4-19-2012 12:00 AM

Fatty Acid ω-Hydroxylases in Soybean

Jessica Koteles, *The University of Western Ontario*

Supervisor: Dr. Mark Bernards, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biology

FATTY ACID ω-HYDROXYLASES IN SOYBEAN

(Thesis format: Monograph)

by

Jessica Koteles

Graduate Program in Biology and
Environment and Sustainability

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
Dr. Mark Bernards	Dr. Sheila Macfie
Supervisory Committee	Dr. Sangeeta Dhaubhadel
Dr. Susanne Kohalmi	Dr. Shiva Singh
Dr. Mark Gijzen	

The thesis by

Jessica Koteles

entitled:

Fatty Acid ω-Hydroxylases in Soybean

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

Date_____ _____
Chair of the Thesis Examination Board

ABSTRACT

Soybean (*Glycine max* (L.) Merr.) is an important and widely cultivated crop. A substantial cause of soybean yield loss is due to root rot caused by the pathogen *Phytophthora sojae*. There is an established correlation between preformed soybean root aliphatic suberin and high levels of partial resistance to *P. sojae*. Since fatty acid ω-hydroxylase (FAωH), which introduces a second functional group to fatty acids, is a critical enzyme in suberin biosynthesis, characterization of soybean FAωHs was undertaken. Six putative soybean FAωHs (Gm FAωHs) were identified using a sequence-based *in silico* approach. Gene-specific primers were developed and relative expression was analyzed across various tissues using RT-PCR. While differential expression patterns were detected for Gm FAωH-1, Gm FAωH-2, Gm FAωH-3 and Gm FAωH-4, neither Gm FAωH-5 nor Gm FAωH-6 were detected in any of the tissues examined. In comparison with RNA-Seq and EST *in silico* data, RT-PCR profiles revealed contrasting expression patterns.

Keywords: Soybean, suberin, *Phytophthora sojae*, resistance, fatty acid ω-hydroxylase

ACKNOWLEDGEMENTS

I am extremely thankful to my supervisor, Dr. Mark Bernards, for his guidance and support throughout this project. His constant encouragement made completing this thesis possible.

Thank you to Dr. Susanne Kohalmi and Dr. Mark Gijzen for sharing their advice and wisdom. To Susanne who endured long meetings discussing molecular biology and to Mark who provided both *in silico* guidance and the soybean seed used in this research.

To the members of my lab and fellow graduate students on the 4th floor of NCB, thank you for your continual friendship, support, coffee breaks, advice and being there to listen whenever I had something to celebrate or complain about.

I am deeply thankful to my family for encouraging me throughout my entire academic career. To my parents whose unconditional love and perspective always made me believe I could succeed.

Finally, thank you to Albert who was there for me every step of the way. His daily support helped me get past seemingly impassible obstacles.

TABLE OF CONTENTS

ABSTRACT iii

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS v

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF APPENDICES x

LIST OF ABBREVIATIONS xi

CHAPTER 1: INTRODUCTION 1

1.1 Soybean 1

1.2 *Phytophthora sojae*: basic life cycle and infection process 1

1.3 Plant defenses against pathogens 2

1.4 Cultivar development 4

1.5 Suberin 6

1.5.1 The suberin poly(aliphatic) domain (SPAD) 7

1.5.2 Fatty acid ω-hydroxylases 9

1.5 Rationale of approach 12

1.6 Objective of thesis 12

CHAPTER 2: MATERIALS AND METHODS 14

2.1 *In silico* analysis 14

2.2 RNA extraction and DNase digestion 15

2.3 DNase digestion of RNA 16

2.4 Generation of cDNA from isolated RNA 17

2.5 Genomic DNA isolation for PCR optimization 17

2.6 Primer design and optimization 18

2.7 PCR purification and sequencing 20

2.8 Semi quantitative RT-PCR of putative *Gm FAωH*s in soybean tissues 21

CHAPTER 3: RESULTS AND DISCUSSION 22

3.1 *In silico* analysis of putative *Gm FAωH*s 22

3.2 *In silico* expression profiles 29

3.3 Designing gene-specific primers 35

3.4 Checking and optimizing primers 38

3.5 RT-PCR analysis of tissue-specific gene expression 45

3.6 A role for *FAωH*s in suberin formation 51

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS 54

BIBLIOGRAPHY 56

CURRICULUM VITAE 87

LIST OF TABLES

Table 1: Primer sets used for PCR amplification of potential *Gm FAωH* genes 19

Table 2: Pairwise nucleotide comparisons of putative *Gm FAωH*s 23

Table 3: Pairwise amino acid comparisons of putative Gm FAωHs.....27

Table 4: RNA-Seq expression data for the six putative *Gm FAωH*s 31

Table 5: EST data for five of the putative *Gm FAωH*s 34

Table 6: Characteristics of the *Gm FAωH* primers designed for RT-PCR 37

Table 7: Tissue expression comparison of the putative *Gm FAωH*s from RT-PCR, EST and RNA-Seq data.....52

LIST OF FIGURES

Figure 1: The relationship between preformed whole-root suberin and partial resistance of soybean to *P. sojae*.....5

Figure 2: The proposed structure of potato suberin.....8

Figure 3: The ω-hydroxylation of fatty acids 11

Figure 4: Homology trees of the six putative Gm FAωHs 25

Figure 5: Amino acid sequence alignment of the related proteins CYP86A1, CYP86A33, CYP86A32, CYP86A9 and the six putative Gm FAωHs.....28

Figure 6: Phylogenetic tree of the predicted full length amino acid sequence comparison between known CYP86As and CYP94As with the six putative Gm FAωHs 30

Figure 7: Nucleotide sequence alignment of the six putative *Gm FAωH*s 36

Figure 8: Optimizing $MgCl_2$ concentration for PCR amplification of soybean *FAωH* genes 39

Figure 9: Optimizing annealing temperature for PCR amplification of soybean *FAωH* genes 40

Figure 10: Optimizing cycle number for RT-PCR amplification of soybean *FAωH* genes 42

Figure 11: PCR products of each putative *Gm FAωH* generated from genomic DNA 44

Figure 12: Analysis of various tissue samples contamination 46

Figure 13: RT-PCR tissue expression of six putative *Gm FAωH*s using 'A' samples.....47

Figure 14: RT-PCR tissue expression of six putative *Gm FAωH*s using 'B' samples.....49

LIST OF APPENDICES

Appendix I: Characteristics of soybean pods and seeds used to designate M2 and L1 stages 63

Appendix II: Extended table of RNA-Seq data.....64

Appendix III: Contig maps of assembled soybean ESTs 65

Appendix IV: Sequence assembly of five of the putative *Gm FAωH*s and their corresponding sequencing results 68

Appendix V: Permission for copyrighted material.....86

LIST OF ABBREVIATIONS

BLAST - Basic Local Alignment Search Tool

CYP86A - Cytochrome P450 Subfamily 86, type A

EST - Expressed Sequence Tag

Gm FAωH - *Glycine max* Fatty Acid Omega Hydroxylase

NCBI - National Center for Biotechnology Information

PCR - Polymerase Chain Reaction

RT-PCR - Reverse Transcription Polymerase Chain Reaction

SPAD - Suberin Poly(Aliphatic) Domain

SPPD - Suberin Poly(Phenolic) Domain

SRS - Substrate Recognition Site

T_m - Melting Temperature

UTR - Untranslated Region

CHAPTER 1: INTRODUCTION

1.1 Soybean

Soybean (*Glycine max* [L.] Merr.), a legume, is one of the most widely harvested crops in the world. The high protein, high oil beans are used for human consumption, animal rations and edible oils as well as many industrial products. Soybean production has experienced an eight-fold increase in Canada since 1976 (Dorff, 2009). In 2006, 1.2 million hectares of soybean were grown in Canada, 73% of which was in Ontario (Dorff, 2009). A substantial cause of soybean yield loss worldwide is root rot, caused by the pathogen *Phytophthora sojae* (Kaufmann and Gerdemann) and significant effort has been expended in improving soybean resistance to this devastating pathogen.

1.2 *Phytophthora sojae*: basic life cycle and infection process

Phytophthora sojae is a soil-borne organism that causes root rot in soybean and is capable of causing substantial economic loss (Tooley and Grau, 1984; Dorrance et al., 2003). *Phytophthora sojae* survives as oospores in soil and plant debris. Oospores can survive for a number of years in a dormant state, and can withstand freezing and long periods of cold temperatures. Oospores of *P. sojae* have endogenous dormancy and not all will germinate at the same time even when conditions are highly favorable (Tyler, 2007). Oospores germinate to form mycelia under high soil moisture conditions, and the mycelia then produce sporangia and motile zoospores under continuous or intermittent saturated soil conditions (Tyler, 2007). Zoospores are attracted to soybean root exudates, which

contain isoflavones (specifically daidzein and genestein) and possibly other compounds secreted from the roots (Zentmyer, 1961; Morris and Ward, 1992; Tyler et al., 1996; Morris et al., 1998). Once they reach a root, the zoospores lose their flagella, produce a cell wall (encyst) and germinate (Dorrance et al., 2007). Germ tubes penetrate into the root by growing through the middle lamellae of the epidermal cells (Tyler, 2007). Continued growth of hyphae occurs through cortical air spaces with haustoria formation to facilitate feeding off nearby cells (biotrophic phase). Eventually the endodermis is breached, allowing the hyphae to colonize the xylem (Enkerli et al., 1997). Herein the pathogen switches to a necrotrophic growth habit and aggressively attacks the plant. *Phytophthora sojae* can spread quickly throughout the plant, causing large, water-soaked, necrotic lesions containing dead cells. Eventually the root and stem tissues become completely colonized, and collapse. New oospores are produced within the dead tissue, and these are returned to the soil as the host plant dies and rots.

1.3 Plant defenses against pathogens

Plants possess multiple types of defenses against pathogens including both structural and chemical defenses. Structural defenses include waxy epidermal cuticles of aboveground tissues, suberized bark and suberin depositions in modified cell walls of underground dermal layers. Chemical defenses mainly consist of secondary metabolites including terpenoids, phenolics (e.g., isoflavonoids) and alkaloids (Freeman and Beattie, 2008). These chemical defenses have high energy costs and nutrient requirements that are associated with production and maintenance (Freeman and Beattie, 2008). Additionally,

defenses can be preformed (innate) or inducible. Preformed defenses can be physical barriers (e.g., cutin, suberin) or secondary metabolites with antimicrobial, antifeedant or insecticidal properties (Arnason and Bernards, 2010). As an example, phytoanticipins are antimicrobial compounds that accumulate in plants prior to an attack (Arnason and Bernards, 2010). Induced defenses can involve the *de novo* synthesis of secondary metabolites, activation of preformed metabolites or reinforcement of existing or deposition of new physical barriers (including new cell wall deposits, lignin and suberin) (Arnason and Bernards, 2010). One example is the induced formation of phytoalexins, which are antimicrobial compounds produced only after infection (Hammerschmidt, 1999). Preformed and induced structural and chemical defenses when combined, comprise the resistance of a plant to a potential pathogen.

In soybean there are two main mechanisms through which resistance to *P. sojae* is manifested: race-specific and partial resistance (also called innate resistance, quantitative resistance, tolerance or field resistance). Race-specific resistance is a gene-for-gene relationship and fourteen different single resistance genes for resistance against *P. sojae* (Rps genes) have been described (Grau et al., 2004). Partial resistance to *P. sojae* in soybean is expressed and inherited quantitatively, effective against all the races of the pathogen, and is exemplified by reduced colonization and slower lesion expansion (Schmitthenner, 1985; Tyler, 2002; Burnham et al., 2003). In general, partial resistance is a multi-gene trait that is believed to involve numerous defense components (e.g., pre-formed or induced physical barriers, phytoanticipins, phytoalexins, etc.), while gene-for-gene resistance involves pre-defense recognition between host and pathogen.

1.4 Cultivar development

Soybean root rot has traditionally been managed by incorporating single gene resistance against *P. sojae* into commercial cultivars. However, the pathogen has the ability to develop new pathotypes that overcome these Rps genes, rendering them ineffective. This is known as pathotype (or race) shift, and generally, new virulent races of the pathogen evolve within an 8 to 15 year window (Olah and Schmitthenner, 1985; Dorrance et al., 2003). The development of new virulent pathogen races in response to race-specific resistant cultivars has led to recorded failures in disease control (Schmitthenner, 1985; Kaitany et al., 2001). Consequently, current disease management recommendations emphasize the use of partial resistance as a more durable strategy to control *P. sojae* (St. Martin et al., 1994; Burnham et al., 2003; Dorrance et al., 2003). Finding additional sources of partial resistance and incorporating this resistance into commercial cultivars will be essential for effective disease management especially as the number and complexity of pathotypes of *P. sojae* continues to increase in areas of soybean production.

Increased partial resistance may be possible through the enhancement of pre-existing plant defenses. For example, Thomas et al. (2007) established that a strong correlation exists between preformed soybean root suberin (especially the poly[aliphatic] domain [SPAD]) and high levels of partial resistance to *P. sojae*, (Figure 1). Two genotypes were studied in detail: 'Conrad,' which has a high degree of partial resistance to *P. sojae*, and line OX760-6, which is susceptible to this pathogen. Conrad had almost twice as much preformed root suberin components as OX760-6, in both the epidermal and

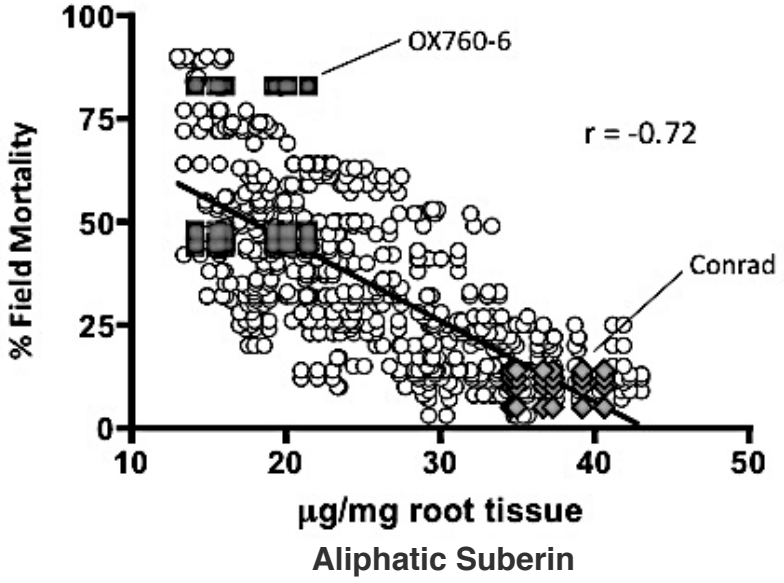


Figure 1: The relationship between preformed whole-root suberin and partial resistance of soybean to *P. sojae* (modified from Thomas et al., 2007). Data obtained for aliphatic suberin components were plotted against plant mortality data obtained from four independent plantings of the same recombinant inbred lines at two different field sites naturally infested with *P. sojae*. The data for the parental lines 'Conrad' (diamonds) and OX760-6 (squares) are highlighted. (With permission from the American Society of Plant Biologists.)

endodermal tissues (Thomas et al., 2007). Additionally, it was subsequently shown that *P. sojae* progression through soybean roots was significantly slower in Conrad compared with OX-760-6, primarily due to growth impediment at the epidermal and endodermal layers of the root (Ranathunge et al., 2008).

These studies suggest that an additional approach to soybean crop protection, in which levels of preformed suberin are increased through breeding, could provide a strategy to reduce the extent of *P. sojae*-mediated root rot in soybean. Furthermore, if appropriate molecular markers for this trait can be developed and used for progeny screening, the development time for new cultivars with enhanced partial resistance to *P. sojae* could be reduced by up to 3 years relative to conventional selection processes (Vaino Poysa, personal communication). Thus, it is essential to identify key elements involved in suberin deposition (especially the poly[aliphatic] component) in order to help facilitate the development of molecular markers for screening.

1.5 Suberin

Suberin is a cell wall modification found in specialized cells of plants, particularly in the epidermis and endodermis of underground components (roots, stolons) (Esau, 1977). It acts as a preformed barrier to pathogens and aids in water retention. Suberin deposition is induced by wounding, where it serves to seal wounded tissue, preventing opportunistic pathogens from entering exposed tissue, as well as the loss of water and nutrients. Suberin is composed of two distinct domains: a poly(phenolic) domain (SPPD) similar to lignin and a poly(aliphatic) domain (SPAD) similar to cutin (reviewed in

Bernards, 2002). The SPAD is thought to reside between the cell wall and plasma membrane, often appearing as repeating lamellae while the SPPD is thought to be in the cell wall (Bernards, 2002) (Figure 2). Since the SPAD appears to correlate with soybean partial resistance to *P. sojae*, its structure and biosynthesis are considered in more detail below.

1.5.1 The suberin poly(aliphatic) domain (SPAD)

The structure of the polymeric matrix of the SPAD derives from the formation of a three dimensional matrix of aliphatic monomers cross-linked by ester bonds with glycerol acting as the linkage between acyl groups (reviewed in Bernards, 2002). The covalently cross-linked matrix is embedded with waxes, which impart the water retention properties of the macromolecule. The main structural components of the SPAD are 1-alkanols, ω-hydroxyalkanoic acids, α,ω-dioic acids, hydroxy-substituted octadecanoates, and glycerol (Kolattukudy, 1980, 1984; Holloway, 1983; Graça and Pereira,1997). Waxes include fatty acids, 1-alkanols and alkanes, which are often very long chains (C22-28). The SPAD is thought to be assembled and deposited on the inner surface of primary walls, and glycerol is hypothesized to covalently link it to the SPPD (Schmutz et al. 1994; Graça and Pereira 2000). The SPPD is composed mainly of hydroxycinnamic acids, hydroxycinnamoyl amides, and hydroxycinnamoyl alcohols (Bernards et al, 1995; Negrel et al., 1996; Kolattukudy, 1980, 1984; Zeier and Schreiber, 1997, 1998; Bernards and Lewis, 1998; Zeier et al, 1999; Bernards, 2002).

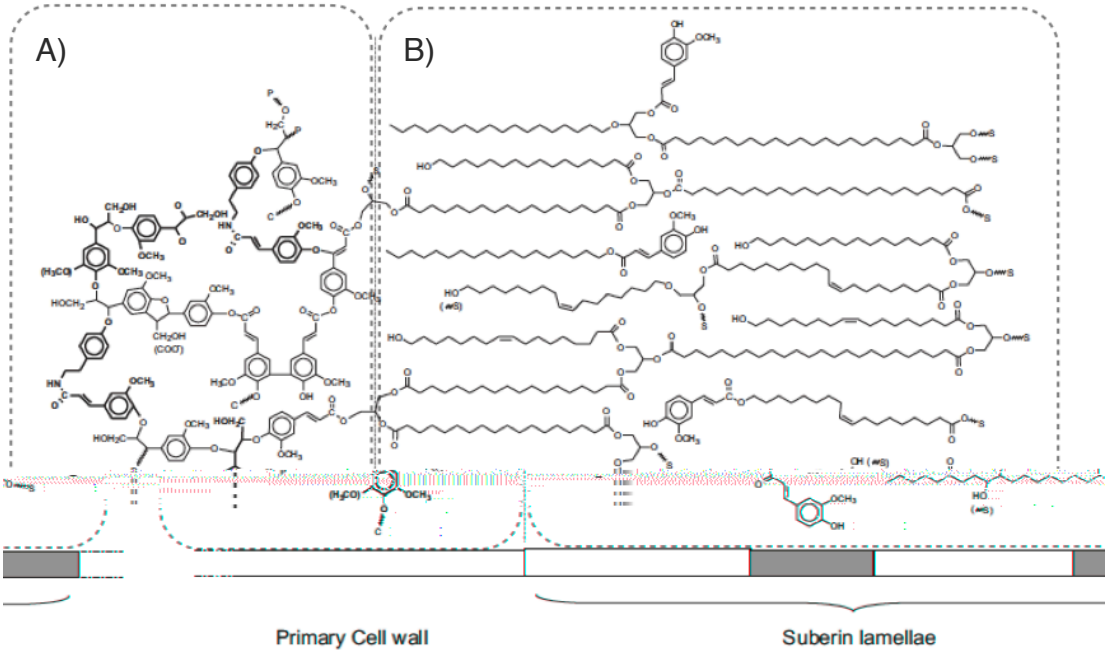


Figure 2: The proposed structure of potato suberin. A) The poly(phenolic) domain (SPPD) is a part of the primary cell wall and is covalently attached to the cell wall. B) The poly(aliphatic) domain (SPAD) extends inwards from the primary cell wall in repeating dark and light bands referred to as the suberin lamellae. P = phenolic, C = carbohydrate, S = aliphatic suberin (Adapted from Bernards, 2002).

The formation of the SPAD involves the modification of fatty acids, including chain elongation to form very long chain fatty acids, reduction to primary alcohols, or oxidation to ω-hydroxy fatty acids and α,ω-dioic acids via fatty acid ω-hydroxylation and subsequent oxidation (Yang and Bernards, 2006).

Fatty acid biosynthesis occurs in the plastid with desaturation occurring before export into the cytosol (Harwood, 1988). In potato, this step is thought to represent a "decision making" step in the allocation of carbon to suberin aliphatic monomers because saturated fatty acids are further generally elongated into very long chain fatty acids at the endoplasmic reticulum before modification to primary alcohols or alkanes (in the case of suberin waxes), while desaturated fatty acids are almost exclusively further oxidized to ω-hydroxy fatty acids and α,ω-dioic acids via fatty acid ω-hydroxylation (Yang and Bernards, 2006). It is this ω-hydroxylation step that is critical to the structure of the SPAD because it introduces a second functional group into the main monomers, which allows them to cross-link into a polymeric matrix.

1.5.2 Fatty acid ω-hydroxylases

The enzymes involved in the oxidation of fatty acids are fatty acid ω-hydroxylases; these are cytochrome P-450 enzymes that catalyze the terminal carbon hydroxylation of fatty acids. Cytochrome P-450s are P-450 monooxygenases which are heme-dependent oxidases that utilize NADPH to reductively cleave atmospheric dioxygen to produce a functional organic substrate and a molecule of water (Schuler and Werck-Reichhart, 2003). For proper function, cytochrome P-450s require a NADPH-

cytochrome P-450 reductase (Gonzalez and Korzekwa, 1995; Schuler and Werck-Reichhart, 2003). The reductase is required as it interacts directly with the cytochrome P-450 enzyme in the transfer of one oxygen atom into a substrate while the other oxygen atom is reduced to H_2O using two electrons from NADPH (Gonzalez and Korzekwa, 1995; De Vetten et al., 1999). The enzyme class is named P-450 because the reduced carbon monoxide complex that is formed during catalysis absorbs light at 450 nm (Gonzalez and Korzekwa, 1995). Cytochrome P-450s are typically located in either the endoplasmic reticulum, the inner mitochondrial membrane, or in chloroplasts (Gonzalez and Korzekwa, 1995).

In plants, multiple cytochrome P-450 subfamilies have been identified as ω-hydroxylases: including CYP86A and CYP94A (Duan and Schuler, 2005), CYP86B (Compagnon et al., 2009) and CYP704B (Dobritsa et al., 2009; Li et al., 2010). Members of the CYP94A subfamily catalyze the initial ω-hydroxylation as well as subsequent oxidation of the ω-hydroxy fatty acid to the corresponding dioic acid (Figure 3). The CYP86A subfamily has been described in *Arabidopsis thaliana* (Duan and Schuler, 2005; Rupasinghe et al., 2007) and CYP86A1 has been shown to be a functional ω-hydroxylase involved in suberin monomer biosynthesis (Hofer et al., 2008). In *Solanum tuberosum*, CYP86A33 was shown to be involved in suberin biosynthesis through RNAi knock-down experiments, where the loss of function resulted in the alteration of suberin composition, and an impairment of the periderm's barrier to water (Serra et al., 2009). The ω-hydroxylase(s) of soybean have not been described in the literature. Due to the inherent low abundance of cytochrome P-450 enzymes and their

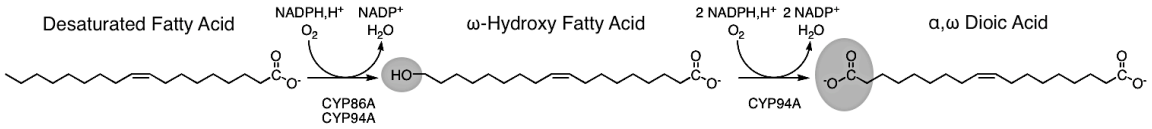


Figure 3: The ω-hydroxylation of fatty acids. The ω-hydroxylation of fatty acids by two cytochrome P-450 subfamilies: CYP86A and CYP94A. Members of the CYP86A subfamily catalyze the ω-hydroxylation of fatty acids. Members of the CYP94A subfamily catalyze the ω-hydroxylation as well as the subsequent oxidation of the terminal carbon to the corresponding dioic acid.

membrane association, traditional protein purification approaches are not feasible for their characterization. Therefore, a molecular biology approach may make the characterization of soybean ω-hydroxylases more feasible. This approach has been made more accessible through the recent completion and annotation of the soybean genome (Schmutz et al., 2010).

1.5 Rationale of approach

Since aliphatic suberin in soybean root correlates to partial resistance to *P. sojae*, suberin is an excellent target for study in the improvement of soybean innate resistance to *P. sojae*. Furthermore, ω-hydroxylases (specifically CYP86As) are vital to the composition of the aliphatic domain of suberin in other plant species (especially *Arabidopsis* and potato). Therefore, the characterization of this subfamily of genes in soybean will be essential to our understanding of suberin deposition in soybean roots. This characterization will not only aid in potential molecular marker development for breeding selection, but will allow further understanding of the suberin/resistance relationship. It will also provide a possible mechanism for altering suberin composition to study the effects on *P. sojae*'s ability to infect soybean roots.

1.6 Objective of thesis

The objective of the research in this thesis is to identify potential CYP86A genes in soybean and establish their tissue expression profiles. Since CYP86A proteins are found in very low abundance in cells and are unstable, traditional protein purification is

unreliable. Therefore, a molecular biological approach will be used. Initially, an *in silico* approach will be used to identify potential candidate CYP86A genes from the soybean genome (Schmutz et al., 2010). Unique primers will then be designed for each gene, and the expression patterns of the putative CYP86A genes will be analyzed in various soybean tissues. This PCR-mediated expression data will be compared to existing EST (various public EST databases) and RNA-Seq (soybase.org) expression data available for soybean.

CHAPTER 2: MATERIALS AND METHODS

2.1 *In silico* analysis

Sequences of characterized fatty acid ω-hydroxylases (*FAωH*s) from *Arabidopsis thaliana* (CYP86A1, accession #: NM_125276) and *Solanum tuberosum* (CYP86A33, accession #: EU293405) were used in a Basic Local Alignment Search Tool (BLAST) search of the soybean genome databases Phytozome.net and compbio.dfci.harvard.edu (Gene Index). Multiple sequence alignments of the putative genes and known CYP86As were performed in DNAMAN to show similarities. CYP86A substrate recognition sites (SRSs) were identified in accordance with Rupasinghe et al. (2007). A phylogenetic tree with 100 bootstrap trials was generated to show the putative genes relationships to each other and to known CYP86As and CYP94As. RNA-Seq and expressed sequence tag (EST) data were utilized to gain insight into the tissue expression patterns of the six putative fatty acid ω-hydroxylases. RNA-Seq data for the six genes was compiled from the atlas at Soybase.org. In collaboration with Dr. Martina Stromvick (McGill University), existing EST data was assembled and analyzed using the software Phrap (P. Green, University of Washington, Seattle, http://www.phrap.org/). Additionally, probe sets that represent the putative soybean *FAωH*s on the Affymetrix soy chip were identified via Soybase.org.

