

2014

INVESTIGATION OF GENES CONTRIBUTING TO WHOLE PLANT SENESCENCE IN SOYBEAN (*GLYCINE MAX*)

Robert Almeida
University of Rhode Island, ralme1@my.uri.edu

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Terms of Use

All rights reserved under copyright.

Recommended Citation

Almeida, Robert, "INVESTIGATION OF GENES CONTRIBUTING TO WHOLE PLANT SENESCENCE IN SOYBEAN (*GLYCINE MAX*)" (2014). *Open Access Master's Theses*. Paper 317.
<https://digitalcommons.uri.edu/theses/317>

This Thesis is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

INVESTIGATION OF GENES CONTRIBUTING TO WHOLE PLANT
SENESCENCE IN SOYBEAN (*GLYCINE MAX*)

BY

ROBERT ALMEIDA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER'S OF SCIENCE

IN

CELL AND MOLECULAR BIOLOGY

UNIVERSITY OF RHODE ISLAND

2014

MASTER OF SCIENCE THESIS
OF
ROBERT ALMEDIA

APPROVED:

Thesis Committee:

Major Professor Joel M. Chandlee
Professor of Cell and Molecular Biology

Marian R. Goldsmith
Professor of Biological Sciences

Albert P. Kausch
Professor of Cell and Molecular Biology

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2014

ABSTRACT

Plant senescence is a genetically determined developmental program characterized by systematic degradative processes that involves activation of new gene activity and down-regulation of other genes that ultimately leads to cell, tissue, organ and whole plant death. Elucidating senescence regulatory pathways and participating genes will allow for the development of strategies to improve crop yields and also curtail post-harvest losses. Three genes are known to be primary regulators of senescence in soybean; namely, *g*, *D1*, and *D2*. In double and triple mutant combinations these genes confer an evergreen leaf and seed phenotype. The double mutation *ggd1d1d2d2* shows an inhibition of degradation of chlorophyll and chlorophyll binding protein, but photosynthesis declines and the leaves still abscise. In the triple mutant *GGd1d1d2d2* the leaves maintain the normal photosynthetic capacity, but still abscise. So, while the senescence program is not entirely blocked in the mutant background, it is altered. Studies have shown that the expression patterns of soybean senescence associated genes (SAGs) are regulated differentially by *g*, *D1* and *D2*. Due to the pivotal regulatory nature of these three genes for senescence, it is important to identify their specific nature. In the first study, an analysis using available soybean genome resources (SoyBase, Phytozome, COGE, etc.) was undertaken. This has resulted in the identification of a gene, Glyma01g41610.2, which encodes a putative transcription factor residing within the marker boundaries of the *D1* locus on chromosome 1 which also shares a high level of synteny with a region on chromosome 11 and includes a paralogous gene, Glyma11g03770.2, within the *D2* marker boundaries. Similarly, another gene, Glyma01g00510.1, also encodes a putative transcription factor and is located within the marker boundaries of the *g* locus on chromosome 1. These genes

were selected as candidates representing *g*, *D1*, and *D2* for RT-PCR analysis. None of the initial candidates exhibited a differential expression profile when comparing wild-type and mutant allelic versions in isogenic genetic backgrounds. The *D2* and *g* genes were then selected for Sanger sequencing to determine if sequence differences were responsible for the observed phenotypic variations. No differences in sequence were observed when comparing wild type and mutant allelic forms. However, sequence variations were observed when comparing *g* in Harosoy versus the reference genome cultivar, Williams 82. A second candidate *g* gene (Glyma01g00520.4) was selected from the defined marker boundary interval on chromosome 1 but did not demonstrate a differential pattern of expression using RT-PCR.

In the second study, additional SAGs up-regulated in various plant species were used to identify candidate soybean ortholog genes that could possibly contribute to whole plant senescence. Some SAGs have been shown to contain a unique senescence response element (SRE) within their promoters that confers a senescence-specific pattern of expression. This is best exemplified by the *SAG12* (cysteine protease) gene of *Arabidopsis thaliana*. The 33 base pair SRE for the *SAG12* gene has been shown to harbor a well-conserved 7 base pair sequence that is also found in SAGs from other plant species. To determine if a related SRE could be responsible for regulating soybean SAGs, a genome-wide study of previously identified as SAGs in other plant species was performed using publically available databases to find related genes in soybean. This search has led to the identification of several soybean genes that harbor this SRE. These genes were bioinformatically analyzed using various structural criteria to identify the best potential soybean ortholog for each gene

type. Structural criteria included measures of alignment similarity with the *Arabidopsis* SRE, proximity of the SRE to the transcription start site, gene architecture, polypeptide sequence identity, and phylogenetic and syntenic relationships. Genes meeting the defined structural criteria underwent evaluation for a functional role in soybean senescence through RT-PCR analysis using a suite of isogenic lines exhibiting normal as well as delayed senescence phenotypes. The selected mutants represent different combinations (single, double and triple mutants) of genes *g*, *D1* and *D2* that give rise to evergreen leaves and green seed phenotypes. None of the genes selected for functional analysis demonstrated evidence of differential expression among the selected isolines. However, many additional genes harboring an SRE have yet to be investigated.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Joel M. Chandlee for his continuous support of my Master's study and research, patience, motivation, enthusiasm, and immense knowledge. His guidance was instrumental during my research and writing of this thesis. I couldn't imagine having a better advisor and mentor for my Master's study. Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Marian R. Goldsmith and Dr. Albert P. Kausch for their encouragement, insightful comments, and hard questions. I'd also like to acknowledge Dr. Jodi Camberg and Dr. Paul Cohen for allowing me to use their laboratories to conduct a majority of this research. Finally, I'd like to thank the many graduate students and friends whose support over the years made this an enjoyable process.

DEDICATION

I would like to dedicate this thesis to my parents, Anna and Mario Almeida, who did everything to provide me with the opportunities they themselves did not have.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	v
DEDICATION.....	vi
LIST OF FIGURES.....	ix
LIST OF TABLES.....	xi

CHAPTER	PAGE
INTRODUCTION.....	1
Literature Cited.....	32
1. Candidate Gene Discovery for Primary Regulatory Genes of Soybean Senescence Using a Map Based Approach.....	41
Abstract.....	42
Introduction.....	44
Materials and Methods.....	47
Results and Discussion.....	54
Literature Cited.....	88
2. Evaluating the Role of a Senescence Response Element (SRE) in Senescence Associated Genes (SAGs) of Soybean.....	90
Abstract.....	91
Introduction.....	93
Results and Discussion.....	96
Materials and Methods.....	99

Literature Cited.....	134
BIBLIOGRAPHY.....	135

LIST OF FIGURES

INTRODUCTION	PAGE
I. Conserved promoter sequences used as selection criteria of soybean SAG's.....	30
CHAPTER I	
1.1a-c. Linkage and sequence map comparisons.....	58
1.2a-c. Physical maps.....	62
1.3. Soybean whole genome syntenic plot.....	66
1.4. Soybean syntenic gene alignment.....	67
1.5. Gel electrophoresis of RNA samples.....	68
1.6. RT-PCR analysis of a putative <i>g</i> gene.....	69
1.7. Genomic sequencing results for candidate <i>g</i> gene.....	70
1.8. RT-PCR analysis of an alternate putative <i>g</i> gene.....	77
1.9. RT-PCR analysis of a putative <i>D1</i> gene.....	78
1.10. RT-PCR analysis of a putative <i>D2</i> gene.....	79
1.11. Genomic sequencing results for candidate <i>D2</i> gene.....	80
CHAPTER II	
2.1. Conserved promoter sequences used for selection criteria.....	112
2.2. RT-PCR analysis of a putative <i>SAG101</i> gene.....	113
2.3. RT-PCR analysis of a putative <i>TMYB</i> gene	114
2.4. RT-PCR analysis of an alternate putative <i>11MADS28</i> gene.....	115
2.5. RT-PCR analysis of an alternate putative <i>18MADS28</i> gene.....	116

