2013

A Study of the Genes *TFL1* and *LFY* as Agents in the Divergent Evolution of the Plant Species *Arabidopsis thaliana* and *Carica papaya*

Ryan Lockard
University of Northern Iowa

A STUDY OF THE GENES *TFL1* AND *LFY* AS AGENTS IN THE DIVERGENT EVOLUTION OF THE PLANT SPECIES *ARABIDOPSIS THALIANA* AND *CARICA PAPAYA*

A Thesis Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors

Ryan Lockard

University of Northern Iowa

May 2013

This Study by: Ryan Lockard

Entitled: A Study of the Genes *TFL1* and *LFY* as Agents in the Divergent Evolution of the Plant Species *Arabidopsis thaliana* and *Carica papaya*

has been approved as meeting the thesis or project requirement for the Designation

University Honors with Distinction or University Honors (select appropriate designation)

May 9, 2013
Date — Dr. Marek Sliwinski, Honors Thesis Advisor

5/10/13
Date — Dr. Jessica Moon, Director, University Honors Program

Table of Contents

List of Tables

List of Figures

Introduction

Over the past several decades, the plant *Arabidopsis thaliana* has been selected by many researchers in the fields of genetics and molecular biology, resulting in the establishment of Arabidopsis as a model organism. There are multiple features of the plant that account for its great popularity. The plant's small size and simple growth requirements allow it to be easily maintained in a laboratory setting, and individual plants produce thousands of seeds with which to work. The full Arabidopsis genome sequence was published in the year 2000, since providing researchers with access to information on the plant's entire genetic makeup. Of prime importance, relative to some other species of plants it is possible to genetically transform Arabidopsis with little difficulty [1].

Two well-studied genes involved in the development of Arabidopsis are *TERMINAL FLOWER 1* (*TFL1*) and LEAFY (*LFY*). *TFL1* is known to control both flower placement and for maintaining indeterminacy in the shoot apical meristem of the wild type plants, producing a raceme inflorescence (Figure 1), while *LFY* has been shown to be responsible for the initiation of flower development in wild type *Arabidopsis thaliana*. Homologs--genes related to one another by descent from common ancestral DNA--of these sequences have been identified in many other plants, including the related but developmentally distinct *Carica papaya*. Unlike Arabidopsis, and through a process unknown, *Carica papaya* plants have evolved to flower determinately and produce a cymose inflorescence with terminal flowers (Figure 1). The exact functions of identified *Carica papaya TFL1* homologs have yet to be determined, though it is predicted that they are involved in the formation of papaya's characteristic cyme. *TFL1* from papaya has not yet been well researched. *LFY* protein (*PFL* in papaya) has been cloned and shown to substitute for Arabidopsis *LFY* [2], but the regulatory DNA sequences for this homolog have not yet been

isolated and studied. Like Arabidopsis, papaya's genome has been sequenced and published [3]. The availability of papaya's full genome combined with its divergent evolution and development make the plant an intriguing system for use in comparative studies with the well-researched Arabidopsis.

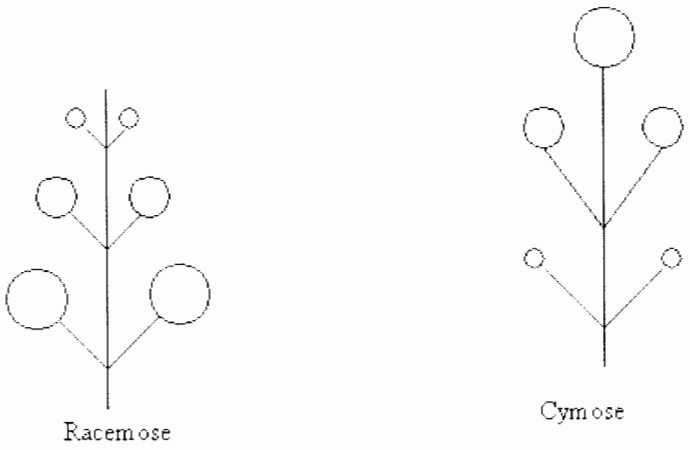

Figure 1. Comparison of racemose and cymose plant inflorescences

The long term goals of this study are to: (1) To determine if the genes *TFL1* and *LFY* perform any role in the divergent evolution and development of these two plant species; (2) If so, begin to assess whether the homologs' functional divergence is a result of gene regulation, protein function, both, or some different genetic mechanism altogether. These goals will be accomplished by genetically transforming *Arabidopsis thaliana* plants containing *TFL1* and *LFY* knockout mutations (those genes removed) with the homologous genes from *Carica papaya* in order to compare differences in development between the plants resulting from the influence of those genes.

This study is significant in that: (1) Findings will better develop *Carica papaya* for use in molecular, genetic and biotechnology research by determining the procedures best suited for use with this organism; (2) Data will expand understanding of the developmental genetics of *Carica*

papaya; (3) Data will have the potential to increase understanding of the molecular basis of various mechanisms that drive the evolution of all life on Earth, as well as allowing for a comparison of the varying contributions of those different mechanisms in evolution (particularly gene function versus gene regulation).