2.2 RNA extraction and DNase digestion

All steps involving handling of RNA were done using RNase-free methods. This included creating an RNase-free zone on the bench top using RNaseZap, using distinct RNase-free microcentrifuge tubes and barrier tips, baking the mortar and pestle at 250°F overnight and frequently changing gloves. In order to isolate RNA, seeds of the Conrad cultivar were grown in the University of Western Ontario greenhouse in Pro-Mix soil. Seeds and pods of the M2 and L1 stages were obtained from Agriculture and Agri-Food Canada (Southern Crop Protection and Food Research Centre, London, ON, Canada) (Appendix I). The QIAGEN RNeasy Plant Mini Kit was used to isolate RNA from 9 distinct tissues: old leaf (> 14 days post germination), young leaf (< 10 days old), stem, flower, M2 (29-35 days post anthesis) and L1 (35-43 days post anthesis) stage pods, M2 and L1 stage seeds, and root. In order to perform a biological replicate, all tissue types (except seed and pod) were collected from two different pots containing multiple soybean plants (‘A’ and ‘B’). These plants were all grown in the same environment and plants within a pot were grouped. Seed and pod tissue samples were separated into two groups after being collected in the field. Between 70 and 90 mg of fresh weight of each tissue was used. Tissue was frozen in liquid nitrogen, ground with a mortar and pestle, and transferred to a RNase-free, liquid nitrogen cooled microcentrifuge tube. Once the liquid nitrogen had evaporated, 450 µL of Buffer RLT was added, and the sample vortexed vigorously and incubated at 56°C for 3 minutes in a water bath to help disrupt the tissue. The lysate was transferred to a QIAshredder spin column and centrifuged for 2 minutes at high speed in a microcentrifuge (17,000 x g). The supernatant of the flow-through was

carefully transferred to a new microcentrifuge tube and mixed with 0.5 volumes of ethanol (96-100%). The sample was then added to an RNeasy spin column and centrifuged for 15 seconds at 8,000 x g. The flow-through was discarded and 700 µL of Buffer RW1 was added before centrifuging for 15 seconds at 8,000 x g. After discarding the flow-through, 500 µL of Buffer RPE was added and the sample centrifuged for 15 seconds at 8,000 x g. Once again the flow-through was discarded and 500 µL of Buffer RPE was added and the sample centrifuged for 2 minutes at 8,000 x g. To eliminate possible carryover of Buffer RPE, the column was centrifuged for an additional minute at full speed. The column was then placed in a clean microcentrifuge tube and the RNA was eluted by adding 30-50 µL of RNase-free H_2O, letting it sit for 5 minutes and then centrifuging for 1 minute at 8,000 x g. In tissues where the RNA yield was lower, the eluate was re-added to the column and centrifuged for an additional minute at 8,000 x g.

2.3 DNase digestion of RNA

All steps involving handling of RNA were done using RNase-free methods as previously described. Prior to generation of cDNA, the RNA samples were treated with DNase to digest contaminating DNA by following the QIAGEN RNeasy Plant Mini Kit protocol. Briefly, each RNA sample was mixed with 10 µL of Buffer RDD, 2.5 µL DNase I and RNase-free H_2O to a final volume of 100 µL. After incubating on the bench top for 10 minutes the QIAGEN RNeasy Plant Mini Kit RNA cleanup protocol was used. To the newly digested sample, 350 µL of Buffer RLT was added. After mixing the reaction, 250 µL of ethanol (96-100%) was added and mixed well by pipetting. The mixture was then

added to a RNeasy spin column and centrifuged for 15 seconds at 8,000 x g. After discarding the flow-through, 500 µL of Buffer RPE was added and the sample was centrifuged at 8,000 x g for 15 seconds. Once again the flow-through was discarded, 500 µL of Buffer RPE was added and the sample was centrifuged at 8,000 x g for 2 minutes. In order to eliminate residual Buffer RPE, the column was subsequently centrifuged at high speed (17,000 x g) for 1 minute. The column was placed in a clean microcentrifuge tube and the RNA was eluted by adding 40 µL of RNase-free H_2O, letting it sit for 5 minutes and then centrifuging for 1 minute at 8,000 x g.

2.4 Generation of cDNA from isolated RNA

All steps involving handling of RNA were done using RNase-free methods as previously described. To prepare cDNA, 2.5 µg of RNA was combined with 1 µL Oligo dT (0.5 µg/µL), 1 µL dNTPs (10 mM) and Milli-Q® H_2O to a final volume of 12 µL and heated at 65°C for 5 minutes. The mixture was cooled on ice for 2 minutes and then 4 µL 5X First Strand buffer, 2 µL DTT (0.1 M) and 1 µL RNase Inhibitor (10 units) were added. After heating the mixture at 42°C for 2 minutes, 1 µL SuperScript Reverse Transcriptase (200 units) was added. The mixture was incubated at 42°C for 50 minutes and then the SuperScript Reverse Transcriptase was inactivated at 70°C for 15 minutes.

2.5 Genomic DNA isolation for PCR optimization

In order to optimize PCR conditions for each gene-specific primer set, genomic DNA was isolated from Conrad stem and leaf tissue using a CTAB (Cetyltrimethyl

Ammonium Bromide) protocol (modified from Kang and Yang, 2004). An extraction buffer (1.4 M NaCl, 100 mM Tris-HCl pH 8.0, 20 mM EDTA pH 8.0 and 2% CTAB) was heated to 65°C with 0.2% v/v β-mercaptoethanol and 2 µL of RNase A per 600 µL aliquot. Plant tissue was ground in a mortar using a pestle and liquid nitrogen. Approximately 200 mg of ground tissue was resuspended in 600 µL of the preheated extraction buffer and incubated at 65°C. After an hour, 600 µL of Chloroform:Isoamyalcohol (24:1) was added and the mixture vortexed before centrifugation at 10,000 x g for 12 minutes. The top liquid phase was transferred to a clean microcentrifuge tube and 0.7 volumes of isopropanol was added in order to precipitate the DNA. The mixture was cooled at 4°C for 30 minutes and then centrifuged at 10,000 x g for 20 minutes at 4°C. After the supernatant was discarded, the pellet was washed with ice cold ethanol (70%) and centrifuged at 10,000 x g for 5 minutes. The pellet was then dried completely using an aspirator and resuspended in 50 µL of Milli-Q® H_2O.

2.6 Primer design and optimization

Primers were designed to unique regions in the 5' and 3' untranslated regions (UTRs) of the six putative *Gm FAωH*s (Table 1). Genomic DNA was used as the template to determine the optimal salt concentration, annealing temperature and cycle number for each primer pair. Salt concentration was optimized for each primer set by running PCR at 48°C, for 33 cycles and a range of different $MgCl_2$ concentrations (1.25 - 5.00 mM $MgCl_2$). Using the chosen optimal salt concentration for each primer set, the annealing

Table 1: Primer sets used for PCR amplification of potential *Gm FAωH* genes. The *ACT-11* primer sequences were obtained from Hu et al. (2009). F = forward and R = reverse. The sizes of the products (bp) are based on using cDNA. $T_m = (w^A+x^T)*2 + (y^G+z^C)*4$ where w^A,x^T,y^G and z^C are the number of the A, T, G and C nucleotide bases in the sequence, respectively.

Gene	Size of Product (bp)	Primer Sequences	T_m (°C)
Gm FAωH-1	1833	F 5' AGTCGCTCTGCTCGTTCGCTC R 5' CAGAGGAAATGATACAGCACCGT	68 68
Gm FAωH-2	1625	F 5' GGTCATATCAGGCTCTAGCTC R 5' GGTTTGAGATCACTTGGGTACACATT	64 74
Gm FAωH-3	2013	F 5' GGAGTATGGTGAGGCATTGTC R 5' CCATCTAGTGAACACCTGGTC	64 64
Gm FAωH-4	2012	F 5' TGGTGCAGCTGATGAGAG R 5' CCTCCAATACCCTATGCTTC	56 60
Gm FAωH-5	1779	F 5' CAAGATATATGTGCATGTCCAC R 5' TAATTCTGCAGCACAATGACG	62 60
Gm FAωH-6	1767	F 5' CTGACAGCGACTCTAGAAGGC R 5' CCATCTTCCAACAACTGTAAACCG	66 70
ACT-11	126	F 5' ATCTTGACTGAGCGTGGTTATTCC R 5' GCTGGTCCTGGCTGTCTCC	70 64

temperature was increased to find one that yielded a single band of expected size. Cycle number was determined by removing PCR samples after a certain number of cycles (17, 19, 21, 23, 25, 27, 29, 31, 33, and 35 cycles) and analyzing them on a 1% agarose gel. The relative brightness of the bands was determined using BioRad Quantity One Software Version 4.5.1. The cycle numbers were plotted against the intensity of the bands on the gel, and a cycle in the middle of the exponential phase was chosen. In order to check the specificity of the primers, each primer pair was checked with genomic DNA using the optimal salt and temperature determined by the above process.

2.7 PCR purification and sequencing

The PCR products generated using optimized salt concentration and annealing temperature with genomic DNA as the template were run on a 1% agarose gel and stained with ethidium bromide to check for specificity. The gels were visualized on a Bio-Rad Gel Doc (Universal Hood II, S76S/01511). PCR reaction products for each putative *Gm FAωH* were purified for sequencing using a QIAGEN QIAquick PCR Purification Kit as per the manufacturer's instructions. Five volumes of Buffer PB were added to 1 volume of the PCR sample and mixed (1 volume = 100 µL). The mixture was added to a QIAquick Spin column and centrifuged for 1 minute at 15,000 x g. After the flow-through was discarded, 750 µL of Buffer QG was added to the column and the sample was centrifuged at 15,000 x g for 1 minute. After the flow-through was discarded again, the column was re-centrifuged for 1 minute at 15,000 x g to remove all traces of ethanol. The column was placed in a clean microcentrifuge tube and the DNA was eluted by adding 30

µL of Buffer EB (10 mM Tris-Cl, pH 8.5), letting the column sit for 5 minutes and then centrifuging for 1 minute at 15,000 x g. The purified PCR reactions were sent to Robarts Research for sequencing via the Big Dye Terminator (BDT) v3.1 from Applied Biosystems, followed by purification via Sephadex spin using Pall Scientific Filter plates and GS 50 Sephadex. The purified samples were loaded onto the Applied Biosystems 3730 DNA Analyzer, and the data analyzed using Sequence Analysis Software v 5.2 from Applied Biosystems.

2.8 Semi quantitative RT-PCR of putative *Gm FAωH*s in soybean tissues

The optimal PCR amplification conditions (Section 2.6) were used to determine the tissue expression of the 6 putative *Gm FAωH* genes in RNA extracts from 9 different tissues: root, flower, M2 and L1 stage seeds, M2 and L1 stage pods, stem, old leaf, and young leaf. PCR was run using the six primer pairs and *ACT-11* primers (94°C 4 min.; 21-33 cycles of [94°C 50 sec., 56-67°C 50 sec., 72°C 2 min.]; 72°C 10 min., 4°C ∞) and the products were run on a 1% agarose gel and stained with ethidium bromide. The gels were visualized on a Bio-Rad Gel Doc (Universal Hood II, S76S/01511) for qualitative analysis.

CHAPTER 3: RESULTS AND DISCUSSION

3.1 *In silico* analysis of putative *Gm FAωH*s

Initial *in silico* analysis revealed a putative soybean ω-hydroxylase, CYP86A24 (*Gm FAωH-1*), which had been tagged due to its high identity with known CYP86As but never further characterized (Li and Yu, unpublished data). When a BLAST search was performed using the sequence of *Gm FAωH-1* against the soybean genome, five additional genes with high sequence similarity were found. These putative soybean ω-hydroxylases were named in terms of their decreasing identity to *Gm FAωH-1* (Table 2). A high number of genes was expected because of the palaeopolyploidic nature of the soybean genome (Schmutz et al., 2010).

The six putative *Gm FAωH*s are scattered over six of the twenty soybean chromosomes. As indicated by their genomic locus numbers (Glyma08g01890, Glyma05g37700, Glyma03g01050, Glyma07g07560, Glyma14g37130, Glyma11g26500), *Gm FAωH-1* is located on the very top of chromosome 8, *Gm FAωH-2* is located very close to the bottom of chromosome 5, *Gm FAωH-3* is located on the very top of chromosome 3, *Gm FAωH-4* is located on the top arm of chromosome 7, *Gm FAωH-5* is located near the bottom of chromosome 14 and *Gm FAωH-6* is located on the lower arm of chromosome 11. The distribution of the genes amongst various chromosomes is consistent with that observed for the *A. thaliana* CYP86A subfamily (Duan and Schuler, 2005).

Table 2: Pairwise nucleotide comparisons of putative *Gm FAωH*s. The percent nucleotide identity between the six putative *Gm FAωH*s is shown. Analysis done in DNAMAN.

	Gm FAωH-1	*Gm FAωH-2*	*Gm FAωH-3*	*Gm FAωH-4*	*Gm FAωH-5*
Gm FAωH-2	93.6				
Gm FAωH-3	72.9	72.7			
Gm FAωH-4	72.3	72.2	93.2		
Gm FAωH-5	64.4	64.2	63.0	62.7	
Gm FAωH-6	64.0	63.5	62.1	62.4	76.3

Comparisons conducted at the nucleotide level indicate that these genes have very different structural organizations, with *Gm FAωH-1* having a single intron interrupting the coding sequence at position -90 nucleotides (nt) (relative to ATG; i.e., in the 5' UTR), *Gm FAωH-5* having a single intron interrupting the coding sequence at position 675 nt (relative to ATG), *Gm FAωH-6* having two introns interrupting the coding sequence at positions 675 nt and 1523 nt (relative to ATG) and *Gm FAωH-2*, *Gm FAωH-3* and *Gm FAωH-4* having no introns. The intron of *Gm FAωH-5* and first intron of *Gm FAωH-6* are relatively large (816 nt and 855 nt, respectively) and are 42.7% identical to one another. The intron of *Gm FAωH-1* (356 nt) and the second intron of *Gm FAωH-6* (148 nt) are smaller and, except for the splice-site junctions, no significant matches exist between these introns and the larger ones. As indicated in homology trees generated for the six putative *Gm FAωH*'s (Figure 4), they group into three pairs with, *Gm FAωH-1* and *Gm FAωH-2*, *Gm FAωH-3* and *Gm FAωH-4*, and *Gm FAωH-5* and *Gm FAωH-6* grouping together. Pairwise comparisons between these sequences indicate that *Gm FAωH-1* and *Gm FAωH-2* have the highest degree of nucleotide identity within their coding sequence (93.6%) (Table 2). *Gm FAωH-3* and *Gm FAωH-4* have the same degree of identity with *Gm FAωH-1* (72.9% and 72.3%, respectively) as well as a high degree of identity with each other (93.2%). *Gm FAωH-5* and *Gm FAωH-6* have the lowest degree of identity with *Gm FAωH-1 (*64.4% and 64.0%, respectively) as well as the lowest degree of identity within their pairing (76.3%).

Pairwise comparisons conducted at the protein level indicate that these proteins are well conserved throughout most of their coding sequence and display the same

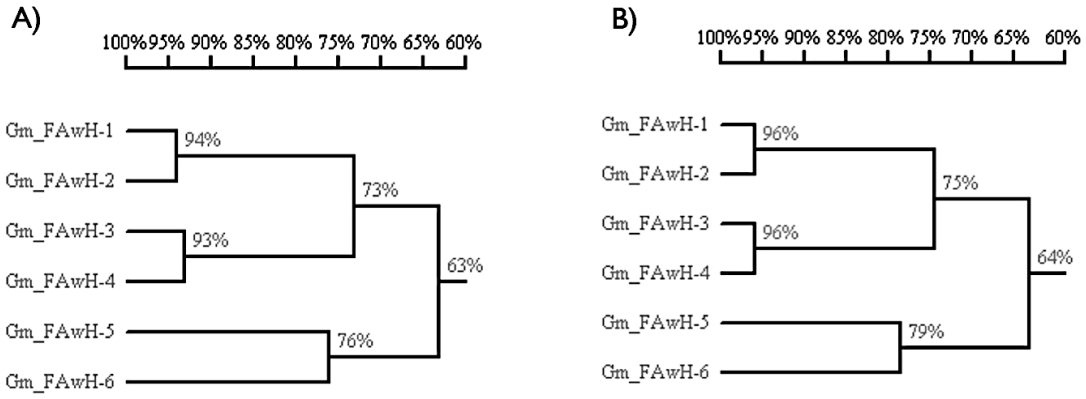


Figure 4: Homology trees of the six putative Gm FAωHs. Trees showing percent identity of A) nucleotide sequence of coding region and B) full length predicted amino acid sequences of the six putative Gm FAωHs.

derived relationships as seen at the nucleotide level (Table 3). Gm FAωH-3 and Gm FAωH-4 have the highest degree of sequence identity (95.9%), with Gm FAωH-1 and Gm FAωH-2 similarly identical (95.8%). Gm FAωH-5 and Gm FAωH-6 have lower similarity to each other (78.7%) and with the four other putative Gm FAωHs (63.3% - 65.3%). Since the six putative Gm FAωHs have >55% amino acid sequence identity they are likely part of the same subfamily of genes. These peptides differ in their number of amino acids, ranging from 508 (Gm FAωH-6), 520 (Gm FAωH-5), 528 (Gm FAωH-1 and Gm FAωH-2), and 532 (Gm FAωH-4), to 533 (Gm FAωH-3). Alignments at the primary sequence level indicate that nearly all of these length differences exist at the C terminus, with Gm FAωH-5 extending 12 residues, Gm FAωH-2 extending 20 residues, Gm FAωH-1 extending 21 residues and Gm FAωH-3 and Gm FAωH-4 extending 24 residues beyond the termination point of Gm FAωH-6 (Figure 5). Based on secondary structure prediction in DNAMAN, these extensions form random coil structures extending from the external surface of these proteins.

The six putative Gm FAωHs were aligned with other CYP86A proteins, including one that has been demonstrated to be a FAωH (CYP86A1 in *A. thaliana*) (Figure 5). This alignment revealed a high degree of conservation in the putative substrate recognition sites defined for CYP86As (Rupasinghe et al., 2007) between the six putative Gm FAωHs and confirmed CYP86As. All six putative Gm FAωHs are nearly identical in SRS1, SRS4, SRS5 and SRS6, more variable in SRS2, and most variable in SRS3 (Figure 5). These substrate recognition sites are thought to contribute to defining catalytic site specificity.

Table 3: Pairwise amino acid comparisons of putative Gm FAωHs. The percent amino acid identity between the six putative Gm FAωHs is shown. Analysis done in DNAMAN.

	Gm FAωH-1	Gm FAωH-2	Gm FAωH-3	Gm FAωH-4	Gm FAωH-5
Gm FAωH-2	95.8				
Gm FAωH-3	75.7	75.3			
Gm FAωH-4	75.5	75.1	95.9		
Gm FAωH-5	65.0	64.5	64.8	65.0	
Gm FAωH-6	65.3	64.6	63.3	63.3	78.7

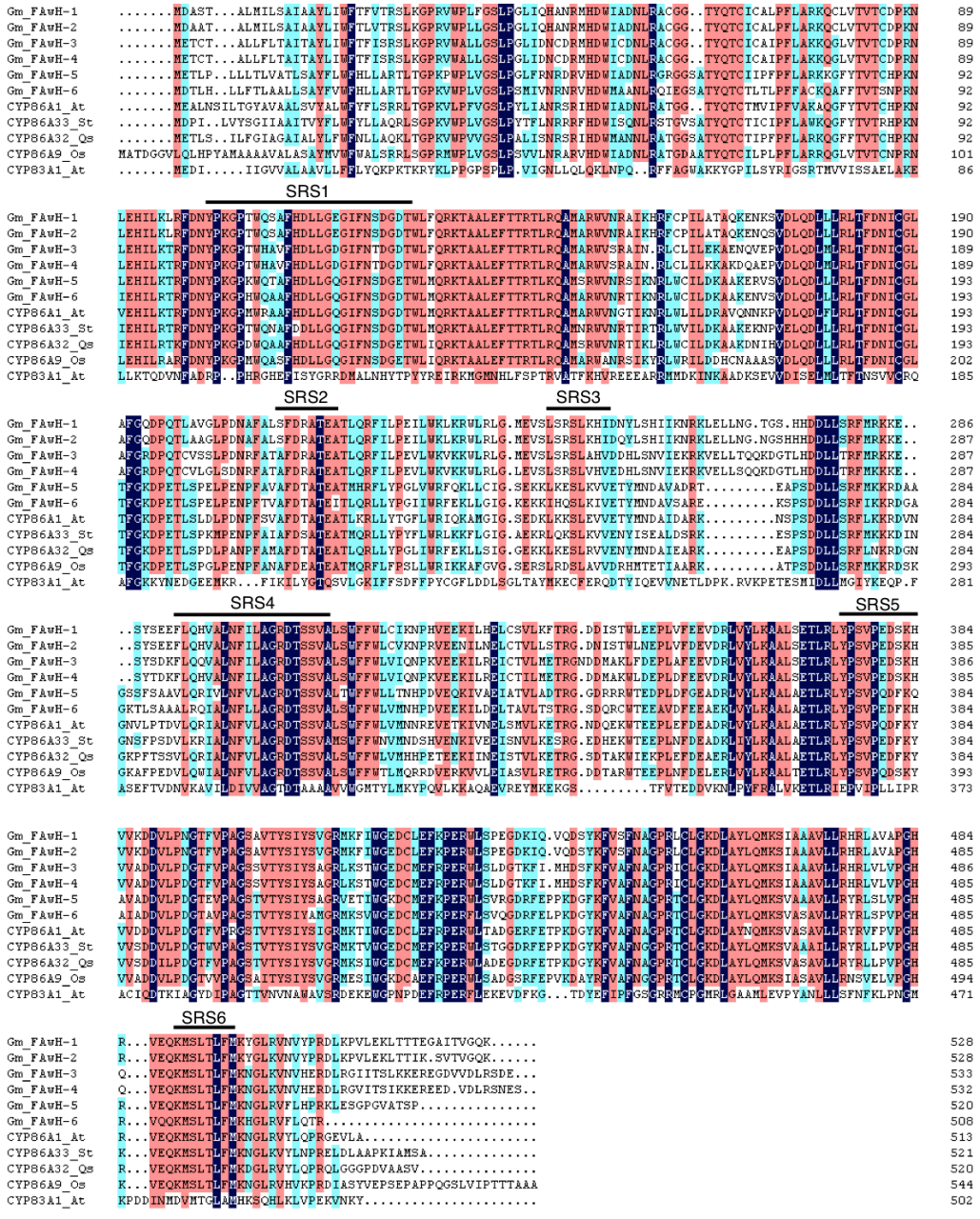


Figure 5: Amino acid sequence alignment of the related proteins CYP86A1, CYP86A33, CYP86A32, CYP86A9 and the six putative Gm FAωHs. Amino acids that are identical, >=75% similar or >=50% similar are shaded in dark blue, red or light blue, respectively. Putative substrate recognition sites are underlined (Rupasinghe et al., 2007). CYP83A1_At is used as an outgroup.

A rooted phylogenetic tree of potentially orthologous fatty acid ω-hydroxylases from different organisms was constructed based on multiple sequence alignments (Figure 6). Included in this comparison were fatty acid ω-hydroxylases known to exist in plants, including both the CYP86A and CYP94A subgroups. Notably, Gm FAωH-5 and Gm FAωH-6 are the closest of the putative Gm FAωHs to align with CYP86A1 from *A. thaliana*, a known fatty acid ω-hydroxylase. The similarities between the six sequences and where they fall in the phylogenetic tree with other known CYP86As suggests that they are all CYP86As.

3.2 *In silico* expression profiles

Existing *in silico* information was utilized to gain insight into the tissue expression patterns of the six putative *Gm FAωH*s. This information came from RNA-Seq (Soybase.org) and EST (various public EST databases) data available for the six putative *Gm FAωH*s. In addition, soybean microarray data was accessed.

RNA-Seq data was compiled from the atlas at Soybase.org (Table 4; Appendix II). This RNA sequencing data (RNA-Seq) was derived from high throughput sequencing of total RNA isolated from various soybean tissues at different developmental stages. The sequences obtained were then mapped to the soybean genome for identification (Severin et al., 2010). This unbiased, broad-based sequencing of the transcriptome allowed for an *in silico* examination of the tissue expression for all six putative *Gm FAωH*s. It indicated that all six putative *Gm FAωH*s were expressed, albeit in different tissues. For example, *Gm FAωH-1* appeared to be most highly expressed in young leaves and M2 stage seeds

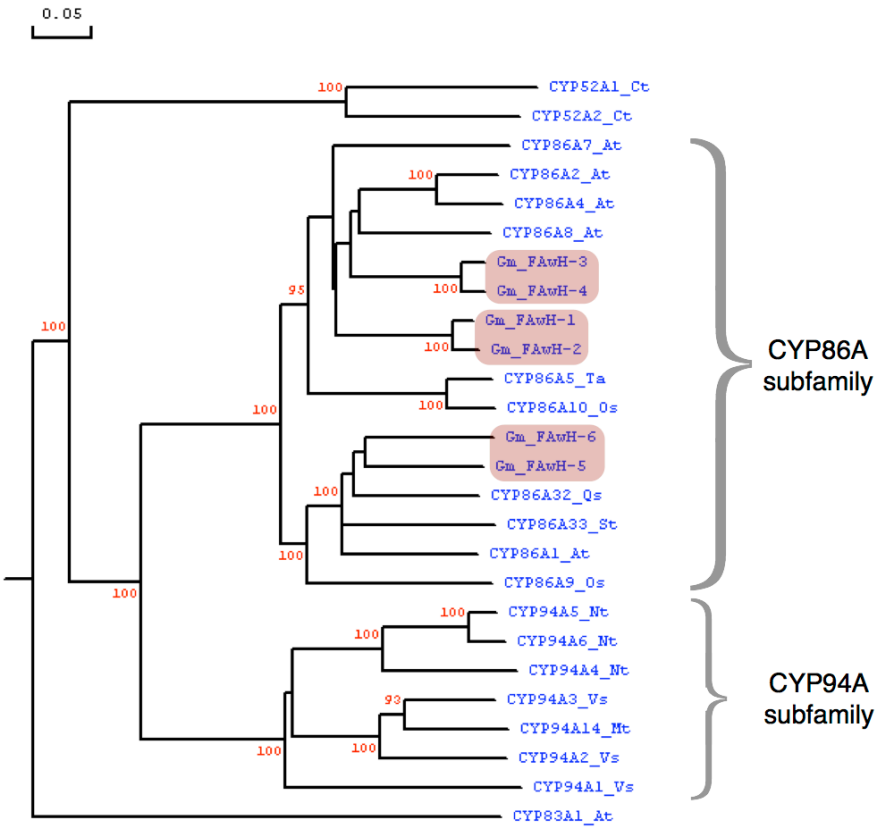


Figure 6: Phylogenetic tree of the predicted full length amino acid sequence comparison between known CYP86As and CYP94As with the six putative Gm FAωHs. Bootstrapped values indicating the level of significance (%) by the separation of the branches are shown in red. The branch length indicates the extent of the difference according to the scale at the top. The alignment used to construct the tree was performed in DNAMAN using ClustalW. CYP83A1_At is used as an outgroup. All sequences were obtained from NCBI. At: *Arabidopsis thaliana*; Ct: *Candida tropicalis*; Gm: *Glycine max*; Mt: *Meticago truncatula*; Nt: *Nicotiana tabacum*; Os: *Oryza sativa*; Qs: *Quercus suber*; St: *Solanum tuberosum*; Ta: *Triticum aestivum*; Vs: *Vivica sativa*.

