THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF CHEMICAL ENGINEERING

METHODS FOR ENHANCING SOMATIC EMBRYOGENESIS IN *THEOBROMA CACAO*

RACHEL LYNN ERWIN
SPRING 2015

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree in Chemical Engineering
with honors in Chemical Engineering

Reviewed and approved* by the following:

Wayne R. Curtis
Professor of Chemical Engineering
Thesis Supervisor

Themis Matsoukas
Professor of Chemical Engineering
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

Theobroma cacao, the chocolate tree, is an important cash crop for many countries in Africa and South America, but many trees have become infected with various pathogenic fungi. In order to combat the crop losses due to the fungi, alternative methods of propagation have been explored in order to clonally propagate superior ecotypes of *cacao*, including somatic embryogenesis (SE). SE is faster and more consistent than previously used methods. Attempts have been made to enhance somatic embryogenesis in plants that are phylogenetically close to *cacao*, including cotton (*Gossypium hirsutum)* and rapeseed (*Brassica napus*). These attempts include embryogenic protein domain (PLA1) addition to the media of cotton and the overexpression of the BABYBOOM (BBM) gene in transgenic rapeseed tissue. *Cacao* SE is very time-consuming, so an alternative method of study was needed. In order to preliminarily test the effectiveness of these methods for enhancing SE, they were tested on *A. thaliana*.

I had three large efforts in the lab. The first was helping Sergio Florez, a PhD candidate, acquire gene expression level data in *cacao* at different timepoints during SE. The second was my own experiment to test the effectiveness of embryogenic genes (BBM, AIL5, and LEC2) at enhancing SE in *Arabidopsis*. The third was another independent experiment to test the effectiveness of PLA1 addition to the media at enhancing SE in *Arabidopsis*.

The results of my experiments show that the overexpression of transcription factors and PLA1 addition did not have a significant effect in enhancing somatic embryogenesis in *Arabidopsis*. The number of embryos per explant and the percent responsiveness of the tissue did not significantly increase with the addition of PLA1 proteins nor overexpression of genes. However, the SE gene expression analysis in *cacao* did yield potential genes to be considered for

use as biomarkers that would indicate a tissue's embryogenic potential. More work remains for enhancing somatic embryogenesis, but these preliminary experiments give some insight into methods of increasing crop production via SE.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Wayne Curtis for his time and effort throughout my years in his lab. Dr. Curtis has a contagious love of learning, an insatiable curiosity, and unmatched determination. I will miss the Curtis Lab atmosphere and all of the past and present members I had the fortune of getting to know. Curtis Lab has made my experience at Penn State an unforgettable one.

I'd also like to thank Sergio Florez for his countless hours of training and mentoring. Sergio always had words of encouragement for me when research was slow, and without him, I may well have given up at times. He showed me what seems like hundreds of lab techniques that allowed me to execute my own research, and he always made sure that I knew the scientific background as to why each step was being done instead of blindly following a procedure. I wish Sergio the best in his career and life as (soon-to-be) Dr. Florez.

I'd like to also thank Tina Lai for establishing the PLA project, for cloning the PLA genes, and for teaching me all the protein purification techniques. Tina made being in lab

Finally, I'd like to thank David Currie, my mother Karen Erwin, my father Gary Erwin, and all my other friends and family. Their endless support has helped me immensely as I navigated through my curriculum and research.

Chapter 1

A Summary and Chronology of My Work in Curtis Lab

I started working in Dr. Wayne Curtis's lab halfway through my freshman year. My first three semesters were spent working with Sydney Shaw, Matt Curtis, and Sergio Florez on the plastic bag temporary immersion bioreactor and also learning basic lab techniques such as plant tissue culture, PCR, *Agro*-mediated transformation of *Arabidopsis*. The summer after my sophomore year, I spent the summer working in the lab, supported by NSF funding for the *cacao* project in Curtis Lab (NSF CBET grant #1035072). During the summer, I primarily assisted Sergio Florez with his work on gene expression analysis in *cacao* somatic embryogenesis. I assisted with *Arabidopsis* and *cacao* transformations, RNA extractions, cDNA preparation, and embryo counts. A paper has recently been accepted to BMC Plant Biology, in which I am a co-author, for the work with BBM in *T. cacao.* (Florez, 2015) To supplement Sergio's work with BBM in *cacao,* I tested the effect of overexpressing transcription factors (BBM, AIL5, LEC2) on *A. thaliana* somatic embryogenesis. This experiment was executed during my junior year, and the results are described in Chapter 5.

Tina Lai, a masters student in Plant Biology, joined Curtis Lab shortly after I did and started the PLA project. Her work involves testing the effect of PLA proteins on *cacao* somatic embryogenesis. To supplement her work, I tested the effect of PLA on *A. thaliana* somatic embryogenesis. This experiment was done during my senior year, and the results are described in Chapter 5.

Chapter 2

Theobroma cacao and Somatic Embryogenesis

Theobroma cacao is a small tree that grows in tropical environments. Its seed pods can be dried, ground, and turned into cocoa powder, which is used to make chocolate. *T. cacao* fuels the $70 billion per year worldwide chocolate industry. It is an important cash crop in West Africa, South America, and more recently, India and Vietnam. Many third-world farmers solely depend on the crop to financially support their families. It is also a replacement crop in Colombia and Peru, allowing farmers to legally farm *cacao* for profits instead of illegal narcotic-producing trees. (Guiltinan, 2007)

Unfortunately, there has been an increase in pathogenic fungi that infect *cacao* trees, such as black pod rot. (Acebo-Guerrero et al., 2012) There are only a few cultivars of *cacao* that are resistant to the fungi. The loss of trees to infection has led to a decrease in supply while the demand for chocolate increases with the growing chocolate industry. New methods had to be explored to try and combat this increasing deficit. Propagation of the resistant cultivars via the seed process leads to undesirable genetic variation, so the deficit lead to the development of clonal tissue culture techniques for *T. cacao*.

There are various techniques for clonal propagation of *cacao* such as root cutting and grafting, but those techniques are labor-intensive, expensive, and the performance is unpredictable. (Guiltinan, 2001) A more consistent and scalable technique is somatic embryogenesis.

Somatic embryogenesis (SE) is a tissue culture method that reprograms somatic cells to become embryogenic. It can be used to clonally propagate plants with desirable characteristics, such as disease resistance. SE was first accomplished in *D. carota* in 1958. (Steward) It has been shown that *cacao* can be propagated via somatic embryogenesis (SE), but with low efficiency and high variation. (Li et al., 1998) Many other crops have also been able to undergo somatic embryogenesis, including cotton (both *cacao* and cotton belong to the Malvaceae family), but there is still much to learn for the optimization of the process for each species and genotype. (Zeng et al. 2006) Plant hormone regulators have been used to induce cell reprogramming for SE, and determining the optimal concentration, time of addition, etc. is very time-consuming. This leads to the exploration of different approaches to enhancing SE.

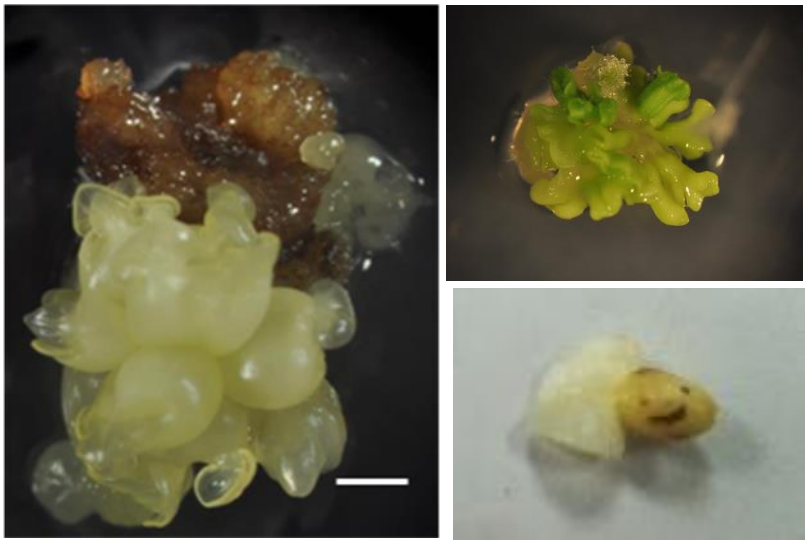

Figure 1 - (Left) *T. cacao* somatic embryos (Image provided by Sergio Florez). (Top Right) *A. thaliana* somatic embryos. (Bottom Right) *T. cacao* somatic embyro with smiley face.

Figure 1 shows *T. cacao* somatic embryos (white, globular bodies). These are a result of specific plant hormones, usually 2,4-D, which help to induce callus formation, reprogram cells,

and form somatic embryos. These embryos can then be converted into plantlets through acclimation to greenhouse conditions and eventually being planted in soil.

This thesis explores two methods to enhance somatic embryogenesis in *Theobroma cacao*, PLA1 addition to the media and the overexpression of transcription factors thought to be involved in SE. Because *cacao* is very slow-growing and tedious to culture, *Arabidopsis thaliana*, a model plant organism, was used to test the methods to enhance SE. This was done to see if *A. thaliana* could serve as an indicator as to how good a method would be at enhancing SE in *cacao*. If *Arabidopsis* could indeed be used as an indicator for a method's performance in *cacao*, it could be a very valuable tool that would save researchers months of time for each experiment. As they say in industry, time is money.

