The Dissertation Committee for Ryan Michael Patrick Certifies that this is the approved version of the following dissertation:

The Roles of Alternative Cap-Binding Proteins of *Arabidopsis thaliana*

Committee:

Karen S. Browning, Supervisor

David W. Hoffman

Enamul Huq

Alan M. Lloyd

Rick Russell

The Roles of Alternative Cap-Binding Proteins of *Arabidopsis thaliana*

by

Ryan Michael Patrick, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2015

Dedication

To my family for their support.

Acknowledgements

I would like to thank Dr. Karen Browning and the members of the Browning lab and for their collaboration and support over the years. This work would not have been possible without the extraordinary mentorship of Dr. Browning, for which I am very grateful. This research was made possible by support from the National Science Foundation.

The Roles of Alternative Cap-binding Proteins of *Arabidopsis thaliana*

Ryan Michael Patrick, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Karen Browning

The mRNA cap-binding complexes eIF4F (made up of the cap-binding protein eIF4E and the large scaffold eIF4G) and eIFiso4F (made up of the plant-specific isoforms eIFiso4E and eIFiso4G) have established roles in translation initiation. However, other cap-binding proteins are known to be encoded in the *Arabidopsis thaliana* genome. We have chosen to investigate the biochemical properties and potential functions of these proteins. We have identified the eIF4E-like proteins, eIF4E1b and eIF4E1c, as Brassicaceae-specific eIF4E isoforms with the ability to form cap-binding complexes. These proteins are able to complement an eIF4E deletion in yeast. However, their limited expression in *A. thaliana* along with their relatively weak binding affinity for eIF4G and more limited ability to promote translation *in vitro* indicate a possible role outside of canonical translation initiation pathways.

The alternative cap-binding protein 4EHP is conserved from animals to plants, but its role and binding partners in plants are not well defined. We demonstrate that a homologous complex to the 4EHP-GIGYF2 cap-binding complex observed in mammals is present in *A. thaliana*. The plant complex appears important to proper development, as double knockouts show a noticeable developmental phenotype and dysregulation of gene expression, but the viability of these knockouts in *A. thaliana* may offer an opportunity to research the complex's function that cannot be performed in animal systems as knockouts

are lethal. RNA immunoprecipitation studies find that 4EHP and GIGYF associate with non-coding RNA in *A. thaliana*, and nucleocytoplasmic fractionation supports a possible nuclear role for the proteins. These findings indicate that the 4EHP-GIGYF complex may have an unexpected role in bridging non-coding RNA to gene expression in plants.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. PROTEIN BIOSYNTHESIS

The central dogma of biology describes the flow of genetic information in a cell from units of DNA, transcribed as messenger RNA (mRNA) and translated by the ribosome, to proteins. Translation rates of mRNA can have a large impact on the level of protein in eukaryotic cells and be highly variable between mRNAs (Schwanhausser et al., 2011), though there is some disagreement over whether translational control is significant, but secondary to transcription (Jovanovic et al., 2015) or if translational control is predominantly responsible for variation in protein levels (Schwanhausser et al., 2011). Protein biosynthesis in eukaryotes is comprised of three steps: initiation, in which the RNA is recruited to the 40S subunit of the ribosome followed by joining with the 60S subunit to form the 80S ribosome; elongation, in which the ribosome decodes the codons present in the open reading frame (ORF) of the mRNA with matching transfer RNAs (tRNA) carrying amino acids in order to synthesize the protein; and termination, in which a release factor enters the decoding site of the ribosome and triggers the release of the newly synthesized protein and disassembly of the ribosome. Translation initiation is a highly regulated process with contributions by greater than a dozen distinct eukaryotic initiation factor (eIF) proteins or protein complexes (Aitken and Lorsch, 2012; Browning and Bailey-Serres, 2015).

1.1.1. Eukaryotic Translation Initiation

Primary RNA transcripts produced by RNA polymerase II (Pol II) are co-transcriptionally capped by the guanylyltransferase activity of capping enzyme complex at

the 5' end of the nascent transcript by guanosine in a 5'-5' linkage with a triphosphate bridge (GpppN) (Martinez-Rucobo et al., 2015), which is then modified by RNA (guanine-N7-)methyltransferase to produce the mature mRNA cap structure (m^7GpppN) (Mao et al., 1995). The presence of the methylated cap on the nascent transcript, which is recognized and bound by the nuclear cap-binding complex, is important in promoting intron splicing and 3' end processing and polyadenylation to produce the mature mRNA transcript (Flaherty et al., 1997; Izaurralde et al., 1994; Topisirovic et al., 2011). The presence of the mRNA cap is critical to mRNA stability, with surveillance mechanisms present to ensure capping (Jiao et al., 2010) and transcript decapping an important step committing to mRNA to degradation (Braun et al., 2012).

Translation initiation requires mRNA recruitment to the pre-initiation complex (PIC), which is made up of the 40S ribosomal subunit, the ternary complex of Met-$tRNA_i$ and eIF2 bound to GTP, and several initiation factors including eIF1, eIF1A, eIF3, and eIF5 (Aitken and Lorsch, 2012). eIF1 and eIF1A stabilize the open conformation of the PIC to facilitate mRNA loading (Aitken and Lorsch, 2012; Passmore et al., 2007) while eIF3 appears to coordinate interactions between PIC factors and is critical to mRNA recruitment *in vitro* and *in vivo* (Aitken and Lorsch, 2012; Jivotovskaya et al., 2006; Mitchell et al., 2010). After the binding of mRNA, the PIC scans the 5' end of the mRNA for the initiation codon, and following its recognition a series of events complete the initiation process including the release of eIF1, the hydrolysis of the GTP bound to eIF2 coinciding with release of eIF2 and eIF5, and finally the release of eIF1A coordinated by eIF5B to facilitate 60S subunit joining (Aitken and Lorsch, 2012).

mRNA recruitment to the PIC is mediated by the binding of the cap-binding complex eIF4F, made up of eIF4E, a relatively small protein (~25 kDa) which recognizes the 7-methylguanosine cap structure, and eIF4G, a large scaffolding protein (~188 kDa);

in mammals and yeast the DEAD-box RNA helicase eIF4A, which binds to eIF4G, is considered part of eIF4F but it is loosely associated with the cap-binding complex in plants (Browning, 2014). The structure of eIF4G includes RNA binding regions, HEAT domains (the number of which vary between organisms) that contact eIF4A, and binding sites for eIF4B (Cheng and Gallie, 2010, 2013). Binding sites on eIF4G for poly(A) binding protein (PABP), which binds the 3' poly(A) tail, may contribute to circularize the transcript so that the 5' and 3' ends are in proximity and protected (Browning, 2014; Gallie and Liu, 2014; Goss and Kleiman, 2013).

The binding interaction of the cap-binding complex to the mRNA promotes unwinding of secondary structure to in turn facilitate scanning by the PIC. The ATP binding and hydrolysis activity of eIF4A leads the protein through a series of conformational changes that cause RNA binding and release, and eIF4G binding enhances the ATP hydrolysis activity by restricting its conformational changes (Hilbert et al., 2011). eIF4B, and its mammalian homolog eIF4H, have RNA binding activity and stimulate eIF4A helicase activity (Andreou and Klostermeier, 2014). eIF4G and eIF4B are required for 5'-3' processivity of RNA unwinding by eIF4A (Garcia-Garcia et al., 2015). Interaction between eIF4G and the PIC is through binding to eIF3 or eIF5 depending on the organism (Aitken and Lorsch, 2012).

The eIF4F complex is an important determinant of mRNA fate: translation, storage or degradation. Stresses and translational inhibition trigger formation of stress granules, sites of translational repression which contain eIF4E, eIF4G, eIF4A, eIF3, and other PIC components (Decker and Parker, 2012). The DEAD-box helicase DDX3 (related to yeast Ded1p) which can also be found in stress granules is necessary for the translation of certain transcripts with complex secondary structure through direct association with eIF4G (Soto-Rifo et al., 2012).

eIF4G is also important in the pioneer round of translation, in which premature termination codons can trigger nonsense-mediated decay (Lejeune et al., 2004). Exon junction complexes (EJC) are present on newly spliced mRNA and recruit the NMD factors UPF3 and UPF2, which can in turn activate NMD through UPF1 if a premature stop codon is detected (Kervestin and Jacobson, 2012). eIF4 G is required for nonsense-mediated decay and associates with both EJC and NMD factors (Lejeune et al., 2004). eIF4G, partly through interaction with PABP, also contributes to suppressing NMD in the presence of long 3' UTR structures (Fatscher et al., 2014) and probably on short ORFs (Pereira et al., 2015).

The eIF4F complex is an important contributor to protecting the mRNA and preventing degradation through competition with decapping enzymes such as Dcp1p (Mitchell and Tollervey, 2001). However, eIF4G can bind to Dcp1p *in vitro* and therefore may potentially modulate decapping *in vivo* (Vilela et al., 2000). A similar observation has been made for eIF4G binding by the yeast decapping activator Scd6, which has a homolog in *Arabidopsis* (DCP5) that also promotes decapping (Rajyaguru et al., 2012). In mammals, eIF4G may also play a role in regulation of AU-rich element (ARE) containing mRNA through binding to AUF1, an RNA binding proteins which recognizes ARE elements and targets them for degradation (Laroia et al., 1999; Lu et al., 2006).

1.2. TRANSLATIONAL REGULATION THROUGH eIF4F

Various crystal structures are available of eIF4E, showing that its principal features are a concave pocket for binding the cap moiety (Marcotrigiano et al., 1997) and N-terminal extension that is the primary region for intermolecular contacts with eIF4G (Gross et al., 2003). Two evolutionarily conserved tryptophan residues in the cap-binding pocket

sandwich the base of the 7-methylguanosine cap while a third conserved tryptophan forms a van der Waals contact with the N7-methyl group and charged residues contact the 5'-5' triphosphate bridge (Marcotrigiano et al., 1997). eIF4E family members can be classified based in part on the conservation of the N-terminal cap-binding tryptophan residue; Class I proteins, which are primarily the eIF4E family members that form complexes with eIF4G and participate in translation initiation, maintain it absolutely (Joshi et al., 2005). Class II proteins, such as 4E-Homologous Protein (4EHP, discussed in later chapters), have this tryptophan substituted with other hydrophobic residues such as tyrosine or phenylalanine, and Class III proteins also lack conservation of this residue, such as human eIF4E3 which has a cysteine at this position and binds by a different mechanism (Joshi et al., 2005; Osborne et al., 2013).

eIF4G has a binding site for eIF4E which is located N-terminal to the HEAT domains (Figure 1.1) and has the consensus motif $Y(x)_4L\Phi$ where x is any amino acid and Φ is a hydrophobic residue (Gross et al., 2003). This consensus region of eIF4G is responsible for forming a hydrophobic structure that makes extensive contacts with the eIF4E N-terminal region but not the cap-binding pocket (Gross et al., 2003). It is well characterized that eIF4G binding to eIF4E allosterically enhances the binding of the latter protein to mRNA cap structure (Haghighat and Sonenberg, 1997; Von der Haar et al., 2000). The formation of the eIF4F complex along with PABP binding to eIF4G leads to a series of additive redundant contacts through enhanced eIF4E cap binding and eIF4G/PABP RNA-binding activity that is believed to stabilize the mRNP and serve as a platform for multiple rounds of ribosomal loading (Gross et al., 2003).

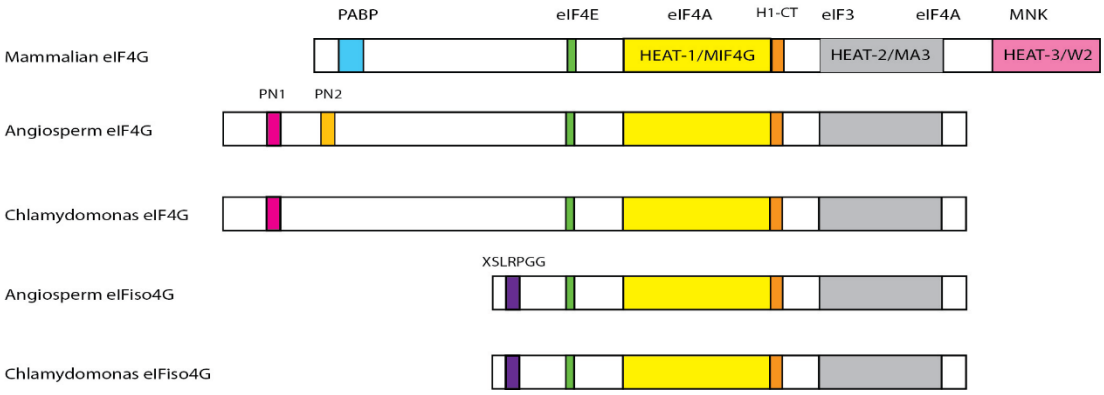


Figure 1.1. **Domain Organization of eIF4G in Mammals and Plants**

Plant eIF4G and eIFiso4G share the core organization of the eIF4E-binding site, HEAT-1/MIF4G domain, H1-CT motif, and HEAT-2/MA3 domain with mammals, but do not have the C-terminal HEAT-3 domain, which is the site of MNK binding. The eIF3 and eIF4A binding regions are thought to be maintained between all shown isoforms. Plant eIF4G has a longer N-terminus than mammals and contain the plant-specific 4G-PN1 and 4G-PN2 motifs as shown. eIFiso4G is remarkably well conserved across plants, with the N-terminal XSLRPGG motif maintained from green algae to angiosperms (Patrick and Browning, 2012).

The eIF4F complex is an important target of global translational regulation, and can be targeted through multiple mechanisms to sequester eIF4G from eIF4E and/or reduce eIF4E cap-binding affinity. In humans and yeast, translational regulation through disruption of eIF4E binding to eIF4G is primarily through 4E-Binding Proteins (4E-BPs) which share the $Y(x)_4L\Phi$ motif present on eIF4G and are modulated through the mammalian target of rapamycin (mTOR) pathway (Topisirovic et al., 2011). Activation of the mTOR pathway through pro-growth signals and nutrient availability leads to

phosphorylation of 4E-BPs which reduces their affinity for eIF4E, while in the absence of signal the non-phosphorylated 4E-BPs are able to compete with eIF4G for eIF4E binding (Richter and Sonenberg, 2005; Robichaud and Sonenberg, 2014). In yeast, the 4E-BPs Caf20p and Eap1p confer translational repression to specific subsets of mRNAs through interaction with RNA-binding PUF proteins and Eap1p may also serve as a bridge to degradation of transcripts through interaction with decapping factor Dhh1p (Blewett and Goldstrohm, 2012; Cridge et al., 2010). Interestingly, however, there is no evidence to date of 4E-BPs or a similar regulatory mechanism in plants (Browning and Bailey-Serres, 2015).

Other eIF4E-binding proteins convey mRNA-specific translational repression through tethering action of an RNA-binding protein (Richter and Sonenberg, 2005). In *Drosophila*, Cup binds through a $Y(x)_4L\Phi$ motif to eIF4E and represses translation of specific developmentally regulated transcripts through interaction with RNA-binding proteins such as Smaug and Bruno that recognize elements in the 3' UTR of their targets (Kinkelin et al., 2012). In animals, eIF4E-binding proteins such as Maskin and Neuroguidin confer translational repression with important contributions to cell cycle progression and neural tube closure respectively, by in turn binding to the RNA-binding cytoplasmic polyadenylation element-binding protein (CPEB) to regulate specific mRNA populations (Cao et al., 2006; Jung et al., 2006). eIF4E-binding proteins can be localized mRNA granules that are stored in a translationally silenced state but are able to be rapidly be translationally derepressed upon an external signal. This paradigm is illustrated by fragile X mental retardation protein (FMRP), which targets certain neuronal mRNAs to synapses and stores them in a translationally repressed state through the action of the eIF4E-binding cytoplasmic FMRP interacting protein 1 (CYFIP1), which releases eIF4E upon neuronal stimulation to promote localized protein synthesis (Napoli et al., 2008).

eIF4E is also a direct target of phosphorylation. In mammals, eIF4G contains a third HEAT domain that binds to mitogen-activated protein kinase (MAPK) signal-integrating kinases called Mnk1 and Mnk2. These kinases phosphorylate a C-terminal serine residue of eIF4E, promoting the translation of a subset of mRNAs which if unchecked can lead to tumorigenesis (Furic et al., 2014). Translational control of specific mRNA transcripts through the eIF4E phosphorylation by the MNK pathway has also been implicated in circadian clock entrainment in mice (Cao et al., 2015). eIF4E can also be targeted for modification by small ubiquitin-like modifier (SUMO) conjugation, which enhances formation of the eIF4F complex and promotes cap-dependent translation of eIF4E-responsive genes (Xu et al., 2010a, b).

eIF4G is also a target of extensive phosphorylation, with as many as 30 phosphorylation sites identified in mammals (Hu et al., 2012). A specific residue (Ser^{1108}) has been identified as a site of phosphorylation regulated by nutrient availability in rat skeletal muscle, with increased phosphorylation correlating to higher eIF4F complex formation and protein synthesis (Bolster et al., 2004); Ser^{1108} phosphorylation is surprisingly under negative regulation from the eIF4G-binding kinase MNK2 (Hu, et al., 2012). Phosphorylation at other sites can also inhibit translation; p21-activated protein kinase 2 (Pak2) has been shown to bind and phosphorylate eIF4G and inhibit eIF4E binding (Ling et al., 2005; Sun et al., 2010). During mitosis, eIF4G phosphorylation by cyclin-dependent kinase 1 (CDK1):cyclin-B disrupts interaction with eIF4A to inhibit translation (Dobrikov et al., 2014). eIF4G has also been proposed to be a target of sumoylation in fission yeast and in humans, and though the effect of this modification is not yet clear it may be related to eIF4G subcellular localization (Jongjitwimol et al., 2014).

Translational repression of specific transcripts by miRNA is accomplished at least in part through targeting eIF4F, as let-7 miRNA has been shown to inhibit cap-dependent

translation (but not cap-independent translation) in an eIF4F level sensitive manner in mouse cell extract (Mathonnet et al., 2007). While the exact mechanism by which this occurs remains controversial, it has been reported that the RNA-induced silencing complex (RISC) member Argonaute1 (Ago1) promotes dissociation of eIF4A from eIF4G in *Drosophila* (Fukaya et al., 2014) and a similar observation was made concurrently in human cell lines (Fukao et al., 2014). Translational repression can also occur through stress-induced tRNA cleavage producing fragments called tiRNA which target eIF4F and promote dissociation from mRNA (Ivanov et al., 2011).

Along with the global mechanisms described above, eIF4G may also serve as a platform for translational regulation of specific genes or subpopulations of mRNAs. It has been observed in yeast that eIF4G contributes to differential translational efficiency of mRNAs genome-wide, as depletion of eIF4G causes a narrowing of the range of translational efficiency (Park et al., 2011a). Isoforms of eIF4G may contribute to gene-specific regulation, as in mammalian spermatogenesis, in which eIF4G3 is required to translate *HSPA2* in order to proceed through meiosis (Sun et al., 2010). A specific eIF4G isoform has also been implicated in *Drosophila* meiosis during spermatogenesis (Ghosh and Lasko, 2015).

Newly synthesized RNA is translationally repressed until after nuclear export and localization (Besse and Ephrussi, 2008; Halstead et al., 2015) and this protection and trafficking is dependent upon formation of mRNP complexes. In yeast, *ASH1* mRNA is translationally repressed during localization through the action of the RNA-binding protein Khd1p which has been proposed to bind to the C-terminal portion of eIF4G (Paquin et al., 2007). In human cells, ribosomal protein L13a, when phosphorylated in response to γ-interferon, dissociates from the 60S subunit and blocks translation of specific transcripts such as *CP* by binding to eIF4G (Kapasi et al., 2007). Translational repression of certain

mRNAs that are localized to the outer membrane of mitochondria is dependent on Pten-induced kinase 1 (PINK1) through which translation is then activated by the E3 ligase Parkin (Gehrke et al., 2015). In *Drosophila*, PINK1 binds to eIF4G which in turn contributes to this mechanism of regulation, an interesting finding given that eIF4G, like PINK1 and Parkin, has recently been linked to Parkinson's disease (Chartier-Harlin et al., 2011).

Another mechanism of translational control is through direct competition with the eIF4F complex. This can be through competition with eIF4E for the 7-methylguanosine cap structure, as has been observed with Pumilio 2 in *Xenopus* (Cao et al., 2006). It can also occur through competition with eIF4G for binding of RNA, as is the case with mammalian YB-1, which binds at the 5' end of transcripts near the cap structure and blocks association with eIF4F (Evdokimova et al., 2001; Nekrasov et al., 2003). Another approach is employed by the mammalian RNA-binding protein Musashi1, which binds to PABP in competition with eIF4G in order to inhibit translation of its target mRNAs (Kawahara et al., 2008).

The eIF4F components may play other, non-translational, roles that are not yet as clearly understood. For example, eIF4G in yeast has been shown to associate with spliceosomal components in the nucleus and may contribute to pre-mRNA processing events (Kafasla et al., 2009), and eIF4E undergoes nucleocytoplasmic shuttling and contributes to nuclear export of a subset of mRNA (Osborne and Borden, 2015). While there is more to learn about the roles of eIF4F, it remains the most well studied and understood mRNA binding platform for post-transcriptional gene regulation.

1.3. PLANT eIF4F AND eIFiso4F

Flowering plants, such as wheat and *Arabidopsis thaliana*, have a conserved eIF4F complex with eIF4G and eIF4E orthologous to cap-binding complexes found in other eukaryotes. However, they also have a distinct eIFiso4F complex which is specific to the flowering plant lineage, made up of eIFiso4G and eIFiso4E. It appears that genes encoding eIFiso4G and eIF4G are both present in nearly green algae and land plant genomes, and likely represent an ancient gene duplication and functional divergence event (Patrick and Browning, 2012). eIFiso4E, on the other hand, is the product of a more recent event and appears in the evolutionary record at the base of flowering plants (Patrick and Browning, 2012). The strong conservation of eIF4F and eIFiso4F genes in plants implies while the genes have overlapping predicted functions in translation initiation; however each may have important specific roles in gene regulation as well. In *A. thaliana*, as in other organisms, cap-binding proteins constitute essential genes; double knockouts of eIF4E and eIFiso4E are not viable (Callot and Gallois, 2014; Patrick et al., 2014), and a knockout of all genes encoding eIF4G/eIFiso4G isoforms has not been possible to obtain (Mayberry and Browning, unpublished observations). Surprisingly, however, knockouts of only eIF4E or eIFiso4E exhibit minor developmental impairment; this also the case for eIF4G and even eIF4G/eIF4E knockouts. Plants lacking eIFiso4G, however, show severe developmental defects including slow growth rate, low fertility, and lowered accumulation of chlorophyll (Lellis et al., 2010). In addition to eIF4E and eIFiso4E, *A. thaliana* encodes two divergent eIF4E-related proteins (discussed in Chapter 2) and an orthologue of metazoan 4EHP (discussed in Chapter 3).

The domain structure of eIF4G in plants is organized similarly to mammals (Figure 1.1), with a shared core structure of an eIF4E-binding site, the HEAT-1/MIF4G and

HEAT-2/MA3 domains which bind eIF4A, and a long N-terminus with little identified structure (Patrick and Browning, 2012). Plant eIF4G differs from mammalian eIF4G in that it lacks the C-terminal HEAT-3/W2 domain that is the site of MNK interaction. Plant eIFiso4G is similar in structure to eIF4G, but lacks the long N-terminus. Several evolutionarily conserved N-terminal motifs of unknown function have been identified in eIF4G and eIFiso4G (Patrick and Browning, 2012). The wheat eIF4E structure is similar to what has been described for mammals and yeast, with one significant exception: the presence of pair of conserved cysteines that have the potential to form a disulfide bond, the formation of which may alter the shape of the cap-binding pocket (Monzingo et al., 2007).

In plants, eIF4G and eIFiso4G bind their preferred partners eIF4E and eIFiso4E with high affinity and specificity in *vitro* and *in vivo*, although it is possible to form mixed complexes of eIF4G/eIFiso4E or eIFiso4G/eIFiso4E in the absence the preferred binding partner and these mixed complexes are functional *in vitro* (Mayberry et al., 2011). eIF4F shows greater ability than eIFiso4F to promote translation of structured mRNAs *in vitro*, and various mRNAs assayed have shown differential requirements for the two complexes in promoting translation; some show similar activity with either complex while others showing preference for eIF4F (Gallie and Browning, 2001; Mayberry et al., 2009).

eIF4F and eIFiso4F proteins are important to the life cycle of many plant viruses, and numerous resistance alleles in different plants have been found to map to constituent proteins (Moury et al., 2013; Robaglia and Caranta, 2006; Sanfacon, 2015; Wang and Krishnaswamy, 2012). The potyviral VPg protein has been shown to directly interact with eIF4F and eIFiso4F proteins (Moury et al., 2013; Robaglia and Caranta, 2006; Sanfacon, 2015; Wang and Krishnaswamy, 2012) and in *A. thaliana* the *lsp1* mutation which prevents potyvirus infection was found to map to nonsense mutation of the *EIFISO4E* gene (Lellis et al., 2002). Cucumber mosaic virus resistance was conferred by either nonsense mutation

of *EIF4E* (*cum1*, *cucumovirus multiplication 1*) or amino acid substitution in eIF4G (*cum2*) (Yoshii et al., 2004). The site of the *cum2* mutation is in an evolutionarily conserved region called H1-CT (HEAT-1 C-terminal) that may affect eIF4A or eIF3 binding (Patrick and Browning, 2012).

1.3.1. Regulation of Translation in Plants

Translational regulation at the level of initiation in plants is poorly characterized. Nothing resembling the 4E-BPs present in yeast and metazoans have been identified. Several other proteins with eIF4E-binding ability have been described, but there is no evidence of translational regulation by these proteins by competing with eIF4G or eIFiso4G (Freire, 2005; Freire et al., 2000). eIF4G is a target of phosphorylation, including differential phosphorylation in response to abscisic acid (Kline et al., 2010; Sugiyama et al., 2008) and photosynthetic activity (Boex-Fontvieille et al., 2013). In maize roots, eIF4E phosphorylation appears to be regulated by hypoxia (Manjunath et al., 1999); eIFiso4E has also been identified in the phosphoproteome of *Arabidopsis* suspension cell culture (Sugiyama et al., 2008). eIF4G has been identified proteomically as a target of SUMOylation under heat stress (Park et al., 2011b). The effect of these identified modifications remains unknown. An area research in the Browning Laboratory is the possibility of regulation of plant eIF4E (or eIFiso4E) via the redox state of the conserved cysteines that are able to form a disulfide bond (Monzingo et al., 2007).

While the mechanisms of regulation of translation in plants are largely unclear, it has been established that translational regulation is an important component of plant developmental processes and stress responses. Seed germination in *A. thaliana* is possible without new transcription, through a program of activating translation of stored mRNAs

(Rajjou et al., 2004). Widespread translational regulation, particularly of genes involved in chloroplast function, is observed during photomorphogenesis (Liu et al., 2013). Diurnal cycles of translation seem to be driven by sucrose levels, a general signal of energy and carbon availability (Pal et al., 2013); the circadian clock has also recently been described as contributing to diurnal translational patterns (Missra et al., 2015). In seedlings, unexpected darkness and re-illumination led to differential translational regulation of specific mRNAs (Juntawong and Bailey-Serres, 2012). Translational responses have been described for stresses including hypoxia (Branco-Price et al., 2008; Mustroph et al., 2009), dehydration (Kawaguchi et al., 2004), and heat stress (Yanguez et al., 2013).

A few plant mRNAs have been described in detail as having specific translational regulatory control. Upstream open reading frames (uORFs) have been identified as translational control mechanisms, such as in *A. thaliana bZIP11* mRNA where a uORF encodes a control peptide that confers translational repression to the main ORF in the presence of sucrose (Rahmani et al., 2009; Von Arnim et al., 2014). Translation of the circadian clock protein LHY is under light-dependent translational control (Kim et al., 2003). The mRNA-specific translational control mechanism described in the best detail in plants involves phytochrome-dependent regulation of *PORA* translation. The *PORA* mRNA is bound in its 5' UTR by an RNA-binding protein, PENTA1, which confers translational repression in a red-light dependent manner through binding of photosensory phytochrome proteins (Paik et al., 2012). While we speculate that eIF4F and/or eIFiso4F subunits are direct contributors to mRNA-specific translational regulation definitive evidence has remained elusive.

Toward the goal of better understanding gene regulatory mechanisms of plant cap-binding protein complexes, we investigated the biochemical properties of the putative eIF4E-related proteins eIF4E1b and eIF4E1c encoded in the *A. thaliana* genome (Chapter

2). Following the verification of eIF4E1b-like proteins as genuine cap-binding proteins, we characterized the role of the evolutionarily conserved 4EHP cap-binding protein with unexpected results implicating gene regulation through interactions with GIGYF and ncRNA in *Arabidopsis* (Chapter 3).

Chapter 2[1]

2.1. INTRODUCTION

Cap-dependent translation in eukaryotes begins with recognition of the 7-methylguanosine cap at the 5' end of an mRNA by the translation initiation factor eIF4E, which forms the eIF4F complex with the scaffolding protein eIF4G. The binding of the RNA helicase eIF4A along with eIF4B promotes unwinding of mRNA secondary structure (Aitken and Lorsch, 2012). The eIF4F complex then serves to circularize mRNA by interaction of eIF4G with poly(A) binding protein (PABP) and recruit the pre-initiation complex through binding of eIF4G to eIF3 and eIF5, ultimately leading to the assembly of the 80S ribosome (Aitken and Lorsch, 2012). eIF4E is an attractive target for global regulation of translational activity through its position at the earliest step in initiation, mRNA cap recognition. In many organisms, eIF4E availability is regulated by 4E-binding proteins (4E-BP) as well as phosphorylation and sumoylation (Jackson et al., 2010; Xu et al., 2010a, b). However, plants appear to lack 4E-BPs, and the role of phosphorylation of eIF4E in translational control is less clear (Browning and Bailey-Serres, 2015; Pierrat et al., 2007).

The eIF4E proteins generally thought to be involved in translation initiation are Class I eIF4E proteins (Joshi et al., 2005), of which two exist in flowering plants: eIF4E, which pairs with eIF4G to form the eIF4F complex, and the plant-specific isoform eIFiso4E, which pairs with eIFiso4G to form eIFiso4F (Mayberry et al., 2011; Patrick and

[1] Based on published work: Patrick, R.M., Mayberry, L.K., Choy, G., Woodard, L.E., Liu, J.S., White, A., Mullen, R.A., Tanavin, T.M., Latz, C.A., and Browning, K.S. (2014). Two *Arabidopsis thaliana* loci encode novel eIF4E isoforms that are functionally distinct from the conserved plant eIF4E. Plant Physiol 164, 1820-1830. Co-authors contributed to cloning of codon-optimized genes and to translation assays.

Browning, 2012). Class I eIF4E family members have conserved Trp residues at positions equivalent to Trp-43 and Trp-56 of *H. sapiens* eIF4E (Joshi et al., 2005), and the canonical members of this class, such as plant eIF4E and eIFiso4E, have the ability to promote translation through binding of mRNA cap structure and eIF4G (or eIFiso4G). In some organisms, secondary Class I isoforms exist with expression patterns and functions divergent from the conserved eIF4E (Rhoads, 2009). *Caenorhabditis elegans* has four isoforms involved in differentiation between mono- and trimethylated mRNA caps (Keiper et al., 2000) and have specialized roles for regulation of certain sets of mRNAs, particularly in the germline (Amiri et al., 2001; Song et al., 2010). *Trypanosoma brucei* has four isoforms with varying ability to bind cap analog and eIF4G isoforms (Freire et al., 2011). *Schizosaccharomyces pombe* has a second eIF4E isoform, eIF4E2, which is nonessential under normal growth conditions, but accumulates in response to high temperatures (Ptushkina et al., 2001). It cannot, however, complement deletion of *EIF4E1*, and while it can bind capped mRNA and promote translation *in vitro*, it has reduced ability to bind an eIF4G-derived peptide.