Table 4: RNA-Seq expression data for the six putative *Gm FAωH*s. The raw data for the gene expression counts for the uniquely mappable reads are presented. The tissues included match those used in this study for RT-PCR verification. A more comprehensive table of RNA-Seq data, including several additional tissue types and developmental stages, is presented in Appendix II. DAF = days after flowering. Data were obtained from Soybase.org/soyseq.

Gene	Root	Flower	Seed M2		Seed L1	
			Seed 28DAF	Seed 35DAF	Seed 42DAF	Young Leaf
Gm FAωH-1	0	293	422	216	32	462
Gm FAωH-2	0	201	162	34	6	155
Gm FAωH-3	14	37	43	47	22	60
Gm FAωH-4	11	33	25	27	14	86
Gm FAωH-5	284	7	41	18	2	1
Gm FAωH-6	152	2	1	0	0	0

and to a lesser extent in flower and L1 stage seeds. *Gm FAωH-2* was shown to be most expressed in flower, M2 stage seeds and young leaves. *Gm FAωH-3* and *4* had similar, and relatively low expression in young leaves, seeds at the M2 and L1 stage, and flowers. *Gm FAωH-5* and *6* showed the least diverse expression with root being the tissue with the highest expression for both, while *Gm FAωH-5* also showed some expression in M2 stage seeds. An extended table of RNA-Seq data (Appendix II) indicates that the highest expression for *Gm FAωH-1* (over 700) was observed for pods that were younger than the M2 stage. Also, *Gm FAωH-6* showed high expression (916) in nodules. Neither of these tissues were represented in our RT-PCR tissue expression analysis (Section 3.5). On the whole, these expression numbers are relatively low for RNA-Seq data (where a highly expressed gene is two orders of magnitude higher), and may be due to the high sequence identity between the six putative *Gm FAωH* genes. Previous studies have indicated the potential for under-representing the total number of counts for genes within closely related gene families as short read sequences that map to multiple locations in a genome are largely discarded (Mortazavi et al., 2008).

Since *Gm FAωH-1* was the only soybean *FAωH* annotated in the NCBI Unigene database, M.A. Bernards and I collaborated with Dr. Martina Stromvick (McGill University) in order to make use of existing expressed sequence tag (EST) data more fully. ESTs that corresponded with a putative *Gm FAωH* were assembled into contiguous sequences (contigs) with the software Phrap (P. Green, University of Washington, Seattle, http://www.phrap.org/). This allowed for a more extensive *in silico* tissue expression analysis of the vast soybean EST data set. In total, 11 contigs were compiled (Appendix

III) which represented five out of the six *Gm FAωH*s. *Gm FAωH-1, Gm FAωH-2, Gm FAωH-3* and *Gm FAωH-5* were represented by two contigs each, while *Gm FAωH-4* was represented by 3 contigs. *Gm FAωH-6* was not represented in any of the EST data. In order to establish the tissue-specific expression of each of the five *Gm FAωH* genes represented in the data, the origin of each library used to generate the EST data was compiled from NCBI (Table 5). In general, the majority of ESTs were from libraries generated from seeds, though ESTs from root, flowers, pods, stems and other tissues were also available (Table 5). Overall, *Gm FAωH-1* was found to be expressed in a variety of tissues, while *Gm FAωH-3, Gm FAωH-4* and *Gm FAωH-5* had EST representation in libraries from seed only. Most ESTs of *Gm FAωH-2* were found in libraries from seed, but a couple were also found in libraries constructed from flowers. While the EST data for *Gm FAωH-1* matched that of the expression profile from Unigene, the fact that a large number of ESTs were from libraries that were generated from whole soybean plant tissue meant that there was limited tissue expression data. In general, the expression profiles predicted by source-cDNA library EST analysis showed corresponding and opposing expression patterns, compared to RNA-Seq, for all six genes. For example, RNA-Seq analysis showed high expression of *Gm FAωH-5* and *Gm FAωH-6* in root tissue, while analysis of EST data only showed a role of *Gm FAωH-1* in this tissue. This variation in root expression data emphasizes the need for RT-PCR-based verification of gene specific expression (Section 3.5).

Soybean microarray data is readily available. Four of the putative *Gm FAωH*s are represented as probe sets on the Affymetrix soy chip, although they are not unique. *Gm*

Table 5: EST data for five of the putative *Gm FAωH*s. A raw count of the number of ESTs generated from libraries from various tissues corresponding to those used in this study is presented. Contigs were assembled from source cDNA libraries and the frequency of occurrence in various tissues established.

Gene	Root	Flower	Seed	Pod	Stem	Leaf	Other*
Gm FAωH-1	2	8	4	3	1	3	25
Gm FAωH-2	0	2	16	0	0	0	12
Gm FAωH-3	0	0	7	0	0	0	1
Gm FAωH-4	0	0	19	0	0	0	1
Gm FAωH-5	0	0	23	0	0	0	0

*Includes EST libraries generated from whole plants, multiple tissues and other tissues not included in the RT-PCR study (shoot tips and hypocotyls).

FAωH-1 and *Gm FAωH-2* are both represented by Gma.11025.1.S1_at and GmaAffx. 49459.1.A1_at and *Gm FAωH-3* and *Gm FAωH-4* are both represented by Gma. 13695.1.A1_at. Since these probe sets correspond to more than one gene, the microarray data for these four genes is not reliable.

3.3 Designing gene-specific primers

To determine tissue expression for each *Gm FAωH* gene *in vitro*, unique gene-specific primers were designed. Due to the *Gm FAωH* genes having high identity in their coding region (Table 2), primers were designed to the 5' and 3' UTRs (Figure 7). An alignment was performed in DNAMAN to find regions of low identity amongst the six *Gm FAωH* genes. Sequences ranging from 18 - 26 base pairs (bp) were chosen because this length is long enough for adequate specificity and short enough for primers to bind easily to the template at the annealing temperature (Table 6). The primers were designed within 363 bp of the start codon in the 5' UTR and within 100 bp of the stop codon in the 3'UTR (Figure 7, Table 6). By necessity, the *Gm FAωH-2* reverse primer is very close to the stop codon albeit still within the coding sequence. The G/C content of each primer was kept between 40.9 - 61.9% (Table 6), which allowed for enough annealing strength (G-C base pairs have three hydrogen bonds) while keeping the melting temperature (T_m) at a favorable level (Li et al., 1997). The T_m is the temperature at which one half of the DNA duplex will dissociate to become single stranded (Rychlik et al., 1990). Single base pair modifications were incorporated within primers, as required, to ensure (1) no self annealing, (2) that two primer complementarity was kept to a minimum between primer

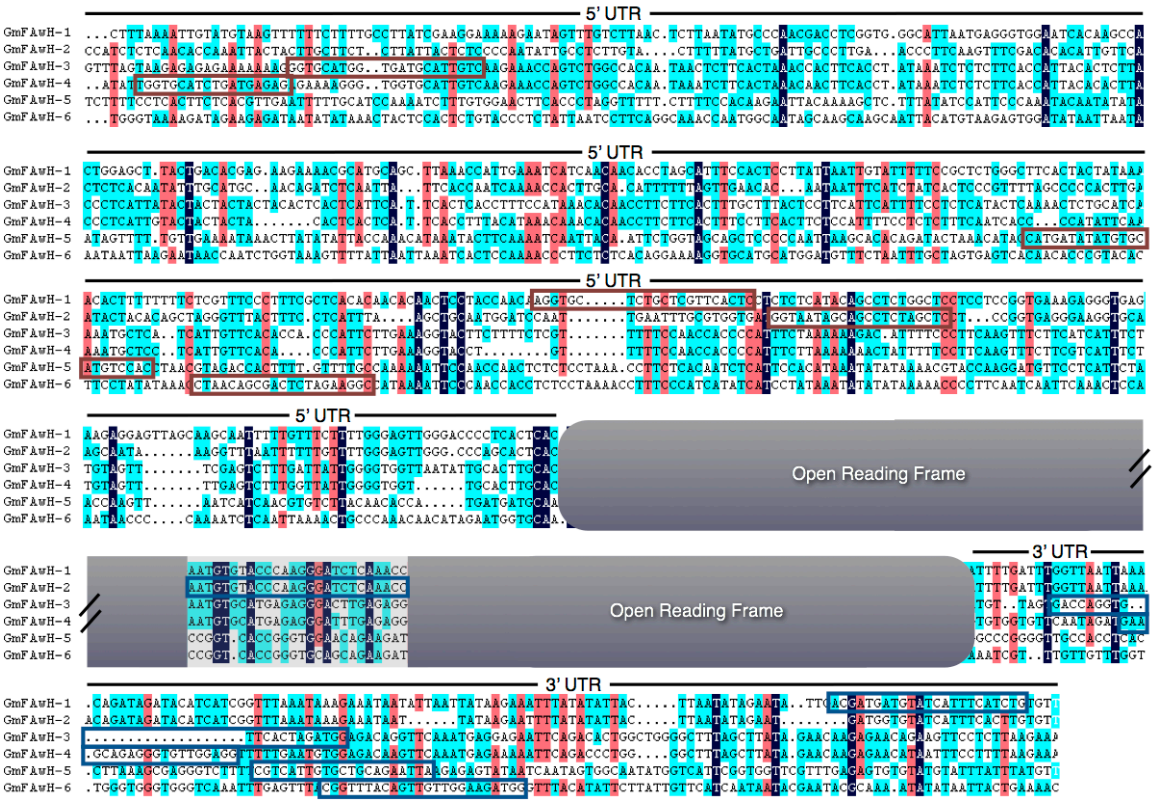


Figure 7: Nucleotide sequence alignment of the six putative *Gm FAωH*s. Nucleotides that are identical, >=75% similar or >=50% similar are shaded in dark blue, red or light blue, respectively. Forward primers are indicated by red boxes and reverse primers by blue boxes. The open reading frame is denoted as a grey box.

Table 6: Characteristics of the *Gm FAωH* primers designed for RT-PCR. The length, distance from the coding region and G/C content are indicated for each primer. FOR = forward and REV = reverse.

Primer	Length (bp)	Distance From Coding Region* (bp)	G/C Content (%)
Gm FAωH-1 FOR	21	102	61.9
Gm FAωH-1 REV	23	100	47.8
Gm FAωH-2 FOR	21	70	52.4
Gm FAωH-2 REV	26	53**	42.3
Gm FAωH-3 FOR	21	363	52.4
Gm FAωH-3 REV	21	7	52.4
Gm FAωH-4 FOR	18	358	55.6
Gm FAωH-4 REV	20	17	50.0
Gm FAωH-5 FOR	22	160	40.9
Gm FAωH-5 REV	21	14	42.9
Gm FAωH-6 FOR	21	143	57.1
Gm FAωH-6 REV	24	52	45.8

*The forward primers are shown as base pairs upstream from the start codon and the reverse primers are shown as downstream from the stop codon.

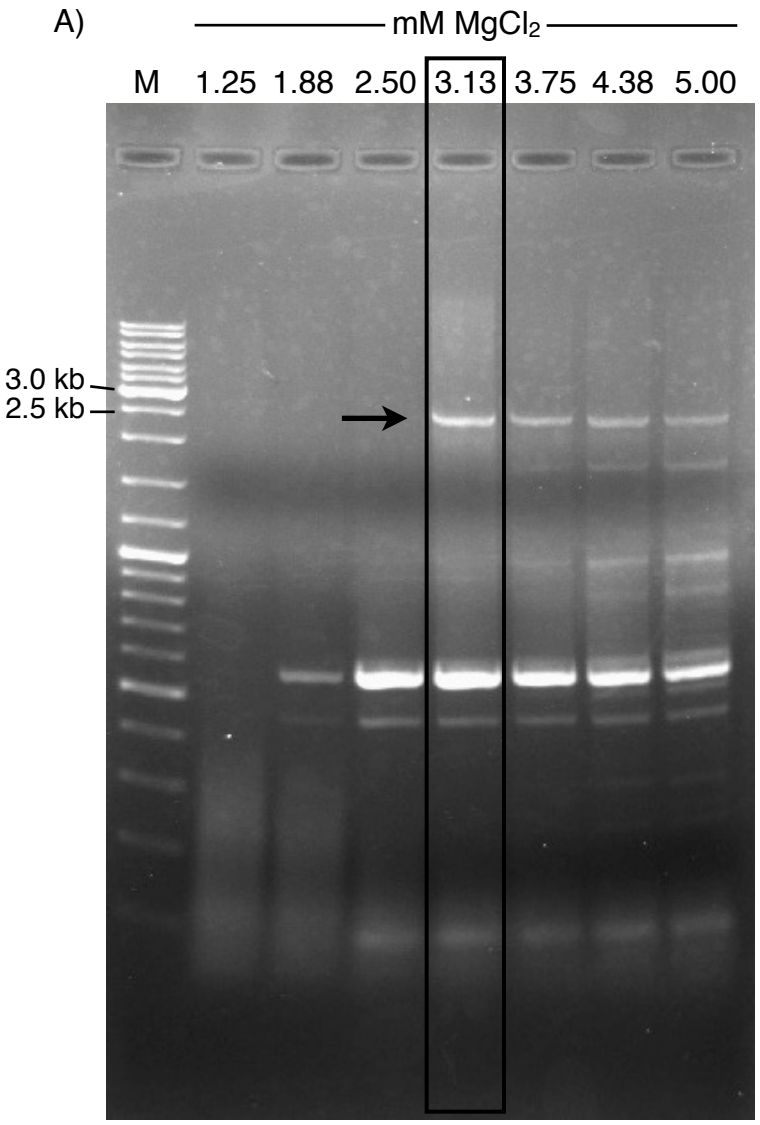
**The *Gm FAωH-2* reverse primer is upstream of the stop codon.

sets and (3) that T_m's were kept within 10°C for primer pairs. To ensure specificity to each gene, final primer sequences were "BLASTed" against the soybean genome.

3.4 Checking and optimizing primers

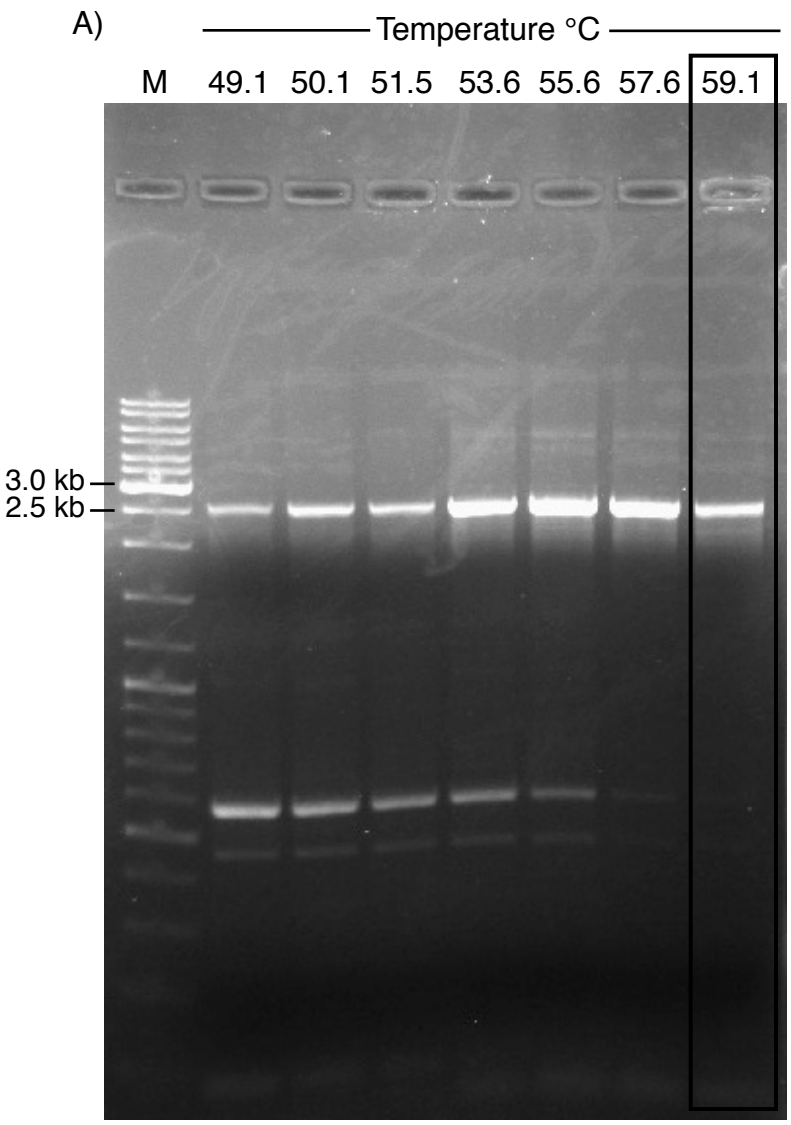
Primers were optimized and checked for specificity. Genomic DNA was used to establish optimal salt ($MgCl_2$) and temperature conditions. Magnesium is a required cofactor for thermostable DNA polymerases (Chien et al., 1976). Excess Mg^{2+} in the reaction can increase non-specific primer binding and increase the non-specific background of the reaction (Williams, 1989). Therefore, Mg^{2+} concentration is important for controlling the specificity of the annealing reaction. In order to find the optimal salt concentration for each primer set, PCR was run at a low temperature, high cycle number and a range of different $MgCl_2$ concentrations (Figure 8a). Salt concentrations with high clarity were chosen for each primer set, including the housekeeping gene actin, *ACT-11* (Figure 8b). The primers worked best at a range of salt concentrations (2.50 - 4.78 mM).

The annealing temperature was also optimized for specificity. Using the T_m calculated for each primer as a guide (lowest T_m per primer pair minus 5°C), each primer set was tested *in vitro* for the optimal annealing temperature to ensure maximum specificity (Figure 9). A temperature that is neither too low nor too high is ideal. A consequence of having an annealing temperature too low is that one or both primers will anneal to sequences other than the true target, as internal single-base mismatches or partial annealing may be permitted (Rychlik et al., 1990). This can lead to nonspecific amplification and, consequently, reduced yield of the desired product (Rychlik et al.,



Gene	$MgCl_2$ (mM)
Gm FAωH-1	4.38
Gm FAωH-2	4.38
Gm FAωH-3	3.13
Gm FAωH-4	2.50
Gm FAωH-5	3.13
Gm FAωH-6	3.75
ACT-11	4.78

Figure 8: Optimizing $MgCl_2$ concentration for PCR amplification of soybean *FAωH* genes. A) 1% agarose gel showing PCR products generated using a range of $MgCl_2$ concentrations for *Gm FAωH-5*. Labels across the top indicate $MgCl_2$ concentration in mM. The arrow shows the expected product band and the black box denotes the chosen concentration. Select bands are labeled in the 10 kb ladder lane (M). B) The chosen $MgCl_2$ concentration for each of the six putative *Gm FAωH*s and the positive control, *ACT-11*. These were found using the same process as with *Gm FAωH-5*.



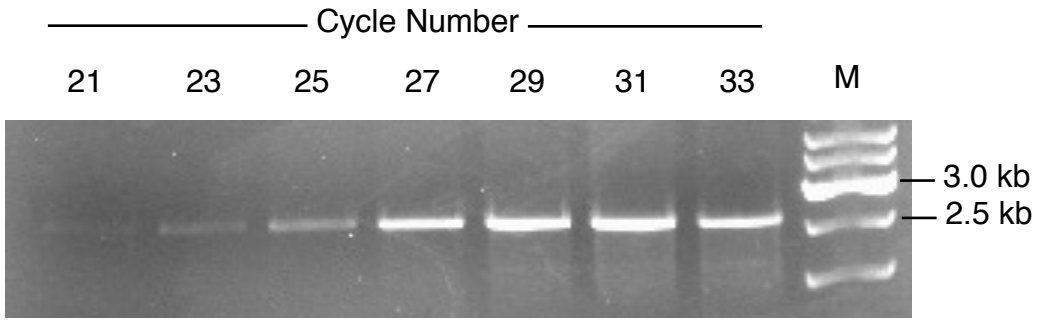
Gene	Temp (°C)
Gm FAωH-1	57
Gm FAωH-2	62
Gm FAωH-3	65
Gm FAωH-4	56
Gm FAωH-5	59
Gm FAωH-6	67
ACT-11	64

Figure 9: Optimizing annealing temperature for PCR amplification of soybean *FAωH* genes. A) 1% agarose gel showing PCR products generated using increasing temperature for *Gm FAωH-5*. Labels across the top indicate annealing temperature in °C. The black box shows the chosen temperature. Select bands are labeled in the 10 kb ladder lane (M). B) The chosen annealing temperature for each of the six putative *Gm FAωH*s and the positive control, *ACT-11*. These were found using the same process as with *Gm FAωH-5*.

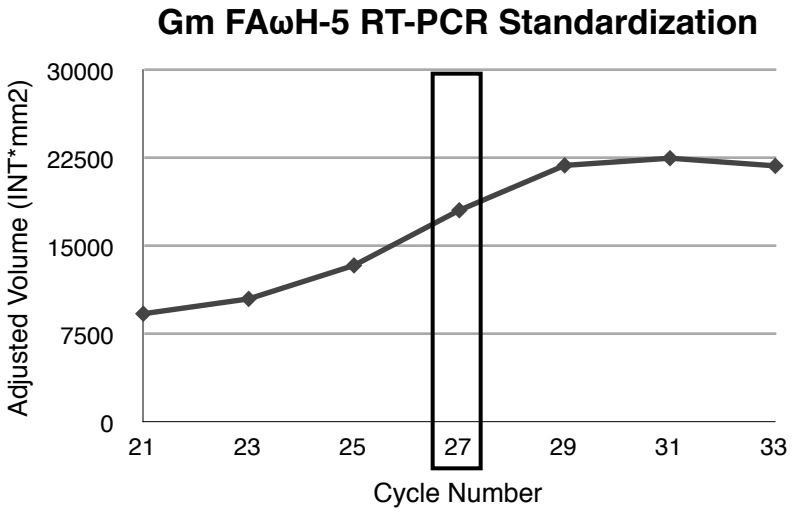
1990). Conversely, too high an annealing temperature may yield little product, as the likelihood of primer annealing is reduced (Rychlik et al., 1990). For each primer set, including the housekeeping gene actin (*ACT-11*), the highest temperature that generated a band was chosen (Figure 9b). The reason for this was the observation that in many cases secondary bands were observed even when a seemingly high temperature was used when cDNA was used as the template (data not shown).

To enable semi-quantitative RT-PCR, the cycle number for each primer pair was chosen within the exponential phase of band intensity. This is because PCR amplification is not logarithmic at a low or high number of amplification cycles (Prediger, 2001). Logarithmic or exponential amplification usually occurs only during the middle cycles (Prediger, 2001). Comparison between gene expression levels and tissues can therefore only be done for bands derived during this phase. This ensures that in each reaction the amount of each reagent is not depleted. PCR samples obtained with each primer set were removed after a prescribed number of cycles and were analyzed on a 1% agarose gel. An example is shown for *Gm FAωH-5* (Figure 10a). The relative intensity of the bands was determined, and the number of cycles plotted against the relative intensity of the bands on the gel (Figure 10b). A cycle within the exponential phase was chosen for each *Gm FAωH* (Figure 10c), and used as the cycle number for subsequent RT-PCR. This analysis was also done with the housekeeping gene, actin (*ACT-11*), as it was ultimately used as a baseline in subsequent RT-PCR analyses. Cycle numbers for the gene specific primer sets, on the whole, were high and ranged from 27 - 33 cycles. *ACT-11*'s cycle number was relatively low at 21 cycles.

A)



B)



C)

Gene	Cycle #
Gm FAωH-1	33
Gm FAωH-2	32
Gm FAωH-3	32
Gm FAωH-4	28
Gm FAωH-5	27
Gm FAωH-6	31
ACT-11	21

Figure 10: Optimizing cycle number for RT-PCR amplification of soybean *FAωH* genes. A) 1% agarose gel showing PCR products generated using increasing cycle number for *Gm FAωH-5*. Labels across the top indicate cycle number. Select bands are labeled in the 10 kb ladder lane (M). B) Graph of intensity of bands at each cycle number for *Gm FAωH-5*. The intensity is measured as a background-adjusted volume (band intensity*area). The black box shows the chosen cycle number within the exponential phase. C) The chosen cycle number for each of the six putative *Gm FAωH*s and the positive control, *ACT-11*. These were found using the same process as with *Gm FAωH-5*.

Since the specificity of these primers was vital to the outcome of tissue-specific expression analyses, each primer pair was checked with genomic DNA using the optimal salt concentration and temperature determined. This analysis resulted in a single band of expected size for each gene; *Gm FAωH-1* (2193 bp), *Gm FAωH-2* (1625 bp), *Gm FAωH-3* (2013 bp), *Gm FAωH-4* (2012 bp), *Gm FAωH-5* (2595 bp) and *Gm FAωH-6* (2770 bp) (Figure 11). The expected band sizes for *Gm FAωH-5* and *6* are relatively large because of introns in genomic DNA. The primers for *Gm FAωH-1* consistently generated smears, especially near the end of the data collection presented herein, and various troubleshooting attempts could not eliminate this problem.

To ensure each PCR reaction only had a single product (i.e., each primer pair amplified only one gene) and to further verify that the primers were amplifying the specific gene they were designed to amplify, the PCR reactions that generated the products shown in Figure 11 were purified and sent for sequencing. The sequencing results (Appendix IV) confirmed that each PCR reaction yielded a single product, except *Gm FAωH-1*. The smearing of *Gm FAωH-1* most likely lead to the observed complex sequencing result. Otherwise, alignments of each of the other PCR products, corresponding to the five other genes, revealed a 99% match to the expected gene. This gave confidence that each primer pair was unique to each gene.

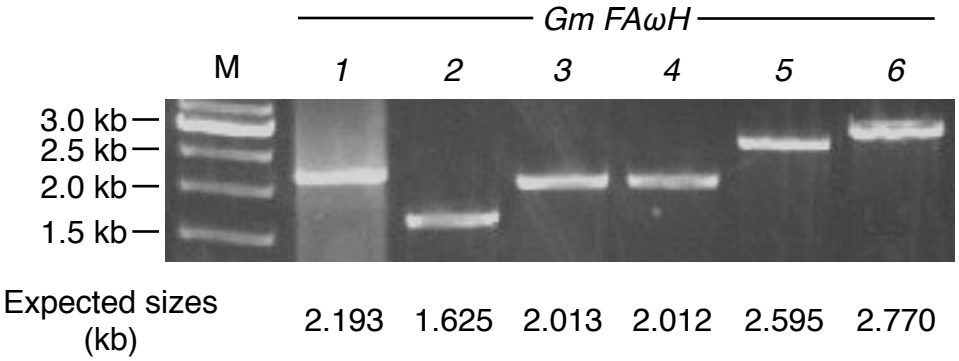


Figure 11: PCR products of each putative *Gm FAωH* generated from genomic DNA. 1% agarose gel showing the PCR products that were produced using the $MgCl_2$ concentration and annealing temperature optimized for each gene. Expected sizes for the genomic bands are indicated below. *GmFAωH-5* and *GmFAωH-6* have relatively large genomic band sizes because of introns. Select bands are labeled in the 10 kb ladder lane (M).