Chapter 3

Methods for Enhancing SE in *cacao*

Phytocyanin-like Arabinogalactan-protein (PLA1) Addition to Media

Embryogenic carrot cell lines have been found to secrete a unique set of proteins. These proteins belong to the arabinogalactan protein family (AGPs). It was found that when these proteins were isolated and added to a cell line that was no longer embryogenic, embryogenicity was restored. This suggests that AGPs play a vital role in somatic embryogenesis. There is a broad range of AGPs that are expressed during different stages of SE, and the AGP levels change until a cell line is no longer embryogenic. (Kreuger, 1993)

It was shown that AGPs were able to promote SE in cotton, a close relative to *cacao*. The active AGPs were shown to be secreted from the cell. Even without glycosylation, the active AGPs retained their SE-promoting activity. This active cotton arabinogalactan protein was sequenced via cDNA analysis and named GhPLA1. SE-promoting activity was found to be greatest at a GhPLA1 concentration of 0.5 mg/L. The GhPLA1 proteins were secreted by Tobacco BY-2 cells and were found to increase the ratio of the likelihood of non-embryogenic calli with added AGPs to generate embryos to that of non-embryogenic calli on control medium. The ratio was found to be 2.27 ± 0.69 for 0.05 mg/L protein concentration and 2.68 ± 0.63 for 0.5 mg/L protein concentration, both with a p-value of <0.001 compared to the control. (Poon et al., 2012) These findings suggest that even a low concentration of extracellular PLA proteins (0.05 mg/L) could have significant effects on the generation of embryos.

Given that the active AGPs are extracellular and do not require glycosylation, they can be produced in bacteria and then added to plant media. The phytocyanin-like domain of GhPLA1 (PL1) was produced in bacteria and added to plant media to test its effectiveness at promoting SE in cotton. It was shown that the addition of PL1 increased the number of embryogenic cotton cell lines. PL1 was added to the media at a concentration of 0.5 mg/L, and it increased the likelihood of restoring embryogenicity to callus by a factor of 2.67 compared to a no-protein control with a p-value of <0.001. (Poon et al., 2012) This is exciting because it opens up the possibility of producing PLA proteins via bacteria at an industrial scale.

Mammalian cell lines are traditionally used when glycosylation of a protein is required, but GhPLA1 is still active without glycosylation, thus, bacteria may be used. The use of bacterial cell lines to produce proteins is much easier and cheaper than using mammalian cell lines. Bacterial cells grow must faster than mammalian cells, resulting in a reduced production time, up to 10-fold. Bacteria are also much more robust than mammalian cells, which means that the operation is less sensitive to disturbances and overall lower maintenance. The media for bacterial cultures is also much cheaper, leading to significantly reduced costs. Mammalian cell cultures require complex media with amino acids, often containing BSA (bovine serum albumin), or synthetic BSA. In contrast, bacterial cell cultures can grow on defined media with simple carbon and nitrogen sources, which are usually much cheaper. It is estimated that the cost of bacterial media is only 10% of the cost of mammalian cell media. (White, 2010) A summary of bacterial versus mammalian cultures is shown in Table 1 below. (SAFC, 2012)

Table 1 - Summary of the advantages of microbial fermentation. (SAFC, 2012)

	Microbial Fermentation	Mammalian Culture
Generation time	20 minutes-hours	Hours-days
Growth length	1-4 days	10-14 days
Product types	- Proteins - Secondary metabolites - Cell wall components - DNA	Proteins
Crude protein titer	1-15 g/L	1-5 g/L
Media cost	Low	High
Growth sensitivity	Low	High
Post translational modifications	Some available in yeast	Yes

After producing PLA in bacterial cells, it must be purified and prepared for storage. The purification could be as simple as using two tangential flow filtration steps, one to purify, and one to concentrate, but that is an engineering design problem not fully solved in this thesis. After purification, the ideal way to store the protein would be lyophilization, or freeze-drying, given that the structure and function of the protein is retained. Freeze-drying proteins allows them to be stored for years as opposed to in solution at 4°C, where the storage time is limited to 1 month. (Pierce Biotech., 2009) If PLA is to be used in industry, companies will want to keep the protein in a form that is as convenient as possible. If the protein is stored in solution, the company will have to worry about expiration dates much more often than if the protein is lyophilized. One downside to lyophilized protein is that it must be reconstituted in order to be used. This could theoretically be bypassed if the protein were kept as a solid and sprinkled into the media.

Transcription Factor Overexpression (BBM, AIL5, LEC2)

One potential method for enhancing SE in *Theobroma cacao* is through the overexpression of key genes known to play a role in cell reprogramming. A factor that makes this method appealing is that it doesn't require the use of expensive hormones. The genes involved are transcription factors that have been suggested to be involved in the cell reprogramming process.

One gene of interest is BABYBOOM (BBM), which was characterized in *Arabidopsis thaliana* and *Brassica napus*. (Boutilier, 2002) It was found that BBM could induce the formation of embryos from specific tissue when the gene was overexpressed. Sergio Florez and I have also recently shown that when the *cacao* ortholog of *Arabidopsis* BBM (AtBBM) is overexpressed in *cacao* tissue, it starts generating embryos from cotyledonary structures without the use of hormones. (Florez, 2015)

Some other genes of interest are AIL5 and LEC2. Like BBM, AIL5 was found to induce embryo formation on cotyledons and increase SE efficiency in *Arabidopsis*. (Tsuwamoto, 2010) LEC2 was found to induce spontaneous embryo formation in *Arabidopsis*, and is thought to be involved in endogenous auxin production. (Wójcikowska, 2013). These genes were tested alongside BBM in this portion of experimentation.

I did some work to try and see if the *cacao* orthologs of BBM, AIL5, and LEC2 expressed in transgenic *Arabidopsis* could enhance SE in *Arabidopsis*. This *Arabidopsis* experiment is similar to the PLA experiment described in the main body of this thesis in that it was designed to act as a shortcut to determine if an effort to enhance SE in *cacao* could be reproduced in *Arabidopsis* and therefore be used to determine its effect in *cacao*. One difference, however, was the ecotype of *Arabidopsis* used. The BBM/AIL5/LEC2 experiment

was executed before the PLA experiment. At this point, the Colombia (Col-0) ecotype was used because that was the type of plant used to make the transgenics. The SE responsiveness of Col-0 ecotype is reported to be 90%. Another ecotype, Wassilewskija (Wass.) was reported to have a 40-60% responsiveness. (Gaj, 2011) Wass. was used in the PLA experiment so that the baseline responsiveness would be lower, allowing PLA to show a greater difference in effect than if the responsiveness was already at 90%.

BBM and other transcription factors that enhance SE are attractive because they cut out the need for added plant hormones. Getting rid of added hormones for plant propagation via SE would save a lot of money on an industrial scale.

Chapter 4

Materials and Methods

PLA Protein Expression and Purification

Tina Lai, a plant biology masters student in Dr. Wayne Curtis's lab, identified two *cacao* PLA domains that are orthologous to GhPLA1. She then cloned these genes into the pET14b vector, and transformed BL21 *E. coli* with the plasmid.

The transformed cells were grown in either auto-induction media or LB media with IPTG induction. Auto-induction contains glucose and lactose and the cells are under the Lac operon. (Studier, 2005) This means that the cells consume glucose and grow until glucose is depleted. Then the cells start consuming lactose, which turns on protein expression. Therefore, cells were grown on auto-induction media overnight at room temperature at 200 rpm. IPTG cultures were grown at 37°C until an OD of 0.4-0.6, then IPTG was added at a concentration of 5 mM. Then, the cultures were grown at room temperature overnight at 200 rpm. After 24 hours, the cells were pelleted at 12,000 rpm at 4°C and stored in -80°C until purification.

Cells were lysed in a buffer containing 8 M urea, 50 mM NaH_2PO_4, 1 mM PMSF, 0.1 mg/mL DNaseI, 0.02 mg/mL RNase, 1 mg/mL lysozyme, adjusted to pH 7.5. 8M urea was used because His-tag purification didn't initially work, suggesting that the polyhistidine tail was buried inside the protein. Urea denatures proteins so that the His-tag could be available for binding. The solution was sonicated in a horn sonicator at 4°C in pulse mode for 3 mins at 50% power, output 4 (Sonicator Cell Disruptor 350, Branson Ultrasonics Corp., Danbury, CT). The

solution was then put at 37°C, 200 rpm for 1 hour and sonicated again five times immediately after. The solution was centrifuged at 12,000 rpm for 20 minutes at 4C (Beckman Coulter Avanti Centrifuge). The supernatant was collected and a Ni-NTA Agarose resin (Qiagen) was added and left to bind protein, stirring, overnight at 4°C.

The slurry was poured into a 25 mL column and the resin was allowed to settle, then washed with 2 mL each of six buffers, shown in Table 2. The buffers had decreasing concentrations of urea to help refold the proteins (which were denatured by the 8 M urea in the lysis buffer). The proteins were eluted using an elution buffer with pH 4.0 (Table 2).

Table 2 - Buffers used for washing and eluting proteins bound to Ni-NTA Agarose resin.

Buffer	**Contents** (All contain 100 mM NaH2PO4 and 100 mM Tris-HCl)
1	8M Urea, pH=8.0
2	6M Urea, pH=8.0
3	4M Urea, pH=8.0
4	2M Urea, pH=8.0
5	0M Urea, pH=8.0
6	0M Urea, pH=6.0
Elution	0M Urea, pH=4.0

2 mL of each buffer was used to wash the proteins, and 10 mL of elution buffer was used to elute. The eluted proteins were run on a 15% SDS gel and quantified via Bradford assay (Thermo Scientific Nanodrop 2000). The his-tag was then cleaved using the Novagen thrombin cleavage enzyme at room temperature for 12 hours. The proteins were then filter sterilized using a 0.2 um acrodisc HT Tuffryn Membrane sterile filter that exhibits low protein binding. After sterilization, the proteins were quantified again via Bradford and brought to the final

concentration of 0.5 mg/L (Poon, 2012) with sterile Milli-Q water. (Protein expression and purification was a joint effort between Tina Lai and me.)