2.6. RT-PCR analysis of an alternate putative <i>SARK</i> gene.....	117
2.7. RT-PCR analysis of a putative <i>AS</i> gene.....	118
2.8. RT-PCR analysis of a putative <i>MYB60</i> gene.....	119
2.9. RT-PCR analysis of an alternate putative <i>TMYB</i> gene.....	120
2.10. Structural criteria guidelines for gene selection.....	121
2.11. TBLASTN results.....	122
2.12. Alignment criteria.....	123
2.13. Gene architecture.....	124
2.14. Polypeptide sequence alignment.....	125
2.15. Phylogenetic analysis.....	126
2.16. Syntenic analysis.....	127
2.17. Gel electrophoresis of RNA samples.....	128

LIST OF TABLES

	PAGE
INTRODUCTION	
I. Selected candidate gene families.....	31
CHAPTER I	
1.1. RT-PCR Primers for four candidate primary regulator genes.....	81
1.2. Primers used for amplifying <i>g</i> target sequence.....	82
1.3. Primers used for amplifying <i>D2</i> target sequence.....	83
1.4. Primers used for sequencing <i>g</i> template.....	84
1.5. Primers used for sequencing <i>D2</i> template.....	85
1.6. RNA samples used for expression analysis.....	86
1.7. Soybean “Evergreen” mutants available as Harosoy isolines.....	87
CHAPTER II	
2.1. Soybean “Evergreen” mutants available as Harosoy isolines.....	129
2.2. Selected candidate gene families.....	130
2.3. List of genes meeting structural criteria.....	131
2.4. PCR primers for eight candidate senescence associated genes (SAG’s).....	132
2.5. RNA samples used for expression analysis.....	133

Introduction

Plant Senescence

One of the more striking hallmarks that signals the arrival of autumn is the magnificently aesthetic display of color changes in tree leaves. This annual, well orchestrated event is the result of a tightly regulated though not yet well understood process known as senescence. Senescence is the sequence of biochemical and physiological events comprising the final stage of development for a plant tissue, organ or whole plant, from the mature, fully developed state until death. The changes that take place in senescence represent a genetically programmed sequence with close coordination at the cell and tissue levels. Cells remain viable and show tight metabolic regulation until the end of senescence (Smart, 1994). During senescence, leaf cells undergo orderly changes in cell structure, metabolism, and gene expression. The earliest and most significant change in cell structure is the breakdown of the chloroplast, the organelle that contains up to 70% of the leaf protein. Metabolically, carbon assimilation is replaced by catabolism of chlorophyll and macromolecules such as proteins, membrane lipids, and RNA. Increased catabolic activity is responsible for converting the cellular materials accumulated during the growth phase of leaves and the redistribution of micro- and macro-nutrients, including nitrogen, sulphur, phosphorus and potassium, to growing and reproductive organs (Balazadeh et al., 2008). Finally, upon reaching maturity, leaves abscise (Rubinstein and Leopold, 1964).

Thus, although senescence is a deleterious process for the leaf organ, it critically contributes to the fitness of the whole plant by ensuring optimal production of offspring and better survival of plants in their given temporal and spatial niches.

Leaf senescence is thus an evolutionarily selected developmental process and comprises an important phase in the plant life cycle (Lim et al., 2007).

Plants exhibit two types of senescence: replicative (or mitotic) senescence and post-mitotic senescence (Gan, 2003). Replicative senescence refers to the loss of capacity of a cell for further division upon aging. Post-mitotic senescence is a degenerative process that occurs after cellular maturation or differentiation and leads to cell death. This form of senescence is generally used when describing leaf senescence at the cell and organ level (Lim et al., 2003). The three main factors that contribute to post-mitotic senescence are developmental age internal and environmental factors. Developmental aging occurs throughout normal development from initiation of a leaf primordium throughout senescence and death. Conceptually, developmental aging would determine when senescence starts but not the progression of senescence itself (Lim et al., 2007). While leaf senescence is basically governed by the developmental age, it is also integrated with various endogenous (i.e., hormonal) and environmental (i.e., external) signals to fine tune the development of the plant. These environmental factors can be divided into two categories, abiotic and biotic. The abiotic factors include drought, nutrient limitation, extreme temperature, and oxidative stress by UV-B irradiation and ozone. The biotic factors include pathogen infection and shading by other plants (Lim et al., 2007). Plants have to respond rapidly to deteriorating environmental conditions since, unlike animals, they cannot move in order to find a more favorable situation. One response that plants can make is to remove those parts of the plant that are not essential. For example, a diseased leaf will senesce, die and drop off the plant, thus helping to prevent spread of disease and

allowing the rest of the plant to continue in its development (Buchanan-Wollaston, 1997).

Soybean as a Model to Study Senescence

In 2011, soybean represented 56 percent of world oilseed production, 33 percent of which was produced in the United States. The United States exported 1.275 billion bushels (34.7 million metric tons) of soybeans, which accounted for 37 percent of the global soybean trade. U.S. soybean and soy product exports exceeded \$21.5 billion in 2011 (<http://www.soystats.com>). A large part of this production is used in the extraction of oil, yielding a cake of high protein quality. Soy products are regarded as economical and nutritious feedstuffs with high crude protein content and a reasonably balanced amino acid profile along with many industrial and practical uses (Gatlin III et al., 2007).

G. max exhibits a monocarpic life pattern (Nooden, 1988), meaning it flowers and fruits once in a life cycle (Simmonds, 1980). During this annual life cycle, soybean senescence occurs primarily in the leaves which appear to be the target of the senescence-inducing influence from the seeds. This system of leaf death may be true for other species as well (Nooden, 1988) making soybean a candidate model for basic/fundamental studies. Evidence that soybean senescence may be delayed naturally has also been observed (Guiamet and Giannibelli, 1996). Previous work with normally aging plants (in the absence of biotic and abiotic stressors) has shown the role hormones play in promoting soybean senescence. It was discovered that *Arabidopsis* lines harboring defects for ethylene, a common plant hormone, showed a delayed senescence phenotype (Guo et al., 2004). Prior research in soybean has also

shown that removal of the epicotyls at 16 or 17 days post-germination reversed the decline in nucleic acid, protein, and chlorophyll content in the cotyledons (modified embryonic leaves that appear in early development). Epicotyl removal at 18 days did not reverse the decline in these components, indicating the cotyledon had passed “the point of no return” developmentally (Krul, 1974). Three mutant genes (g, *D1*, *D2*) that delay senescence have also been identified (Guiamet and Giannibelli, 1996).

In an agricultural setting, leaf senescence may limit yield, contributing to the postharvest loss of vegetable crops. Therefore, studying leaf senescence will not only contribute to our knowledge about this fundamental developmental process, but may also lead to ways of manipulating the senescence process for agricultural applications (Gan and Amasino, 1997) such as improving stress tolerance (Lim et al., 2003).