3.5 RT-PCR analysis of tissue-specific gene expression

The differential expression of each putative *Gm FAωH* was further examined *in vitro* using PCR based on the conditions optimized for each primer pair (Section 3.4) and equal amounts of cDNA from various soybean tissues (cv. Conrad), including root, flower, seed (M2 and L1 stages), pod (M2 and L1 stages), stem and old and young leaves. This was done with both non-DNase-treated (data not shown) and DNase-treated samples. A biological replicate was also performed using two groups of samples ('A' and 'B'). All RT-PCR results were qualitatively assessed against the expression level of a housekeeping gene transcript (*ACT-11*). Genomic DNA was used as a positive control to show that the conditions worked with each primer and a negative control (no transcript) was also included to make sure there wasn't any contamination or nonspecific amplification. Additionally, a control experiment was run with tissue samples generated without reverse transcriptase to ensure that any amplification that occurred in the sample was derived from the synthesized cDNA and not genomic DNA or other amplicon contamination. These are illustrated for *ACT-11* (Figure 12).

While smearing continued to be a problem for the *Gm FAωH-1* primer set, faint bands of the correct size were visible with cDNA from flower and M2 and L1 stage pod tissues (Figure 13). This corresponded well with both the RNA-Seq and EST data in that flower tissue showed high expression levels with both of these *in silico* analyses. Conversely, RT-PCR did not reveal *Gm FAωH-1* expression in young leaves or seeds, which was where the highest expression was found with RNA-Seq and, at least moderate expression in EST data. Pods at the M2 or L1 stage were not included in the

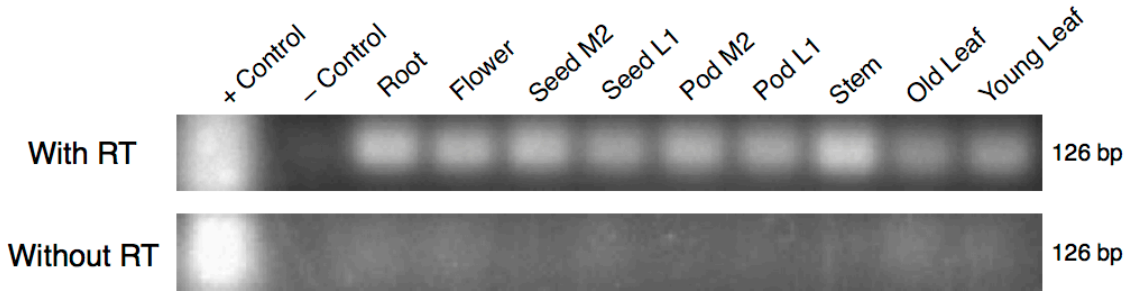


Figure 12: Analysis of various tissue samples contamination. 1% agarose gels showing PCR products generated with a housekeeping gene, *ACT-11*, using cDNA involved in RT-PCR tissue expression analysis ('B' samples). Labels across the top indicate tissues from which the cDNA was generated. The gel on top represents samples using cDNA generated using reverse transcriptase (with RT), while the gel below represents samples generated without reverse transcriptase (without RT). The positive control (+) is genomic DNA sample and the negative control (-) is a reaction mixture without added transcript (cDNA). Molecular weights (bp) are shown on the right hand side of each gel.

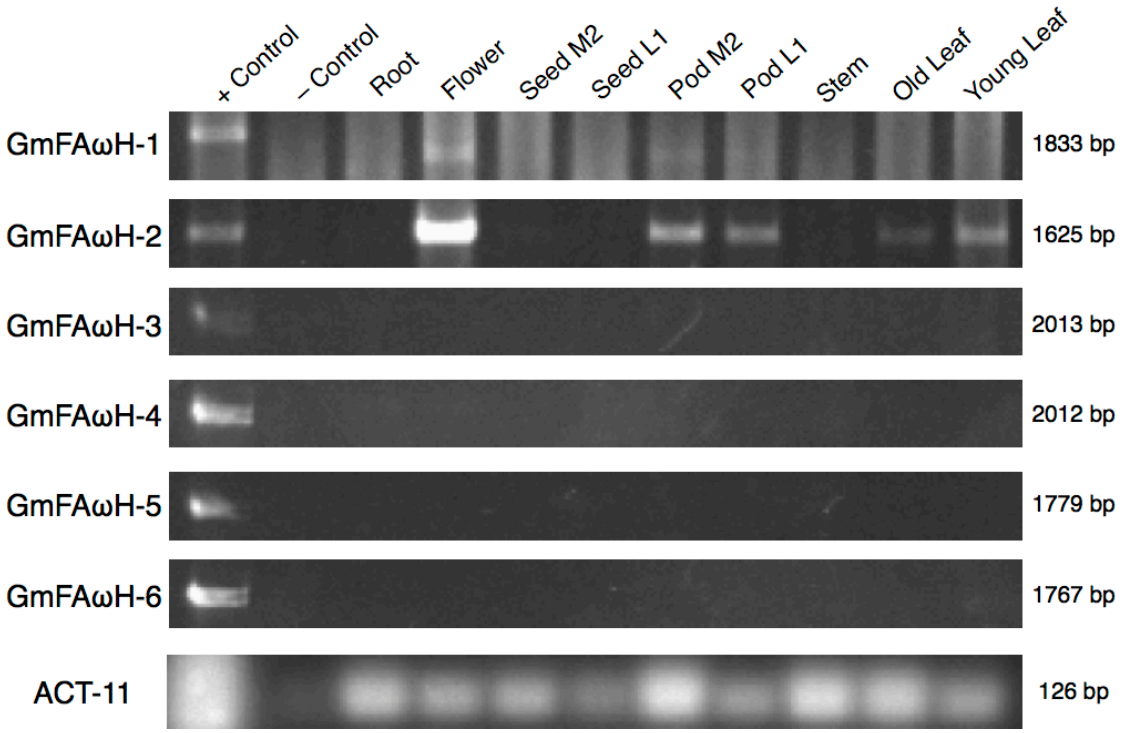


Figure 13: RT-PCR tissue expression of six putative *Gm FAωH*s using 'A' samples. Segments from 1% agarose gels in the molecular weight region of expected PCR products generated using the primer-pair optimized conditions ($MgCl_2$ concentration, temperature profile and cycle number) are shown for each primer pair. Labels across the top indicate tissues from which the cDNA was obtained. *ACT-11* is a housekeeping gene transcript that is used to qualitatively assess the results and ensure adequate cDNA is present for each preparation. The positive control (+) is a genomic DNA sample and the negative control (-) is a reaction mixture without added transcript (cDNA). Molecular weights (bp) are shown on the right hand side of each gel.

RNA-Seq data; however, my RT-PCR data indicate that expression would be higher in younger pods than older pods. As a consequence of the gel smearing obtained with *Gm FAωH-1* primers, it was hard to determine specific *Gm FAωH-1* tissue expression.

Gm FAωH-2 expression was observed in flower, M2 and L1 pods (Figures 13 and 14), and both old and young leaves (Figure 13). This corresponded to the RNA-Seq data in that flower showed the highest gene expression and young leaf showed moderate gene expression. Alternatively, both the RNA-Seq and EST data predicted *Gm FAωH-2* expression in seeds whereas this was not observed with RT-PCR. Again, M2 and L1 pods were not represented by RNA-Seq but my RT-PCR data suggests that *Gm FAωH-2* would be expressed in younger pods.

Gm FAωH-3 showed low expression in root (Figure 14), which corresponded with the RNA-Seq data. While both the RNA-Seq and EST data suggested that *Gm FAωH-3* should be expressed in seeds, my RT-PCR data did not detect any gene expression in seed. Similarly, RNA-Seq predicted *Gm FAωH-3* expression in flowers and young leaves, though this was not observed with my RT-PCR analysis.

Like its most closely matched potential homolog *Gm FAωH-3, Gm FAωH-4* appeared to be expressed in root and flower (Figure 14), which corresponded with the RNA-Seq data. However, *Gm FAωH-4* expression was found to be higher in young leaves, flowers and seeds with RNA-Seq, than predicted by my RT-PCR analysis. Similarly, EST data indicates relatively high expression in seed, while I saw little, if any, via RT-PCR.

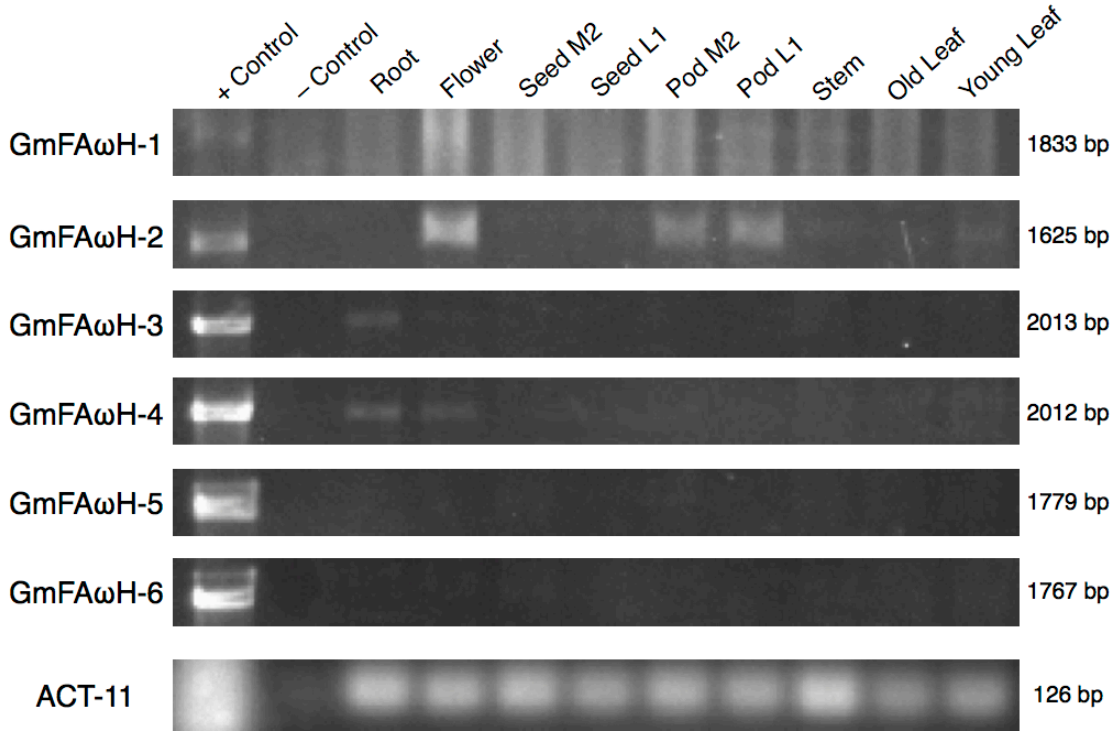


Figure 14: RT-PCR tissue expression of six putative *Gm FAωH*s using 'B' samples. Segments from 1% agarose gels in the molecular weight region of expected PCR products generated using the primer-pair optimized conditions ($MgCl_2$ concentration, temperature profile and cycle number) are shown for each primer pair. Labels across the top indicate tissues from which the cDNA was obtained. *ACT-11* is a housekeeping gene transcript that is used to qualitatively assess the results and ensure adequate cDNA is present for each preparation. The positive control (+) is a genomic DNA sample and the negative control (-) is a reaction mixture without added transcript (cDNA). Molecular weights (bp) are shown on the right hand side of each gel.

Gm FAωH-5 and *6* were not detected in any tissue via RT-PCR (Figure 13 and 14). While *Gm FAωH-5* was shown to have high expression in seeds (EST) and roots (RNA-Seq), the highest level of expression of *Gm FAωH-6* in RNA-Seq was in nodules, a tissue that wasn't studied in RT-PCR.

Gm FAωH gene expression in leaves is consistent with what is expected of suberin and cutin deposition. Suberin deposition is expected within the bundle sheath cells in leaves and the cuticle is actively deposited in young leaves (Esau, 1977). Additionally, some *Gm FAωH* genes should also contribute to hydroxy fatty acid formation for incorporation into the cutin found on the surface of flowers, seed coats and leaf surfaces (Kolattukudy, 2001). However, my primary interest is in the *Gm FAωH* genes expressed in root tissue, because suberin deposition in underground tissues correlates with *P. sojae* resistance (Thomas et al., 2007; Ranathunge et al., 2008). My RT-PCR data indicated that *Gm FAωH-3* and *Gm FAωH-4* are expressed in roots while the EST data indicated *Gm FAωH-1* expression in root. By contrast, RNA-Seq data showed that *Gm FAωH-5* and *6* had high levels of expression of all the *Gm FAωH*s in root. At this stage, it is not clear which *Gm FAωH*s are most abundantly expressed in soybean roots, and therefore responsible for the formation of the hydroxylated fatty acids found in soybean root suberin.

Variation in expression amongst the six *Gm FAωH* genes was anticipated because often genes within a gene family show differential patterns of expression dependent on tissue type and as a result of environmental and chemical stresses (Duan and Schuler, 2005). Overall, expression patterns from RT-PCR, RNA-Seq and ESTs matched for some

genes and tissues and varied for others. Because of the demonstrated specificity of the gene-specific primers (using genomic DNA template; Section 3.4), and the positive control analysis of the RNA extracts from my tissue samples (demonstrated with actin; Figure 12), I have confidence in my RT-PCR results. However, there remains a need to replicate my results beyond the two independent analyses completed herein, both to confirm the patterns I have observed and ensure that I have captured the true expression patterns of the family of *Gm FAωH*s. Both the RNA-Seq and EST data are dependent on computer models of sequence specificity. These six identified putative *Gm FAωH* genes have such high identity between them that it is possible that sequence counts of expression could be misallocated. Nevertheless, variation amongst the three expression analyses is expected because of the variation of tissue harvesting. I.e., tissues may have been harvested during different developmental stages leading to a variation in expression patterns.

3.6 A role for *FAωH*s in suberin formation

The goal of this research was to find soybean *FAωH* genes that are likely involved in suberin biosynthesis. From the RT-PCR, EST and RNA-Seq expression data (Table 7), tissue-specific roles for each of the six putative *Gm FAωH*s can be hypothesized.

Gm FAωH-1 and *Gm FAωH-2* showed expression in flower in all three types of expression analyses indicating that these genes are likely involved in cutin biosynthesis. Expression was also shown in seed, pod, stem (*Gm FAωH-1* only) and leaf furthering their involvement in cutin biosynthesis. Additionally, *Gm FAωH-1* was shown to be

Table 7: Tissue expression comparison of the putative *Gm FAωH*s from RT-PCR, EST and RNA-Seq data. The tissue expression of each putative *Gm FAωH* for six broad tissue types is indicated by the presence of the type of analysis.

Gene	Root	Flower	Seed	Pod	Stem	Leaf
Gm FAωH-1	EST	RT-PCR EST RNA-Seq	EST RNA-Seq	RT-PCR EST	EST	EST RNA-Seq
Gm FAωH-2		RT-PCR EST RNA-Seq	EST RNA-Seq	RT-PCR		RT-PCR RNA-Seq
Gm FAωH-3	RT-PCR RNA-Seq	RNA-Seq	EST RNA-Seq			RNA-Seq
Gm FAωH-4	RT-PCR RNA-Seq	RT-PCR RNA-Seq	EST RNA-Seq			RNA-Seq
Gm FAωH-5	RNA-Seq	RNA-Seq	EST RNA-Seq			
Gm FAωH-6	RNA-Seq					

expressed in root in EST analysis which gives evidence that it may be involved in suberin biosynthesis as well.

The expression analysis of *Gm FAωH-3* and *Gm FAωH-4* gives evidence for involvement in both suberin and cutin biosynthesis. Both the RT-PCR and RNA-Seq data showed expression in root indicating a connection with suberin biosynthesis whereas expression in flower, seed and leaf from a variety of data demonstrates involvement in cutin biosynthesis.

Lastly, RNA-Seq analysis showed expression of *Gm FAωH-5* in root, flower and seed indicating an involvement in both suberin and cutin biosynthesis. *Gm FAωH-6* only showed expression in root in RNA-Seq analysis giving evidence that it may be exclusively involved in suberin biosynthesis.

This research indicates that *Gm FAωH-1, Gm FAωH-3, Gm FAωH-4, Gm FAωH-5* and *Gm FAωH-6* are expressed in root and are likely involved in suberin biosynthesis.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

Using a comparative sequence-based *in silico* approach, six putative *FAωH*s were identified in the soybean genome. These six genes were shown to have high sequence identity making them a probable gene family, likely arising from the palaeopolyploidic nature of the soybean genome. Gene-specific primers were developed, RT-PCR conditions optimized and the specificity shown using genomic DNA. RT-PCR was then performed with various tissues, and revealed differential expression patterns for *Gm FAωH-1*, *Gm FAωH-2, Gm FAωH-3* and *Gm FAωH-4*. Neither *Gm FAωH-5* nor *Gm FAωH-6* were detected in any of the tissues used. RNA-Seq and EST *in silico* data showed both corresponding and opposing expression patterns for all six genes. The expression patterns of the six putative *Gm* FAωHs indicated by these analyses gives evidence as to their tissue-specific roles. *Gm FAωH-1, Gm FAωH-3, Gm FAωH-4, Gm FAωH-5* and *Gm FAωH-6* were shown to be expressed in roots by one or more of these expression analyses and therefore could be involved in suberin biosynthesis.

Due to the differing expression data noted above, tissue expression of the six putative *Gm FAωH*s will need to be confirmed in order to determine which specific gene(s) are expressed in roots; root expression of *Gm FAωH* genes is expected to be linked to the deposition of suberin in this tissue. Subsequently, an RNAi (interference) knockout approach, targeting the root specific *Gm FAωH* genes can be developed. The latter can be accomplished using *Agrobacterium rhizogenes*-derived hairy roots, which are readily generated and avoid the necessity for transgenic plant regeneration (Nilsson

and Olsson, 1997). Hairy roots as a vehicle for studying suberin formation and its role in *P. sojae* resistance is currently being characterized in the Bernards lab. Knocking out selective *Gm FAωH*s in root tissue can address two important hypotheses. First, the hypothesis that specific *Gm FAωH* genes are involved in root suberin biosynthesis can be tested by selectively inhibiting individual *Gm FAωH* genes and analyzing the impact of suberin deposition in the knockout mutants using chemical analysis (e.g., MeOH-HCl depolymerization of aliphatic suberin; Meyer et al*.,* 2011). If suberin is compromised in either quantity or quality, it will demonstrate the importance of specific fatty acid ω-hydroxylases in the formation of root suberin. The second hypothesis, that the quantity of root suberin is correlated with resistance to *P. sojae* can be tested with an infection study, wherein the ability of soybean knockout mutants to slow down infection by *P. sojae* is monitored (Subramanian et al., 2005; Ranathunge et al., 2008). If mutants with compromised suberin show enhanced susceptibility to *P. sojae*, it will support the hypothesis that suberin is involved in the resistance to *P. sojae*.

BIBLIOGRAPHY

Arnason, J. T., and Bernards, M. A. 2010. Impact of constitutive plant natural products on herbivores and pathogens. Can. J. Zool. **88**: 615–627. doi:10.1139/Z10-038

Bernards, M A. 2002. Demystifying suberin. Can. J. Bot. **80**: 227–240. doi: 10.1139/B02-017.

Bernards, M. A., and Lewis, N. G. 1998. The macromolecular aromatic domain in suberized tissue: a changing paradigm. Phytochemistry, **47**: 915–933. doi:10.1016/S0031-9422(98)80052-6.

Bernards, M. A., Lopez, M. L., Zajicek, J., and Lewis, N. G. 1995. Hydroxycinnamic acid-derived polymers constitute the polyaromatic domain of suberin. J. Biol. Chem. **270**: 7382–7386.

Burnham, K., Dorrance, A., VanToai, T., and St. Martin, S. K. 2003. Quantitative trait loci for partial resistance to *Phytophthora sojae* in soybean. Crop Sci. **43**: 1610–1617. doi:10.1111/j.1439-0523.2010.01799.x.

Chien, A., Edgar, D. B., and Trela, J. M. 1976. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. J Bacteriol. **127**: 1550–1557.

Compagnon, V., Diehl, P., Benveniste, I., Meyer, D., Schaller, H., Schreiber, L., Franke, R., et al. 2009. CYP86B1 is required for very long chain omega-hydroxyacid and α, ω -dicarboxylic acid synthesis in root and seed suberin polyester. Plant Physiol. **150**: 1831–1843. doi:10.1104/pp.109.141408.

De Vetten, N., ter Horst, J., van Schaik, H. P., de Boer, A., Mol, J., and Koes, R. 1999. A cytochrome *b*5 is required for full activity of flavonoid 3', 5'-hydroxylase, a cytochrome P450 involved in the formation of blue flower colors. Proc. Natl. Acad. Sci. **96**: 778–783.

Dobritsa, A. A., Shrestha, J., Morant, M., Pinot, F., Matsuno, M., Swanson, R., Møller, B. L., et al. 2009. CYP704B1 is a long-chain fatty acid ω-hydroxylase essential for sporopollenin synthesis in pollen of *Arabidopsis*. Plant Physiol. **151**: 574–589. doi: 10.1104/pp.109.144469.

Dorff, E. 2009. The soybean, agriculture's jack-of-all-trades, is gaining ground across Canada. Canadian Agriculture at a Glance. Statistics Canada Catalogue no. 96-325-XIE.

Dorrance, A. E., McClure, S. A., and St. Martin, S. K. 2003. Effect of partial resistance on phytophthora stem rot incidence and yield of soybean in Ohio. Plant Dis. **87**: 308–312.

Dorrance, A.E., Mills, D., Robertson, A. E., Draper, M. A., Giesler, L. and Tenuta, A. 2007. Phytophthora root and stem rot of soybean. The Plant Health Instructor. doi: 10.1094/PHI-I-2007-0830-07.

Duan, H., and Schuler, M. 2005. Differential expression and evolution of the *Arabidopsis* CYP86A subfamily. Plant Physiol. **137**: 1067–1081. doi:10.1104/pp.104.055715.

Enkerli, K., Hahn, M. G., and Mims, C. W. 1997. Ultrastructure of compatible and incompatible interactions of soybean roots infected with the plant pathogenic oomycete *Phytophthora sojae*. Can. J. Bot. **75**: 1493–1508.

Esau, K. 1977. Anatomy of seed plants. John Wiley and Sons, New York.

Freeman, B.C., and Beattie, G. A. 2008. An Overview of plant defenses against pathogens and herbivores. The Plant Health Instructor. doi: 10.1094/PHI-I-2008-0226-01.

Gonzalez, F. J., and Korzekwa, K. R. 1995. Cytochromes P450 expression systems. Annu. Rev. Pharmacol. Toxicol. **35**: 369–390. doi:10.1146/annurev.pa.35.040195.002101.

Graca, J., and Pereira, H. 1997. Cork suberin: A glyceryl based polyester. Holzforschung, **51**: 225–234. doi:10.1515/hfsg.1997.51.3.225.

Graça, J., and Pereira, H. 2000. Diglycerol alkendioates in suberin: building units of a poly(acylglycerol) polyester. Biomacromolecules, **1**: 519–522.

Grau, C. R., Dorrance, A. E., Bond, J. and Russin, J. S. 2004. Fungal Disease. *In* Soybeans: Improvement, production, and uses. Agronomy Monograph, 3rd ed. Vol. 16. *Edited by* H.R. Boerma and J.E. Specht. Madison, WI: Amercian Society of Agronomy, Crop Science Society of America, Soil Science Society of America. Pp. 679–763.

Hammerschmidt, R. 1999. Phytoalexins: What have we learned after 60 years? Annu. Rev. Phytopathol. **37**: 285–306. doi:10.1146/annurev.phyto.37.1.285.

Harwood, J. L. 1988. Fatty acid metabolism. Ann. Rev. Plant Physiol. Plant Mol. Biol. **39**: 101–138. doi:10.1146/annurev.pp.39.060188.000533.

Holloway, P. 1983. Some variations in the composition of suberin from the cork layers of higher plants. Phytochemistry, **22**: 495–502. doi:10.1016/0031-9422(83)83033-7.

Höfer, R., Briesen, I., Beck, M., Pinot, F., Schreiber, L., and Franke, R. 2008. The *Arabidopsis* cytochrome P450 CYP86A1 encodes a fatty acid ω-hydroxylase involved in suberin monomer biosynthesis. J. Exp. Bot. **59**: 2347–2360. doi:10.1093/jxb/ern101.

Hu, R., Chengming, F., Hongyu, L., Qingzhu, Z. And Fu, Y-F. 2009. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Mol. Biol. **10**: 93. doi:10.1186/1471-2199-10-93.

Kang, T. J., and Yang, M. S. 2004. Rapid and reliable extraction of genomic DNA from various wild-type and transgenic plants. BMC Biotech. **4**: 20. doi: 10.1186/1472-6750-4-20.

Kaitany, R. C., Hart, L. P., and Safir, G. R. 2001. Virulence composition of *Phytophthora sojae* in Michigan. Plant Disease, **85**: 1103–1106. doi:10.1094/PDIS.2001.85.10.1103

Kolattukudy, P. E. 1980. Biopolyester membranes of plants: Cutin and suberin. Science, **208**: 990–1000. doi:10.1126/science.208.4447.990.

Kolattukudy, P. E. 1984. Biochemistry and function of cutin and suberin. Can. J. Bot. **62**: 2918–2933. doi:10.1139/b84-391.

Kolattukudy, P. E. 2001 Polyesters in higher plants. *In* Advances in biochemical engineering biotechnology. Vol. 71: Biopolyesters. *Edited by* T. Scheper. Springer-Verlag, Berlin. pp. 1–19.

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., Franke, R., et al. 2010. Cytochrome P450 family member CYP704B2 catalyzes the ω-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. Plant Cell, **22**: 173–190. doi:10.1105/tpc.109.070326.

Li, P., Kupfer, K. C., Davies, C. J., Burbee, D., Evans, G. A., and Garner, H. R. 1997. Primo: A primer design program that applies base quality statistics for automated large-scale DNA sequencing. Genomics, **40**: 476–485. doi:10.1006/geno.1996.4560.

Meyer, C., Peterson, C. A., and Bernards, M. A. 2011. Spatial and temporal deposition of suberin during maturation of the onion root exodermis. Botany, **89**: 119–131. doi: 10.1139/B10-092.

Morris, P. F., and Ward, E. W. B. 1992. Chemoattraction of zoospores of the soybean pathogen, *Phytophthora sojae*, by isoflavones. Physiol. Mol. Plant Pathol. **40**: 17–22. doi:10.1016/0885-5765(92)90067-6.