PLA *Arabidopsis* Growth and SE Prep

The *Arabidopsis* SE method used for this thesis was taken from the Gaj methods paper on somatic embryogenesis in *Arabidopsis* from immature zygotic embryos. (Gaj, 2011) *Arabidopsis thaliana* (Wassilewskija ecotype, TAIR) were grown from seeds for 8 weeks in a conviron with a 16-hr photoperiod, watered once every two days. In the eighth week, siliques were picked from the plants and surface sterilized for 10 minutes with a 20% bleach (Clorox) in water solution with 2 drops of Tween 80 per 50 mL. The bleach solution was removed via pipetter, and the siliques were then rinsed with sterile Milli-Q water three times. Figure 2 shows siliques at different stages of development. For this thesis, siliques that resembled the fifth or sixth specimens from the left of Figure 2 were chosen for embryo excision.

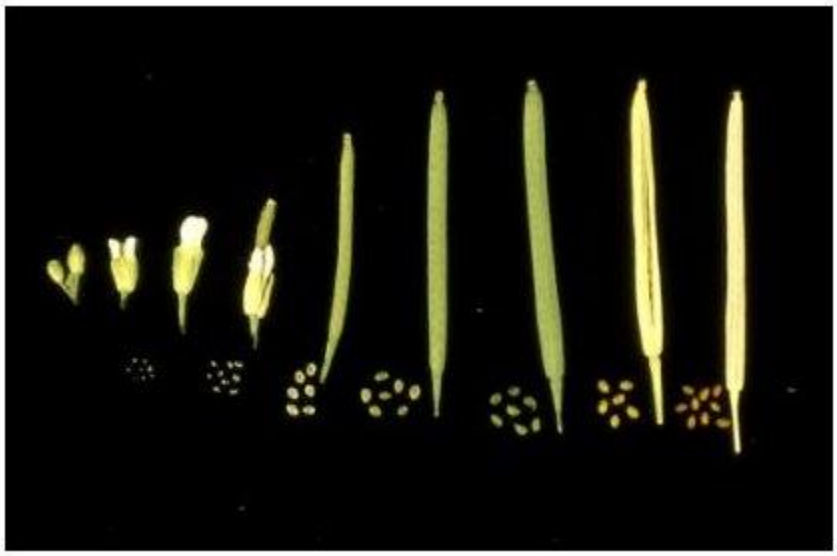

Figure 2 - Different stages of silique development in *Arabidopsis*. Siliques that resembled the fifth and sixth specimens from the left were picked for embryo excision in this thesis. (Meinke, 2008)

Immature zygotic embryos were then excised from the sterile siliques under a microscope in a laminar flow sterile hood using 29 gauge syringe needles. The siliques were placed on sterile petri dishes during excision. The microscope was sprayed liberally with 70% ethanol and wiped dry in the hood in the hopes of it staying sterile enough to not contaminate the cultures. A picture of a silique with immature zygotic embryos is shown below in Figure 3.

Figure 3 - Immature zygotic embryos in seed coats with the silique sliced open. (Meinke, 2008)

The embryos were removed from their underdeveloped seed coats and placed on 60 mm x 15 mm polystyrene petri dishes (Sigma-Aldrich, P5481) with E5 media (Appendix B). A picture of the embryos on the plate is shown in Figure 4.

Figure 4 - Immature zygotic embryos immediately after excision and placement on E5 media. The cotyledons are the top and left 'arms' in the magnified picture.

After 15 days in the conviron (same 16-hr photoperiod), the callus was transferred to plates containing MS20 media (Appendix B). The explants at this point in time are shown in Figure 5.

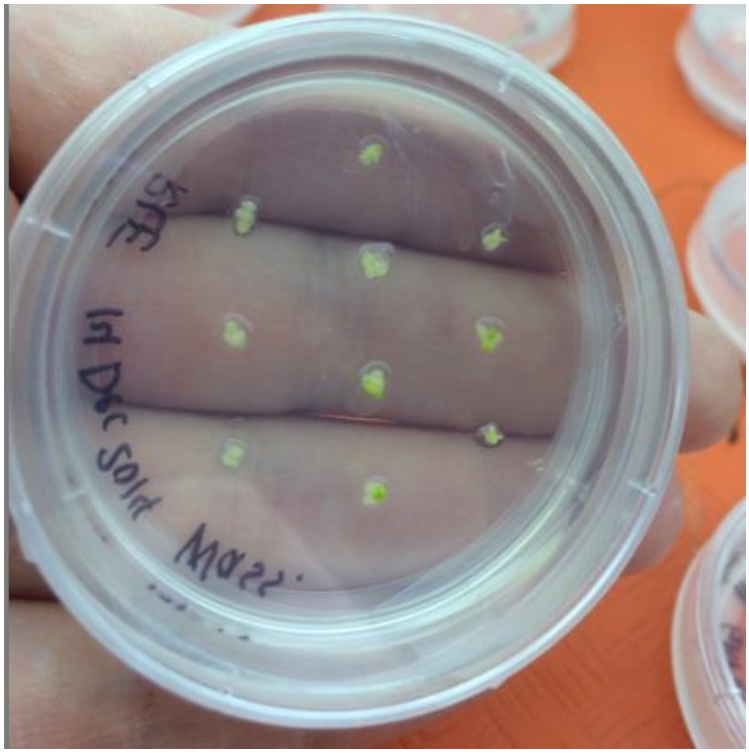

Figure 5 - Callused explants after 15 days on E5 medium, transferred to MS20 medium. The explants no longer have a defined structure. Now, they are simply masses of undifferentiated cells.

After transfer, 7 μL of protein solution were pipetted over the callus so that the solution fully covered the plant tissue. The plates were put back in the conviron so that embryos could form. After 3 and 6 days, the embryos were counted.

Making Transgenic BBM/AIL5/LEC2 *Arabidopsis*

Sergio Florez cloned the genes of interest into *Agrobacterium tumefaciens*, and transgenic *Arabidopsis* Col-0 were produced via floral dip (described in more detail in our recent paper). The seeds from the transformed plants were collected, sterilized the same way that the siliques were in the body of this thesis (20% bleach solution), and plated on MS basal salts medium. The use of antibiotic resistance was used to determine gene insertion.

Figure 6 shows the method used for screening the *Arabidopsis* plants for gene insertion. The AIL5 and LEC2 genes are in the 126 vector (Guiltinan Lab), which has a kanamycin resistance gene. The BBM gene was in the pCAMBIA vector, which contains a hygromycin resistance gene. *Arabidopsis* seeds were sterilized and plated on either MS+kan or MS+hyg plates, and the resistant plants contain the gene of interest. Since the antibiotic resistance gene is present in the transgenic *Arabidopsis*, the plants with gene insertion grow dark green while the wildtype plants turn yellow. The transgenic plants are marked with white circles.

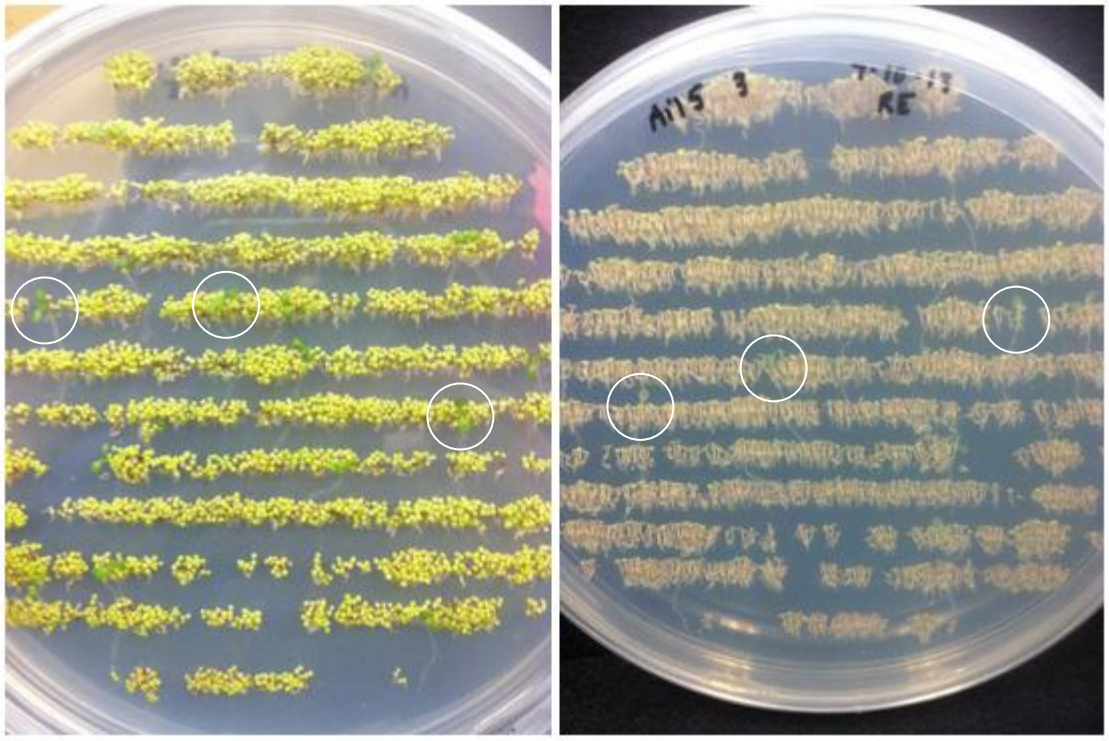

Figure 6 - Screening for transgenic *Arabidopsis*. The Ail5 gene is in the 126 plasmid (Guiltinan Lab), which has a knamycin resistance gene. *Arabidopsis* seeds were sterilizedand plated on MS+kan plates, and the resistant plants contain the Ail5 gene. Resistant plants grow well (dark green), while seeds without the gene turn yellow.

BBM/AIL5/LEC2 *Arabidopsis* Growth and SE Prep

After transgenic *Arabidopsis* were selected via selection plating, the procedure was identical except for the addition of 7 uL of protein. The plants were grown for 7-8 weeks in a conviron, and siliques were picked from each corresponding transgenic plant and sterilized in a bleach solution. The immature embryos were excised and plated on E5 plates for 15 days, then transferred to MS20 plates. Embryos were counted after 10 days, and the results are shown in Chapter 5.