Genetic Regulation of Senescence

Transcription Factors in Plant Development

The regulation of gene transcription is central both to tissue specific-gene expression and to the regulation of gene activity in response to specific stimuli. While instances of posttranscriptional regulation do exist (miRNA, RNAi, etc.), in most cases regulation occurs at the level of transcription by deciding which genes will be transcribed into the primary RNA transcript. Once this has occurred, the remaining stages of gene expression, such as RNA splicing, ultimately result in the production of the corresponding protein (Latchman, 1997). One of the largest and most diverse classes of DNA-binding proteins responsible for regulating gene expression are transcription factors. Transcription factors, largely confined to the nucleus, regulate cell development, differentiation, and cell growth by binding to a specific DNA site

(or set of sites) and regulating gene expression (Pabo and Sauer, 1992). Eukaryotic transcription factors usually consist of several domains. The DNA-binding domain binds to regulatory sequences that can either be adjacent to the promoter or at some distance from it. Most commonly, transcription factors include additional domains that help activate transcription. When a transcription factor is bound to DNA, its activation domain promotes transcription by interacting with RNA polymerase II, by interacting with other associated proteins, or by modifying the local structure of chromatin (Berg et al., 2012).

The DNA site(s) or response elements targeted by transcription factors are conserved DNA bases residing adjacent to the genes they regulate that may repress or activate gene expression. These include promoter elements and enhancers that form a complete set of regulators for each gene that is unique ensuring the right amount of the right protein is expressed at the right time as development proceeds. Transcription factors are aided by crucial proteins such as coactivators, corepressors, chromatin remodelers, histone acetylases, deacetylases, kinases, and methylases, which are present in all eukaryotic cells and contribute to the initiation of every RNA polymerase II primary transcript that eventually becomes messenger RNA (Brivanlou and Darnell, 2002).

The major families of plant transcription factors are: MYB, AP2/EREBP, NAC, bHLH/MYC, bZIP, HB, Z-C2H2, MADS, WRKY, ARF-Aux/IAA, and Dof (Riechmann and Ratcliffe, 2000). These transcription factors employ various structural motifs such as the helix-turn-helix, basic-leucine zipper and Cys2His2 zinc-finger to achieve binding of their particular recognition sequences. A vast majority of these

motifs bind in the major groove of DNA and interact with DNA bases through different combinations of electrostatic and Van der Waals forces (Berg et al., 2012).

Transcription Factors in Plant Senescence

Leaf senescence is an active process involving the differential expression of hundreds of genes and therefore it is presumed that numerous transcription factors are involved as central elements of the regulatory network (Woo et al., 2010). Genes for 96 transcription factors have been identified in *Arabidopsis* as being upregulated at least threefold in senescing leaves. These belong to 20 different transcription factor families, the largest groups being NAC, WRKY, C2H2-type zinc finger, AP2/EREBP, Aux/IAA, and MYB proteins (Lim et al., 2007), with only few examples having been further demonstrated as having a specific functional role in senescence.

NAC proteins are one of the largest families of plant-specific transcription factors with more than 100 members in *Arabidopsis*. NAC family genes play a role in embryo and shoot meristem development, lateral root formation, auxin signaling, and defense response. A total of 20 genes encoding NAC transcription factors, representing almost one fifth of the NAC family members, showed enhanced expression during natural senescence and dark-induced senescence (Lim et al., 2007). Recently, a T-DNA knockout mutation of one of these genes, *AtNAP*, was shown to delay leaf senescence significantly. Induced overexpression caused early senescence, suggesting that *AtNAP* functions as a positive element in leaf senescence (Guo and Gan, 2006).

Among the plant-specific WRKY transcription factor gene family, *AtWRKY53* and *WRKY6* have been further characterized in relation to leaf senescence. *WRKY53* is

upregulated at a very early stage of leaf senescence but decreases again at later stages, implying that *WRKY53* might play a regulatory role in the early events of leaf senescence (Hinderhofer and Zentgraf, 2001). A knockout line of the *WRKY53* gene showed delayed leaf senescence, whereas induced overexpression caused premature senescence, showing that it functions as a positive element in leaf senescence (Miao et al., 2004). *WRKY6* is strongly up-regulated during leaf senescence as well as during pathogen infection. However, although the *wrky6* knockout mutation alters expression of SAGs it does not have any apparent effect on leaf senescence. *SIRK*, a gene encoding a receptor-like protein kinase whose developmental expression is strongly induced specifically during leaf senescence, is dependent on *WRKY6* function. Senescing leaves of *wrky6* knockout mutants showed a reduction in *SIRK* transcript levels while green leaves of *WRKY6* overexpression lines showed elevated *SIRK* transcript levels. Furthermore, the *SIRK* gene promoter was specifically activated by *WRKY6 in vivo* (Robatzek and Somssich, 2002).

Instances of transcription factors with domains responsible for interactions with phytohormones have also been implicated in senescence. AP2 (APETALA2) and EREBPs (ethylene-responsive element binding proteins) are members of a family of transcription factors unique to plants whose distinguishing characteristic is that they contain the so-called AP2 DNA-binding domain. AP2/ REBP genes form a large multigene family, which play a variety of roles throughout the plant life cycle, from being key regulators of several developmental processes, like floral organ identity determination or control of leaf epidermal cell identity, to forming part of the

mechanisms used by plants to respond to various types of biotic and environmental stress (Riechmann and Meyerowitz, 1998).

The RAV family transcription factor in *Arabidopsis*, *RAVI*, has an N-terminal region containing an AP2 DNA-binding domain. *Rav1* mRNA increases at a later stage of leaf maturation and reached a maximal level early in senescence, but decreases again during late senescence. This profile indicates that *RAVI* could play an important regulatory role in the early events of leaf senescence. Furthermore, constitutive and inducible overexpression of *RAVI* causes premature leaf senescence. These data strongly suggest that *RAVI* is sufficient to cause leaf senescence and functions as a positive regulator in this process (Woo et al., 2010).

Another class of plant specific transcription factors are proteins of two related families called Aux/IAA and auxin response factors (ARFs) that regulate auxin-induced gene expression. Aux/IAA proteins share four conserved amino acid sequence motifs called domains I, II, III and IV which are localized to the nucleus. Domains I–IV are conserved in multiple Aux/IAA proteins because they presumably have important structural or regulatory functions. Studies have shown that domains III and IV can mediate homodimerization (Aux/IAA) and heterodimerization between Aux/IAA proteins and ARF proteins (which share these domains). ARF proteins also have a highly conserved N-terminal DNA-binding domain that binds to auxin-response elements in promoters of auxin regulated genes, and a divergent middle region that can activate transcription in some ARFs but has unknown function in others. Together, these results suggest that Aux/IAA proteins regulate gene expression by interacting with ARF proteins to alter their activity (Reed, 2001).

ORE14 is an *Arabidopsis* gene identified by screening populations using ethyl methane sulphonate (EMS) for delayed senescence mutants. *ORE14* was identified by map based cloning which showed it encoded auxin response factor 2 (*ARF2*). A study of the *ore14/arf2* mutant showed a highly significant delay in the senescence parameters examined which included chlorophyll content, the photochemical efficiency of photosystem II, membrane ion leakage, and the expression of senescence-associated genes. A delay of senescence symptoms was also observed under various senescence-accelerating conditions, where detached leaves were treated with darkness, phytohormones, or oxidative stress along with age dependant senescence. These results indicate that the gene defined by these mutations might be a key regulatory genetic component controlling functional leaf senescence (Lim et al. 2010). Further analysis of lines containing T-DNA insertions in *ARF2* and the *ore14/arf2* mutant suggests that the repression of auxin signalling by *ARF2* might positively regulate the onset and progression of leaf senescence in *Arabidopsis* and thus may be essential in controlling auxin-mediated leaf longevity (Ellis et al. 2005, Lim et al., 2010).

