EVIDENCE FOR SUBFUNCTIONALIZATION OF THE *FLIP4* GENE FAMILY IN *ARABIDOPSIS THALIANA*

by

Michael Thomas Judge

Honors Thesis

Appalachian State University

Submitted to the Department of Biology
and The Honors College
in partial fulfillment of the requirements for the degree of

Bachelor of Science

August, 2015

Approved by:

_____
Annkatrin Rose, Ph.D., Thesis Director

_____
Ross Gosky, Ph.D., Second Reader

_____
Lynn Siefferman, Ph.D., Biology Department Honors Director

_____
Leslie Sargent Jones, Ph.D., Director, The Honors College

ABSTRACT

Gene duplication events are thought to be major drivers of diversity in angiosperm lineages. For most duplicated genes, one paralog is quickly lost to fractionation or pseudogenization. Alternatively, random mutations in one or both paralogs sometimes allow them to develop new or specialized roles. We hypothesize a current shift from redundancy to subfunctionalization between the paralogs *AtFLIP4-1* and *AtFLIP4-2* (Filament-like Protein 4), two genes coding for uncharacterized coiled-coil proteins in *Arabidopsis thaliana*. Phenotypes differ for single knock-out mutants (*flip4-1* plants are sickly with low fertility; *flip4-2* plants are indistinguishable from the wild-type). Furthermore, *FLIP4-1* is highly expressed in sperm and pollen while localization of *FLIP4-2* to the chloroplast has been observed in green protoplast culture. On the other hand, double knock-out mutants appear to be inviable, suggesting a redundant and vital FLIP4 function. To explore potential genetic interactions between the genes and their product functions, reverse-transcription quantitative polymerase chain reaction (RT-qPCR) was used to quantify the expression of both genes throughout the development of the plant in the different mutants. Furthermore, network analysis tools were used to explore their potential functions. Both genes were expressed at relatively low levels in most leaf and reproductive organ tissues, while *FLIP4*-2 appeared to be upregulated late in leaf development and in buds and siliques. The expression profiles of the two paralogs were not found to correlate with each other well, and their interaction partners differ, suggesting subfunctionalization. Furthermore, the expression pattern of *FLIP4-2* is consistent with a role in senescence or stress responses.

ACKNOWLEDGEMENTS

I owe innumerable thanks to those who have led, supported, and surrounded me throughout this project. Firstly, I am deeply indebted to Dr. Annkatrin Rose for her mentorship, support, and patience. I entered her lab nearly three years ago as a new Honors student with no experience; nonetheless, she taught me the basics of molecular biology and plant genetics and made herself available for constant guidance and interaction without which my efforts would have been null. She has also been completely supportive of my endeavors in other labs, and her guidance has been pivotal for me though these and on to a graduate program. Dr. Ross Gosky has been a helpful source of guidance and support throughout my university experience, and I would like to thank him for his comments in writing this thesis. Dr. Matt Estep has been instrumental in my academic and research career thus far. Although I did not work in his lab, Dr. Estep has made himself available for discussion of research, academia, and balance with personal life at every step of this project, and I have benefitted greatly from his efforts to cultivate a cohesive and vibrant atmosphere for undergraduate and graduate researchers in the department. Dr. Sarah Marshburn has been my go-to reference for work involving RNA and RT-qPCR and has worked extensively with me on obtaining and interpreting these results. She has also helped me to orient my career and has provided me with some of my first discussions in the philosophy of science and with 'bench psychology.' Likewise, Dr. Mary U. Connell's advice on research and academics has been vital for me over the last six months, and much of the equipment used in these experiments was obtained by her for use by the ASU Department of Biology. She is missed dearly after her recent passing. Thanks to Jim Sobieraj, Will Brennan, the Karatan Lab, Tongji Xing and various ASU Biology faculty for helpful discussions about

experiments, RNA quality control, and to Dr. Ece Karatan for the occasional pipette tips, optical qPCR strips, and use of the ABI 7300. Thanks to Dr. Howie Neufeld for our discussions about graduate school, science, and for his help with the statistical aspects of experimental design and analysis. I cannot thank Dr. Leslie Sargeant-Jones of the Heltzer Honors College enough. She has pushed me to go farther than I imagined was possible since my entrance into the Honors College after my freshman year, and her advice led to my first conversations with Dr. Rose and was the beginning of many other opportunities throughout my university experience. Dr. Angela Mead, Jess Yandow, and the Honors College staff have been a huge help with various administrational and funding aspects of this project. Special thanks to Dr. Johnny Walker for his support of this project through the Honors College Partnership Board Research Fund, and thanks to the Department of Biology and the Office of Student Research at Appalachian State University for various means of support for this project. Thanks to my mentors at NCSU and UGA (Drs. Linda Robles, Jose Alonso, Anna Stepanova, Javier Brumos, Jonathan Arnold, and Rich Meagher, and James Griffith) for major contributions to my research education and for continued support on aspects of this project. A warm thanks to the St. Elizabeth of the Hill Country and St. Bernadette's parishes, the members of the Catholic Campus Ministry, and my family and friends for their support and encouragement. Lastly, this project was initiated and largely conceived and planned by Keri Bear Reel; my work stands as an attempt to continue her major efforts. Thanks to her, Amanda Havighorst, and Alison Deshields for helping me with initial integration into the lab, and to other Rose Lab members throughout the years, including Dominic Balcon, Colby Richardson, and Matt Cole, whose data and analyses are mentioned in this report, as well as Becca Rzasa and Erica Stewart for their encouragement and support.

TABLE OF CONTENTS

ABSTRACT *ii*

ACKNOWLEDGEMENTS *iii*

FIGURES AND TABLES *vi*

INTRODUCTION

Background Information **1**
FLIP4 *Genes, FLIP4 Proteins, and* flip4 *Mutant Characteristics* **4**
Project Design **7**

MATERIALS AND METHODS

Plant Growth Conditions for RT-qPCR and Phenotypic Observations **8**
Tissue Collection and Homogenization **9**
RNA Extraction and Purification **9**
Assessment of RNA Quantity and Quality **10**
Reverse Transcription and quantitative Polymerase Chain Reaction (RT-qPCR) **11**
Primer Design and Validation **11**
SYBR® Green qPCR Reactions **14**
qPCR Data Analysis **14**
Network Analysis of FLIP4-2 **16**

RESULTS

Assessment of RNA Quality and Validation of RT-qPCR Reactions **17**
FLIP4 *Transcriptional Profiles in WT,* flip4-1*, and* flip4-2 **22**
Pearson and Spearman Correlations between Profiles in Leaves **24**
Analysis of Integrated FLIP4-2 Interaction Networks **25**

DISCUSSION

Development of a System for Measuring FLIP4 Expression **29**
Transcripts were Constitutively Detected, Even in Null Mutants **30**
Insights into the FLIP4 *Evolutionary Process* **31**
Potential Roles of the FLIP4 Proteins in Leaf Senescence or Stress Responses **34**
Concluding Remarks **37**

REFERENCES **39**

LIST OF TABLES

Table 1	Information and Specifications for Primers Used in qPCR	**14**
Table 2	Correlation Matrices for *FLIP4-1* and *FLIP4-2* Expression Profiles across Genotypes in Leaves	**24**

LIST OF FIGURES

Figure 1	*FLIP4* Mutant Phenotypes	**6**
Figure 2	Structure of qPCR Primers, UTRs, Introns, and T-DNA Insertions in the *FLIP4* Genes	**13**
Figure 3	Assessment of RNA Integrity by Native Gel Electrophoresis	**17**
Figure 4	Reaction Kinetics and Specificity of *FLIP4* qPCR Primers	**18**
Figure 5	Standard Curves of *FLIP4* qPCR Primers	**19**
Figure 6	*PP2A* Transcript Abundance across All Samples Normalized to Profile Average by Genotype	**20**
Figure 7	Expression Profiles of *FLIP4-1* and *FLIP4-2* Transcripts in WT	**22**
Figure 8	Expression Profiles of *FLIP4-1* (a) and *FLIP4-2* (b) in WT, *flip4-1* and *flip4-2* Leaves and Reproductive Organs	**23**
Figure 9	Network of 50 Most Closely Associated Genes Surrounding *FLIP4-2* and its Protein	**26**
Figure 10	Functional Annotations of 50 Genes Associated with FLIP4-2 and its Protein	**28**

INTRODUCTION

Background Information

One characteristic of higher plant species is their relative genetic plasticity, or ability to cope with changes to the genome in individuals. These changes can happen through a number of processes which range from single nucleotide polymorphisms (SNPs) to entire genome duplication events. Random mutations such as SNPs, transposition, and incomplete/nondisjunction during meiosis may affect traits immediately (between two generations) or accumulate, giving rise to changes in expression such as transcription and splicing, translation, or protein localization. It is by the nonlethal or beneficial accumulation of these changes over evolutionary time that new genes are thought to emerge. Any of these mutations can be harmful, neutral, or beneficial to the organism when taken together and may ultimately affect the biological fitness of the individuals in generations thereafter. In the case of gene or genome duplication, one copy of a gene may 'relax' and accumulate mutations, while function of the original gene product (i.e. protein or functional RNA) is conserved in the other copy. This may lead to the formation of a pseudogene or new/enhanced specialization via neofunctionalization, subfunctionalization, or gene dosage. This appears to be a major mechanism for the enhancement of genetic and functional diversity with decreased pressure for the individual to cope with the loss of a potentially vital gene product, as well as including regulatory sequences outside the genes themselves (Flagel and Wendel, 2009; Monson 2003; Reel 2013).

Gene duplication in flowering plants classically occurs by polyploidization (Leitch and Leitch 2008; Adams and Wendel 2005). For instance, *Arabidopsis* has undergone three

major genome duplication events since the divergence of the Brassicaceae, as well as many smaller events (Barker et al. 2009; The *Arabidopsis* Genome Initiative 2000). Furthermore, ancestral whole-genome duplications led to a great increase in regulatory genes involved in flower and seed development, suggesting a defining importance of this mechanism to angiosperms (Jiao et al. 2011). Allopolyploidization is also a major force in flowering plant genome diversification and expansion, and results in more rapid genome evolution and speciation (Feldman and Levy 2009). Evidence has emerged in *Arabidopsis* that many homologs share regulation patterns and machinery under cold stress, suggesting that duplication via allopolyploidization may result in an advantage in robustness of stress response networks (Akama et al. 2014). While transposition serves as a major rearrangement and duplication mechanism in larger plant genomes such as maize (Bennetzen 2000), only ~4% of complete transposed DNA elements in *Arabidopsis* appear to contain transcriptionally active genes. It is still possible that transposons play other roles in the deletion or nonfunctionalization of duplicated genes in *Arabidopsis*, given their mutagenic properties.

When a gene is duplicated, a number of fates are theoretically possible for the new paralogs, including nonfunctionalization, neofunctionalization, and subfunctionalization. Ohno initially theorized that selective pressures were alleviated for one paralog which could then acquire mutations which may be beneficial or deleterious, while the other copy could maintain the original function (Ohno 1970). The availability of several genomes at the turn of the century (e.g. The *Arabidopsis* Genome Initiative 2001) allowed for experimental analysis of mutation rates in all predicted duplicates. A critical point in the evolution of a new duplicate pair is preservation (i.e. prevention of nonfunctionalization and pseudogenization)

in the genome (Innan and Kondrashov 2010). Using substitution rates, selective pressures were found in fact lower for most genes when they first appear (Lynch and Conery 2000). Duplicates are most often silenced within a few million years (Lynch and Conery 2000), either by genome defense mechanisms such as antisense RNA (Adams 2007), or by epigenetic control via DNA methylation and histone modifications (Wang et al. 2014; Adams 2007).

While Ohno suggested that variation in gene function may arise due to a lack of selective pressure, it is actually both positive and negative pressure which causes loss of genes *and* their fixation in the genome (Rodin and Riggs 2003). For instance, while purifying selection is thought to act strongly on paralogs (Lynch and Conery 2000), paralogs have been found to evolve more rapidly than orthologs with comparable rates of divergence. This may be due to either immediate positive selection or by relaxation of purifying selection (Kondrashov et al. 2002). Epigenetic processes appear to aid in subfunctionalization as well as silencing by applying selective pressures, and may add variety. For example, DNA methylation appears to influence exon number, expression level, and mutation rate (Wang et al. 2014), and epigenetic silencing by cytosine methylation can protect duplicates from the fate of pseudogenization and may even contribute to differential expression in different tissues and developmental stages. This nascent form of subfunctionalization can lead to negative selective pressures, which would provide a route to evolution of new gene functions (Riggs and Rodin 2003). Interestingly, retention of genes encoding certain types of peptide functions seems to be favored after duplication. For instance, in *Arabidopsis*, proteins with predicted signal peptide regions are significantly overrepresented in young paralogs, pointing to a possible link between selection processes and gene-environment interactions.