Chapter 5

Results and Discussion

PLA Addition

The *A. thaliana* somatic embryogenesis experiment was run with a wildtype control (Wass. ecotype for decreased baseline reponsiveness), an empty vector control, GhPLA1 with the polyhistidine tail still attached to the protein, GhPLA1 that had its polyhistidine tail cleaved, and TcPLA2 with a polyhistidine tail. The wildtype callus was covered with sterile Milli-Q water. The empty vector control was covered in a solution that came from *E. coli* with the pET14b vector without a gene. The cell lysate was processed in the same way the GhPLA1 and TcPLA2 were with the His-tag purification procedure. Theoretically, there is no protein in this control, but it was kept as a control nonetheless. The GhPLA1 without a His-tag was purified using Nickel resin, but was incubated with thrombin cleavage enzyme. The embryos were counted at day 3 and day 6, but only day 6 is shown because the change between the two counts was insignificant. The average number of embryos per explant and the average number of embryos per responsive explant with standard deviation error bars are shown in Figure 7. Figure 8 shows the percent responsiveness of each protein treatment. The responsiveness was defined as the percentage of explants that produced at least 1 embryo.

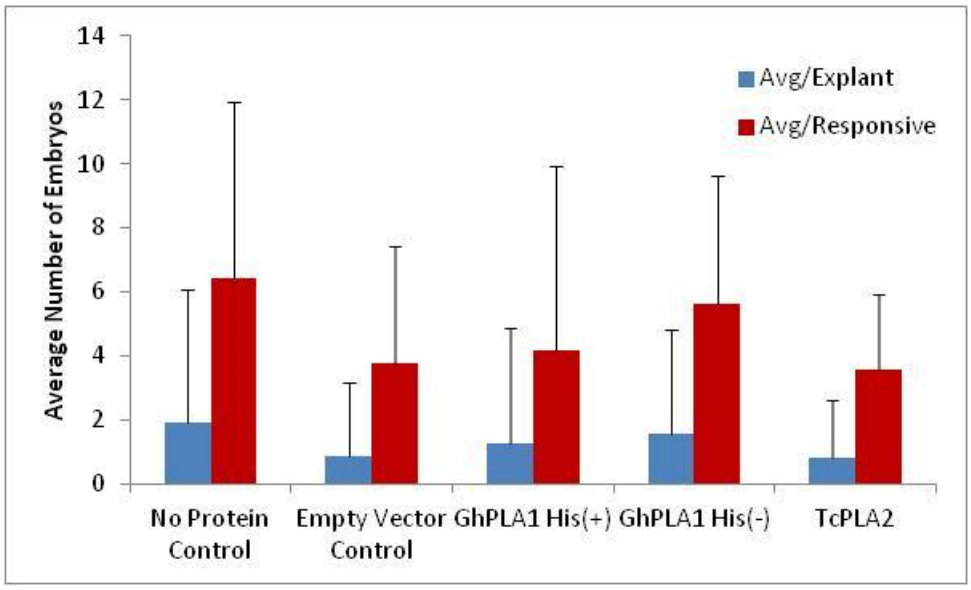

Figure 7 - Average number of embryos per explant and average number of embryos per responsive explant (PLA SE experiment).

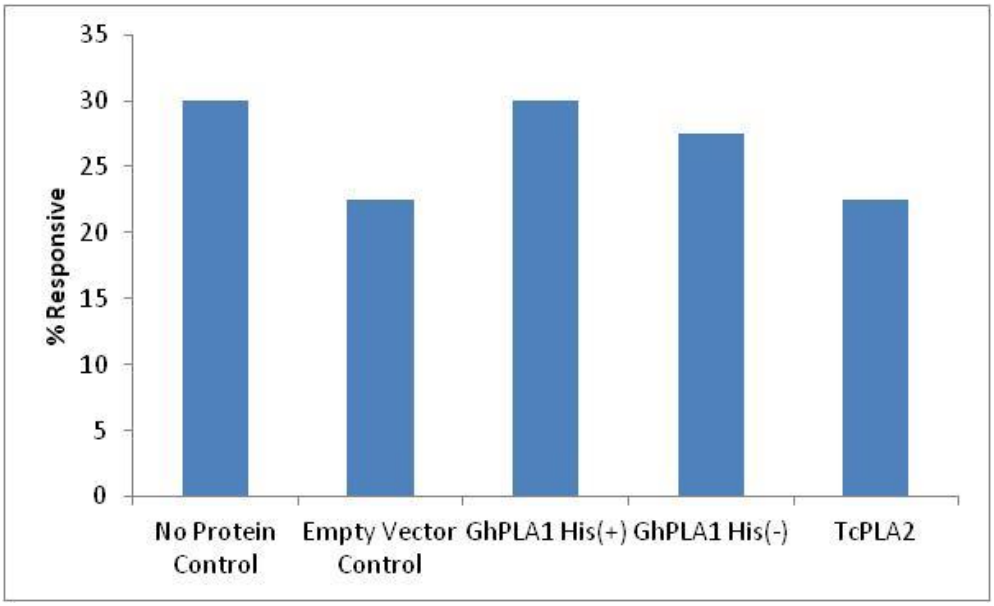

Figure 8 - Percent responsiveness of each protein treatment. A reponsive explant produced at least one embryo during the six days (PLA SE experiment).

The standard error bars are quite large for this experiment, indicating the high variability that often comes with plant tissue experiments. The average number of embryos per explant for each protein treatment does not vary much, indicating that there isn't a very strong effect on the SE system. *Arabidopsis* somatic embryogenesis happens much more rapidly than *cacao* (~2 days in *Arabidopsis* compared to ~40 days in *cacao* to see embryos). It's hard to say why there wasn't a great effect, but perhaps the rapidness played a part.

Pictures of *Arabidopsis* somatic embryos are shown in Figure 9. The left is a no protein control explant with an embryo clearly visible on the top right of the tissue. The explant originated from an immature zygotic embryo excised from its silique and put on 2,4-D to create callus. When the callus was taken off 2,4-D and put on MS20 medium, the embryos were formed. The right is an explant treated with TcPLA2 protein. This explant was unresponsive because it produced no embryos.

Figure 9 - (Left) Empty vector control *Arabidopsis* explant with somatic embryos. (Right) Unresponsive TcPLA2-treated *Arabidopsis* explant (no somatic embryos).

Though the PLA protein addition experiment didn't yield particularly useful results in *Arabidopsis*, does not mean that PLA proteins won't have a better effect on *T. cacao* somatic embryogenesis. Tina Lai, a masters student in Dr. Wayne Curtis's lab, has been researching the effect of PLA proteins on *cacao* somatic embryogenesis.

Transcription Factor Overexpression

This experiment was similar to the PLA experiment in the way it was executed. The main difference is that instead of adding protein to the explants, the explant tissue comes from transgenic *Arabidopsis* with specific transcription factors (BBM, AIL5, LEC2) overexpressed. The Col-0 ecotype was used, and there was an empty vector control (*Arabidopsis* transformed with 126 plasmid without gene insertion, but containing antibiotic resistance). The different BBM categories represent different BBM gene insertion events, meaning that three *Arabidopsis* transgenic plants had the BBM gene inserted into the chromosome, but at different locations.

The embryos were counted at 10 days off hormones. The average number of embryos per explant and the average number of embryos per responsive explant with standard deviation error bars are shown in Figure 10. Figure 11 shows the percent responsiveness of each protein treatment. The responsiveness was defined as the percentage of explants that produced at least 1 embryo.

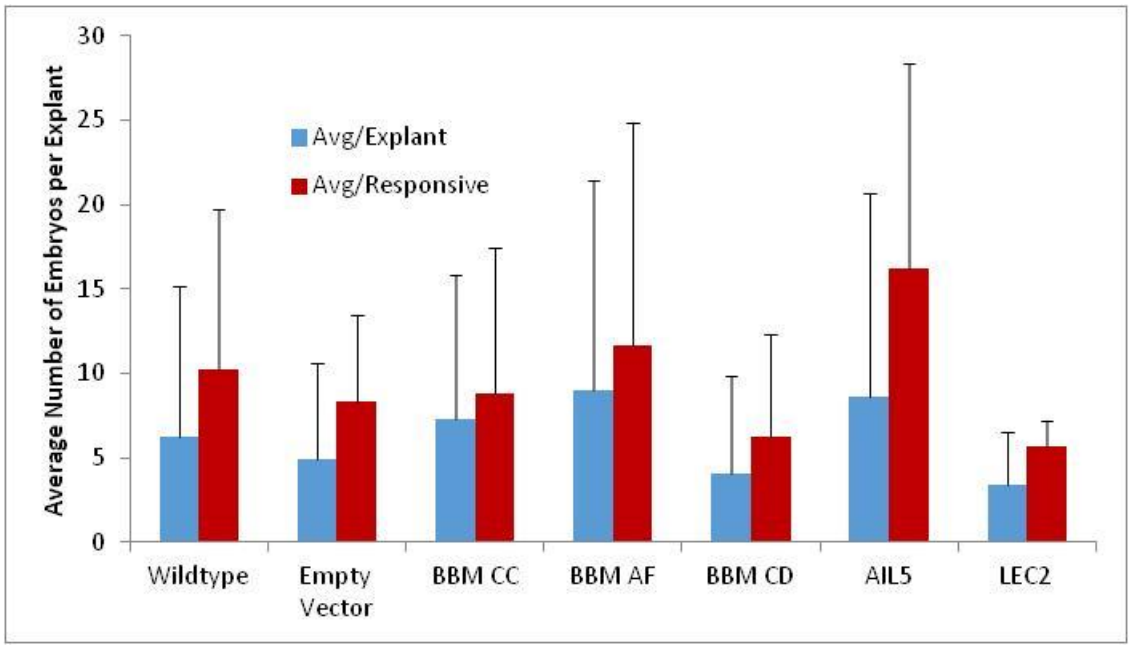

Figure 10 - Average number of embryos per explant and average number of embryos per responsive explant (Transcription Factor SE experiment).

