,7], colour="7"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,6], colour="6"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,5], colour="5"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,4], colour="4"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,3], colour="3"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,2], colour="2"), size=01, geom="path", na.rm=T) +
stat_density(aes(x=xdata[,1], colour="1"), size=01, geom="path", na.rm=T) +
scale_y_continuous(limits = c(0,0.02), breaks=seq(from=0, to=0.02, 0.005)) +
xlab(paste("Position of Max ", nt, " Composition (nt)", sep="")) +
ylab("Density") +
scale_colour_manual("Cluster", values=c("1"=linecolors[20],
                                         "2"=linecolors[19],
                                         "3"=linecolors[18],
                                         "4"=linecolors[17],
                                         "5"=linecolors[16],
                                         "6"=linecolors[15],
                                         "7"=linecolors[14],
                                         "8"=linecolors[13],
                                         "9"=linecolors[12],
                                         "10"=linecolors[11],
                                         "11"=linecolors[10],
                                         "12"=linecolors[9],
                                         "13"=linecolors[8],
                                         "14"=linecolors[7],
                                         "15"=linecolors[6],
                                         "16"=linecolors[5],
                                         "17"=linecolors[4],
                                         "18"=linecolors[3],
                                         "19"=linecolors[2],
                                         "20"=linecolors[1]),

```

```

breaks=c("1","2","3","4","5",
         "6","7","8","9","10",
         "11","12","13","14","15",
         "16","17","18","19","20")+
  opts(title=paste(Feature,sep=""),
       plot.title.position=c(0.3,0.9),
       panel.background = theme_blank(),
       theme_rect(colour = 'white',
                 fill = 'white',
                 size = 3,
                 linetype='solid'),
       panel.grid.minor = theme_blank(),
       axis.line = theme_segment(colour = 'black', size =
1, linetype = 'solid'),

       plot.margin = unit(c(1,1,1,1),"lines"),
       axis.title.x = theme_text(size=16,face="bold"),
       axis.text.x = theme_text(colour="black",size=16),
       axis.title.y =
theme_text(size=16,face="bold",angle=90),
       axis.text.y = theme_text(colour="black",size=16),
       legend.background = theme_rect(colour="white",fill
= 'white'),

       legend.key=theme_blank(),
       #legend.position = c(0.9, 0.9),
       legend.text = theme_text(size = 12),
       legend.position = 'right',
       legend.key.height = unit(3.25/20, "in")
)+
  scale_x_continuous(limits =
c(0,100),breaks=seq(from=0,to=100,by=20))

##### repeat above for each nt and then turn off device #####
dev.off()

```