Furthermore, the distribution of paralog functions appears to be skewed toward genes encoding membrane-bound and secreted proteins, and finer distinctions within these functions appear to vary with organism complexity (Kondrashov et al. 2002). Carvunis et al. provided evidence in 2011 that evolution, including that following from gene duplication events, is largely driven by functional system-level properties like coexpression and protein-protein interactions. Paralogs in *Arabidopsis* evolve less quickly than their interactions, and a correlation between coexpression and interaction partners exists for paralogous gene pairs (*Arabidopsis* Interactome Mapping Consortium 2011).

FLIP4 *Genes, FLIP4 Proteins, and* flip4 *Mutant Characteristics*

One of the most recent gene duplication events in *Arabidopsis* gave rise to two surviving paralogs (*FLIP4-1*, At3g50910 and *FLIP4-2*, At5g66480) encoding the FILAMENT-LIKE PROTEIN4 (FLIP4) family; and FLIP4-2, respectively, both of which contain coiled-coil domains (Cole 2014, Reel 2013). Given the abundant and vital roles of coiled-coils in plants, including structural/physical roles, trafficking, protein-protein interactions, and signal transduction, the roles of the FLIP4 proteins seem likely to be important to normal cell function. The FLIP4 proteins are also predicted to be membrane-associated. We hypothesize that this pair is in the process of diverging specialization in function and/or expression patterns, possibly including critical roles in key plant functions such as photosynthesis and developmental processes such as seed setting.

FLIP4-1 has a full genomic nucleotide sequence length of 2726 bp and corresponding polypeptide length of 447 amino acids, while *FLIP4-2* is 2446 bp and codes for 444 amino acids (TAIR). The gene structures are similar, with a large exon, followed by two smaller

exons. Evidence also exists for alternative splicing in FLIP4-2 transcripts (NCBI). The functions of the FLIP4 proteins are unknown (TAIR), although Yeast Two-Hybrid assays (thebiogrid.org; Chatr-Aryamontri et al. 2015) and software predictions have indicated putative interaction partners, especially for FLIP4-1. The FLIP4 proteins do not likely interact with each other (thebiogrid.org; Chatr-Aryamontri et al. 2015).

To investigate the genetics and potential phenotypes related to the FLIP4 genes, seed lines have been obtained containing T-DNA insertion mutant alleles for each gene and genotyped by PCR (Reel 2013). *flip4-2* plants are phenotypically indistinguishable from the wild-type (WT), while *flip4-1* plants exhibit reduced growth (Figure 1) and reduced seed set in constant light conditions. Notable phenotype characteristics are long petioles and wrinkled, revolute leaves which sometimes yellow and senesce prematurely. Interestingly, the reduction in fitness in *flip4-1* is not observed in short or long day conditions, which may allude to a relationship between *FLIP4-1* expression and response to diurnal cycles or other light-dependent phenomena. Furthermore, *flip4-1* mutants appear to maintain color longer than the *flip4-2* mutant or WT in unusual drought or dark conditions (personal observation). Of seeds set from *flip4-1* x *flip4-2* crosses, planted with 100% germination rate, and grown to maturity in our lab, no double mutants were found.

Microarray data suggest that *FLIP4-1* is highly expressed in pollen and in sperm cells (Toufighi et al. 2005). Combined with double-mutant inviability thus far, this suggests shared roles between FLIP4-1 and FLIP4-2 and a critical role for the former in reproductive processes. A probe for *FLIP4-2* was not included on the ATH1 genechip (Affymetrix) used in most of the microarray studies we found, but immunoblot/immunohistochemistry studies (Richardson 2012), and targeting domain prediction/localization studies place FLIP4-2 in

Figure 1. *FLIP4* Mutant Phenotypes. Seed lines containing knockout mutations for both FLIP4 protein functions by T-DNA insertion into exons have been grown in our lab. The *flip4-1* mutant is sickly-looking and notably smaller than the WT. The *flip4-2* mutant, on the other hand, is consistently healthy and often appears to be larger than the WT. These trends are observed primarily in rosettes and from Days 5-10 to about Day 25-30. All images are scaled equally. Plants shown are representative of 10 pots with relatively even growth conditions. Images were processed (for clarity) in Adobe Photoshop Elements 8.

chloroplasts, suggesting that it may be preferentially expressed in green tissues and involved in photosynthesis. When the phenotypic observations and microarray data are taken together, it appears that the complete loss of FLIP4 may result in male-specific sterility derived from gametophyte inviability or a disruption of pollination (Reel 2013).

Although the FLIP4 proteins are very similar in sequence, they may have acquired different developmental or spatially-oriented roles. If FLIP4-1 and FLIP4-2 have begun to diverge in function and/or localization, this should be evident in their gene expression patterns. More specifically, an existing association between expression levels of these two genes (directly dependent, differential or redundant, or independent) in a given tissue at a given time could elucidate spatial expression differences in the proteins as a result of diverging promoter function. Elevation or lack of expression of each gene might also be associated with developmental stages (particularly *FLIP4-1* and organism-level processes leading up to pollen and seed production), potentially leading to clues about the roles and functions of these proteins as such. mRNA abundance for individual genes is generally

accepted as a proxy for their expression levels, as quantified in microarray, RNA-seq, or RT-qPCR.

Project Design

In order to further assess potential correlations between *FLIP4-1* and *FLIP4-2* expression, quantitative expression profiles were generated for *FLIP4-1* and *FLIP4-2*. Total RNA was extracted and purified from specific tissues or organs of *flip4-1*, *flip4-2*, and WT plants throughout development, followed by reverse transcription (RT) and quantitative polymerase chain reaction (qPCR) of the resulting total cDNA with gene-specific primers, with particular attention to quality control (Bustin et al. 2009). Data were analyzed using the Pfaffl (2001) method and profiles were compared statistically using Spearman and Pearson correlations. To gain insight on potential FLIP4 functions, a basic interaction/co-expression network analysis was also carried out in addition considering qualitative aspects of the profiles in the WT.

MATERIALS AND METHODS

Plant Growth Conditions for RT-qPCR and Phenotypic Observations

Wild-type (Columbia ecotype, WT), and *FLIP4-1* and *FLIP4-2* T-DNA insertion mutant seeds (Salk lines 074693 and 033887; *flip4-1*and *flip4-2*, respectively) were obtained from the *Arabidopsis* Biological Information Resource (ABRC, The Ohio State University). Plant growth density was planned according to plant age at harvesting (45-60 seeds per pot, Days 10-20; 5 per pot, Days 25-40; 1 per pot, reproductive organs) as suggested by methods released by the ABRC (ABRC) and following preliminary phenotypic observations. Plants for each genotype and development-time point were grown in three pots constituting a triplicate of each sample. Furthermore, triplicates were divided so that a 'group' consisted of one pot of each genotype for a given timepoint. Seeds kept long-term at 4°C were sowed on pre-moistened MetroMix 360 (Sun-Gro, Agawam, MA USA) mixed with Osmocote 14-14-14 Slow-Release Fertilizer beads (Scotts, Marysville, OH USA; presoaked in hot water as described in Reel 2013; 8.33g per 64 pots) in 6.5-cm diameter pots, which in turn were placed in 8 x 4 –pot flats. Seeds were stratified for 4 days at 4°C then transferred to constant light (150-200umol/m2/s; Li-Cor photometer, Lincoln, NE USA) at 20-23°C in a Conviron Adaptis-A1000 Growth Chamber with *Arabidopsis* Chamber Kit (Hendersonville, NC USA). Humidity was decreased by offsetting the lid of the flat from cotyledon emergence until just after the emergence of the first true leaves (Days 0-5) and was not controlled thereafter. To compensate for variation in light intensity within the growth chamber, group position was randomized every 3-4 days until Day 30-35 (positions determined using a random number generator in Microsoft Excel 2010). After 35 days, plants were grown at a constant 21°C.

Plants were watered as needed (no less than once every four days). Collections were performed in late afternoon for plants younger than 35 days.

Plants grown for phenotypic observations were grown under the conditions stated above. Pots containing five or more plants were thinned by carefully removing all but two well-spaced, representative plants (Day 18, WT and *flip4-2* mutant; Day 20, *flip4-1* mutants). Images were recorded with a Canon Powershot SX40 (Annkatrin Rose) or an iPhone 6 (only for Day 15) camera and cropped in Adobe Photoshop Elements 8.0 using the pot diameter for scale.

Tissue Collection and Homogenization

Three samples from the leaves (every five days, Days 10-40) and mature buds, flowers, and green siliques of different individuals were excised with clean scissors from plants of each genotype (Columbia ecotype, WT; *flip4-1* knockout; and *flip4-2* knockout), immediately flash-frozen in liquid nitrogen and stored at -80°C until homogenization. Petioles and cotyledons were excluded. Samples were finely ground with clean, pre-chilled mortar and pestles on dry ice and under liquid nitrogen such that the material was never thawed. Aliquots of 100 mg of dry (i.e. liquid nitrogen was evaporated) powdered tissue were weighed in pre-chilled microfuge tubes and kept at -80°C until RNA extraction.

RNA Extraction and Purification

Homogenized tissue (100 mg) was quickly lysed by constant rigorous vortexing (VWR Analog Vortex Mixer; Radnor, PA USA) in 1.5-mL microfuge tubes in 300uL RNA Extraction Buffer reported in Box et al. 2011 (0.1 M Tris pH 8.0, 5 mM EDTA pH 8.0, 0.1 M

NaCl, 0.5% SDS; 1% 2-mercaptoethanol added and heated to 60°C directly before use) until tissue was thawed (<30 s). An equal volume of acidic phenol:chloroform (P:C; 1:1 v/v phenol pH 4.5:chloroform; Amresco, Solon, OH, USA) was immediately added, followed by alternate vortexing/shaking at 1400 rpm and room temperature for 10 minutes (Eppendorf Thermomixer R; Hauppauge, NY USA) and debris was pelleted by centrifugation (room temperature, 20000 x g, 5min; VWR Galaxy 20R; Radnor, PA USA). The aqueous layer was removed and added to 300uL P:C, followed by multiple rounds of rigorous vortexing and incubation on ice (alternating ~30 s each) and centrifugation at 4°C to encourage phase separation. The resulting aqueous phase was collected and washed of residual phenol in 300uL of chloroform, then precipitated in 1.5 volumes of isopropanol and 0.2 volumes of sodium acetate (3M NaOAc, pH 5.2) at -80°C until frozen (> 15min). The procedure was either continued or samples were kept until a later date at -80°C, after which RNA was pelleted by centrifugation (room temperature, 20000 x g, 30min; VWR Galaxy 20R), washed twice in an excess of ethanol (70%, 600uL; Sigma-Aldrich, St. Louis, MO USA), pelleted by centrifugation, dried under a laminar flow hood, and resuspended in 300uL nuclease-free water (UltraPure™ distilled water, Life Technologies, Grand Island, NY USA) by rigorous vortexing. The RNA was then extracted twice in P:C, precipitated, washed, and resuspended as described above.

Assessment of RNA Quantity and Quality

Quality and quantity were verified spectrophotometrically using a NanoDrop 2000 (Thermo Scientific, Waltham, MA USA). For each RNA sample, 0.5 ug were electrophoresed on a native agarose gel (2% in TAE, 0.5 ug/mL ethidium bromide, 100 V,

~90 min) and assessed for quality by 18S and 28S ribosomal RNA bands following manufacturer's recommendations sans heating (RNA Loading Dye, 0.5 ug RNA sample; Thermo Scientific, Waltham, MA USA) and compared to a 2-log DNA ladder (New England BioLabs, Inc., Ipswich, MA USA). Gels were visualized under UV light and imaged in a ProteinSimple® AlphaImager HP lightbox (San Jose, CA USA).

Reverse Transcription and quantitative Polymerase Chain Reaction (RT-qPCR)

Reverse Transcription (RT, cDNA synthesis) reactions were carried out using the High-Capacity cDNA Reverse Transcription kit (20-uL reactions, 1 ug per reaction) according to the manufacturer's protocol (Applied Biosystems, Foster City, CA). A reaction lacking only reverse transcriptase (RTase; RTase-minus control) was carried out for each sample in order to assess the extent of genomic DNA (gDNA) carryover. Aliquots of all cDNA synthesis reactions were kept at -20°C and thawed no more than twice.

Primer Design and Validation

Primers for the *FLIP4-1* and *FLIP4-2* cDNA were designed to be specific to each paralogous transcript, but not splice variants, using sites identified by PrimerQuest (qPCR Intercalating Dyes setting; Integrated DNA Technologies, Coralville, IA USA), Primer-BLAST (NCBI), and manual sequence screening assisted by L-Align (Figure 2; Huang and Miller 1991). *PP2A* normalization gene primers (designed for the A3 subunit of protein phosphatase 2A) were reported in Czechowski et al. (2005), and were ordered commercially (10 nM, desalted lyophilized; MWG Operon, Huntsville, AL USA). Primer pair candidates were screened for target specificity, efficiency, and optimum annealing temperature against a

representative set of pooled cDNA samples in standard PCR. The efficiency of each pair was assessed by triplicate standard curves across six serial decimal dilutions in SYBR® Select I assays (1:2.5-1:250,000 x cDNA) and kinetics of all reactions were found to be linear ($R^2 \geq 0.98$) and relatively efficient (*FLIP4-1*, E = 0.98; *FLIP4-2*, E = 0.88; *PP2A*, E = 0.87, data not shown), with the exception of *FLIP4-1*. The *PP2A* primer pair efficiency reported in Czechowski et al. was 0.93. The formula used for calculating efficiency E was $E = 10^{1/slope} - 1$. Primer specifications are summarized in Table 1. Amplicon sizes were confirmed by agarose gel electrophoresis. To confirm discrimination between the paralogs, each primer was tested against the purified cloned cDNA of the other paralog in qPCR reactions. Although both sets primed amplification of the cDNA for both paralogs individually (a good example of the high sequence similarity between them), each primer set showed strong preference for the cDNA of the paralog for which it was designed versus that of the other. In fact, the difference in primer sensitivity for the cDNAs of the different paralogs was consistently and signficantly nonzero ($p < 0.001$), with a typical observed C_t difference of 17-18 cycles (C_t; cycle or iteration of the reaction during which the fluorescence of the reaction rises above a threshold value assigned across all reactions of a particular assay type to a log-linear part of the reaction curve by ABI 7300 qPCR machine software).