Review of Literature

The modern understanding evolution at the genetic level, based around an evolutionary developmental biology framework as described by Carroll [5], supports the idea that form and body plan evolved primarily through alteration in *expression* of highly conserved proteins. Carroll describes a common toolkit of highly conserved and functionally-equivalent proteins operating in much the same way in many distinct species. Evolution of plant form likely occurs by way of mutations to *cis*-regulatory sequences of these toolkit genes or to transcription factors in larger regulatory networks as opposed to mutations in coding sequences of structural genes. That proteins with similar sequences influence the development of distinct traits between separate species suggests an important role for regulation in the production of those proteins. Furthermore, there is evidence that changes in genetic regulatory sequences are capable of explaining the observed evolutionary divergence of phenotypic characteristics and gene regulation across distinct organisms [5]. This suggests that changes to gene coding sequences and protein primary structure (amino acid sequence) are not the sole explanation for observed phenotypic differences in vastly different species; variations in the mechanisms at work behind the production/expression of proteins are important as well.

A subsequent study by Przemyslaw et al. [6], used computer modeling to predict the floral architecture of inflorescences, the structures on plants that produce flowers. According to this model only two genes from the common toolkit of highly conserved plant proteins are necessary to produce the diversity of inflorescences seen in the plant kingdom. Based on this model, this study analyzes two genes that may correspond to this model and may be responsible for converting a racemous inflorescence to a cymous inflorescence.

Materials and Methods

Collection of *Arabidopsis thaliana* Seeds

Before a full study of *Carica papaya TFL1* homologs and *PFL* can be possible, a mutant Arabidopsis plant must be produced that contains both *TFL1* and *LFY* knockout mutations. In order to obtain organisms displaying the target mutations, Arabidopsis seeds were planted and allowed to mature. The collection of seeds began as soon as plants began producing seeds. *TFL1* and *LFY* knockout seed stocks have since been obtained.

Extraction of *Carica papaya* DNA

In order to isolate genomic DNA from *Carica papaya*, young shoot tips and stems were collected from papaya plants located in the University of Northern Iowa Botanical Center. The extraction of genomic DNA was accomplished using the GE Healthcare Illustra Nucleon Phytopure Genomic DNA Extraction Kit (Fisher Scientific, USA). This produced two samples, referred to as *Carica papaya* DNA samples 1 and 2.

Agarose gel electrophoresis was used to visualize and assess the size and quality of the extracted *Carica papaya* DNA. 100 mL 1X TBE (Tris base/borate/EDTA) was used along with 0.7 g Genetic Analysis Grade Agarose powder (Fisher Scientific, USA) to make a 0.7 % gel. 2 µl ethidium bromide 10 mg / ml was added to stain the DNA. The gel ran for 35 minutes at 150 volts.

Additionally, the concentration of the extracted papaya DNA samples were measured using a NanoVue brand spectrophotometer. The following values were obtained:

(1) Concentration of *Carica papaya* DNA Sample 1: 500 ng / µl

(2) Concentration of *Carica papaya* DNA Sample 2: 595 ng / µl

Design of Primers for the Polymerase Chain Reaction

Primers for use in polymerase chain reaction (PCR) were designed based on the DNA coding sequences for the *Carica papaya TFLI* homologs c32595.1, c45053.1, sc3387.3 and sc107.26 (Table 1). This sequence information was obtained from an online database maintained by the National Center for Biotechnology Information. See the Appendix for the genomic DNA sequences of all papaya homologs used in this study. This data was entered into the molecular biology/bioinformatic software Geneious (Biomatters, New Zealand), and primers were constructed using the software's primer design tools. The primers were synthesized by Integrated DNA Technologies (IDT, USA).

c32595.1 Forward	AGAGACCCTTTGGTTGTAGGGCGA
c32595.1 Reverse	CCACCTGAACCACTTTCTCTTTGGCA
c3387.2 Forward	ATGACCGACCCGGACGCACC
c3387.2 Reverse	GCTCGCCGGCTCCTTCTGAG
c45053.1 Forward	AGGGAATTATGAGAGTTGCAGAGCCT
c45053.1 Reverse	TGTTGCATCGGTTGTGCCTGGA
sc107.26 Forward	TGGAGCCACTGACTATAGGGAGAGTT
sc107.26 Reverse	TCTTGCAGCAGTTTCTCTTTGGGCA

Table 1. PCR Primers for *Carica papaya TFL1* Homologs

Design of Primers for the GenomeWalker Protocol

Primers for use with the GenomeWalker protocol (Clontech, USA) were designed for the homologs sc32.18 and *PFL*. Like the primers constructed for PCR, these GW primer sequences were designed using Geneious, based on specifications described in the kit's protocol. The sequence used to design the sc32.18 GenomeWalker primers was taken from PCR products amplified and analyzed by a former Honor's student researcher from Dr. Sliwinski's lab. The *PFL* sequence used was obtained from the National Center for Biotechnology Information. The sequences of GenomeWalker primers can be found in Table 2.