Primary Regulator Genes That Regulate the Soybean Senescence Developmental Program

Three different genes (*g*, *D1*, *D2*) are thought to be involved with regulating the whole plant/leaf senescence program of soybean because mutations contribute to an evergreen (non-senescing) phenotype. Homozygous *d1d1d2d2* lines remain green and show an inhibition of chlorophyll degradation and chlorophyll-binding proteins yet they still undergo a decline in photosynthetic activity and leaf abscission (Guiamet et al.1991, Canfield et al. 1995). When combined with the dominant mutant *G*

(*GGd1d1d2d2*) a decline in photosynthetic activity does not take place, but the leaves still abscise. It has been suggested that *d1d2* may control a central regulatory process in the senescence program and that homozygosity at both nuclear loci is required because the two are homeologous (duplicate) loci in the ancient tetraploid soybean genome (Guamet and Giannibelli 1996). The genes responsible for these primary regulators remain unknown; however, because of their important role in the upstream regulation of the senescence program it is plausible they could encode transcription factors. Their characterization represents a logical and organized approach to better understand the mechanisms that control senescence. Near-isogenic lines are available as single, double and triple mutant combinations for these three genes and can serve as the basis for expression analysis of possible transcription factor genes representing *g*, *D1*, and *D2* during senescence. As patterns of altered expression are characterized, a better understanding of the genetic regulation of the senescence pathways will unfold.

Map Based Methods of Gene Identification

A genetic map is a list of genetic elements ordered with regard to their chromosomal position according to their inheritance patterns. Previously these elements were inferred to be genes underlying phenotypic characters such as seed shape in pea or eye color in *Drosophila*, though not having to be restricted to morphological traits. More recently, DNA markers such as restriction length polymorphisms (RFLPs), simple sequence repeat (SSRs), and single nucleotide polymorphisms (SNPs) have become prominent in genetic mapping studies. Currently, genetic maps are thought of as ordered sets of markers, together with inter-marker distances representing milestones along a chromosome or region of a chromosome.

A genetic map serves many practical biological purposes and is a key tool in both classical and modern plant research. For the large majority of plants whose genomes are yet to be sequenced it provides an important resource to understand the order and spacing of markers (and relative order when compared to those of other plants), and for additional genetic information through comparative mapping with genetic maps and genome sequences of other plant species. Once developed, the genetic map underpins studies of plant genes, including quantitative trait loci (QTL) implicated in complex plant traits. For plant species that have been sequenced, it provides a scaffold for genome sequence assembly and validation. The bridge between the genetic map and assembled genome sequence then enables the prediction of candidate genes corresponding to QTL. Finally, it forms an essential tool in marker-assisted plant breeding programs, enabling plant breeders to develop in a targeted fashion new plant varieties in response to demands such as increased yield and resistance to pests and pathogens (Cheema and Dicks, 2009). Genetic maps covering the 20 soybean chromosomes have employed various polymorphic markers. Currently there are 5,500 markers spanning 2296.4 centimorgans (cM). This figure includes 664 RFLP, 3792 SNP, 1006 SSR and 38 other markers available with an average of 275 per chromosome (Hyten et al., 2010)

To facilitate the identification of the genetic basis of many traits and accelerate the creation of improved plant varieties an accurate genome sequence is needed. The first plant physical map generated was that of the *Arabidopsis* genome in 2000 due to its suitability as a model species for plant research and its small genome (157Mb) (Hamilton and Buell, 2012). In 2012, the 1.1-gigabase genome of *Glycine max* var.

Williams 82 was assembled by an 8x whole-genome shotgun assembly (Glyma 1.01) and is currently annotated with 54,175 protein-coding loci (v1.1) (Schmutz et al., 2010). The sequence is freely available through the Joint Genome Institute Phytozome website (<http://phytozome.net>).

Another free and publically available soybean database, SoyBase, has made available a tool that conveniently allows for a comparison of genetic and sequence maps. While there is a strong association with distance in both genetic and physical maps, genetic map distances are calculated by means of recombinant frequencies and do not represent actual physical distances on chromosomes. However, cytogenetic and molecular analysis has shown that genetic distances are, in fact, roughly proportional to chromosome distances (Griffiths et al., 2000), and this assumption is reflected in the current version of SoyBase maps. The markers available on these maps have been utilized by researchers aiming to discover new genes that may be responsible for soybean mosaic virus resistance (*Rsv4*) (Maroof et al., 2010), aphid resistance (*Rag2*) (Kim et al. 2010), and soybean cyst nematode resistance (*rhg1-b*) (Kim, et al., 2010) using fine mapping techniques. It is by employing a map based approach that this current study aims to identify transcription factors that may ultimately represent the primary soybean regulator genes, *g*, *D1* and *D2*.

A Conserved Promoter Element Identified in a Senescence Associated Gene (SAG) of *Arabidopsis thaliana*

Manipulating Senescence through Biotechnology

The ultimate downstream targets of these primary regulator genes are important in effecting the degradative processes of senescence. Efforts to inhibit or delay the effects of senescence have been made using fusion proteins produced from

gene constructs encoding a cysteine protease that incorporate the *SAG12* gene senescence response element (SRE) (Gan and Amasino, 1995) originally identified in the model organism *A. thaliana* (Noh and Amasino, 1999). The *SAG12* SRE is 33 base pairs in length and included within an essential promoter element located -472 to -784 upstream of the transcription start site. It is required for basal level promoter activity and for full *SAG12* expression when in conjunction with an additional upstream enhancer (-1181 to -1345) and basal promoter (-66 to the start codon) (Figure I). The SRE (and basal promoter) is highly conserved in the orthologous *SAG12* gene of rapeseed; however, it does not have significant similarity to any known consensus binding sequences of transcription factors (Noh and Amasino, 1999). The lack of similarity indicates that the developmental regulation of *SAG12* may involve a new or divergent class of transcription factors that specifically recognize this SRE. That only partial promoter activity was conferred by the -603 to -571 region also implies that the other parts of the conserved essential promoter region are required for full *SAG12* promoter activity (Noh and Amasino, 1999).

In tobacco (*Nicotiana tabacum*) the SRE was effectively used with a maize homeobox gene, knotted (*kn1*), and isopentenyl transferase (*ipt*), a cytokinin-producing gene known to inhibit senescence (Gan and Amasino, 1995), to delay senescence. Tobacco plants harboring the *SAG:kn1* and *SAG:ipt* constructs developed with normal morphology but had delayed senescence in both intact and detached leaves. Cytokinin levels were significantly raised in leaves of both experimental constructs compared to wild type, suggesting that the delay in leaf senescence may be mediated through changes in cytokinin metabolism (Ori, 1999). An *ipt* gene under

control of the senescence-specific SAG12 promoter (pSAG12-*IPT*) significantly delayed developmental and postharvest leaf senescence in mature heads of transgenic lettuce (*Lactuca sativa*) homozygous for the transgene. Apart from retardation of leaf senescence, mature, 60-day-old plants exhibited normal morphology with no significant differences in head diameter or fresh weight of leaves and roots (McCabe, 2001). Similar results were achieved in other distantly related plant species such as rapeseed (*Brassica napus*) (Noh and Amasino, 1999), tomato (*Solanum lycopersicum*) (Swartzberg et al., 2006), rice (*Oryza sativa*) (Liu et al., 2010), broccoli (*Brassica oleracea* var. *italica*) (Long-Fang, 2001) and bok choy (*Brassica chinensis*) (Yuan et al., 2002).