FLIP4-1 mRNA

ATAGAGGAAGAAGCAAAGAAGAAGAAAAAAAAAAAAACAGAGCGTTTCAAATATTTGCGGAGATACAAAGCTGAGAAGAGAT
AAAAAAAGCCTTGGAGAAGAGGTTTTCTTCTTCTTCTTCTTCTTCTTCTTCCCTCCTCTCCTCTTAACGATGTCGTATCAC
GCTTCTTCTTATCTTCTCCGCCGTCGTCTTACTTCCGTATAACCTCGATCTCTCACCGTTGAATCTTCCTCAGCTTTCTTCCGTGGA
GATCTCCTTCGCGCTCGAGAGCCATGCCTACTTTTTCTGCTATTGCTTTGGATAGAATGCTAGAGCCAGGGGCTTCCACGTCTGT
CGAGAGTGTTCCAAGCACAACTAATTTGTTTTATTCAAAGCCACCGATATCAAAACTGGAGAAAGGTAAAGGTAAATTACCTAA
TGAGAGGACAGTTACTCGTCCTCTGATGTCACCTGCTCTCTACGCCACACCTGATGCGATTCCGCTTCCTAATTCGCCATCGTCT
TTCCCACCGTCGCCTTATATCATCAACCACAAATCACGTGGACCTCCACGTCTTCTTAAAAGCTCTTCTGAGGCTAATGTTGTGT
CTTCTTCC<SALK_074693>CACCAGAAGACGCTGGAAGAGGAAACTATTACTGCTGAAACAGATGTAAAGGTTTCCCCTCGGA
GGAGGTCCACTTCCTTCTCGTTTCCTATCACTGAGGTTACTGAAGATGATTATTCCAATGGAGTTCATGCTCGTACGGTTGGAA
ACTATAATTTTGATGGTATTGTTGATGGTCCTGTTGGAAATTGGAGTCCGCTTGATGGGAAAAGTGGGAATGGGAAG<SALK_
074693>TCAGAGTTGGATAACGCTGCTAATGGTTTAGAACGGGTAAATGGTCTGTCTGAGCCAGTCCCTATTAAAACAGACAA
GGAGAGTGAGTCAGAAGATTTCTATGACCCAGGTGAATCAGCAAGCTTCACAAGTAACACAGACGTAGAGGGTGATGCTGGG
GACGAAAGTTCACAAAGACTTGCCACACCTGTGGGAGAGTTTTATGATGCTTGGGATG(18920954)intron(18921328)AACT
ATCAACCGACAGTGGAATGCAGAGTTCAGTTAATAACATCGAATCTGAATTAAGTGAGATAAGGCTGAGTTTACTAATGGAAA
TCGAGAAGAGAAAACAGACAGAGGAGGCTTTGGAGCAGATGCAAATTCACTGGCAGAGACTCCGTGAGCAGATGGCTCAGG
TGGGTCTGTTCGTTCCTATTGATCCTACCGCCTCCACCAACAACATGAATCTATCAGAAGAACTACGCTGTCAACTCGAGATTGC
GAGGTTTGTCTCTGACTCTTTAGGCAGAGGTATGGCAAAGGCCGAGGTAGAGATGGAGATGGAATCTATGCTCGAAACTAAG
AACTTTGAAATCACACGGTTGTCTGACCGTCTACACTACTACGAGGCCGTGAACAGAGAGATGTCTCAACGAAACCAAGAAGC
CATAG(18921752)intron(18921845)AGGTGGCACGACGAGAGAGGCAGAAGAGAAAGAAGAGGCAGAGATGGATTTGGG
GATCAATTGCAGCAACCATTACTCTAGGAAGCGCGGCATTGGCATGGTCTTACATTCCTGCAGCCAAACCATCATCTGAAGTAT
CTCTTAAGGATGAATGATTAAGTAAGTCTCTTGTATACGTAGTGGGTTTGATGCAAGCAAGAGAAATAAAGATACAGGACGTT
TAGCGTACGCCTTGTGAATAGAAATATATACATATTTGGTTATTGTTTTAACTGTTCTCAAGAGGATGTATAGCTCCGAGTAGCT
GCAAACACATTTTTCTTGCGGTGGCATACTGATGAGTGGCAGCCACAGTCACTAATATTTGGTTCGGTTTGGTTCTTATTAACAT
TCA

FLIP4-2 mRNA

GGAGAAGAGGTTTCTCTTTTTTCTCTCCTCCTCAGTATATATCTAATCTCTGTTTGCTCTTTTTTCGCACCATGTCATATCAATCTA
CCATCAATCGATCGTCTCTTCTTCCTCCGTCTGCGTTGATCTTTTCGTTTCGTCTTCTCCGATTGATGCTCTGTCGCGAGAGTCAT
GCCTACATTTTCTGCTGCCGCTTTGGGTAGATCACTAAATTCAGGGACGTCCTTATCTAGTAAATTTCCAAGCACACTGCAGTCA
AAGCCGTCTATTTTGAATGATGAAAGTAAGCAACCTAAAGAAAAGACATTCACTCGTCCTCAGATGTCACCTTCTCTATATGCC
ACTACTAAAGAAATCCCTCATCCCAATTCGCCTTCTTCCTACCCTCCATCTCCCTATATCATCAATCACAAAGCTCGTGGGCCTGT
TCTTTTCAATAGGGACTCTGAGGTGGATGGGCCTTCTCACCCGATAACCTCTGGTGAGGAAAAGATTAGTGGTAATGTAGATG
TGGAGGCTACTGCCTCCTTGAGCAAGTCCACATCCCTATCATTTCCTATCACCGAGGCTATTGCGGTGGATCACACCAATGGTG
TTCACACCCAAGGCATCCATGAACGTCCTGTTTGGGATTGCAGTCCCCCTCTTGGGACATTTTTGAACGAGAAGTCGGGAAGA
GATATCAGCAATGGTGGGATTGGAAGTAATAATGCTACCAGTAATTTAGAATGGCAAAGTTACTTATTGGAGCCGGTAAGAAT
AAAAGCAGACAAGGAGCTTGAGCCTGAAAATTTCTATAACCCAGGTGAATTAGTAAGCTTCACAAGTAACACAGAGGTAGAA
GATTTCGAGAGAGCTGAAAGCTCACACAGTCTTGCTACACATGTTGGAGAGTTTTATGATGCTTGCGATG(26545735)intron(
26545412)AACTATCAACTGACAGTGGAATGCAGAGTTCAGCTAACAACATTGAATCTGAGGTAAGGGAGATGAGATTGGGTT
TACTAATGGAAATCGAAAGGAGGAGGCAGGCAGAGGCGACTCTAGAGCAAATGCAAGTTCATTGGCGGAGGCTCCGTGACC
AGTTGGCCGATGTGGGTATGTTCCTTCCTCTTGATCCCACCAGAAGCCAATACAGCATGAACCTAGCTGATGAGCTACGTTGTC
AACTTGAGGTCACCAGGTTTGTCTCCGACACACTAGGCAGTGACTTGGCCAAGACAGAGGTAGAGATGGAGATGGAAGCTGA
GCTCGAAGCTAA<SALK_033887c>GAACTTTGAAATCACCCGGTTGTCTGACCGTCTTCATTATTACGAGACTGTGAACCAGG
AAATGTCCCAGCGCAACCAAGAGGCCATAG(26544990)intron(26544906)AGGTGGCAAGGCGAGACGGGCAAAAGAGAA
AGAGGAGGCAGAGGTGGATTTGGGGATCAATAGCTGCAACCATCACACTCGGGAGTGGAGTCTTGGCCTGGTCGTACCTTCC
TCCCGGAATGTTATCATCAGATGAAGCTCAACCTCAGCTGTCTCCAAAAGATTCATAATGAGGTTGAATCAAAGCTTCAG<SAL
K_033887c>GCCTTGTGAATAAAAATAGGAAAATGGGTTTTGTTTTAACCTTTTCTAGAGAATGTATAGAACATTCATGATGAC
GACATAACTCAGCTTCATACGATCAAGGTCATACCAATTTGGTTCTGTTTGTGTCGCTTATTGACTTGCAATGTTATTATTAGCC
AGAAGTTCCGTGCC

Figure 2. Structure of qPCR Primers, UTRs, Introns, and T-DNA Insertions in the *FLIP4* Genes. There is considerable similarity in the sequences of *FLIP4-1* and *FLIP4-2*. The structures of the genes are also very similar. T-DNA insertions (Purple) are ~4484bp in size (Alonso et al. 2003) and were therefore not shown, and only apply to mutants. The two T-DNA insertions in each mRNA represent the ambiguity in the two locations of the insertion based on sequences from NCBI. In *flip4-1* mutant mRNA the first exon contains the insertion, and primers (Dark Yellow) are found downstream in the sequence in the second exon. In the *flip4-2* mutant it is unclear whether the insertion falls between the primers in the second exon or after the coding region altogether in the 3' untranslated region (UTR; UTRs, red text). The latter is more likely, since the amplicon obtained is of the expected size, even in the *flip4-2* mutant. Intron locations, light blue; Start Codons, Green; Stop Codons; Red. Images were processed (for clarity) in Adobe Photoshop Elements 8.

Table 1. Information and Specifications for Primers used in qPCR

AGI (Gene)	Sequence (5'-> 3")	Efficiency	Primer T_m (C)	Amplicon Length (bp)	Amplicon T_m (C)
At1g13320	F - TAACGTGGCCAAAATGATGC	0.87	60	61	76.0
(*PP2A*)	R - GTTCTCCACAACCGCTTGGT				
At3g50910	F - GTGGGTCTGTTCGTTCCTATT	0.98	59.4	98	79.7
(*FLIP4-1*)	R - AGAGACAAACCTCGCAATCTC				
At5g66480	F - ACACTAGGCAGTGACTTGGC	0.88	59.4	136	78.4
(*FLIP4-2*)	R - TGAAGACGGTCAGACAACCG				

SYBR® Green qPCR Reactions

Quantitative PCR reactions for *FLIP4-1* and *FLIP4-2* were prepared following the manufacturer's protocols and consisted of 7.5 uL SYBR® Select Master Mix (2X; Life Technologies, Grand Island, NY USA), 100 or 200 uM of each primer (forward and reverse), and 4 uL cDNA (diluted 1:40 in Ultrapure™) water) and were carried out on an ABI 7300 in MicroAmp® Optical 96-Well Reaction plates with Optical Adhesive Covers (Life Technologies, Grand Island, NY USA). *FLIP4-1* and *FLIP4-2* reactions were carried out at 59.4°C using 100 uM of each primer (200 uM total), while *PP2A* reactions were carried out at 60°C using 200 uM each primer (400 uM total) as reported in Czechowski et al. No-template controls substituting water for cDNA were run on each plate. A standard melting curve program was run at the end of each sample to validate the specificity and success of each reaction.

qPCR Data Analysis

Three technical replicates for amplification of *FLIP4-1* and *FLIP4-2*, and *PP2A* in each biological sample were averaged. Means were then normalized to *PP2A* expression

using the Pfaffl method ($R = \frac{E_{Target}^{\Delta C_t\, target(control-treatment)}}{E_{Reference}^{\Delta C_t\, reference\,(control-\,treatment)}}$, where 'E' is the Pfaffl efficiency, 'control' is the average C_t across all reactions of a profile in a given genotype, and 'treatment' is the average C_t of technical replicates for each biological replicate in that profile; Pfaffl 2001). Note that, in this case, the treatment is time or tissue, and R=1 is the mean for a given profile. Since this method assumes a slightly different primer efficiency calculation, primer efficiencies were first re-calculated accordingly ($E = 10^{1/slope}$; Pfaffl 2001). Normalized data for each biological replicate were averaged and the standard error was taken to assess biological variation. Linear and nonlinear similarity between expression profiles was measured using Pearson product moment and Spearman-order correlations, respectively (Spearman 1904; D'haeseleer 2005; de Siquiera Santos et al. 2013; Brazma and Vilo 2000), with corresponding hypothesis tests in Minitab 17 (Minitab, Inc.). P-values of 0.05 and less were considered significant, while p-values of 0.1 or less were also noted. Plots were produced and normalization calculations were performed in Microsoft Excel 2010.