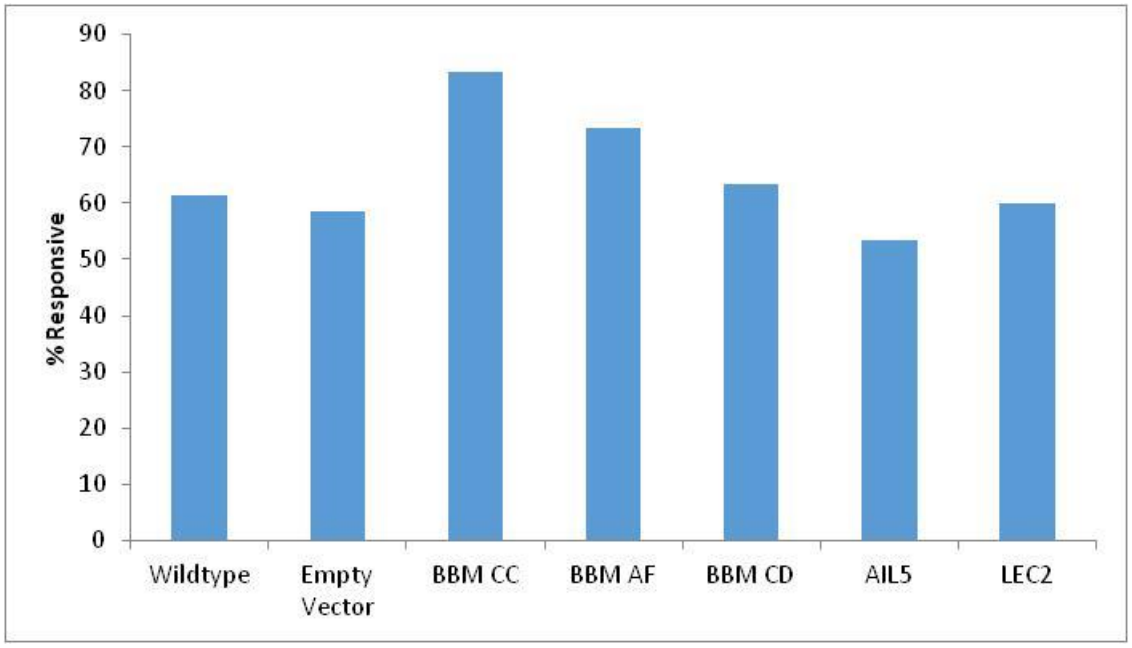

Figure 11 - Percent responsiveness of each protein treatment. A reponsive explant produced at least one embryo during the ten days (Transcription Factor SE experiment).

Compared to the wildtype control, it seems as though AIL5 is increasing the number of embryos per responsive explant, but reducing the percentage of responsive explants. A t-test was run between the average number of embryos per explant for AIL5 and the wildtype control, and yielded a p-value of 0.63, indicating that the results are not statistically significant. Another t-test was run between LEC2 and the wildtype control which yielded a p-value of 0.80, indicating that the results are not statistically significant.

Figure 12 shows explants with embryos from the transcription factor overexpression experiment. The left shows a wildtype Col-0 explant with many embryos while the right shows a transgenic LEC2 explant with less, smaller embryos.

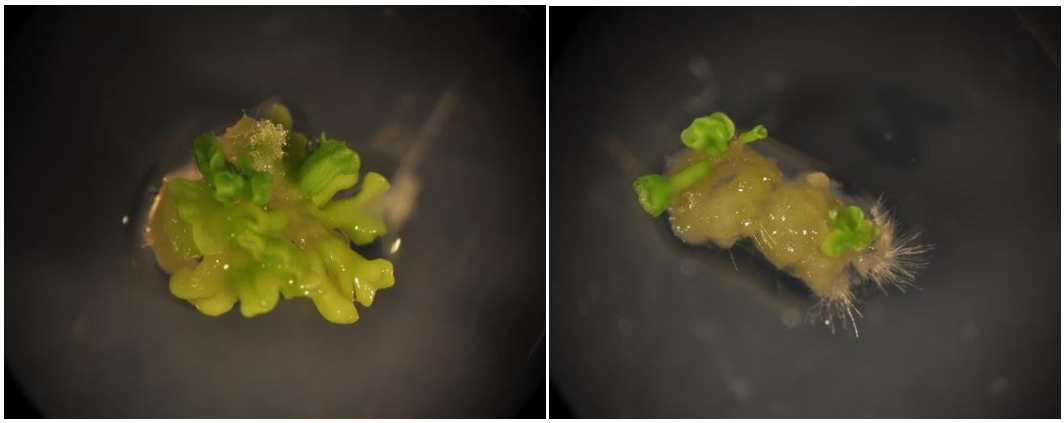

Figure 12 - (Left) Wildtype (Col-0) *Arabidopsis* explant with many somatic embryos. (Right) Transgenic LEC2 *Arabidopsis* explant with small somatic embryos.

Chapter 6

Conclusions and Future Work

The addition of PLA protein domains to *Arabidopsis* tissue did not significantly enhance somatic embryogenesis. It did not increase the number of embryos per explant or the responsiveness of the explants. There is work being done with PLA addition to *cacao* tissue to characterize the effectiveness in that system. Because there was an increase in cotton SE with GhPLA1 (Poon, 2012), there is still a good chance that TcPLA1 or TcPLA2 could enhance SE in *cacao* and that *Arabidopsis* does not serve as a good indicator.

I did the work for this thesis as an undergraduate, so there was just not enough time to carry out a full *cacao* experiment with PLA proteins. This is why I tried to develop an indicator experiment for SE enhancement in *Arabidopsis*. The hope was that the results seen in *Arabidopsis* would directly correlate to how well the protein addition would perform in *cacao*. The results have shown that *Arabidopsis* do not respond in the same way as cotton, and that the system cannot necessarily be used as a quick indicator of SE enhancement.

Potential Biomarkers for Enhanced SE

Sergio Florez is a PhD student in Dr. Wayne Curtis's lab. His main research is trying to enhance SE in *cacao* using a *T. cacao* ortholog of the babyboom (BBM) gene. It was found that BBM expression in specific tissues triggers the formation of embryos on seedlings of

Arabidopsis thaliana and *Brassica napus*. (Boutilier, 2002) Sergio and I have recently expressed BBM in stable transgenic *cacao* tissue, and have shown that the expression causes the spontaneous formation of embryos on cotyledonary structures without the use of plant hormones. (Florez, 2015) Using plant hormones is expensive at an industrial scale, so cutting out the use of them would prove to be very economical for companies using the SE process for plant propagation.

Sergio and I have also done experiments that measure the gene expression of a few key genes thought to be involved in the embryogenesis process. *Cacao* tissue was put on hormones to induce callus formation and taken off hormones at 14 days when embryos were formed. Some *cacao* tissue was sampled at different stages of the process, cDNA was made for each sample, and qPCR was performed to determine the relative expression of the BBM, LEC1, PKL, LEC2, AGL15, ABI3, and CLF, all genes which are thought to be involved in SE (Boutilier, 2012; Lotan, 1998; Ogas, 1999; Wójcikowska, 2013; Heck, 1995; Parcy, 1997; Kim, 1998) . I helped with collecting and transferring the *cacao* tissue, RNA extractions, and cDNA preparation. The levels of the expression of each gene at different time points were normalized to the expression at day 0, and are shown in Figure 13.