Morris, P. F., Bone, E., and Tyler, B. M. 1998. Chemotropic and contact responses of *Phytophthora sojae* hyphae to soybean isoflavonoids and artificial substrates. Plant Physiol. **117**: 1171–1178.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods, **5**: 621–628. doi:10.1038/nmeth.1226.

Negrel, J., Pollet, B., and Lapierre, C. 1996. Ether-linked ferulic acid amides in natural and wound periderms of potato tuber. Phytochemistry, **43**: 1195–1199. doi:10.1016/ S0031-9422(96)00500-6.

Nilsson, O., and Olsson, O. 1997. Getting to the root: The role of the *Agrobacterium rhizogenes* rol genes in the formation of hairy roots. Physiol. Plant. **100**: 463–473. doi:10.1111/j.1399-3054.1997.tb03050.x.

Olah, A. F., and Schmitthenner, A. F. 1985. A growth chamber test for measuring phytophthora root rot tolerance in soybean seedlings. Phytopathology, **75**: 546–548. doi:10.1094/Phyto-75-546.

Prediger, E. A. 2001. Quantitating mRNAs with relative and competitive RT-PCR. Methods in Mol. Biol. **160**: 49–63. doi:10.1385/1-59259-233-3:049.

Ranathunge, K., Shao, S., Qutob, D., Gijzen, M., Peterson, C. A., and Bernards, M. A. 2010. Properties of the soybean seed coat cuticle change during development. Planta, **231**: 1171–1188. doi:10.1007/s00425-010-1118-9.

Ranathunge, K., Thomas, R. H., Fang, X., Peterson, C. A., Gijzen, M., and Bernards, M. A. 2008. Soybean root suberin and partial resistance to root rot caused by *Phytophthora sojae*. Phytopathology, **98**: 1179–1189. doi:10.1094/PHYTO-98-11-1179.

Rupasinghe, S. G., Duan, H., and Schuler, M. A. 2007. Molecular definitions of fatty acid hydroxylases in *Arabidopsis thaliana*. Proteins, **68**: 279–293. doi:10.1002/prot.21335.

Rychlik, W., Spencer, W. J., and Rhoads, R. E. 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. Nucleic Acids Res. **18**: 6409–6412.

Schmitthenner, A. F. 1985. Problems and progress in control of phytophthora root rot of soybean. Plant Disease, **69**: 362–368. doi:10.1094/PD-69-362.

Schmutz, A., Jenny, T., and Ryser, U. 1994. A caffeoyl fatty-acid glycerol ester from wax associated with green cotton fiber suberin. Phytochemistry, **36**: 1343–1346. doi:10.1016/S0031-9422(00)89721-6.

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature, **463**: 178–183. doi:10.1038/nature08670.

Schuler, M. A., and Werck-Reichhart, D. 2003. Functional genomics of P450s. Annu. Rev. Plant Biol., **54**: 629–667. doi:10.1146/annurev.arplant.54.031902.134840.

Serra, O., Soler, M., Hohn, C., Sauveplane, V., Pinot, F., Franke, R., Schreiber, L., et al. 2009. CYP86A33-targeted gene silencing in potato tuber alters suberin composition, distorts suberin lamellae, and impairs the periderm's water barrier function. Plant Physiol. **149**: 1050–1060. doi:10.1104/pp.108.127183.

Severin, A. J., Woody, J. L., Bolon, Y.-T., Joseph, B., Diers, B. W., Farmer, A. D., Muehlbauer, G. J., et al. 2010. RNA-Seq atlas of *Glycine max*: A guide to the soybean transcriptome. BMC Plant Biology, **10**: 1–16. doi:10.1186/1471-2229-10-160.

St. Martin, S. K., Scott, D. R., Sschmitthenner, A. F., and McBlain, B. A. 1994. Relationship between tolerance to phytophthora rot and soybean yield. Plant Breed. **113**: 331–334. doi:10.1111/j.1439-0523.1994.tb00745.x.

Subramanian, S., Graham, M. Y., Yu, O., and Graham, T. L. 2005. RNA interference of soybean isoflavone synthase genes leads to silencing in tissues distal to the transformation site and to enhanced susceptibility to *Phytophthora sojae*. Plant Physiol. **137**: 1345–1353. doi:10.1104/pp.104.057257.

Thomas, R., Fang, X., Ranathunge, K., Anderson, T. R., Peterson, C. A., and Bernards, M. A. 2007. Soybean root suberin: Anatomical distribution, chemical composition, and relationship to partial resistance to *Phytophthora sojae*. Plant Physiol. **144**: 299–311. doi:10.1104/pp.106.091090.

Tooley, P. W., and Grau, C. R. 1984. Field characterization of rate-reducing resistance to *Phytophthora megasperma* f. sp. *glycinea* in soybean. Phytopathology, **74**: 1201–1208. doi:10.1094/Phyto-74-1201.

Tyler, B. M. 2002. Molecular basis of recognition between phytophthora pathogens and their hosts. Annu. Rev. Phytopathol. **40**: 137–167. doi:10.1146/annurev.phyto.40.120601.125310.

Tyler, B. M. 2007. *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. Mol. Plant Pathol. **8**: 1–8. doi:10.1111/j.1364-3703.2006.00373.x.

Tyler, B. M., Wu, M-H., Wang, J-M., Cheung, W., and Morris, P. F. 1996. Chemotactic preferences and strain variation in the response of *Phytophthora sojae* zoospores to host isoflavones. Appl. Environ. Microbiol. **62**: 2811–2817.

Williams, J.F. 1989. Optimization strategies for the polymerase chain reaction. Biotechniques, **7**: 762–769.

Yang, W-L., and Bernards, M. A. 2006. Wound-induced metabolism in potato (*Solanum tuberosum*) tubers: Biosynthesis of aliphatic domain monomers. Plant Signal. Behav. **1**: 59–66.

Zeier, J., and Schreiber, L. 1997. Chemical composition of hypodermal and endodermal cell walls and xylem vessels isolated from *Clivia miniata* (identification of the biopolymers lignin and suberin). Plant Physiol. **113**: 1223–1231. doi:10.1104/pp.113.4.1223.

Zeier, J., and Schreiber, L. 1998. Comparative investigation of primary and tertiary endodermal cell walls isolated from the roots of five monocotyledoneous species: Chemical composition in relation to fine structure. Planta, **206**: 349–361. doi: 10.1007/s004250050410.

Zeier, J., Goll, A., Yokoyama, M., Karahara, I., and Schreiber, L. 1999. Structure and chemical composition of endodermal and rhizodermal/hypodermal walls of several species. Plant Cell Environ. **22**: 271–279. doi:10.1046/j.1365-3040.1999.00401.x.

Zentmyer, G. A. (1961). Chemotaxis of zoospores for root exudates. Science, **133**: 1595–1596. doi:10.1126/science.133.3464.1595.

APPENDICES

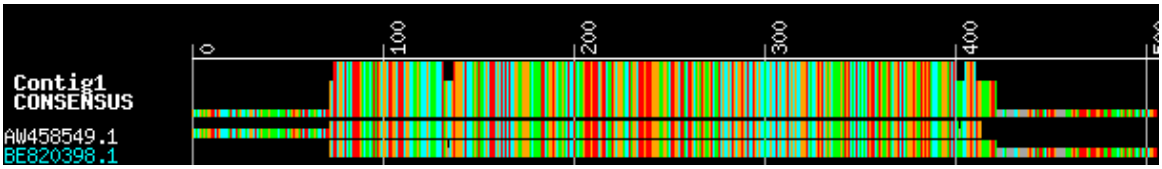
Appendix I: Characteristics of soybean pods and seeds used to designate M2 and L1 stages. Tissue was collected in the field by using the visual attributes of M2 and L1 stages. Once in the lab, five samples of each group were measured and weighed to verify their developmental stage. DAP = days after pollination. Samples were obtained from Agriculture and Agri-Food Canada (Southern Crop Protection and Food Research Centre, London, ON, Canada). Guidance for characterization was received from Agriculture Canada and Ranathunge et al. (2010).

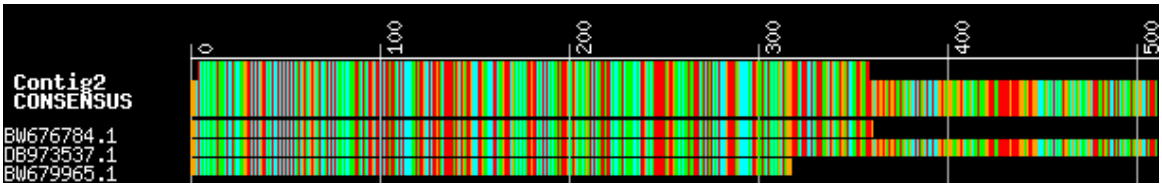
Stage	DAP (days)	Size (cm)	Whole weight (g)	Pod weight (g)	Number of seeds	Weight of seeds (g)	Average weight/ seed	Seed shape
M2	29-35	4.0	1.10	0.53	3	0.5730	0.1910	Flat
M2		4.0	0.92	0.43		0.4770	0.1590	
M2		3.8	1.07	0.47		0.5650	0.1883	
M2		3.9	1.12	0.49		0.5751	0.1917	
M2		4.0	0.95	0.41		0.5025	0.1675	
L1	35-43	3.6	1.38	0.46	3	0.8475	0.2825	Round
L1		4.0	1.47	0.56		0.8352	0.2784	
L1		3.7	1.37	0.52		0.8379	0.2793	
L1		3.6	1.32	0.49		0.8113	0.2704	
L1		3.6	1.37	0.51		0.8513	0.2838	

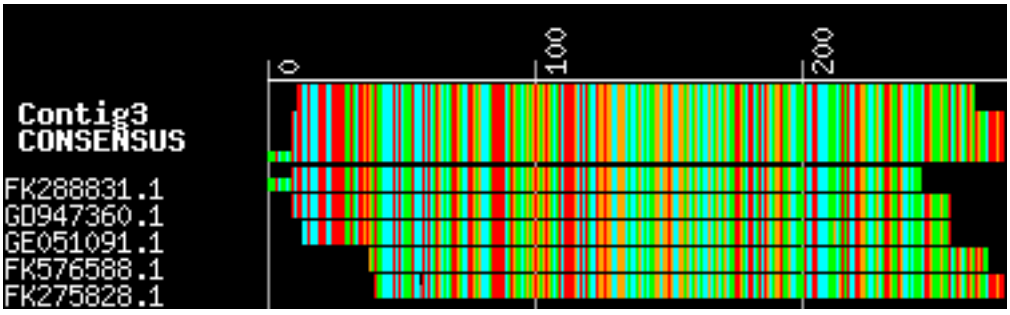
Appendix II: Extended table of RNA-Seq data. Generated from Soybase.org/Soyseq. DAF = days after flowering.

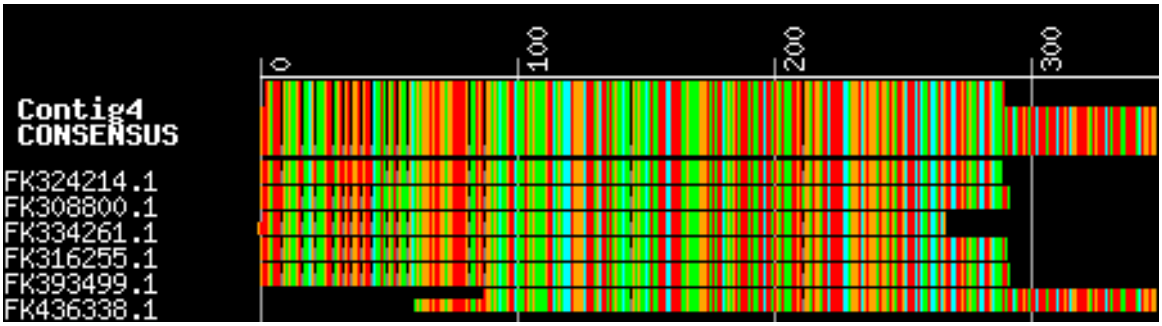
	Young Leaf	Flower	One cm Pod	Pod 10DAF	Pod 14DAF	Seed 10DAF	Seed 14DAF	Seed 21DAF	Seed 25DAF	Seed 28DAF	Seed 35DAF	Seed 42DAF	Root	Nodule
Gm FAwH-1	462	293	754	740	362	10	40	35	224	422	216	32	0	0
Gm FAwH-2	155	201	357	346	167	44	91	47	174	162	34	6	0	7
Gm FAwH-3	60	37	48	42	49	45	44	19	77	43	47	22	14	2
Gm FAwH-4	86	33	34	44	34	25	35	15	70	25	27	14	11	6
Gm FAwH-5	1	7	14	8	0	16	37	7	35	41	18	2	284	67
Gm FAwH-6	0	2	2	2	1	0	0	0	1	1	0	0	152	916

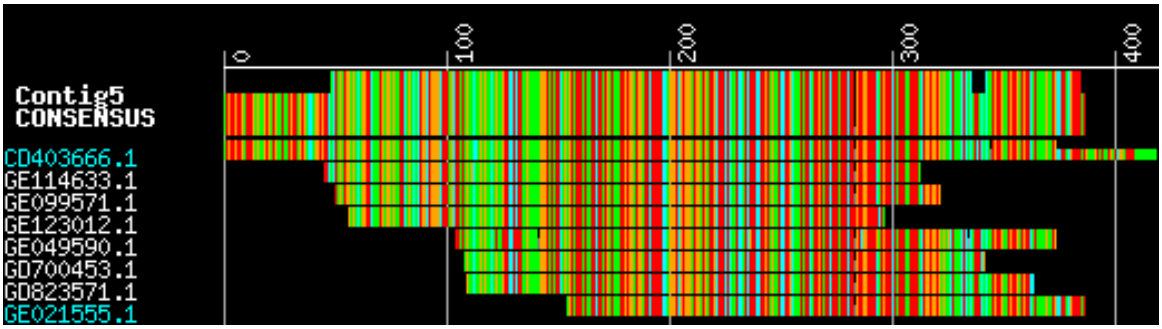
Appendix III: Contig maps of assembled soybean ESTs. The ESTs were assembled into contigs with the software Phrap (P. Green, University of Washington, Seattle, http:// www.phrap.org/). Images of the contigs were created with the contigimage software from the BioData suite (J. Freeman, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://www.ccgb.umn.edu/). The name of each EST is on the left (in cyan if it was reverse complemented) and the base pair position is on the X axis. Nucleotide bases are indicated as follows: A = green, C = cyan, G = orange, T = red, N = gray and * = dark gray. Data was obtained from Dr. Martina Stromvick (McGill University).

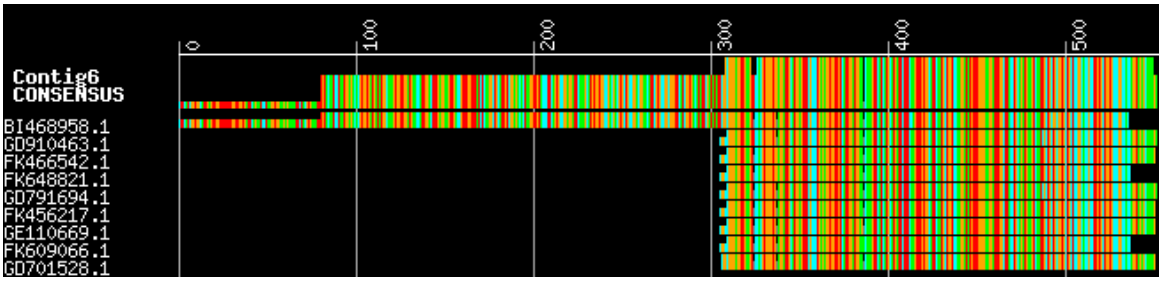


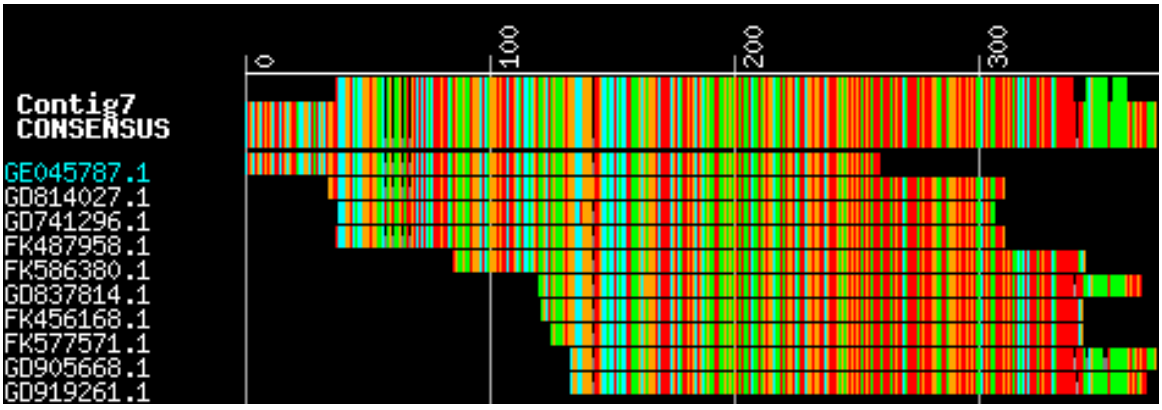


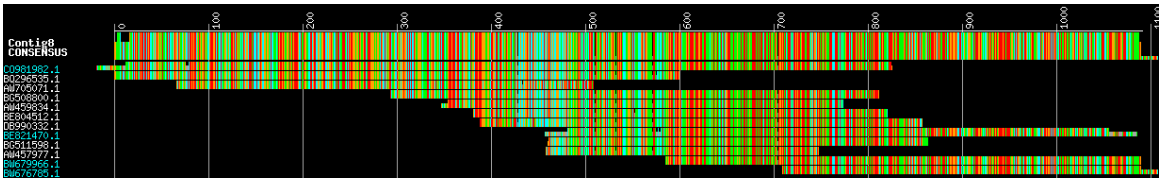


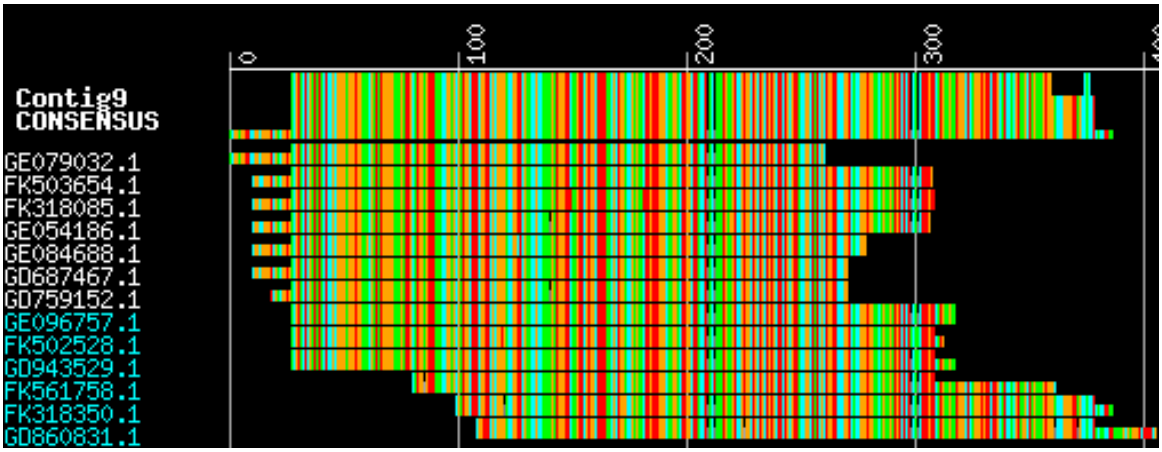

Contig4
CONSENSUS
FK324214.1
FK308800.1
FK334261.1
FK316255.1
FK393499.1
FK436338.1
0
100
200
300

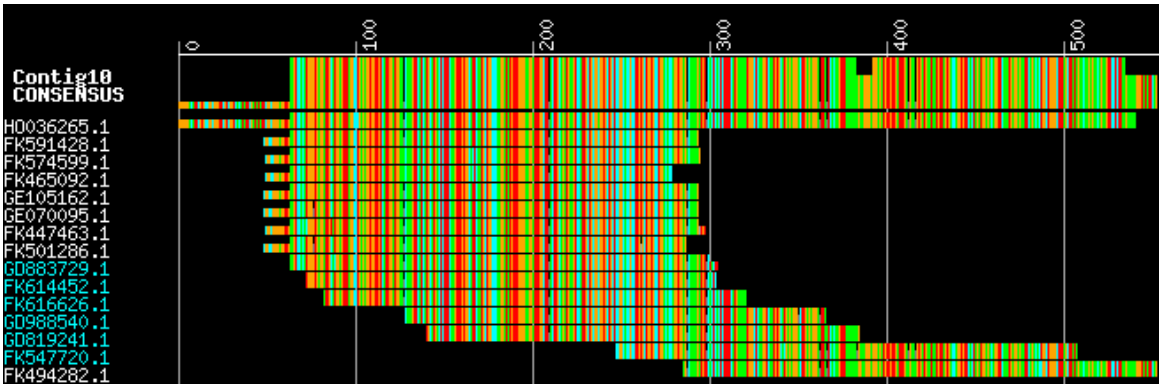

Contig5
CONSENSUS
CD403666.1
GE114633.1
GE099571.1
GE123012.1
GE049590.1
GD700453.1
GD823571.1
GE021555.1
0
100
200
300
400

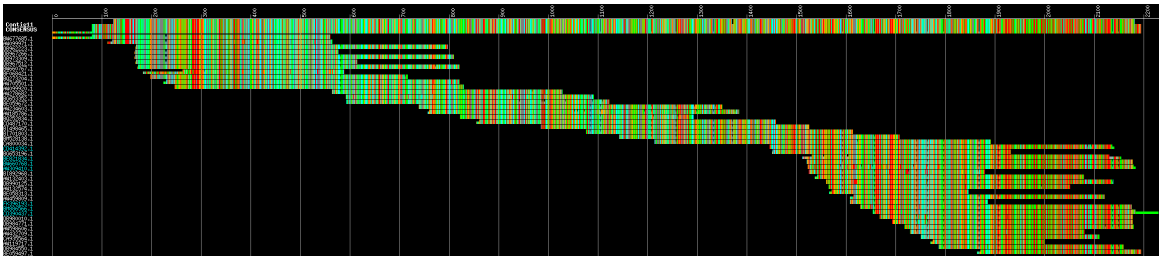

Contig6
CONSENSUS
BI468958.1
GD910463.1
FK466542.1
FK648821.1
GD791694.1
FK456217.1
GE110669.1
FK609066.1
GD701528.1
0
100
200
300
400
500


Contig7
CONSENSUS
GE045787.1
GD814027.1
GD741296.1
FK487958.1
FK586380.1
GD837814.1
FK456168.1
FK577571.1
GD905668.1
GD919261.1
0
100
200
300

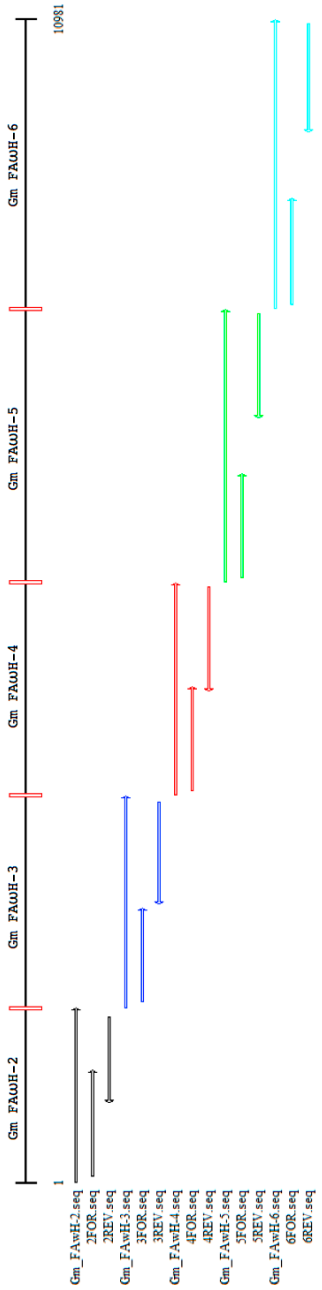

Contig8
CONSENSUS


Contig9
CONSENSUS
GE079032.1
FK503654.1
FK318085.1
GE054186.1
GE084688.1
GD687467.1
GD759152.1
GE096757.1
FK502528.1
GD943529.1
FK561758.1
FK318350.1
GD860831.1


Contig10
CONSENSUS
HO036265.1
FK591428.1
FK574599.1
FK465092.1
GE105162.1
GE070095.1
FK447463.1
FK501286.1
GD883729.1
FK614452.1
FK616626.1
GD988540.1
GD819241.1
FK547720.1
FK494282.1



Appendix IV: Sequence assembly of five of the putative *Gm FAωH*s and their corresponding sequencing results. Differences between sequence results and primary sequence were checked in the corresponding electropherogram and those nucleotides that were clearly misreads were changed and indicated with lowercase. The different colors indicate distinct contigs. Assembly performed in DNAMAN.