RNA quality and success of cDNA synthesis were also assessed by normalizing the transcript levels of *PP2A* across all samples of a given genotype. Reactions which showed an unusually low *PP2A* abundance were excluded from further analysis (> 2–fold decrease after normalization to the average C_t for *PP2A* by the $\Delta\Delta C_t$ method, given that 2-fold changes in transcript abundance are generally accepted as up- or down-regulation; Xiong 2006; Livak and Schmittgen 2001). Since the same amount of RNA was used in each cDNA synthesis reaction, and *PP2A* is stably expressed across development, it stands that its relative abundance as quantified in qPCR reactions should be comparable across all samples of the same genotype. Particularly concerning samples were those which showed very low transcript abundance; these tended to be partially degraded as observed by RNA native

electrophoresis. Additionally, a reaction type (sample-primer combination) was excluded if the C_t of the RTase-minus control was less than 4 cycles away from that of the technical average of the same standard RT reaction type.

Network Analysis of FLIP4-2

The online GeneMANIA tool (Warde-Farley et al. 2010; Zuberi et al. 2013) was used to search for the 50 genes and 10 attributes most closely connected to each paralog *FLIP4-1* and *FLIP4-2*, entered respectively as AGI numbers At3g50910 and At5g66480, according to the default weighting method. All available GeneMANIA Networks for *Arabidopsis thaliana* were used, including 316 sources for Attribute, Co-expression, Co-localization, Genetic interactions, Physical interactions, and Shared protein domain relationships. The associated functions of the genes and their respective False Discovery Rates (FDR) were obtained from the tool, and functional annotations of FDR > 0.01 were considered. The FDR is a multiple comparisons method used to control the higher chance of *significant* false positives within a large family of comparisons (Benjamini and Hochberg 1995). The 50-gene network was imported and expanded by searching for the AGIs against all *Arabidopsis* databases accessible within Cytoscape 3.2.1 (Shannon et al. 2003) and merging by taking the union of the expanded networks. Roughly half of the resulting nodes (~1500) were unclustered. Individual clusters were identified using MCODE with Fluff. Enriched functions were identified for each cluster using BiNGO (Maere et al. 2005) under default settings for *Arabidopsis thaliana*. Clusters were identified for the whole network and significantly overrepresented functions were considered (hypergeometric, FDR-corrected p-value < 0.05; Breeze et al. 2011; Maere et al. 2005; Benjamini and Hochberg 1995).

RESULTS

Assessment of RNA Quality and Validation of RT-qPCR Reactions

RNA quality and quantity was assessed spectrophotometrically and integrity was assessed by visualizing ribosomal (18S and 28S), transfer, and other unknown RNA bands by native gel electrophoresis (Figure 3; $1.86 < A_{260}:A_{280} < 2.15$; $2.24 < A_{260}:A_{230} < 2.62$; Yield = 45.9 ± 35.1 ug Nucleic Acid/100mg tissue; Mean ± SDM). Extractions were found to be reasonably consistent with regard to RNA size selection and most samples did not exhibit observable degradation, e.g. due to contaminating endonucleases (RNases) and freeze-thaw cycles. RNA samples which were missing either ribosomal RNA band or showed obvious smearing (e.g. red stars, Figure 3) or otherwise inconsistent selection of size during extraction were excluded from data analysis.

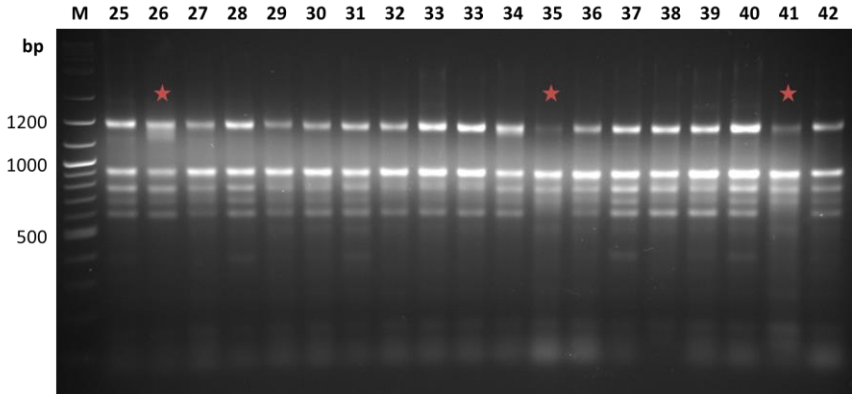

Figure 3. Assessment of RNA Integrity by Native Gel Electrophoresis. After spectrophotometry, an aliquot of each RNA sample (0.5ug each) was electrophoresed on a 2% Native RNA gel in RNA Loading buffer containing formamide (Thermo Scientific) and compared to a 2-log DNA ladder (NEB, far left lane 'M'). This image shows a representative set of samples. Note that 18S and 28S rRNA bands (~900 bp and ~1200 bp, respectively) and other bands are well-resolved, indicating high-quality RNA. Also note the presence of low-molecular-weight bands (likely tRNA). Starred samples were excluded because of suspected degradation (i.e., smearing, sample # 26 or disproportionate rRNA bands, # 35 and # 41) or dilution error in cDNA synthesis (not shown). Image was processed (for clarity) in Adobe Photoshop Elements 8.

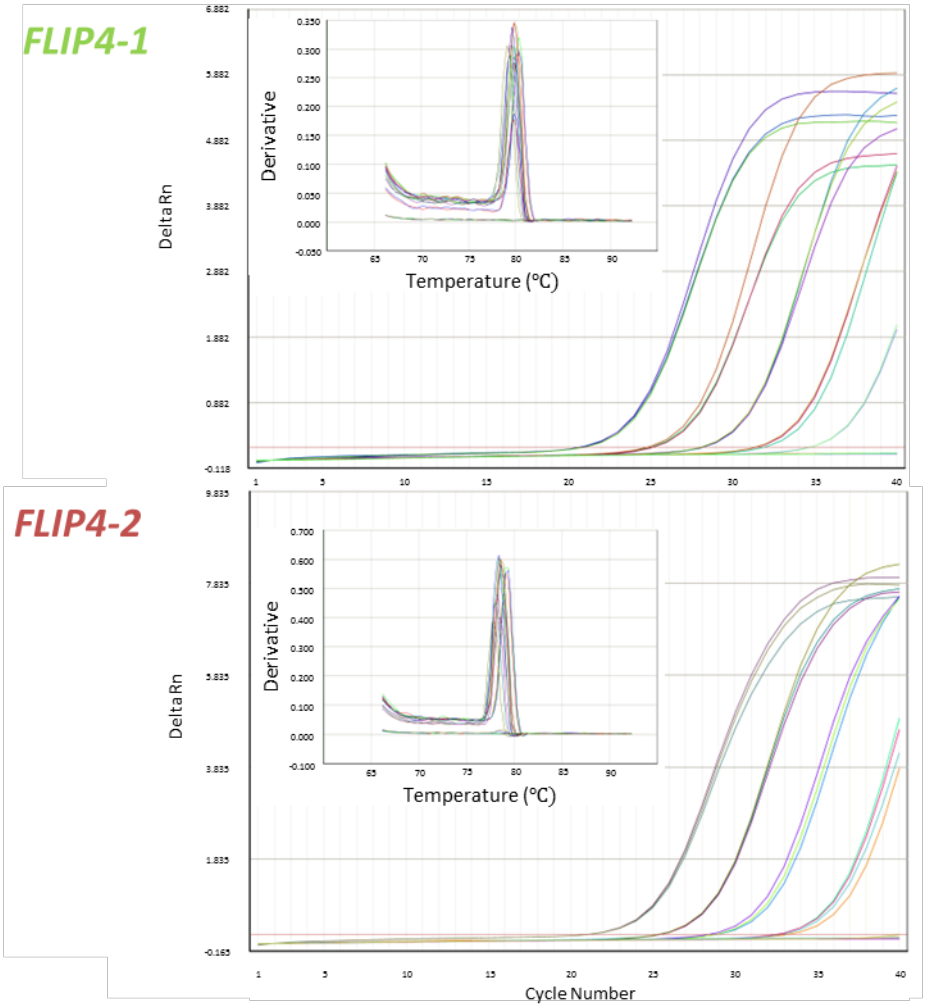

Figure 4. Reaction Kinetics and Specificity of *FLIP4* qPCR Primers. Standard curves of cDNA concentration relative to cDNA synthesis (RT) reaction were conducted in triplicate on pooled cDNA samples from varying tissue types, genotypes, and time points for primer sets for *FLIP4-1* (blue) and *FLIP4-2* (red). Amplification plots ($\Delta_{\text{Reaction fluorescence}}$ versus Cycle Number) for the standard curves. Note the tight triplicates of each reaction and the even spacing between each group, left-to-right corresponding to the highest and lowest pooled cDNA concentrations. Inlays: melting/dissociation curves (inverse of $\Delta_{\text{fluorescence}}$ versus temperature in°C) conducted for all reactions. Note the single peaks for each assay which are consistently around the same respective melting temperatures, although the intensity of the *FLIP4-1* assay is lower than that of *FLIP4-2*.

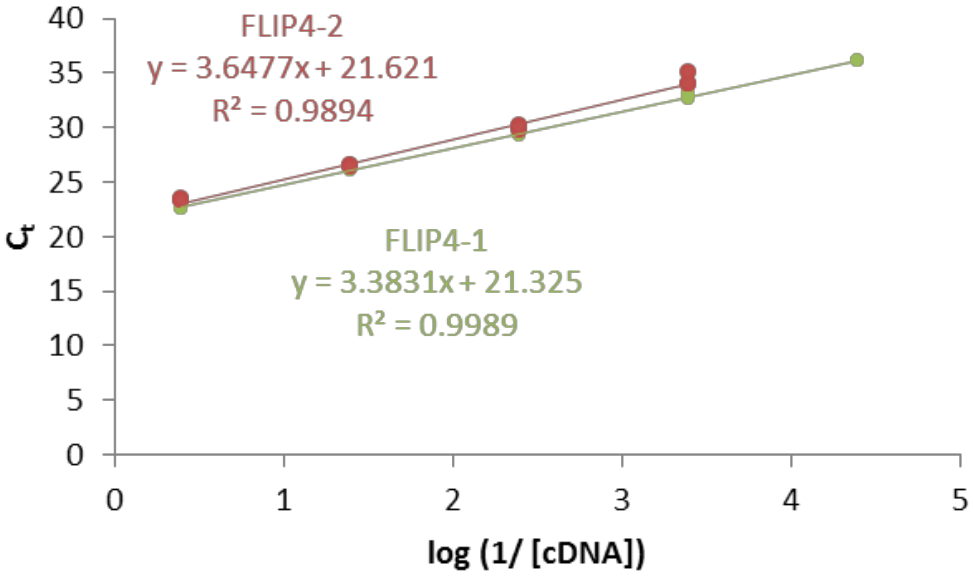

Figure 5. Standard Curves of *FLIP4* qPCR Primers. The data from the standard curves in Figure 4 were analyzed by plotting the log of the inverse of the initial template concentration versus the average Threshold Cycle (C_t) for each concentration. Note the close similarity in intercept and slope for the kinetics of the two assays, highlighting their comparability. Plots and linear regression were performed in Microsoft Excel 2010 to obtain the slope so that efficiencies could be calculated by the Pfaffl method. Images were processed (for clarity) in Adobe Photoshop Elements 8.

The primers used to amplify the FLIP4 transcripts in qPCR required a great deal of design and optimization. Analysis of reaction kinetics by serial decimal standard curves found that all assays were log-linear with regard to initial template amount, a trend that was reliable across multiple orders of magnitude of cDNA concentration with the exception of only the lowest concentrations for *FLIP4-1* reactions ($R^2 \geq 0.98$, Figures 4 and 5). The reactions were found to amplify specific amplicons based on melting-curve analysis of the PCR product in every reaction (inlays, Figure 4), as well as by gel electrophoresis (not shown).