Vertebrates encode a novel Class I isoform called *EIF4E1B* with oocyte-specific expression and functions (Evsikov and Marin de, 2009). Zebrafish (*Danio rerio*) *EIF4E1B*, with expression limited to muscle and reproductive tissue, has conserved residues identified as necessary for binding cap analog and eIF4G, yet fails to bind either and cannot functionally complement deletion of yeast eIF4E (Robalino et al., 2004). In *Xenopus* oocytes, the eIF4E1b protein was found to bind eIF4E-transporter (4E-T) protein and cytoplasmic polyadenylation element binding protein (CPEB) to form a translation repressing complex (Minshall et al., 2007). *Drosophila* species have undergone extensive expansion of *EIF4E*-encoding loci to as many as seven different Class I eIF4E isoforms (Tettweiler et al., 2012). The seven *EIF4E* isoforms of *D. melanogaster* are differentially

expressed, with only five able to bind to eIF4G and complement deletion of yeast eIF4E (Hernandez et al., 2005). The eIF4E-3 isoform of *D. melanogaster* was recently described as having a specific role in spermatogenesis (Hernandez, 2012).

Upon completion of sequencing of the *Arabidopsis thaliana* genome (Rhee et al., 2003), it was discovered that in addition to the conserved plant *EIF4E* (*At4g18040*) and *EIFISO4E* (*At5g35620*), there existed a tandem pair of genes of high sequence similarity on chromosome 1 which also encoded two Class I eIF4E proteins, *EIF4E1B* (*At1g29550,* also known as *EIF4E3*) and *EIF4E1C* (*At1g29590,* also known as *EIF4E2*). Published microarray and RNA-Seq data indicate little to no *EIF4E1C* gene expression; however, the *EIF4E1B* gene appears to be expressed at low levels in most tissues and enriched in tissues involved in reproduction. The protein sequences contain the residues predicted to be involved in regular eIF4E function, but also showed some divergence at highly conserved residues of the canonical plant eIF4E. Genome sequencing data indicate that these genes are part of a divergent eIF4E clade specific to Brassicaceae.

The biochemical properties of the eIF4E1b and eIF4E1c proteins were investigated in this work, and it was found that while they can bind mRNA cap analog, bind eIF4G, and support translation in yeast lacking eIF4E, their eIF4G-binding and translation initiation enhancing capabilities *in vitro* were less robust when compared to the conserved *Arabidopsis* eIF4E. In addition, it appears that these *EIF4E1B*-type genes cannot substitute for *EIF4E* or *EIFISO4E in planta* since deletion of both of these genes appears to be lethal. Taken together, these findings indicate the *EIF4E1B*-type genes represent a divergent eIF4E whose specialized roles should be considered separately from the canonical eIF4E in plant translation initiation.

2.2. METHODS

2.2.1. In silico Analysis

eIF4E and eIFiso4E protein sequences for alignment and analysis were collected from Phytozome (Goodstein et al., 2012) and BLAST searches to GenBank sequences (Benson et al., 2012). Alignment was performed by ClustalW2 (Larkin et al., 2007). Residues defined as conserved in eIF4E were those with 90% or greater identity in the canonical coding from among the aligned sequences. The Phylogeny.fr pipeline (Dereeper et al., 2008) was used for alignment and phylogenetic tree generation with alignment by MUSCLE and tree construction by PhyML using 500 bootstrap replicates.

2.2.2. eIF4E and eIFiso4E Cross

Mutant lines for eIF4E (At4g18040; *cum1*, a nonsense point mutation) and eIFiso4E (AT5g35620; *iso4E-1*, Sainsbury Laboratory's SLAT library) have been previously described (Duprat et al., 2002; Yoshii et al., 2004). Crosses between these two lines were performed in both directions and the T2 progeny screened by PCR to identify wild type, heterozygous, or double homozygous lines. *iso4E-1* and *cum1* lines were screened with primers as described in Table 2.4.

2.2.3. Construction of eIF4E1, eIF4E1b, eIF4E1c, eIF4E1c(long), eIF4G$_{322-1727}$, eIF4G$_{1-1727}$, and eIF4F, eIF4F1b, eIF4F1c Expression Constructs

Initial attempts to express *Arabidopsis* eIF4G protein from cDNA clones were unsuccessful. Using DNAWorks (Hoover and Lubkowski, 2002), *A. thaliana* eIF4G$_{322-}$

$_{1727}$, eIF4E1, eIF4E1b, eIF4E1c and eIF4E1c(long) were designed with codon optimization for expression in *Escherichia coli* and assembled by overlap PCR of oligonucleotides (Horton et al., 1989). Initial cloning of eIF4G$_{322-1727}$ was into pCR-Blunt-II-TOPO (Invitrogen) followed by subcloning into pSB1AC3 (Shetty et al., 2008) and pET22b vectors (Novagen). eIF4E genes were cloned in one step. eIF4G$_{322-1727}$ was cloned into pCR-Blunt-II-TOPO in four sections and then assembled in pSB1AC3. Full length eIF4G$_{1-1727}$ was created by cloning a synthetic DNA sequence (GenScript) to provide the missing N-terminal sequence to eIF4G$_{322-1727}$; the restriction site used to ligate the synthetic DNA was then altered to match wild-type protein sequence by site-directed mutagenesis (Mutagenex). The pET22b eIF4G$_{322-1727}$ vector was used to clone eIF4E1, eIF4E1b, and eIF4E1c genes at a site 3' of the eIF4G coding region to create dicistronic plasmids for expression of eIF4F, eIF4F1b, and eIF4F1c complexes.

2.2.4. Purification of Recombinant Proteins

eIF4E proteins were expressed in BL21(DE3) *E. coli*, and purified as previously described by m^{7}GTP-Sepharose affinity chromatography (Mayberry et al., 2007). eIF4F complexes were expressed in Tuner(DE3) *E. coli* (Novagen), and purified as previously described for wheat eIFiso4F (Mayberry et al., 2011; Mayberry et al., 2007). eIF4G$_{322-1727}$ and eIF4G$_{1-1727}$ were expressed in Tuner(DE3) *E. coli*, and purified as previously described for wheat eIF4G (Mayberry et al., 2011; Mayberry et al., 2007). Recombinant wheat eIF4A and native wheat eIF3 were purified as previously described (Lax et al., 1986; Mayberry et al., 2007).

2.2.5 In vitro Assay

A. thaliana eIF4G and cap-binding proteins were assayed in an *in vitro* translation assay using wheat germ S30 extract that had been depleted of cap-binding proteins and complexes. Three 4 mL portions of m^7GTP-Sepharose (GE Biosciences) were equilibrated in 20 mM Hepes, pH 7.6, 120 mM KAc, 5 mM $MgAc_2$, 10% glycerol, and 7.15 mM β-mercaptoethanol. A 2 mL aliquot of S30 extract was used to exchange the buffer from each of the three 4 mL portions of the m^7GTP-Sepharose. 25 mL of wheat germ S30 extract (Lax et al., 1986; Mayberry and Browning, 2006) was mixed for 15 min. with 4 mL of m^7GTP-Sepharose by rocking on ice; the supernatant was collected, and the process was repeated with the remaining two portions of 4 mL of m^7GTP-Sepharose. The eIF4F/eIFiso4F depleted S30 extract was aliquoted, flash frozen and stored at -80 °C.

The 50 µl translation assay reaction mixture contained 24 mM Hepes-KOH, pH 7.6, 2 mM $MgAc_2$, 100 mM KAc, 30 mM KCl, 2.4 mM DTT, 0.1 mM spermine, 1 mM ATP, 0.2 mM GTP, 34 µM [^{14}C]leucine, 50 µM 19 amino acids (-leucine), 7.8 µM creatine phosphate, 3 µg creatine kinase, 0.75 A_{260} units of yeast tRNA, 15 µl of depleted S30 extract, 4-5 pmol barley α-amylase mRNA, 10 µg of recombinant wheat eIF4A, 0.5 µg of recombinant wheat eIF4B and 6 µg of native wheat eIF3 and the indicated amounts of eIF4F, eIF4G and/or cap-binding proteins. Incubation was for 30 min at 27 °C and the amount of [^{14}C]leucine incorporated into protein was determined as previously described (Lax et al., 1986; Mayberry and Browning, 2006; Mayberry et al., 2007).

2.2.5. SPR Analysis

Surface Plasmon Resonance (Biacore) experiments were carried out as described previously (Mayberry et al., 2011) at Biosensor Tools LLC (Salt Lake City, Utah) by Dr. David Myszka. Briefly, protein binding was measured at 25°C using a Biacore 2000 optical biosensor equipped with a CM4 sensor chip in running buffer (20 mM HEPES, 100 mM KCl, 1.5 mM TCEP, 0.1 mM EDTA, 100 μM m^7GTP, 5% glycerol, 0.01% Tween-20, 0.1 mg/mL BSA, pH 7.6). $eIF4G_{322-1727}$ was amine-coupled at three surface densities (500, 1370, and 4430 Resonance Units [RU]). eIF4E1, eIF4E1b, and eIF4E1c proteins were tested for eIF4G binding in three-fold dilution series performed in triplicate. For eIF4E1b and eIF4E1c the highest concentration tested was 1.5 μM and for eIF4E it was 66.7 nM. Response data for each protein were fit globally to a 1:1 interaction model using Scrubber2 (Biologic Software Pty Ltd) across the three eIF4G surface densities.

2.2.6. Yeast Complementation of eIF4E

The *Saccharomyces cerevisiae* strain T93C (Altmann et al., 1989), containing a chromosomal deletion of eIF4E and a plasmid with eIF4E under control of a Gal promoter, (*eIF4E::LEU2 ura3 trp1 leu2* [pGal1-eIF4E URA3]), was transformed with pG-1 vectors (with an added NcoI site N-terminal to BamHI in the cloning region) containing *A. thaliana* eIF4E constructs. pG-1 provides constitutive gene expression and a TRP1 marker (Schena et al., 1991). Positive transformants were verified by plasmid re-isolation and sequencing.

Yeast strains were grown overnight, then diluted to 0.1 OD, grown to mid-log phase (~0.3 OD) and plated in serial dilutions on Synthetic Complete Medium (SCM) -Trp + 2% Galactose and SCM –Trp + 2% Glucose. Plates were incubated at 30°C for 48 hours and

screened for ability of *A. thaliana* eIF4E proteins to rescue yeast growth. This experiment was performed in triplicate.

2.3. RESULTS

2.3.1. In silico Analysis

BLAST searches of available genomic and EST (Expressed Sequence Tag) data using NCBI and Phytozome (Benson et al., 2012; Goodstein et al., 2012) find that *EIF4E1B*-type genes are present in close *Arabidopsis* relatives including *Capsella rubella*, *Brassica* and *Raphanus* species, and *Thellungiella* species. However, there is no evidence of these genes outside of Brassicaceae, including the closest relative species sequenced, *Carica papaya*. It therefore appears that the *EIF4E1B*-type genes are the result of a Brassicaceae-specific gene duplication and specialization. The genomes of *Eucalyptus grandis* and *Fragaria vesca* also encode predicted divergent eIF4E protein forms (*EIF4E1BL* genes in Figure 2.1), though it remains to be determined whether these genes are expressed or conserved in other related species.

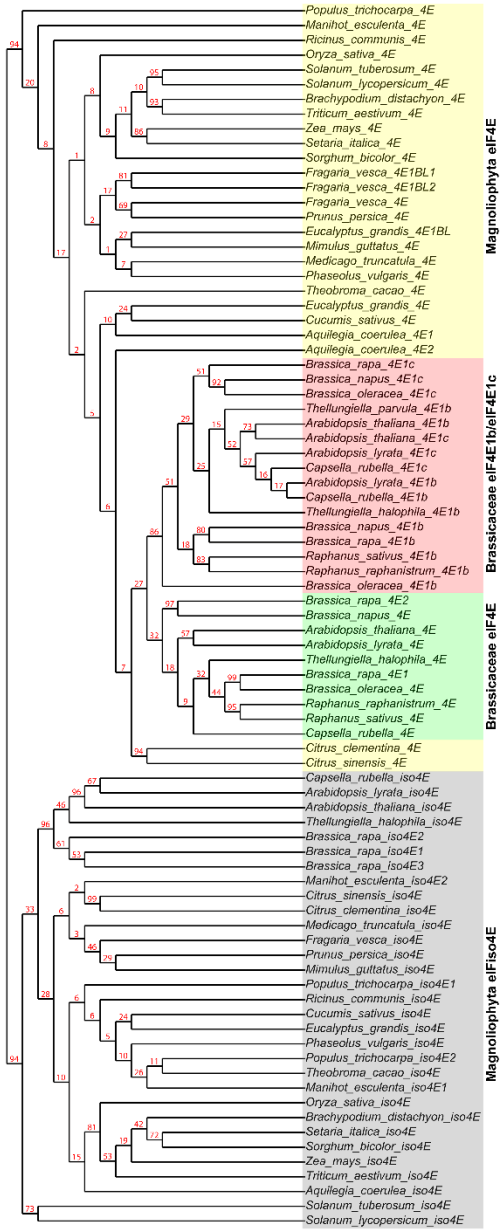


Figure 2.1. **Cladogram of Brassicaceae eIF4E1b-like proteins in relation to the conserved eIF4E and eIFiso4E proteins of flowering plants.**

The Phylogeny.fr pipeline (Dereeper et al., 2008) was used for alignment and phylogenetic tree generation with alignment by MUSCLE and tree construction by PhyML using 500 bootstrap replicates.

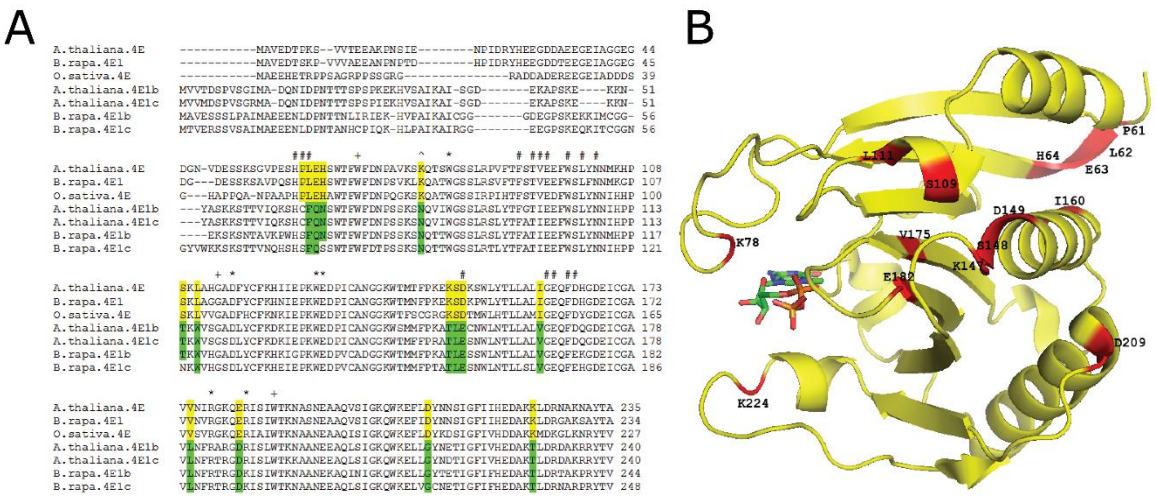


Figure 2.2. **Alignment of Brassicaceae eIF4E**

(A) Representative flowering plant eIF4E and Brassica-specific eIF4E1b/c sequences aligned by ClustalW2 (See Figure 2.7 for full alignment). Residues highlighted in yellow have 90% or greater identity in conserved flowering plant eIF4E sequences but are consistently altered in eIF4E1b-type sequences, residues highlighted in green are conserved divergent residues at these locations in eIF4E1b-type sequences. Residues marked (#) are conserved residues involved in interaction with eIF4G identified in yeast (Gross et al., 2003). Residues marked (+) are experimentally determined to be essential for cap-binding (German-Retana et al., 2008; Yeam et al., 2007). Residues marked (*) are predicted to be involved in cap-binding in plant eIF4E crystal structures (Ashby et al., 2011; Monzingo et al., 2007) and the residue marked (^) is predicted from the wheat eIF4E structure to form a salt bridge with the negatively charged phosphate backbone of the cap. **(B)** Diverging residues in *Arabidopsis* eIF4E1b/c were modeled with PyMOL (DeLano, 2002) on the wheat eIF4E structure (Monzingo et al., 2007) as indicated.

The genomes of *A. thaliana*, *A. lyrata*, *C. rubella*, and *B. rapa* encode two *EIF4E1B*-type loci, called *EIF4E1B* and *EIF4E1C*, while *T. halophila* and *T. parvula* only have evidence for one copy of the gene. Alignment and phylogenetic tree construction of

eIF4E and eIFiso4E sequences (Figure 2.1) shows that the eIF4E1b-type protein sequences cluster together, separately from the conserved eIF4E of flowering plants and from eIFiso4E, which diverged from eIF4E early in the flowering plant lineage (Patrick and Browning, 2012). In addition to completed and draft genomes, there is EST evidence of *EIF4E1B* and/or *EIF4E1C* in *B. oleracea*, *B. napus*, *R. raphanistrum*, and *R. sativus*, as well as GSS (Genome Survey Sequence) support for the presence of an *EIF4E1B*-type gene in *Sisymbrium irio* (Figure 2.7).

Interestingly, the *EIF4E1B* and *EIF4E1C* genes of *Arabidopsis* and *Capsella* are more closely conserved at the sequence level to each other than to *EIF4E1B* and *EIF4E1C* of *Brassica* (Figure 2.1), indicating that there has been recent duplication of *EIF4E1B* separately in each lineage. This is supported by the fact that while *Arabidopsis* and *Capsella* have *EIF4E1B* and *EIF4E1C* as a tandem duplication on one chromosome, *B. rapa* has *EIF4E1B* and *EIF4E1C* genes on separate chromosomes. *Thellungiella* species, meanwhile, have only one copy of the gene.

Fifteen residues within the protein have been identified as 90% conserved in flowering plant canonical eIF4E while consistently altered in eIF4E1b-type proteins; many of these residues are conserved as a specific divergent amino acid in eIF4E1b and eIF4E1c (Figure 2.2, Figure 2.7). Residues that have been identified as being involved in cap binding from crystal structures of *Triticum aesestivum* (wheat) (Monzingo et al., 2007) and *Pisum sativum* (pea) eIF4E (Ashby et al., 2011) or by mutational analysis (German-Retana et al., 2008; Yeam et al., 2007) are well conserved in eIF4E1b-type proteins. One exception is the conserved positively charged residue at K78, which is predicted in the wheat eIF4E crystal structure to stabilize the negatively charged phosphate backbone of the cap structure. In eIF4E1b/c-type proteins, this residue is changed to asparagine which may

weaken this interaction. However, mutation of this position in *P. sativum* eIF4E had no effect on the ability to promote translation in yeast (Ashby et al., 2011).

Several mutations in eIF4E1b-type proteins occur at locations that are both well conserved between eIF4E of plants and mammals and predicted to be involved in eIF4E binding to eIF4G from a co-crystal structure of *S. cerevisiae* eIF4E with a fragment of eIF4G (Gross et al., 2003). eIF4E residues P61, L62, and D149 (Figure 2.2) are all predicted to be part of the eIF4G binding interface and are altered in eIF4E1b-type proteins. These changes appear to have altered the ability of eIF4E1b/c to interact with eIF4G compared to eIF4E (see below). While mutations in eIF4E that confer viral resistance are naturally occurring (Robaglia and Caranta, 2006) and directed mutagenesis has further identified residues conferring virus resistance (Ashby et al., 2011; German-Retana et al., 2008), the fifteen conserved flowering plant eIF4E residues differing in eIF4E1b-type proteins do not overlap with these residues, with the exception of a K78 eIF4E mutation, which confers virus resistance in *P. sativum* (Ashby et al., 2011). Interestingly, T-DNA insertion mutants for *EIF4E1B* or *EIF4E1C* do not have any effect on turnip mosaic virus infection in *Arabidopsis* (Gallois et al., 2010). To date, neither *EIF4E1B* nor *EIF4E1C* has been reported to be a virus resistance gene in contrast to numerous reports of virus resistance attributed to *EIF4E* and *EIFISO4E* alleles (Wang and Krishnaswamy, 2012).

Analysis of homozygous T-DNA insertion lines for *EIF4E1B* (GK-874C07) or *EIF4E1C* (GK-361E12) do not show an obvious phenotype. Due to their close proximity on chromosome 1, it has not been possible to obtain a double mutant even after screening >5000 plants.

2.3.2. *EIF4E1B/C* expression

Due to their sequence similarity, *EIF4E1B* and *EIF4E1C* share a spot on many commonly used microarrays, limiting data as to whether either or both are expressed. However, RNA-Seq data can distinguish between the two genes and indicate that in flower tissue (Jiao and Meyerowitz, 2010; Niederhuth et al., 2013), shoot apical meristem (Torti et al., 2012), developing embryos (Nodine and Bartel, 2012), and the central cell of the female gamete (Schmid et al., 2012), *EIF4E1B* mRNA is expressed and associates with polysomes in flowers (Jiao and Meyerowitz, 2010); however, *EIF4E1C* mRNA was at much lower to undetectable levels in these tissues. In an analysis of 80 genomes released by the 1001 Genomes Project, *EIF4E1C* was predicted to be spontaneously deleted in 12 strains suggesting that it is likely not providing any advantage to promote its retention in the genome (Cao et al., 2011).

Microarray data from the Arabidopsis eFP Browser (Winter et al., 2007) suggest that *EIF4E1B* is most highly expressed in developing flowers, while Genevestigator (Zimmermann et al., 2004) supports expression in shoots and reproductive tissue. *EIF4E1B* was identified as a gene up-regulated during pollen tube growth (Wang et al., 2008). Additionally, *EIF4E1B* was significantly enriched in pollen tubes grown by a semi-*in vivo* method (Qin et al., 2009). *EIF4E1B* was identified as a sperm-enriched gene, while *EIF4E* was sperm-depleted (Borges et al., 2008). In a microarray experiment investigating developing embryos, *EIF4E1B* was found to be expressed at high levels at the zygote stage of development relative to *EIF4E*, while *EIF4E1C* levels were near background levels (Xiang et al., 2011). Taken together, these findings may indicate a role for *EIF4E1B* in reproduction in Brassicaceae similar to that in *Drosophila* or zebrafish.

2.3.3. eIFiso4E or eIF4E is Required for Viability

The T-DNA knockout line for *EIFISO4E* (*iso4e-1*, Duprat et al., 2002) and the nonsense mutant for *EIF4E* (*cum1*, Yoshii et al., 2004) are viable and do not exhibit major developmental phenotypes individually. The *cum1* mutation does have smaller flowers and a slight delay in flowering time. However, extensive attempts have been made by this laboratory to isolate an *iso4e-1*/*cum1* double mutant without success. *iso4e-1* plants heterozygous for *cum1* are viable, but self-fertilized plants do not yield viable double mutants in the ratio expected for normal progeny (Table 2.1). The defect appears to be embryo-lethal as nearly all planted seeds germinated and were successfully screened. Similarly, *cum1* plants heterozygous for *iso4E-1* do not yield viable double mutants (Table 2.3). A male gametophyte lethal phenotype has also been associated with the *iso4e/4e1* genotype, preventing the recovery of a double homozygous mutant (Callot and Gallois, 2014). These results suggest that *EIF4E1B* and *EIF4E1C* gene products are not sufficient to fulfill the necessary role for a canonical Class I eIF4E protein, either due to low or localized expression, or loss of properties that contribute to translation initiation such as binding eIF4G and/or mRNA cap structure.

Since *EIF4E1B* and *EIF4E1C* genes are products of a Brassicaceae-specific gene duplication, they may have specialized functions and/or lost the ability to function as cap-binding proteins. We sought to determine if eIF4E1b and eIF4E1c proteins are biochemically capable of performing the functions of cap-binding proteins *in vitro,* as well as *in vivo* using a yeast complementation system.

Table 2.1. **Screening of *iso4e-1 cum1*/*EIF4E* progeny from self-fertilization.** 107 seeds were planted on MS agar plates, with 105 germinating and 103 successfully transplanted and screened. Recovery of *iso4e-1 cum1*/*EIF4E* was lower than expected (35%) and double homozygous mutant plants were not recovered.

	Number with Genotype	% of Total	Expected Mendelian[2] %
iso4e-1 cum1/*cum1*	0/107	0%	25%
iso4e-1 cum1/*EIF4E*	37/107	35%	50%
iso4e-1 EIF4E/*EIF4E*	66/107	62%	25%
not screened[1]	4/107	4%	NA

[1]Seed did not germinate or died before screening.

[2]Expected amounts if normal Mendelian inheritance.

2.3.4. eIF4E1b and eIF4E1c Bind to m^7GTP and Form Complexes with eIF4G

A. thaliana eIF4E, eIF4E1b, and eIF4E1c were expressed in *E. coli* and purified by affinity chromatography using m^7GTP-Sepharose (Figure 2.3A). Both eIF4E1b and eIF4E1c were found to bind and elute from m^7GTP-Sepharose in comparable yields to eIF4E (data not shown). eIF4E1b and eIF4E1c therefore seem to have biologically relevant cap-binding ability.

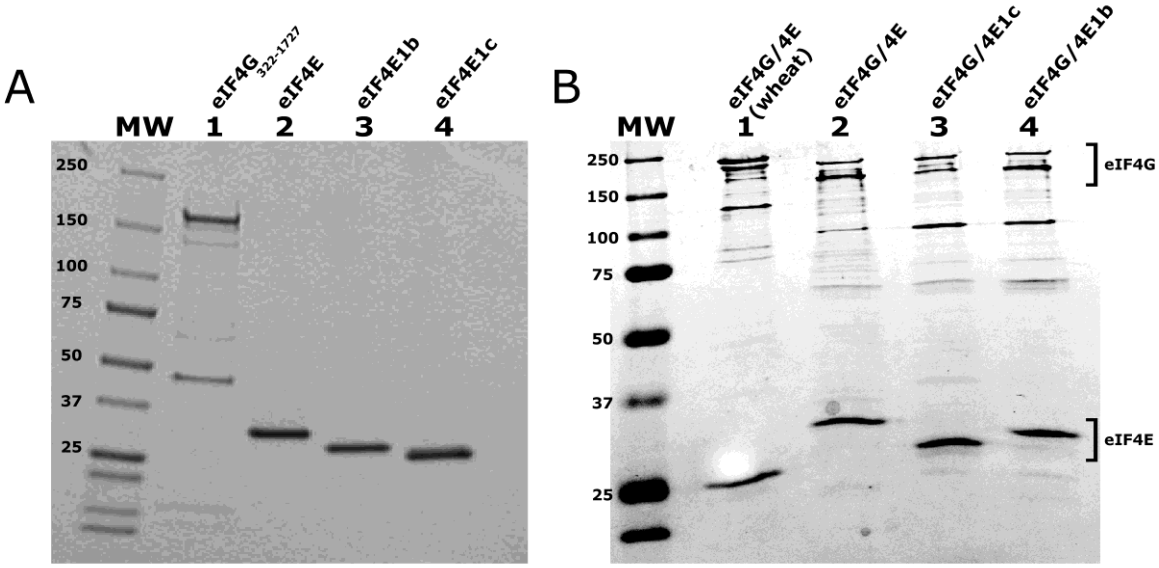


Figure 2.3. PAGE Analysis of Purified Proteins.

(A) Purified *A. thaliana* eIF4G and eIF4E isoforms. Lane 1, $eIF4G_{322-1727}$ (3 µg), Lane 2, eIF4E (1.5 µg); Lane 3 eIF4E1b (1.5 µg); Lane 4 eIF4E1c (1.5 µg) were separated by 12.5% SDS-PAGE and stained with Coomassie brilliant blue. **(B)** Recombinant *A. thaliana* eIF4F complexes from dicistronic constructs were expressed in *E. coli* and purified by m^7GTP-Sepharose affinity and phosphocellulose chromatography. Lane 1, wheat eIF4F (3 µg); Lane 2. *A. thaliana* $eIF4G_{1-1727}$ /eIF4E (3 µg); Lane 3, $eIF4G_{1-1727}$ /eIF4E1c (3 µg); Lane 4, $eIF4G_{1-1727}$ /eIF4E1b (3 µg). Proteins were separated by SDS-PAGE and stained with Coomassie brilliant blue.

The initial selection of the coding sequence of *Arabidopsis* eIF4G (*$eIF4G_{322-1727}$)* to express was made based on protein similarity to the N-terminus of wheat eIF4G (Mayberry et al., 2007). However, subsequent peptide sequence data in the pep2pro database (http://fgcz-pep2pro.uzh.ch/index.php) suggested that upstream initiation codons are utilized. Since the precise initiation codon or if there are multiple start sites, as occurs for mammalian

eIF4G (Coldwell et al., 2012), is not known for *Arabidopsis*, the codon selected by The Arabidopsis Information Resource (TAIR) was used to generate another expression construct for $eIF4G_{1-1727}$ which includes all the peptides identified in the pep2pro database.

eIF4E, eIF4E1b, and eIF4E1c were co-expressed with *A. thaliana* $eIF4G_{322-1727}$, which retains the eIF4E binding site, to form eIF4F complexes which were purified by m^7GTP-Sepharose affinity chromatography (Figure 2.3B). eIF4E1b co-purified with $eIF4G_{322-1727}$ to form the eIF4F1b complex, and eIF4E1c co-purified with $eIF4G_{322-1727}$ to form the eIF4F1c complex. Both eIF4F1b and eIF4F1c purified with comparable yield to eIF4F, confirming that eIF4E1b and eIF4E1c were able to bind eIF4G and form a stable complex.

Table 2.2. SPR Analysis. The binding affinity of purified *A. thaliana* cap-binding proteins to $eIF4G_{322-1727}$. Affinities were measured by surface plasmon resonance (SPR) using a Biacore 2000 with CM4 sensor chip. $eIF4G_{322-1727}$ was amine-coupled at three surface densities (500, 1370, and 4430 resonance units), with eIF4E1, eIF4E1b, and eIF4E1c proteins tested for eIF4G binding in three-fold dilution series performed in triplicate (for eIF4E1b and eIF4E1c the highest concentration tested was 1.5 μM and for eIF4E it was 66.7 nM). Response data for each protein was fit globally to a 1:1 interaction model using Scrubber2 (Biologic Software Pty Ltd) across the three eIF4G surface densities.

	K_D (nM)
eIF4E1	0.275 ± 0.002
eIF4E1b	451 ± 2
eIF4E1c	970 ± 10

Purified eIF4E, eIF4E1b, and eIF4E1c were assayed by surface plasmon resonance (SPR) for their binding affinity to purified $eIF4G_{322-1727}$ (Table 2.2). The K_D for eIF4E binding to $eIF4G_{322-1727}$ was extremely tight, at 0.275 ± 0.002 nM. This finding is consistent with the measurement of the wheat eIF4G and eIF4E binding K_D of 0.181 ± 0.002 nM (Mayberry et al., 2011). Surprisingly, eIF4E1b binding to $eIF4G_{322-1727}$ was 1640-fold weaker than eIF4E (451 ± 2 nM), while eIF4E1c binding was weaker still (970 ± 10 nM).

The eIF4E1b/eIF4E1c binding affinity to $eIF4G_{322-1727}$ is lower than was observed in the wheat system for a mixed complex of eIFiso4E binding to eIF4G (14.3 ± 0.2 nM) (Mayberry et al., 2011). It was previously shown that eIFiso4E was displaced from a mixed complex with wheat eIF4G by eIF4E, thus the correct binding partner is selectively favored (Mayberry et al., 2011). Based on the observed lower binding affinity, despite the ability of eIF4E1b or eIF4E1c to co-purify with $eIF4G_{322-1727}$ *in vitro*, it is unlikely either could form a complex with eIF4G *in vivo* in *Arabidopsis* unless eIF4E is absent (see Figure 2.5 below).