The *Arabidopsis thaliana* SAG12 SRE is Found in SAG's from Other Plant Species

Studies by Davies and King (1993) and King et al. (1995) on asparagus found levels of asparagine and asparagine synthetase (*AS*) transcripts increase following the harvest of asparagus spears and during natural foliar senescence. Increased *AS* transcript levels during the course of senescence have also been observed in *Arabidopsis*, sunflower, *M. truncatula*, rice and corn (Gaufichona et al., 2010). Asparagine is thought to be the major transport product in conditions of excess nitrogen or limited carbon supply, which may occur in detached tissues, during senescence, and in photosynthetic tissues during extended dark periods. Thus, the activity of the enzyme responsible for producing asparagine and *AS* must be controlled in response to a complex combination of metabolic, environmental and developmental signals (Winichayakul et al., 2004). Interestingly, upstream promoter analysis identified a conserved senescence-specific sequence motif regulating the

expression of the *AS* gene in asparagus (Winichayakul et al., 2004) similar in sequence to the SRE of the *Arabidopsis SAG12* gene.

This senescence-specific sequence (Noh and Amasino, 1999) up-regulates the expression of the *SAG12* gene during senescence. The upstream region of the *AS* gene was aligned to the *SAG12* upstream region of both *Arabidopsis* and rapeseed and a highly conserved seven base pair region was identified within the well conserved 33 base pair region. This upstream promoter element was also shown to be functionally relevant with deletion assays showing senescence induction occurring when present and delayed when deleted (Winichayakul et al., 2004). An attempt to identify a soybean ortholog of this gene that behaves in a similar pattern is investigated in the present study.

Transcription Factor Genes in Soybean Senescence

Previous work by Chandlee and colleagues employed microarray analysis to identify genes that are differentially expressed in senescing leaf tissue in soybean (Schreier and Chandlee, 2009). Although problems can arise with this method from background interference caused by similar enzymes that are also active, or lack of instrument sensitivity due to low transcript levels (Buchanan-Wollaston, 1997), a MYB transcription factor was identified by this technique as highly differentially expressed (Schreier and Chandlee, 2009).

The structural characteristic common to all known MYB proteins is the DNA-binding domain which has been shown to bind DNA in a sequence-specific manner. Additionally, these proteins usually contain a negatively charged activator domain that has been implicated in transcriptional activation in certain cases (Martin and Paz-Ares,

1997). The MYB DNA-binding domain of plants usually consists of two imperfect repeats of about 52 residues (R2, R3) whereas MYB proteins from animals contain three (R1, R2, R3). These MYB repeats fold into a variant of the helix-turn-helix motif and contain 3 regularly spaced tryptophan residues that play a role in the folding of the hydrophobic core (Dubos et. al, 2010). Although these proteins share the homologous MYB domain, differences in the DNA base contacting residues produce distinct DNA-binding specificities in different members of the family. This gene superfamily participates in a host of processes including regulation of gene expression (Yanhui et al., 2006), secondary metabolism (Mehrtens et al., 2005), hormone signal transduction (Abe et al. 2003), response to environmental stresses (Jung et al., 2008), cell shape, and organ development (Higginson et al., 2003).

Higher plant species usually contain a large number of MYB proteins with *Arabidopsis* genome encoding 196 different MYB proteins spread across four classes (Dubos et. al, 2010) and the soybean genome harboring 252 MYB genes divided into 3 classes (Du et. al, 2012). The number and pervasiveness of MYB-related genes in all major groups of eukaryotic organisms suggests that proteins with MYB-like DNA-binding domains developed early in evolution to regulate gene expression. However, plants appear to have used R2R3-type MYB transcription factors selectively to control their specialized physiological functions, while in contrast, vertebrates have developed only one small group of MYB proteins to control cellular proliferation and differentiation. Because of the large number of genes involved and their roles in plant-specific processes it has been suggested that the MYB superfamily is very important in transcriptional control of higher plants (Martin and Paz-Ares, 1997).

MYB transcription factors have been linked to senescence. In *A. thaliana*, AtMYBL-overexpressing plants displayed a markedly enhanced leaf senescence phenotype. Physiological processes of leaf senescence began earlier in the AtMYBL-overexpressing line than in WT and ATMYBL RNAi plants. The senescence parameters investigated included chlorophyll content, membrane ion leakage and the expression of senescence related genes such as SAG12 (Zhang et al., 2010).

A MYB transcription factor has also been shown to regulate anthocyanin biosynthesis in *A. thaliana*. Wild type lettuce accumulates anthocyanin, predominantly cyanidin and traces of delphinidin, through the course of normal development, and develops a red pigmentation. A transgene construct that over expresses MYB60 from *Arabidopsis* has been shown to inhibit anthocyanin biosynthesis in the lettuce plant (Park et al., 2008) and thus acts as a negative regulator of anthocyanin synthesis. While soybean evergreen mutants exhibit green leaves throughout the senescence process, anthocyanin accumulation persists and leaves continue to abscise (personal communication, Joel M. Chandlee). Identification of a suitable *MYB60* soybean ortholog candidate derived from the *Arabidopsis MYB60* gene was achieved. This allowed for functional analysis to first determine if the ortholog was differentially expressed in wild type and the evergreen genetic triple mutant background before warranting further experiments to determine how it regulates this process. Analyses of the upstream promoter elements (UPEs) of the MYB identified by microarray and the *MYB60* soybean ortholog genes are included in the present study.

Some genes that regulate normal developmental processes of flowering plants play a role in senescence. Specifically, *MADS*-box genes, encoding the *MADS*-

domain family of transcription factors, are involved in controlling all major aspects of the life in land plants. This family is characterized by a highly conserved DNA-binding MADS domain which is about 58 amino acids long and encoded by a DNA sequence termed the MADS box (Gramzow and Theissen, 2010). These genes are well known for their importance in plant development. For example, loss-of-function of some flowering plant *MADS*-box genes causes homeotic transformations of floral organs, indicating that these genes work as organ identity genes during the ontogeny of flowers. *MADS*-box genes also govern reproductive development in eudicotyledonous flowering plants and the developmental processes that follow fertilization of the flower, i.e., seed and fruit development. They also control “flowering time genes” which, depending on internal or environmental factors such as plant age, day-length, and cold, repress or promote the floral transition (Becker and Theissen, 2003). *MADS*-box genes have also been implicated in senescence. The *MADS*-box gene *AGL15* is preferentially expressed during embryogenesis and seed development (Perry et al., 1996) and when constitutively expressed in *Arabidopsis* it strongly delays abscission and senescence in reproductive tissues (Fernandez et al., 2000).

Previous work by Chandlee and colleagues (personal communication) has identified a *MADS*-box gene (*GmSEP3.3/4*) that showed differential expression in tissues of the wild type strain of soybean, Harosoy, at different stages in development as well as in various genetic backgrounds (i.e., wild type vs. assorted senescence mutants), indicating a direct relationship with senescence. Other candidate *MADS*-box genes are differentially expressed in Harosoy during senescence (*GmSEP1.1/2*,

GmSOC.1.2, and GmAG.1/.2), but further analysis of expression patterns of these genes in altered senescence mutant lines is needed (Schreier and Chandlee, 2009). Some preliminary patterns have been elucidated that further characterize the known conserved consensus UPEs of these differentially expressed senescence proteins and are evaluated as part of this study.

Additional Genes of Interest for Promoter Analysis

SAG12 and SAG2 are putative cysteine proteases that show senescence-associated (SAG2) and senescence-specific (SAG12) mRNA expression in *A. thaliana*. They show sequence similarity to a family of cysteine proteases, SAG2 to cathepsin H and SAG12 to cathepsin L. The basal level of SAG2 expression in young leaves indicates that this protease functions in protein turnover throughout the life span of the leaf. The specific induction of SAG12 during leaf senescence may indicate that it has a more specialized role in protein breakdown during senescence. Cathepsin cysteine proteases are active at acidic pH, and are therefore assumed to be localized to lysosomes or vacuoles (Grbic, 2003).