In order to screen for errors in cDNA synthesis, including dilution errors, reaction inefficiency, or other unknown errors, the technical triplicates for each normalization gene reaction (using *PP2A* primers) were averaged and normalized to their respective averages per

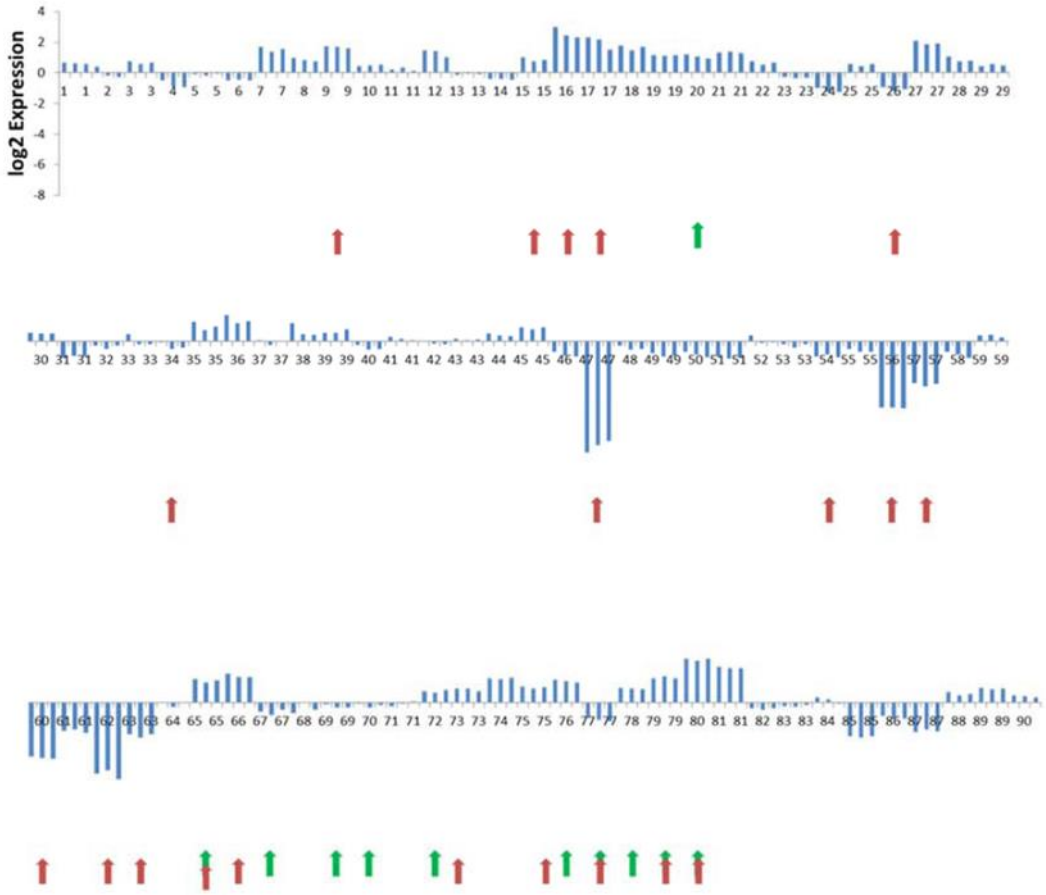

Figure 6. *PP2A* Transcript Abundance across All Samples Normalized to Profile Average by Genotype. *PP2A* transcript levels in all cDNA samples were normalized to the average for the *PP2A* developmental series using the $2^{\Delta\Delta Ct}$ method (since the same primer was used) and underwent a $\log_2$ transformation. Triplicate reactions for each sample are shown labeled by Sample #, and scale is conserved across rows. Samples # 47, 56, 57, 60, 62, 63, and 85 were suspect, as they exhibited > 2-fold "decrease", which is a generally accepted minimum difference to qualify as down-regulation in comparison to reference genes (Xiong 2006). Samples which either showed RNA degradation when natively electrophoresed or had extremely low transcript abundance were not included in downstream profile analyses. Red arrows indicate degraded RNA/low-abundance *PP2A* transcript (20/90), while green arrows indicate suspected gDNA contamination (11/90) and overlapping arrows indicate both issues (4/90).

genotype using the $2^{\Delta\Delta Ct}$ method. If the same amount of RNA was used in all cDNA synthesis reactions, and RNA was converted to cDNA equally efficiently, then comparable threshold cycles would be expected across similar samples. In fact, this was not always observed. At least seven cDNA synthesis reactions showed abnormally low levels of transcript detection (Figure 6). These samples were also excluded from data analysis for

expression profiles for all primers.

FLIP4 *Transcriptional Profiles in WT*, flip4-1, *and* flip4-2

In WT leaves, *FLIP4-1* transcripts were generally more abundant than those of its paralog, *FLIP4-2* (Figure 7). Day 15 and Day 40 were exceptions, although it should be noted that Day 40 measurements showed a slightly higher (but not higher than *FLIP4-2*) abundance in the WT for *FLIP4-1* transcripts before two of three data biological replicates were deemed to be unreliable. At Day 30, transcript levels rose above an expression ratio of 2 for *FLIP4-1* then appeared to decrease thereafter in leaves. Interestingly, a small increase in *FLIP4-1* transcript was found in flower samples, although these data points were based on only one biological replicate (Figure 7).

In the mutant phenotypes, *FLIP4-1* transcript abundance was variable (Figure 8a). In *flip4-1* leaves, transcript level was quite high ($R \approx 3.5$) at Day 10, very low at Day 15 ($R < 0.5$), and variable but present between Day 20 and Day 35 ($1.5 < R < 3.0$). *FLIP4-1* mRNA levels returned to $R < 0.5$x at Day 40. In *flip4-1* buds, *FLIP4-1* transcripts reached a maximum for all genotypes ($R \approx 4.4$) unique to the *flip4-1* genotype, and were present in siliques ($R \approx 1.6$). All flower samples for *flip4-1* were lost in RNA processing. In *flip4-2* mutant leaves, on the other hand, transcript levels were somewhat periodic, followed by a steady increase from $R \approx 0.5$ at Day 20 to $R \approx 3$ at Day 30. Thereafter, transcript abundance remained low ($0.7 < R < 1.5$). In buds and siliques, *FLIP4-1* transcript levels were relatively low ($R = 1.1, 1.2$, respectively) but were slightly elevated ($R \approx 2.5$) in flowers.

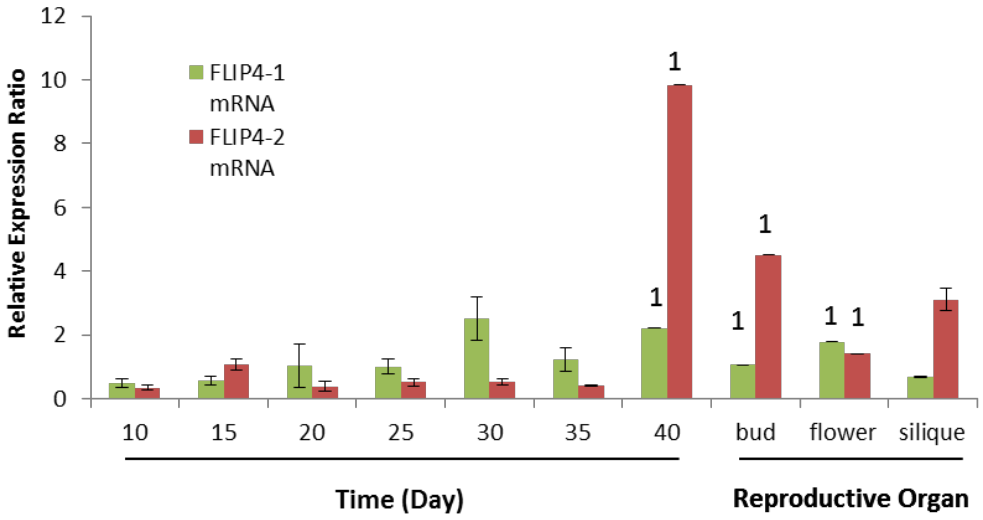

Figure 7. Expression Profiles of *FLIP4-1* and *FLIP4-2* transcripts in WT. Total RNA was isolated from leaf tissue from days 10-40, and from mature buds, flowers, and siliques. RNA was reverse transcribed and *FLIP4-1* and *FLIP4-2* transcripts were quantified in qPCR relative to *PP2A* transcript. Both genes were lowly expressed at comparable levels in leaves until Day 30 of growth, when *FLIP4-1* was upregulated roughly two-fold, then subsided. *FLIP4-2* was slightly more abundant in leaves at Day 15, and increased roughly 10-fold between Days 35 and 40. Both transcripts were expressed in all of the reproductive tissues examined, although *FLIP4-2* transcripts were much more abundant in buds and green siliques while transcript levels did not appear to differ in flowers. Mean expression ratio (R; difference in transcript abundance from the profile mean relative to that of *PP2A* for a given sample) ± SEM of one (1; no true SE), two, or three biological replicates. $R < 1$ indicates lower expression than average, whereas $R > 1$ indicates greater expression.

FLIP4-2 *Transcriptional Profiles in WT*, flip4-1, *and* flip4-2

The expression profile of *FLIP4-2* was found to be relatively consistent across the WT, *flip4-1* and *flip4-2* mutant genotypes (Figure 8b). Transcript levels were relatively stable and low ($R < 2$) in all genotypes until Day 35, with the exception of a slight increase at Day 15 in WT leaves. Transcript levels rose greatly in all genotypes after Day 35. This increase was more pronounced at Day 40 in WT and *flip4-2* mutant leaves ($R \approx 9.5$, WT; $R \approx 6.5$, *flip4-2*) than in *flip4-1* mutants ($R \approx 3$). It is interesting that *FLIP4-2* expression appears to be higher in the *flip4* mutants than in the WT from Day 20 to Day 35. Moreover, a reversal of

this relationship was observed at Day 40, when *FLIP4-2* transcript levels increased drastically (nearly 10-fold from Day 35 to Day 40).

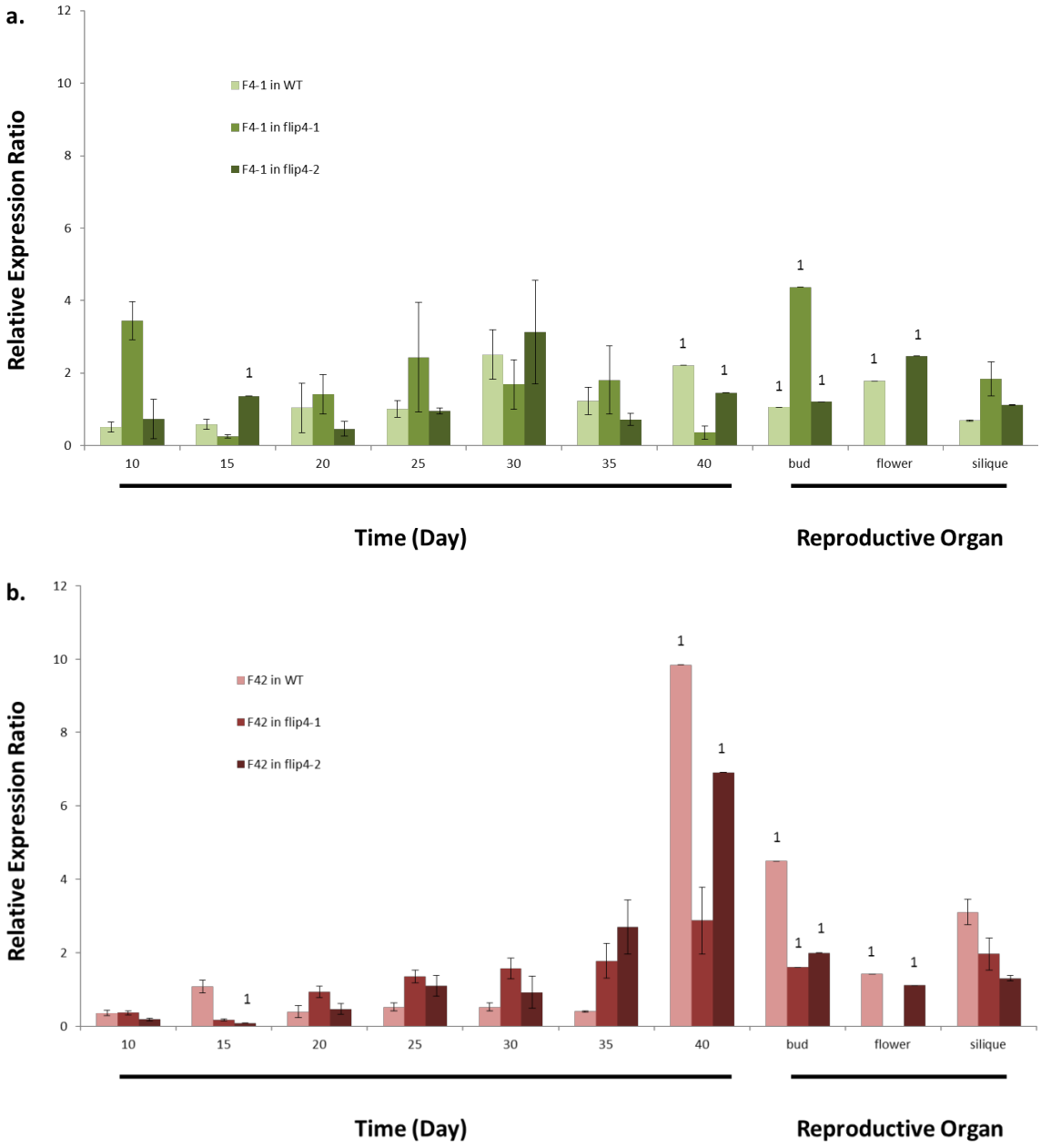

Figure 8. Expression Profiles of *FLIP4-1* (a) **and *FLIP4-2*** (b) **in WT, *flip4-1* and *flip4-2* leaves and reproductive organs.** Mean ± SE of one (1; no true SE), two, or three biological replicates; Expression Ratio (R), difference in transcript abundance from the profile mean relative to that of *PP2A*. Note that mRNA is detected for both genes, even in T-DNA insertion mutants, and generally higher expression of *FLIP4-2* in mutant leaves than in WT leaves. R = 1 is the profile average; R < 1 indicates lower expression than average, whereas R > 1 indicates greater expression.