2.3.5. eIF4E1b and eIF4E1c Have Translation-Enhancing Activity *in vitro* but are Displaced by eIF4E

To measure the ability of eIF4E1b and eIF4E1c to function in the initiation of translation, *in vitro* translation assays in wheat germ S30 depleted of cap-binding complexes were carried out. Recombinant *Arabidopsis* $eIF4G_{1-1727}$ was mixed with equimolar amounts of cap-binding proteins to form eIF4F complexes, and these were tested for their ability to translate mRNA compared to recombinant wheat eIF4F (Figure 2.4). *Arabidopsis* eIF4G alone provided little stimulation of translation, while the conserved *Arabidopsis* eIF4F complex of eIF4G with eIF4E performed similarly to wheat eIF4F.

eIF4G paired with either eIFE1b or eIF4E1c showed similar activity, but required significantly higher concentrations of the complexes (~5-10 fold) to approach the extent of stimulation of the eIF4F complex from either wheat or *Arabidopsis.*

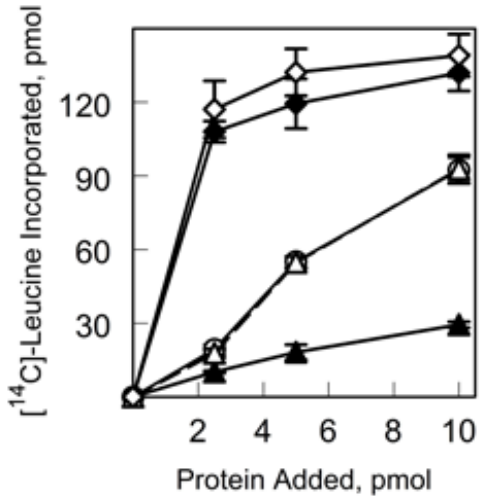


Figure 2.4. ***In vitro*** **Assay of eIF4E Isoforms with eIF4G.**

Equal molar amounts of At eIF4G$_{1-1727}$ and the respective eIF4E isoform were mixed and then added as indicated. The reaction mixture contained 5 pmol of barley α-amylase mRNA and 15 µl of a wheat germ S30 extract that had been depleted of eIF4F and eIFiso4F by passage over a m^{7}GTP Sepharose column as described in *Materials and Methods*. Wheat eIF4F (◊); At eIF4G/eIF4E (♦); At eIF4G/eIF4E1b (○); At eIF4G/eIF4E1c (Δ); At eIF4G (▲). Note that curves for eIF4G/eIF4E1b and eIF4G/eIF4E1c overlap. Experiments were done in triplicate and averaged. The incorporation of [^{14}C]Leucine in the absence of any added factor was 5 pmol.

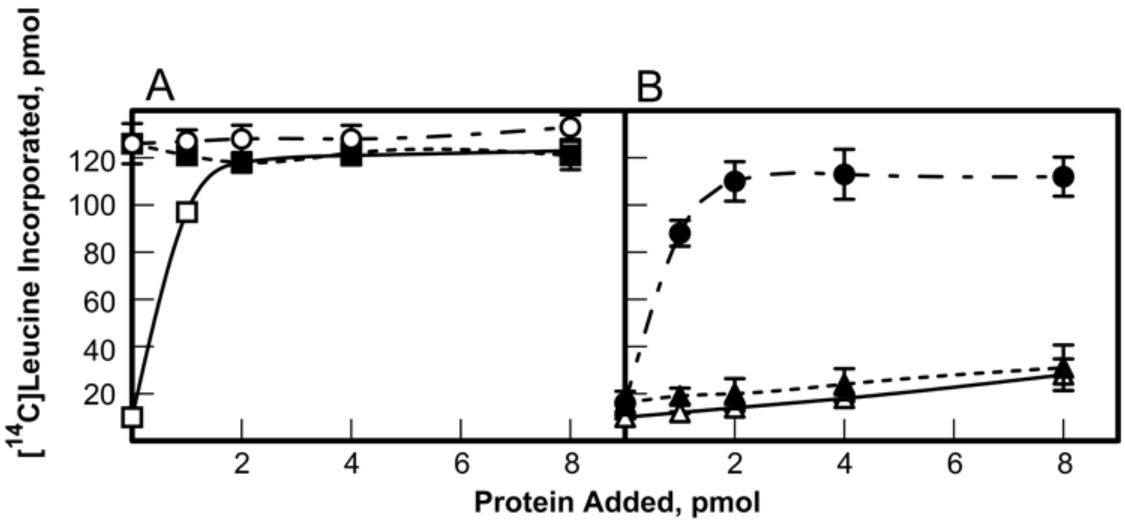


Figure 2.5. Displacement of eIF4E1b from Complex with eIF4G by eIF4E as Measured by *in vitro* Translation Activity.

The complex of eIF4G/eIF4E was presented with increasing amounts of eIF4E1b to determine if activity was reduced to the level of the eIF4G/eIF4E1b complex (panel A). Alternatively, a mixed complex of eIF4G/eIF4E1b was presented with increasing amounts of eIF4E to determine if eIF4E1b could be displaced by eIF4E and form the more active eIF4G/eIF4E complex (panel B). The reaction mixture contained 4 pmol of barley α-amylase mRNA and 15 µl of a wheat germ S30 extract that had been depleted of eIF4F and eIFiso4F by passage over a m^7GTP Sepharose column as described in *Materials and Methods*. **(A).** 2 pmol eIF4G titrated with eIF4E (□); 2 pmol eIF4G/eIF4E titrated with additional eIF4E as indicated (■); 2 pmol eIF4G/eIF4E titrated with additional eIF4E1b as indicated (○). **(B).** 2 pmol eIF4G titrated with eIF4E1b as indicated (Δ); 2 pmol eIF4G/eIF4E1b titrated with additional eIF4E1b as indicated (▲); 2 pmol eIF4G/eIF4E1b titrated with additional eIF4E as indicated (●). Experiments were done in triplicate and averaged. The incorporation of [^{14}C]Leucine in the absence of any added factor was 10 pmol.

The contribution of eIF4E1b to *in vitro* translation was further examined with an assay placing eIF4E1b in competition with eIF4E for eIF4G to observe changes in activity

of *in vitro* translation (Figure 2.5). The complex of eIF4E/eIF4G was challenged with either additional eIF4E or eIF4E1b and there was no significant change in the translational activity observed in either case (Figure 2.5A); however, a complex of eIF4E1b/eIF4G when challenged with eIF4E clearly led to displacement of eIF4E1b to form the more stable and more translationally active complex of eIF4E/eIF4G (Figure 2.5B). These results support the observed binding constants (Table 2.2) and suggest that eIF4E must be absent for any complex to form *in vivo* between eIF4G and eIF4E1b/c. It is also evident that at least *in vitro* eIF4E1b does not act as a general translational repressor.

2.3.6. eIF4E1b and eIF4E1c Complement eIF4E Deletion in Yeast

A. thaliana eIF4E has previously been shown to complement for the deletion of eIF4E in *S. cerevisiae* (Rodriguez et al., 1998). The *in vitro* data suggest that eIF4E1b and eIF4E1c are functional in that they bind to m^7GTP-Sepharose and eIF4G. To further investigate the ability of eIF4E1b and eIF4E1c to function as *bona fide* cap-binding proteins, they were tested for their ability to complement eIF4E deletion in yeast. As shown in Figure 2.6, both *EIF4E1B* and *EIF4E1C* are able to substitute for yeast *EIF4E* gene deletion *in vivo*. This finding implies that eIF4E1b and eIF4E1c have biologically relevant ability to promote translation and, as shown in the *in vitro* experiments, retain sufficient cap-binding and eIF4G-binding properties in spite of their sequence differences from eIF4E. Thus their inability to function *in planta* in a background lacking *EIF4E* and *EIFISO4E* suggests that their expression is highly localized or controlled.

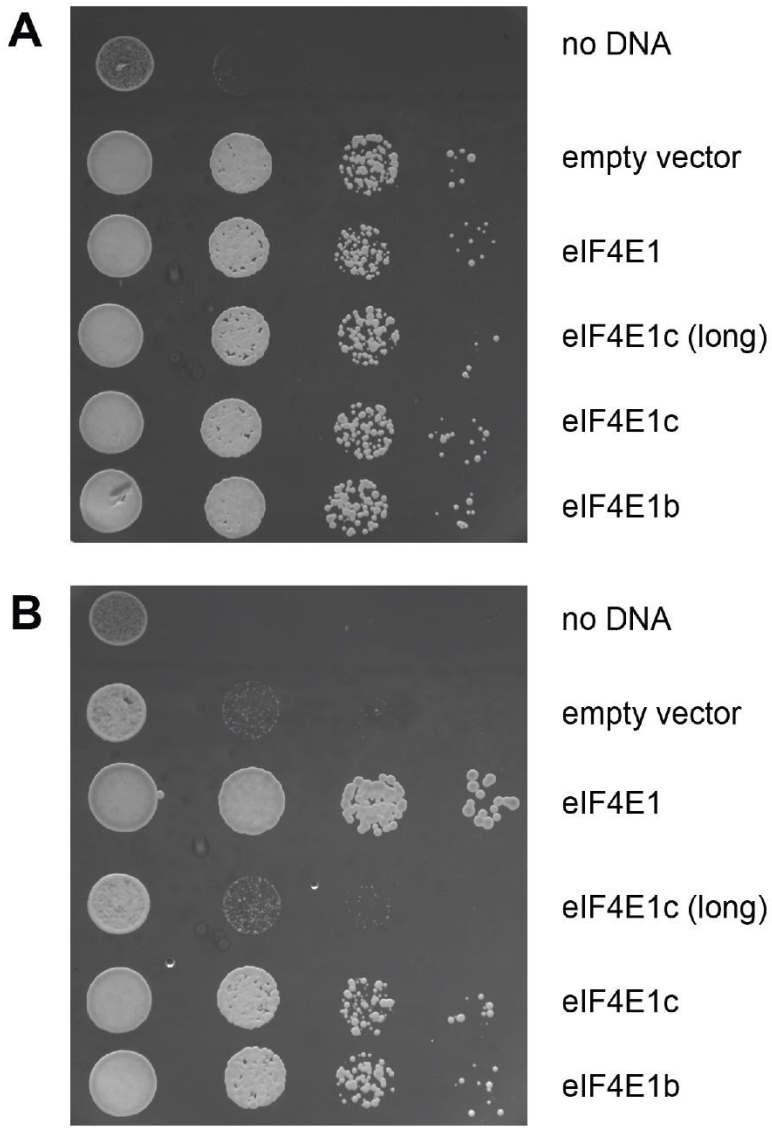


Figure 2.6. ***A. thaliana*** **eIF4E Proteins Complement eIF4E gene deletion in *S. cerevisiae.***

Complementation was tested by introducing pG-1 plasmids for constitutive expression of *A. thaliana* eIF4E genes into a yeast strain (T93C, Altmann et al., 1989) with eIF4E under control of a GAL promoter. Serial dilutions of mid-log phase yeast were plated in 10-fold serial dilutions on Synthetic Complete Medium (SCM) -Trp plates containing 2% galactose **(A)** or 2% glucose **(B)**, and incubated at 30°C for 48 hours. The experiment was performed in three biological replicates; representative results are shown.

A version of eIF4E1c protein with an additional N-terminal sequence predicted by TAIR (Rhee et al., 2003) was also tested as *EIF4E1C(long)*. This additional sequence is likely to be an artifact of gene assignment due to an incorrect prediction for the start site. The extra peptide sequence has no similarity to any known eIF4E peptide sequences, and would be unique among the *EIF4E1B*-like genes as well as plant *EIF4E* genes. The *EIF4E1C(long)* gene was not able to complement the deletion of *EIF4E* in yeast, indicating the additional N-terminal amino acid residues may interfere with either cap recognition or eIF4G binding, preventing productive translation initiation. This is consistent with observations by our laboratory that N-terminal or C-terminal fusions of plant cap-binding proteins are not viable *in vivo* in yeast or in *Arabidopsis* (Levin, Tseng, Browning, manuscript in preparation).

2.4. CONCLUSIONS

Arabidopsis thaliana is the best model system for plant translation initiation currently available due to the availability of knockout lines of many translation initiation factors for *in vivo* study as well as the successful purification of many recombinant proteins for these factors. *Arabidopsis thaliana* and other members of the Brassicaceae family have non-canonical *eIF4E* related genes present in their genomes. There is little evidence for *EIF4E1C* gene expression, and the *EIF4E1B* gene is expressed at low levels in most tissues, though microarray and RNA-Seq data support enrichment in reproductive tissue. Unfortunately, α-AteIF4E antibody cross-reacts poorly with eIF4E1b and eIF4E1c, so it is not possible to confirm that these proteins are actually produced *in vivo*. The eIF4E1b and eIF4E1c are *bona fide* cap-binding proteins sufficient to promote translation initiation in yeast, and display translation initiation activity *in vitro*. However, due to their low binding

affinity for eIF4G relative to eIF4E and their low level of expression, it seems unlikely that these genes contribute substantially to translation initiation in most plant tissues. Increasing numbers of non-canonical eIF4E family members have been described in eukaryotes (Rhoads, 2009). As more plant genomes are sequenced, other events similar to the *EIF4E* duplication and divergence in Brassicaceae may be observed. *Eucalyptus grandis* and *Fragaria vesca* both encode apparent divergent *EIF4E* genes, though there is not yet available data to tell whether they are expressed or if the genes are conserved among close relatives. Given the data for poor interaction with eIF4G and its identification as a sperm-enriched gene, one might expect the role of eIF4E1b to be similar to what has been described in vertebrate oocytes: binding the 7-methylguanosine cap and excluding eIF4G binding in order to repress translation. The germline enrichment of *A. thaliana EIF4E1B* seems in line with findings from vertebrates (Minshall et al., 2007), *C. elegans* (Amiri et al., 2001), and *Drosophila* (Hernandez et al., 2012) of specialized eIF4E isoforms with roles in reproductive tissue. However, the *in vitro* data from this work suggest that eIF4E1b does not contribute to translational repression in this manner, though mRNA-specific repression or interaction with other proteins cannot be ruled out. *Arabidopsis thaliana EIF4E1B* and *EIF4E1C* seem non-essential, as T-DNA insertion plants develop normally; however, the strong conservation of the *EIF4E1B*-type genes within the Brassicaceae family implies that they provide some as-yet-unidentified contribution. In addition, crosses between *EIF4E* (*cum1*) and either *EIF4E1B* or *EIF4E1C* T-DNA lines, do not have any observable phenotype or issues with reproduction. Thus, although eIF4E1b and eIF4E1c appear to be able to function as cap-binding proteins *in vitro* and in yeast, it remains to be determined if *EIF4E1C* is even expressed *in planta* and the actual levels of eIF4E1b protein expression and localization remain to be determined. Based on the large difference (~1600 fold) in binding affinity of eIF4E1b for eIF4G relative to eIF4E, it is unlikely eIF4E1b

plays any role in canonical translation, but perhaps has a Brassicaceae-specific role in some tissues where eIF4E protein is not expressed and would allow eIF4E1b to interact with eIF4G or other proteins.

Chapter 3

3.1. INTRODUCTION

After research had begun to define the role of eIF4E in cap-binding and regulation of translation, a search began for related proteins which might also have a cap-binding function. This led to the concurrent discovery of 4EHP (4E Homologous Protein) in humans (also sometimes called eIF4E2) and nCBP (novel Cap-Binding Protein) in *A. thaliana* (Rom et al., 1998; Ruud et al., 1998). These related proteins were subsequently assigned to Class II eIF4E, differentiated from Class I by hydrophobic substitutions for the tryptophan residues at positions equivalent to Trp-43 and Trp-56 of human eIF4E (Joshi, et al., 2005). While this family of proteins appears conserved throughout animals and plants, there are no representative homologs in model fungi species such as *S. cerevisiae*. After cloning based on homology, it was confirmed by recombinant protein purification that 4EHP of humans and nCBP of *A. thaliana* have cap-binding activity *in vitro* (Rom, et al., 1998; Ruud, et al., 1998). Following this point, plant nCBP will be referred to as 4EHP in order to simplify nomenclature of these proteins that look increasingly homologous.

Previous to this work, little was known about the function of 4EHP in plants. It is highly conserved across land plants and most green algae (Patrick and Browning, 2012). 4EHP can be purified from *Arabidopsis* cell culture by m^7GTP-Sepharose, as confirmed by western and mass spectrometry (Bush et al., 2009; Ruud et al., 1998). Recombinant eIFiso4G co-purifies on m^7GTP-Sepharose with 4EHP, and together they have a modest ability to stimulate translation *in vitro* (Ruud, et al., 1998). Immunolocalization experiments found that 4EHP is diffusely present in the nucleus and cytoplasm in *Arabidopsis* cell culture (Bush, et al., 2009). While eIF4E or eIFiso4E are often essential

to the pathogenesis of various plant viruses, *4EHP* has yet to be identified as a viral resistance allele.

3.1.1. 4EHP

Human 4EHP has been crystalized, and has a structure very similar to eIF4E with the exception of the eIF4G/4E-BP-binding interface (Rosettani et al., 2007). 4EHP has not been observed to bind eIF4G, and binds to 4E-BP 100-fold weaker than eIF4E (Rom, et al., 1998; Rosettani, et al., 2007). Human 4EHP has been described as weakly binding to cap analogs, showing 100- to 200-fold weaker affinity for m^7GTP relative to eIF4E (Rosettani et al., 2007; Zuberek et al., 2007). This is in contrast with recombinant *A. thaliana* 4EHP, which appears to bind to m^7GTP-Sepharose more strongly than eIF4E or eIFiso4E and requires higher levels of m^7GTP to elute during affinity purification (Ruud et al., 1998).

Interestingly, the weak cap-binding activity of human 4EHP appears to be enhanced by ISGylation (Okumura et al., 2007). ISG15 is a ubiquitin-like modifier protein, which is targeted to certain substrates upon type I interferon (IFN) stimulation; 4EHP is one such target of ISGylation by the E3 ligase human homolog of Ariadne (HHARI), which had previously been reported to ubiquitylate 4EHP (Okumura et al., 2007; Tan et al., 2003). More recently, it has been observed that in response to genotoxic stress 4EHP binds HHARI and is ubiquinated (von Stechow et al., 2015). These HHARI-dependent modifications of 4EHP in response to stress which result in apparent increased cap-binding activity have been speculated to contribute to repression of translation by increasing competitiveness with eIF4E (Okumura et al., 2007; von Stechow et al., 2015).

4EHP is an essential gene for mammalian development, as homozygous knockout of the gene causes aberrant embryonic development and perinatal lethality in mice (Morita et al., 2012). *4EHP* appears to be expressed across a wide range of mammalian tissue types but has particularly high expression in testes (Joshi et al., 2004). It has been reported that 4EHP is primarily cytosolic in human cell culture and mouse oocytes (Rom et al., 1998; Villaescusa et al., 2009); however, it has been recently reported that 4EHP undergoes nucleocytoplasmic shuttling (Kamenska et al., 2013).

The primary cellular function assigned to 4EHP to date is its function in *Drosophila* as an embryonic translational repressor. The first 4EHP binding partner to be discovered was the RNA-binding protein Bicoid, which binds to 4EHP through a 4E-BP-like motif and binds the 3' region of *caudal* mRNA; the tethering of 4EHP to *caudal* through Bicoid is presumed to promote 4EHP competition with eIF4E for the 5' cap of *caudal* mRNA and lead to a translationally inactive mRNP (Cho et al., 2005). A similar translational repression paradigm has been reported for *hunchback* mRNA through interaction of 4EHP and Brain tumor (Brat), a subunit of an RNA-binding complex that includes Nanos and Pumilio (Cho et al., 2006). In mouse oocytes, it has been proposed that 4EHP unexpectedly binds to the transcription factor Prep1 in the cytosol in order to repress translation of *Hoxb4* (Villaescusa et al., 2009). There is also recent evidence that 4EHP interacts with the zinc-finger protein Tristetraprolin to repress translation of ARE-containing mRNAs in human cell culture (Tao and Gao, 2015).

4EHP has also been ascribed positive translational regulatory roles. In *C. elegans*, the homologous IFE-4 protein co-fractionates with 40S ribosomal subunit containing fractions and contributes to polysomal loading of a small number of transcripts (Dinkova et al., 2005). More recently it has been discovered that in human cell culture, 4EHP contributes to escape of certain mRNAs from translational repression under hypoxia

through a complex including hypoxia-inducible factor 2α (HIF-2α) and the RNA-binding protein RBM4 (Uniacke et al., 2012). This mechanism of escape of hypoxic translational repression by 4EHP contributes to tumorigenesis (Uniacke et al., 2014).

3.1.2. GIGYF

In 2012, a new 4EHP binding partner was identified in mammals by immunoprecipitation and mass spectrometry, Grb10-interacting GYF protein 2 (GIGYF2) (Morita et al., 2012). GIGFY2 is a large (~180 kDa in humans) protein which binds 4EHP through an N-terminal 4E-BP-like motif; GIGYF1 shares this motif and is also a likely 4EHP binding partner based on mass spectrometry (Morita, et al., 2012). 4EHP and GIGYF can be co-purified from human cell culture lysate by m^7GDP-agarose beads, indicating they likely form a cap-binding complex *in vivo*, and knockdown of either subunit caused an apparent increase in overall rate of translation as measured by radioactive methionine incorporation (Morita, et al., 2012).

GIGYF2 was first identified in a screen to find interacting partners of Grb10, an important mediator of insulin-like growth factor (IGF-I) signaling (Giovannone et al., 2003). GIGYF2 was found to interact with Grb10 through its GYF domain, a protein-protein interaction domain specific for proline-rich sequences; the GIGYF2 protein also features a bipartite nuclear localization signal (Giovannone, et al., 2003). *GIGYF2* was identified as a candidate for the PARK11 locus, a chromosomal region linked to familial Parkinsons disease (Lautier et al., 2008). Subsequent analyses cast some skepticism on this finding, but more recently there has been support for *GIGYF2* mutations contributing to Parkinsons disease in some populations (Ruiz-Martinez et al., 2015; Zhang et al., 2015). *GIGYF2* mutations have also recently been connected to autism (Krumm et al., 2015) and

endometrial cancer (Le Gallo and Bell, 2014). Like *4EHP*, the homozygous knockout of *GIGYF2* leads to perinatal lethality in mice; heterozygous mutants show premature neurodegeneration and motor dysfunction (Giovannone et al., 2009).

GIGYF2 has been shown to interact with required for cell differentiation 1 homolog (RQCD1), also known as CNOT9 of the CCR-NOT4 complex, which contributes to cell growth signal transduction (Ajiro et al., 2009). Another interactor is zinc finger protein 598 (ZNF598), a homolog of yeast histone H4 E3 ligase Hel2p (Morita et al., 2012; Singh et al., 2012). GIGYF1, which has a predicted 14-3-3 interaction site (Giovannone, et al., 2003), has recently been identified by mass spectrometry in 14-3-3-interacting protein complexes following DNA damage, along with 4EHP and ZNF598 (Blasius et al., 2014). GIGYF2 has also been identified by mass spectrometry in complexes containing hepatocyte nuclear factor-4α (HNF4α) along with many transcriptional regulatory proteins (Daigo et al., 2011).

The GYF domain of the related yeast protein Smy2p has been crystalized in a ligand-bound form that reveals a hydrophobic cleft where the motif PPGΦ can bind, where Φ is a hydrophobic residue the identify of which is limited by the shape of the hydrophobic pocket in different GYF domain proteins (Ash et al., 2010). The preferred binding motif of GIGYF2 was determined by phage display to be PPG[FILMW] and the GYF domain was shown to bind to splicing-related proteins by yeast two-hybrid (Kofler et al., 2005). GIGYF2 localization has been observed to be cytoplasmic (Ajiro et al., 2009), endosomal (Higashi et al., 2010), or endoplasmic reticulum/Golgi-localized with inducible redistribution to stress granules (Ash, et al., 2010), but curiously pulldown experiments from human cell culture with the GYF domain of GIGYF2 identify a large number of nuclear transcriptional and splicing regulatory proteins (Ash, et al., 2010).

An *A. thaliana* homolog of *GIGYF2*, *AT5G42950*, first came to our attention when it was identified by mass spectrometry of cap-binding complexes purified by m^7GTP-Sepharose beads from *A. thaliana* cell culture (Bush, et al., 2009). Little is known about *AT5G42950*, but it is has been identified in proteomic studies as showing differential phosphorylation in response to a number of stimuli, including abscisic acid (Kline et al., 2010), auxin (Zhang et al., 2013a), and osmotic stress (Stecker et al., 2014). The GYF domain of *AT5G42950* has, like the human and yeast homologs, been tested by phage display to determine that its preferred binding motif is PPGF (Kofler et al., 2005). The *Arabidopsis* GIGYF protein also has the curious property of containing a PPGF motif itself, which could potentially promote self-oligomerization.

3.1.3. NON-CODING RNA

Long non-coding RNA (lncRNA), RNA transcripts of roughly 200 or more nucleotides which encode no protein, have recently been recognized as major players in human development and gene regulation. Many of these lncRNAs serve as scaffolds for protein complexes; well-studied examples include those lncRNAs that guide histone modifying proteins to specific chromatin loci, such as the *cis*-acting *X-inactive specific transcript* (*Xist*) which is required for X chromosome inactivation, and the *trans*-acting *HOX transcript antisense RNA* (*HOTAIR*) which contributes to silencing of the HOXD gene cluster (Fatica and Bozzoni, 2014). LncRNAs have many roles outside of chromatin regulation, such as regulating mRNA stability and translation (Fatica and Bozzoni, 2014) or nucleating and organizing non-membranous organelles like nuclear paraspeckles (Clemson et al., 2009).

An increasing variety of lncRNA species have been recognized in plants. Many are involved in the generation and targeting of small interfering RNA (siRNA), with precursor transcripts produced by RNA polymerase Pol IV that are used to produce siRNA that are targeted to Pol V-produced scaffold transcripts to effect gene silencing (Chekanova, 2015). Another well-studied lncRNA with a *cis*-acting chromatin modification function is *COLDAIR*, which through vernalization-induced transcription recruits Polycomb-Repressive Complex 2 (PRC2) to repress transcription of the flowering repressor gene *Flowering Locus C* (*FLC*); antisense non-coding transcripts such as *COOLAIR* also contribute to *FLC* repression (Chekanova, 2015). Most lncRNA in animals and plants have been observed or are assumed to have a 5' cap structure; 5' caps have been observed for *COLDAIR*, *COOLAIR*, many Arabidopsis small nucleolar RNAs (snoRNAs), and a large class of recently identified intermediate-sized non-coding RNA (Chekanova, 2015; Kim et al., 2010; Wang et al., 2014b; Zhang et al., 2013b).

A recently described *trans*-acting chromatin-associated non-coding RNA with transcriptional regulatory activity is *hidden treasure 1* (*HID1*), which is conserved from rice to *A. thaliana* . *HID1* cofractionates with large mRNP complexes and appears to bind the promoter region of *phytochrome interacting factor 3* (*PIF3*); *hid1* knockdown mutants show longer hypocotyl growth specifically under red light conditions in a *PIF3*-dependent manner, and it was proposed that *HID1* serves as a negative regulator of *PIF3* transcription (Wang et al., 2014a).

Observations made in this work support an evolutionarily conserved 4EHP-GIGYF cap-binding complex being present in plants that is homologous to the described mammalian complex. Surprisingly, however, this complex may have functions outside of translational regulation. We propose that the 4EHP-GIGYF complex of *A. thaliana* binds

to non-coding RNA and may therefore be involved in a gene regulatory mechanism independent of translation.

3.2. METHODS

3.2.1. Cloning of 4EHP, GIGYF, and Dicistronic Expression Plasmids

Arabidopsis 4EHP (*AT5G18110*) and GIGYF (*AT542950*) were codon optimized for expression in *E. coli* using DNAWorks (Hoover and Lubkowski, 2002) and assembled by overlap PCR of oligonucleotides (Horton et al., 1989). Initial cloning of PCR products was into pCR-Blunt-II-TOPO (Invitrogen) followed by subcloning into pSB1AC3 (Shetty et al., 2008) and pET22b vectors (Novagen). 4EHP was cloned in one step into pCR-Blunt-II-TOPO and subcloned into pSB1AC3, while GIGYF was cloned into pCR-Blunt-II-TOPO in three sections; and then assembled in pSB1AC3. 4EHP and GIGYF expression plasmids were created by subcloning into pET22b. The pET22b GIGYF vector was used to clone 4EHP at a site 3' of the GIGYF coding region to create a dicistronic expression plasmid.

3.2.2. Expression and Affinity Purification of 4EHP and 4EHP-GIGYF Complexes

4EHP protein was expressed in BL21(DE3) E. coli and purified as previously described by m^7GTP-Sepharose (GE Healthcare) affinity chromatography (Ruud et al., 2005) with the following modifications: $N^1_{7.6}$ buffer (20 mM Hepes, pH 7.6, 0.1 mM EDTA, 2 mM DTT, 10% glycerol) with 0.1M KCl was used for cell resuspension and purification; the sonicated cell lysate was centrifuged for 45 minutes at 48,000 RPM in a 50.2 Ti rotor at 4°C. 4EHP-GIGYF complexes were expressed in TUNER(DE3) *E. coli*

(Novagen) and resuspended in $N^1_{7.6}$ buffer containing 0.3M KCl and purified using a 2 mL phosphocellulose P11 (Whatman) column, and eluted with $N^1_{7.6}$ buffer containing 0.5 M KCl. Fractions containing the highest concentration of protein were pooled, diluted to 0.1 M KCl with $N^1_{7.6}$ buffer, and further purified by m^7GTP-Sepharose affinity chromatography similarly to 4EHP, with the exception of elution, which was performed by adjusting the buffer to 1M KCl. Copurification of GIGYF with 4EHP by m^7GTP-Sepharose affinity was confirmed by mass spectrometry. Eluted proteins were separated by SDS-PAGE followed by Coomassie stain, band excision, and washing with deionized water. The gel slice was submitted to The University of Texas at Austin Proteomics Facility for trypsin digest and LC-MS/MS analysis by Orbitrap Fusion (Thermo Scientific). The spectra were visualized by Scaffold (Proteome Software).

3.2.3. Cloning of GST-GIGYF$_{1-605}$ and GST-GIGYF$_{1059-1658}$ Expression Plasmids

Cloning of GST-GIGYF$_{1-605}$ was performed by PCR using the primers GN1 (5'-GAGAGGAATTCCCGATGGCCAATAGTAGCGCAG-3') and GN2 (5'-GAGAGTCGACTTATTAAGCGCGCAGATGGGGCATC-3') with pET22b codon-optimized GIGYF as a template, with cloning of the PCR product into pCR-Blunt-II-TOPO (Invitrogen) followed by subcloning into pGEX4T-1 (GE Healthcare). Cloning of GST-GIGYF$_{1059-1658}$ for antibody production was performed similarly with the primers GAB1 (5'-GAGACATATGATTGTGGAAAGCAAACTGCTCG-3') and GAB2 (5'-GAGAGGATCCTCATCAAGTGTTATTCCGTGTGC-3').

3.2.4. Expression and Purification of GST-GIGYF$_{1-605}$, and GST-GIGYF$_{1059-1658}$, and GST

Purification of GST and the two GST-tagged proteins was performed similarly: proteins were expressed in BL21(DE3) *E. coli*, with initial pellet resuspension in binding buffer, phosphate-buffered saline (PBS) pH 7.3 (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, adjusted to pH 7.3 with HCl) with 1 small tablet per 10 mL buffer Pierce Protease Inhibitor (Thermo Fisher Scientific) and 1 mM phenylmethylsulfonyl fluoride (PMSF). Resuspended cells were sonicated followed by ultracentrifugation for 45 minutes at 48,000 RPM in a 50.2 Ti rotor at 4°C. Supernatants were passaged through 2 mL of Sephadex G-25 (Sigma-Aldrich) to remove debris followed by chromatography on 0.5 mL of Glutathione Sepharose 4B (GE Biosciences). Resin was washed with binding buffer followed by elution with GST elution buffer (50 mM Tris-HCl pH 7.6, 10 mM glutathione). Eluate fractions containing protein as measured by Protein Assay Dye Reagent (Bio-Rad) were pooled and dialyzed overnight at 4°C in PBS, pH 7.3. Purity and intactness of the dialyzed protein was confirmed by SDS-PAGE and Coomassie staining. The protein concentration was determined by Bradford assay.

3.2.5. GIGYF Antibody

The internal section GIGYF$_{1059-1658}$ was selected based upon alignments of plant GIGYF and GIGYF-like proteins as being likely to contain epitopes solely present in GIGYF. Rabbit antibody was raised to purified GST-GIGYF$_{1059-1658}$ (Pocono Rabbit Farm & Laboratory). Final antiserum was tested against extracts from Col0 and *gigyf-1* (see below) plants to confirm antibody specificity.