Primer	Sequence
Adaptor primer 1 (AP1; 22-mer)	GTAATACGACTCACTATAGGGC
Nested adaptor primer 2 (AP2; 19-mer)	ACTATAGGGCACGCGTGGT
sc32.18 specific primer 1	CACCTCCATGAACCTCCACT
sc32.18 specific primer 2	GGGAAGATTTTGATAAGATGAGATAGA
GenomeWalker adaptor	5'- GTAATACGACTCACTATAGGGCACGCGTGGT CGACGGCCCGGGCTGGT -3'
	5'-PO_4-ACCAGCCC-NH_2-'3
PFL specific primer 1	GCCAACTTTGCCATGCCGAGCGCAATAATA
PFL specific primer 2	TCAGCGCGAACAGCAGCCTTAATCCCATAC

Table 2. GenomeWalker primers for *Carica papaya* homologs sc32.18 and *PFL*

Polymerase Chain Reaction of *Carica papaya* DNA

50 µl PCR reactions were used in an attempt to amplify the papaya *TFLI* homologs c4505.31, c32595.1, s3387.2 and sc107.26, as well as the papaya *LFY* homolog *PFL*. Each reaction contained the following:

(1) 25 µl GoTaq Green Master Mix (Promega, USA)

(2) 10 µl DNA grade, protease and DNase free water (Fisher Scientific, USA)

(3) 5 µl *Carica papaya* genomic DNA template (55 µg/µL)

(4) 5 µl forward primer

(5) 5 µl reverse primer

The primers were run using in a Bio Rad MyCycler Thermal Cycler using a custom-based PCR program with the following settings:

(1) 95° C for 2 minutes

(2) 94° C for 30 seconds

(3) 50° C for 30 seconds

(4) 72° C for 3 minutes

(5) 72° C for 10 minutes

(6) 8° C

Agarose gel electrophoresis was used to visualize the PCR products. 100 mL 1X TBE (Tris base/borate/EDTA) was used along with 1.2 g Genetic Analysis Grade Agarose powder (Fisher Scientific, USA) to make a 1.2 % gel. 2 µl ethidium bromide 10 mg / ml was added to stain the DNA. The gel ran for 15 minutes at 220 volts.

Restriction Digest of *Carica papaya* DNA

The following components were combined in a 200 µl PCR reaction tube and mixed by inversion:

(1) 25 µl *Carica papaya* genomic DNA

(2) 8 µl one of four restriction enzyme

(3) 10 µl 10x restriction enzyme buffer

(4) 57 µl deionized water

The reactions were incubated for 2 hours at 37° C in Bio Rad MyCycler Thermal Cycler. Following this incubation, the reactions were vortexed for 10 seconds and then placed back into 37° C incubation for an additional 17 hours.

Agarose gel electrophoresis was used to visualize the test reactions. 100 mL 1X TBE (Tris base/borate/EDTA) was used along with 0.6 g Genetic Analysis Grade Agarose powder (Fisher Scientific, USA) to make a 0.6% gel. 2 µl 10 mg / ml ethidium bromide was added to stain the DNA. The gel ran for 15 minutes at 220 volts.

Restriction Digest Test: pGLO plasmid and PCR Product 55

In one 200 µl PCR reaction tube, 5 µl pGLO plasmid and 15 µl of master mix (2 µl DraI restriction enzyme with 4 µl NE 4 enzyme buffer and 24 µl nuclease-free water) were combined. The remaining 15 µl of master mix was added to a second tube containing 5 µl of PCR Product 55. These restrictions were incubated for 1 hour at 37° C in a Bio Rad MyCycler Thermal Cycler. Following this incubation, the reactions were vortexed for 10 seconds.

Agarose gel electrophoresis was used to visualize the test reactions. 100 mL 1X TBE (Tris base/borate/EDTA) was used along with 0.8 g Genetic Analysis Grade Agarose powder

(Fisher Scientific, USA) to make a 0.8% gel. 2 µl ethidium bromide 10 mg / ml was added to stain the DNA. The gel ran for 15 minutes at 220 volts.

TRIzol Purification of *Carica papaya* DNA

The two samples of *Carica papaya* DNA was further purified using TRIzol reagent (Ambion, USA) following the manufacture's protocol.

Results

Carica papaya DNA Extraction

Samples of papaya DNA were obtained and evaluated for size and quantity using agarose / EtBr gel electrophoresis (Figure 2). Gel lane M contained 1 Kb(+) DNA ladder and lanes 1 and 2 contained duplicate papaya DNA extractions. The papaya samples contained genomic DNA at the expected size of greater than 12 kb (the largest band in the ladder lane). Additionally, the samples were found to contain a high level of smaller nucleic acids at approximately 1 Kb and smaller. This likely represents co-extracted mRNA and other RNAs.