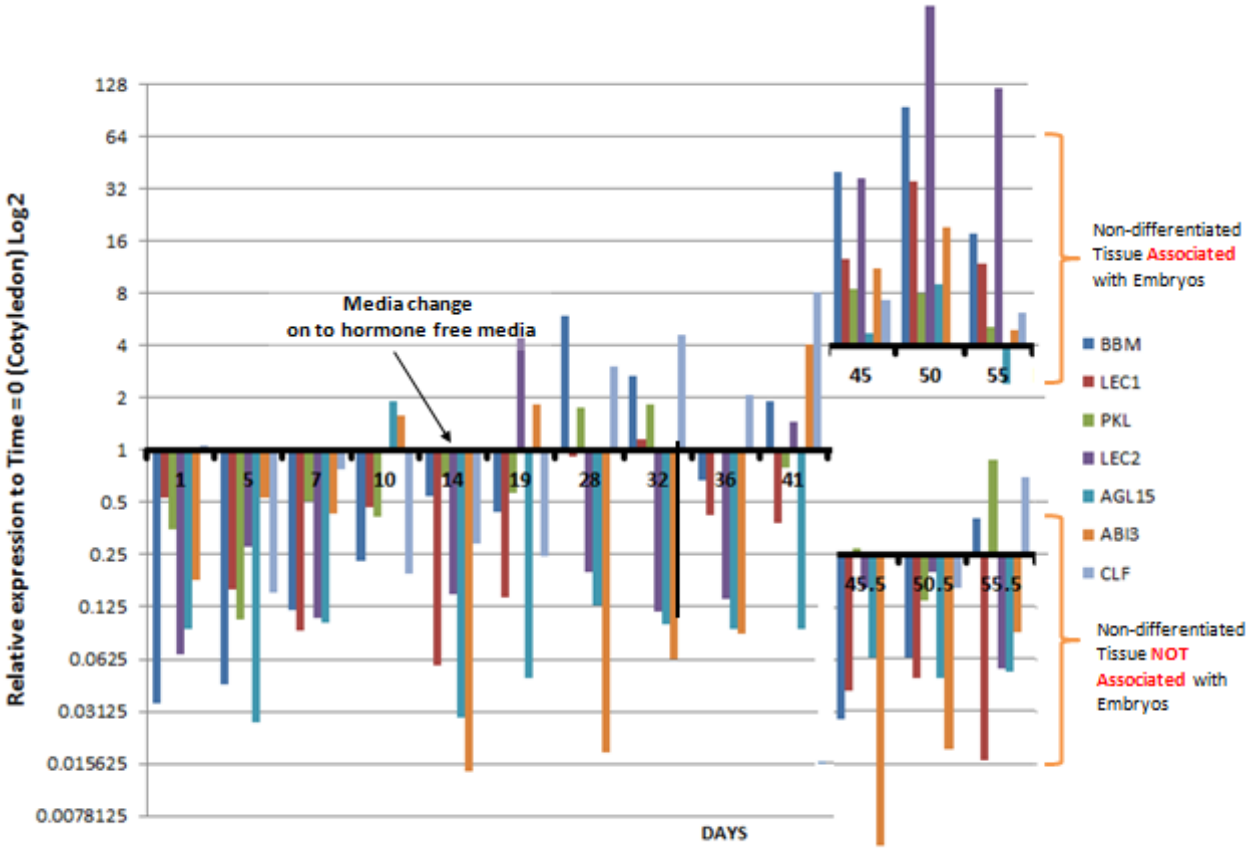

Figure 13 - Expressions of key genes in *cacao* tissue at different time points during the SE process. Tissue was taken off hormones at day 14. The top right section represents gene expression levels of embryogenic tissues, and the bottom right represents gene expression levels of non-embryogenic tissues. All gene levels are normalized to the respective levels at day 0. (Image provided by Sergio Florez)

The expression levels change throughout the SE process, and it is thought that some of these genes can then be used as a biomarker for embryogenic tissue. When the *cacao* tissue is put on hormones (day 1 to 14), the genes are mostly all down-regulated. However, when the tissue is taken off hormones on day 14, some gene expression levels are up-regulated. The tissue does not start producing visible embryos until past day 40, so it was not possible to tell if the tissues being sampled from day 14 to day 40 were embryogenic or not, so the gene levels are an

average of sorts between the time the tissue is taken off hormones and the time visible embryos are produced. It is therefore assumed that these spikes in gene levels are due to the fraction of up-regulation in pre-embryogenic tissues.

If proven to be true, these genes could then be used to predict which tissues are embryogenic before waiting to actually see embryos, theoretically cutting the experimental time from 40 days to a little over 14. Then, researchers (even undergraduates) could perform experiments on *cacao* somatic embryogenesis and see results much quicker instead of using the inconclusive *Arabidopsis* system. A researcher could prepare an experiment exactly like they would for a normal *cacao* SE run, but instead of waiting for embryos to become visible, they could prepare cDNA samples from RNA extractions and analyze the gene expression levels of each sample. If the genes selected as biomarkers are overexpressed or repressed (depending on the gene), then it could be determined if the treatment were working or not. This could be done with the PLA protein addition experiment. The *cacao* tissue could be prepared and the PLA protein added identically to previous experiments, but at some point after 14 days, the expression levels of the tissue could be assessed to determine if the right genes are expressed at the right levels for increased SE. In doing this, we could easily determine if PLA is working in the *cacao* system.

Pilot-Scale Production of PLA proteins

If PLA proteins are proven to be useful in the *cacao* system, there would be a need to then scale up the production and purification of the proteins. First, the PLA genes of interest

would need to be cloned in *E. coli* expression vectors without the polyhistidine tag. After successful cloning *E. coli* would then need to be grown in a bioreactor to get higher yields of protein than in the shake flask system currently utilized for lab scale production. There have been many papers on high-density growth of *E. coli*, and media for fed-batch growth in a bioreactor would be chosen from one of these papers. Most bioreactors required to be operated in a fed-batch manner due to the growth inhibition from high levels of nutrients. Some substrates that become inhibitors at high levels are glucose (50 g/L), ammonium (3 g/L), iron (1.15 g/L), magnesium (8.7 g/L), phosphorus (10 g/L), and zinc (0.038 g/L). (Shiloach, 2005) Therefore, the medium must start with concentrations below these values, and the substrates must be gradually added to the culture. Another substrate that is inhibitory is acetate. As glucose is consumed by *E. coli*, it produces acetate if there is an excess of a carbon source, such as glucose. To keep this acetate production to a minimum, the *E. coli* is kept carbon-limited with a feed containing glucose/glycerol (glycerol takes more time for *E. coli* to metabolize, thus maintaining carbon-limitation).

Once protein is produced via bioreactor, it must be purified. The cells can be lysed via sonication or a microfluidizer. The lysate can be put through a centrifuge to clarify the solution, and the supernatant can be run through purification steps. One of the most common protein purification processes is tangential flow filtration (TFF), sometimes called crossflow filtration (CFF). A diagram of a TFF membrane is shown below in Figure 14.

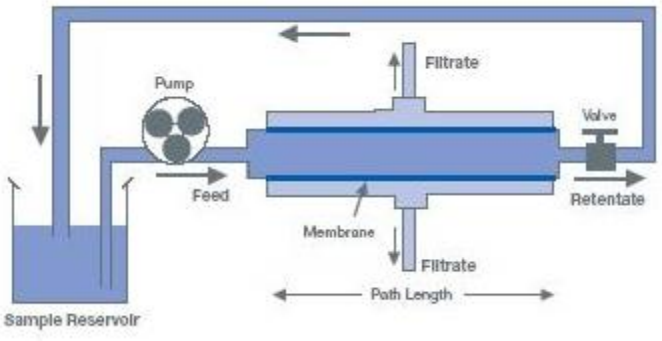

Figure 14 - Flow path through a simple TFF device. (Taken from Schwartz, Pall Corp., 2015)

The main idea behind TFF is that the flow path is normal to the membrane. The does two things: 1) it provides turbulence near the membrane so that there is effective separation of particles based on their sizes, and 2) it limits the buildup of a cake layer, allowing for more flux of material and therefore shorter run time and longer lifetime of the membranes. PLA proteins are around 11.7 kDa without a polyhistidine tag, and there are many larger proteins native to *E. coli* that are produced and need to be filtered out. Therefore, a 30 kDa MWCO membrane was selected as a first step of purification. The PLA proteins will permeate through the membrane, while the native proteins will be kept in the retentate. An SDS gel with *E. coli* native proteins (~20-180 kDa) and PLA protein (~12 kDa, dark black band), is shown in Figure 15.

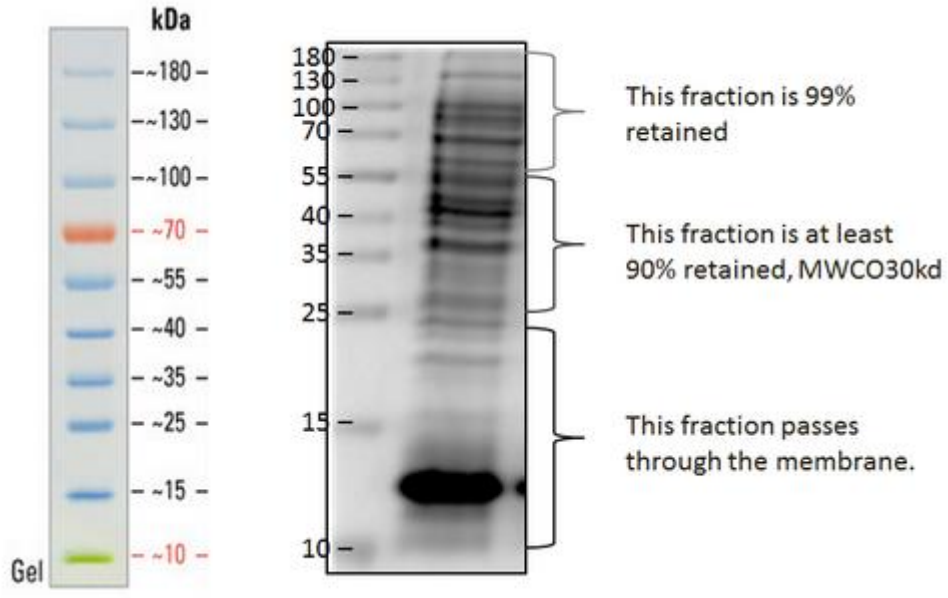

Figure 15 - SDS gel with corresponding size ladder showing large *E. coli* native proteins (~20-180 kDa) and PLA (~12 kDa). The largest proteins are retained by the membrane very effectively (>55 kDa has 99% retention), while PLA passes through the membrane easily. (Image provided by Tina Lai)

After this first step, a second step is needed to then diafiltrate and concentrate the proteins. A smaller membrane must therefore be used. To retain an acceptable amount of protein (~95%), a membrane must be selected that has pores that are 3-5 times smaller than our PLA proteins. Therefore, a 2-4 kDa MWCO membrane would be acceptable. After the buffer is replaced and the proteins concentrated, the protein must be able to be stored. Ideally, the protein would be freeze-dried to maintain stability while it is stored, but freeze-drying has not been tested with PLA proteins before, so it is unknown whether or not the protein would still be functional after lyophilization.

If PLA is proven to be successful in enhancing somatic embryogenesis in *cacao* tissue, it could be a game-changer in the world of plant propagation. Improving the efficiency of plant propagation could be as easy as sprinkling powder into your media (given that PLA can indeed

be stably lyophilized). More work definitely needs to be done regarding the effectiveness and scale-up of PLA proteins, but there are great rewards to reap if it is shown to be a viable option for SE improvement.