Pearson and Spearman Correlations between Profiles in Leaves

For the purpose of comparing expression profiles as done here, the Pearson statistic is far more descriptive than the Spearman correlation, which was simply included as an added measure to catch any nonlinear but monotonic relationships between different profiles. Cases in which a significant Spearman correlation (ρ) existed without a corresponding Pearson correlation (R) are not necessarily useful for this study.

Table 2. Correlation Matrices for *FLIP4-1* and *FLIP4-2* Expression Profiles across Genotypes in Leaves.

	mRNA	*FLIP4-1*			*FLIP4-2*	
	Geno	WT	*flip4-1*	*flip4-2*	WT	*flip4-1*
Pearson Correlation*	*FLIP4-1* *flip4-1*	-0.333 (0.465)				
	FLIP4-1 *flip4-2*	0.748 (0.053)	-0.235 (0.611)			
	FLIP4-2 WT	0.506 (0.246)	-0.545 (0.206)	0.113 (0.809)		
	FLIP4-2 *flip4-1*	0.788 (0.035)	-0.313 (0.494)	0.245 (0.597)	0.735 (0.060)	
	FLIP4-2 *flip4-2*	0.584 (0.168)	-0.442 (0.321)	0.063 (0.893)	0.920 (0.003)	0.906 (0.005)
Spearman Correlation*	*FLIP4-1* *flip4-1*	-0.286 (0.535)				
	FLIP4-1 *flip4-2*	0.429 (0.337)	-0.321 (0.482)			
	FLIP4-2 WT	0.464 (0.294)	-0.714 (0.071)	0.786 (0.036)		
	FLIP4-2 *flip4-1*	0.821 (0.023)	0.000 (1.000)	0.250 (0.589)	0.357 (0.432)	
	FLIP4-2 *flip4-2*	0.679 (0.094)	0.071 (0.879)	0.143 (0.760)	0.321 (0.482)	0.964 (0.000)

*** Correlation, (p-value). Yellow, p < 0.05; Pink, p < 0.1.**

The profile shape of *FLIP4-1* in WT leaves correlated with that of *FLIP4-1* in the *flip4-2* mutant leaves, and the correlation was significant at the 0.1 level (R = 0.748, p = 0.053; Table 2). However, Spearman correlation between the profiles was weak and insignificant (ρ = 0.429, p = 0.337). This simply means that the data were not related

monotonically. No other significant correlations could be drawn between *FLIP4-1* profiles across genotypes, although *FLIP4-1* profiles did correlate with *FLIP4-2* profiles in different genotypes. *FLIP4-1* in WT leaves correlated significantly with *FLIP4-2* in the *flip4-1* mutant by Pearson and Spearman tests ($R = 0.788$, $p = 0.035$; $\rho = 0.821$, $p = 0.023$). A weak Spearman correlation between *FLIP4-1* in WT and *FLIP4-2* in *flip4-2* genotype was also detected and was significant at the 0.1 level ($\rho = 0.679$, $p = 0.094$). Likewise, a negative correlation was detected between *FLIP4-1* in the *flip4-2* mutant ($\rho = -0.714$, $p = 0.071$). Interestingly, the Spearman test described a perfectly random relationship between *FLIP4-1* in the *flip4-1* mutant and *FLIP4-2* in the *flip4-1* mutant, but with absolutely no significance ($\rho = 0.000$, $p = 1.000$). Finally, the *FLIP4-1* profile in *flip4-2* was correlated with *FLIP4-2* in WT by the Spearman test ($\rho = 0.786$, $p = 0.036$).

FLIP4-2 in WT was weakly but significantly correlated with *FLIP4-2* in the *flip4-1* mutant at the 0.1 level ($R = 0.735$, $p = 0.060$), although this was not the case for the Spearman correlation (Table 2). *FLIP4-2* in WT was also correlated significantly with *FLIP4-2* in the *flip4-2* mutant ($R = 0.920$, $p = 0.003$). Lastly, *FLIP4-2* profiles in the *flip4-1* and *flip4-2* mutants were significantly correlated ($R = 0.906$, $p = 0.005$; $\rho = 0.964$, $p < 0.001$), as would be expected from the similarity of the independent mutant profiles with the WT.

Analysis of Integrated FLIP4-2 *Interaction Networks*

Network analysis of a 50-gene physical interaction (44.66% of network weight), co-expression (32.77%), predicted interaction (13.94%), genetic interaction (5.56%), shared protein motif/domain (2.75%), and co-localization (0.32%) network surrounding *FLIP4-2*

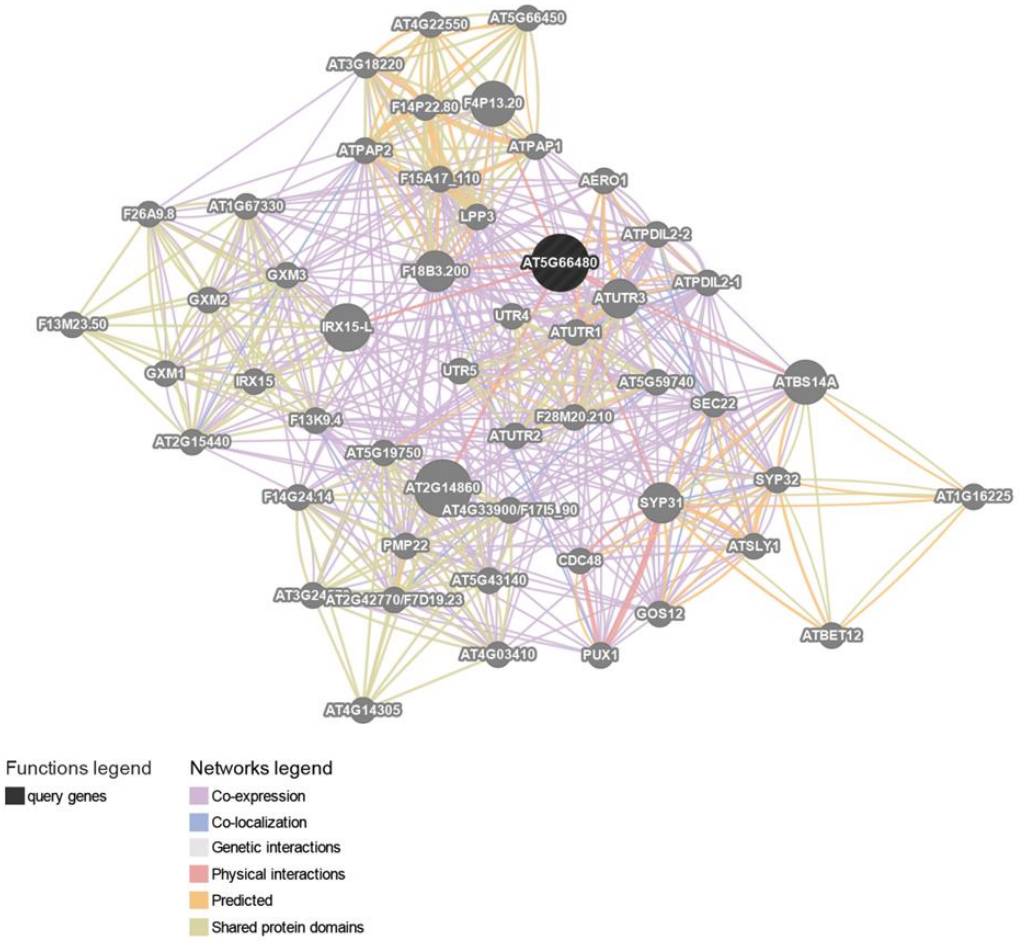

Figure 9. Network of 50 most closely associated genes surrounding *FLIP4-2* and its protein. GeneMANIA Network of 50 genes surrounding *FLIP4-2* (query gene), implemented based on physical interaction (44.66% of network weight), co-expression (32.77%), predicted interaction (13.94%), genetic interaction (5.56%), shared protein motif/domain (2.75%), and co-localization (0.32%). Note that FLIP4-2 is connected only by physical interactions to the densely co-expressed cluster.

and its protein (Figure 9) included mostly proteins enriched for peroxisome and lipid membrane, or cell membrane-related functions (Zuberi et al. 2013; Figures 10). Roughly half of the nodes comprising the imported GeneMANIA network (~1500) were unclustered. Five clusters were identified for the whole network and significantly overrepresented (Benjamini and Hochberg corrected p-value < 0.05; Breeze et al. 2011; Maere et al. 2005; data not shown). The first two clusters were relatively small (6 and 5 nodes, respectively) and were

nearly identical in enrichment, showing significant overrepresentation for response to arsenic. The third cluster was quite large (222 nodes) and significantly overrepresented for a number of functions, including histone modification, negative regulation of flower development, GDP-mannose transport, intracellular protein transport, intra-Golgi vesicle-mediated transport, and cellular membrane fusion. The relatively low p-values found in cluster 3 are likely due to the large number of genes. The fourth cluster was made up of a more reasonable 21 nodes, and pointed more weakly to connections to nickel ion response, thylakoid membrane organization, endoplasmic reticulum unfolded protein response, regulation of Rab-GTPase activity, and arsenic, nucleotide sugar, and protein transport. Finally, the fifth cluster was overrepresented for arsenic transport, cellular membrane fusion, response to inorganic substance, and brassinosteroid mediated signaling pathways.

When the original 50-gene network surrounding *FLIP4-2* was pulled out of the larger network and clustered, only two clusters were found, comprised of 7 and 4 nodes (data not shown). The first was enriched significantly for intracellular protein transport, protein secretion, Golgi vesicle transport processes, vesicle docking involved in exocytosis, and very significantly overrepresented for cellular membrane fusion, while the second pointed towards pollen germination, organ growth, protein complex disassembly, and cellular membrane fusion. The original GeneMANIA network (Figure 9) showed a densely co-expressed network of genes with functions connected to the peroxisomal and microbody membranes, endoplasmic reticulum stress response, and cell wall metabolism (Figure 10). No functions were enriched above a FDR of 0.01 in the network surrounding *FLIP4-1* and its protein, and functional enrichment analysis in Cytoscape was not performed.

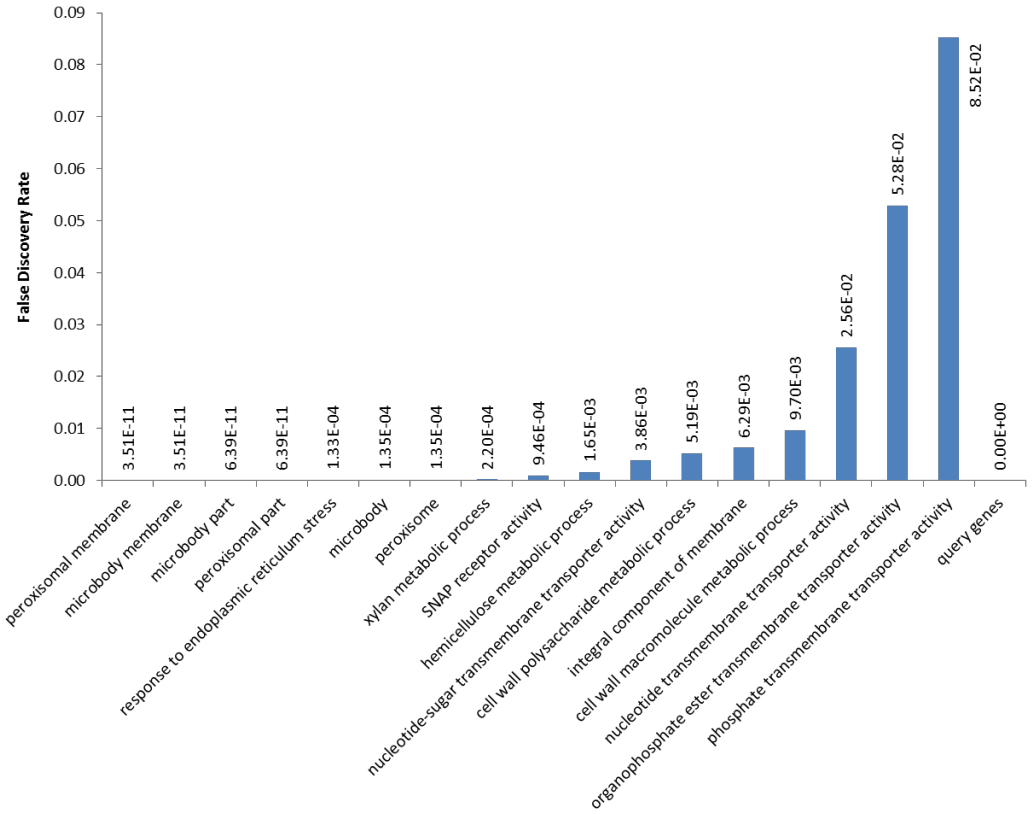

Figure 10. Functional Annotations of 50 Genes Associated with *FLIP4-2* and its protein. Note the low false discovery rates of a variety of peroxisomal, membrane, transporter, and cell wall metabolic processes suggested by this analysis. Lower FDRs are associated with greater confidence that a group of genes belongs under a given functional category.