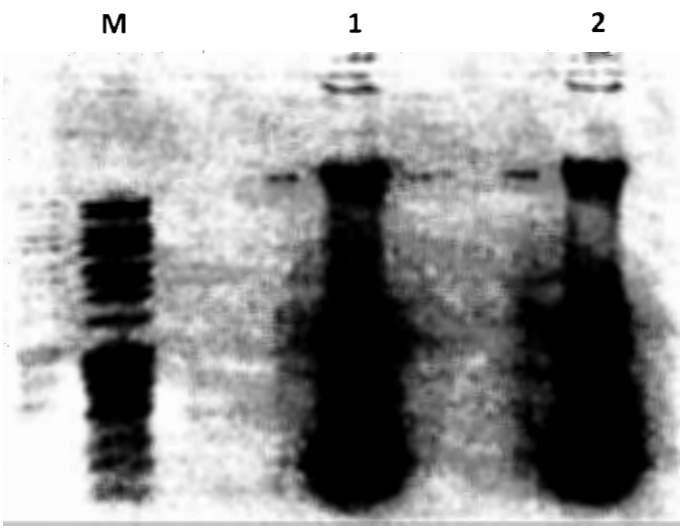

Figure 2. Samples of extracted *Carica papaya* DNA.

Polymerase Chain Reaction of *Carica papaya* DNA

PCR amplification of three papaya *TFLI* homologs was attempted using forward and reverse primers designed based on the genes' sequences in the papaya genome (Figure 3). Table 3 provides the predicted sizes of the *TFLI* homologs. The homologs c45053.1 and sc107.26 may

have amplified, as indicated by a bands at roughly 333 bp and 510 respectively, but all reactions also contained bands at inappropriate sizes indicating non-optimal PCR conditions. This could be caused by a number of different factors including mis-priming of the PCR primers due to problems with the papaya genome sequence used to design the primers. To alleviate this problem new degenerate primers could be designed based on conserved sequences in the TFL1 genes found in all plants.

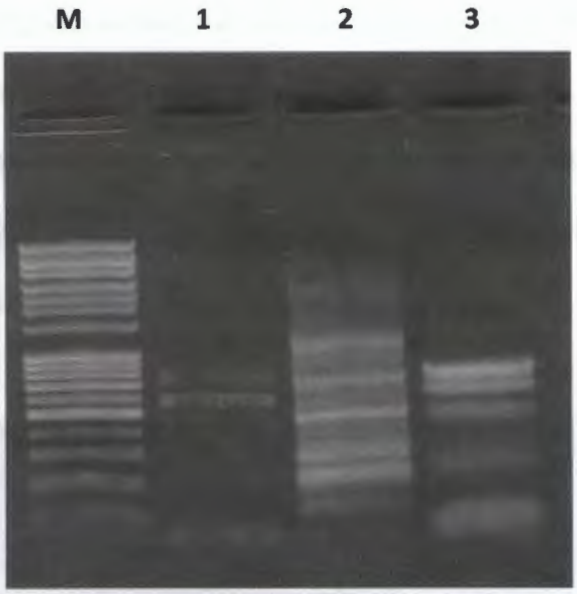

Figure 3. PCR amplification of *Carica papaya* homologs. M: 1kb (+) DNA Ladder; 1: c45053.1; 2: sc3387.2; 3: sc107.26

Homolog	Predicted Band Size (BP)
c325955.1	525
c3387.2	321
c4505.1	333
sc107.26	510

Table 3. The expected band sizes of PCR amplified *Carica papaya TFL1* homologs

Genome Walking to Clone *Carica papaya* Genes

Portions of the coding sequence of two papaya genes of interest, sc32.18 and *PFL*, have been successfully cloned. To obtain the complete gene sequence including *cis*-regulatory regions and the remaining coding sequences, the GenomeWalker (GW) protocol was used. This involves restriction digests of genomic DNA followed by ligation of adaptors to the cut genomic DNA. PCR primers specific to the gene of interest can be used in combination with these adaptors to sequence fragments of the gene one step at a time. With enough steps the full sequence of the gene can be attained by 'walking' along the genome.

First it was necessary to test the papaya DNA extractions for suitability in restriction digests. DNA was incubated with the restriction enzyme AvaII but uncut DNA was still present indicating inadequate DNA purity for this procedure (Figure 4).