3.2.6. Phylogeny and Alignment of GIGYF Proteins

Sequences of GIGYF and related proteins were collected from Phytozome (Goodstein, et al., 2012) and NCBI RefSeq (Pruitt et al., 2007). The Phylogeny.fr pipeline (Dereeper et al., 2008) was used for GIGYF alignment and phylogenetic tree generation with alignment by MUSCLE and tree construction by PhyML using 100 bootstrap replicates. Alignments of subdomains were performed in ClustalW2 (Larkin, et al., 2007).

4.2.7. 4EHP Pulldown by GST-GIGYF$_{1-605}$

Equimolar amounts of purified 4EHP and either GST or GST-GIGYF$_{1-605}$ (the fraction of GIGYF containing the predicted 4EHP binding site) were mixed in PBS pH 7.3 buffer in a total volume of 120 μL. Two input aliquots of 10 μL each were set aside, and the remainder was mixed with 100 μL of Glutathione Sepharose 4B beads that had been pre-equilibrated in PBS pH 7.3 in a 1.5 mL microfuge tube. The tubes were placed on ice for 10 minutes with regular mixing (every 2 minutes), and then centrifuged at 500 x *g* at room temperature for 2 minutes. The supernatant was removed and the beads were washed with 100 μL of cold PBS pH 7.3 for 5 minutes on ice with regular mixing, then centrifuged at 500 x *g* at room temperature for 2 minutes. The wash step was repeated 2 more times, with a final elution of the bound protein with 100 μL of GST elution buffer (50 mM Tris-HCl pH 7.6, 10 mM glutathione). Input and bead pulldown samples (10 μL each) were subjected to SDS-PAGE, transferred to PVDF and probed with 1:1000 α-4EHP antisera (Ruud, et al., 2007) or 1:1000 α-GST antibody (Thermo Fisher Scientific, CAB4169). Secondary antibody incubation with 1:20,000 Goat anti-Rabbit IgG, peroxidase labeled (Kirkegaard & Perry Laboratories) was followed by detection with SuperSignal West Pico

or Femto reagent (Thermo Fischer Scientific) and imaging by ChemiDoc MP Imager (Bio-Rad).

3.2.8. GST-GIGYF$_{1-605}$ Pulldown by 4EHP on m^{7}GTP-Sepharose

Verification of ability to form a cap-binding complex was performed similarly to GST pulldown: equimolar amounts of purified 4EHP and GST-GIGYF$_{1-605}$ were mixed in with N$^1_{7.6}$ buffer containing 0.1 M KCl in a total volume of 120 µL. A control was performed with GST-GIGYF$_{1-605}$ only. Two input aliquots of 10 µL each were set aside, and the remainder was mixed with 40 µL m^{7}GTP-Sepharose beads that had been pre-equilibrated in N$^1_{7.6}$ buffer containing 0.1 M KCl in a 1.5 mL microfuge tube. The tubes were placed on ice for 10 minutes with regular mixing (every 2 minutes), and then centrifuged at 500 x *g* at room temperature for 2 minutes. The supernatant was removed and the beads were washed with cold 100 µL N$^1_{7.6}$ buffer containing 0.1 M KCl for 5 minutes on ice with regular mixing, then centrifuged at 500 x *g* at room temperature for 2 minutes. The wash step was repeated 2 more times, with a final elution of the bound protein with 100 µL 2x Laemmli Sample Buffer. Input and bead pulldown samples (10 µL each) were subjected to SDS-PAGE, transferred to PVDF and then probed with 1:1000 α-4EHP antisera or 1:1000 α-GST antibody.

3.2.9. Nuclear Localization of 4EHP and GIGYF

Nuclear-cytoplasmic fractionation was performed as previously described (Wang et al., 2011) with some modifications. 0.2 g of tissue from 14-day-old *A. thaliana* seedlings or 0.3 g of tissue from *A. thaliana* callus (courtesy of Dr. Anne Simon, University of

Maryland) was flash frozen in liquid nitrogen and homogenized followed by resuspension in 1 mL lysis buffer (20 mM Tris-HCl pH 7.4, 20 mM KCl, 2 mM EDTA, 2.5 mM $MgCl_2$, 25% glycerol, 250 mM sucrose, 1 mM DTT) and supplemented with 1 small tablet per 10 mL buffer Pierce Protease Inhibitor (Thermo Fisher Scientific). The homogenate was filtered through a Falcon 40 μm Cell Strainer (Corning) and 50 μL of filtered homogenate was mixed with an equal volume of 2x Laemmli Sample Buffer (Bio-Rad) as an input sample. The filtered homogenate was centrifuged at 1500 x *g* for 10 minutes at 4°C, and the supernatant (cytoplasmic fraction) was then centrifuged at 10,000 x *g* for 10 minutes at 4°C. 50 μL of the cleared supernatant was mixed with 2x Laemmli Sample Buffer as a cytoplasmic fraction sample. The pellet from the original centrifugation (1500 x *g*) was resuspended in 1 mL nuclear resuspension buffer with Triton X-100 (NRBT, 20 mM Tris-HCl pH 7.4, 2.5 mM $MgCl_2$, 25% glycerol, 0.2% Triton X-100) and centrifuged at 1500 x *g* for 10 minutes at 4°C. The supernatant was discarded and the wash step in NRBT buffer was repeated twice more. A final wash was performed by resuspension in nuclear resuspension buffer (NRB, 20 mM Tris-HCl pH 7.4, 2.5 mM $MgCl_2$, 25% glycerol), and centrifugation at 1500 x *g* for 10 minutes at 4°C. The supernatant was discarded and the nuclei pellet was resuspended in 80 μL of 2x Laemmli Sample Buffer. SDS-PAGE and western blots were performed with 10 μL aliquots of the input (total), cytoplasmic, and nuclear fractions and probed with control antibodies for cytoplasm, 1:5000 α-PEP Carboxylase (Abcam, ab34793), and for nuclei, 1:2000 α-Histone H3 (Abcam, ab1791) followed by 1:1000 α-4EHP and 1:1000 α-GIGYF antisera.

3.2.10. Immunoprecipitation of 4EHP

The IgG fraction was purified from 3 mL of α-4EHP serum by Protein A Sepharose CL-4B beads (GE Healthcare). Recombinant purified 4EHP protein crosslinked to agarose beads using AminoLink Plus Immobilization Kit (Thermo Fisher Scientific) and the IgG fraction was affinity purified following manufacturer instructions. Affinity purified α-4EHP was crosslinked to Pierce Protein A/G Plus Agarose resin (Thermo Fisher Scientific) following manufacturer instructions.

300 mg of tissue from 14-day-old *A. thaliana* seedlings that had been flash frozen in liquid nitrogen and ground was homogenized in 500 µL Pierce IP buffer (Thermo Fisher Scientific) with 1 small tablet per 10 mL buffer Pierce Protease Inhibitor (Thermo Fisher Scientific) and 1 mM PMSF. The homogenate was centrifuged 16,000 x *g* for 10 minutes at 4°C and the supernatant was filtered through a Falcon 40 µm Cell Strainer (Corning). The filtered supernatant was pre-cleared for 30 minutes with Protein A Sepharose CL-4B beads and then incubated with α-4EHP crosslinked beads (prepared as described above) at room temperature for 1.5 hours with rotation. The supernatant was recovered and the resin was washed following instructions of the Pierce Crosslink Immunoprecipitation Kit (Thermo Fisher Scientific). Bound protein was then eluted with 2x Laemmli Sample Buffer (Bio-Rad). SDS-PAGE and western blot was performed with 10 µL aliquots of α-4EHP IP, input, and IP supernatant, followed by probing with 1:1000 α-4EHP and 1:1000 α-GIGYF antisera. For mass spectrometry analysis, 10 µL of α-4EHP IP was ran briefly into a Novex 4-12% Tris-Glycine gel (Thermo Fisher Scientific) by SDS-PAGE followed by Coomassie stain, band excision, and washing with deionized water. The gel slice was submitted to The University of Texas at Austin Proteomics Facility for trypsin digest and LC-MS/MS analysis by Orbitrap Fusion (Thermo Scientific). The spectra were visualized by Scaffold (Proteome Software).

3.2.11. *4ehp* and *gigyf-1* T-DNA Insertion Lines

T-DNA insertion lines *4ehp* (salk_131503) and *gigyf-1* (salk_135013) were obtained from the Arabidopsis Biological Resource Center. Screening primers were designed using Salk T-DNA Express (Alonso et al., 2003) and Primer3 (Untergasser et al., 2012). *4ehp* plants were screened with primers RT2 (5'-GGATTTTCGTATTGGTGGTATTGA-3'), LF2 (5'-TGATGGTCAGAAGCATTGCGA-3') and LB1nested. *gigyf-1* plants were screened with LP1 (5'-TGGCGTCAGTCTTCTGGGGG-3'), RP1 (5'-TGCAGATTCCCGACCAGGCCA-3') and LBb1.3 (5'-ATTTTGCCGATTTCGGAAC-3').

3.2.12. RNA Sequencing

Col0, *4ehp*, *gigyf-1*, or *4ehp x gigyf-1* seedlings were grown on 1% sucrose MS media plates (4.3 g/L Murashige and Skoog Powder [PhytoTechnology], 2.3 mM MES, 1x vitamins [Sigma]) adjusted to pH 5.7 with KOH, with 8 g/L Phytagel (Sigma-Aldrich). Tissue from 7-day-old seedlings was harvested, flash frozen with liquid nitrogen, and ground to a powder. 80 mg of tissue was subjected to RNA extraction by RNeasy Plant Mini Kit (Qiagen), with final elution in 30 μL RNase-free water. 3 μL of 10x buffer and 3 μL of DNase I (recombinant, RNase-free) (Roche) was added and mixed by pipetting, then incubated for 1 hour at 37°C. RNA cleanup of digested samples was performed with RNeasy Mini Kit (Qiagen) with elution in 30 μL RNase-free water. RNA samples, in biological triplicate, were submitted to The University of Texas at Austin Genome Sequencing and Analysis Facility for library preparation and paired end sequencing

(Illumina NextSeq 500). Paired end reads were aligned to the annotated *A. thaliana* reference transcriptome from The Arabidopsis Information Resource version 10 (Lamesch et al., 2012) with TopHat (Trapnell et al., 2012), and transcript quantitation and differential expression were produced with Cuffdiff (Trapnell, et al., 2012). GO Term Enrichment was analyzed with AmiGO (Carbon et al., 2009). Integrative Genomics Viewer (IGV) was used for read coverage visualization (Thorvaldsdottir et al., 2013).

3.2.13. RNA Immunoprecipitation and Sequencing

Affinity purified α-4EHP antibody was cross-linked to 60 μL of Protein A MagBeads (GenScript, L00273) following the manufacturer's instructions. *A. thaliana* Col0 14-day-old seedlings were grown on 1.5% sucrose MS media plates (as described above), harvested, and washed in cold sterile water. Seedlings were cross-linked in 1% formaldehyde for 15 minutes under vacuum infiltration (Terzi and Simpson, 2009). The cross-linking reaction was stopped with a 1/10 volume of 1M glycine pH 2.5 for 5 minutes under vacuum. Cross-linked seedlings were then washed with cold sterile water, briefly dried at room temperature, and flash frozen in liquid nitrogen and ground to a powder. 15g of cross-linked Col0 tissue was homogenized in 20 mL of RNA immunopreciptation (RIP) Buffer A (50 mM Tris-HCl pH 7.4, 100 mM KCl, 2.5 mM $MgCl_2$, 0.1% NP-40, 20U/mL RNaseOUT [Invitrogen]) containing 1 small tablet per 10 mL buffer cOmplete EDTA-free protease inhibitor (Roche) and 1 mM PMSF, and then centrifuged at 15,000 x *g* for 15 minutes at 4°C (adapted from (Carbonell et al., 2012). The supernatant was then filtered through Miracloth (EMD Millipore) and incubated with α-4EHP MagBeads (pre-equilibrated in RNase-free PBS pH 7.0) with rotation for 1 hour at 4°C. Following incubation, beads were washed 3 times with RNase-free PBS pH 7.0 and then eluted under

non-denaturing conditions per manufacturer instructions with 0.1 M glycine pH 2.5 in a total volume of 120 μL and neutralized with 12 μL of 1 M Tris-HCl pH 8.5. Eluted protein-RNA complexes were spiked with additional RNaseOUT, subjected to 0.4 mg/mL proteinase K (Ambion) and decross-linked with 0.2 M NaCl for 1 hour at 42°C followed by 1 hour at 65°C (Terzi and Simpson, 2009). DNA was removed by incubation with DNase I recombinant, RNase-free (Roche) for 1 hour at 37°C. RNA was isolated by TRIzol reagent (Thermo Fisher Scientific) following the manufacturer's instructions and resuspended in 20 μL of RNase-free water. The immunoprecipitated RNA sample was submitted to The University of Texas at Austin Genome Sequencing and Analysis Facility for library preparation and paired end sequencing (Illumina HiSeq 2500). For a control, three biological repeats of Col0 total RNA were prepared through TRIzol extraction of 400 mg frozen and homogenized tissue from 14-day-old seedlings, followed by ribosomal RNA removal with Ribo-Zero rRNA Removal Kit (Epicentre) and library preparation and sequencing. Paired end reads were aligned to the annotated *A. thaliana* reference transcriptome from The Arabidopsis Information Resource version 10 (Lamesch, et al., 2012) with TopHat (Trapnell, et al., 2012), and transcript quantitation and differential expression were produced with Cuffdiff (Trapnell, et al., 2012).

3.2.14. RNA Immunoprecipitation with RT-PCR

For investigation of binding targets by RIP-RT-PCR, 300 mg of formaldehyde cross-linked, frozen, and ground Col0 tissue (compared with control *4ehp* or *gigyf-1* RNA immunoprecipitation) was homogenized in 1 mL RIP Buffer B (50 mM HEPES pH 7.6, 150 mM NaCl, 1 mM EDTA, 2.5 mM $MgCl_2$, 10% glycerol, 0.1% Tween 20 [Sigma-Aldrich], 20U/mL RNaseOUT) with Pierce protease inhibitor (Thermo Fisher Scientific)

and 1 mM PMSF and then centrifuged at 16,000 x g for 15 minutes at 4°C. The supernatant was filtered through Falcon 40 μm Cell Strainer (Corning) followed by the addition of 50 μL of affinity purified α-4EHP or α-GIGYF antibody and rotation for 1.5 hours at 4°C. 30 μL of Protein A MagBeads (GenScript, L00273), pre-equilibrated in RNase-free PBS pH 7.0, were then added, followed by rotation for 1.5 hours at 4°C. Beads were washed 4 times with RNase-free PBS pH 7.0 and RNA was extracted with TRIzol reagent (Thermo Fisher Scientific) as per manufacturer instructions. RNA was resuspended in a volume of 15 μL of RNase-free water, with 2.5 μL of 10x buffer and 2.5 μL of DNase I recombinant, RNase-free (Roche) then added, mixed by pipetting and incubated for 1.5 hour at 37°C. RNA cleanup was performed by RNeasy Mini Kit (Qiagen) with elution in 25 μL of RNase-free water. cDNA was generated using SuperScript II Reverse Transcriptase (Thermo Fisher Scientific) and random hexamers. PCR was performed using Taq DNA Polymerase (New England Biolabs).

3.2.15. Cloning of *HID1* and *HID1as*

Cloning of *HID1* was performed by PCR using the primers HIDB1 (5'-GAGGATCCTAATACGACTCACTATAGGGGGACGTGACGAGATTTCTCTCCA-3') to add a 5' BamHI site and T7 promoter, and HIDH1 (5'-GAAAGCTTGGAATATAGATCTCATGGCCG-3') to add a 3' HindIII site, with *A. thaliana* cDNA as a template, with cloning of the PCR product into pCR-Blunt-II-TOPO (Invitrogen). Cloning of the antisense RNA, *asHID1,* was performed similarly with the primers AHIDB1 (5'-GAGGATCCTAATACGACTCACTATAGGGGGAATATAGATCTCATGGCCG-3') to add a 5' BamHI site and T7 promoter, and AHIDH1 (5'-

GAAAGCTTACGTGACGAGATTTCTCTCC-3') to add a 3' HindIII site, with *A.thaliana* DNA as a template. Plasmids were linearized with HindIII for use as a transcription template.

3.2.16. Fluorescent RNA EMSA

Fluorescently labeled *HID1* and *asHID1* RNA were generated using the MEGAshortscript kit (Ambion) in a 40 μL reaction with 2 μL Fluorescein-12-UTP (Perkin-Elmer) in a reaction at 37°C for 3 hours followed by 15 minutes of DNase treatment and cleanup with RNeasy Mini Kit. Electrophoretic Mobility Shift Assay (Hellman and Fried, 2007) was performed by mixing labeled RNA with indicated proteins (GST-GIGYF$_{1\text{-}605}$, GST-GIGYF$_{1059\text{-}1658}$, or GST) in a 10 μL reaction with N1$_{7.6}$ buffer, 0.1 M KCl. Mixtures were incubated for 30 minutes on ice and then separated on 1% agarose in 1x TBE buffer (Invitrogen) for 45 minutes at 70V. Fluorescence imaging was performed on ChemiDoc MP Imager (Bio-Rad).

3.3. RESULTS

3.3.1. GIGYF in Plants

Based on the report of GIGYF2 as a 4EHP-binding protein in mammals (Morita, et al., 2012), and our observation that a GIGYF-related protein, AT5G42950, had been observed in mass spectrometry of cap-binding complexes isolated from *A. thaliana* cell culture (Bush, et al., 2009), we chose to investigate AT5G42950 as a potential 4EHP-binding protein in *A. thaliana*. NCBI Blast results (Johnson et al., 2008) indicated that AT5G42950 is the closest related protein to GIGYF2 in *A. thaliana*, and was therefore

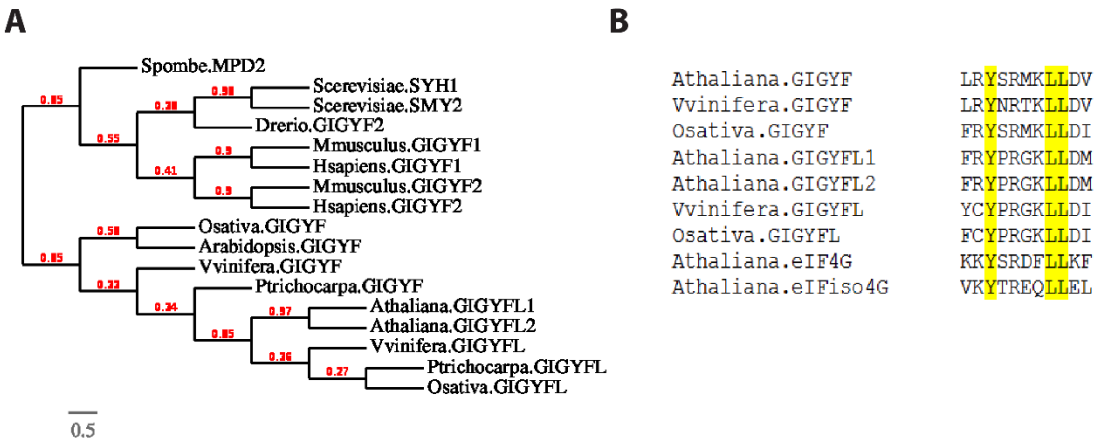


Figure 3.1. *A. thaliana* GIGYF and GIGYF-Like Proteins

(A) Phylogenetic tree of GYF-domain containing genes from yeast, animals, and plants. The Phylogeny.fr pipeline (Dereeper et al., 2008) was used for alignment and phylogenetic tree generation with alignment by MUSCLE and tree construction by PhyML using 100 bootstrap replicates. **(B)** Alignment of the 4E-Binding motif present in GIGYF, GIGYFL, and eIF4G/eIFiso4G.

assigned the designation GIGYF. However, two GIGYF-Like (GIGYFL) proteins (AT1G24300 and AT1G27430) are also encoded in the *A. thaliana* genome. Sequences from representative mammalian and plant GIGYF and GIGYFL proteins, as well as the yeast GYF-domain containing proteins (SMY2 and SYH1 in *S. cerevisiae*, MPD2 in *S. pombe*) were aligned with MUSCLE and a phylogeny was generated using PhyML (Figure 3.1a) (Dereeper, et al., 2008). Based on this alignment and gene information for plant homologues gathered in Phytozome (Goodstein, et al., 2012), it appears that GIGYF is conserved throughout flowering plants and is somewhat more closely related to the GIGYF2 protein present in animals, while a second GIGYFL clade is also present and

conserved in flowering plants. A putative 4E-BP-like sequence (the Y(x)$_4$LΦ motif) is present in the N-terminal portion of GIGYF and GIGYFL proteins, similar to what has been reported for 4EHP-binding proteins in animals and for mammalian GIGYF2. Alignment of this motif present in GIGYF and GIGYFL proteins (and compared to 4E-BP motifs of eIF4G and eIFiso4G) show strong conservation of the Y(x)$_4$LL sequence.

3.3.2. Formation of a 4EHP-GIGYF Complex

To demonstrate the possibility of a 4EHP-GIGYF cap-binding complex forming, codon-optimized versions of the *A. thaliana* proteins were cloned in single or dicistronic expression vectors. The dicistronic vector was expressed in *E. coli* in order to co-purify the putative complex. The cell lysate was subjected first to a phosphocellulose column purification step similarly to what is performed for eIF4F (Mayberry, et al., 2007) on the assumption that GIGYF, like eIF4G, would bind to phosphocellulose and elute in a salt-dependent manner due to the presence of large stretches of positively charged residues. After observing bands of the expected size in the eluate, fractions were pooled and subjected to m^{7}GTP-Sepharose affinity column purification. Several attempts led to the surprising realization that 4EHP-GIGYF complex appears to bind to the column very tightly and normal competition off the column with excess m^{7}GTP as is performed for eIF4F and for 4EHP alone was not sufficient for elution. Finally, it was possible to show binding of 4EHP-GIGYF to the m^{7}GTP-Sepharose column by eluting with 1 M KCl (Figure 3.2). The large band expected to be GIGYF was excised from an SDS-PAGE and its identity was confirmed by mass spectrometry (University of Texas at Austin Proteomics Facility). However, in attempting follow-up experiments with purified 4EHP-GIGYF complex, we found that GIGYF appeared to have solubility issues and was difficult to work

with, and suspicion fell to the previously identified self-interaction motif (PPGF) (Kofler et al., 2005) present in GIGYF which may lead to formation of aggregates *in vitro*. We attempted alternative approaches in order to gather more evidence for formation of a 4EHP-GIGYF complex.

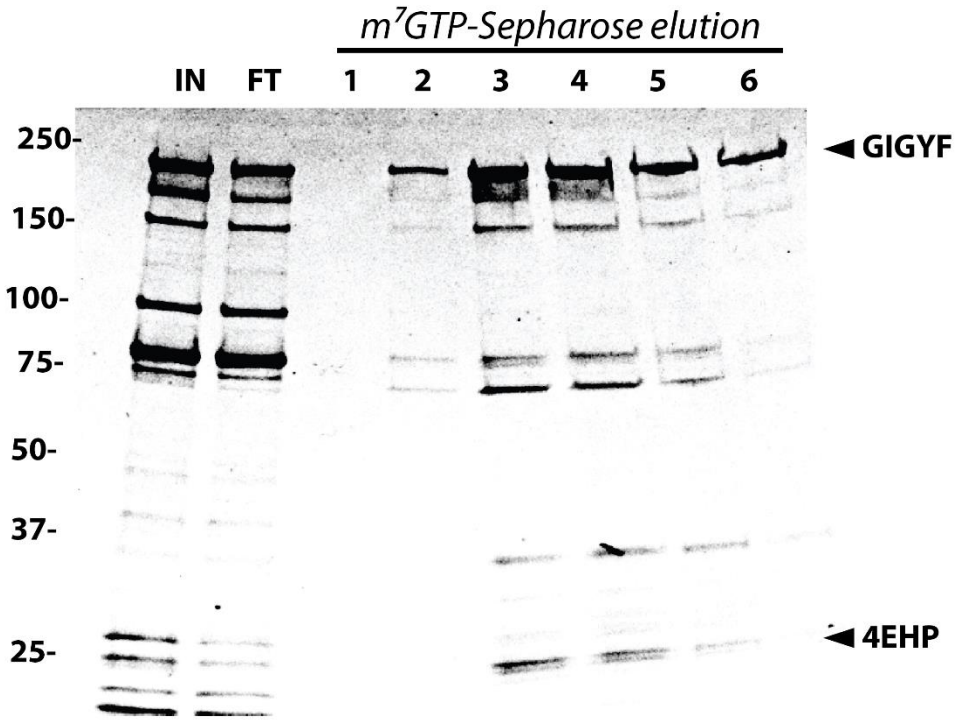


Figure 3.2. Binding of 4EHP-GIGYF Complex to m^{7}GTP-Sepharose

Recombinant *A. thaliana* 4EHP and GIGYF were co-expressed in *E. coli* and purified by phosphocellulose and m^{7}GTP-Sepharose affinity chromatography to confirm their ability to form a complex with biochemically relevant cap-binding activity. Input (IN) and flowthrough (FT) fractions were loaded (10 μL) along with elution fractions 1-6 (10 μL), separated by SDS-PAGE and Coomassie stained.

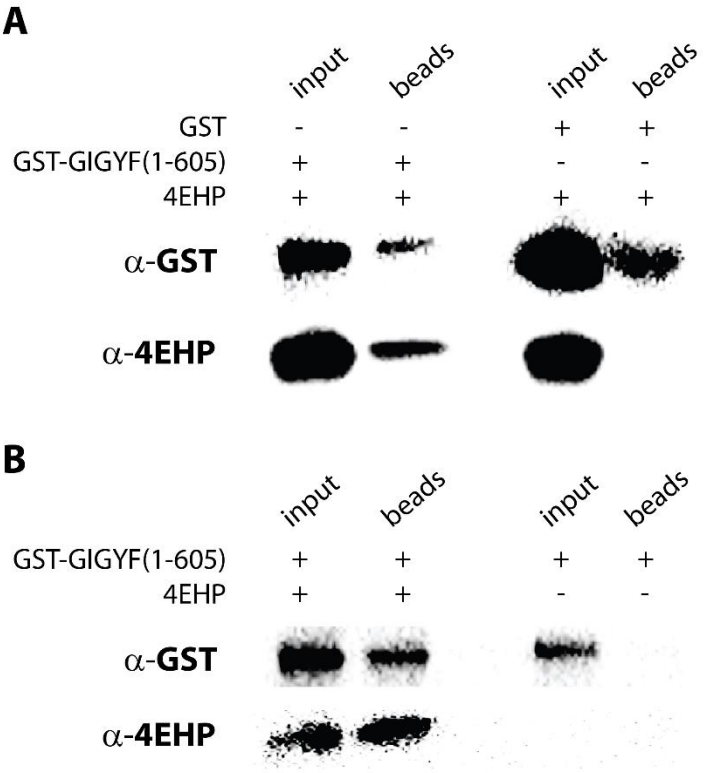
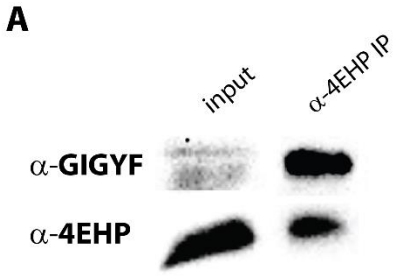


Figure 3.3. Pulldown of 4EHP and GST-GIGYF$_{1-605}$ Complexes

(A) The N-terminal portion of GIGYF, containing the putative binding site for 4EHP, was fused to GST for expression and purification. Equimolar amounts of 4EHP and either GST-GIGYF$_{1-605}$ or GST were tested by Glutathione Sepharose 4B pulldown to show the specific binding of 4EHP to GST-GIGYF$_{1-605}$. 10 μL aliquots of the input and eluate fractions were loaded and separated by SDS-PAGE followed by blotting to PVDF and Western blot detection. **(B)** GST-GIGYF$_{1-605}$ was tested for its ability to bind m^{7}GTP-Sepharose beads in the presence or absence of an equimolar amount of 4EHP. 10 μL aliquots of the input and eluate fractions were loaded and separated by SDS-PAGE followed by blotting to PVDF and Western blot detection.

The N-terminal portion of GIGYF, residues 1-605, containing the putative 4EHP binding site and the GYF domain, but removing the self-interaction site, was cloned into a GST fusion vector, expressed, and purified. In order to demonstrate the ability of the N-terminal portion of GIGYF to bind to 4EHP, purified 4EHP was mixed with either GST-GIGYF$_{1-605}$ or GST and was subjected to pulldown by Glutathione Sepharose 4B beads (Figure 3.3a). 4EHP was present in the GST-GIGYF$_{1-605}$ pulldown but not the GST control. To test for the ability of the proteins to form a cap-binding complex, GST-GIGYF$_{1-605}$ was incubated with m^{7}GTP-Sepharose beads in either the presence or absence of 4EHP. Beads were washed and then protein was eluted with 2X Laemmli Sample Buffer (Figure 3.3b). GST-GIGYF$_{1-605}$ bound and eluted from the m^{7}GTP-Sepharose beads only when 4EHP was present, indicating that the GIGYF only binds to m^{7}GTP-Sepharose in a 4EHP-dependent manner.



B

protein name	UniProt ID	peptides
GIGYF	Q9FMM3	13
4EHP	IF4E5	7

Figure 3.4. Co-immunoprecipitation of GIGYF with 4EHP

(A) Cell lysate of *A. thaliana* Col0 seedlings were subjected to immunoprecipitation with α-4EHP antibody. Input and precipitate fraction aliquots (10 μL) were tested by SDS-PAGE and western blot for the presence of 4EHP and GIGYF protein. **(B)** α-4EHP immunoprecipitate (10 μL) was analyzed by mass spectrometry for the presence of potential binding proteins. Among candidate 4EHP proteins, only GIGYF was detected.

In order to learn more about the *in vivo* role of GIGYF, antibody was raised to a portion of the C-terminus that was chosen due to its dissimilarity to the related GIGYFL proteins. The presence of 4EHP-GIGYF complex *in vivo* was tested by immunoprecipitation of 4EHP from Col0 seedling extract (Figure 3.4a). GIGYF was detected by western blot as being enriched by 4EHP immunoprecipitation. To confirm that GIGYF was the primary partner of 4EHP in *A. thaliana* seedlings, the 4EHP immunoprecipitation sample was examined by mass spectrometry (Figure 3.4b). Peptides matching 4EHP and GIGYF were detected, however, there was no evidence of eIF4G, eIFiso4G, or GIGYFL proteins in the 4EHP immunoprecipitation suggesting that these proteins are not interaction partners *in vivo*.

3.3.3. *4ehp* and *gigyf-1* Mutant Lines

In order to better understand the function of the 4EHP-GIGYF complex, T-DNA insertion lines were procured from ABRC and verified for *4ehp* (SALK_131503, insertion in the 4th exon) and *gigyf-1* (SALK_135013, insertion in the 6th exon). *4ehp* lines are robustly green and slightly late flowering but do not have a significant phenotype, while *gigyf-1* mutants are nearly indistinguishable from Col0. A double mutant line, *4ehp x gigyf-1*, was generated, which shows a much more distinct developmental phenotype (Figure 3.5). Double mutant plants are pale, slow growing, somewhat late flowering, and display premature leaf senescence, but are viable and fertile.