Group#1

```
Consensus   GTAATAGCAGCCTCTAGCTCCTCCGGTGAGGGAAGGTGCAAGCAATAAAGGTTTAATTTT    60
Gm_FAwH-2   GTAATAGCAGCCTCTAGCTCCTCCGGTGAGGGAAGGTGCAAGCAATAAAGGTTTAATTTT
2FOR                                                          ttTaATTTT
2REV

Consensus   TTGTTTTGGGAGTTGGGCCCAGCACTCACATGGACGCAGCAACGGCTTTAATGATCCTAT   120
Gm_FAwH-2   TTGTTTTGGGAGTTGGGCCCAGCACTCACATGGACGCAGCAACGGCTTTAATGATCCTAT
2FOR        TTGTTTTGGGAGTTGGGCCCAGCACTCACATGGACGCAGCAacggctTTAATGATCCTAT
2REV

Consensus   CAGCAATAGCGGCCTATTTAATATGGTTCACCTTGGTCACTCGCTCCCTCAAAGGTCCAC   180
Gm_FAwH-2   CAGCAATAGCGGCCTATTTAATATGGTTCACCTTGGTCACTCGCTCCCTCAAAGGTCCAC
2FOR        CAGCAATAGCGGCCTATTTAATATGGTTCACCTTGGTCACTCGCTCCCTCAAAGGTCCAC
2REV

Consensus   GTGTCTGGCCCCTATTGGGCAGCCTCCCTGGCCTCATCCAACACGCCAACCGCATGCACG   240
Gm_FAwH-2   GTGTCTGGCCCCTATTGGGCAGCCTCCCTGGCCTCATCCAACACGCCAACCGCATGCACG
2FOR        GTGTCTGGCCCCTATTGGGCAGCCTCCCTGGCCTCATCCAACACGCCAACCGCATGCACG
2REV

Consensus   ACTGGATCGCTGACAACCTCCGCGCGTGCGGCGGCACGTACCAGACCTGCATCTGCGCCC   300
Gm_FAwH-2   ACTGGATCGCTGACAACCTCCGCGCGTGCGGCGGCACGTACCAGACCTGCATCTGCGCCC
2FOR        ACTGGATCGCTGACAACCTCCGCGCGTGCGGCGGCACGTACCAGACCTGCATCTGCGCCC
2REV

Consensus   TCCCYTTCCTCGCCCGAAAACAGTGCCTGGTGACTGTCACGTGCGACCCCAAAAACCTCG   360
Gm_FAwH-2   TCCCCTTCCTCGCCCGAAAACAGTGCCTGGTGACTGTCACGTGCGACCCCAAAAACCTCG
2FOR        TCCCTTTCCTCGCCCGAAAACAGTGCCTGGTGACTGTCACGTGCGACCCCAAAAACCTCG
2REV

Consensus   AGCACATCCTCAAGCTTCGCTTCGACAACTACCCAAAAGGCCCCACGTGGCAGTCCGCCT   420
Gm_FAwH-2   AGCACATCCTCAAGCTTCGCTTCGACAACTACCCAAAAGGCCCCACGTGGCAGTCCGCCT
2FOR        AGCACATCCTCAAGCTTCGCTTCGACAACTACCCAAAAGGCCCCACGTGGCAGTCCGCCT
2REV

Consensus   TCCACGATTTGCTCGGCGAAGGCATCTTCAACTCCGACGGCGACACGTGGCTCTTCCAGC   480
Gm_FAwH-2   TCCACGATTTGCTCGGCGAAGGCATCTTCAACTCCGACGGCGACACGTGGCTCTTCCAGC
2FOR        TCCACGATTTGCTCGGCGAAGGCATCTTCAACTCCGACGGCGACACGTGGCTCTTCCAGC
2REV

Consensus   GTAAGACTGCCGCACTGGAATTCACCACTCGCACCCTGCGCCAAGCCATGGCGCGCTGGG   540
Gm_FAwH-2   GTAAGACTGCCGCACTGGAATTCACCACTCGCACCCTGCGCCAAGCCATGGCGCGCTGGG
2FOR        GTAAGACTGCCGCACTGGAATTCACCACTCGCACCCTGCGCCAAGCCATGGCGCGCTGGG
2REV

Consensus   TGAACCGAGCCATAAAGCACAGGTTCTGCCCCATCCTAGCCACGGCTCAGAAAGAAAACC   600
Gm_FAwH-2   TGAACCGAGCCATAAAGCACAGGTTCTGCCCCATCCTAGCCACGGCTCAGAAAGAAAACC
2FOR        TGAACCGAGCCATAAAGCACAGGTTCTGCCCCATCCTAGCCACGGCTCAGAAAGAAAACC
2REV

Consensus   AATCAGTGGATCTTCAAGACCTCTTGCTTCGGCTCACTTTTGACAACATATGCGGCTTGG   660
Gm_FAwH-2   AATCAGTGGATCTTCAAGACCTCTTGCTTCGGCTCACTTTTGACAACATATGCGGCTTGG
2FOR        AATCAGTGGATCTTCAAGACCTCTTGCTTCGGCTCACTTTTGACAACATATGCGGCTTGG
2REV
```

```
Consensus   CTTTCGGCCAAGACCCGCAAACCCTAGCCGCGGGGCTTCCAGATAACGCTTTCGCTCTCT   720
Gm_FAwH-2   CTTTCGGCCAAGACCCGCAAACCCTAGCCGCGGGGCTTCCAGATAACGCTTTCGCTCTCT
2FOR        CTTTCGGCCAAGACCCGCAAACCCTAGCCGCGGGGCTTCCAGATAACGCTTTCGCTCTCT
2REV

Consensus   CCTTCGACCGAGCCACTGAAGCCACTCTACAACGCTTTATTCTGCCGGAGATCTTGTGGA   780
Gm_FAwH-2   CCTTCGACCGAGCCACTGAAGCCACTCTACAACGCTTTATTCTGCCGGAGATCTTGTGGA
2FOR        CCTTCGACCGAGCCACTGAAGCCACTCTACAACGCTTTATTCTGCCGGAGATCTTGTGGA
2REV                                          GCTTtATTcTGCCGGAGATcTTGtGGA

Consensus   AGCTGAAGCGATGGCTCAGGCTCGGAATGGAAGTGAGCCTGAGCCGGAGCCTCAAACACA   840
Gm_FAwH-2   AGCTGAAGCGATGGCTCAGGCTCGGAATGGAAGTGAGCCTGAGCCGGAGCCTCAAACACA
2FOR        AGCTGAAGCGATGGCTCAGGCTCGGAATGgAAGTGAGCCTGAGCCGGAgCCTCAAACACA
2REV        AGCTGAAGcGAtGgcTCAGgCTCGGAATGGAAGTGAgCcTGAGCcGGAGCcTCAAACACA

Consensus   TCGACCAGTACCTCTCCCACATCATCAAGAACCGCAAGCTGGAGCTGCTGAATGGTAATG   900
Gm_FAwH-2   TCGACCAGTACCTCTCCCACATCATCAAGAACCGCAAGCTGGAGCTGCTGAATGGTAATG
2FOR        TCGACCAGTACCTCTCCCACATCATCAAGAACCGCAAGCTGGAGCTGCTGAATGGTAATG
2REV        TCGACCAGTACCTCTCcCACATCATCAAGAACCGCAAGCTGGAGCTGCTGAATGGTAATG

Consensus   GGTCCCACCACCACGACGACCTCTTGTCCCGCTTCATGAGGAAGAAAGAATCCTACTCGG   960
Gm_FAwH-2   GGTCCCACCACCACGACGACCTCTTGTCCCGCTTCATGAGGAAGAAAGAATCCTACTCGG
2FOR        GGTCCCACCACCACGACGACCTCTTGTCCCGCTTCATGAgGAAGAAAGAATCCTACTCGG
2REV        GGTCCCACCACCACGACGACCTCTTGTCCCGCTTCATGAGGAAGAAAGAATCCTACTCGG

Consensus   AGGAGTTCCTCCAACACGTGGCACTCAACTTCATCCTCGCTGGGCGCGACACGTCATCGG  1020
Gm_FAwH-2   AGGAGTTCCTCCAACACGTGGCACTCAACTTCATCCTCGCTGGGCGCGACACGTCATCGG
2FOR        AGGAGTTCCTCCaACACGtgGcaCTCAACTTCATCCTCGCTGGGCGCGACACGTCATCGG
2REV        AGGAGTTCCTCCAACACGTGGCACTCAACTTCATCCTCGCTGGGCGCGACACGTCATCGG

Consensus   TCGCGCTCAgCTGGTTCTTCTGGcTCTGCGTCAAGAACCCGCGCGtGGAGGAAAACATCC  1080
Gm_FAwH-2   TCGCGCTCAGCTGGTTCTTCTGGCTCTGCGTCAAGAACCCGCGCGTGGAGGAAAACATCC
2FOR        TCGCGCTCANCTGGttCTTCTGGNTCtGCGTCAAgAACCCgcgcGNg
2REV        TCGCGCTCAGCTGGTTCTTCTGGCTCTGCGTCAAGAACCCGCGCGTGGAGGAAAACATCC

Consensus   TAAACGAACTCTGTACTGTTCTTCTGTCCACGCGTGGAGATAACATCTCCACGTGGCTGA  1140
Gm_FAwH-2   TAAACGAACTCTGTACTGTTCTTCTGTCCACGCGTGGAGATAACATCTCCACGTGGCTGA
2FOR
2REV        TAAACGAACTCTGTACTGTTCTTCTGTCCACGCGTGGAGATAACATCTCCACGTGGCTGA

Consensus   ACGAGCCTCTTGTGTTCGATGAGGTAGACCGTTTGGTCTACCTCAAGGCAGCATTGTCTG  1200
Gm_FAwH-2   ACGAGCCTCTTGTGTTCGATGAGGTAGACCGTTTGGTCTACCTCAAGGCAGCATTGTCTG
2FOR
2REV        ACGAGCCTCTTGTGTTCGATGAGGTAGACCGTTTGGTCTACCTCAAGGCAGCATTGTCTG

Consensus   AGACACTGCGGCTTTATCCCTCCGTGCCGGAGGACTCCAAACATGTTGTGAAGGATGATG  1260
Gm_FAwH-2   AGACACTGCGGCTTTATCCCTCCGTGCCGGAGGACTCCAAACATGTTGTGAAGGATGATG
2FOR
2REV        AGACACTGCGGCTTTATCCCTCCGTGCCGGAGGACTCCAAACATGTTGTGAAGGATGATG

Consensus   TTTTGCCTAATGGAACCTTCGTTCCGGCAGGTTCCGCGGTGACCTATTCAATTTACTCTG  1320
Gm_FAwH-2   TTTTGCCTAATGGAACCTTCGTTCCGGCAGGTTCCGCGGTGACCTATTCAATTTACTCTG
2FOR
2REV        TTTTGCCTAATGGAACCTTCGTTCCGGCAGGTTCCGCGGTGACCTATTCAATTTACTCTG
```

```
Consensus   TTGGGAGGATGAAGTTCATTTGGGGAGAGGACTGCCTCGAGTTCAAGCCTGAGCGGTGGC  1380
Gm_FAwH-2   TTGGGAGGATGAAGTTCATTTGGGGAGAGGACTGCCTCGAGTTCAAGCCTGAGCGGTGGC
2FOR
2REV        TTGGGAGGATGAAGTTCATTTGGGGAGAGGACTGCCTCGAGTTCAAGCCTGAGCGGTGGC

Consensus   TCTCTCCTGAAGGGGACAAAATTCAGGTGCAAGACTCTTACAAGTTTGTTTCGTTCAATG  1440
Gm_FAwH-2   TCTCTCCTGAAGGGGACAAAATTCAGGTGCAAGACTCTTACAAGTTTGTTTCGTTCAATG
2FOR
2REV        TCTCTCCTGAAGGGGACAAAATTCAGGTGCAAGACTCTTACAAGTTTGTTTCGTTCAATG

Consensus   CGGGACCTAGGCTTTGTTTGGGGAAGGACTTGGCTTACTTGCAGATGAAGTCAATCGCCG  1500
Gm_FAwH-2   CGGGACCTAGGCTTTGTTTGGGGAAGGACTTGGCTTACTTGCAGATGAAGTCAATCGCCG
2FOR
2REV        CGGGACCTAGGCTTTGTTTGGGGAAGGACTTGGCTTACTTGCAGATGAAGTCAATCGCCG

Consensus   CCGCGGTGCTCCTGCGCCACCGCCTCGCTGTCGCGCCGGGCCACCGCGTCGAACAGAAGA  1560
Gm_FAwH-2   CCGCGGTGCTCCTGCGCCACCGCCTCGCTGTCGCGCCGGGCCACCGCGTCGAACAGAAGA
2FOR
2REV        CCGCGGTGCTCCTGCGCCACCGCCTCGCTGTCGCGCCGGGCCACCGCGTCGAaCAGAAGA

Consensus   TGTCACTCACACTGTTCATGAAGTACGGTCTCAAGGTGAATGTGTACCCAAGGGATCTCA  1620
Gm_FAwH-2   TGTCACTCACACTGTTCATGAAGTACGGTCTCAAGGTGAATGTGTACCCAAGGGATCTCA
2FOR
2REV        TGTCACTCACACt

Consensus   AACCCGTGCTGGAAAAATTAACCACCATCA                               1650
Gm_FAwH-2   AACCCGTGCTGGAAAAATTAACCACCATCA
2FOR
2REV
```

Group#2

```
Consensus   CATGGTGATGCATTGTCAAGAAACCAGTCTGGCCACAATAACTCTTCACTAAACCACTTC    60
Gm_FAwH-3   CATGGTGATGCATTGTCAAGAAACCAGTCTGGCCACAATAACTCTTCACTAAACCACTTC
3FOR                                                        AAACCaCTTC
3REV

Consensus   ACCTATAAATCTCTCTTCACCATTACACTCTTACCCTCATTATACTACTACTACTACACT   120
Gm_FAwH-3   ACCTATAAATCTCTCTTCACCATTACACTCTTACCCTCATTATACTACTACTACTACACT
3FOR        ACCTATAAATCTCTCTTCACCATTACACTCTTACCCTCATTATACTACTACTACTACACT
3REV

Consensus   CACTCATTCATTCACTCACCTTTCCATAAACACAACCTTCTTCACTTTGCTTTACTCCTT   180
Gm_FAwH-3   CACTCATTCATTCACTCACCTTTCCATAAACACAACCTTCTTCACTTTGCTTTACTCCTT
3FOR        CACTCATTCATTCACTCACCTTTCCATAAACACAACCTTCTTCACTTTGCTTTACTCCTT
3REV

Consensus   CATTCATTTTCCTCTCATACTCAAAACTCTGCATCAAAATGCTCATCATTGTTCACACCA   240
Gm_FAwH-3   CATTCATTTTCCTCTCATACTCAAAACTCTGCATCAAAATGCTCATCATTGTTCACACCA
3FOR        CATTCATTTTCCTCTCATACTCAAAACTCTGCATCAAAATGCTCATCATTGTTCACACCA
3REV

Consensus   CCCATTCTTGAAAGGTACTTCTTTTCTCGTTTTTCCAACCACCCCATTTCTAAAAAAGAC   300
Gm_FAwH-3   CCCATTCTTGAAAGGTACTTCTTTTCTCGTTTTTCCAACCACCCCATTTCTAAAAAAGAC
3FOR        CCCATTCTTGAAAGGTACTTCTTTTCTCGTTTTTCCAACCACCCCATTTCTAAAAAAGAC
3REV
```

```
Consensus   ATTTTCCCTTCAAGTTTCTTCATCATTTCTTGTAGTTTCGAGTCTTTGATTATTGGGGTG   360
Gm_FAwH-3   ATTTTCCCTTCAAGTTTCTTCATCATTTCTTGTAGTTTCGAGTCTTTGATTATTGGGGTG
3FOR        ATTTTCCCTTCAAGTTTCTTCATCATTTCTTGTAGTTTCGAGTCTTTGATTATTGGGGTG
3REV

Consensus   GTTAATATTGCACTTGCACTATGGAAACTTGCACGGCTCTATTGTTTTTAACGGCGATCA   420
Gm_FAwH-3   GTTAATATTGCACTTGCACTATGGAAACTTGCACGGCTCTATTGTTTTTAACGGCGATCA
3FOR        GTTAATATTGCACTTGCACTATGGAAACTTGCACGGCTCTATTGTTTTTAACGGCGATCA
3REV

Consensus   CAGCATACTTGATATGGTTCACGTTCATCTCACGGTCGCTGAAGGGTCCACGTGTCTGGG   480
Gm_FAwH-3   CAGCATACTTGATATGGTTCACGTTCATCTCACGGTCGCTGAAGGGTCCACGTGTCTGGG
3FOR        CAGCATACTTGATATGGTTCACGTTCATCTCACGGTCGCTGAAGGGTCCACGTGTCTGGG
3REV

Consensus   CCTTATTGGGCAGTCTCCCAGGCCTCATAGACAACTGTGACCGCATGCATGACTGGATCT   540
Gm_FAwH-3   CCTTATTGGGCAGTCTCCCAGGCCTCATAGACAACTGTGACCGCATGCATGACTGGATCT
3FOR        CCTTATTGGGCAGTCTCCCAGGCCTCATAGACAACTGTGACCGCATGCATGACTGGATCT
3REV

Consensus   GCGACAACCTACGCGCGTGTGGCGGCACGTACCAGACCTGCATCTGTGCAATCCCCTTCC   600
Gm_FAwH-3   GCGACAACCTACGCGCGTGTGGCGGCACGTACCAGACCTGCATCTGTGCAATCCCCTTCC
3FOR        GCGACAACCTACGCGCGTGTGGCGGCACGTACCAGACCTGCATCTGTGCAATCCCCTTCC
3REV

Consensus   TCGCCAAGAAGCAGGGTCTCGTGACTGTCACGTGCGACCCGAGGAACTTGGAGCACATAC   660
Gm_FAwH-3   TCGCCAAGAAGCAGGGTCTCGTGACTGTCACGTGCGACCCGAGGAACTTGGAGCACATAC
3FOR        TCGCCAAGAAGCAGGGTCTCGTGACTGTCACGTGCGACCCGAGGAACTTGGAGCACATAC
3REV

Consensus   TCAAGACACGCTTCGACAATTACCCTAAAGGACCCACGTGGCATGCTGTCTTTCATGATC   720
Gm_FAwH-3   TCAAGACACGCTTCGACAATTACCCTAAAGGACCCACGTGGCATGCTGTCTTTCATGATC
3FOR        TCAAGACACGCTTCGACAATTACCCTAAAGGACCCACGTGGCATGCTGTCTTTCATGATC
3REV

Consensus   TGTTGGGTGATGGGATCTTTAACACCGATGGTGACACATGGCTGTTCCAGCGCAAGACAG   780
Gm_FAwH-3   TGTTGGGTGATGGGATCTTTAACACCGATGGTGACACATGGCTGTTCCAGCGCAAGACAG
3FOR        TGTTGGGTGATGGGATCTTTAACACCGATGGTGACACATGGCTGTTCCAGCGCAAGACAG
3REV

Consensus   CTGCGCTGGAATTTACCACCCGGACGCTGCGCCAAGCCATGGCCAGGTGGGTGAGTCGAG   840
Gm_FAwH-3   CTGCGCTGGAATTTACCACCCGGACGCTGCGCCAAGCCATGGCCAGGTGGGTGAGTCGAG
3FOR        CTGCGCTGGAATTTACCACCCGGACGCTGCGCCAAGCCATGGcCAGGTGGGTGAGTCGAG
3REV

Consensus   CCATCAACCGGCTCTGTCTGATTCTCGAAAAAGCCGAGAATCAAGTCcAGCCGGTTGATC   900
Gm_FAwH-3   CCATCAACCGGCTCTGTCTGATTCTCGAAAAAGCCGAGAATCAAGTCGAGCCGGTTGATC
3FOR        CCATCAACCGGcTCTGTCTGAtTCTCgaAAAAGCCgAGAATCAAGTCNAGCCGGTTGAtC
3REV

Consensus   TGCAAGACTTAATGCTTCaGCTCAtTTTTGATAATATTTGTGGGTTGGaTTTCGGGCGAG   960
Gm_FAwH-3   TGCAAGACTTAATGCTTCGGCTCACTTTTGATAATATTTGTGGGTTGGCTTTCGGGCGAG
3FOR        TGCAAgActTAAtGctTCNGCTCANTTTTgatAAtatTTGtGGGTtGGNTTT
3REV
```

```
Consensus   ACCCACAGACTTGTGTGTCAAGTCTGCCCGATAACCGATTTGgCACGGCTTTCGATCGAG  1020
Gm_FAwH-3   ACCCACAGACTTGTGTGTCAAGTCTGCCCGATAACCGATTTGCCACGGCTTTCGATCGAG
3FOR
3REV                          cAagtcTgCCCGATaAccGATTTg.CaCGGCTTTCGATCGAG

Consensus   CCACTGAAGCCACGCTCCAACGATTCATTTTACCTGAGGTATTGTGGAAGGTGAAAAAAT  1080
Gm_FAwH-3   CCACTGAAGCCACGCTCCAACGATTCATTTTACCTGAGGTATTGTGGAAGGTGAAAAAAT
3FOR
3REV        CCACTGAAgccaCGcTCCAACGATTCATTTTACcTGAGgTATTGTGGAAGGTGAAAAAAT

Consensus   GGCTTCGGCTTGGGATGGAAGTCAGCTTGAGCCGAAGCCTTGCCCACGTGGACGACCATT  1140
Gm_FAwH-3   GGCTTCGGCTTGGGATGGAAGTCAGCTTGAGCCGAAGCCTTGCCCACGTGGACGACCATT
3FOR
3REV        GGCTTCGGCtTGGGATGGAAGTCAGCTTGAGCCGAAGCCTTGCCCACGTGGACGACCATT

Consensus   TGTCAAATGTGATTGAGAAACGCAAGGTAGAGTTGTTGACTCAGCAGAAAGATGGGACTC  1200
Gm_FAwH-3   TGTCAAATGTGATTGAGAAACGCAAGGTAGAGTTGTTGACTCAGCAGAAAGATGGGACTC
3FOR
3REV        TGTCAAATGTGATTGAGAAACGCAAGGTAGAGTTGTTGACTCAGCAGAAAGATGGGACTC

Consensus   TTCATGATGACTTGTTGACAAGGTTTATGAGGAAAAAAGAATCCTATTCAGACAAGTTTC  1260
Gm_FAwH-3   TTCATGATGACTTGTTGACAAGGTTTATGAGGAAAAAAGAATCCTATTCAGACAAGTTTC
3FOR
3REV        TTCATGATGACTTGTTGACAAGGTTTATGAGGAAAAAAGAATCCTATTCAGACAAGTTTC

Consensus   TCCAACAAGTGGCGTTGAATTTTATCCTAGCTGGTCGTGACACCTCATCGGTGGCATTAA  1320
Gm_FAwH-3   TCCAACAAGTGGCGTTGAATTTTATCCTAGCTGGTCGTGACACCTCATCGGTGGCATTAA
3FOR
3REV        TCCAACAAGTGGCGTTGAATTTTATCCTAGCTGGTCGTGACACCTCATCGGTGGCATTAA

Consensus   GTTGGTTTTTTTGGTTGGTGATTCAGAACCCTAAGGTGGAAGAGAAAATTCTACGTGAAA  1380
Gm_FAwH-3   GTTGGTTTTTTTGGTTGGTGATTCAGAACCCTAAGGTGGAAGAGAAAATTCTACGTGAAA
3FOR
3REV        GTTGGTTTTTTTGGTTGGTGATTCAGAACCCTAAGGTGGAAGAGAAAATTCTACGTGAAA

Consensus   TTTGTACAGTCCTGATGGAGACACGTGGCAATGATGACATGGCAAAGTTGTTTGATGAGC  1440
Gm_FAwH-3   TTTGTACAGTCCTGATGGAGACACGTGGCAATGATGACATGGCAAAGTTGTTTGATGAGC
3FOR
3REV        TTTGTACAGTCCTGATGGAGACACGTGGCAATGATGACATGGCAAAGTTGTTTGATGAGC

Consensus   CTTTGGCCTTTGAGGAAGTTGACCGTTTGGTTTATCTCAAGGCTGCATTGTCGGAGACAC  1500
Gm_FAwH-3   CTTTGGCCTTTGAGGAAGTTGACCGTTTGGTTTATCTCAAGGCTGCATTGTCGGAGACAC
3FOR
3REV        CTTTGGCCTTTGAGGAAGTTGACCGTTTGGTTTATCTCAAGGCTGCATTGTCGGAGACAC

Consensus   TAAGGTTGTACCCTTCAGTGCCGGAGGACTCAAAGCATGTGGTGGCCGACGACGTGCTGC  1560
Gm_FAwH-3   TAAGGTTGTACCCTTCAGTGCCGGAGGACTCAAAGCATGTGGTGGCCGACGACGTGCTGC
3FOR
3REV        TAAGGTTGTACCCTTCAGTGCCGGAGGACTCAAAGCATGTGGTGGCCGACGACGTGCTGC

Consensus   CGGACGGGACATTCGTGCCGGCGGGTTCGTCGGTCACATATTCGATATATTCGGCGGGGA  1620
Gm_FAwH-3   CGGACGGGACATTCGTGCCGGCGGGTTCGTCGGTCACATATTCGATATATTCGGCGGGGA
3FOR
3REV        CGGACGGGACATTCGTGCCGGCGGGTTCGTCGGTCACATATTCGATATATTCGGCGGGGA
```

```
Consensus   GGTTGAAGTCCACATGGGGTGAGGATTGCATGGAGTTTAGGCCTGAGAGGTGGTTGTCAT  1680
Gm_FAwH-3   GGTTGAAGTCCACATGGGGTGAGGATTGCATGGAGTTTAGGCCTGAGAGGTGGTTGTCAT
3FOR
3REV        GGTTGAAGTCCACATGGGGTGAGGATTGCATGGAGTTTAGGCCTGAGAGGTGGTTGTCAT

Consensus   TGGATGGAACAAAGTTCATCATGCATGACTCTTTCAAGTTTGTGGCATTCAATGCTGGCC  1740
Gm_FAwH-3   TGGATGGAACAAAGTTCATCATGCATGACTCTTTCAAGTTTGTGGCATTCAATGCTGGCC
3FOR
3REV        TGGATGGAACAAAGTTCATCATGCATGACTCTTTCAAGTTTGTGGCATTCAATGCTGGCC

Consensus   CAAGGATATGTTTGGGGAAGGATTTGGCTTACTTGCAGATGAAGTCCATTGCTGCGGCGG  1800
Gm_FAwH-3   CAAGGATATGTTTGGGGAAGGATTTGGCTTACTTGCAGATGAAGTCCATTGCTGCGGCGG
3FOR
3REV        CAAGGATATGTTTGGGGAAGGATTTGGCTTACTTGCAGATGAAGTCCATTGCTGCGGCGG

Consensus   TGCTGCTCCGGCACCGCCTTGTGCTGGTGCCTGGACACCAGGTGGAGCAAAAGATGTCAC  1860
Gm_FAwH-3   TGCTGCTCCGGCACCGCCTTGTGCTGGTGCCTGGACACCAGGTGGAGCAAAAGATGTCAC
3FOR
3REV        TGCTGCTCCGGCACCGCCTTGTGCTGGTGCCTGGACACCAGGTGGAGCAAAAGATGTCAC

Consensus   TCACTCTTTTCATGAAAAATGGGCTCAAGGTCAATGTGCATGAGAGGGACTTGAGAGGGA  1920
Gm_FAwH-3   TCACTCTTTTCATGAAAAATGGGCTCAAGGTCAATGTGCATGAGAGGGACTTGAGAGGGA
3FOR
3REV        TCACTCTTTTCATGAAAAATGGGCTCAAGGTCAATGTGCATGAGAGGGACTTGAGAGGGA

Consensus   TTATCACAAGTTTAAAAAAAGAAAGGGAGGGAGATGTTGTTGATTTGAGAAGTGATGAAT  1980
Gm_FAwH-3   TTATCACAAGTTTAAAAAAAGAAAGGGAGGGAGATGTTGTTGATTTGAGAAGTGATGAAT
3FOR
3REV        TTATCACAAGTTTAAAAAAAGAAAGGGAGGGAGA

Consensus   AATGTTAGTGACCAGGTGTTCACTAGATGG                               2010
Gm_FAwH-3   AATGTTAGTGACCAGGTGTTCACTAGATGG
3FOR
3REV
```