DISCUSSION

Development of a System for Measuring FLIP4 *Expression*

RT-qPCR has been referred to as the "gold standard" for measuring the effects that treatments, or time and development, have on the expression of specific genes (Derveaux et al. 2010). The method is sensitive to small changes in transcript abundance over a large range of copy numbers, which makes it an ideal tool for assessing the potentially subtle yet biologically impactful shifts in expression which might be expected in subfunctionalizing paralogs. However, the method can be challenging due to the many factors involved and the strong influence each can have on the results obtained (Bustin et al. 2009). These factors include RNA quality, efficiency and consistency in cDNA synthesis and qPCR, selectivity of the reaction for the intended transcript, and the appropriateness of the normalization gene(s) and data analysis methods (Nolan et al. 2006; Bustin et al. 2009).

A further challenge arises in the case of paralogs, given their sequence similarity, especially in the coding regions. Although it is not yet clear if sequence divergence in introns occurs more freely in general (intron and exon divergence may in fact be linked in rice; Zhu et al. 2009), the FLIP4 exons align well while their introns do not (simple alignment, data not shown). At the same time, secondary structures were common in the mRNA sequences for the FLIP4 genes, making the selection of an accessible priming site very difficult. As such, multiple attempts at primer design and reaction optimization were required to develop a dynamic, sensitive assay which could distinguish between the FLIP4 paralogs. The assays developed for this project are reliable and should be useful in further studies for the elucidation of FLIP4 protein functions.

Transcripts were Constitutively Detected, Even in Null Mutants

Interestingly, amplicons were amplified even in T-DNA insertion mutants. As T-DNA insertions are generally considered functional knockouts (Krysan et al. 1999), especially when the insertion is in an exon, this seems strange. While the loss of the gene function can result from simple changes to protein or functional RNA sequence, the transcript could still be produced in some form as long as promoter and enhancer sequences are not disrupted or repositioned relative to the gene. Moreover, the loss of a protein or RNA function does not imply changes in the rate of transcription. On the other hand, an insertion of T-DNA magnitude (~4484 bp; Alonso et al. 2003) into an exon might be expected to reduce the efficacy of the RNA polymerase during transcription, resulting in an incomplete transcript (e.g. lack of a poly-A tail and therefore expedited degradation or compromised processing), or the excision or silencing of a larger genetic region by genome defense mechanisms. In these cases, the transcript should be detected in greatly reduced numbers, if at all in the mutant in comparison to the WT. Alternatively, the RNA could be transcribed and processed, including the insert in whole or part. In the case that the insert falls within the amplicon intended to be produced in qPCR, there would be a much larger amplification product. This could be detectable as a substantial increase in melting temperature (Ririe et al. 1997), but the more likely scenario is that the DNA polymerase used in the qPCR reaction cannot synthesize the entire amplicon, and that a wide range of products or none at all are produced.

All of these are possible but unlikely in the case of the *flip4* mutants used in this study since reaction success and amplicon melting temperatures and sizes were consistent across genotypes for both assays. It is certain that the T-DNA insertion into the *flip4-1* gene is

within the first exon and likely causes a catastrophic disruption of the coding sequence. In the *flip4-2* mutant, it is not certain whether the insertion lies within the amplicon in the second exon, or in the 3' UTR. For the same reasons that the above scenarios are unlikely, the insert seems more likely to lie in the 3' UTR. While this could still result in loss of *FLIP4-2* function due to disruption of key motifs related to the coordination of RNA processing and stability, the region is not translated and protein function may not be affected. As such, there is a reasonable explanation for the detection of each transcript in every genotype, but the mutation in *flip4-2* may not result in a true loss of function.

A failed knockout mutation in *flip4-2* presents an alternative to the original hypothesis that *FLIP4-1* is able to compensate for nonfunctional FLIP4-2, which is based on the observation of an altered phenotype in *flip4-1* while there is not a clear one in *flip4-2*. If this is the case, the failure to identify double mutants would be more difficult to explain. An alternative explanation, then, is that the insert has been lost between generations in the strain used since initial testing in our lab. This could be confirmed by a simple PCR using insert and gene-specific primers followed by gel purification and sequencing. If the insert is confirmed, loss of FLIP4-2 function simply may not have a similar or observable phenotypic outcome to *flip4-1*, or the necessary stresses on the plant may not have been tested yet.

Insights into the FLIP4 *Evolutionary Process*

There seems to be little to no expression compensation by *FLIP4-1*, although elevated *FLIP4-2* transcript levels in Day 15-35 leaves in both mutants could be a sign of some sort of unidirectional compensation, i.e. *FLIP4-2* responding to changes in *FLIP4-1* function. The expression levels of FLIP4-1 in the Wild-type and *flip4-2* mutant are similar ($R = .748$, $p =$

.053) but different from the *flip4-1* mutant profile, suggesting that the *FLIP4-1* knockout likely affects *FLIP4-1* transcript level, indirectly or directly. This sort of disruption would imply a feedback loop on the level of protein function or metabolism. The hypothetical mechanism is apparently not extended from *FLIP4-2* or its function, as the *FLIP4-2* knockout seems to have little effect on *FLIP4-1* expression. On the other hand, *FLIP4-2* expression appears to be affected similarly and nontrivially by loss of either protein function. This spells a scenario of different regulation schemes and suggests that the FLIP4-1 promoter sequence may contain different or altered regulatory elements. If the paralogs are in the process of subfunctionalizing, some level of compensation with promoter evolution might be expected (Hanada et al. 2009).

Following up on the Cole's promoter analysis (2014) a search using the AGRIS AtcisDB (Yilmaz et al. 2011) transcription factor (TF) binding site prediction tool shows that indeed multiple common and differing stress and senescence-associated TFs are predicted to bind *FLIP4-1* and *FLIP4-2* promoters. Fifteen putative binding sites were found for the *FLIP4-1* promoter, including Light-responsive element (LRE) and motifs suggestive of potential BHLH, bZIP, WRKY, and ARF control. The *FLIP4-2* promoter also had predicted LRE motifs, as well as potential bZIP, MYB, LFY, VOZ proteins, and ABI binding sites. Both of the predicted promoters thus appear to imply roles in stress responses or senescence. It is interesting, however, that only two TF families mentioned in Breeze et al. 2011 (versus the three for *FLIP4-1*) are predicted to bind *FLIP4-2* regulatory regions when *FLIP4-2* expression appears to be more consistent with a role in senescence according to the data presented in this report. Taken alongside the observation that *flip4-1* knockout mutant exhibits a sickly phenotype under constant light, *FLIP4-1* appears to be implemented in a

light-induced stress response or mediation of light's role in the senescence process. For instance, darkness in individual leaves induces leaf-specific senescence, whereas darkening of the entire plant rather simply appears to slow development (Weaver and Amasino 2001). The *flip4-1* phenotype appears to emerge in constant light or dark conditions (personal observation).

If both paralogs served the same functions, their proteins would be expected to share many, if not all, interaction partners. In fact, this does not appear to be the case, as the two share only two predicted protein interactions directly, with SYP32 (syntaxin of plants 32) and RanGAP1 (Ran GTPase activating protein) (thebiogrid.org; Chatr-Aryamontri et al. 2015), although they have nearly forty physical interactions combined and a number of secondary or tertiary connections exist, especially by co-expression (Warde-Farley et al. 2010; Zuberi et al. 2013). Remarkably, FLIP4-1 has three times the number of predicted interaction partners for FLIP4-2, but GeneMANIA did not detect significantly enriched functions for its interaction neighborhood, even when the network was expanded to include over 100 genes and proteins. On the other hand, *FLIP4-2* interaction partners appear to associate with the peroxisomal membrane with a very low false discovery rate (FDR). Although it seems odd that FLIP4-1 and FLIP4-2 do not appear to share many interaction partners, recent paralogs were found to share only 41% of their interactions on average in the *Arabidopsis* interactome (*Arabidopsis* Interactome Mapping Consortium 2011). Furthermore, divergence from this average is associated with functional divergence, providing evolutionary insights for *FLIP4-1* and *FLIP4-2* (*Arabidopsis* Interactome Mapping Consortium 2011). For instance, the genes may have taken on roles more closely related to different interaction partners.

It should also be noted that the sudden drop in transcript level at Day 15 in the *flip4-1*

mutant is unexpected and difficult to explain, potentially raising questions about the quality of the data or the validity of the biological samples. Moreover, there is not much expression data published for *FLIP4-2*. For this reason, more expression correlation interactions may apparently exist in databases with *FLIP4-1*, highlighting the putative nature of conclusions derived from computational analyses.

Potential Roles of the FLIP4 Proteins in Leaf Senescence or Stress Responses

The fact that *FLIP4-1* is generally more highly expressed early in leaves and was co-expressed or potentially interacts with organ growth genes may suggest a developmental role in leaves. Oddly, heightened expression was not observed consistently in flowers and buds, in contrast to the expectations based on pollen-specific high expression levels reported and the association of pollen with these tissues (Toufighi et al. 2005). This was the case even before potentially erroneous data were excluded (the exclusion of similar data for *FLIP4-1* in WT leaves resulted in a decrease in observed expression at Day 40). Early in leaves, the log-transformed expression levels of both genes were in fact relatively flat and in reasonable agreement with Breeze et al. data (2011). Higher levels of *FLIP4-1* and *FLIP4-2* were observed at in this study at Days 30 and 40, respectively. *FLIP4-1* transcript levels may correlate with generally higher photosynthetic rates that are typical of that developmental stage and potentially suggests a link to metabolism (Breeze et al. 2011).

On the other hand, the expression of *FLIP4-2* appeared to increase greatly at Day 40, which is well into flower production stage, and in siliques (Boyes et al. 2011). As a membrane-associated protein which is predicted to be localized to the chloroplast (previous lab data), FLIP4-2 could play an organizational role or may be involved with transport of

metabolites or proteins, which were parts of our original hypothesis. These connections also apply, however, to the process of senescence in leaves if the genes in fact lead to the negative regulation of photosynthesis or development; the answer to this question is critical to the interpretation of these connections. Oddly, Breeze et al. found in 2011 that both *FLIP4* transcript levels changed little (< 1 fold) between Days 19 and 39 in senescing WT (Col.) leaves under similar conditions, although *FLIP4-1* expression increased through Day 39. The reason for this discrepancy is not known, although normalization, the sensitivity of microarrays versus qPCR, and the quality/selection of RNA resulting from different extraction methods could all have great effects on relative amounts of these transcripts.

Photosynthesis-related genes were found to be expressed early in senescing tissues, although genes involved in chlorophyll, carotenoid and amino acid biosynthesis pathways were downregulated by Day 23 (Breeze et al. 2011). This is consistent with a role for FLIP4-1 in photosynthesis. Upregulated genes included transcription factors implemented in drought, heat, and oxide stresses. Other upregulated groups were involved in autophagy and chlorophyll degradation pathways (Breeze et al. 2011).

Breeze et al. reported upregulation of the transcription factor MYC2 and its JAZ protein repressors early in senescence, with implied effects throughout the process. Interestingly, Cole (2014) detected putative ABRE-like (drought response, both), AtMYC2 (enhanced ABA sensitivity to drought, *FLIP4-1*), and SORLIP2 (light-regulated transcription) elements in the *FLIP4* promoters. Upregulation by MYC2 could explain the higher levels of *FLIP4-1* transcript during Days 25-35, while *FLIP4-2* transcript levels at Day 40 correspond with the increasing levels of ABA leading up to Day 39 reported by Breeze et al. in 2011.

Balazadeh et al. (2008) review the major points of leaf senescence. The tightly formulaic process depends on sugar, nutrient, hormone, and light availability, and consists of a redistribution of nutrients, the catabolism of chlorophyll and other pigments, proteins and complexes, lipids, and RNA. TF families implemented in the regulation of senescence include the NAC, WRKY, MYB, C2H2 zinc-finger, bZIP, and AP2/EREBP families, while members of the NAC family are also involved in ABA response, secondary cell wall synthesis, pathogen response, and salt tolerance. Aside from the classic abscisic acid and ethylene (Kim et al. 2014; Kim et al. 2015) induction of senescence, salicylic and jasmonic acid (via WRKY, ESR/ESP), and auxin (via ARFs and ARRs) contribute to control of the senescence program (Evans et al. 2010 and Breeze et al. 2011).