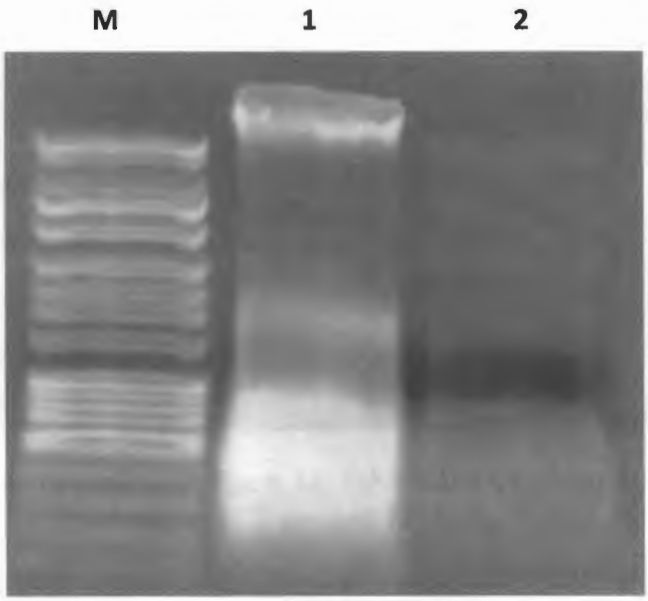

Figure 4. Unsuccessful restriction digest of genomic *Carica papaya* DNA and restriction enzyme AvaII. 1: control without AvaII; 2: AvaII incubated DNA.

Restriction Digest Test: pGlo plasmid and PCR Product 55

In order to determine whether the failures of several attempts at papaya restriction digests were due to impure DNA sample or non-functional restriction enzymes, a test digest reaction was set up using two controls--pGLO plasmid and PCR product 55--and the restriction enzyme DraI. These DNA samples were from an entirely different experiment conducted by another researcher in the lab and were previously shown to be of sufficient purity for restriction digests. The pGLO plasmid was cut as expected in this trial, indicating that the enzyme still function (Figure 5). The results of this digest indicate that the *Carica papaya* DNA extracted earlier during the study seems to fail to meet a necessary base level of purity, preventing the proper digestion of the DNA by restriction enzymes. Subsequent sequence of PCR product 55 showed that this DNA fragment does not contain a DraI cut site.

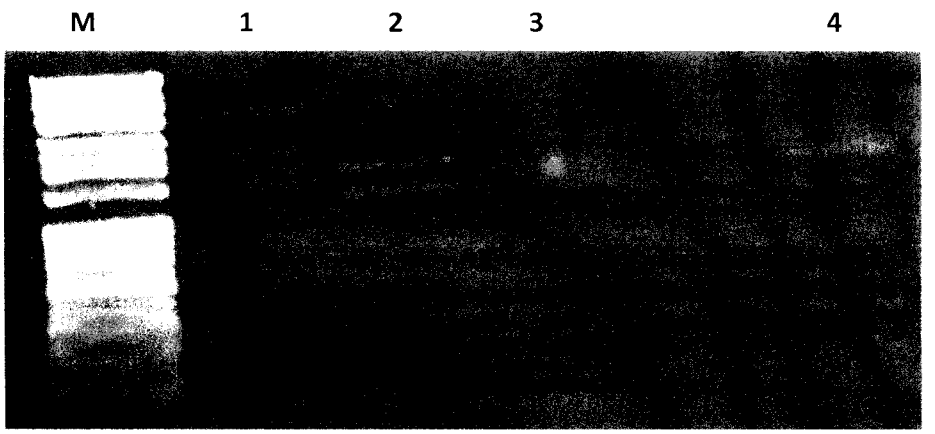

Figure 5. Successful test of restriction digest. M: 1kb (+) DNA Ladder; 1: Non-digest pGLO plasmid; 2: digest pGLO; 3: non-digest PCR product 55; 4: digest PCR 55

TRIzol Purification of *Carica papaya* DNA

Despite a series of attempts, a successful TRIzol reagent-purified sample of papaya DNA was not obtained, as phase separation of target DNA and residual cellular was never achieved, preventing a pure sample.

Discussion and Conclusions

This thesis completed a number of steps in determining the genetic mechanism of inflorescence evolution using the plants *Arabidopsis thaliana* and *Carica papaya.* A DNA extraction from tissue of a mature papaya plant was conducted, supplying the project with DNA stocks for use in various molecular biology techniques and procedures. Using the published genomic data for *Carica papaya*, standard polymerase chain reaction primers were designed for the homologs c32595.1, c45053.1, sc3387.3 and sc107.26 with the goal of amplifying those genes for later DNA sequencing. This will allow for better design of primers for use in a second set of modified PCRs (the GenomeWalker protocol), the products of which can then be used to amplify both coding and regulatory sequences of the homologs, a step upon which future cloning and transformation experiments are entirely contingent. It has already been possible to construct these GenomeWalker primers for the homologs sc32.18 and PFL, and initial attempts have been made to adequately prepare extracted papaya DNA for the component steps of the GenomeWalker protocol.

The project has not been without complications. The GenomeWalker protocol's restriction digest and the TRIzol reagent purification have caused most of the impediments. To test purity of DNA, digest reactions were run using two different, non-papaya sources of DNA: pGLO plasmid and a PCR product 55 originating from a separate study in the lab. These samples were used as control conditions to determine the source of trouble in digest experiments. When reacted with the DraI restriction enzyme (one of four enzymes used in GW protocol to prepare template DNA for ligation to GW adaptors), pGLO plasmid analyzed with agarose / EtBr gel electrophoresis produced the results typical of a successful digest of pGLO by DraI.