SAG101 has been identified as an acyl hydrolase. During senescence this protein is involved in degrading lipids (He and Gan, 2002). Small RNAs (sRNAs) are key regulators of gene expression in many eukaryotic organisms. These molecules, mostly ranging from about 20 to 30 nucleotides in length, affect all levels of genetic information in plants. A special class of sRNAs, known as microRNAs (miRNAs), can regulate both the chromatin state of their targets and the availability of the encoded transcripts for translation into functional proteins (Rubio-Somoza and Weigel, 2011). The Leaf Senescence Database (LSD,

<http://www.eplantsenescence.org/>) has identified a potential miRNA site of regulation for the *SAG101* gene in *Arabidopsis* which, coupled with regulation by SRE, allows for the possibility of two levels of regulation.

Another gene of interest, *GmSARK* (Senescence Associated Receptor Kinase), has a miRNA regulatory region that was identified in the LSD. This gene was selected for promoter analysis in the present study based on (RNAi)-mediated knock-down of *GmSARK* which dramatically retarded soybean leaf senescence (Li et al., 2006). Protein kinases and especially membrane-associated receptor-like kinases (RLKs) have been found to be involved in many developmental and stress signal transduction pathways. Each RLK consists of three domains, an extracellular receptor domain, a single-pass transmembrane domain and an intracellular kinase domain. In the absence of signal molecules RLKs are usually localized in cell membranes in the form of monomers; once signals emerge the extracellular domain will recognize and bind the signal molecules, resulting in the dimerization of RLKs. The dimerization usually causes the intracellular domains to be autophosphorylated or transphosphorylated, eventually activating the RLKs. Thus, the extracellular signals are transduced into the inside of cells (Li, 2006). In two studies on senescence-regulated IPT gene expression, the promoter of a gene encoding a receptor protein kinase upregulated during senescence of *Phaseolus vulgaris* (common bean) leaves was used (Hajouj et al., 2000). Tobacco plants transformed with the *PSARK:IPT* construct showed a delay in senescence and an exceptional drought tolerance (Rivero et al., 2007). Recently, peanut plants transformed with the same construct were shown to maintain higher photosynthetic rates and higher transpiration under reduced irrigation conditions. In

the field, the transgenic peanut plants produced significantly higher yields than the control plants (Qin et al., 2011). However, control of this *IPT* gene by *PSARK* also demonstrated an increase in expression of brassinoste related genes and repression of jasmonate genes causing the development of enhanced root biomass in these transgenic plants (Peleg et al. 2011; Rivero et al. 2010). This signifies that the increased levels of cytokinin during senescence may not be responsible for the longevity of the green pigmentation in the leaves or photosynthetic capabilities but rather due to additional changes in the metabolism of the plant (Gregersen et al., 2013).

The collection of senescence related genes that includes an AS, transcription factors (MYB, MADS and MYB60), proteases (SAG12 and SAG2), a lipase (SAG101), and a transmembrane receptor kinase (SARK) (Table I) provide a basis for a comprehensive bioinformatic analysis of soybean orthologs to identify a similar function in soybean senescence. Evidence for a conserved UPE, namely the SRE, were examined in soybean orthologs of these various gene types along with a host of other structural criteria including gene architecture, amino acid composition, phylogenetics, and syntenic evidence. Functional analyses of candidate genes were then explored. The best candidates determined by structural analysis were functionally analyzed using RT-PCR to determine whether the conserved SRE regulates their expression under normal (wild-type) and altered genetic backgrounds (mutant) at different developmental time points.

Plant and Soybean Bioinformatic Resources

Phytozome Overview

Phytozome (www.phytozome.net) is a database and graphical user interface enabling comparative genomic studies among land plants and is provided as a joint project of the Department of Energy's Joint Genome Institute and the Center for Integrative Genomics. As of version 9.1, the database houses thirty-one sequenced land plant genomes and is constantly growing as new genomes become available. Each gene has been annotated with PFAM, KOG, and PANTHER assignments and publicly available annotations from RefSeq, SwissProt, Ensembl, and JGI are hyper-linked and searchable. For comparative studies, various clustering methods have been applied to construct orthologous groups of genes that represent the modern descendents of ancestral gene sets at key phylogenetic nodes. These clusterings allow easy access to clade specific orthology/paralogy relationships as well as clade specific genes and gene expansions (Goodstein et al., 2012). The soybean release used in work presented in this thesis is Glyma1.1.

The *Arabidopsis* Information Resources Overview (TAIR)

The *Arabidopsis* Information Resource (TAIR) (www.arabidopsis.org) maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*. Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, metabolism, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the *Arabidopsis* research community. Gene product function data is updated every two weeks from the latest published research

literature and community data submissions. Gene structures are updated 1-2 times per year using computational and manual methods as well as community submissions of new and updated genes. TAIR also provides extensive linkouts from data pages to other *Arabidopsis* resources (Lamesch et al., 2011). The *Arabidopsis* release used in work presented here is v10.1.

SoyBase Overview

SoyBase (www.soybase.org), the USDA-ARS soybean genetic database, is a comprehensive repository for professionally curated genetics, genomics and related data resources for soybean. SoyBase contains the most current genetic, physical and genomic sequence maps integrated with qualitative and quantitative traits. The quantitative trait loci (QTL) represent more than 18 years of QTL mapping of more than 90 unique traits. SoyBase also contains a well-annotated ‘Williams 82’ genomic sequence and associated data mining tools. The genetic and sequence views of the soybean chromosomes and the extensive data on traits and phenotypes are extensively interlinked. This allows entry to the database using almost any kind of available information, such as genetic map symbols, soybean gene names or phenotypic traits. SoyBase is the repository for controlled vocabularies for soybean growth, development and trait terms, which are also linked to the more general plant ontologies (Grant et al. 2010).

Leaf Senescence Database (LSD)

The leaf senescence database (LSD, <http://www.eplantsenescence.org/>) was developed by compiling a broad literature survey that contains a total of 1145 senescence

associated genes (SAGs) from 21 plant species. These SAGs were retrieved based on genetic, genomic, proteomic, physiological or other experimental evidence, and were classified into different categories according to their functions in leaf senescence or morphological phenotypes when mutated. Extensive annotations for these SAGs have been made by both manual and computational approaches, and users can browse or search the database to obtain information including literature, mutants, phenotypes, expression profiles, miRNA interactions, orthologs in other plants and cross links to other databases. Also integrated into LSD is a bioinformatics analysis platform, WebLab, which allows users to perform extensive sequence analysis of their interested SAGs of interest. The SAG sequences in LSD can also be downloaded readily for bulk analysis. The LSD contains the largest number of SAGs to date and represents the most comprehensive and informative plant senescence-related database, which aids in facilitates systems biology research and comparative studies on plant aging (Liu et al., 2010).

Comparative Genomics Platform Overview (CoGe)

CoGe is publicly available at <http://genomeevolution.org>. This resource contains four major systems: a data engine storing thousands of genomes, a suite of interconnected web-based tools, a wiki documentation system with hundreds of pages on comparative genomics, and a TinyURL resource for storing links to CoGe to regenerate data and analyses. The data in CoGe is constantly growing as new genomes and new versions of existing genomes become available. Currently, there are nearly 20,000 genomes from 15,000 organisms. There are over 20 tools in CoGe; each of these performs one general task, such as searching for genomes, displaying FASTA sequences, querying

genomes, comparing genomic regions, etc. These tools are all interlinked with one another so that results generated in one tool may be seamlessly sent to another tool for downstream analyses. Due to the interlinking of these tools, following no specific workflow or analytical pipeline is not required (Lyons and Freeling, 2008).