Appendix A

Protein Sequences of GhPLA1, TcPLA1, and TcPLA2

GhPLA1

MAAKAFSRSITPLVLLFIFLSFAQGKEIMVGGKTGAWKIPSSESDSLNKWAEKARFQIGD
SLVWKYDGGKDSVLQVSKEDYTSCNTSNPIAEYKDGNTKVKLEKSGPYFFMSGAKGHC
EQGQKMIVVVMSQKHRYIGISPAPSPVDFEGPAVAPTSGVAGLKAGLLVTVGVLGLF

TcPLA1

MAAFARTLTSSLLLLFLFLNYIEAKEILVGGKTDAWKIPSSESDSLNKWAENSRFRIGDSL
AWKYDGSKDSVLQVTKEDYASCNTSNPTAEYKDGNTKVKLEKSGPFYFISGAKGHCEQ
GQKLVVVVLSSRHKYTGISPAPSPAEIEGPAIAPTSNAAGLKAGFLVTLGVLVLGLF

TcPLA2

MASLGTLVRLLAIGLVLFSFCEAKERLIGGSEDAWKIPVNSSDSLNQWAGKTRFKVGDF
LIWKYDGKVDSVLQVTKEDYESCNTSKPIKEYKDGHTKVELDKSGPFYFISGADGHCQK
GQKLVIVVMSEKHWDHPDSPTPTAPAPAPAKNSASLIKPLRDGFLFLGLASFEVAFAFF

Appendix B

E5B and MS20 Media Recipes

Table 3 - Recipes for MS20 and E5 media, as specified for SE in *Arabidopsis*. (Gaj, 2011)

Component	MS Medium (mg/L)	E5 Medium (mg/L)
Macronutrients		
NH_4NO_3	1650	-
KNO_3	1900	3000
$MgSO_4•7H_2O$	370	500
KH_2PO_4	170	-
$NaH_2PO_4•2H_2O$	-	150
$(NH_4)_2SO_4$	-	134
$CaCl_2$	440	150
Micronutrients		
KI	0.83	0.75
H_3BO_3	6.2	3.0
$MnSO_4•4H_2O$	22.3	13.2
$ZnSO_4•7H_2O$	8.6	2.0
$CuSO_4•5H_2O$	0.025	0.025
$Na_2MoO_4•2H_2O$	0.25	0.25
$CoCl_2•6H_2O$	0.025	0.025
$FeSO_4•7H_2O$	27.8	27.8
$Na_2EDTA•2H_2O$	37.3	37.3
Organics		
Myo-inositol	100.0	100.0
Nicotinic acid	0.5	1.0
Pyridoxine-HCl	0.5	1.0
Thiamine-HCl	0.1	10.0
Glycine	2.0	-
2,4-D		5uM
Sucrose	20 g/L	20 g/L
Agar	3.5 g/L	8.0 g/L
pH	5.8	5.5

Appendix C

Finding a *T. cacao* Ortholog to GhPLA1

The analysis below was performed by Tina Lai for her master's thesis on the effect of PLA1 on *cacao* somatic embryogenesis.

The *cacao* genome has been sequenced, so the GhPLA1 gene sequence can be compared to the native genes of *cacao*. To find a *cacao* orthologue of GhPLA1, a tBLASTn analysis was performed in the CocoaGen DB. The top hits are shown in Table 4 below.

Table 4 - Top tBLASTn results for GhPLA1 in *cacao* genome database. (Provided by Tina Lai)

Query	Expect	% Identity	Locus ID	Chr.	Start	End
Query_1	4E-80	78.47	Tc09_g001400	Tc09	755190	756058
Query_1	4E-69	68.03	Tc09_g002630	Tc09	1473302	1474300
Query_1	6E-54	60.31	Tc09_g002640	Tc09	1480768	1481775
Query_1	2E-33	48.54	Tc02_g016390	Tc02	9837535	9839444
Query_1	3E-28	49.45	Tc07_g004760	Tc07	2684030	2685102

These top hits were then analyzed in a neighbor-joining phylogenetic tree using MEGA5 program to see which *cacao* gene hit was phylogenetically closest to GhPLA1. Also included in the phylogenetic analysis was GhPLA1 and some of its orthologues, including At2g25060 (*Arabidopsis*), Pn14 (Morning glory) (Yoshizaki, 2000), OsENODL1 (Rice) (Mashiguchi, 2004),

MtENOD16 (M. trunculata) (Greene, 1998), GmN315 (Soybean) (Kouchi and Hata, 1993), and NtEPc (Tobacco) (Kyo, 2000). Figure 16 shows the analysis results.

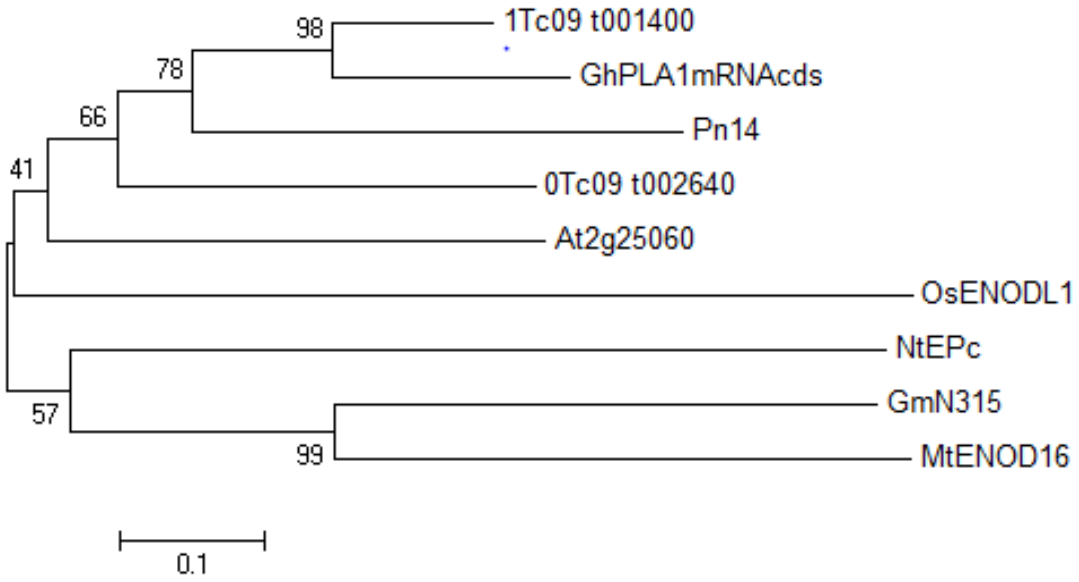

Figure 16- Phylogenetic analysis of GhPLA1, top *cacao* gene hits, and other similar gene orthologs. (Provided by Tina Lai)

1Tc09t001400 and 0Tc09t002640 are the phylogenetically closest hits to GhPLA1. A higher number to the left of the bracket indicates a more similar gene. 1Tc09t001400 was named TcPLA1, and 0Tc09t002640 was named TcPLA2.

The domains of each GhPLA1, TcPLA1, and TcPLA2 were then aligned, and the results are shown in Figure 17.

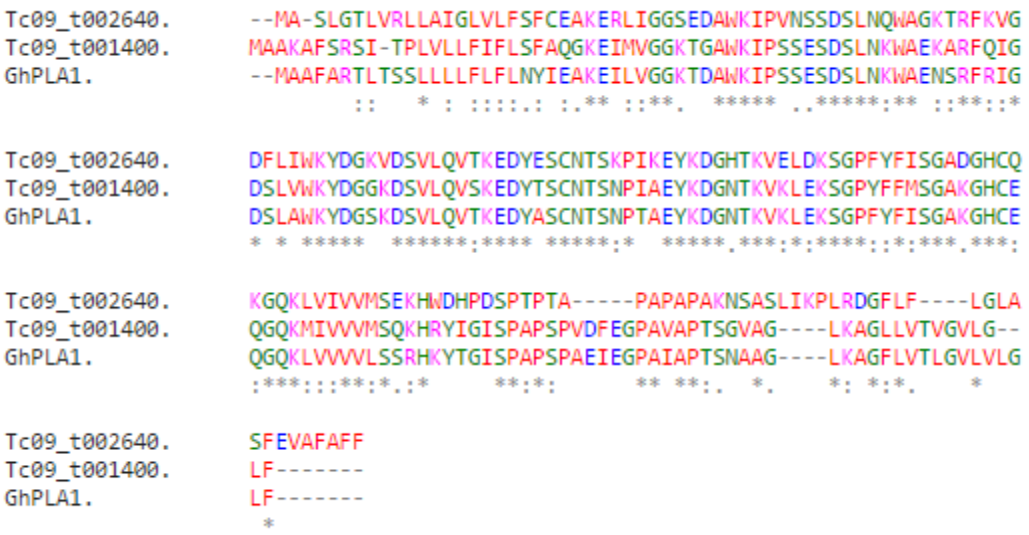

Figure 17 - Alignment of TcPLA2 (top), TcPLA1 (middle), and GhPLA1. "*" means all three sequences have the same amino acid, ":" means all three have amino acids with similar properties, and "." means two sequences have amino acids with similar properties. Each amino acid is color-coded for its properties.

The three protein sequences are very similar, especially in the middle region. TcPLA1 has a 55.21% identity to GhPLA1, and TcPLA2 has a 60.24% identity to GhPLA1. For a protein, ~50% identity is usually a good ballpark for an ortholog, so these proteins make sense to be *cacao's* orthologues to cotton's PLA1.

These two proteins, TcPLA1 and TcPLA2, will be cloned into *E. coli* and tested on *Arabidopsis* to see if they enhance somatic embryogenesis like GhPLA1 did in cotton.

Cloning of TcPLA1, TcPLA2, and GhPLA1

The PLA proteins do not require post-translational modification in order to be effective, so they may be expressed in bacteria. In this case, *E. coli* was used because of the ease of growth, high protein production, and ease of genetic engineering.

The Poon Lab from Latrobe University in Sydney, Australia sent the GhPLA1 gene in a pDEST17 vector in the form of plasmid DNA. The vector was electroporated into BL21 *E. coli* cells. Transformed cells were selected for with 50 mg/L ampicillin, and were verified via colony PCR and sequencing.