PCR product 55 showed no reaction on the gel, though it was later learned that this particular DNA sequence contains no DraI restriction sites, so these results do in fact make sense. Based on the results of this test, it was decided that the most likely explanation for the failure of the papaya digest reactions is the insufficient purity of extracted papaya DNA. To address this, a purification of sample DNA was necessary, and the TRIzol reagent purification was chosen as the method to do so. However, the use of this protocol presented its own difficulties, as after several failed attempts it was still not possible to separate DNA in solution from other cellular components. As time was limited for this project, it has not yet been determined what could be the source of difficulties in this purification. Still, several possibilities have been considered, including the idea that TRIzol reagent is a poor choice for purifying *Carica papaya* DNA. A second possible cause could be poor technique and/or the accumulation of multiple human errors. With these considerations in mind, the next attempt to address this issue will be to test the efficacy of TRIzol reagent using DNA of the organism *Escherichia coli* as a control. This should allow for the determination of whether this purification technique simply does not work well with papaya DNA, or if failure is more likely a result of poor technique and error. If DNA samples from both *Carica papaya* and *E. coli* are successfully purified using TRIzol reagent, this would suggest that prior attempts which ended in failure likely resulted as they did due to procedural error. If, however, when run as a part of the same test papaya DNA sample fail to purify and *E. coli* DNA samples are successfully cleaned, it would indicate that TRIzol may not be the correct reagent to use along with papaya. Completion of these tests will allow for progression from preliminary experimental steps to cloning and genetic transformation phases of this study.

Future Direction

The next step in this project will be to finish the GenomeWalker protocol in order to generate both coding and *cis*-regulatory sequence data for each of the five papaya *TFL* homologs and *PFL*. The steps remaining in this procedure include the ligation of GenomeWalker Adaptors and two rounds of PCR to amplify the genes of interest (a Primary PCR and a "nested" PCR).

However, it is still necessary to obtain coding sequences for several of the homologs (c32595.1, c45053.1, sc3387.3, and sc107.26) by way of a standard PCR reaction and DNA sequencing. New PCR primers will need to be designed for those homologs, as not all were successfully amplified in this project. Once these steps have been completed, it will be possible to design GenomeWalker primers for those sequences, as has already been done for the papaya homolog sc32.18.

The GenomeWalker protocol will allow for the amplification of both the coding and *cis*-regulatory sequences of the homologs. Once these sequences have been obtained it will be possible to clone, and then sub-clone, these genes using two cloning vectors in order to study their functions. The first vector will include a 35S Cassette, which will be used because it contains a plant promoter which will cause the over-expression of *Carica papaya* homologs once those genes have successfully transformed into Arabidopsis plants with a *TFL1* knockout mutation. Once transformed, gene expression for the homologs will be switched on throughout the plant, resulting in Arabidopsis plants expressing a papaya *TFL1* homolog in a plant that would otherwise lack *TFL1* function. In the case of *PFL*--the papaya *LFY* homolog, transformed Arabidopsis plants that would otherwise be without *LFY* function will instead express the homolog. The second cloning vector to be used in this study is pGreen0229. This vector will be used because it makes possible the transformation of Arabidopsis mutants through infection by

Agrobacterium, which will allow for a papaya homolog to be incorporated into the genome of Arabidopsis.

Following transformation, the Arabidopsis plants will be allowed to mature, and their seeds will be collected and then planted. Basta herbicide resistance will be used to select for transgenic plants. Mature transformed Arabidopsis plants will display variable flowering based on the function of the papaya *TFL1* homolog (or *PFL*). As the functions of *TFL1* in papaya are as of yet not understood, the homologs studied in this project may or may not be responsible for producing a determinate meristem, as found in wild type papaya plants. If papaya *TFL1* functions similarly to that of wild type Arabidopsis *TFL1*, phenotypic expression of the papaya homolog transgenic Arabidopsis will display an indeterminate meristem. Conversely, if the papaya *TFL1* homolog is responsible for the production of the determinant meristem observed in wild type *Carica papaya* plants, transgenic Arabidopsis will flower as if the plant had not been transformed, resulting in a determinant meristem identity.

However, before *Carica papaya* homologs can be studied as outlined above, a mutant Arabidopsis plant must be obtained that contains both *TFL1* and *LFY* knockout mutations. *TFL1* knockout seed stocks have been obtained, as have *LFY* knockout seeds. These will be planted, and mature individuals displaying either one of the two desired knockout mutations will then be collected and crossed in order to produce Arabidopsis plants with both gene knockouts. It is the progeny of these crosses that will later serve as targets for transformation of Arabidopsis via *Agrobacterium*.