Group#3

```
Consensus   ATCTGATGAGAGAGAAAAGGGTGGTGCATTGTCAAGAAACCAGTCTGGCCACAATAAATC    60
Gm_FAwH-4   ATCTGATGAGAGAGAAAAGGGTGGTGCATTGTCAAGAAACCAGTCTGGCCACAATAAATC
4FOR                                                aGTCTGGCCaCaaTAAATC
4REV

Consensus   TTCACTAAACAACTTCACCTATAAATCTCTCTTCACCATTACACACTTACCCTCATTGTA   120
Gm_FAwH-4   TTCACTAAACAACTTCACCTATAAATCTCTCTTCACCATTACACACTTACCCTCATTGTA
4FOR        TTCaCTAAACAACTTCACCTATAAATCTCTCTTCACCATTACACACTTACCCTCATTGTA
4REV

Consensus   CTACTACTACACTCACTCATTCACCTTTACATAAACAAACACAACCTTCTTCACTTTCCT   180
Gm_FAwH-4   CTACTACTACACTCACTCATTCACCTTTACATAAACAAACACAACCTTCTTCACTTTCCT
4FOR        CTACTACTACACTCACTCATTCACCTTTACATAAACAAACACAACCTTCTTCACTTTCCT
4REV

Consensus   TCACTTCTCCATTTTCCTCTCTTTCAATCACCCCATATTCAAAAATGCTCCTCATTGTTC   240
Gm_FAwH-4   TCACTTCTCCATTTTCCTCTCTTTCAATCACCCCATATTCAAAAATGCTCCTCATTGTTC
4FOR        TCACTTCTCCATTTTCCTCTCTTTCAATCACCCCATATTCAAAAATGCTCCTCATTGTTC
4REV
```

```
Consensus   ACACCCATTCTTGAAAGGTACCTGTTTTTCCAACCACCCCATTTCTTAAAAAAACTATTT   300
Gm_FAwH-4   ACACCCATTCTTGAAAGGTACCTGTTTTTCCAACCACCCCATTTCTTAAAAAAACTATTT
4FOR        ACACCCATTCTTGAAAGGTACCTGTTTTTCCAACCACCCCATTTCTTAAAAAAACTATTT
4REV

Consensus   TTCCTTCAAGTTTCTTCGTCATTTCTTGTAGTTTTGAGTCTTTGGTTATTGGGGTGGTTG   360
Gm_FAwH-4   TTCCTTCAAGTTTCTTCGTCATTTCTTGTAGTTTTGAGTCTTTGGTTATTGGGGTGGTTG
4FOR        TTCCTTCAAGTTTCTTCGTCATTTCTTGTAGTTTTGAGTCTTTGGTTATTGGGGTGGTTG
4REV

Consensus   CACTTGCACTATGGAAACTTGCACGGCTCTGTTGTTTTTAACGGCCATCACAGCATACTT   420
Gm_FAwH-4   CACTTGCACTATGGAAACTTGCACGGCTCTGTTGTTTTTAACGGCCATCACAGCATACTT
4FOR        CACTTGCACTATGGAAACTTGCACGGCTCTGTTGTTTTTAACGGCCATCACAGCATACTT
4REV

Consensus   GATATGGTTCACGTTCATCTCACGGTCGCTGAAGGGTCCACGTGTCTGGGCCTTATTGGG   480
Gm_FAwH-4   GATATGGTTCACGTTCATCTCACGGTCGCTGAAGGGTCCACGTGTCTGGGCCTTATTGGG
4FOR        GATATGGTTCACGTTCATCTCACGGTCGCTGAAGGGTCCACGTGTCTGGGCCTTATTGGG
4REV

Consensus   CAGTCTCCCGGGCCTCATTGACAACTGTGACCGCATGCATGACTGGATCTGCGACAACCT   540
Gm_FAwH-4   CAGTCTCCCGGGCCTCATTGACAACTGTGACCGCATGCATGACTGGATCTGCGACAACCT
4FOR        CAGTCTCCCGGGCCTCATTGACAACTGTGACCGCATGCATGACTGGATCTGCGACAACCT
4REV

Consensus   ACGCGCGTGTGGCGGCACGTACCAGACATGCATCTGTGCAATCCCCTTCCTCGCCAAGAA   600
Gm_FAwH-4   ACGCGCGTGTGGCGGCACGTACCAGACATGCATCTGTGCAATCCCCTTCCTCGCCAAGAA
4FOR        ACGCGCGTGTGGCGGCACGTACCAGACATGCATCTGTGCAATCCCCTTCCTCGCCAAGAA
4REV

Consensus   GCAGGGTCTCGTGACTGTCACGTGCGACCCGAGGAACTTAGAGCACATACTCAAGACACG   660
Gm_FAwH-4   GCAGGGTCTCGTGACTGTCACGTGCGACCCGAGGAACTTAGAGCACATACTCAAGACACG
4FOR        GCAGGGTCTCGTGACTGTCACGTGCGACCCGAGGAACTTAGAGCACATACTCAAGACACG
4REV

Consensus   CTTCGACAATTACCCTAAAGGGCCCACGTGGCATGCTGTCTTTCATGATCTGTTGGGTGA   720
Gm_FAwH-4   CTTCGACAATTACCCTAAAGGGCCCACGTGGCATGCTGTCTTTCATGATCTGTTGGGTGA
4FOR        CTTCGACAATTACCCTAAAGGGCCCACGTGGCATGCTGTCTTTCATGATCTGTTGGGTGA
4REV

Consensus   TGGGATCTTTAACACCGACGGTGACACGTGGCTGTTCCAGCGCAAGACAGCTGCGCTGGA   780
Gm_FAwH-4   TGGGATCTTTAACACCGACGGTGACACGTGGCTGTTCCAGCGCAAGACAGCTGCGCTGGA
4FOR        TGGGATCTTTAACACCGACGGTGACACGTGGCTGTTCCAGCGCAAGACAGCTGCGCTGGA
4REV

Consensus   ATTCACTACCCGAACGCTGCGCCAAGCCATGGCTCGGTGGGTGAGCCGAGCTATCAACCG   840
Gm_FAwH-4   ATTCACTACCCGAACGCTGCGCCAAGCCATGGCTCGGTGGGTGAGCCGAGCTATCAACCG
4FOR        ATTCACTACCCGAACGCTGCGCCAAGCCATGGCTCGGTGGGTGAGCCGAGCTATCAACCG
4REV

Consensus   GCTCTGCTTGATTCTTAAGAAAGCTAAAGATCAAGCAGAGCCGGTTGATCTTCAAGATTT   900
Gm_FAwH-4   GCTCTGCTTGATTCTTAAGAAAGCTAAAGATCAAGCAGAGCCGGTTGATCTTCAAGATTT
4FOR        GCTCTGCTTGAtTCTTAAGAAAGCTAAAGATCAAGCAgAgCCGGTTGATCTTCAAGATTT
4REV
```

```
Consensus   GATGCTTCGGCTCACTTTTGATAACATTTGTGGTCTCGCTTTTGGGCGAGACCCACAAAC   960
Gm_FAwH-4   GATGCTTCGGCTCACTTTTGATAACATTTGTGGTCTCGCTTTTGGGCGAGACCCACAAAC
4FOR        GATGCTTCGGCTCACTTTTGAtAACaTTTGTGGTCTCGCTTTTGGGCgAGAcCCACAAAC
4REV

Consensus   TTGTGTGTTGGGTCTGTCAGATAACCGATTTGCtACGGCTTTTGACCGAGCCACCGAAGC  1020
Gm_FAwH-4   TTGTGTGTTGGGTCTGTCAGATAACCGATTTGCCACGGCTTTTGACCGAGCCACCGAAGC
4FOR        TTGTGTGTTGGGTCTGTCAgAtAa.CGATTTGCNaCGGCTTTtGACCGAGCCAcCgAAgC
4REV                          ATaACCGATTTgCCACGGCTTTTGACCGAGCCaccgAagC

Consensus   CACGTTACAACGGTTCATTTTGCCTGAGGTGTTGTGGAAGGTTAAGAAATGGCTTCGGCT  1080
Gm_FAwH-4   CACGTTACAACGGTTCATTTTGCCTGAGGTGTTGTGGAAGGTTAAGAAATGGCTTCGGCT
4FOR        Ca
4REV        CaCGTTACaACGGTTCATTTTGCcTGAGGTGTTGTGgAAGGTTAAGAAATGGCTTCGGCT

Consensus   TGGACTGGAAGTGAGCTTGAGCCGAAGCCTTGTCCATGTGGAGGACCATTTGTCAAATGT  1140
Gm_FAwH-4   TGGACTGGAAGTGAGCTTGAGCCGAAGCCTTGTCCATGTGGAGGACCATTTGTCAAATGT
4FOR
4REV        TGGACTGGAAGTGAGCTTGAGCCGAAGCCTTGTCCATGTGGAGGACCATTTGTCAAATGT

Consensus   TATTGAGAAACGCAAGGTGGAGTTGTTGAGTCAACAAAAAGATGGTACTCTTCATGATGA  1200
Gm_FAwH-4   TATTGAGAAACGCAAGGTGGAGTTGTTGAGTCAACAAAAAGATGGTACTCTTCATGATGA
4FOR
4REV        TATTGAGAAACGCAAGGTGGAGTTGTTGAGTCAACAAAAAGATGGTACTCTTCATGATGA

Consensus   CTTGTTGACTAGGTTTATGAAGAAAAAGGAATCCTACACGGACAAGTTTCTCCAACATGT  1260
Gm_FAwH-4   CTTGTTGACTAGGTTTATGAAGAAAAAGGAATCCTACACGGACAAGTTTCTCCAACATGT
4FOR
4REV        CTTGTTGACTAGGTTTATGAAGAAAAAGGAATCCTACACGGACAAGTTTCTCCAACATGT

Consensus   GGCGTTGAATTTTATCCTAGCTGGACGTGACACTTCATCGGTGGCATTGAGTTGGTTTTT  1320
Gm_FAwH-4   GGCGTTGAATTTTATCCTAGCTGGACGTGACACTTCATCGGTGGCATTGAGTTGGTTTTT
4FOR
4REV        GGCGTTGAATTTTATCCTAGCTGGACGTGACACTTCATCGGTGGCATTGAGTTGGTTTTT

Consensus   TTGGTTGGTGATTCAGAATCCTAAGGTGGAAGAGAAAATTCTACGTGAAATTTGTACAAT  1380
Gm_FAwH-4   TTGGTTGGTGATTCAGAATCCTAAGGTGGAAGAGAAAATTCTACGTGAAATTTGTACAAT
4FOR
4REV        TTGGTTGGTGATTCAGAATCCTAAGGTGGAAGAGAAAATTCTACGTGAAATTTGTACAAT

Consensus   CCTAATGGAGACACGTGGTGATGATATGGCAAAGTGGTTGGATGAACCATTGGACTTTGA  1440
Gm_FAwH-4   CCTAATGGAGACACGTGGTGATGATATGGCAAAGTGGTTGGATGAACCATTGGACTTTGA
4FOR
4REV        CCTAATGGAGACACGTGGTGATGATATGGCAAAGTGGTTGGATGAACCATTGGACTTTGA

Consensus   GGAAGTTGACCGTTTGGTTTATCTCAAGGCTGCATTGTCGGAGACACTAAGGTTGTACCC  1500
Gm_FAwH-4   GGAAGTTGACCGTTTGGTTTATCTCAAGGCTGCATTGTCGGAGACACTAAGGTTGTACCC
4FOR
4REV        GGAAGTTGACCGTTTGGTTTATCTCAAGGCTGCATTGTCGGAGACACTAAGGTTGTACCC

Consensus   TTCAGTGCCGGAGGATTCAAAGCATGTGGTGGCCGACGACGTGCTGCCGGACGGAACATT  1560
Gm_FAwH-4   TTCAGTGCCGGAGGATTCAAAGCATGTGGTGGCCGACGACGTGCTGCCGGACGGAACATT
4FOR
4REV        TTCAGTGCCGGAGGATTCAAAGCATGTGGTGGCCGACGACGTGCTGCCGGACGGAACATT
```

```
Consensus   CGTGCCGGCGGGTTCATCGGTGACATATTCGATATATTCGGCGGGGAGGTTGAAGTCCAC  1620
Gm_FAwH-4   CGTGCCGGCGGGTTCATCGGTGACATATTCGATATATTCGGCGGGGAGGTTGAAGTCCAC
4FOR
4REV        CGTGCCGGCGGGTTCATCGGTGACATATTCGATATATTCGGCGGGGAGGTTGAAGTCCAC

Consensus   GTGGGGTGAGGATTGCATGGAGTTTCGCCCTGAGAGGTGGTTGTCATTGGATGGAACAAA  1680
Gm_FAwH-4   GTGGGGTGAGGATTGCATGGAGTTTCGCCCTGAGAGGTGGTTGTCATTGGATGGAACAAA
4FOR
4REV        GTGGGGTGAGGATTGCATGGAGTTTCGCCCTGAGAGGTGGTTGTCATTGGATGGAACAAA

Consensus   GTTCATCATGCATGATTCTTTCAAGTTTGTGGCATTCAATGCTGGTCCAAGGATATGTTT  1740
Gm_FAwH-4   GTTCATCATGCATGATTCTTTCAAGTTTGTGGCATTCAATGCTGGTCCAAGGATATGTTT
4FOR
4REV        GTTCATCATGCATGATTCTTTCAAGTTTGTGGCATTCAATGCTGGTCCAAGGATATGTTT

Consensus   GGGGAAGGATTTGGCTTACCTGCAGATGAAGTCAATTGCTGCTGCGGTGCTGCTCCGGCA  1800
Gm_FAwH-4   GGGGAAGGATTTGGCTTACCTGCAGATGAAGTCAATTGCTGCTGCGGTGCTGCTCCGGCA
4FOR
4REV        GGGGAAGGATTTGGCTTACCTGCAGATGAAGTCAATTGCTGCTGCGGTGCTGCTCCGGCA

Consensus   CCGCCTTGTGCTGGTGCCCGGCCACCAGGTAGAGCAAAAGATGTCACTCACTCTGTTCAT  1860
Gm_FAwH-4   CCGCCTTGTGCTGGTGCCCGGCCACCAGGTAGAGCAAAAGATGTCACTCACTCTGTTCAT
4FOR
4REV        CCGCCTTGTGCTGGTGCCCGGCCACCAGGTAGAGCAAAAGATGTCACTCACTCTGTTCAT

Consensus   GAAAAATGGGCTCAAGGTCAATGTGCATGAGAGGGATTTGAGAGGGGTTATCACAAGTAT  1920
Gm_FAwH-4   GAAAAATGGGCTCAAGGTCAATGTGCATGAGAGGGATTTGAGAGGGGTTATCACAAGTAT
4FOR
4REV        GAAAAATGGGCTCAAGGTCAATGTGCATGAGAGGGATTTGAGAGGGGTTATCACAAGTAT

Consensus   TAAAAAGGAAAGGGAGGAAGATGTTGATTTGAGAAGTAACGAAAGTTAGTGTGGTGTTCA  1980
Gm_FAwH-4   TAAAAAGGAAAGGGAGGAAGATGTTGATTTGAGAAGTAACGAAAGTTAGTGTGGTGTTCA
4FOR
4REV        TAAAAAGGAAAGGGAGGAAGATGTTGAtTTGAGAAGTAACGAAA

Consensus   ATAGATGAAGCAGAGGGTGTTGGA                                      2004
Gm_FAwH-4   ATAGATGAAGCAGAGGGTGTTGGA
4FOR
4REV
```

Group#4

```
Consensus   GCATGTCCACCTAACGTAGACCACTTTTGTTTTGCCAAAAAATTCCAACCAACTCTCTCC    60
Gm_FAwH-5   GCATGTCCACCTAACGTAGACCACTTTTGTTTTGCCAAAAAATTCCAACCAACTCTCTCC
5FOR                                                aatTCcaACCAACTCTCTCC
5REV

Consensus   TAAACCTTCTCACAATCTCATTCCACATAAATATATAAAACGTACCAAGGATGTTCCTCA   120
Gm_FAwH-5   TAAACCTTCTCACAATCTCATTCCACATAAATATATAAAACGTACCAAGGATGTTCCTCA
5FOR        TAAACCTTCTCACAATCTCATTCCACATAAATATATAAAACGTACCAAGGATGTTCCTCA
5REV

Consensus   TTCTAACCAAGTTAATCATCAACGTGTCTTACAACACCATGATGATGCAAATGGAAACCC   180
Gm_FAwH-5   TTCTAACCAAGTTAATCATCAACGTGTCTTACAACACCATGATGATGCAAATGGAAACCC
5FOR        TTCTAACCAAGTTAATCATCAACGTGTCTTACAACACCATGATGATGCAAATGGAAACCC
5REV
```

```
Consensus   TCCCTCTCCTCCTCACTCTAGTAGCAACCCTATCAGCCTACTTCCTTTGGTTCCACCTCC   240
Gm_FAwH-5   TCCCTCTCCTCCTCACTCTAGTAGCAACCCTATCAGCCTACTTCCTTTGGTTCCACCTCC
5FOR        TCCCTCTCCTCCTCACTCTAGTAGCAACCCTATCAGCCTACTTCCTTTGGTTCCACCTCC
5REV

Consensus   TCGCCCGAACCCTAACCGGGCCAAAGCCCTGGCCATTAGTGGGGAGCCTCCCAGGCCTTT   300
Gm_FAwH-5   TCGCCCGAACCCTAACCGGGCCAAAGCCCTGGCCATTAGTGGGGAGCCTCCCAGGCCTTT
5FOR        TCGCCCGAACCCTAACCGGGCCAAAGCCCTGGCCATTAGTGGGGAGCCTCCCAGGCCTTT
5REV

Consensus   TCAGGAACCGCGATCGAGTCCACGACTGGATCGCGGATAACCTCCGTGGCAGGGGAGGCT   360
Gm_FAwH-5   TCAGGAACCGCGATCGAGTCCACGACTGGATCGCGGATAACCTCCGTGGCAGGGGAGGCT
5FOR        TCAGGAACCGCGATCGAGTCCACGACTGGATCGCGGATAACCTCCGTGGCAGGGGAGGCT
5REV

Consensus   CCGCAACGTACCAGACGTGCATCATCCCATTCCCTTTCTTGGCACGCAAGAAAGGGTTCT   420
Gm_FAwH-5   CCGCAACGTACCAGACGTGCATCATCCCATTCCCTTTCTTGGCACGCAAGAAAGGGTTCT
5FOR        CCGCAACGTACCAGACGTGCATCATCCCATTCCCTTTCTTGGCACGCAAGAAAGGGTTCT
5REV

Consensus   ACACGGTCACGTGCCATCCAAAGAACCTCGAGCACATCCTCAAGACGCGCTTCGACAACT   480
Gm_FAwH-5   ACACGGTCACGTGCCATCCAAAGAACCTCGAGCACATCCTCAAGACGCGCTTCGACAACT
5FOR        ACACGGTCACGTGCCATCCAAAGAACCTCGAGCACATCCTCAAGACGCGCTTCGACAACT
5REV

Consensus   ACCCAAAAGGCCCCAAGTGGCAAACCGCGTTCCATGATCTTTTGGGCCAAGGAATCTTCA   540
Gm_FAwH-5   ACCCAAAAGGCCCCAAGTGGCAAACCGCGTTCCATGATCTTTTGGGCCAAGGAATCTTCA
5FOR        ACCCAAAAGGCCCCAAGTGGCAAACCGCGTTCCATGATCTTTTGGGCCAAGGAATCTTCA
5REV

Consensus   ACAGCGATGGCGAGACATGGCTCATGCAACGTAAAACCGCGGCATTAGAGTTCACCACGC   600
Gm_FAwH-5   ACAGCGATGGCGAGACATGGCTCATGCAACGTAAAACCGCGGCATTAGAGTTCACCACGC
5FOR        ACAGCGATGGCGAGACATGGCTCATGCAACGTAAAACCGCGGCATTAGAGTTCACCACGC
5REV

Consensus   GAACGTTGAAGCAAGCCATGTCTCGTTGGGTTAACCGATCCATTAAGAACCGGTTGTGGT   660
Gm_FAwH-5   GAACGTTGAAGCAAGCCATGTCTCGTTGGGTTAACCGATCCATTAAGAACCGGTTGTGGT
5FOR        GAACGTTGAAGCAAGCCATGTCTCGTTGGGTTAACCGATCCATTAAGAACCGGTTGTGGT
5REV

Consensus   GCATACTTGACAAAGCGGCGAAAGAACGCGTCTCGGTGGATTTGCAAGACCTTCTTCTGC   720
Gm_FAwH-5   GCATACTTGACAAAGCGGCGAAAGAACGCGTCTCGGTGGATTTGCAAGACCTTCTTCTGC
5FOR        GCATACTTGACAAAGCGGCGAAAGAACGCGTCTCGGTGGATTTGCAAGACCTTCTTCTGC
5REV

Consensus   GGTTAACCTTTGATAATATTTGTGGACTCACGTTCGGCAAAGACCCAGAGACTCTTTCTC   780
Gm_FAwH-5   GGTTAACCTTTGATAATATTTGTGGACTCACGTTCGGCAAAGACCCAGAGACTCTTTCTC
5FOR        GGTTAACCTTTGATAATATTTGTGGACTCACGTTCGGCAAAGACCCAGAGACTCTTTCTC
5REV

Consensus   CGGAGCTACCAGAAAACCCGTTCGCTGTTGCTTTTGACACTGCCACTGAAGCCACCATGC   840
Gm_FAwH-5   CGGAGCTACCAGAAAACCCGTTCGCTGTTGCTTTTGACACTGCCACTGAAGCCACCATGC
5FOR        CGGAGCTACCAGAaAACCCGTTCGCTGTTGCTTTTGACACTGCCACTGAAGCCACCATGC
5REV
```

```
Consensus   ACAGGTTTGTATAATTACACAATACACCTAATTAAATTTTTTTGGCAATACTATAGTTTT   900
Gm_FAwH-5   ACAGGTTTGTATAATTACACAATACACCTAATTAAATTTTTTTGGCAATACTATAGTTTT
5FOR        ACAGGTTTGTATAATTACACAATACACCTAATTAAATTTTTTTGGCAATACTATAGTTTT
5REV

Consensus   AGAGATTTTCATGACCGCATTCACAACTATAATTTAAAACCTCAAGTAATGTATCCAATA   960
Gm_FAwH-5   AGAGATTTTCATGACCGCATTCACAACTATAATTTAAAACCTCAAGTAATGTATCCAATA
5FOR        AGAGATTTTCATGACCGCATTCACAACTATAATTTAAAACCTCAAGTAATGtATCCAATA
5REV

Consensus   GCAAAATTTAGACACAGTTTCTTAATTTTCTTTACACTATTTTTCTATTAAACTTGAAAT  1020
Gm_FAwH-5   GCAAAATTTAGACACAGTTTCTTAATTTTCTTTACACTATTTTTCTATTAAACTTGAAAT
5FOR        gCAAAATTTAgACACAGTTTCtTaatTTTCTTTACACTATTTTTCTAtTAAACTTGaAat
5REV

Consensus   TTTATTTTAATATATAATCTATATTATGAACATTTATTATGCACATTATGTATATATATA  1080
Gm_FAwH-5   TTTATTTTAATATATAATCTATATTATGAACATTTATTATGCACATTATGTATATATATA
5FOR        TTtatTTTaat
5REV

Consensus   AGGTGGAATTTTATTTCTAGCTGATTGAAGAATAATAATAGAGAAAGTTTACGATATACT  1140
Gm_FAwH-5   AGGTGGAATTTTATTTCTAGCTGATTGAAGAATAATAATAGAGAAAGTTTACGATATACT
5FOR
5REV

Consensus   TTCTAAATTTCAAAGGTTTTGTCAATATCTCTCACTTTCTTAGTGTATCAAAACTAATTC  1200
Gm_FAwH-5   TTCTAAATTTCAAAGGTTTTGTCAATATCTCTCACTTTCTTAGTGTATCAAAACTAATTC
5FOR
5REV

Consensus   AACAACTATTAACAAAAGAAAAAAAAATTGTGCCAAATCTTTAAACCTATGAGTACGCAA  1260
Gm_FAwH-5   AACAACTATTAACAAAAGAAAAAAAAATTGTGCCAAATCTTTAAACCTATGAGTACGCAA
5FOR
5REV

Consensus   CGTAACAACTCTCACAATAAAGTTAACGATGTTAGATATTTAGACACCCACAACTTATAA  1320
Gm_FAwH-5   CGTAACAACTCTCACAATAAAGTTAACGATGTTAGATATTTAGACACCCACAACTTATAA
5FOR
5REV

Consensus   ATACTTCATCAACATATTTCTCTCTTTCTATCATATAGTAACCATATAAATAACTTAATT  1380
Gm_FAwH-5   ATACTTCATCAACATATTTCTCTCTTTCTATCATATAGTAACCATATAAATAACTTAATT
5FOR
5REV

Consensus   TTTTTTTTTTTTTTATATATATCTCTCTAATCAAGATCTAAACACTTTGTATACAAGAAT  1440
Gm_FAwH-5   TTTTTTTTTTTTTTATATATATCTCTCTAATCAAGATCTAAACACTTTGTATACAAGAAT
5FOR
5REV

Consensus   CATTTTGCCAGTATTAAAAACACTTAAAACAAATGTATCTAAGTGTTGTTGTATATAAAA  1500
Gm_FAwH-5   CATTTTGCCAGTATTAAAAACACTTAAAACAAATGTATCTAAGTGTTGTTGTATATAAAA
5FOR
5REV
```

```
Consensus   TTAATTCTTAAGATTATTAATGCTATTGTTTGTATAACCACGTACATTCACAATTACAAT  1560
Gm_FAwH-5   TTAATTCTTAAGATTATTAATGCTATTGTTTGTATAACCACGTACATTCACAATTACAAT
5FOR
5REV                                                              CaaT

Consensus   CTAGAGCTCTATGgTATATAGAACTCAGAAATTTTAATATAATTAATTATGTACTACTAT  1620
Gm_FAwH-5   CTAGAGCTCTATGTTATATAGAACTCAGAAATTTTAATATAATTAATTATGTACTACTAT
5FOR
5REV        CTAGAGCTcTATg.TaTATaGaACTCagAAaTTTTAAtaTAAttAATtATGTAcTACTAT

Consensus   ATATTTTTGCTCCCCTTGTTAAACGTTTTATGGAGTACAGGTTCCTCTACCCGGGACTAG  1680
Gm_FAwH-5   ATATTTTTGCTCCCCTTGTTAAACGTTTTATGGAGTACAGGTTCCTCTACCCGGGACTAG
5FOR
5REV        ATATTTTTGCTCCCCTTGTTAAACGTTTTATGGAGTACAGGTTCCTCTACCCGGGACTAG

Consensus   TGTGGAGATTCCAGAAGCTTTTGTGCATTGGATCGGAGAAAAAGTTAAAGGAGAGCCTCA  1740
Gm_FAwH-5   TGTGGAGATTCCAGAAGCTTTTGTGCATTGGATCGGAGAAAAAGTTAAAGGAGAGCCTCA
5FOR
5REV        TGTGGAGATTCCAGAAGCTTTTGTGCATTGGATCGGAGAAAAAGTTAAAGGAGAGCCTCA

Consensus   AAGTTGTGGAAACCTACATGAACGACGCCGTGGCGGATCGCACAGAAGCTCCCTCCGACG  1800
Gm_FAwH-5   AAGTTGTGGAAACCTACATGAACGACGCCGTGGCGGATCGCACAGAAGCTCCCTCCGACG
5FOR
5REV        AAGTTGTGGAAACCTACATGAACGACGCCGTGGCGGATCGCACAGAAGCTCCCTCCGACG

Consensus   ACTTGCTCTCGCGTTTCATGAAGAAGCGCGACGCCGCCGGCAGCTCCTTCTCCGCCGCGG  1860
Gm_FAwH-5   ACTTGCTCTCGCGTTTCATGAAGAAGCGCGACGCCGCCGGCAGCTCCTTCTCCGCCGCGG
5FOR
5REV        ACTTGCTCTCGCGTTTCATGAAGAAGCGCGACGCCGCCGGCAGCTCCTTCTCCGCCGCGG

Consensus   TGCTTCAGCGCATCGTCCTCAACTTCGTCCTCGCCGGCAGGGACACGTCATCGGTGGCCC  1920
Gm_FAwH-5   TGCTTCAGCGCATCGTCCTCAACTTCGTCCTCGCCGGCAGGGACACGTCATCGGTGGCCC
5FOR
5REV        TGCTTCAGCGCATCGTCCTCAACTTCGTCCTCGCCGGCAGGGACACGTCATCGGTGGCCC

Consensus   TCACCTGGTTCTTCTGGCTCCTCACGAACCACCCCGACGTGGAGCAAAAAATAGTGGCTG  1980
Gm_FAwH-5   TCACCTGGTTCTTCTGGCTCCTCACGAACCACCCCGACGTGGAGCAAAAAATAGTGGCTG
5FOR
5REV        TCACCTGGTTCTTCTGGCTCCTCACGAACCACCCCGACGTGGAGCAAAAAATAGTGGCTG

Consensus   AGATCGCAACCGTCCTGGCCGACACGCGCGGCGGAGACCGACGCCGCTGGACGGAGGACC  2040
Gm_FAwH-5   AGATCGCAACCGTCCTGGCCGACACGCGCGGCGGAGACCGACGCCGCTGGACGGAGGACC
5FOR
5REV        AGATCGCAACCGTCCTGGCCGACACGCGCGGCGGAGACCGACGCCGCTGGACGGAGGACC

Consensus   CTCTCGACTTCGGCGAGGCTGACAGGCTCGTGTACCTCAAGGCGGCGCTGGCCGAAACGC  2100
Gm_FAwH-5   CTCTCGACTTCGGCGAGGCTGACAGGCTCGTGTACCTCAAGGCGGCGCTGGCCGAAACGC
5FOR
5REV        CTCTCGACTTCGGCGAGGCTGACAGGCTCGTGTACCTCAAGGCGGCGCTGGCCGAAACGC

Consensus   TGCGTTTATACCCGTCCGTTCCGCAGGACTTCAAGCAGGCCGTGGCCGACGACGTGCTGC  2160
Gm_FAwH-5   TGCGTTTATACCCGTCCGTTCCGCAGGACTTCAAGCAGGCCGTGGCCGACGACGTGCTGC
5FOR
5REV        TGCGTTTATACCCGTCCGTTCCGCAGGACTTCAAGCAGGCCGTGGCCGACGACGTGCTGC
```

```
Consensus   CGGACGGCACGGAGGTTCCGGCGGGGTCGACGGTGACGTACTCGATATACTCCGCGGGGA  2220
Gm_FAwH-5   CGGACGGCACGGAGGTTCCGGCGGGGTCGACGGTGACGTACTCGATATACTCCGCGGGGA
5FOR
5REV        CGGACGGCACGGAGGTTCCGGCGGGGTCGACGGTGACGTACTCGATATACTCCGCGGGGA

Consensus   GGGTGGAAACGATATGGGGGAAAGATTGCATGGAGTTTAAACCGGAGCGATGGCTTTCGG  2280
Gm_FAwH-5   GGGTGGAAACGATATGGGGGAAAGATTGCATGGAGTTTAAACCGGAGCGATGGCTTTCGG
5FOR
5REV        GGGTGGAAACGATATGGGGGAAAGATTGCATGGAGTTTAAACCGGAGCGATGGCTTTCGG

Consensus   TTCGCGGGGACCGGTTCGAACCGCCAAAAGATGGGTTCAAGTTCGTGGCTTTTAACGCTG  2340
Gm_FAwH-5   TTCGCGGGGACCGGTTCGAACCGCCAAAAGATGGGTTCAAGTTCGTGGCTTTTAACGCTG
5FOR
5REV        TTCGCGGGGACCGGTTCGAACCGCCAAAAGATGGGTTCAAGTTCGTGGCTTTTAACGCTG

Consensus   GACCGAGAACTTGTTTGGGCAAGGACTTGGCTTACCTACAAATGAAGTCCGTGGCTGCTG  2400
Gm_FAwH-5   GACCGAGAACTTGTTTGGGCAAGGACTTGGCTTACCTACAAATGAAGTCCGTGGCTGCTG
5FOR
5REV        GACCGAGAACTTGTTTGGGCAAGGACTTGGCTTACCTACAAATGAAGTCCGTGGCTGCTG

Consensus   CTGTGCTTCTGCGTTACCGGCTATCGCTGGTTCCCGGTCACCGGGTGGAACAGAAGATGT  2460
Gm_FAwH-5   CTGTGCTTCTGCGTTACCGGCTATCGCTGGTTCCCGGTCACCGGGTGGAACAGAAGATGT
5FOR
5REV        CTGTGCTTCTGCGTTACCGGCTATCGCTGGTTCCCGGTCACCGGGTGGAACAGAAGATGT

Consensus   CTCTCACTTTGTTCATGAAGAATGGGCTCAGGGTGTTCTTGCATCCACGTAAGCTAGAAA  2520
Gm_FAwH-5   CTCTCACTTTGTTCATGAAGAATGGGCTCAGGGTGTTCTTGCATCCACGTAAGCTAGAAA
5FOR
5REV        CTCTCACTTTGTTCATGAAGAATGGGCTCAGGGTGTTCttGCATCCACGTAAGCTAGAAA

Consensus   GTGGGCCCGGGGTTGCCACCTCACCTTAAAGCGAGGGTCTTTTCGTCATTGTGCTGCAGA  2580
Gm_FAwH-5   GTGGGCCCGGGGTTGCCACCTCACCTTAAAGCGAGGGTCTTTTCGTCATTGTGCTGCAGA
5FOR
5REV        GTGGGCCCGGGGTtgCca
```