TcPLA1 and TcPLA2 were amplified from *Theobroma cacao* tissue via PCR. Each set of primers was designed to have Nde1 and Xho1 restriction enzyme cut sites for cloning into the pET14b vector (Invitrogen). Both genes were first cloned into a pcr2.1 TOPO sequencing vector utilizing TA cloning (Invitrogen) and electroporated into TOP10 E.coli cells. After sequence verification, plasmid isolation was performed on positive colonies using a Qiagen miniprep kit, and the gene was cloned into the pET14b vector, containing a 6-histidine tag. The plasmids were transformed into BL21 *E. coli* cells, selected for with 50 mg/L ampicillin, and used for protein expression. (All cloning was done by Tina Lai.)

REFERENCES

Acebo-Guerrero, Y.; Hernandez-Rodriguez, A.; Heydruch-Perez, M.; El Jaziri, M.; Hernandez-Lauzardo, A. N.; Management of black pod rot in *cacao* (*Theobroma cacao* L.): a review. *Fruits* 2012, 67: 41-48.

Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu C-M, van Lammeren A a M, Miki BL a, Custers JBM, van Lookeren Campagne MM: Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 2002, 14:1737–49.

Florez, S.L.; Erwin, R.L.; Maximova, S.N.; Guiltinan, M.J.; Curtis, W.R.: Enhanced somatic embryogenesis in *Theobroma cacao* using the homologous BABYBOOM transcription factor. *BMC Plant Biol* 2015.

Gaj, M.D.; Somatic Embryogenesis and Plant Regeneration in the Culture of *Arabidopsis thaliana* (L.) Heynh. Immature Zygotic Embryos. *Plant Embryo Culture: Methods and Protocols.* 2011, 710: 257-265.

Greene, E.A.; Erard, M.; Dedieu, A.; Barker, D.G.: MtENOD16 and 20 are members of a family of phytocyanin-related early nodulins. *Plant. Mol. Biol.* 1998, 36: 775–783.

Guiltinan, MJ: *Cacao*. In *Biotechnology in Agriculture and Forestry : Transgenic Crops V*. *Volume 60*. Edited by Pau EC, Michael RD. Berlin, Heidelberg: Berlin Heidelbelg: Springer-Verlag; 2007:497–518.

Guiltinan, Mark J., Zhijian Li, Abdoulaye Traore, and Siela Maximova. Methods and Tissue Culture Media for Inducing Somatic Embryogenesis, Agrobacterium-medicated Transformation and Efficient Regeneration of *Cacao* Plants. The Penn State Research Foundation, assignee. Patent US6197587 B1. 6 Mar. 2001. Print.

Heck, G.R.; Perry, S.E.; Nichols, K.W.; Fernandez, D.E.: AGL15, a MADS domain protein expressed in developing embryos. *Plant Cell* 1995, 7:1271-82.

Kim, G.T.; Tsukaya, H.; Uchimiya, H.: The CURLY LEAF gene control both division and elongation of cells during the expansion of the leaf blade in *Arabidopsis thaliana*. *Planta* 1998, 206: 175-83.

Kouchi, H.; Hata, S.: Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol. Gen. Genet.* 1993, 238: 106–119.

Kreuger, M.; van Holst, G. J.: Arabinogalactan proteins are essential in somatic embryogenesis of Daucus carota L. *Planta* 1993, 189: 243-248.

Kyo M, Miyatake H, Mamezuka K, Amagata K: Cloning of cDNA encoding NtEPc, a marker protein for the embryogenic dedifferentiation of immature tobacco pollen grains cultured in vitro. *Plant Cell Physiol.* 2000, 41: 129–137.

Li Z, Traore A, Maximova S, Guiltinan MJ: Somatic embryogenesis and plant regeneration from floral explants of *cacao* (*Theobroma cacao* L.) using thidiazuron. *Vitr Cell Dev Biol - Plant* 1998, 34:293–299.

Lotan T, Ohto M, Yee KM, West M, Lo R, Kwong RW, et al. *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell.* 1998, 93: 1195–205.

Mashiguchi K, Asami T, Suzuki Y: Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes in *Arabidopsis*. *Biosci. Biotechnol. Biochem*. 2009, 73: 2452–2459.

Meinke, D.; Muralla, R.; Sweeney, C.; Dickerman, A.: Identifying essential genes in *Arabidopsis thaliana. Trends Plant Sci* 2008, 13: 483-491.

Ogas, J.; Kaufmann, S.; Henderson, J.; Somerville, C.: PICKLE is a CHD3 chromatic-remodeling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 1999, 96: 13839-44.

Parcy, F.; Valon, C.; Kohara, A.; Misera, S.; Giraudat, J.: The ABSCISIC ACID-INSENSITIVE, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of *Arabidopsis* seed development. *J. Plant Cell* 1997, 9: 1265-77.

Passarinho P, Ketelaar T, Xing M, van Arkel J, Maliepaard C, Hendriks MW, Joosen R, Lammers M, Herdies L, den Boer B, van der Geest L, Boutilier K: BABY BOOM target genes provide diverse entry points into cell proliferation and cell growth pathways. *Plant Mol Biol* 2008, 68: 225–37.

Pierce Biotechnology: Tech Tip #43: Protein stability and storage. Thermo Fisher Scientific Inc. 2009 https://tools.lifetechnologies.com/content/sfs/brochures/TR0043-Protein-storage.pdf. Accessed April 7, 2015.

Poon, S.; Heath, R. L.; Clarke, A. E.; A chimeric arabinogalactan protein promotes somatic embryogenesis in cotton cell culture. *Plant Physiol.* 2012, 160: 684-95.

SAFC: Contract Manufacturing Services. Production of pharmaceutical compounds through microbial fermentation. Sigma-Aldrich Co. LLC. http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/SAFC/General_Information/1/production-of-pharmaceutical-compounds.pdf. Accessed April 7, 2015.

Schwartz, L.; Introduction to Tangential Flow Filtration for Laboratory and Process Development Applications. Pall Corporation Website.

http://www.pall.com/main/laboratory/literature-library-details.page?id=34212. Accessed April 3, 2015.

Shiloach, J.; Fass, R.: Growing *E. coli* to high cell density--A historical perspective on method development. *Biotech. Adv.* 2005, 23: 345-357.

Steward F, Mapes M, Mears K: Growth and Organized Development of Cultured Cells. II. Organization in Cultures Grown from Freely Suspended Cells. *Am J Bot* 1958, 45:705–708.

Studier, F. W.; Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 2005, 41: 207-234.

Thakare D, Tang W, Hill K, Perry SE. The MADS-domain transcriptional regulator AGAMOUS-LIKE15 promotes somatic embryo development in *Arabidopsis* and soybean. *Plant Physiol*. 2008, 146: 1663–72.

Tsuwamoto, R.; Yokoi, S.; Takahata, Y.: *Arabidopsis* EMBRYOMAKER encoding an AP2 domain transcription factor plays a key role in developmental change from vegetative to embryonic phase. *Plant Mol Biol*. 2010, 73: 481–92.

White, J.; The Business of Recombinant Protein Production. Drug Discovery & Development Magazine. http://www.dddmag.com/articles/2010/10/business-recombinant-protein-production. Accessed April 7, 2015.

Wójcikowska, B.; Jaskóła, K.; Gąsiorek, P.; Meus, M.; Nowak, K.; Gaj, M.D.: LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of *Arabidopsis*, via YUCCA-mediated auxin biosynthesis. *Planta*. 2013, 238: 425–40.

Yoshizaki M, Furumoto T, Hata S, Shinozaki M, Izui K: Characterization of a novel gene encoding a phytocyanin-related protein in morning glory (Pharbitis nil). *Biochem. Biophys. Res. Commun.* 2000, 268: 466–470.

Zeng F, Zhang X, Zhu L, Tu L, Guo X, Nie Y: Isolation and characterization of genes associated to cotton somatic embryogenesis by suppression subtractive hybridization and macroarray. *Plant Mol Biol* 2006, 60: 167–83.

ACADEMIC VITA

Rachel Erwin
racherwin@gmail.com

Education

The Pennsylvania State University
Schreyer Honors College
B.S. Chemical Engineering
The Bioprocess and Biomolecular Engineering Option
Minor in Biochemistry and Molecular Biology

Honors and Awards

The Schreyer Honors College Academic Excellence Scholarship
The Leonhard Engineering Honors Program Scholarship
College of Engineering Academic Scholarship
Dean's List Recipient for all completed semesters

Research Experience

Undergraduate Research Assistant
Dr. Wayne Curt University Park, PA December 2011-Present

- Honors thesis work, started Fall 2013
  - Propagation of plants via somatic embryo formation: A model system utilizing the manipulation of transcription factors in *Arabidopsis thaliana*
- Experience with bench-scale bioreactors, aseptic tissue culture, bacterial cell culture, molecular biology, and protein purification

NSF-Sponsored REU Participant
Dr. Wayne Curtis University Park, PA May 2013-August 2013

- Used various molecular biology techniques to quantitate the effectiveness of a gene used to induce somatic embryogenesis in *Theobroma cacao*

Work Experience

Pharmaceutical Delivery Systems Research Intern
West Pharmaceutical Services Exton, PA May 2014-August 2014

- Developed a method to study adsorption and stability of protein drugs in varying container types. Generated a formal lab report documenting the new method.
  - Experience with HPLC, UV/Vis Spectrophotometry, and Fluorometry
- Collaborated with another intern to carry out performance testing on medical devices and statistically analyze the results using Minitab. A formal summary report was generated and put into a company-wide knowledge database.

Publications

Florez, S.; Erwin, R.; Maximova, S.; Guiltinan, M.; Curtis, W. "Enhanced somatic embryogenesis in *Theobroma cacao* using the homologous BABYBOOM transcription factor." (Accepted - pending publication)

Association Memberships

Tau Beta Pi, The Engineering Honor Society, PA Beta Chapter
Corresponding Secretary 2013-2015
Society of Women Engineers, Penn State Chapter