That FLIP4-2, a coiled-coil protein, seems to associate with membrane-bound proteins and those annotated for membrane fusion is not surprising, due to the organizational and structural roles that coiled-coils are known to play in the plant cell (Rose and Meier 2004). Connections to signaling pathways, to inorganic molecule/protein/metabolite transport, to the ABA, ethylene, and brassinosteroid signaling pathways, and to protein complex disassembly found in this network were not anticipated (but were mentioned in Evans et al. 2010). In 2006 van der Graaff et al. found that a number of transmembrane proteins and transporters were upregulated during leaf senescence, as many processes in programmed cell death require the export of cell contents, e.g. via vesiculation, which depends on transmembrane proteins such as DMP-1 (Kasaras et al. 2011). Transporter protein expression is also observed due to the mass export of nutrients from senescing tissues (Evans et al. 2010). Another cluster in the present analysis also included genes involved in regulation of chlorophyll biosynthesis, histone modification, and defense response to biotic

stress. These processes are all drawn on during senescence (Breeze et al. 2011; Khan et al. 2014). In all, a sudden increase in FLIP4-2 transcript between Days 35 and 40 and in siliques, the predicted interaction network and functional enrichments, and senescence-related response elements predicted in its promoter, the evidence presented here is consistent with a FLIP4-2 role in senescence.

It should be noted that no functions were enriched above a FDR of 0.01 in the network surrounding *FLIP4-1* and its protein, and functional enrichment analysis in Cytoscape was not performed, although this is a conservative threshold. The lack of enrichment may indicate that the functions of the genes related to *FLIP4-1* and its protein are more diverse than those surrounding *FLIP4-2*, or that it has less-specific role(s). The fact that FLIP4-1 is predicted to interact with many more proteins than FLIP4-2 gives weight to this explanation. It is also possible that high-quality data is simply not available for the bulk of putative FLIP4-1 interaction partners.

Concluding Remarks

Based on observations indicating that the *flip4-1* mutant phenotype is enhanced by drought, light and other stress conditions, and Cole's finding (2014) of light and ABA response element motifs in the promoter regions of the *FLIP4* genes and their associations presented here with senescence-related genes, a rational next experiment would be to measure the expression of *FLIP4-1* and *FLIP4-2* in WT, *flip4-1* and *flip4-2* under normal, high light, and drought conditions. Additionally, treatment with ABA and other stress or senescence-related signaling molecules could solidify connections and differences between the FLIP4 proteins and these processes. To verify that transcript level changes are in fact due

to transcriptional regulation, protein expression or RNA stability assays may be necessary. Discrepancies between data presented here and in Breeze et al. 2011 may be rectified by generating expression profiles for senescence-associated genes as well for a positive control. It would also be helpful to confirm T-DNA insertion sites by sequencing or restriction digest. The connection between high expression specific to pollen in *FLIP4-1* (Toufighi et al. 2005), *flip4-1* reduced seed set (Cole 2014) and stress response/senescence is still unclear. ER, Golgi, peroxisome, and cell membrane fusion and protein transport are necessary in cell growth and division, while the membranes of pollen tubes must be constantly remodeled as they expand. Similar network analyses of *FLIP4-1* expression should be carried out. Lastly, the knockout mutation in the *flip4-2* genotype needs to be confirmed by Northern blot, Western blot, or PCR before further testing is done.

These evidences point to a few new future avenues of research in determining the evolutionary dynamics and functions of the *FLIP4* gene family. Given the lack of overlap between the interaction partners of this protein family and the differences in expression, it appears that one has taken on a new and important role. Taken together, this work has presented lines of evidence to be followed up on pointing to the subfunctionalization of the FLIP4 proteins, and it allows our hypothesis to evolve to include FLIP4-2 involvement in stress response or senescence processes.

REFERENCES

Adams K.L. (2007). Evolution of duplicate gene expression in polyploid and hybrid plants. J. Hered. **98**: 136-141.

Adams, K.L., and Wendel J.F. (2005). Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. **8**: 135-41.

Akama, S., Shimizu-Inatsugi, R., Shimizu, K.K., and Sese, J. (2014). Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid *Arabidopsis*. Nucleic Acids Res. **42**(6).

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., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry , C.C., and Ecker, J.R. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science **301**: 653-657.

***Arabidopsis* Biological Resource Center.** The Ohio State University. Handling *Arabidopsis* plants and seeds: Methods used by the *Arabidopsis* Biological Resource Center. Retreived from < https://abrc.osu.edu/seed-handling> 27 July 2015.

***Arabidopsis* Interactome Mapping Consortium** (2011). Evidence for Network Evolution in an *Arabidopsis* Interactome Map. Science **333**: 601-607.

Balazadeh, S., Riaño-Pachón, D.M., and Mueller-Roeber, B. (2008). Plant Biol. (Stuttg.) **10**: S63-75.

Barker, M.S., Vogel, H., and Schranz, M. E. (2009). Paleopolyploidy in the *Brassicales*: analyses of the cleome transcriptome elucidate the history of genome duplications in *Arabidopsis* and other *Brassicales*. Genome Biol. Evol. **1**: 391-399.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B. **57**: 289-300.

Bennetzen, J.L. (2000). Transposable element contributions to plant gene and genome evolution. Plant Mol. Biol. **42**: 251-269.

Box, M.S., Coustham, V., Dean, C., and Mylne, J.S. (2011). A simple phenol-based method for 96-well extraction of high quality RNA from *Arabidopsis*. Plant Methods **7**:7.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., and Görlach, J. (2001). Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. Plant Cell **13**: 1499-510.

Brazma, A., and Vilo, J. (2000). Gene expression data analysis. FEBS Letters **480**: 17-24.

Breeze, E., Harrison, E., McHattie, S., Hughes, L., Hickman, R., Hill, C., Kiddle, S., Kim, Y., Penfold, C.A., Jenkins, D., Zhang, C., Morris, K., Jenner, C., Jackson, S., Thomas, B., Tabrett, A., Legaie, R., Moore, J.D., Wild, D.L., Ott, S., Rand, D., Beynon, J., Denby, K., Mead, A., and Buchanan-Wollaston, V. (2011). High-resolution temporal profiling of transcripts during *Arabidopsis* leaf senescence reveals a distinct chronology of processes and regulation. Plant Cell **23**: 873–894.

Bustin S.A., Benes V., Garson J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., and Wittwer, C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. **55**: 611-622.

Chatr-Aryamontri, A., Breitkreutz, B.J., Oughtred, R., Boucher, L., Heinicke, S., Chen, D., Stark, C., Breitkreutz, A., Kolas, N., O'Donnell, L., Reguly, T., Nixon, J., Ramage, L., Winter, A., Sellam, A., Chang, C., Hirschman, J., Theesfeld, C., Rust, J., Livstone, M.S., Dolinski, K., and Tyers, M. (2015). The BioGRID interaction database: 2015 update. Nucleic Acids Res. **43**: D470-8.

Cole, M. A. (2014). Characterization of the promoter regions of FLIP4-1 and FLIP4-2 in *Arabidopsis thaliana*. Undergraduate honors thesis, Appalachian State University, Department of Biology.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.-R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiol. **139**: 5-17.

D'haeseleer, P. (2005). How does gene expression clustering work? Nat. Biotechnol. **23**: 1499-501.

de Siquiera Santos, S., Takahashi, D.Y., Nakata, A., and Fujita, A. (2013). A comparative study of statistical methods used to identify dependencies between gene expression signals. Brief. Bioinform. **15**: 906-18.

Derveaux, S., Vandesompele, J., and Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. Methods **50**: 227-230.

Evans, I.M., Rus, A.M., Belanger, E.M., Kimoto, M., and Brusslan, J.A. (2010). Dismantling of *Arabidopsis thaliana* mesophyll cell chloroplasts during natural leaf senescence. Plant Biol. (Stuttg.) **12**: 1-12.

Feldman, M., and Levy, A.A. (2009). Genome evolution in allopolyploid wheat--a revolutionary reprogramming followed by gradual changes. J. Genet. Genomics. **36**: 511-8.

Flagel, L.E., and Wendel, J.F. (2009). Gene duplication and evolutionary novelty in plants. New Phytol. **183**: 557-564.

Hanada, K., Kuromori, T., Myouga, F., Toyoda, T., Li, W.H., and Shinozaki, K. (2009). Evolutionary persistence of functional compensation by duplicate genes in *Arabidopsis*. Genome Biol Evol. **1**: 409-14.

Huang, X., and Miller, W. (1991). A time-efficient, linear-space local similarity algorithm. Adv. Appl. Math. **12**: 337-357.

Innan, H., and Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. Nat. Rev. Genet. **11**: 97-108.

Jiao Y., Wickett N.J., Ayyampalayam S., Chanderbali A.S., Landherr L., Ralph P.E., Tomsho L.P., Hu Y., Liang H., Soltis P.S., Soltis D.E., Clifton S.W., Schlarbaum S.E., Schuster S.C., Ma H., Leebens-Mack J., and dePamphilis C.W. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature **473**: 97-100.

Kasaras, A., Melzer, M., and Kunze, R. (2012). *Arabidopsis* senescence-associated protein DMP1 is involved in membrane remodeling of the ER and tonoplast. BMC Plant Biol. **12**:54.

Khan, M., Rozhon, W., and Poppenberger, B. (2014). The role of hormones in the aging of plants - a mini-review. Gerontology **60**: 49-55.

Kim, H.J., Hong, S.H., Kim, Y.W., Lee, I.H., Jun, J.H., Phee, B.K., Rupak, T., Jeong, H., Lee, Y., Hong, B.S., Nam, H.G., Woo, H.R., and Lim, P.O. (2014). Gene regulatory cascade of senescence-associated NAC transcription factors activated by ETHYLENE-INSENSITIVE2-mediated leaf senescence signalling in *Arabidopsis*. J. Exp. Bot. **65**: 4023-36.

Kim, J., Chang, C., and Tucker, M.L. (2015). To grow old: regulatory role of ethylene and jasmonic acid in senescence. Front. Plant Sci. **29**: 20.

Kondrashov, A.F., Rogozin, I.B., Wolf, Y.I., and Koonin, E.V. (2002). Selection in the evolution of gene duplications. Genome Biology **3**(2).

Krysan, P.J., Young, J.C., and Sussman, M.R. (1999). T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell **11**: 2283-90.

Leitch, A.R., and Leitch, I.J. (2008). Genomic plasticity and the diversity of polyploid

plants. Science **320**: 481-3.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^-$\Delta\Delta C_t$ Method. Methods **25**: 402-8.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science **290**: 1151-1155.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. **15**:21: 3448-9.

Monson, R.K. (2003). Gene duplication, neofunctionalization, and the evolution of C_4 photosynthesis. Int. J. Plant Sci. **164**: S43-S54.

Nolan, T., Hands, R.E., and Bustin, S.A. (2006). Quantification of mRNA using real-time RT-PCR. Nat. Protoc. **1**: 1559-82.

Ohno, S. (1970). Evolution by gene duplication. Springer-Verlag, New York, NY, USA.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29**: e45.

Reel, K.B. (2013). Redundancy and specialization in the *Arabidopsis thaliana* FLIP4 gene family. Undergraduate honors thesis, Appalachian State University, Department of Biology.

Richardson, C.A. (2012). Localization of FLIP4-2 via immunofluorescence microscopy in *Arabidopsis thaliana* leaf cross sections. Undergraduate honors thesis, Appalachian State University, Department of Biology.

Ririe, K.M., Rasmussen, R.P., and Wittwer, C.T. (1997). Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal. Biochem. **245**: 154-60.

Rodin, S.N., and Riggs, A.D. (2003). Epigenetic silencing may aid evolution by gene duplication. J. Mol. Evol. **56**: 718-29.

Rose, A., and Meier, I. (2004). Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. Cell Mol. Life Sci. **61**: 1996-2009.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **13**: 2498-504.

Spearman, C. (1904). The proof and measurement of association between two things. Amer. J. Psychol. **15**: 72-101.

The *Arabidopsis* Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**: 796-815.

Toufighi, K., Brady, S.M., Austin, R., Ly, E., and Provart, N.J. (2005). The Botany Array Resource: e-Northerns, expression angling, and promoter analyses. Plant J. **43**: 153-163.

van der Graaff, E., Schwacke, R., Schneider, A., Desimone, R., Flügge, U.-I., and Kunze, R. (2006). Transcription analysis of *Arabidopsis* membrane transporters and hormone pathways during developmental and induced leaf senescence. Plant Physiol. **141**: 776–792.

Wang, J., Marowsky, N.C., and Fan, C. (2014). Divergence of gene body DNA methylation and evolution of plant duplicate genes. PLoS One **9**(10).

Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G.D., and Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. **38**: W214-20.

Weaver, L.M., and Amasino, R.M. (2001). Senescence is induced in individually darkened *Arabidopsis* leaves, but inhibited in whole darkened plants. Plant Physiol. **127**: 876–886.

Xiong, J. (2006). Essential Bioinformatics. Cambridge University Press. New York, NY USA.

Yilmaz, A., Mejia-Guerra, M.K., Kurz, K., Liang, X., Welch, L., and Grotewold, E. (2011). AGRIS: Arabidopsis Gene Regulatory Information Server, an update. Nucleic Acids Res. **39**: D1118-D1122.

Zhu, L., Zhang, Y., Zhang, W., Yang, S., Chen, J.Q., and Tian, D. (2009). Patterns of exon-intron architecture variation of genes in eukaryotic genomes. BMC Genomics **10**:47.

Zuberi, K., Franz, M., Rodriguez, H., Montojo, J., Lopes, C.T., Bader, G.D., and Morris, Q. (2013). GeneMANIA prediction server 2013 update. Nucleic Acids Res. **41**: W115-22.