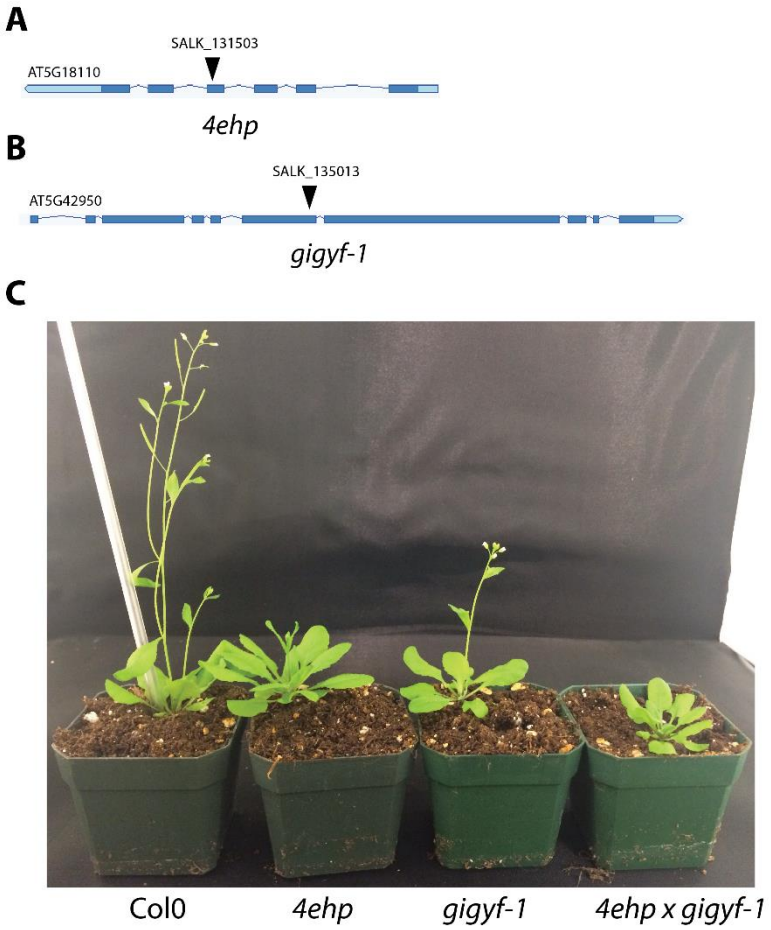


Figure 3.5. *4ehp* and *gigyf-1* T-DNA Lines

(A) The T-DNA insertion site of *4ehp*. **(B)** The T-DNA insertion site of *gigyf-1*. **(C)** 36-day old seedlings. While *gigyf-1* plants show only a minor phenotype and *4ehp* plants are slightly delayed in flowering, *4ehp x gigyf-1* plants have a more severe phenotype including slow growth, later flowering, and premature leaf senescence.

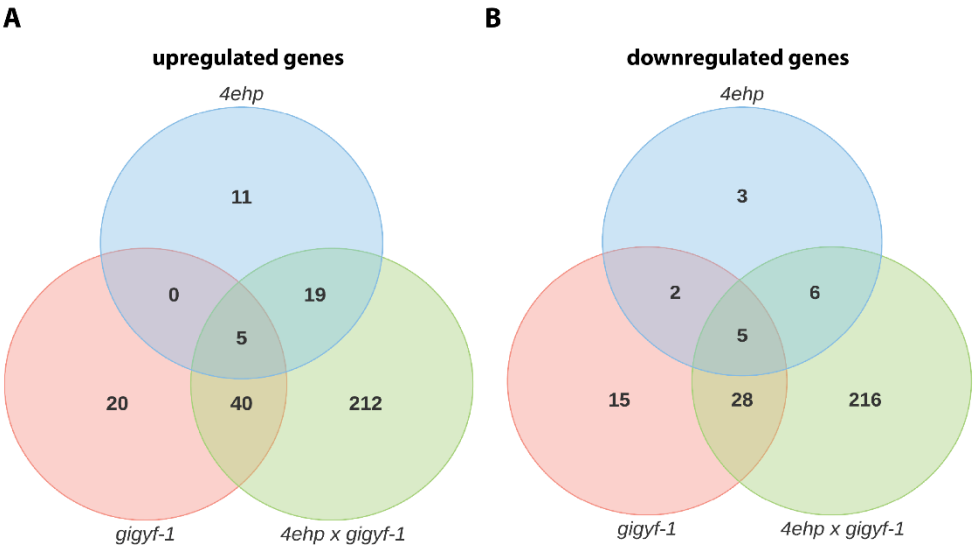


Figure 3.6. Differentially Regulated Genes in Mutant Lines

(A) Genes found to be up-regulated (FDR < 0.05, at least 2-fold change in expression) in *4ehp*, *gigyf-1*, and *4ehp x gigyf-1* lines. **(B)** Genes found to be downregulated in *4ehp*, *gigyf-1*, and *4ehp x gigyf-1* lines under the same criteria.

In order to better understand the function of 4EHP and GIGYF in gene regulation, mRNA from 7-day-old seedlings of mutant plants grown on MS plates was sequenced and tested for differential expression at a cutoff of FDR < 0.05 and 2-fold change in expression. *4ehp* showed upregulation of 35 genes and downregulation of 16 genes, while *gigyf-1* showed 65 genes up and 50 genes down (Figure 3.6). The double mutant, however, showed significantly more dysregulation, with 276 genes upregulated and 255 down. Overlap between all 3 mutant lines was not extensive, with only 5 shared upregulated genes and 5 shared downregulated genes (Figure 3.6, Table 3.1). *4ehp x gigyf-1* upregulated genes were tested for GO Term Enrichment (Figure 3.7) and showed dysregulation of genes related to

multiple stresses and to seed development. The highest term enrichment was for genes involved in seed dormancy (Table 3.2) including well known genes such as *DELAY OF GERMINATION 1* (*DOG1*) (Bentsink et al., 2006). This indicates developmental dysregulation in the absence of the 4EHP-GIGYF complex results in the ectopic expression of genes normally found to be expressed in seeds in seedlings.

Table 3.1. Differential gene expression. Differential expression of genes with overlapping dysregulation in *4ehp*, *gigyf-1*, and *4ehp x gigyf-1* mutants.

		Fold-Change (log2)		
Gene ID	**Gene Name**	*4ehp*	*gigyf-1*	*4ehp x gigyf-1*
AT2G05510	Putative glycine-rich protein	1.34	1.04	1.76
AT2G47770	AtTSPO	1.37	1.14	2.38
AT4G12480	EARLI1	1.70	1.41	2.01
AT4G23700	AtCHX17	1.99	2.76	3.43
AT5G35935	Copia-like retrotransposon family	5.29	5.55	5.53
AT5G54740	SESA5	*-6.29*	*-5.72*	*-2.78*
AT1G75830	PDF1.1	*-3.14*	*-5.45*	*-2.44*
AT4G27160	SESA3	*-4.42*	*-3.42*	*-3.16*
AT3G57645	U2.2	*-1.03*	*-1.41*	*-2.73*
AT5G02540	Short-chain dehydrogenase/reductase family	*-1.33*	*-1.13*	*-1.62*

Table 3.2. Seed dormancy pathway. Genes upregulated in the *4ehp x gigyf-1* double mutant implicated in seed development [Seed dormancy process (GO:0010162)] with GO Term Enrichment identified by AmiGO (Carbon, et al., 2009).

		Fold Change (log2)
Gene ID	**Gene Name**	*4ehp x gigyf-1*
AT3G15670	LEA76	1.32
AT2G25890	Oleosin family protein	2.29
AT5G40420	OLEO2	2.76
AT5G44120	CRA1	2.27
AT4G25140	OLEO1	2.48
AT3G20210	DELTA-VPE	2.06
AT5G45830	DOG1	1.64
AT5G01300	phosphatidylethanolamine-binding family	2.42
AT3G51810	ATEM1	1.84
AT5G62490	HVA22B	2.17
AT2G31980	CYS2	2.01
AT4G28520	CRU3	2.42
AT1G72100	LEA domain-containing protein	2.30
AT2G35300	LEA18	2.27

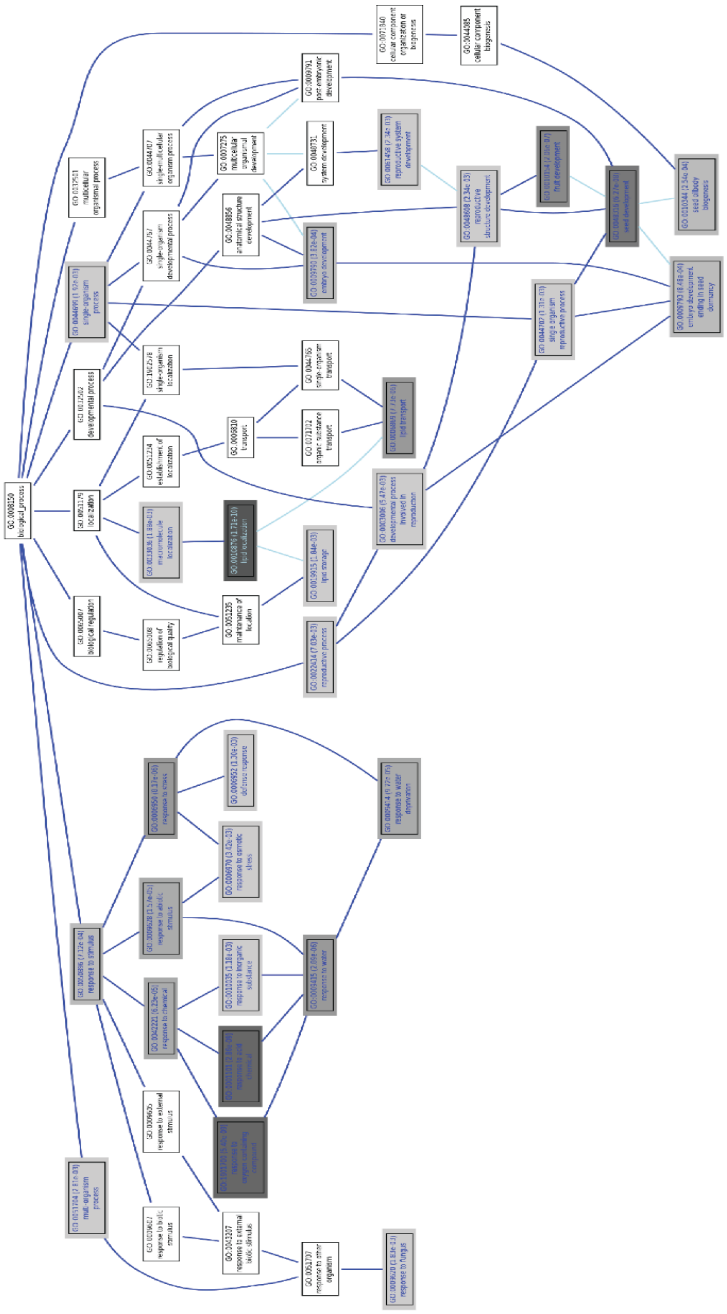


Figure 3.7 (previous page). GO Term Enrichment for *4ehp x gigyf-1* Upregulated Genes

Map of enriched GO Term categories for genes and pathways upregulated in *4ehp x gigyf-1* double mutant plants by AmiGO (Carbon, et al., 2009).

3.3.4. 4EHP RNA Binding Targets

To further examine how 4EHP cap-binding complexes contribute to gene regulation, 4EHP protein was formaldehyde crosslinked to bound RNA in 14-day-old seedlings and immunoprecipitated followed by RNA extraction (RNA IP, or RIP). An RNA sequencing library prepared from the 4EHP RIP was mapped to the *A. thaliana* genome and compared with 3 replicates of Col0 total RNA that had been depleted of rRNA. Differential expression testing showed the enrichment of many ncRNA species including proposed intergenic ncRNA (Zheng et al., 2010), snoRNA (Kim et al., 2010) and recently identified intermediate-sized ncRNA (Wang et al., 2014b) as well as probable pseudogenes and spliced protein-encoding genes. ncRNA were identified in their processed, end form, which together with the lack of reads representing unspliced protein-coding genes, implies 4EHP binds to its target RNAs post-transcriptionally rather than cotranscriptionally.

3.3.5. Subcellular Localization of 4EHP and GIGYF

Based on the results of the 4EHP RIP-Seq showing enrichment of non-coding RNA and snoRNA, we were interested in whether the 4EHP-GIGYF complex had nuclear functions. The only previous report of 4EHP localization in plants, from *A. thaliana* cell culture, showed diffuse localization in both the cytoplasm and nucleus (Bush, et al., 2009). We performed nucleocytoplasmic fractionation followed by western blot for both callus, representing undifferentiated cells, and 14-day-old seedling tissue (Figure 3.8). In both cases, GIGYF was found to primarily localize to the nucleus. Interestingly, 4EHP showed different localization between the two samples: in seedlings, 4EHP was observed to be primarily cytoplasmic, while in callus, it was found in the cytoplasm but appeared to

localize strongly to the nucleus as well. This observation implies that the role of 4EHP and its binding partners may differ between undifferentiated tissue and quiescent, differentiated cells. 4EHP and GIGYF may not be constitutively bound to each other, as in seedlings they appear to localize to different cellular compartments, though they will co-immunoprecipitate as a complex. Given these observations, it seems possible that 4EHP has a nucleocytoplasmic shuttling role similar to what has been recently observed in human cell culture (Kubacka et al., 2013).

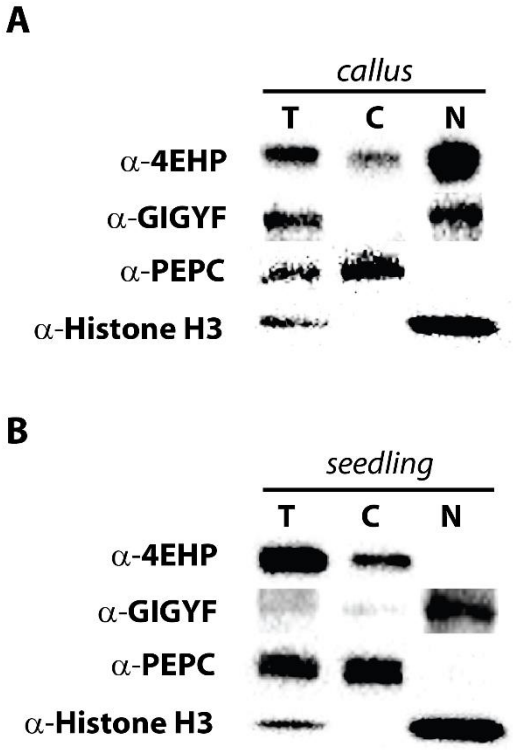


Figure 3.8. Nucleocytoplasmic Fractionation.

(A) Fractionation of *A. thaliana* callus into Total (**T**), Cytoplasmic (**C**), and Nuclear (**N**) fractions. Fractions (10 μL aliquots) were tested for compartmentalization of 4EHP and GIGYF by SDS-PAGE and western blot. Phosphoenolpyruvate carboxylase (PEPC) and Histone H3 serve as cytoplasmic and nuclear controls, respectively. **(B)** Nucleocytoplasmic fractionation of 14-day-old seedling tissue.

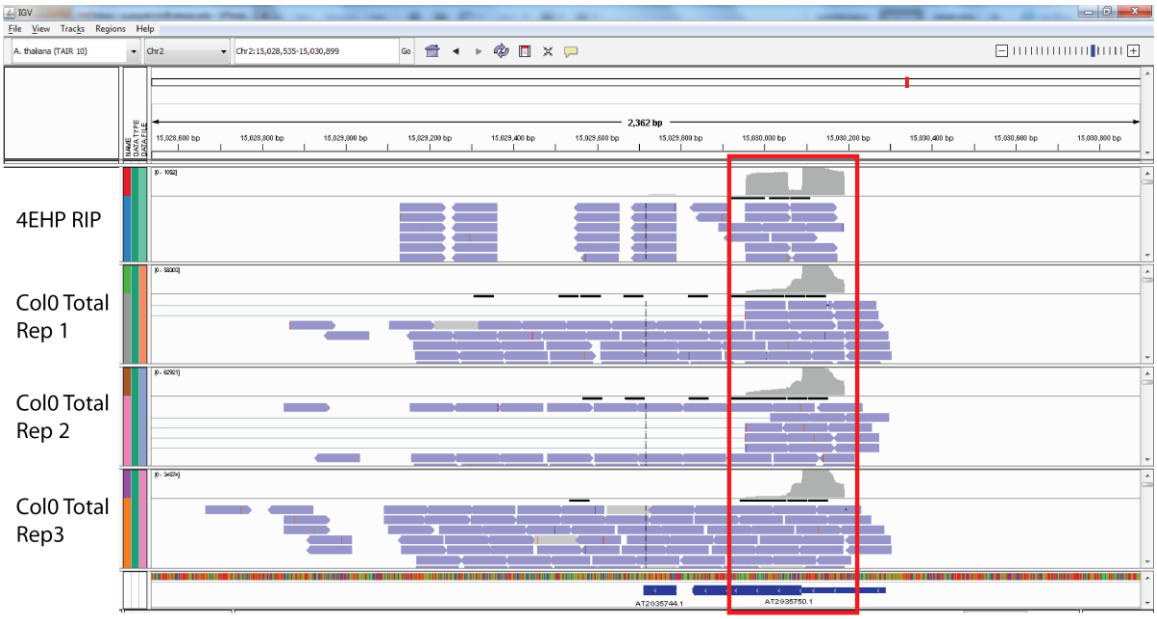


Figure 3.9. 4EHP RNA Immunoprecipitation of *HID1*

4EHP RIP-Seq (top) read coverage compared to three Col0 total RNA replicates in IGV (Thorvaldsdóttir, et al., 2012). The *HID1* transcript is highlighted.

3.3.6. ncRNA binding partners of 4EHP-GIGYF

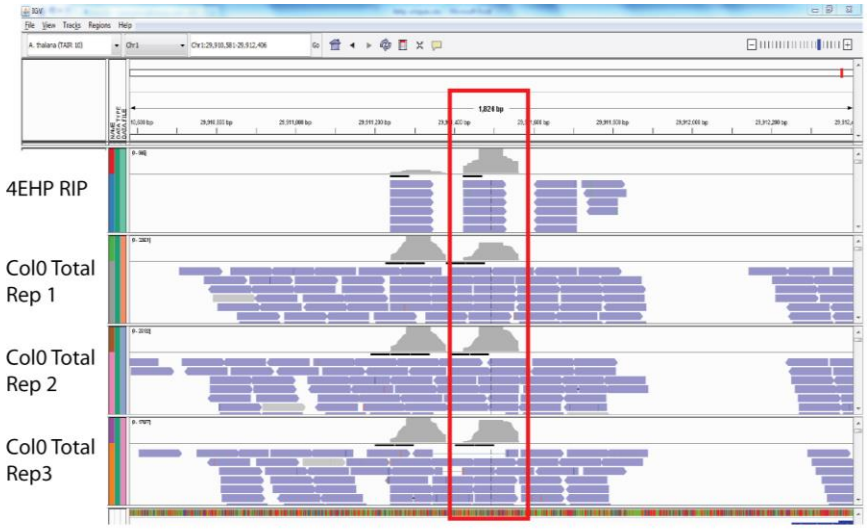
We looked for targets in the 4EHP RIP-Seq of interest that might be able to shed light on the role of the 4EHP-GIGYF complex. While many of the binding targets are presumed to be snoRNA or snRNA, several non-coding RNA of unknown function are present and highly enriched by 4EHP RIP. One identified RNA is *HID1* (Figure 3.9). *HID1* is a 236 nucleotide ncRNA that is reported to have nuclear localization and bind to chromatin as part of a larger RNP that presumably regulates gene expression (Wang et al., 2014a). In the interest of identifying other ncRNA, we turned to a recently published study

which identified and annotated 838 intermediate-sized ncRNA of a similar class to *HID1* (Wang, et al., 2014b). Cross referencing with this ncRNA list, another target we found of interest is *nc2749* which was strongly enriched by the 4EHP RIP-Seq (Figure 3.10a). *nc2749* is 140 nucleotides in length and BLAST searches of other plant genomes and alignment indicate strong sequence conservation as well as lack of a consistent ORF or snRNA/snoRNA characteristics (Figure 3.10b). Additionally, *HID1* and *nc2749* share the important characteristic of being at the 5' end of a transcript that is processed into multiple smaller ncRNA; it is presumed that in such cases the 5'-most ncRNA is cotransciptionally capped and 3' processed ncRNAs may be less likely to have a cap structure (Kim, et al., 2010; Wang, et al., 2014b). Therefore, the specific enrichment of the likely capped *HID1* and *nc2749,* but not the downstream ncRNA is a positive sign for the accuracy of the RIP-Seq results representing direct targets of 4EHP, of which the presence of a cap is presumably an important determinant.

To confirm the RIP-Seq we performed RIP followed by RT-PCR of putative binding targets compared to the non-specific control transcript *chlorophyll A/B binding protein (CAB)* (Figure 3.11). Additionally, to control for off target or promiscuous binding of antibody that may lead to non-specific presence of RNA in the RNA-Seq, we performed control crosslinking and RIP in knockout lines for the proteins. 4EHP RIP-RT-PCR revealed the presence of *HID1* and *nc2749* transcript only in the reverse transcribed RNA IP of Col0 extract (Figure 3.11A), supporting their status as true 4EHP binding targets. Another third binding target, *nc2506*, was either not present or too low to detect, while a fourth possible target, *snor82*, showed non-specific RNA IP in the absence of 4EHP protein. GIGYF RIP-RT-PCR gave similar results to 4EHP, as *HID1* and *nc2749* transcripts were again only detected in the reverse transcribed RNA IP of Col0 extract (Figure 3.11B). In contrast with 4EHP, however, *snor82* showed much weaker non-specific

RNA IP in the absence of GIGYF protein, while *nc2506* was again not detected. These observations support *HID1* and *nc2749* as being true binding targets and constituents of 4EHP-GIGYF RNP complexes in *A. thaliana*.

A



B

```
Osa.nc2749      -AAGGCAATAAATTAGAGTCTCTGATGCGCT---CATTTCAATAGCGTTGTCAGCGCTGT 56
Ptr.nc2749      AAAGGCAACAAATTAGAGTCTCTGAAGTGCTTTAAATTTCAATAGCGTTTTCAGTGCTGT 60
Ath.nc2749      AAAGGCAACAAATTAGAGTCTCTGATTCACT---TATGTCAATAGTGGATAAAGCGCTGT 57
                 ******* ****************:  .**    ** ******* * : :.** *****

Osa.nc2749      ACGTTGCCGAAACGAAT--CGG--GCAAAGATTCTGGCACTCG---CCATCTCATATATG 109
Ptr.nc2749      ACGTTGCCGACACCAATTTTGGGTGCTGAGATTCTGGCATTTGCTTCCGTTTCATTGAAG 120
Ath.nc2749      ACGTTGCCGAGATAGAT-ATGG----CGATATTCTGGTACTTGGACTAGTTTTGATAGTT 112
                ********** *  .**   **     .* ******* * * *    ..* * .:: .:

Osa.nc2749      GCTTGTTGCATTCAACTTGTCCATACAGTT 139
Ptr.nc2749      -CTT-ATGCATTCAGATGTTCC-------- 140
Ath.nc2749      CCAT-GTGCATTCAAAAG-CCCTTACATGT 140
                 *:*  ********..:   **
```

Figure 3.10. 4EHP RNA immunoprecipitation of *nc2749*

(A) 4EHP RIP (top) read coverage compared to three Col0 total RNA replicates in IGV (Thorvaldsdóttir, et al., 2012). The *nc2749* transcript is highlighted. **(B)** Alignment of the *A. thaliana nc2749* sequence with conserved sequences present in rice and poplar genomes using ClustalW2 (Larkin, et al., 2007).

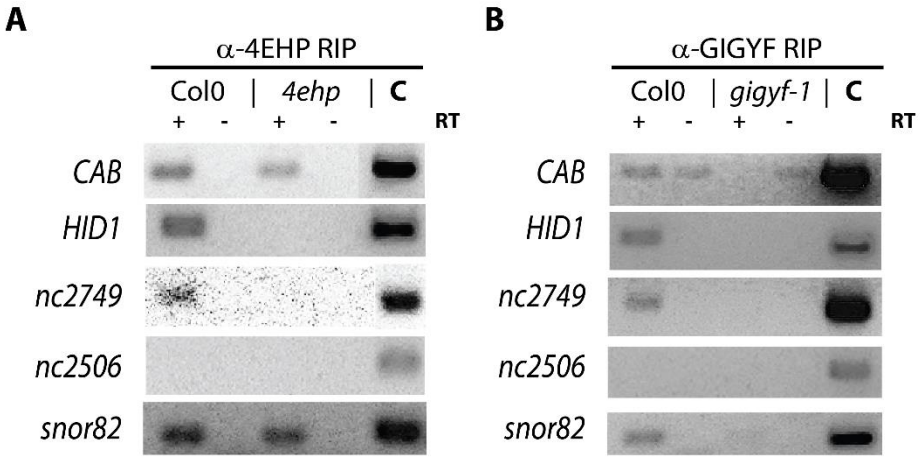


Figure 3.11. RIP-RT-PCR

(A) RT-PCR of 4EHP RNA immunoprecipitation. RNA immunoprecipitated with 4EHP antibody from crosslinked Col0 or *4ehp* seedling extract was reverse transcribed (**+ or - RT**) with random oligos into cDNA and tested by PCR with primers to indicated transcripts. Col0 oligo(dT) primed cDNA was used as control, (**C**). **(B)** RT-PCR of GIGYF RNA immunoprecipitation, performed similarly.

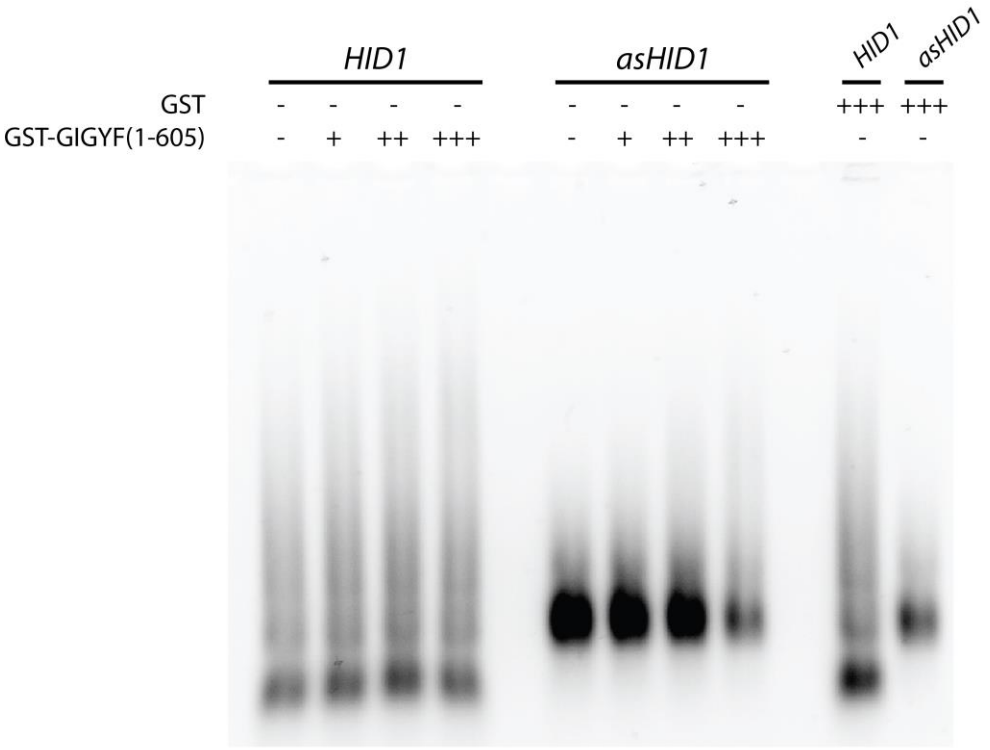


Figure 3.12. GST-GIGYF$_{1-605}$ EMSA of Fluorescent RNA

Fluorescently labeled *HID1* or its antisense counterpart *asHID1* was mixed with increasing amounts of GST-GIGYF$_{1-605}$ and then separated by electrophoresis on a 1% agarose gel in TBE buffer. GST was tested as a control for non-specific binding.

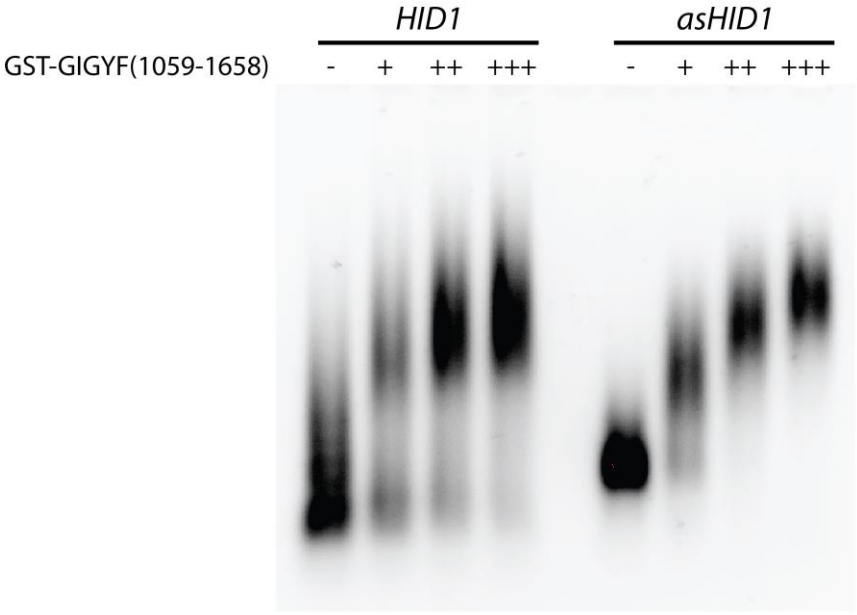


Figure 3.13. GST-GIGYF$_{1059-1658}$ EMSA of Fluorescent RNA

Fluorescently labeled *HID1* or its antisense counterpart *asHID1* was mixed with increasing amounts of GST-GIGYF$_{1059-1658}$ and then separated by electrophoresis on a 1% agarose gel in TBE buffer.

3.3.7. GIGYF RNA Binding Activity

The activity and specificity of 4EHP is assumed to be fairly limited to binding 7-methylguanosine cap structures and is unlikely to explain the choice of *in vivo* binding partners. Based on our observations, we believed that GIGYF may play a role in specificity to ncRNA targets of the complex, as it appears to have two large stretches that have

predicted RNA binding ability; one surrounding the 4EHP-binding site at the N-terminal end, and another C-terminal of the GYF domain. To test the RNA binding ability and specificity of GIGYF, we generated fluorescently labeled *HID1* and antisense *asHID1* RNA *in vitro* using Fluorescein-12-UTP. Binding of the RNA to increasing concentrations of GST-tagged subdomains of GIGYF and to GST was tested by electrophoretic mobility shift assay (EMSA). Neither the GST control nor the GST-tagged N-terminal portion, GST-GIGYF$_{1\text{-}605}$ showed RNA binding ability *in vitro* (Figure 3.12). GST-GIGYF$_{1059\text{-}1658}$ bound to *HID1*, but also bound equally well to its antisense counterpart *asHID1* (Figure 3.13), indicating non-specific binding ability for GIGYF *in vitro*.

3.4 CONCLUSIONS

This work sought to better understand the role of 4EHP in plants, which do date was ill-defined in spite of its strong conservation throughout the plant lineage and relationship to the animal 4EHP protein. We have demonstrated that 4EHP forms a homologous cap-binding complex to the 4EHP-GIGF2 complex observed in mammals. The 4EHP protein of animals has been primarily been implicated in regulation of translation: a negative regulator in embryogenesis (Cho et al., 2006), and a positive regulator under stress in cell culture (Uniacke et al., 2012). However, recent work has also opened up the possibility of other roles, such as the observation that 4EHP undergoes nucleocytoplasmic shuttling (Kubacka et al., 2013). This work supports nuclear localization of 4EHP in a tissue or developmental stage specific manner.

The mechanisms of translational regulation by cap-binding complexes in plants remain unclear. While plants have two separate conserved cap-binding complexes in eIF4F

and eIFiso4F, they lack the 4E-BPs that are a crucial mechanism of regulation in animals and yeast. 4EHP was (and remains) a candidate for a competing translational regulator in plant systems, however several observations made in this work make this seem less likely this is either the primary or sole operating mechanism of 4EHP. First, it forms a complex with GIGYF, which is observed to primarily localize to the nucleus by nucleocytoplasmic fractionation. Second, GIGYF also has the curious property of potentially self-oligomerizing, which does not rule out negative translational regulation but does point away from a role in positive translational regulation as it would tend to form mRNP aggregates that would not be accessible to other initiation factors or to ribosomes. Third, RIP-Seq results did not return attractive candidates for mRNAs that are subject to translational regulation but instead indicated that 4EHP binds a sizable number of ncRNAs. It remains possible that some of these ncRNA species may have roles in translational regulation; however, the ncRNAs observed in 4EHP RIP-Seq such as *HID1* have primarily been assigned nuclear roles, though a dual role cannot be ruled out.