It was expected that this research project would provide new insights and understandings into the developmental genetics of the plant species *Carica papaya*. Specifically, it was anticipated that data obtained from this project would help elucidate the functions of *Carica*

papaya homologs of the *Arabidopsis thaliana TFL1* and *LFY* genes, which in *Carica papaya* are predicted to be associated with the growth of the plant's cymose inflorescence and terminal flowers. Though complications during this stage of research limited the conclusions that can be made about these genes' functions in the development of papaya, the progress made towards these larger goals has been well established in this thesis. This project has also produced useful information on the efficacy of various molecular techniques and procedures as applied to papaya; this should assist in directing protocol choices in the future as there is now a better understanding of which techniques work well with papaya.

It was expected that, based on the findings of this research, novel insights into the evolution of related plants and families of plants may be uncovered, providing a more fully realized picture of a small chapter in life's long evolutionary drama. Again, though methodological difficulties have slowed data collection, the steps taken in this thesis provide knowledge of how to better work with and study papaya at the molecular level. This increased knowledge will likely prevent the waste of time and resources in the future of the study.

Furthermore, it was predicted that answers to the highly relevant question of whether the influence of these genes in *Carica papaya* as compared to Arabidopsis is associated more so with gene/protein expression or protein function may be uncovered. Explanations provided by this particular portion of the project may also serve as valuable tools of extrapolation, assisting in progressing the current understanding of the various processes and mechanisms that drive the evolution of life on Earth. The work completed for this thesis project makes possible the further study of the genes, and when finished should provide insight on this topic.

References

1. Somerville C, Koornneef M (2002) A fortunate choice: the history of Arabidopsis as a model plant. Nature Reviews 3: 883-889

2. Qingyi Y, Moore PH, Albert HH, Roader AH, and Ming R (2005) Cloning and characterization of a FLORICAULA/LEAFY ortholog, PFL, in polygamous *Carica papaya*. Cell Research 15: 576-584

3. Fusheng W, Wing RA (2008) A fruitful outcome to the papaya genome project. Genome Biology 9: 277

4. Image credit for Figure 1:
http://www.excellup.com/InterBiology/elevenbioimage/11bio31.GIF

5. Carroll SB (2005) Evolution at Two Levels: On Genes and Form. PLoS Biol 3(7): e245

6. Przemyslaw P, Erasmus Y, Lane B, Harden LD, Coen E (2007) Evolution and Development of Inflorescence Architectures. Science 316: 1452-1456.

Appendix

Genomic DNA sequences of *Carica papaya TFL1* homologs and *PFL*

c32595.1

ATGACAAGGGATAGAGACCCTTTGGTTGTAGGGCGAGTTATTGGAGATGTTCTAGAT
CCTTTTATCAGGTCTATTTCTCTCAGAGTGAACTACAACTCAAGGGAGGTTAACAAT
GGCTGCGAGCTCAAACCCTCTCAAGTTGTCTCACAACCAAGGGTTGATATTGGTGGA
GATGATCTTAGGACTTTCTATACTTTGGTGATGGTGGATCCTGATGCACCCAGCCCA
AGTGATCCAAACCTCAGGGAGTACTTGCATTGGTTGGTCACTGATATTCCTGCAACT
ACTGGGGCATCCTTTGGACAAGAGATCGTTTGTTATGAAAATCCAAGACCAACAGTG
GGGATTCATCGATTTGTGTTCGTACTGTTTCGTCAACTGGGGAGACAAACAGTATAT
GCACCAGGTTGGCGTCAAAACTTCAATACAAGAGACTTCGCAGAGCTTTACAATCTT
GGATTACCTGTGGCATCCGTTTATTTCAATTGCCAAAGAGAAAGTGGTTCAGGTGGA
AGGAGAAGATAA

c3387.2

ATGACCGACCCGGACGCACCCAGCCCGAGTGAGCCTAGTATGAGAGAATGGGTCCA
TTGGGTCGTGGTGGACATTCCTGGAGGCACCAACCCGACTCGAGGGAAAGAGATCC
TGCCCTACATGGGTCCACGTCCGCCGGTGGGAATACACCGATACATAATGGTGCTTT
TTCAGCAGAAGCGTCCACTGGGGCTGATCGAGCAGCCGCCATCTCGTGCCAACTTCA
ACACTCGACTGTTCGCCGGGCATTTGGACCTGGGGCTTCCGGTAGCCACCGTTTACT
TCAATGCTCAGAAGGAGCCGGCGAGCCGCAGGCGTTGA

c45053.1

ATGGAAAAGGGAATTATGAGAGTTGCAGAGCCTCTGATTGTGGGAAGAGTTGTGGG
AGACGTTCTGGATTCTTTCAGTCCCACCATTAAAATCTCTGTAACTTACAACAGCAA
CAAGCATGTCTCTAATGGCCATGAAATCTTCCCTTCCACACTTGGTTTCAAACCTAG
GGTTGAGATTCACGGAGGTGACATGAGATCCTTCTTCACACTGGTTATGACAGATCC
AGATGTTCCTGGACCTAGTGATCCTTATCTAAGAGAGCACCTGCACTGGATGGTGAC
CAACATTCCAGGCACAACCGATGCAACATTTGGTAGGGTTACCAATTAA