Main CoGe Entrance Pages

- OrganismView - Search and get an overview of an organism and its genomic information.
- CoGeBlast - Blast against any number of organisms using the CoGe Blast interface (supports Blast, and BlastZ).
- FeatView - Search for genomic features by name or description. Phylogenetic trees can be created through this page.
- SynMap - Generate syntenic dotplots of any two genomes.
- Synfind - SynFind identifies syntenic regions against any set of genomes given a gene in one genome, even if that gene is not present in a target genome. In the process of these analyses, SynFind identifies all syntenic regions to all genes in the query genome. Complete syntenic gene-sets can be downloaded, and syntenic depth tables are generated to access the polyploidy level between the query genome and each target genome.
- GEvo - Compare multiple genomic regions using a variety of sequence comparison algorithms for high-resolution analysis to quickly identify patterns of genome evolution.

Plant Genome Duplication Database Overview (PGDD)

PGDD (<http://chibba.agtec.uga.edu/duplication/>) is a web service providing synteny information in terms of colinearity between chromosomes. At present, PGDD contains data for 26 plants including bryophytes and chlorophyta, as well as angiosperms with draft genome sequences. In addition to the inclusion of new genomes are included as they become available along with new functions to enhance PGDD (Lee et al., 2013).

A Database of Plant Cis-acting Regulatory DNA Elements Overview (PLACE)

PLACE (<http://www.dna.affrc.go.jp/PLACE/>) is a database of motifs found in plant cis-acting regulatory DNA elements, all from previously published reports. It covers vascular plants only. In addition to the motifs originally reported, their variations in other genes or in other plant species reported later are also compiled. The PLACE database also contains a brief description of each motif and relevant literature with PubMed ID numbers and DDBJ/EMBL/GenBank nucleotide sequence databases accession numbers also included (Higo et al., 1999).

Gramene

Gramene's (www.gramene.org) purpose is to facilitate researchers' ability to understand the grass genomes and take advantage of genomic sequence known in one species for identifying and understanding corresponding genes, pathways and phenotypes in other grass species. This is achieved by building automated and curated relationships between cereals for both sequence and biology. The automated and curated relationships are queried and displayed using controlled vocabularies and web-based displays. The controlled vocabularies (Ontologies), currently being used include

Gene ontology, Plant ontology, Trait ontology, Environment ontology and Gramene Taxonomy ontology. The web-based displays for phenotypes include the Genes and Quantitative Trait Loci (QTL) modules. Sequence based relationships are displayed in the Genomes module using the genome browser adapted from Ensembl, in the Maps module using the comparative map viewer (CMap) from GMOD, and in the Proteins module displays. BLAST is used to search for similar sequences. Literature supporting all the above data is organized in the Literature database (Jaiswal et al., 2006).

Primer-Blast

Primer-BLAST was developed at NCBI to help users make primers that are specific to intended PCR target. It uses Primer3 to design PCR primers and then uses BLAST and global alignment algorithm to screen primers against user-selected database in order to avoid primer pairs (all combinations including forward-reverse primer pair, forward-forward as well as reverse-reverse pairs) that can cause non-specific amplifications (Ye et al., 2012).

Geneious

Geneious Pro is a commercial, integrated, cross-platform bioinformatics software suite for manipulating, finding, sharing, and exploring biological data such as DNA sequences or proteins, phylogenies, 3D structure information, publications, etc. It features sequence alignment and phylogenetic analysis, contig assembly, primer design and cloning, access to NCBI and UniProt, BLAST, protein structure viewing, automated PubMed searching, and many more applications (Biomatters New Zealand, 2013).

Significance of Project and Outline

Soybean is a major source of food worldwide for humans and livestock. Improving yields is a desirable objective because of this major economic significance. Whole plant/leaf senescence, an orderly degenerative process leading to death, is a developmental program known to be genetically controlled in soybean. The manipulation of this process can potentially improve yields. An understanding of soybean genes that regulate senescence both upstream, early events in the developmental program and downstream, later events in the developmental program, will be useful in this regard (Schreier and Chandler, 2009). Currently little is known about the genes that serve as the primary regulators of senescence or about promoter elements for the downstream responding SAGs in the soybean senescence program. The auxin response factor and MYB gene families represent two of the largest families of transcription factors that play important roles in many aspects of growth and development in most, if not, all eukaryotes. New knowledge will allow researchers to continue to decipher the many roles of these important transcription factors in soybean development. The correlation of transcription factors that lead to whole plant and leaf senescence and the possibility of their role as major regulators of senescence make them attractive targets to elucidate the regulatory framework of this developmental process. The regulatory regions of genes that usher in the changes normally associated with senescence are also of interest because of the opportunities to exploit them for agricultural benefits. The availability of evergreen mutants as soybean near-isogenic lines provides an important resource to assist in the study of these important processes. Any further understanding of senescence, especially its regulators, will add to the

understanding of the genetic basis for developmental regulation and has direct application to the improvement of crop plants by reduction of spoilage and increased yields.

This project targets three objectives:

1) Identification of primary regulatory genes of soybean senescence using a map based approach

Genetic and physical maps offer genetic markers that are used as boundaries to create a database that can be screened for transcription factors representing *g*, *D1* and *D2*.

2) Genome-wide identification of soybean genes that harbor a conserved UPE and possibly contribute to whole plant senescence

Screen publically available databases using various structural criteria to identify candidate genes for further functional analysis.

3) Expression analysis of primary regulator and ortholog genes harboring an SRE

Evaluate the expression of candidate genes using RT-PCR to monitor their expression in WT and mutant soybean strains.

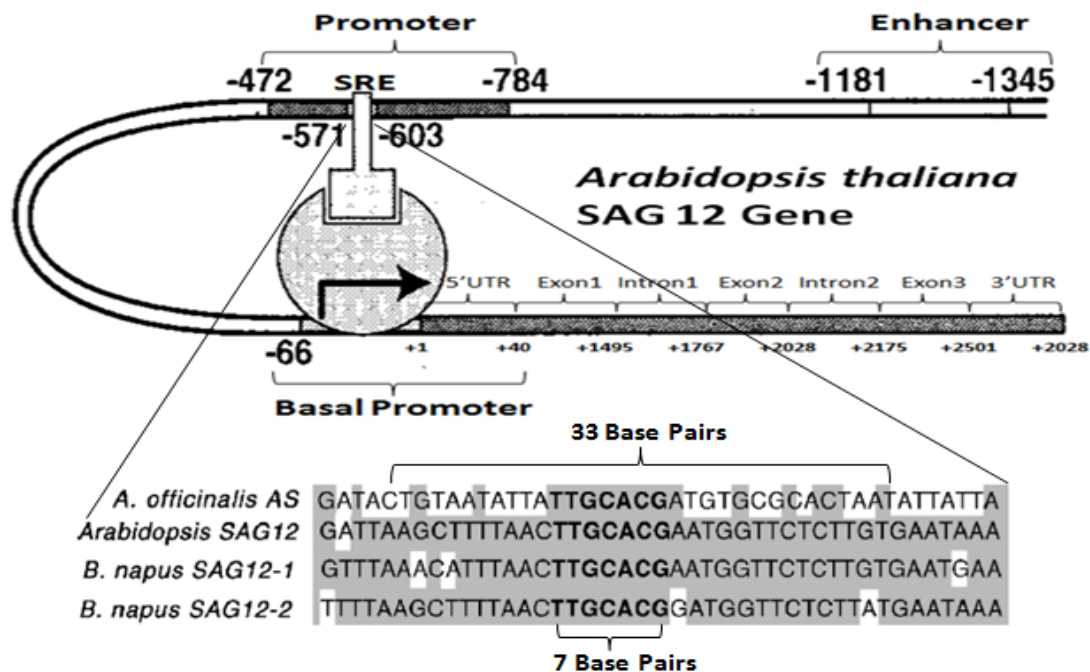


Figure I. Conserved promoter sequences used as selection criteria of soybean SAG's.