The 4EHP-GIGYF complex is an important developmental regulator. Single mutants of *4ehp* show only a mild phenotype, implying that GIGYF is somewhat functional in its absence, either operating without the cap-binding subunit or forming a mixed complex with another cap-binding protein. Single mutants of *gigyf* also show minor effects, implying that 4EHP may operate somewhat redundantly through other binding partners; however, it is also possible that *gigyf-1* is hypomorphic, as the 5' portion of the *GIGYF* transcript is still present in the RNA sequencing. Double mutants have an additive and more severe phenotype and a large degree of gene dysregulation, indicating the intact complex plays important developmental roles.

The 4EHP RIP-Seq indicates that 4EHP binds to a number of intermediate-sized ncRNA in *A. thaliana*. One, which we confirmed by RIP-RT-PCR, is *HID1*, a recently

described ncRNA with a nuclear gene regulatory role . *HID1* has been proposed to bind to the promoter of *PIF3* to contribute to repression of *PIF3* transcription under red light. A second ncRNA we validated as a 4EHP binding target is *nc2749*, an undescribed but conserved ncRNA of unknown function. We attempted to determine whether GIGYF is involved in discriminating the binding partners of the 4EHP-GIGYF complex through EMSA. We found that GIGYF has RNA binding capacity in its C-terminal portion, but that this activity for binding *HID1* was not sequence or secondary structure specific, as it also bound the antisense transcript *asHID1*.

This work was able to validate a second form of 7-methylguanosine cap-binding complex that is present and conserved from animals to plants, paralleling the eIF4F complex. While the roles of the 4EHP-GIGYF complex remain less clear than those of eIF4F, this work makes significant progress toward determining the properties of the complex in plants.

3.5. FUTURE DIRECTIONS

The observation of 4EHP as potentially having nuclear and cytoplasmic roles leaves many areas open for future investigation. One potential role that would be consistent with these observations is the possibility that 4EHP is involved in trafficking of mRNPs, possibly in both directions. Another possibility is that 4EHP is involved in some post-transcriptional processing step for ncRNA. Given the observation that 4EHP binds to *HID1*, which is reported to bind to chromatin to regulate transcription, a primary area of future investigation will be whether 4EHP functions together with *HID1* in regulating its transcriptional targets, such as *PIF3*. It is possible that 4EHP is solely responsible for some type of transport or processing step of the *HID1* ncRNA. It is also possible that the 4EHP-

GIGYF complex together with *HID1* is targeted to chromatin. Future work will focus on whether 4EHP and GIGYF work in a gene regulatory pathway with *HID1* and whether they are indeed bound at the promoters of *HID1*-regulated genes.

It will also be of interest to investigate the *nc2749* ncRNA identified by this work as an evolutionarily conserved transcript that presumably has an important function. Toward this goal, a knockdown or knockout line for this ncRNA should be generated in order to observe its function and probe for gene dysregulation in its absence. It will also be possible to observe the subcellular localization of *nc2749* to determine whether it may have similar roles in transcriptional regulation as *HID1*. If these lines of investigation are promising, it should be possible to find whether the RNP of 4EHP-GIGYF and *nc2749* behaves similarly to *HID1*, possibly at the chromatin level, representing a ubiquitous mechanism for gene regulation in which different regulatory targets are controlled by different ncRNA (Wang et al., 2014b) .

An interesting possibility given these observations of a putative RNP of 4EHP-GIGYF and ncRNA that could be present at chromatin is that GIGYF may serve as an adaptor scaffold for the ncRNA through its protein-protein interacting GYF domain. The putative [PPGF] motif bound by GIGYF (Kofler et al., 2005), is found to be conserved in a number of proteins that have roles in transcriptional regulation, including histone modifying factors and transcription factors. Future investigation into the binding partners of GIGYF should give insight into how the 4EHP-GIGYF ncRNA complex may link the chromatin to transcriptional regulatory effects through a non-coding RNA bridge.

Further experiments along these lines would also have ample opportunity to discover more unexpected findings and branch off into new directions. As 4EHP and GIGYF have relevance to human health, and have been implicated in the pathogenesis of Parkinsons disease, autism, and endometrial cancer (Le Gallo and Bell, 2014), and

knockouts in mice cause lethality (Giovannone et al., 2009) (Morita et al., 2012), the advantage to studying the role of this conserved complex in plants is that *4ehp x gigyf-1* plants are developmentally impacted but viable. Thus, *A. thaliana* may provide a powerful model system for investigating how the 4EHP-GIGYF complex and its ncRNA binding partners contribute to development and gene regulation across all eukaryotes.

Appendix

```
Arabidopsis.thaliana.4E           ------------------------------------MAVEDTPKS--VV 11
Arabidopsis.lyrata.4E             ------------------------------------MAVEDTPKS--VV 11
Capsella.rubella.4E               ------------------------------------MAVEDTPKS--VV 11
Thellungiella.halophila.4E        ------------------------------------MAVEDTLKP--VV 11
Brassica.rapa.4E1                 ------------------------------------MAVEDTSKP-VVV 12
Brassica.oleracea.4E              ------------------------------------MAVEDTSKP-VVV 12
Raphanus.raphanistrum.4E          ------------------------------------MAVEDTSKPLVVV 13
Raphanus.sativus.4E               ------------------------------------MAVEDTSKPVVVV 13
Brassica.rapa.4E2                 ------------------------------------MAVEDTLKP-NVA 12
Brassica.napus.4E                 ------------------------------------MAVEDTLKP-NVP 12
Sisymbrium.irio.4E                --------------------------------------------------
Solanum.tuberosum.4E              ------------------------------------MAAAEMERTTSFD 13
Solanum.lycopersicum.4E           ------------------------------------MAAAEMERTMSFD 13
Zea.mays.4E                       ------------------------------------MAEETDTRPASAG 13
Setaria.italica.4E                ------------------------------------MADETDTRPTSAG 13
Sorghum.bicolor.4E                ------------------------------------MADEIDTRPASAG 13
Brachypodium.distachyon.4E        ------------------------------------MAEDTETRPASAG 13
Triticum.aestivum.4E              ------------------------------------MAEDTETRPASAG 13
Oryza.sativa.4E                   ------------------------------------MAEEHETRPPSAG 13
Citrus.sinensis.4E                ------------------------------------MAGESDNPRTE-- 11
Citrus.clementina.4E              ------------------------------------MAGESDNPRTE-- 11
Manihot.esculenta.4E              MDTLGQINLDLSPLSLSITYRTESKKQQQAAEKINAKMAAEEPLKSTT-- 48
Ricinus.communis.4E               ------------------------------------MAVEEPQKLTIPN 13
Populus.trichocarpa.4E            ------------------------------------MDVEDPQKLAT-- 11
Glycine.max.4E                    ------------------------------------MVVEDAQKSAITE 13
Medicago.truncatula.4E            ------------------------------------MVVEETPKS--ID 11
Prunus.persica.4E                 ------------------------------------MVVEDAPKT-SAS 12
Fragaria.vesca.4E                 ------------------------------------MVVEEAAKTISAS 13
Cucumis.sativus.4E                ------------------------------------MVVEDTIK----- 8
Cucumis.melo.4E                   ------------------------------------MVVEDSMK----- 8
Eucalyptus.grandis.4E             ------------------------------------MVVEETVKSAAAA 13
Mimulus.guttatus.4E               -------------------------------------MVEETEKS---- 8
Aquilegia.coerulea.4E             ------------------------------------MVVAE--ES---- 7
Theobroma.cacao.4E                ------------------------------------MGVEENLKS---- 9
Eucalyptus.grandis.4E1BL          ------------------------------------MSIGDAATSRPAA 13
Fragaria.vesca.4E1BL1             ------------------------------------MASFSTLRGSAEW 13
Fragaria.vesca.4E1BL2             ----------------------------------------------HSWTLW 6
Arabidopsis.thaliana.4E1b         ------------------------MVVT-DSPVSGIMA-DQNIDPNTTT 23
Arabidopsis.thaliana.4E1c         ------------------------MVVM-DSPVSGRMA-DQNIDPNTTT 23
Arabidopsis.lyrata.4E1b           ------------------------MVVV-DSSVSAIMA-DQNIDPNTAT 23
Arabidopsis.lyrata.4E1c           --------------------------------------------------
Capsella.rubella.4E1c             ------------------------MVFE-DSSVSALLA-VENIDPNTTT 23
Capsella.rubella.4E1b             ------------------------MVVE-DSSVSVILA-VENIDPNTTT 23
Raphanus.sativus.4E1b             ------------------------MAVE-SSSLPAIMAEEENLDPNTTS 24
Brassica.oleracea.4E1b            ------------------------MAVE-SSSLPAIMAEEENLDPNTPN 24
Brassica.rapa.4E1b                ------------------------MAVE-SSSLPAIMAEEENLDPNTTN 24
Brassica.napus.4E1b               ------------------------------------MAEEENLDPNTTN 13
Raphanus.raphanistrum.4E1b        ------------------------------------MAEEENLDPNTTS 13
Brassica.napus.4E1c               ------------------------MAVE-GSSVSAIMAEEENLDPNTAN 24
Brassica.oleracea.4e1c            ------------------------MAVE-GSSVSAIMAEEENLDPNTAN 24
Brassica.rapa.4E1c                ------------------------MTVE-RSSVSAIMAEEENLDPNTAN 24
Thellungiella.parvula.4E1b        ------------------------MAVE-GSSVSANMAEEENLDPNTAK 24
Thellungiella.halophila.4E1b      ------------------------MVVEEDSSVSTIMA-EENLDPNTAN 24
Sisymbrium.irio.4E1b              --------------------------------------------------
```

```
Arabidopsis.thaliana.4E            TEEAKPNSIE--------NPIDRYHEEGDDAEEGEIAGGEGDGN---VDE 50
Arabidopsis.lyrata.4E              TEEANPNSIE--------NPIDRYHEEGDDAEEGEIVGGDGDGN---VDE 50
Capsella.rubella.4E                NEEANPNSKE--------NPIDRYHEEGDDAEEGEIVGGEGHGD---VDE 50
Thellungiella.halophila.4E         TEEANPNSTE--------KPIDRYHEEGDDAEEGEIVGGEGDGD---VDE 50
Brassica.rapa.4E1                  AEEANPNPTD--------HPIDRYHEEGDDTEEGEIAGGEGDG-----DE 49
Brassica.oleracea.4E               AEEANPNPTD--------HPIDRYHEEGDDAEEGEIAGGETDG-----DE 49
Raphanus.raphanistrum.4E           AEEANPNPTD--------QPIDRYHERGDDAEEGEIG----DG-----DE 46
Raphanus.sativus.4E                AEEANPNPTD--------QPIDRYHEQGDDAEEGEIG----DG-----DE 46
Brassica.rapa.4E2                  TEESNPNSAD--------HPIDRYHEEGDDAEEGATV-----------DE 43
Brassica.napus.4E                  TEESNPNSVD--------HPIDRYHEEGDDAEEGAIV-----------DE 43
Sisymbrium.irio.4E                 --------------------------------------------------
Solanum.tuberosum.4E               AAEKLKAADG------------GGGEVDDELEEGEIVEESND-----TAS 46
Solanum.lycopersicum.4E            AAEKLKAADG------------GGGEVDDELEEGEIVEESND-----TAS 46
Zea.mays.4E                        SRGRP--------------------APEDDDREEGEITDLA-------CAP 37
Setaria.italica.4E                 SRGRP--------------------APDDDDREEGEIADDS-------SAP 37
Sorghum.bicolor.4E                 SRGRPAH-----------------ATEDDDREEGEIADDT-------PAP 39
Brachypodium.distachyon.4E         ---------------------------AEEREEGEIADDD-------SAA 29
Triticum.aestivum.4E               ---------------------------AEEREEGEIADDGDG-----SSA 31
Oryza.sativa.4E                    RPPSSGRG---------------RADDADEREEGEIADDDSG-----HAP 43
Citrus.sinensis.4E                 -DQTNTKNPN-------PS-------EEEELEEGEIVG--DD---ES-SK 40
Citrus.clementina.4E               -DQTNTKNPN-------PS-------EEEELEEGEIVG--DD---EP-SK 40
Manihot.esculenta.4E               -EETPNPNLN-------SNPRAQDDVNDDEPEEGEIVG--DE---ESSAK 85
Ricinus.communis.4E                SEETPNRNPN-------PNP---SDVKEDEVEEGEIVGGEGE---EESTA 50
Populus.trichocarpa.4E             -EKTPN--PN----------------TEDDLEEGEIVAGGDD---DSSLK 39
Glycine.max.4E                     DQNPSRVAND-------NN-------DDEDLEEGEIPDDGDD-GASATSK 48
Medicago.truncatula.4E             DQITNPKIED-------IN---------NDLEEGEIN--DDD--SSAISK 41
Prunus.persica.4E                  EDQAKTETN--------PKP----REEDDEPEEGEIVGD-EE----SASK 45
Fragaria.vesca.4E                  EDQSKTESNN-------PKS----MQDENELEEGEIVGGGDD----ESSK 48
Cucumis.sativus.4E                 -ATSAEDLSN---SIANQNPRGRGGEEDEELEEGEIVGDDDL---DSSNL 51
Cucumis.melo.4E                    -ATSAEDLSN---SIANQNPRGRGGDEDEELEEGEIVGDDDL---DSSNL 51
Eucalyptus.grandis.4E              AAATADDQNNPAAANPNPNPKGVAGSGDPELEEGEIAGEEDL---AEQSA 60
Mimulus.guttatus.4E                -ASTAIKEEAENLTSMMKNHLNLG-EVEPEAEEVEIVGGDPN---HDAKA 53
Aquilegia.coerulea.4E              -LEMR----------------NKEEEEEEMPIEDKRVEKILE---SSSSS 37
Theobroma.cacao.4E                 -LSISEEGNK-------NPNPNVKEDEEEEPEEGEIVGEEDD---STSSS 48
Eucalyptus.grandis.4E1BL           HEDRP-----------------------PVNADSDPRRFLHGGGGNRSGR 40
Fragaria.vesca.4E1BL1              VKEEV-----------------------AALLRRDPT------------- 27
Fragaria.vesca.4E1BL2              VGNPM-----------------------ANSEQED--------------- 18
Arabidopsis.thaliana.4E1b          SPSPKEKHVSAIKAI-SGD-------EKAPSKE----KKN------YASK 55
Arabidopsis.thaliana.4E1c          SPSPIEKHVSAIKAI-SGD-------EKAPSKE----KKN------YASK 55
Arabidopsis.lyrata.4E1b            NPSRQEKHVPAIKAI-SGD-------EKSPSKE----KKNDA----YASK 57
Arabidopsis.lyrata.4E1c            -------------------------------MEK----KSNDY----YALK 12
Capsella.rubella.4E1c              NSSLKEKHFPAIKGAISGD-------EEGLSKE----KKRNE----DASK 58
Capsella.rubella.4E1b              NPILKEKYVPAIKAISSGD-------YEGPSKEETIISSGNY-----FASK 62
Raphanus.sativus.4E1b              LIRIEK-YVPAIKAICGGGG-----GEEGPSKGKGIMCGG-------KKS 61
Brassica.oleracea.4E1b             LIRIEK-HVPAIKAICGG-------GDEGPSKEKKIMCGG-------KKS 59
Brassica.rapa.4E1b                 LIRIEK-HVPAIKAICGG-------GDEGPSKEKKIMCGG-------EKS 59
Brassica.napus.4E1b                LIRIEK-HVPAIKAICGG-------GDEGPSKEKKIMCGG-------EKS 48
Raphanus.raphanistrum.4E1b         LIRLEK-HVPAIKAICGG--------GEGPSKEK----GI-------KKS 43
Brassica.napus.4E1c                HCPIQK-HLPAIKAIGGG--------EEGPSKEQKITCGGNGYV--WKKS 63
Brassica.oleracea.4e1c             HCPIQK-HLPAIKAIGGG--------EEGPSKEQKITCGGNGYV--WKKS 63
Brassica.rapa.4E1c                 HCPIQK-HLPAIKAIRGG--------EEGPSKEQKITCGGNGYV--WKKS 63
Thellungiella.parvula.4E1b         PSFVEK-YFPAIRAICGD--------DEGPSKEKKIMCGGKDNE--SVKS 63
Thellungiella.halophila.4E1b       PSRIEKQHVPAIKAIRGG--------EEGPSKGKTILCGGKDN----VSK 62
Sisymbrium.irio.4E1b               --------------------------------------------------
```

```
Arabidopsis.thaliana.4E            SSKSGVPES----------------HPLEHSWTFWFDNPAVK---SKQT 80
Arabidopsis.lyrata.4E              SSKSAVPQS----------------HPLEHSWTFWFDNPSVK---SKQT 80
Capsella.rubella.4E                SNKSAAPQS----------------HPLEHSWTFWFDNPAVK---SKQT 80
Thellungiella.halophila.4E         SGKSAVPES----------------HQLEHSWTFWFDNPSVK---SKQT 80
Brassica.rapa.4E1                  SSKSAVPQS----------------HPLEHSWTFWFDNPSVK---LKQA 79
Brassica.oleracea.4E               SSKSAVPQS----------------HPLEHSWTFWFDNPSVK---LKQA 79
Raphanus.raphanistrum.4E           SSKSAVPQS----------------HPLEHSWTFWFDNPSVK---LKQT 76
Raphanus.sativus.4E                SSKSAVPQS----------------HPLEHSWTFWFDNPSVK---LKQT 76
Brassica.rapa.4E2                  SSKSAVPES----------------HPLEHSWTLWFDNPSVK---SKQT 73
Brassica.napus.4E                  SSKSAVPES----------------HPLEHSWTLWFDNPSVK---SKQT 73
Sisymbrium.irio.4E                 --------------------------------------------------
Solanum.tuberosum.4E               YLGKEITVK----------------HPLEHSWTFWFDSPIAK---SRQT 76
Solanum.lycopersicum.4E            YLGKEITVK----------------HPLEHSWTFWFDNPTTK---SRQT 76
Zea.mays.4E                        SP----PAT----------------HPLEHSWTFWFDNPQSK---SKQA 63
Setaria.italica.4E                 APP-LQPAT----------------HPLEHSWTFWFDNPQGK---SKQA 66
Sorghum.bicolor.4E                 AL----PVT----------------HPLEHSWTFWFDNPQGK---NKQA 65
Brachypodium.distachyon.4E         LAQ-ANNGP----------------HPLEHAWTFWFDNPQGK---SRNA 58
Triticum.aestivum.4E               AAA-GRITA----------------HPLENAWTFWFDNPQGK---SRQV 60
Oryza.sativa.4E                    PQA-NPAAP----------------HPLEHAWTFWFDNPQGK---SKQA 72
Citrus.sinensis.4E                 NSTAVMQQP----------------HPLEHSWTFWFDNPFAK---SKQA 70
Citrus.clementina.4E               NSTAVMQQP----------------HPLEHSWTFWFDNPFAK---SKQA 70
Manihot.esculenta.4E               KSSAVTYQP----------------HPLEHQWTFWFDNPTAK---SKQA 115
Ricinus.communis.4E                KS-ALIYEA----------------HPLEHQWTFWFDNPSAK---SKQA 79
Populus.trichocarpa.4E             KSVSLPYQP----------------HPLEHQWTFWFDNPSAK---SKQA 69
Glycine.max.4E                     PPSALVRNP----------------HPLENSWTFWFDNPSAK---SKQA 78
Medicago.truncatula.4E             PLTAGHQS-----------------HPLENSWTFWFDNPQTK---SKQQ 70
Prunus.persica.4E                  PSKGIAPQS----------------HALEHSWTFWFDSPAAKSAKTKQE 78
Fragaria.vesca.4E                  MSK---PQQ----------------HPLEHSWTFWFDIPSSKPGKSKQE 78
Cucumis.sativus.4E                 T-AALVHQP----------------HPLEHSWTFWFDNPSAK---SKQA 80
Cucumis.melo.4E                    S-ASLVHQP----------------HPLEHSWTFWFDNPSAK---SKQA 80
Eucalyptus.grandis.4E              KEEPKADQP----------------HPLEHSWTFWFDNPAAK---SKQM 90
Mimulus.guttatus.4E                MAAAAAAPPP---------------RHPLEHSWTFWFDNPSAK---NKQA 85
Aquilegia.coerulea.4E              SGLVVNQQPSILF------------LHPLEHSWTFWFDNPSGK---SKQK 72
Theobroma.cacao.4E                 SKKGVVEQP----------------HPLEHSWTFWFDNPSAK---SKQA 78
Eucalyptus.grandis.4E1BL           AEPEIAREARSLSKRAPGEGVTVGRPHQLGHSWTFWFDSPAAK---SSQA 87
Fragaria.vesca.4E1BL1              ---------------------------IPYKVATTIAYELFGFP---YIIL 48
Fragaria.vesca.4E1BL2              ----------------------------------------WGG-------- 21
Arabidopsis.thaliana.4E1b          KSTTVIQKS----------------HCFQNSWTFWFDNPSSK---SNQV 85
Arabidopsis.thaliana.4E1c          KSTTVIQKS----------------HCFQNSWTFWFDNPSSK---SNQV 85
Arabidopsis.lyrata.4E1b            KSTTVIQKS----------------HCFQNSWTFWFDNPSSK---SNQV 87
Arabidopsis.lyrata.4E1c            KSTTVIQKS----------------HFFQSSWTFWFDNPSSK---SNQV 42
Capsella.rubella.4E1c              KSTTVIHKS----------------HLFENSWTFWFDNPSSK---SNQV 88
Capsella.rubella.4E1b              KSTTVIQKS----------------HLFENSWTFWFDTPSSK---SNQV 92
Raphanus.sativus.4E1b              KSTTAIEPS----------------HSFQNSWTFWFDNPSSK---SSQA 91
Brassica.oleracea.4E1b             KSNTAVHPS----------------HSFRNSWTFWFDNPSSK---SNQA 89
Brassica.rapa.4E1b                 KSNTAVKPW----------------HSFQNSWTFWFDNPSSK---SNQT 89
Brassica.napus.4E1b                KSNTAVKPW----------------HSFQNSWTFWFDNPSSK---SNQT 78
Raphanus.raphanistrum.4E1b         RSTSAIEPS----------------HSFQNSWTFWFDNPSSK---SNQA 73
Brassica.napus.4E1c                KSTTVIEHS----------------HALQSSWTFWFDNPSSK---SNQA 93
Brassica.oleracea.4e1c             KSTTVIEHS----------------HALQSSWTFWFDNPSSK---SNQA 93
Brassica.rapa.4E1c                 KSTTVNQHS----------------HSFQSSWTFWFDTPSSK---SNQT 93
Thellungiella.parvula.4E1b         KSTTVIQRS----------------HSFQNSWTFWFDNPSSK---SNQA 93
Thellungiella.halophila.4E1b       KSTTVIQNS----------------HSFQNSWTIWFDNPSSK---SYQA 92
Sisymbrium.irio.4E1b               --------------------------------------------------
```

```
Arabidopsis.thaliana.4E            SWGSSLRPVFTFSTVEEFWSLYNNMKHPSKLAHGADFYCFKHIIEPKWED 130
Arabidopsis.lyrata.4E              TWGSSLRPVFTFSTVEEFWSLYNNMKHPSKLAHGADFYCFKHNIEPKWED 130
Capsella.rubella.4E                TWGSSLRSVFTFSTVEEFWSLYNNMRHPSKLAHGADFYCFKHNIEPKWED 130
Thellungiella.halophila.4E         TWGSSLRSVFTFSTVEEFWSLYNNMRHPSKLAGGADFYCFKHNIEPKWED 130
Brassica.rapa.4E1                  TWGSSLRSVFTFSTVEEFWSLFNNMKGPSKLAGGADFYCFKHNIEPKWED 129
Brassica.oleracea.4E               TWGSSLRSVFTFSTVEEFWSLFNNMKGPSKLAGGADFYCFKQNIEPKWED 129
Raphanus.raphanistrum.4E           TWGSSLRSVFTFSTVEEFWGLYNNMRVPSKLVTGADFYCFKHNIEPKWED 126
Raphanus.sativus.4E                AWGSSLRSVFTFSTVEEFWGLYNNMRVPSKLVTGADFYCFKHNIEPKWED 126
Brassica.rapa.4E2                  TWGSSLRSVFTFSTVEEFWSLYNNIRHPSKLANGADLYCFKHNIEPKWED 123
Brassica.napus.4E                  TWGSSLRSVFTFSTVEEFWSLYNNIRHPSKLANGADLYYFKHNIEPKWED 123
Sisymbrium.irio.4E                 ----------------FCSLYNNMRHPSKLAHGADFYCFKHNIEPKWED 33
Solanum.tuberosum.4E               AWGSSLRNVYTFSTVEDFWGAYNNIHHPSKLVMGADFHCFKHKIEPKWED 126
Solanum.lycopersicum.4E            AWGSSLRNVYTFSTVEDFWGAYNNIHHPSKLIMGADFHCFKHKIEPKWED 126
Zea.mays.4E                        AWGSSIRPIHTFSTVEEFWGLYNNINHPSKLIVGADFHCFKNKIEPKWED 113
Setaria.italica.4E                 AWGSSIRPIHTFSTVEDFWGLYNNIHHPSKLIVGADFHCFKNKIEPKWED 116
Sorghum.bicolor.4E                 AWGSSIRPIHTFSTVEDFWGLYNNIHHPSKLVIGADFHCFKNKIEPKWED 115
Brachypodium.distachyon.4E         AWGSTIHPIHTFSTVEDFWSLYNNIHQPSKLNVGSDFHCFKNKIEPKWED 108
Triticum.aestivum.4E               AWGSTIHPIHTFSTVEDFWGLYNNIHNPSKLNVGADFHCFKNKIEPKWED 110
Oryza.sativa.4E                    TWGSSIRPIHTFSTVEDFWSLYNNIHHPSKLVVGADFHCFKNKIEPKWED 122
Citrus.sinensis.4E                 TWGSSMRSIYTFSSVEEFWSLYNNIHHPSKLAVGADFYCFKNKIEPKWED 120
Citrus.clementina.4E               TWGSSMRSIYTFSSVEEFWSLYNNIHHPSKLAVGADFYCFKNKIEPKWED 120
Manihot.esculenta.4E               TWGSSMRSIYTFATVEEFWSIYNNIHHPSKLAVGADFHCFKYKIEPKWED 165
Ricinus.communis.4E                TWGSSMRPIYTFATVEEFWSIYNNIHHPSKLAVGADFHCFKHKIEPKWED 129
Populus.trichocarpa.4E             SWGSSLRSIFTFSTIEEFWSVYNNIHHPSKLAVGADFHCFKDKIEPKWED 119
Glycine.max.4E                     AWGSSIRPIYTFATVEEFWSIYNNIHHPSKLGVGADFHCFKHKIEPKWED 128
Medicago.truncatula.4E             AWGSSIRPVYTFSTVEEFWSIYNNIHHPSKLAIGADFHCFKHKIEPKWED 120
Prunus.persica.4E                  DWGSSIRPIYTFSTVEEFWSIYNNIRHPSKLALGTDFHCFKYKIEPKWED 128
Fragaria.vesca.4E                  DWGSSLRPIYTFSTVEEFWGIYNNIRHPSKLNVGTDFHCFKNKIEPKWED 128
Cucumis.sativus.4E                 TWGASMRPIYTFSTVEEFWSVYNNIHHPSKLALRADLYCFKHKIEPKWED 130
Cucumis.melo.4E                    TWGASIRPIYTFSTVEEFWSVYNNIHHPSKLAMRADLYCFKHKIEPKWED 130
Eucalyptus.grandis.4E              AWGASMRPIYTFSTVEEFWSLYNNIHHPGKLAVGADFYCFKNKIEPKWED 140
Mimulus.guttatus.4E                AWGSSIRPIFTFSSVEDFWSVYNNVHHPSKLAVGADFHCFKNKIEPKWED 135
Aquilegia.coerulea.4E              TWGNSIRPVHNCSTVEDFWCLYNNIHHPSKLAVGADLYCFKHIIQPKWED 122
Theobroma.cacao.4E                 IWGSSMRPIYTFASVEQFWSLYNNIHHPSKLAVGADFHCFKYKIEPKWED 128
Eucalyptus.grandis.4E1BL           AWGSSIRPIYTFATVEEFWSVYNNLHQPSKLVVGADLHCFKNKVEPKWED 137
Fragaria.vesca.4E1BL1              PEVPAIVTSRSCGGCQHGRSMYNNMHHPSKLTPGYDFYLFKENIEPKWED 98
Fragaria.vesca.4E1BL2              -SLCPIYTFSTVKGFWR---VYNNMHHPSKLTPGYDYYIFKENIEPELED 67
Arabidopsis.thaliana.4E1b          IWGSSLRSLYTFGTIEEFWSLYNNIHPPTKWVSGADLYCFKDKIEPKWED 135
Arabidopsis.thaliana.4E1c          IWGSSLRSLYTFATIEEFWSLYNNIHPPTKWVSGSDLYCFKDKIEPKWED 135
Arabidopsis.lyrata.4E1b            TWGSSLRSLYTFATIEEFWSLYNNIHPPTKWVPGADLYCFKHKIEPKWED 137
Arabidopsis.lyrata.4E1c            TWGSSLRSLYTFATIEEFWSLYNNIHPPTKWVPGADLYCFKHKIEPKWED 92
Capsella.rubella.4E1c              TWGSSLRSLYTFSSVEEFWSLYNNIHPPTKWVPGADLYCFKHKIEPKWED 138
Capsella.rubella.4E1b              TWGSSLRSLYTFASVEEFWSLYNNIHPPTKWVPGADLYCFKHKIEPKWED 142
Raphanus.sativus.4E1b              TWGSSLRSLYTFATIEEFWSLYNNMHPPTKCVHGADIYCFKD-------- 133
Brassica.oleracea.4E1b             TWGSSLRSLYTIATIEEFWSLYNNMHPPSKWVHGADLYCFKIEP--KWED 137
Brassica.rapa.4E1b                 TWGSSLRSLYTFATIEEFWSLYNNMHPPTKWVHGADLYCFKHKIGPKWED 139
Brassica.napus.4E1b                TWGSSLRSLYTFATIEEFWSLYNNMHPPTKWVHGADLYCFKHKIGPKWED 128
Raphanus.raphanistrum.4E1b         TWGSSLRSLYTFATIEEFWSLYNNMHPPTKCVHGGDTYCFKDKIDPKWED 123
Brassica.napus.4E1c                AWGSSLRTLYTFATIEEFWSLYNNIHPPTKWVHGSDLYCFKHEIEPKWED 143
Brassica.oleracea.4e1c             AWGSSLRTLYTFATIEEFWSLYNNIHPPTKWVHGSDLYCFKHEIEPKWED 143
Brassica.rapa.4E1c                 TWGSSLRTLYTFATIEEFWSLYNNIHPPNKWVHGSDLYCFKHEIEPKWED 143
Thellungiella.parvula.4E1b         TWGSSLRTLYTLATVEEFWSLYNNIHPPTKWVPGADLYCFKHKIDPKWED 143
Thellungiella.halophila.4E1b       TWGSSLRTLYTFATIEEFWSLYNNMHPPTKWVPGADLYCFKHKIEPKWED 142
Sisymbrium.irio.4E1b               ----------------------------------------KIEAKWED 8
```

```
Arabidopsis.thaliana.4E             PICANGGKWTMTFPK----EKSDKSWLY-TLLALIGEQFDHGDEICGAVV 175
Arabidopsis.lyrata.4E               PICANGGKWTMNFPK----EKSDKSWLY-TLLALIGEQFDHGDEICGAVV 175
Capsella.rubella.4E                 PICANGGKWTMNFPK----EKADKSWLY-TLLALIGEQFDHGDEICGAVV 175
Thellungiella.halophila.4E          PICANGGKWTMTFPK----EKSDKCWLY-TLLALIGEQFDHGDEICGAVV 175
Brassica.rapa.4E1                   PICANGGKWTMNFPK----EKSDKPWLY-TLLALIGEQFDHGDEICGAVV 174
Brassica.oleracea.4E                PICANGGKWTMNFPK----EKSDKPWLY-TLLALIGEQFDHGDEICGAVV 174
Raphanus.raphanistrum.4E            PICANGGKWTMNFSK----EKSDKPWLY-TLLALIGEQFDHGDEICGVVV 171
Raphanus.sativus.4E                 PICANGGKWTMNFSK----EKSDKPWLY-TLLALIGEQFDHGDEICGVVV 171
Brassica.rapa.4E2                   PICANGGKWTMNFSR----EKSDKPFLY-TLLALIGEQFDHGDEICGVVV 168
Brassica.napus.4E                   PICANGGKWTMNFSR----EKSDKPFLY-TLLALIGEQFDHGDEICGVVV 168
Sisymbrium.irio.4E                  PICANGGKWTMNFPK----EKSDKPWLY-TLLALIGEQFDHGDEICGAVV 78
Solanum.tuberosum.4E                PVCANGGTWKMSFLK----GKSDTSWLY-TLLAMIGHQFDHGDEICGAVV 171
Solanum.lycopersicum.4E             PVCANGGTWKMSFSK----GKSDTSWLY-TLLAMIGHQFDHGDEICGAVV 171
Zea.mays.4E                         PICANGGKWTISCGR----GKSDTFWLH-TLLAMIGEQFDYGDEICGAVV 158
Setaria.italica.4E                  PICANGGKWTISCGR----GKSDTMWLH-TLLAMIGEQFDYGDEICGAVV 161
Sorghum.bicolor.4E                  PICANGGKWTISCGR----GKSDTLWLH-TLLAMIGEQFDYGDEICGAVV 160
Brachypodium.distachyon.4E          PICANGGKWTISCGK----GKSDTFWLH-TLLALIGEQFDYGDEICGAVV 153
Triticum.aestivum.4E                PICANGGKWTISCGR----GKSDTFWLH-TLLAMIGEQFDFGDEICGAVV 155
Oryza.sativa.4E                     PICANGGKWTFSCGR----GKSDTMWLH-TLLAMIGEQFDYGDEICGAVV 167
Citrus.sinensis.4E                  PVCANGGKWTVIFPK----GKSDTSWLY-TLLAMIGEQFDHGDEICGAVV 165
Citrus.clementina.4E                PVCANGGKWTVIFPK----GKSDTSWLY-TLLAMIGEQFDHGDEICGAVV 165
Manihot.esculenta.4E                PVCANGGKWTVTFGR----GKSDTSWLY-TLLAMIGEQFDHGDEICGAVV 210
Ricinus.communis.4E                 PVCANGGKWTLTFQK----GKSDGSWLN-TLLAMIGEQFERGDEICGAVV 174
Populus.trichocarpa.4E              PICANGGKWSVTLSK----GKSDTFWLN-TLLALIGEQFDHGDEICGAVV 164
Glycine.max.4E                      PICANGGKWTMTFQR----GKSDTSWLY-TLLAMIGEQFDHGDEICGAVV 173
Medicago.truncatula.4E              PICANGGKWTVQFSR----GKSDTSWLY-TLLAMIGEQFDHGDEICGAVV 165
Prunus.persica.4E                   PVCANGGKWTVTFPK----GKSDTSWLY-TLLGMIGEQFDHGDEICGAVV 173
Fragaria.vesca.4E                   PVCANGGKWTLTFPK----GKSDNSWLH-TLLALIGEQFDHGDEICGAVV 173
Cucumis.sativus.4E                  PVCASGGKWTVNFSR----GKSDNGWLY-TLLAMIGEQFDCGDEICGAVV 175
Cucumis.melo.4E                     PVCANGGKWTVNFPR----GKSDNGWLY-TLLAMIGEQFDCGDEICGAVV 175
Eucalyptus.grandis.4E               PVCANGGKWTVNFNR----GKSDTCWLY-TLLAMIGEQFDHGDEICGAVV 185
Mimulus.guttatus.4E                 PVCANGGKWTVNFSR----GKSDTAWLY-TLLAMIGEQFDYGDEICGAVV 180
Aquilegia.coerulea.4E               PVCANGGKWTIGFSR----GKADSCWLY-TLLAMIGEQFDHGDEICGAVV 167
Theobroma.cacao.4E                  PVCANGGKWTVTLPR----GKSDTCWLY-TLLALIGEQFEYGDEICGAVV 173
Eucalyptus.grandis.4E1BL            PVCANGGKWTVTFSR----GKSDTSWLY-TLLAMVGEQFDHGDEICGAVV 182
Fragaria.vesca.4E1BL1               PVCATGGKWTMTFSN----GESDRSWMD-MLQALVKEEFKHRDEICGAVV 143
Fragaria.vesca.4E1BL2               PACSDGGKWTLTFST----GRSDQSWLH-TIQALVQEQFNHRYEICGAVI 112
Arabidopsis.thaliana.4E1b           PICANGGKWSMMFPK----ATLECNWLN-TLLALVGEQFDQGDEICGAVL 180
Arabidopsis.thaliana.4E1c           PICANGGKWTMFFPR----ATLESNWLN-TLLALVGEQFDQGDEICGAVL 180
Arabidopsis.lyrata.4E1b             PICGNGGKWTMFFPKA---ATLESNWLN-TLLALVGEQFDQGDEICGAVL 183
Arabidopsis.lyrata.4E1c             PICANGGKWTMFFPKA---ATLESNWLN-TLLALVGEQFDQGDEICGTVL 138
Capsella.rubella.4E1c               PICANGGKWTMMFPK----ATLESNWLN-TLLALVGEQFDHGDEICGAVL 183
Capsella.rubella.4E1b               PICANGGKWTMMFPK----ATLESNWLN-TLLALVGEQFDHGDEICGAVL 187
Raphanus.sativus.4E1b               ---------------------------------------------------
Brassica.oleracea.4E1b              PVCADGGKWTMMFXK----ATLESNWLN-TLLALVGEQFEQGDDICGAVL 182
Brassica.rapa.4E1b                  PVCADGGKWTMMFPK----ATLESSWLN-TLLALVGEQFEKGDEICGAVL 184
Brassica.napus.4E1b                 PVCADGGKWTMMFPK----ATLESSWLN-TLLALVGEQFEKGDEICGAVL 173
Raphanus.raphanistrum.4E1b          PVCANGGKWTMMFPK----ATLESSWLN-TLLALVGEQFEQGDEICGAVL 168
Brassica.napus.4E1c                 PVCANGGKWTMMLPK----ATLESNWLN-TLLSLVGEQFEHGDEICGAVL 188
Brassica.oleracea.4e1c              PVCANGGKWTMMLPK----ATLESNWLN-TVL------------------- 170
Brassica.rapa.4E1c                  PVCANGGKWTMMFPT----ATLESNWLN-TLLSLVGEQFEHGDEICGAVL 188
Thellungiella.parvula.4E1b          PICANGGKWTMMFPK----PTLESNWLDTVLLALVGEQFEQGDEICGAVL 189
Thellungiella.halophila.4E1b        PTCADGGKWTMMFPK----AKLESNWLN-TLLALVGEQFEQGDEICGAVL 187
Sisymbrium.irio.4E1b                PICGNGRKRTIMFPK----ATVEYNWLN-TLLALVGEQFEHGDEICGAVL 53
```

```
Arabidopsis.thaliana.4E            NIRG--KQERISIWTKNASNEAAQVSIGKQWKEFLDYNNSIGFIIHEDAK 223
Arabidopsis.lyrata.4E              NIRG--KQERISIWTKNASNEAAQVSIGKQWKEFLDYNNSIGFIIHEDAK 223
Capsella.rubella.4E                NIRG--K-ERISIWTKNASNEAAQVSIGKQWKEILDYNNSIGFIIHEDAR 222
Thellungiella.halophila.4E         NIRG--KQERISIWTKNASNEAAQVSIGKQWKEFLDYNNSIGFIIHEDAM 223
Brassica.rapa.4E1                  NVRG--KQERISIWTKNASNEAAQVSIGKQWKEFIDYNNSIGFIIHEDAK 222
Brassica.oleracea.4E               NVRG--KQERISIWTKNASNEAAQVSIGRQWKEFIDYNNSIGFIIHEDAK 222
Raphanus.raphanistrum.4E           NVRV--KQERISIWTKNASNEAAQVSIGRQWKEIIDHNNSIGFIIHEDTK 219
Raphanus.sativus.4E                NVRV--KQERISIWTKNATNEAAQVSIGRQWKEIIDYNNSIGFIIHEDAK 219
Brassica.rapa.4E2                  NVRA--KQERISIWTKNSSNEAAQVSIGRQWKEFLDYNSSIGFIIHEDAK 216
Brassica.napus.4E                  NVRA--KQERISIWTKNSSNEAAQVSIGRQWKEFLDYNSSIGFIIHEDAK 216
Sisymbrium.irio.4E                 NIRG--KQERISIWTKNSSNEAAQVSIGRQWKEFLDYNNSIGFIIHDDAK 126
Solanum.tuberosum.4E               SVRS--KGEKIALWTKNAANETAQVSIGKQWKQFLDHSDSVGFIFHDDAK 219
Solanum.lycopersicum.4E            SVRA--KGEKIALWTKNAANETAQVSIGKQWKQFLDYSDSVGFIFHDDAK 219
Zea.mays.4E                        SVRG--KQERIAIWTKNAANEAAQVSIGKQWKELLDYKDSIGFIVHDDAK 206
Setaria.italica.4E                 SVRG--KQERIAIWTKNAGNEAAQVSIGKQWKEFLDYKDSIGFIVHDDAK 209
Sorghum.bicolor.4E                 SVRG--KQERIAIWTKNAANEAAQISIGKQWKEFLDYKDSIGFIVHDDAK 208
Brachypodium.distachyon.4E         SIRG--KQERVAIWTKNAANEAAQISIGKQWKEFLDYKDSIGFIVHDDAK 201
Triticum.aestivum.4E               SVRQ--KQERVAIWTKNAANEAAQISIGKQWKEFLDYKDSIGFIVHEDAK 203
Oryza.sativa.4E                    SVRG--KQERIAIWTKNAANEAAQISIGKQWKEFLDYKDSIGFIVHDDAK 215
Citrus.sinensis.4E                 NVRA--RQEKISLWTKNASNEAAQMSIGKQWKELLDYSDTIGFIFHEDAK 213
Citrus.clementina.4E               NVRA--RQEKISLWTKNASNEAAQMSIGKQWKELLDYSDTIGFIFHEDAK 213
Manihot.esculenta.4E               NVRI--KQEKIALWTKNASNEAAQLSIGKQWKEFLDYNDTIGFIFHVRTQ 258
Ricinus.communis.4E                SVRA--RQEKIALWTKNAANEAAQLSIGKQWKELLDYNDTIGFIFHDDAK 222
Populus.trichocarpa.4E             SVRA--RQEKIAIWTKNASNETAQVSIGKQWKEFLDYNETIGFIFHDDAK 212
Glycine.max.4E                     NVRS--RQEKIAIWTKNASNEAAQVSIGKQWKEFLDYNDTIGFIFHEDAK 221
Medicago.truncatula.4E             NVRS--RAEKISIWTKNAANEAAQVSIGKQWKEFLDYNETIGFIFHDDAK 213
Prunus.persica.4E                  NVRN--RQEKISIWTKNAINEAAQLSIGKQWKGFLDYNETIGFIFHEDAM 221
Fragaria.vesca.4E                  NVRG--RQEKISIWTKNAENEAAQMSIGKQWKSFLDSNENIGFIFHEDAR 221
Cucumis.sativus.4E                 NVRS--GQDKISIWTKNASNEAAQASIGKQWKEFLDYNESIGFIFHDDAK 223
Cucumis.melo.4E                    NVRS--GQDKISIWTKNASNEAAQASIGKQWKEFLDYNESIGFIFHDDAK 223
Eucalyptus.grandis.4E              NVRN--RQEKVSIWTKNAANEVAQVSIGKQWKEFLDYNENIGFIFHEDAK 233
Mimulus.guttatus.4E                NVRA--RQEKISIWTKNAANEVAQVSIGKQWKEFLDYNESIGFIFHDDAK 228
Aquilegia.coerulea.4E              NVRA--KQEKISLWTKNAFNETAQISIGKQWKEFLDHNDNIGFIFHEDAK 215
Theobroma.cacao.4E                 SVRG--RQEKIALWTKNAANETAQISIGKQWKELLDCNDTIGFIFHEDAK 221
Eucalyptus.grandis.4E1BL           SVRA--KQEKIALWTRDASDEAAQMSIGQQWKEFLNYSQKAGFILHEDAK 230
Fragaria.vesca.4E1BL1              NVRN--GQESIALWTKNATNEAAQVSIGKQWKGFLDSKETIVFTIHAG-- 189
Fragaria.vesca.4E1BL2              SLRD--GQEKIALWTKNAANEAVQVSIGKQWKGFLDSNETIWFTFH---- 156
Arabidopsis.thaliana.4E1b          NFRA--RGDRISLWTKNAANEEAQLSIGKQWKELLGYNETIGFIVHEDAK 228
Arabidopsis.thaliana.4E1c          NFRT--RGDRISLWTKKAANEEAQLSIGKQWKELLGYNDTIGFIVHEDAK 228
Arabidopsis.lyrata.4E1b            NFRA--RGDRISLWTKNAANEEAQLSIGKQWKELLGYNDTIGFIVHEDAK 231
Arabidopsis.lyrata.4E1c            NFRT--RGDRISLWTKNAANEEAQLSIGKQWKELLGYNEKIGFIVHEDAK 186
Capsella.rubella.4E1c              NFRA--RGDRISLWTKNAADEDAQLSIGKQWKELVGYNDTIGFIVHEDAK 231
Capsella.rubella.4E1b              NFRA--RGDRISLWTKNAADEDAQLSIGKQWKELVGYNDTIGFIVHEDAK 235
Raphanus.sativus.4E1b              --------------------------------------------------
Brassica.oleracea.4E1b             NFRT--RGDKISIWTKNAADEKAQ-------------------------- 204
Brassica.rapa.4E1b                 NFRT--RGDKISIWTKNAANEKAQINIGKQWKELLGYTETIGFIFHEDAK 232
Brassica.napus.4E1b                NFRT--RGDKISIWTKNAANEKAQINIGKQWKELLGYTETIGFIFHEDAK 221
Raphanus.raphanistrum.4E1b         NFRT--RGDKISLWTKNAANKEAQLSIGKQWKELLGYTETIGFIFHEDAK 216
Brassica.napus.4E1c                NFRT--RGDKISLWTKNAANEEAQISIGKQWKERL--------------- 221
Brassica.oleracea.4e1c             --------------------------------------------------
Brassica.rapa.4E1c                 NFRT--RGDKISIWTKNAANEEAQLSIGKQWKELVGCNETIGFIFHEDAK 236
Thellungiella.parvula.4E1b         NFRT--RGDKISIWTKNAGDEEAQVSIGKQWKELLGYNETIGFIFHEDAK 237
Thellungiella.halophila.4E1b       NFRT--RGDKISLWTKNAANEKAQLSIGKQWKELLGYNESIGFIYHEDAK 235
Sisymbrium.irio.4E1b               NFRT--KGDKISIWTKNAANEEAHVN-GNSGRNFLGYNETIGFIFQEDAK 100
```

```
Arabidopsis.thaliana.4E            K-LDRNAKN------AYTA------------------ 235
Arabidopsis.lyrata.4E              K-LDRNAKN------AYTA------------------ 235
Capsella.rubella.4E                K-LDRNAKN------AYTA------------------ 234
Thellungiella.halophila.4E         K-LDRNAKS------AYTA------------------ 235
Brassica.rapa.4E1                  K-LDRGAKS------AYTA------------------ 234
Brassica.oleracea.4E               K-LDRGAKS------AYTA------------------ 234
Raphanus.raphanistrum.4E           K-LDRGAKS------AYTA------------------ 231
Raphanus.sativus.4E                K-LDRGARS------AYTA------------------ 231
Brassica.rapa.4E2                  K-LDRGAKS------AYTA------------------ 228
Brassica.napus.4E                  K-LDRGAKS------AYTA------------------ 228
Sisymbrium.irio.4E                 K-LDRGAKS------AYTV------------------ 138
Solanum.tuberosum.4E               R-LDRNAKN------RYTV------------------ 231
Solanum.lycopersicum.4E            R-LDRNAKN------RYTV------------------ 231
Zea.mays.4E                        K-MDKGLKN------RYTV------------------ 218
Setaria.italica.4E                 K-MDKGPKN------RYTV------------------ 221
Sorghum.bicolor.4E                 K-ADKGPRN------RYTV------------------ 220
Brachypodium.distachyon.4E         KDNTKGPKN------RYTV------------------ 214
Triticum.aestivum.4E               R-SDKGPKN------RYTV------------------ 215
Oryza.sativa.4E                    K-MDKGLKN------RYTV------------------ 227
Citrus.sinensis.4E                 H--DNRSKN------RYTT------------------ 224
Citrus.clementina.4E               H--DNRSKN------RYTT------------------ 224
Manihot.esculenta.4E               R--SLTEVP------RIATQYELYRVYESVSMSPAYY 287
Ricinus.communis.4E                K-HERSAKN------RYTI------------------ 234
Populus.trichocarpa.4E             K--DRNAKN------RYSV------------------ 223
Glycine.max.4E                     K-LDRGAKN------KYVV------------------ 233
Medicago.truncatula.4E             K-LDRAAKN------KYVV------------------ 225
Prunus.persica.4E                  R-QERSAKN------KYVA------------------ 233
Fragaria.vesca.4E                  R--ERNPKN------SYTV------------------ 232
Cucumis.sativus.4E                 K-FDRHAKN------KYMV------------------ 235
Cucumis.melo.4E                    K-FDRLAKN------KYMV------------------ 235
Eucalyptus.grandis.4E              K-LDKAAKN------RYNA------------------ 245
Mimulus.guttatus.4E                K-LDRGAKN------RYSV------------------ 240
Aquilegia.coerulea.4E              K-HDRFAKN------RYSI------------------ 227
Theobroma.cacao.4E                 K-LDKAAKN------RYTI------------------ 233
Eucalyptus.grandis.4E1BL           K-LDRAARN------RYTV------------------ 242
Fragaria.vesca.4E1BL1              --QDDASQVGSTATNLYTV------------------ 206
Fragaria.vesca.4E1BL2              ---EDARQVGSYANDRYTV------------------ 172
Arabidopsis.thaliana.4E1b          T-LDRDAKR------RYTV------------------ 240
Arabidopsis.thaliana.4E1c          T-LDRDAKR------RYTV------------------ 240
Arabidopsis.lyrata.4E1b            T-LDRDAKR------RYTV------------------ 243
Arabidopsis.lyrata.4E1c            T-LDRHAKS------RYTV------------------ 198
Capsella.rubella.4E1c              T-LDRDAKN------RYTV------------------ 243
Capsella.rubella.4E1b              T-LDRDAKN------RYTV------------------ 247
Raphanus.sativus.4E1b              ------------------------------------
Brassica.oleracea.4E1b             ------------------------------------
Brassica.rapa.4E1b                 T-LDRTAKP------RYTV------------------ 244
Brassica.napus.4E1b                T-LDRTAKP------RYTV------------------ 233
Raphanus.raphanistrum.4E1b         T-LDRNAKP------QYTV------------------ 228
Brassica.napus.4E1c                ------------------------------------
Brassica.oleracea.4e1c             ------------------------------------
Brassica.rapa.4E1c                 T-LDRNARP------RYTV------------------ 248
Thellungiella.parvula.4E1b         I-LDRNAKP------RYTI------------------ 249
Thellungiella.halophila.4E1b       T-LDRNAKP------RYTV------------------ 247
Sisymbrium.irio.4E1b               T-LDRNAKP------RYTI------------------ 112
```