sc107.26

ATGGAGCCACTGACTATAGGGAGAGTTATAGGAGAAGTTGTGGATGTTTTTACGCCC
AGTGTCAAAATGACGGTCACTTACAATTCTAATAAGCAGGTTGCTAATGGTCATGAG
CTTATGCCTTCTGTCATTGCTGCAAAACCTCGTGTTGAGATTGGTGGAGATGACATG
AGATCTGCTTATACATTAATCATGACAGATCCAGATGCTCCAAGCCCAAGTGATCCG
TACTTGAGAGAACATCTGCACTGGATGGTGACAGATATTCCTGGAACCACTAATGCT
TCCTTTGGAAGAGACATTATAAGCTACGAGACTCCAAAGCCGGTGGTGGGAATCCA
TAGATATGCTTTCATATTATTCAAGCAAAAAGGAAGACGAACTGTGAGGCCGCCGG
CTTCCAGGGACCATTTCAACACAAGGAGCTTCTCTGCAGAAAACGGTCTCGGCCTGC
CCGTCGCTGCCGTCTACTTCAATGCCCAAAGAGAAACTGCTGCAAGAAGACGATAA

sc32.18

ATGATGAGTGATCCTCTTGTGGTTGGTAGAGTGATTGGAGAAGTGGTTGATCCTTTC
ACTCCAAACGTGAAAATGGCCGTCACTTATAATTCCAACAAGCAGGTTTTTAATGGC
CATGAACTGTTTCCTTCTGCTGTCACTCATAAGCCCAAAGTGGAGGTTCATGGAGGT
GATATGAGATCATTTTTCACTCTGGTGATGACAGACCCAGATGTTCCTGGTCCAAGT
GACCCTTACCTGAGGGAACACCTGCACTGGATTGTAACCGACATACCAGGCACAAC
AGATGCAACATTTGGAAAGGAGGTGGTGAGCTATGAAATGCCTCGACCAAACATAG
GGATTCACAGGTTTGTGTTCCTTCTGTTCAAGCAAAAACGGAGGAATACAGTGATGA
GTACTCCACGTTCAAGGGATAAGTTTAATACAAGAAAGTTTGCAGAAGACAGTGAA
CTTGGTCTCCCTGTTGCTGCCGTCTTCTTCAACTGTCAAAGAGAAACAGCTGCTAGA
AGACGCTAA

PFL

ATGGATCCAGACGGCTTTTCTTCCAGCTTGTTCAAGTGGGACCCAACGAGGGGAATA
GTGCAGGCGCCAGTCAGGTTGCTGGAGGCGGTAGCTGCGGCGCCTACGCAGGCGGC
GTACGGAGTGAGGCCGAGGGAGCTGGGTGGTCTAGAGGAGCTTTTTCAAGATTACG
GCATCAGGTACTTCACCGCTGCGAAGATCGCCGAGCTGGGTTTCACGGCTAGCACGC
TGGTGGATATGAAGGATGAGGAACTGGACGAGATGATGAACAGCTTGAGCCAGATT
TTTAGGTGGGAGCTTCTGGTGGGAGAGAGGTATGGGATTAAGGCTGCTGTTCGCG
CTGAAAGGAGGCGGCTTGACGACGACGATTCCAGAAGAAGACA GACCCTCTCTA
CTGACACTACCCACGCTCTCGATGCTCTCTCCCAGGAAGGTCTA
AAGGTGGTGGGTGAGGAGGAGCAGGAGGAGGAAAACGGTGGTGGAAGCGAGAG
ACAGCGCGAGCACCCTTTCATCGTGACAGAGCCTGGGGAGGTGGCACGTGGCAAAA
AAAATGGCCTTGATTATCTCTTCCACTTGTACGAGCAGTGTCGTGACTTCTTGATCCA
AGTCCAGAACATCGCCAAGGAGCGAGGAGAAAAGTGTCCCACGAAGGTGACGAAC
CAGGTGTTTAGATATGCAAAGAAAGCTGGGGCGAGTTATATAAACAAGCCAAAAAT
GCGACACTATGTGCACTGCTATGCTTTACACTGTCTTGACGAGAAGG
AATCAAATGCGTTGAGGACAGCATTTAAGGAGAGAGGAGAAAATGTAGGGTCGTGG
AGACAGGCGTGTTATAAGCCTCTTGTCGCCATTGCAGCACGCCAAGGTTGGGACATT
GATGCCATTTTCAATGCACATCCTCGTCTTGCCATTTGGTATGTCCCCAACAAGCTTC
GCCAACTTTGCCATGCCGAGCGCAATAATACTGCCATTGCTTCTACCTCCGCGGCTG
CTCATCATCTTCCATTCTAA