Group#5

```
Consensus   AGAAGGCCATAAAATTCCCAACCACCTCTCCTAAAACCTTTCCCATCATATCATCCTATA    60
Gm_FAwH-6   AGAAGGCCATAAAATTCCCAACCACCTCTCCTAAAACCTTTCCCATCATATCATCCTATA
6FOR                                          ccTTTCCcaTCATATCATCCTATA
6REV

Consensus   AATATATATAAAAACCCCTTCAATCAATTCAAACTCCAAATAACCCCAAAATCTCAATTA   120
Gm_FAwH-6   AATATATATAAAAACCCCTTCAATCAATTCAAACTCCAAATAACCCCAAAATCTCAATTA
6FOR        AATATATATAAAAACCCCTTCAATCAATTCAAACTCCAAATAACCCCAAAATCTCAATTA
6REV

Consensus   AAACTGCCCAAACAACATAGAATGGTGCAAATGGATACACTCCATCTTCTCTTCACCCTA   180
Gm_FAwH-6   AAACTGCCCAAACAACATAGAATGGTGCAAATGGATACACTCCATCTTCTCTTCACCCTA
6FOR        AAACTGCCCAAACAACATAGAATGGTGCAAATGGATACACTCCATCTTCTCTTCACCCTA
6REV

Consensus   GCTGCTCTTTTATCCGCATATTTCGTTTGGTTCCATCTCTTGGCTCGAACCCTAACCGGT   240
Gm_FAwH-6   GCTGCTCTTTTATCCGCATATTTCGTTTGGTTCCATCTCTTGGCTCGAACCCTAACCGGT
6FOR        GCTGCTCTTTTATCCGCATATTTCGTTTGGTTCCATCTCTTGGCTCGAACCCTAACCGGT
6REV
```

```
Consensus   CCAAAAGTATGGCCCTTGGTTGGTAGCCTTCCAAGCATGATCGTGAACCGCAATCGGGTT   300
Gm_FAwH-6   CCAAAAGTATGGCCCTTGGTTGGTAGCCTTCCAAGCATGATCGTGAACCGCAATCGGGTT
6FOR        CCAAAAGTATGGCCCTTGGTTGGTAGCCTTCCAAGCATGATCGTGAACCGCAATCGGGTT
6REV

Consensus   CACGACTGGATGGCTGCCAACCTCCGCCAAATCGAGGGATCGGCCACCTACCAAACCTGC   360
Gm_FAwH-6   CACGACTGGATGGCTGCCAACCTCCGCCAAATCGAGGGATCGGCCACCTACCAAACCTGC
6FOR        CACGACTGGATGGCTGCCAACCTCCGCCAAATCGAGGGATCGGCCACCTACCAAACCTGC
6REV

Consensus   ACCCTCACCCTCCCCTTCTTTGCTTGCAAGCAGGCTTTCTTTACCGTCACTAGCAATCCT   420
Gm_FAwH-6   ACCCTCACCCTCCCCTTCTTTGCTTGCAAGCAGGCTTTCTTTACCGTCACTAGCAATCCT
6FOR        ACCCTCACCCTCCCCTTCTTTGCTTGCAAGCAGGCTTTCTTTACCGTCACTAGCAATCCT
6REV

Consensus   AGAAACATAGAACACATCCTCCGAACCCGGTTTGACAACTACCCCAAGGGACCCCACTGG   480
Gm_FAwH-6   AGAAACATAGAACACATCCTCCGAACCCGGTTTGACAACTACCCCAAGGGACCCCACTGG
6FOR        AGAAACATAGAACACATCCTCCGAACCCGGTTTGACAACTACCCCAAGGGACCCCACTGG
6REV

Consensus   CAGGCAGCCTTCCACGACCTCCTAGGCCAAGGAATCTTCAACAGCGACGGTGACACGTGG   540
Gm_FAwH-6   CAGGCAGCCTTCCACGACCTCCTAGGCCAAGGAATCTTCAACAGCGACGGTGACACGTGG
6FOR        CAGGCAGCCTTCCACGACCTCCTAGGCCAAGGAATCTTCAACAGCGACGGTGACACGTGG
6REV

Consensus   CTGATGCAACGTAAAACCGCTGCCCTAGAGTTCACCACCCGAACCCTAAGACAAGCCATG   600
Gm_FAwH-6   CTGATGCAACGTAAAACCGCTGCCCTAGAGTTCACCACCCGAACCCTAAGACAAGCCATG
6FOR        CTGATGCAACGTAAAACCGCTGCCCTAGAGTTCACCACCCGAACCCTAAGACAAGCCATG
6REV

Consensus   GCTCGCTGGGTGAACCGGACCATAAAAAACCGTCTTTGGTGCATTTTGGATAAAGCTGCT   660
Gm_FAwH-6   GCTCGCTGGGTGAACCGGACCATAAAAAACCGTCTTTGGTGCATTTTGGATAAAGCTGCT
6FOR        GCTCGCTGGGTGAACCGGACCATAAAAAACCGTCTTTGGTGCATTTTGGATAAAGCTGCT
6REV

Consensus   AAGGAAAATGTCTCGGTAGATTTGCAAGACCTTCTTTTGCGTTTAACCTTTGATAATATT   720
Gm_FAwH-6   AAGGAAAATGTCTCGGTAGATTTGCAAGACCTTCTTTTGCGTTTAACCTTTGATAATATT
6FOR        AAGGAAAATGTCTCGGTAGATTTGCAAGACCTTCTTTTGCGTTTAACCTTTGATAATATT
6REV

Consensus   TGTGGACTCACCTTCGGTAAAGACCCCGAAACTCTCTCGCCGGAACTACCCGAAAACCCC   780
Gm_FAwH-6   TGTGGACTCACCTTCGGTAAAGACCCCGAAACTCTCTCGCCGGAACTACCCGAAAACCCC
6FOR        TGTGGACTCACCTTCGGTAAAGACCCCGAAACTCTCTCGCCGGAACTACCCGAAAACCCC
6REV

Consensus   TTCACTGTTGCCTTTGACACTGCCACAGAAATCACCTTGCAAAGGTACCACTAAATTTTT   840
Gm_FAwH-6   TTCACTGTTGCCTTTGACACTGCCACAGAAATCACCTTGCAAAGGTACCACTAAATTTTT
6FOR        TTCACTGTTGCCTTTGACACTGCCACAGAAATCACCTTGCAAAGGTACCACTAAATTTTT
6REV

Consensus   GCTGTGTTCACATAATTTCAATTTTTTTCCTTATTTTTATTTAAATTTTTTTATTAGAAG   900
Gm_FAwH-6   GCTGTGTTCACATAATTTCAATTTTTTTCCTTATTTTTATTTAAATTTTTTTATTAGAAG
6FOR        GCTGTGTTCACATAATTTCAATTTTTTTCCTTATTTTTATTTAAATTTTTTTATTAgAAG
6REV
```

```
Consensus   TTTCATCATATATGACAATAACCATAATATCTTGTGAGCATGAGAGGTCAGTGTTATAGA   960
Gm_FAwH-6   TTTCATCATATATGACAATAACCATAATATCTTGTGAGCATGAGAGGTCAGTGTTATAGA
6FOR        TTTCATCATATATGACAATAACCATAATATCTTGtGAGCATGAGAGGTCAGTGTtATAGA
6REV

Consensus   ATGATAATAATTATTTTAATTTTAACCACTCATTTAATAAATTGTACTATTAAAATTATT  1020
Gm_FAwH-6   ATGATAATAATTATTTTAATTTTAACCACTCATTTAATAAATTGTACTATTAAAATTATT
6FOR        ATGATAATAATTATTTTaaTTTTAaCcACTCATTTaATAAAttGTACTaT
6REV

Consensus   TCCTCTCATCTTATTTGTTGGGAAACGTATCACCTGAAAAACATAAGAATATAAAATACT  1080
Gm_FAwH-6   TCCTCTCATCTTATTTGTTGGGAAACGTATCACCTGAAAAACATAAGAATATAAAATACT
6FOR
6REV

Consensus   AATCTCAATTGTATAATTTGAATAAATAGTTAAGATCAAAATATTCTTTTTAATATGTGT  1140
Gm_FAwH-6   AATCTCAATTGTATAATTTGAATAAATAGTTAAGATCAAAATATTCTTTTTAATATGTGT
6FOR
6REV

Consensus   GTATAATTATAAGTTCAAAATTAAACAAATTCATGCCGCTATTATATTTATATTATCTTA  1200
Gm_FAwH-6   GTATAATTATAAGTTCAAAATTAAACAAATTCATGCCGCTATTATATTTATATTATCTTA
6FOR
6REV

Consensus   TTTGATTCTATTTTATGTACTCTAAGTGTAAGGACATACAGTAAGAAAAGAGAGTTATTA  1260
Gm_FAwH-6   TTTGATTCTATTTTATGTACTCTAAGTGTAAGGACATACAGTAAGAAAAGAGAGTTATTA
6FOR
6REV

Consensus   TGAAACTTTGAAAAGGCAAGCTAGATGTTAATTGTTTCTTTTATACCAATAATTAAATGA  1320
Gm_FAwH-6   TGAAACTTTGAAAAGGCAAGCTAGATGTTAATTGTTTCTTTTATACCAATAATTAAATGA
6FOR
6REV

Consensus   ATTTAGACACAACAAAAAGAAAGCTTCTTAATAAGAATATAAATCATCATAGGATTTCTC  1380
Gm_FAwH-6   ATTTAGACACAACAAAAAGAAAGCTTCTTAATAAGAATATAAATCATCATAGGATTTCTC
6FOR
6REV

Consensus   ATAATAATTTCTACTAGCACAAATTAATAAAAAAAAATTGTAGAGATATGATGAAATGGGA  1440
Gm_FAwH-6   ATAATAATTTCTACTAGCACAAATTAATAAAAAAAAATTGTAGAGATATGATGAAATGGGA
6FOR
6REV

Consensus   ATACCCTTATTGGATATTTGTTATTTTGTTATCATTATAAATTAAATTCTTCAAAGTTAT  1500
Gm_FAwH-6   ATACCCTTATTGGATATTTGTTATTTTGTTATCATTATAAATTAAATTCTTCAAAGTTAT
6FOR
6REV

Consensus   ACATACTAAAATTTTATATATTTTTAGTTTAATATCTGTTTACCTGACAATTTTTGTAAT  1560
Gm_FAwH-6   ACATACTAAAATTTTATATATTTTTAGTTTAATATCTGTTTACCTGACAATTTTTGTAAT
6FOR
6REV
```

```
Consensus   ATTAACGATAAAATATTTAAAATAATTATTATAAAAATTCACTAATGATAATTTTCCGTT  1620
Gm_FAwH-6   ATTAACGATAAAATATTTAAAATAATTATTATAAAAATTCACTAATGATAATTTTCCGTT
6FOR
6REV

Consensus   GATATAATGTGAAAATTCTTCTTTTTTTTTTTACATTTTTTTTATACATCTATGATTAGG  1680
Gm_FAwH-6   GATATAATGTGAAAATTCTTCTTTTTTTTTTTACATTTTTTTTATACATCTATGATTAGG
6FOR
6REV                                                TTTTtAtaCATcTATGaTtAGG

Consensus   CTTCTGTACCCaGtCATAATATGGAGATTCGAGAAGCTTCTTGGCATCGGTAAGGAGAAG  1740
Gm_FAwH-6   CTTCTGTACCCTGGCATAATATGGAGATTCGAGAAGCTTCTTGGCATCGGTAAGGAGAAG
6FOR
6REV        cTtctgTacccGg.CaTAATAtGGAGATTCGAGAAGCTTCTtGgCATCggTAaGGAGaAG

Consensus   AAAATACATCAGAGTTTGAAGATCGTCGAAACTTACATGAACGACGCCGTTTCAGCTCGT  1800
Gm_FAwH-6   AAAATACATCAGAGTTTGAAGATCGTCGAAACTTACATGAACGACGCCGTTTCAGCTCGT
6FOR
6REV        AAAATACATCAGAGTTTGAAGATCGTCGAAACTTACATGAACGACGCCGTTTCAGCTCGT

Consensus   GAGAAATCGCCTTCCGACGACTTGCTATCAAGGTTCATAAAGAAGCGTGACGGTGCCGGA  1860
Gm_FAwH-6   GAGAAATCGCCTTCCGACGACTTGCTATCAAGGTTCATAAAGAAGCGTGACGGTGCCGGA
6FOR
6REV        GAGAAATCgCcTTCCGACGACTTGCTATCAAGGTTCATAAAGAAGCGTGACGGTGCCGGA

Consensus   AAAACACTGAGTGCTGCCGCTCTGCGACAAATCGCGTTGAACTTCCTCCTCGCCGGCAGG  1920
Gm_FAwH-6   AAAACACTGAGTGCTGCCGCTCTGCGACAAATCGCGTTGAACTTCCTCCTCGCCGGCAGG
6FOR
6REV        AAAACACTGAGTGCTGCCGcTCTGCGACAAATCGCGTTGAACTTCcTCCTCGCCGGCAGG

Consensus   GACACGTCGTCGGTGGCGTTAAGCTGGTTCTTCTGGCTCGTCATGAACCACCCCGACGTG  1980
Gm_FAwH-6   GACACGTCGTCGGTGGCGTTAAGCTGGTTCTTCTGGCTCGTCATGAACCACCCCGACGTG
6FOR
6REV        GACACGTCGTCGGTGGCGTTAAGCTGGTTCTTCTGGCTCGTCATGAACCACCCCGACGTG

Consensus   GAAGAGAAGATTCTCGACGAACTAACGGCGGTGCTTACTTCCACACGCGGCAGCGACCAG  2040
Gm_FAwH-6   GAAGAGAAGATTCTCGACGAACTAACGGCGGTGCTTACTTCCACACGCGGCAGCGACCAG
6FOR
6REV        GAAGAGAAGATTCTCGACGAACTAACGGCGGTGCTTACTTCCACACGCGGCAGCGACCAG

Consensus   CGGTGCTGGACAGAGGAAGCAGTGGACTTCGAAGAAGCGGAGAAACTGGTTTACTTAAAA  2100
Gm_FAwH-6   CGGTGCTGGACAGAGGAAGCAGTGGACTTCGAAGAAGCGGAGAAACTGGTTTACTTAAAA
6FOR
6REV        CGGTGCTGGACAGAGGAAGCAGTGGACTTCGAAGAAGCGGAGAAACTGGTTTACTTAAAA

Consensus   GCAGCCTTAGCTGAGACACTGCGTTTGTATCCGTCGGTGCCGGAGGATTTCAAGCACGCG  2160
Gm_FAwH-6   GCAGCCTTAGCTGAGACACTGCGTTTGTATCCGTCGGTGCCGGAGGATTTCAAGCACGCG
6FOR
6REV        GCAGCCTTAGCTGAGACACTGCGTTTGTATCCGTCGGTGCCGGAGGATTTCAAGCACGCG

Consensus   ATCGCCGACGATGTGTTGCCGGACGGCACGGCGGTTCCGGCAGGTTCAACGGTGACGTAT  2220
Gm_FAwH-6   ATCGCCGACGATGTGTTGCCGGACGGCACGGCGGTTCCGGCAGGTTCAACGGTGACGTAT
6FOR
6REV        ATCGCCGACGATGTGTTGCCGGACGGCACGGCGGTTCCGGCAGGTTCAACGGTGACGTAT
```

```
Consensus   TCGATTTACGCAATGGGGAGAATGAAGAGCGTGTGGGGCGAGGATTGCATGGAGTTTAAA  2280
Gm_FAwH-6   TCGATTTACGCAATGGGGAGAATGAAGAGCGTGTGGGGCGAGGATTGCATGGAGTTTAAA
6FOR
6REV        TCGATTTACGCAATGGGGAGAATGAAGAGCGTGTGGGGCGAGGATTGCATGGAGTTTAAA

Consensus   CCGGAGCGGTTTTTATCGGTTCAGGGAGACCGGTTCGAACTGCCAAAAGACGGTTACAAG  2340
Gm_FAwH-6   CCGGAGCGGTTTTTATCGGTTCAGGGAGACCGGTTCGAACTGCCAAAAGACGGTTACAAG
6FOR
6REV        CCGGAGCGGTTTTTATCGGTTCAGGGAGACCGGTTCGAACTGCCAAAAGACGGTTACAAG

Consensus   TTCGTTGCGTTTAACGCTGGACCGAGAACCTGTTTGGGGAAGGACTTAGCTTACCTTCAG  2400
Gm_FAwH-6   TTCGTTGCGTTTAACGCTGGACCGAGAACCTGTTTGGGGAAGGACTTAGCTTACCTTCAG
6FOR
6REV        TTCGTTGCGTTTAACGCTGGACCGAGAACCTGTTTGGGGAAGGACTTAGCTTACCTTCAG

Consensus   ATGAAGTCTGTGGCTTCCGCTGTGCTTCTGCGTTACCGGCTGTCGCCGGTTCCCGGTCAC  2460
Gm_FAwH-6   ATGAAGTCTGTGGCTTCCGCTGTGCTTCTGCGTTACCGGCTGTCGCCGGTTCCCGGTCAC
6FOR
6REV        ATGAAGTCTGTGGCTTCCGCTGTGCTTCTGCGTTACCGGCTGTCGCCGGTTCCCGGTCAC

Consensus   CGGGTGCAGCAGAAGATGTCGCTCACTCTCTTCATGAAGCATGGGCTCCGCGTGTTCTTG  2520
Gm_FAwH-6   CGGGTGCAGCAGAAGATGTCGCTCACTCTCTTCATGAAGCATGGGCTCCGCGTGTTCTTG
6FOR
6REV        CGGGTGCAGCAGAAGATGTCGCTCACTCTCTTCATGAAGCATGGGCTCCGCGTGTTCTTG

Consensus   CAGACACGTCAGCTTCAACCGCAGGCCACCACGTCTGCATAGAAAGAATGGGTGGATTTG  2580
Gm_FAwH-6   CAGACACGTCAGCTTCAACCGCAGGCCACCACGTCTGCATAGAAAGAATGGGTGGATTTG
6FOR
6REV        CAGACACGTCAGCTTCAACCGCAGGCCACCACGTCTGCATAGAAAGAATGGGTGGATTTG

Consensus   GTGATGCACAGTGTAAAGGGCAACATGGTCATTTCATTGTTATACATGTCAATTTATTTG  2640
Gm_FAwH-6   GTGATGCACAGTGTAAAGGGCAACATGGTCATTTCATTGTTATACATGTCAATTTATTTG
6FOR
6REV        GTGATGCACAGTGTAAAGGGCAACATGGTCATTTCATTGTTATACATGTCAATTTATTTG

Consensus   TGTTTCTAAATTTTCTGTATATTTACTTCTGGGAGGGTGAAGAGTGAAATCGTTTGTTGT  2700
Gm_FAwH-6   TGTTTCTAAATTTTCTGTATATTTACTTCTGGGAGGGTGAAGAGTGAAATCGTTTGTTGT
6FOR
6REV        TGTTTCTAAATTTTCTGTATATTTACTTCTGGGAGGGTGAAGAGTGAAATCGTTTGT

Consensus   TTGGTTGGGTGGGTGGGTCAAATTTGAGTT                               2730
Gm_FAwH-6   TTGGTTGGGTGGGTGGGTCAAATTTGAGTT
6FOR
6REV
```

Appendix V: Permission for copyrighted material.

Copyright Clearance Center

Confirmation Number: 10992200
Order Date: 04/22/2012

Customer Information

Customer: Jessica Koteles

Order Details

Plant physiology

Billing Status:
Charged to Credit Card

Order detail ID:	62410446
ISSN:	1532-2548
Publication year:	2007
Publication Type:	e-Journal
Publisher:	AMERICAN SOCIETY OF PLANT BIOLOGISTS
Rightsholder:	AMERICAN SOCIETY OF PLANT BIOLOGISTS
Author/Editor:	Raymond Thomas
Your reference:	Jessica Koteles' thesis, Section 1.4

Permission Status: **Granted**

Permission type:	Republish or display content
Type of use:	Republish in a dissertation
Requested use:	Dissertation
Republication title:	Fatty Acid ω-Hydroxylases in Soybean
Republishing organization:	University of Western Ontario
Organization status:	Non-profit 501(c)(3)
Republication date:	05/21/2012
Circulation/ Distribution:	100
Type of content:	Graph
Description of requested content:	Cosegregation of preformed whole-root aliphatic suberin and soybean partial resistance to P. sojae.
Page range(s):	306
Translating to:	No Translation
Requested content's publication date:	05/01/2007
Payment Method:	$ 3.50

Total order items: 1

Order Total: $3.50

CURRICULUM VITAE

Name: Jessica Koteles

Post Secondary Education and Degrees

University of Waterloo, Honours Biology Specialization in Molecular Biology and Biotechnology, 2009, B.Sc.

Honours and Awards

Environment and Sustainability Collaborative Research Seminar Award 2010 - 2011

Western Graduate Research Scholarship 2009 - 2011

Department of Biology Graduate Student Teaching Award 2009 - 2010

Related Work Experience

Teaching Assistant, University of Western Ontario 2009 - 2011

Science Outreach Instructor, University of Toronto 2007

Lab Assistant, Department of Biology, University of Waterloo 2006 - 2007

Conferences

Plant Canada, Plant Adaptation to Environmental Change Meeting, 2011

St. Mary's University, Halifax, Nova Scotia

Poster Presentation: Fatty Acid ω-Hydroxylases in Soybean

Canadian Society of Plant Physiologists Eastern Regional Meeting, 2010

Brock University, St. Catherine's, Ontario

Poster Presentation: Fatty Acid ω-Hydroxylases in Soybean