Figure A.1. ClustalW2 alignment of eIF4E genes of flowering plants. Residues highlighted in yellow have 90% or greater identity in conserved flowering plant eIF4E sequences but are consistently altered in eIF4E1b-type sequences, while highlighting in green marks conserved divergent residues at these locations in eIF4E1b-type sequences. *R. sativus* eIF4E1b, and *B. napus* and *B. oleracea* eIF4E1c, and *S. irio* eIF4E1b are partial sequences.

Table A.1. Screening of *cum1 iso4e-1*/*EIFISO4E* self-fertilization progeny. 60 seeds germinated on dirt, with 59 surviving to be screened. Recovery of *cum1 iso4E-1*/*EIFISO4E* was lower than expected (35%) and double homozygous mutant plants were not recovered.

	Number with Genotype	% of Total	Expected Mendelian %
cum1 iso4e-1/iso4e-1	0/60	0%	25%
cum1 iso4e-1/EIFISO4E	21/60	35%	50%
cum1 EIFISO4E/EIFISO4E	38/60	63%	25%
not screened	1/60	2%	

Table A.2. DNA oligonucleotides used for plant screening.

Primer	Sequence	Purpose
4EAnnes	TTGACCCAATAGAGTCCAGAAAT	eIFiso4E screen
4E2KO4	CTCTCCAATCAAAGCCATCAACTA	eIFiso4E screen -- WT band with 4EAnnes
DSMP8	GTTTTGGCCGACACTCCTTACC	eIFiso4E screen -- insert band with 4EAnnes
Cum1 F	GTCGGAAATAAAATAAAATCAAAAACCTAAGCT	eIF4E screen
Cum1 R	AAGCCTAATTCAATAGAATCCGA	eIF4E screen
		Cum1 F+R PCR product is digested with HindIII (Invitrogen) -- *cum1* is digested, WT is not

References

Aitken, C.E., and Lorsch, J.R. (2012). A mechanistic overview of translation initiation in eukaryotes. Nat Struct Mol Biol *19*, 568-576.

Ajiro, M., Katagiri, T., Ueda, K., Nakagawa, H., Fukukawa, C., Lin, M.L., Park, J.H., Nishidate, T., Daigo, Y., and Nakamura, Y. (2009). Involvement of RQCD1 overexpression, a novel cancer-testis antigen, in the Akt pathway in breast cancer cells. Int J Oncol *35*, 673-681.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., *et al.* (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science *301*, 653-657.

Altmann, M., Müller, P.P., Pelletier, J., Sonenberg, N., and Trachsel, H. (1989). A mammalian translation initiation factor can substitute for its yeast homologue *in vivo*. J Biol Chem *264*, 12145-12147.

Amiri, A., Keiper, B.D., Kawasaki, I., Fan, Y., Kohara, Y., Rhoads, R.E., and Strome, S. (2001). An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in *C-elegans*. Development *128*, 3899-3912.

Andreou, A.Z., and Klostermeier, D. (2014). eIF4B and eIF4G jointly stimulate eIF4A ATPase and unwinding activities by modulation of the eIF4A conformational cycle. J Mol Biol *426*, 51-61.

Ash, M.R., Faelber, K., Kosslick, D., Albert, G.I., Roske, Y., Kofler, M., Schuemann, M., Krause, E., and Freund, C. (2010). Conserved beta-hairpin recognition by the GYF domains of Smy2 and GIGYF2 in mRNA surveillance and vesicular transport complexes. Structure *18*, 944-954.

Ashby, J.A., Stevenson, C.E., Jarvis, G.E., Lawson, D.M., and Maule, A.J. (2011). Structure-based mutational analysis of eIF4E in relation to sbm1 resistance to pea seed-borne mosaic virus in pea. PLoS ONE *6*, e15873.

Benson, D.A., Karsch-Mizrachi, I., Clark, K., Lipman, D.J., Ostell, J., and Sayers, E.W. (2012). GenBank. Nucleic Acids Res *40*, D48-D53.

Bentsink, L., Jowett, J., Hanhart, C.J., and Koornneef, M. (2006). Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. Proc Natl Acad Sci U S A *103*, 17042-17047.

Besse, F., and Ephrussi, A. (2008). Translational control of localized mRNAs: restricting protein synthesis in space and time. Nat Rev Mol Cell Biol *9*, 971-980.

Blasius, M., Wagner, S.A., Choudhary, C., Bartek, J., and Jackson, S.P. (2014). A quantitative 14-3-3 interaction screen connects the nuclear exosome targeting complex to the DNA damage response. Genes Dev *28*, 1977-1982.

Blewett, N.H., and Goldstrohm, A.C. (2012). An eIF4E-binding protein promotes mRNA decapping and is required for PUF repression. Mol Cell Biol *32*, 4181-4194.

Boex-Fontvieille, E., Daventure, M., Jossier, M., Zivy, M., Hodges, M., and Tcherkez, G. (2013). Photosynthetic control of Arabidopsis leaf cytoplasmic translation initiation by protein phosphorylation. PLoS ONE *8*, e70692.

Bolster, D.R., Vary, T.C., Kimball, S.R., and Jefferson, L.S. (2004). Leucine regulates translation initiation in rat skeletal muscle via enhanced eIF4G phosphorylation. J Nutr *134*, 1704-1710.

Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijo, J.A., and Becker, J.D. (2008). Comparative transcriptomics of Arabidopsis sperm cells. Plant Physiol *148*, 1168-1181.

Branco-Price, C., Kaiser, K.A., Jang, C.J., 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. Plant J *56*, 743-755.

Braun, J.E., Truffault, V., Boland, A., Huntzinger, E., Chang, C.T., Haas, G., Weichenrieder, O., Coles, M., and Izaurralde, E. (2012). A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation. Nat Struct Mol Biol *19*, 1324-1331.

Browning, K.S. (2014). Plant Translational Machinery. In Molecular Biology, S.H. Howell, ed. (Springer New York), pp. 129-151.

Browning, K.S., and Bailey-Serres, J. (2015). Mechanism of cytoplasmic mRNA translation. Arabidopsis Book *13*, e0176.

Bush, M.S., Hutchins, A.P., Jones, A.M., Naldrett, M.J., Jarmolowski, A., Lloyd, C.W., and Doonan, J.H. (2009). Selective recruitment of proteins to 5' cap complexes during the growth cycle in Arabidopsis. Plant J *59*, 400-412.

Callot, C., and Gallois, J.L. (2014). Pyramiding resistances based on translation initiation factors in Arabidopsis is impaired by male gametophyte lethality. Plant Signal Behav *9*, e27940.

Cao, J., Schneeberger, K., Ossowski, S., Gunther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle, O., Lippert, C., *et al.* (2011). Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet *43*, 956-963.

Cao, Q., Kim, J.H., and Richter, J.D. (2006). CDK1 and calcineurin regulate Maskin association with eIF4E and translational control of cell cycle progression
239. Nat Struct Mol Biol *13*, 1128-1134.

Cao, R., Gkogkas, C.G., de, Z.N., Blum, I.D., Yanagiya, A., Tsukumo, Y., Xu, H., Lee, C., Storch, K.F., Liu, A.C., *et al.* (2015). Light-regulated translational control of circadian behavior by eIF4E phosphorylation. Nat Neurosci *18*, 855-862.

Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S., Ami, G.O.H., and Web Presence Working, G. (2009). AmiGO: online access to ontology and annotation data. Bioinformatics *25*, 288-289.

Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K.B., Montgomery, T.A., Nguyen, T., Cuperus, J.T., and Carrington, J.C. (2012). Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. Plant Cell *24*, 3613-3629.

Chartier-Harlin, M.C., Dachsel, J.C., Vilarino-Guell, C., Lincoln, S.J., Lepretre, F., Hulihan, M.M., Kachergus, J., Milnerwood, A.J., Tapia, L., Song, M.S., *et al.* (2011). Translation initiator EIF4G1 mutations in familial Parkinson disease. Am J Hum Genet *89*, 398-406.

Chekanova, J.A. (2015). Long non-coding RNAs and their functions in plants. Curr Opin Plant Biol *27*, 207-216.

Cheng, S., and Gallie, D.R. (2010). Competitive and noncompetitive binding of eIF4B, eIF4A, and the poly(A) binding protein to wheat translation initiation factor eIFiso4G. Biochemistry *49*, 8251-8265.

Cheng, S., and Gallie, D.R. (2013). Eukaryotic initiation factor 4B and the poly(A)-binding protein bind eIF4G competitively. Translation *1*, e24038.

Cho, P.F., Gamberi, C., Cho-Park, Y.A., Cho-Park, I.B., Lasko, P., and Sonenberg, N. (2006). Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. Curr Biol *16*, 2035-2041.

Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., and Sonenberg, N. (2005). A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell *121*, 411-423.

Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., and Lawrence, J.B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell *33*, 717-726.

Coldwell, M.J., Sack, U., Cowan, J.L., Barrett, R.M., Vlasak, M., Sivakumaran, K., and Morley, S.J. (2012). Multiple isoforms of the translation initiation factor eIF4GII are generated via use of alternative promoters, splice sites and a non-canonical initiation codon. Biochem J *448*, 1-11.

Cridge, A.G., Castelli, L.M., Smirnova, J.B., Selley, J.N., Rowe, W., Hubbard, S.J., McCarthy, J.E., Ashe, M.P., Grant, C.M., and Pavitt, G.D. (2010). Identifying eIF4E-binding protein translationally-controlled transcripts reveals links to mRNAs bound by specific PUF proteins. Nucleic Acids Res *38*, 8039-8050.

Daigo, K., Kawamura, T., Ohta, Y., Ohashi, R., Katayose, S., Tanaka, T., Aburatani, H., Naito, M., Kodama, T., Ihara, S., *et al.* (2011). Proteomic analysis of native hepatocyte nuclear factor-4alpha (HNF4alpha) isoforms, phosphorylation status, and interactive cofactors. J Biol Chem *286*, 674-686.

Decker, C.J., and Parker, R. (2012). P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. Cold Spring Harb Perspect Biol *4*, a012286.

DeLano, W.L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol *12*, 14-20.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., *et al.* (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res *36*, W465-469.

Dinkova, T.D., Keiper, B.D., Korneeva, N.L., Aamodt, E.J., and Rhoads, R.E. (2005). Translation of a small subset of Caenorhabditis elegans mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform. Mol Cell Biol *25*, 100-113.

Dobrikov, M.I., Shveygert, M., Brown, M.C., and Gromeier, M. (2014). Mitotic phosphorylation of eukaryotic initiation factor 4G1 (eIF4G1) at Ser1232 by Cdk1:cyclin B inhibits eIF4A helicase complex binding with RNA. Mol Cell Biol *34*, 439-451.

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K.S., and Robaglia, C. (2002). The *Arabidopsis* eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. Plant J *32*, 927-934.

Evdokimova, V., Ruzanov, P., Imataka, H., Raught, B., Svitkin, Y., Ovchinnikov, L.P., and Sonenberg, N. (2001). The major mRNA-associated protein YBA is a potent 5' cap-dependent mRNA stabilizer. EMBO J *20*, 5491-5502.

Evsikov, A.V., and Marin de, E.C. (2009). Evolutionary origin and phylogenetic analysis of the novel oocyte-specific eukaryotic translation initiation factor 4E in Tetrapoda. Dev Genes Evol *219*, 111-118.

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet *15*, 7-21.

Fatscher, T., Boehm, V., Weiche, B., and Gehring, N.H. (2014). The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. RNA *20*, 1579-1592.

Flaherty, S.M., Fortes, P., Izaurralde, E., Mattaj, I.W., and Gilmartin, G.M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3' processing. Proc Natl Acad Sci USA *94*, 11893-11898.

Freire, E.R., Dhalia, R., Moura, D.M., da Costa Lima, T.D., Lima, R.P., Reis, C.R., Hughes, K., Figueiredo, R.C., Standart, N., Carrington, M., *et al.* (2011). The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. Mol Biochem Parasitol *176*, 25-36.

Freire, M.A. (2005). Translation initiation factor (iso) 4E interacts with BTF3, the beta subunit of the nascent polypeptide-associated complex. Gene *345*, 271-277.

Freire, M.A., Tourneur, C., Granier, F., Camonis, J., El Amrani, A., Browning, K.S., and Robaglia, C. (2000). Plant lipoxygenase 2 is a translation initiation factor-4E- binding protein. Plant Molecular Biology *44*, 129-140.

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., Sonenberg, N., Thoma, C., and Fujiwara, T. (2014). MicroRNAs Trigger Dissociation of eIF4AI and eIF4AII from Target mRNAs in Humans. Mol Cell *56*, 79-89.

Fukaya, T., Iwakawa, H.O., and Tomari, Y. (2014). MicroRNAs Block Assembly of eIF4F Translation Initiation Complex in Drosophila. Mol Cell *56*, 67-78.

Furic, L., Beardsley, E., and Topisirovic, I. (2014). eIF4E Phosphorylation Downstream of MAPK Pathway. In Translation and Its Regulation in Cancer Biology and Medicine, A. Parsyan, ed. (Netherlands: Springer Netherlands), pp. 363-374.

Gallie, D.R., and Browning, K.S. (2001). eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. J Biol Chem *276*, 36951-36960.

Gallie, D.R., and Liu, R. (2014). Phylogenetic analysis reveals dynamic evolution of the poly(A)-binding protein gene family in plants. BMC Evol Biol *14*, 238.

Gallois, J.L., Charron, C., Sanchez, F., Pagny, G., Houvenaghel, M.C., Moretti, A., Ponz, F., Revers, F., Caranta, C., and German-Retana, S. (2010). Single amino acid changes in the turnip mosaic virus viral genome-linked protein (VPg) confer virulence towards Arabidopsis thaliana mutants knocked out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G. J Gen Virol *91*, 288-293.

Garcia-Garcia, C., Frieda, K.L., Feoktistova, K., Fraser, C.S., and Block, S.M. (2015). Factor-dependent processivity in human eIF4A DEAD-box helicase. Science *348*, 1486-1488.

Gehrke, S., Wu, Z., Klinkenberg, M., Sun, Y., Auburger, G., Guo, S., and Lu, B. (2015). PINK1 and Parkin control localized translation of respiratory chain component mRNAs on mitochondria outer membrane. Cell Metab *21*, 95-108.

German-Retana, S., Walter, J., Doublet, B., Roudet-Tavert, G., Nicaise, V., Lecampion, C., Houvenaghel, M.C., Robaglia, C., Michon, T., and Le, G.O. (2008). Mutational analysis of a plant cap-binding protein eIF4E reveals key amino-acids involved in biochemical functions and potyvirus infection. J Virol.

Ghosh, S., and Lasko, P. (2015). Loss-of-function analysis reveals distinct requirements of the translation initiation factors eIF4E, eIF4E-3, eIF4G and eIF4G2 in Drosophila spermatogenesis. PLoS One *10*, e0122519.

Giovannone, B., Lee, E., Laviola, L., Giorgino, F., Cleveland, K.A., and Smith, R.J. (2003). Two novel proteins that are linked to insulin-like growth factor (IGF-I) receptors by the Grb10 adapter and modulate IGF-I signaling. J Biol Chem *278*, 31564-31573.

Giovannone, B., Tsiaras, W.G., de la Monte, S., Klysik, J., Lautier, C., Karashchuk, G., Goldwurm, S., and Smith, R.J. (2009). GIGYF2 gene disruption in mice results in

neurodegeneration and altered insulin-like growth factor signaling. Hum Mol Genet *18*, 4629-4639.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., *et al.* (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res *40*, D1178-D1186.

Goss, D.J., and Kleiman, F.E. (2013). Poly(A) binding proteins: are they all created equal? Wiley Interdiscip Rev RNA *4*, 167-179.

Gross, J.D., Moerke, N.J., Von der Haar, T., Lugovskoy, A.A., Sachs, A.B., McCarthy, J.E.G., and Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. Cell *115*, 739-750.

Haghighat, A., and Sonenberg, N. (1997). eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'- cap structure. J Biol Chem *272*, 21677-21680.

Halstead, J.M., Lionnet, T., Wilbertz, J.H., Wippich, F., Ephrussi, A., Singer, R.H., and Chao, J.A. (2015). An RNA biosensor for imaging the first round of translation from single cells to living animals. Science *347*, 1367-1671.

Hellman, L.M., and Fried, M.G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. Nat Protoc *2*, 1849-1861.

Hernandez, G. (2012). Translational Control across Eukaryotes. Comp Funct Genomics *2012*, 317690.

Hernandez, G., Altmann, M., Sierra, J.M., Urlaub, H., Del Corral, R.D., Schwartz, P., and Rivera-Pomar, R. (2005). Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. Mech Dev *122*, 529-543.

Higashi, S., Iseki, E., Minegishi, M., Togo, T., Kabuta, T., and Wada, K. (2010). GIGYF2 is present in endosomal compartments in the mammalian brains and enhances IGF-1-induced ERK1/2 activation. J Neurochem *115*, 423-437.

Hilbert, M., Kebbel, F., Gubaev, A., and Klostermeier, D. (2011). eIF4G stimulates the activity of the DEAD box protein eIF4A by a conformational guidance mechanism. Nucleic Acids Res *39*, 2260-2270.

Hoover, D.M., and Lubkowski, J. (2002). DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res *30*, e43.

Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene *77*, 61-68.

Hu, S.I., Katz, M., Chin, S., Qi, X., Cruz, J., Ibebunjo, C., Zhao, S., Chen, A., and Glass, D.J. (2012). MNK2 Inhibits eIF4G Activation Through a Pathway Involving Serine-Arginine-Rich Protein Kinase in Skeletal Muscle. Sci Signal *5*, ra14.

Ivanov, P., Emara, M.M., Villen, J., Gygi, S.P., and Anderson, P. (2011). Angiogenin-induced tRNA fragments inhibit translation initiation. Mol Cell *43*, 613-623.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657-668.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol *11*, 113-127.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. Nature *467*, 608-611.

Jiao, Y., and Meyerowitz, E.M. (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol Syst Biol *6*, 419.

Jivotovskaya, A.V., Valasek, L., Hinnebusch, A.G., and Nielsen, K.H. (2006). Eukaryotic translation initiation factor 3 (eIF3) and eIF2 can promote mRNA binding to 40S subunits independently of eIF4G in yeast. Mol Cell Biol *26*, 1355-1372.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T.L. (2008). NCBI BLAST: a better web interface. Nucleic Acids Res *36*, W5-9.

Jongjitwimol, J., Feng, M., Zhou, L., Wilkinson, O., Small, L., Baldock, R., Taylor, D.L., Smith, D., Bowler, L.D., Morley, S.J., *et al.* (2014). The S. pombe Translation Initiation Factor eIF4G Is Sumoylated and Associates with the SUMO Protease Ulp2. PLoS ONE *9*, e94182.

Joshi, B., Cameron, A., and Jagus, R. (2004). Characterization of mammalian eIF4E-family members. Eur J Biochem *271*, 2189-2203.

Joshi, B., Lee, K., Maeder, D.L., and Jagus, R. (2005). Phylogenetic analysis of eIF4E-family members. BMC Evol Biol *5*, 48.

Jovanovic, M., Rooney, M.S., Mertins, P., Przybylski, D., Chevrier, N., Satija, R., Rodriguez, E.H., Fields, A.P., Schwartz, S., Raychowdhury, R., *et al.* (2015). Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. Science *347*, 1259038.

Jung, M.Y., Lorenz, L., and Richter, J.D. (2006). Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. Mol Cell Biol *26*, 4277-4287.

Juntawong, P., and Bailey-Serres, J. (2012). Dynamic Light Regulation of Translation Status in Arabidopsis thaliana. Front Plant Sci *3*, 66.

Kafasla, P., Barrass, J.D., Thompson, E., Fromont-Racine, M., Jacquier, A., Beggs, J.D., and Lewis, J. (2009). Interaction of yeast eIF4G with spliceosome components: implications in pre-mRNA processing events. RNA Biol *6*, 563-574.

Kamenska, A., Lu, W.T., Kubacka, D., Broomhead, H., Minshall, N., Bushell, M., and Standart, N. (2013). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. Nucleic Acids Res *42*, 3298-3313.

Kapasi, P., Chaudhuri, S., Vyas, K., Baus, D., Komar, A.A., Fox, P.L., Merrick, W.C., and Mazumder, B. (2007). L13a blocks 48S assembly: role of a general initiation factor in mRNA-specific translational control. Mol Cell *25*, 113-126.

Kawaguchi, R., Girke, T., Bray, E.A., and Bailey-Serres, J. (2004). Differential mRNA translation contributes to gene regulation under non-stress and dehydration stress conditions in Arabidopsis thaliana. Plant J *38*, 823-839.

Kawahara, H., Imai, T., Imataka, H., Tsujimoto, M., Matsumoto, K., and Okano, H. (2008). Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. J Cell Biol *181*, 639-653.

Keiper, B.D., Lamphear, B.J., Deshpande, A.M., Jankowska-Anyszka, M., Aamodt, E.J., Blumenthal, T., and Rhoads, R.E. (2000). Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. J Biol Chem *275*, 10590-10596.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nat Rev Mol Cell Biol *13*, 700-712.

Kim, J.Y., Song, H.R., Taylor, B.L., and Carré, I.A. (2003). Light-regulated translation mediates gated induction of the *Arabidopsis* clock protein LHY. EMBO J *22*, 935-944.

Kim, S.H., Spensley, M., Choi, S.K., Calixto, C.P., Pendle, A.F., Koroleva, O., Shaw, P.J., and Brown, J.W. (2010). Plant U13 orthologues and orphan snoRNAs identified by RNomics of RNA from Arabidopsis nucleoli. Nucleic Acids Res *38*, 3054-3067.

Kinkelin, K., Veith, K., Grunwald, M., and Bono, F. (2012). Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. RNA *18*, 1624-1634.

Kline, K.G., Barrett-Wilt, G.A., and Sussman, M.R. (2010). In planta changes in protein phosphorylation induced by the plant hormone abscisic acid. Proc Natl Acad Sci U S A *107*, 15986-15991.

Kofler, M., Motzny, K., and Freund, C. (2005). GYF domain proteomics reveals interaction sites in known and novel target proteins. Mol Cell Proteomics *4*, 1797-1811.

Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X., *et al.* (2015). Excess of rare, inherited truncating mutations in autism. Nat Genet *47*, 582-588.

Kubacka, D., Kamenska, A., Broomhead, H., Minshall, N., Darzynkiewicz, E., and Standart, N. (2013). Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. PLoS One *8*, e72761.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., *et al.* (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res *40*, D1202-1210.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Laroia, G., Cuesta, R., Brewer, G., and Schneider, R.J. (1999). Control of mRNA decay by heat shock-ubiquitin-proteasome pathway. Science *284*, 499-502.

Lautier, C., Goldwurm, S., Durr, A., Giovannone, B., Tsiaras, W.G., Pezzoli, G., Brice, A., and Smith, R.J. (2008). Mutations in the GIGYF2 (TNRC15) gene at the PARK11 locus in familial Parkinson disease. Am J Hum Genet *82*, 822-833.

Lax, S.R., Lauer, S.J., Browning, K.S., and Ravel, J.M. (1986). Purification and Properties of Protein Synthesis Initiation and Elongation Factors from Wheat Germ. Methods Enzymol *118*, 109-128.

Le Gallo, M., and Bell, D.W. (2014). The emerging genomic landscape of endometrial cancer. Clin Chem *60*, 98-110.

Lejeune, F., Ranganathan, A.C., and Maquat, L.E. (2004). eIF4G is required for the pioneer round of translation in mammalian cells. Nat Struct Mol Biol *11*, 992-1000.

Lellis, A.D., Allen, M.L., Aertker, A.W., Tran, J.K., Hillis, D.M., Harbin, C.R., Caldwell, C., Gallie, D.R., and Browning, K.S. (2010). Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability. Plant Mol Biol *74*, 249-263.

Lellis, A.D., Kasschau, K.D., Whitham, S.A., and Carrington, J.C. (2002). Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during potyvirus infection. Curr Biol *12*, 1046-1051.

Ling, J., Morley, S.J., and Traugh, J.A. (2005). Inhibition of cap-dependent translation via phosphorylation of eIF4G by protein kinase Pak2. EMBO J *24*, 4094-4105.

Liu, M.J., Wu, S.H., Wu, J.F., Lin, W.D., Wu, Y.C., Tsai, T.Y., Tsai, H.L., and Wu, S.H. (2013). Translational landscape of photomorphogenic Arabidopsis. Plant Cell *25*, 3699-3710.

Lu, J.Y., Bergman, N., Sadri, N., and Schneider, R.J. (2006). Assembly of AUF1 with eIF4G-poly(A) binding protein complex suggests a translation function in AU-rich mRNA decay. RNA *12*, 883-893.

Manjunath, S., Williams, A.J., and Bailey-Serres, J. (1999). Oxygen deprivation stimulates Ca^{2+}-mediated phosphorylation of mRNA cap-binding protein elF4E in maize roots. Plant J *19*, 21-30.

Mao, X., Schwer, B., and Shuman, S. (1995). Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene. Mol Cell Biol *15*, 4167-4174.

Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell *89*, 951-961.

Martinez-Rucobo, F.W., Kohler, R., van de Waterbeemd, M., Heck, A.J., Hemann, M., Herzog, F., Stark, H., and Cramer, P. (2015). Molecular Basis of Transcription-Coupled Pre-mRNA Capping. Mol Cell *58*, 1079-1089.

Mathonnet, G., Fabian, M.R., Svitkin, Y.V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W.C., Darzynkiewicz, E., Pillai, R.S., *et al.* (2007). MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. Science *317*, 1764-1767.

Mayberry, L.K., Allen, M.L., Dennis, M.D., and Browning, K.S. (2009). Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs. Plant Physiol *150*, 1844-1854.

Mayberry, L.K., Allen, M.L., Nitka, K.R., Campbell, L., Murphy, P.A., and Browning, K.S. (2011). Plant Cap-binding Complexes Eukaryotic Initiation Factors eIF4F and eIFISO4F: Molecular Specificity of Subunit Binding. J Biol Chem *286*, 42566-42574.

Mayberry, L.K., and Browning, K.S. (2006). In vitro translation of plant viral RNA. Current Protocols in Microbiology *16K*, 1-13.

Mayberry, L.K., Dennis, M.D., Allen, M.L., Nitka, K.A., Murphy, P.A., Campbell, L., and Browning, K.S. (2007). Expression and Purification of Recombinant Wheat Translation Initiation Factors eIF1, eIF1A, eIF4A, eIF4B, eIF4F, eIF(iso)4F, and eIF5. Methods Enzymol *430*, 397-408.

Minshall, N., Reiter, M.H., Weil, D., and Standart, N. (2007). CPEB interacts with an ovary-specific eIF4E and 4E-T in early Xenopus oocytes
420. J Biol Chem *282*, 37389-37401.

Missra, A., Ernest, B., Lohoff, T., Jia, Q., Satterlee, J., Ke, K., and von Arnim, A.G. (2015). The Circadian Clock Modulates Global Daily Cycles of mRNA Ribosome Loading. Plant Cell *27*, 2582-2599.

Mitchell, P., and Tollervey, D. (2001). mRNA turnover. Curr Opin Cell Biol *13*, 320-325.

Mitchell, S.F., Walker, S.E., Algire, M.A., Park, E.H., Hinnebusch, A.G., and Lorsch, J.R. (2010). The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. Mol Cell *39*, 950-962.

Monzingo, A.F., Dhaliwal, S., Dutt-Chaudhuri, A., Lyon, A., Sadow, J.H., Hoffman, D.W., Robertus, J.D., and Browning, K.S. (2007). The Structure of Eukaryotic Translation Initiation Factor-4E from Wheat Reveals a Novel Disulfide Bond. Plant Physiol *143*, 1504-1518.

Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F., *et al.* (2012). A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol Cell Biol *32*, 3585-3593.

Moury, B., Charron, C., Janzac, B., Simon, V., Gallois, J.L., Palloix, A., and Caranta, C. (2013). Evolution of plant eukaryotic initiation factor 4E (eIF4E) and potyvirus genome-linked protein (VPg): A game of mirrors impacting resistance spectrum and durability. Infect Genet Evol *27*, 472-480.

Mustroph, A., Zanetti, M.E., Jang, C.J., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J. (2009). Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proc Natl Acad Sci U S A *106*, 18843-18848.

Napoli, I., Mercaldo, V., Boyl, P.P., Eleuteri, B., Zalfa, F., De, R.S., Di, M.D., Mohr, E., Massimi, M., Falconi, M., *et al.* (2008). The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. Cell *134*, 1042-1054.

Nekrasov, M.P., Ivshina, M.P., Chernov, K.G., Kovrigina, E.A., Evdokimova, V.M., Thomas, A.A.M., Hershey, J.W.B., and Ovchinnikov, L.P. (2003). The mRNA-binding protein YB-1 (p50) prevents association of the eukaryotic initiation factor eIF4G with mRNA and inhibits protein synthesis at the initiation stage. J Biol Chem *278*, 13936-13943.

Niederhuth, C.E., Patharkar, O.R., and Walker, J.C. (2013). Transcriptional profiling of the Arabidopsis abscission mutant hae hsl2 by RNA-Seq. BMC Genomics *14*, 37.

Nodine, M.D., and Bartel, D.P. (2012). Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. Nature *482*, 94-97.

Okumura, F., Zou, W., and Zhang, D.E. (2007). ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. Genes Dev *21*, 255-260.

Osborne, M.J., and Borden, K.L. (2015). The eukaryotic translation initiation factor eIF4E in the nucleus: taking the road less traveled. Immunol Rev *263*, 210-223.

Osborne, M.J., Volpon, L., Kornblatt, J.A., Culjkovic-Kraljacic, B., Baguet, A., and Borden, K.L. (2013). eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. Proc Natl Acad Sci U S A *110*, 3877-3882.

Paik, I., Yang, S., and Choi, G. (2012). Phytochrome regulates translation of mRNA in the cytosol. Proc Natl Acad Sci U S A *109*, 1335-1340.

Pal, S.K., Liput, M., Piques, M., Ishihara, H., Obata, T., Martins, M.C., Sulpice, R., van Dongen, J.T., Fernie, A.R., Yadav, U.P., *et al.* (2013). Diurnal changes of polysome loading track sucrose content in the rosette of wild-type arabidopsis and the starchless pgm mutant. Plant Physiol *162*, 1246-1265.

Paquin, N., Menade, M., Poirier, G., Donato, D., Drouet, E., and Chartrand, P. (2007). Local activation of yeast ASH1 mRNA translation through phosphorylation of Khd1p by the casein kinase Yck1p. Mol Cell *26*, 795-809.

Park, E.H., Zhang, F., Warringer, J., Sunnerhagen, P., and Hinnebusch, A.G. (2011a). Depletion of eIF4G from yeast cells narrows the range of translational efficiencies genome-wide. BMC Genomics *12*, 68.

Park, H.C., Choi, W., Park, H.J., Cheong, M.S., Koo, Y.D., Shin, G., Chung, W.S., Kim, W.Y., Kim, M.G., Bressan, R.A., *et al.* (2011b). Identification and molecular properties of SUMO-binding proteins in Arabidopsis. Mol Cells *32*, 143-151.

Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., and Ramakrishnan, V. (2007). The Eukaryotic Translation Initiation Factors eIF1 and eIF1A Induce an Open Conformation of the 40S Ribosome. Mol Cell *26*, 41-50.

Patrick, R.M., and Browning, K.S. (2012). The eIF4F and eIFiso4F Complexes of Plants: An Evolutionary Perspective. Comp Funct Genomics *2012*, 287814.

Patrick, R.M., Mayberry, L.K., Choy, G., Woodard, L.E., Liu, J.S., White, A., Mullen, R.A., Tanavin, T.M., Latz, C.A., and Browning, K.S. (2014). Two Arabidopsis thaliana loci encode novel eIF4E isoforms that are functionally distinct from the conserved plant eIF4E. Plant Physiol *164*, 1820-1830.

Pereira, F.J., Teixeira, A., Kong, J., Barbosa, C., Silva, A.L., Marques-Ramos, A., Liebhaber, S.A., and Romao, L. (2015). Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity. Nucleic Acids Res *43*, 6528-6544.

Pierrat, O.A., Mikitova, V., Bush, M.S., Browning, K.S., and Doonan, J.H. (2007). Control of protein translation by phosphorylation of the mRNA 5'-cap-binding complex
256. Biochem Soc Trans *35*, 1634-1637.

Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res *35*, D61-65.

Ptushkina, M., Berthelot, K., Von der Haar, T., Geffers, L., Warwicker, J., and McCarthy, J.E.G. (2001). A second eIF4E protein in *Schizosaccharomyces pombe* has distinct eIF4G-binding properties. Nucleic Acids Res *29*, 4561-4569.

Qin, Y., Leydon, A.R., Manziello, A., Pandey, R., Mount, D., Denic, S., Vasic, B., Johnson, M.A., and Palanivelu, R. (2009). Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. PLoS Genet *5*, e1000621.

Rahmani, F., Hummel, M., Schuurmans, J., Wiese-Klinkenberg, A., Smeekens, S., and Hanson, J. (2009). Sucrose control of translation mediated by an upstream open reading frame-encoded peptide. Plant Physiol *150*, 1356-1367.

Rajjou, L., Gallardo, K., Debeaujon, I., Vandekerckhove, J., Job, C., and Job, D. (2004). The effect of alpha-amanitin on the Arabidopsis seed proteome highlights the distinct roles of stored and neosynthesized mRNAs during germination. Plant Physiol *134*, 1598-1613.

Rajyaguru, P., She, M., and Parker, R. (2012). Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins. Mol Cell *45*, 244-254.

Rhee, S.Y., Beavis, W., Berardini, T.Z., Chen, G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., *et al.* (2003). The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. Nucleic Acids Res *31*, 224-228.

Rhoads, R.E. (2009). eIF4E: new family members, new binding partners, new roles. J Biol Chem *284*, 16711-16715.

Richter, J.D., and Sonenberg, N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature *433*, 477-480.

Robaglia, C., and Caranta, C. (2006). Translation initiation factors: a weak link in plant RNA virus infection. Trends Plant Sci *11*, 40-45.

Robalino, J., Joshi, B., Fahrenkrug, S.C., and Jagus, R. (2004). Two zebrafish eIF4E family members are differentially expressed and functionally divergent. J Biol Chem *279*, 10532-10541.

Robichaud, N., and Sonenberg, N. (2014). eIF4E and Its Binding Proteins. In Translation and Its Regulation in Cancer Biology and Medicine, A. Parsyan, ed. (Netherlands: Springer Netherlands), pp. 73-113.

Rodriguez, C.M., Freire, M.A., Camilleri, C., and Robaglia, C. (1998). The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast

complementation and are differentially expressed during plant development. Plant J *13*, 465-473.

Rom, E., Kim, H.C., Gingras, A.C., Marcotrigiano, J., Favre, D., Olsen, H., Burley, S.K., and Sonenberg, N. (1998). Cloning and characterization of 4EHP, a novel mammalian eIF4E- related cap-binding protein. J Biol Chem *273*, 13104-13109.

Rosettani, P., Knapp, S., Vismara, M.G., Rusconi, L., and Cameron, A.D. (2007). Structures of the Human eIF4E Homologous Protein, h4EHP, in its m(7)GTP-bound and Unliganded Forms. J Mol Biol *368*, 691-705.

Ruiz-Martinez, J., Krebs, C.E., Makarov, V., Gorostidi, A., Marti-Masso, J.F., and Paisan-Ruiz, C. (2015). GIGYF2 mutation in late-onset Parkinson's disease with cognitive impairment. J Hum Genet *60*, 637-640.

Ruud, K.A., Kuhlow, C., Goss, D.J., and Browning, K.S. (1998). Identification and characterization of a novel cap-binding protein from Arabidopsis thaliana. J Biol Chem *273*, 10325-10330.

Sanfacon, H. (2015). Plant Translation Factors and Virus Resistance. Viruses *7*, 3392-3419.

Schmid, M.W., Schmidt, A., Klostermeier, U.C., Barann, M., Rosenstiel, P., and Grossniklaus, U. (2012). A powerful method for transcriptional profiling of specific cell types in eukaryotes: laser-assisted microdissection and RNA sequencing. PLoS ONE *7*, e29685.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.

Shetty, R.P., Endy, D., and Knight, T.F., Jr. (2008). Engineering BioBrick vectors from BioBrick parts. J Biol Eng *2*, 5.

Singh, R.K., Gonzalez, M., Kabbaj, M.H., and Gunjan, A. (2012). Novel E3 ubiquitin ligases that regulate histone protein levels in the budding yeast Saccharomyces cerevisiae. PLoS One *7*, e36295.

Song, A., Labella, S., Korneeva, N.L., Keiper, B.D., Aamodt, E.J., Zetka, M., and Rhoads, R.E. (2010). A C. elegans eIF4E-family member upregulates translation at elevated temperatures of mRNAs encoding MSH-5 and other meiotic crossover proteins. J Cell Sci *123*, 2228-2237.

Soto-Rifo, R., Rubilar, P.S., Limousin, T., de, B.S., Decimo, D., and Ohlmann, T. (2012). DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. EMBO J *31*, 3745-3756.

Stecker, K.E., Minkoff, B.B., and Sussman, M.R. (2014). Phosphoproteomic Analyses Reveal Early Signaling Events in the Osmotic Stress Response. Plant Physiol *165*, 1171-1187.

Sugiyama, N., Nakagami, H., Mochida, K., Daudi, A., Tomita, M., Shirasu, K., and Ishihama, Y. (2008). Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in Arabidopsis. Mol Syst Biol *4*, 193.

Sun, F., Palmer, K., and Handel, M.A. (2010). Mutation of eIF4g3, encoding a eukaryotic translation initiation factor, causes male infertility and meiotic arrest of mouse spermatocytes. Development *137*, 1699-1707.

Tan, N.G.S., Ardley, H.C., Scott, G.B., Rose, S.A., Markham, A.F., and Robinson, P.A. (2003). Human homologue of ariadne promotes the ubiquitylation of translation initiation factor 4E homologous protein, 4EHP. FEBS Lett *554*, 501-504.

Tao, X., and Gao, G. (2015). Tristetraprolin Recruits Eukaryotic Initiation Factor 4E2 To Repress Translation of AU-Rich Element-Containing mRNAs. Mol Cell Biol *35*, 3921-3932.

Terzi, L.C., and Simpson, G.G. (2009). Arabidopsis RNA immunoprecipitation. Plant J *59*, 163-168.

Tettweiler, G., Kowanda, M., Lasko, P., Sonenberg, N., and Hernandez, G. (2012). The Distribution of eIF4E-Family Members across Insecta. Comp Funct Genomics *2012*, 960420.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192.

Topisirovic, I., Svitkin, Y.V., Sonenberg, N., and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip Rev RNA *2*, 277-298.

Torti, S., Fornara, F., Vincent, C., Andres, F., Nordstrom, K., Gobel, U., Knoll, D., Schoof, H., and Coupland, G. (2012). Analysis of the Arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. Plant Cell *24*, 444-462.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Uniacke, J., Holterman, C.E., Lachance, G., Franovic, A., Jacob, M.D., Fabian, M.R., Payette, J., Holcik, M., Pause, A., and Lee, S. (2012). An oxygen-regulated switch in the protein synthesis machinery. Nature *486*, 126-129.

Uniacke, J., Kishan, P.J., Lachance, G., Francisco, C.B., and Lee, S. (2014). Cancer Cells Exploit eIF4E2-Directed Synthesis of Hypoxia Response Proteins to Drive Tumor Progression. Cancer Res *74*, 1379-1389.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res *40*, e115.

Vilela, C., Velasco, C., Ptushkina, M., and McCarthy, J.E.G. (2000). The eukaryotic mRNA decapping protein Dcp1 interacts physically and functionally with the eIF4F translation initiation complex. EMBO J *19*, 4372-4382.

Villaescusa, J.C., Buratti, C., Penkov, D., Mathiasen, L., Planaguma, J., Ferretti, E., and Blasi, F. (2009). Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. PLoS ONE *4*, e5213.

Von Arnim, A.G., Jia, Q., and Vaughn, J.N. (2014). Regulation of plant translation by upstream open reading frames. Plant Sci *214*, 1-12.

Von der Haar, T., Ball, P.D., and McCarthy, J.E.G. (2000). Stabilization of eukaryotic initiation factor 4E binding to the mRNA 5'-cap by domains of eIF4G. J Biol Chem *275*, 30551-30555.

von Stechow, L., Typas, D., Carreras Puigvert, J., Oort, L., Siddappa, R., Pines, A., Vrieling, H., van de Water, B., Mullenders, L.H., and Danen, E.H. (2015). The E3 ubiquitin ligase ARIH1 protects against genotoxic stress by initiating a 4EHP-mediated mRNA translation arrest. Mol Cell Biol *35*, 1254-1268.

Wang, A., and Krishnaswamy, S. (2012). Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. Mol Plant Pathol *13*, 795-803.

Wang, W., Ye, R., Xin, Y., Fang, X., Li, C., Shi, H., Zhou, X., and Qi, Y. (2011). An importin beta protein negatively regulates MicroRNA activity in Arabidopsis. Plant Cell *23*, 3565-3576.

Wang, Y., Fan, X., Lin, F., He, G., Terzaghi, W., Zhu, D., and Deng, X.W. (2014a). Arabidopsis noncoding RNA mediates control of photomorphogenesis by red light. Proc Natl Acad Sci U S A *111*, 10359-10364.

Wang, Y., Wang, X., Deng, W., Fan, X., Liu, T.T., He, G., Chen, R., Terzaghi, W., Zhu, D., and Deng, X.W. (2014b). Genomic features and regulatory roles of intermediate-sized non-coding RNAs in Arabidopsis. Mol Plant *7*, 514-527.

Wang, Y., Zhang, W.Z., Song, L.F., Zou, J.J., Su, Z., and Wu, W.H. (2008). Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in Arabidopsis. Plant Physiol *148*, 1201-1211.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE *2*, e718.

Xu, X., Vatsyayan, J., Gao, C., Bakkenist, C.J., and Hu, J. (2010a). HDAC2 promotes eIF4E sumoylation and activates mRNA translation gene specifically. J Biol Chem *285*, 18139-18143.

Xu, X., Vatsyayan, J., Gao, C., Bakkenist, C.J., and Hu, J. (2010b). Sumoylation of eIF4E activates mRNA translation. EMBO Rep *11*, 299-304.

Yanguez, E., Castro-Sanz, A.B., Fernandez-Bautista, N., Oliveros, J.C., and Castellano, M.M. (2013). Analysis of genome-wide changes in the translatome of Arabidopsis seedlings subjected to heat stress. PLoS ONE *8*, e71425.

Yeam, I., Cavatorta, J.R., Ripoll, D.R., Kang, B.C., and Jahn, M.M. (2007). Functional dissection of naturally occurring amino acid substitutions in eIF4E that confers recessive potyvirus resistance in plants. Plant Cell *19*, 2913-2928.

Yoshii, M., Nishikiori, M., Tomita, K., Yoshioka, N., Kozuka, R., Naito, S., and Ishikawa, M. (2004). The Arabidopsis cucumovirus multiplication 1 and 2 loci encode translation initiation factors 4E and 4G. J Virol *78*, 6102-6111.

Zhang, H., Zhou, H., Berke, L., Heck, A.J., Mohammed, S., Scheres, B., and Menke, F.L. (2013a). Quantitative phosphoproteomics after auxin-stimulated lateral root induction identifies an SNX1 protein phosphorylation site required for growth. Mol Cell Proteomics *12*, 1158-1169.

Zhang, J., Mujahid, H., Hou, Y., Nallamilli, B., and Peng, Z.H. (2013b). Plant Long ncRNAs: A New Frontier for Gene Regulatory Control. American Journal of Plant Sciences *4*, 1038-1045.

Zhang, Y., Sun, Q.Y., Yu, R.H., Guo, J.F., Tang, B.S., and Yan, X.X. (2015). The contribution of GIGYF2 to Parkinson's disease: a meta-analysis. Neurol Sci *36*, 2073-2079.

Zheng, Q., Ryvkin, P., Li, F., Dragomir, I., Valladares, O., Yang, J., Cao, K., Wang, L.S., and Gregory, B.D. (2010). Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. PLoS Genet *6*, e1001141.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W. (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol *136*, 2621-2632.

Zuberek, J., Kubacka, D., Jablonowska, A., Jemielity, J., Stepinski, J., Sonenberg, N., and Darzynkiewicz, E. (2007). Weak binding affinity of human 4EHP for mRNA cap analogs. RNA *13*, 691-697.