A highly senescence-specific promoter identified in the *Arabidopsis thaliana* SAG12 gene. A subsequent study identified a conserved 7 base pair sequence within the 33 base pair Senescence Response Element (SRE). These two sequences serve as the basis for selection of candidate genes.

Table I: Selected candidate gene families

Gene of Interest	Function
SIMYB	Senescence-Specific Transcription Factor
AoAS	Nitrogen Remobilization
AtSARK	Senescence-Associated Receptor Kinase (Transmembrane Protein, Leucine Rich)
AtSAG12	Senescence-Specific Cysteine Protease
AtSAG 2	Senescence-associated Cysteine Protease
AtSAG101	Acyl Hydrolase
AtMYB60	Anthocyanin Biosynthesis Transcription Factor
GmSEP1.1/.2	MADS-box Transcription Factor
GmSEP 3.3/.4	MADS-box Transcription Factor
GmSOC.1/.2	MADS-box Transcription Factor
GmAG.1/.2	MADS-box Transcription Factor

A database and literature search for genes experimentally implicated with senescence in soybean and non-soybean plant species was conducted. A list was compiled and screened for soybean paralogues and orthologs that then had 3000 base pairs upstream evaluated for the presence of an SRE and conserved element.

Literature Cited

Abe H., Urao T., Ito T., Seki M., Shinozaki K., and Yamaguchi-Shinozaki K. ***Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling** *Plant Cell* (2003) **15**: 63-78.

Balazadeh S., Riano-Pachon D.M. and Mueller-Roeber B. **Transcription factors regulating leaf senescence in *Arabidopsis thaliana*** *Plant Biology* (2008) **10**: 63–75.

Becker, A. and Theissen, G. **The major clades of MADS-box genes and their role in the development and evolution of flowering plants** *Molecular Phylogenetics and Evolution* (2003) **29**: 464–489.

Geneious version (6.1.5) created by Biomatters. available from <http://www.geneious.com/>

Berg J.M., Tymoczko J.L., Stryer L., and Gatto G.J. **Biochemistry 7th ed.** W. H. Freeman and Company (2012) 941-943.

Brivanlou A.H. and Darnell J.E. Jr. **Signal transduction and the control of gene expression** *Science* (2002) **295** (5556): 813-81.

Buchanan-Wollaston V. **The molecular biology of leaf senescence** *Journal of Experimental Botany* (1997) **48** (307): 181-199.

Canfield M.R., Guiamét J.J., and Noodén L.D. **Alteration of Soybean Seedling Development in Darkness and Light by the Stay-green Mutation *cytG* and *Gd1d2*** *Annals of Botany* (1995) **75** (2): 143-150.

Chandlee J.M. **Current molecular understanding of the genetically programmed process of leaf senescence** *Physiologia Plantarum* (2001) **113**: 1–8.

Cheema J. and Dicks J. **Computational approaches and software tools for genetic linkage map estimation in plants** *Briefings in Bioinformatics* (2009) **10** (6): 595-608.

Chen L-F. O., Hwang J-Y., Charng Y_Y., Sun C-W. and Yang S-F. **Transformation of broccoli (*Brassica oleracea* var. *italica*) with isopentenyltransferase gene via *Agrobacterium tumefaciens* for post-harvest yellowing retardation** *Molecular Breeding* (2001) **7**: 243–257.

College of Life Sciences National Laboratory of Protein Engineering and Plant Genetic Engineering and Center for Bioinformatics (2012) Leaf Senescence Database 2012 <http://www.eplantsenescence.org/>

Davies KM. and King GA. **Isolation and characterization of a cDNA clone for harvest induced asparagine synthetase from *Asparagus officinalis* L.** *Plant Physiology* (1993) **102**: 1337-1340.

Du H., Yang S.S., Liang Z., Feng B.R., Liu L., Huang Y.B., and Tang Y.X. **Genome-wide analysis of the MYB transcription factor superfamily in soybean** *BMC Plant Biology* (2012) **12** (106): 1-22.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. and Lepiniec, L. **MYB transcription factors in *Arabidopsis*** *Trends in Plant Science* (2010) **15** (10): 573-581.

Ellis C.M., Nagpal P., Young J.C., Hagen G., Guilfoyle T.J. and Reed J.W. **AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*** *Development* (2005) **132** (20): 4563-4574.

Fernandez, D., Heck, G.R., Perry, S.E., Patterson, S.E., Bleecker, A.B., and Fang, S.C. **The embryo MADS domain factor AGL15 acts postembryonically: inhibition of perianth senescence and abscission via constitutive expression** *Plant Cell* (2000) **12**: 183–197.

Gan S. **Mitotic and Postmitotic Senescence in Plants** *Science of Aging Knowledge Environment* (2003) **38** (7): 1-10.

Gan S. and Amasino R.M. **Inhibition of Leaf Senescence by Autoregulated Production of Cytokinin** *Science* (1995) **270** (5244): 1986-1988.

Gan S. and Amasino R.M. **Making Sense of Senescence** *Plant Physiology* (1997) **113**: 313-319.

Gatlin III D.M., Barrows F.T., Brown P, Dabrowski D., Gaylord T.G, Hardy R.W., Herman E., Hu G., — Krogdahl A., Nelson R., Overturf K., Rust M., Sealey W., Skonberg D., Souza E.J., Stone D., Wilson R., and Wurtele E. **Expanding the utilization of sustainable plant products in aquafeeds: a review** *Aquaculture Research* (2007) **38**: 551-579.

Gaufichona L., Reisdorf-Crena M.,b, Rothsteinc SJ., Chardona F., and Suzukia A. **Biological functions of asparagine synthetase in plants** *Plant Science* (2010) **179**: 141–153.

Goodstein D.M., Shu S., Howson R., Neupane R., Hayes R.D., Fazo J., Mitros T., Dirks W., Hellsten U., Putnam N., and Rokhsar D.S. **Phytozome: a comparative**

platform for green plant genomics *Nucleic Acids Research* (2012) **40** (Database issue): D1178-D1186.

Gramzow, L. and Theissen, G. **A hitchhiker's guide to the MADS world of plants** *Genome Biology* (2010) **11**: 1-11.

Grant D., Nelson R.T., Cannon S.B., and Shoemaker R.C. **SoyBase, the USDA-ARS soybean genetics and genomics database** *Nucleic Acids Research* (2010) **38** (Database issue): D843–D846.

Grbic V. **SAG2 and SAG12 protein expression in senescing *Arabidopsis* plants** *Physiologia Plantarum* (2003) **119**: 263–269.

Gregersen P.L., Culetic A., Boschian L. and Krupinska K. **Plant senescence and crop productivity** *Plant Molecular Biology* (2013) [Epub ahead of print]: 1-20.

Griffiths A.J.F., Miller J.H., Suzuki D.T., Lewontin R.C. and Gelbart W.M. **An Introduction to Genetic Analysis, 7th ed.** New York: W. H. Freeman (2000) Linkage maps.

Guamet J.J. and Giannibelli C.M. **Nuclear and cytoplasmic “stay-green” mutations of soybean alter the loss of leaf soluble proteins during senescence** *Physiologia Plantarum* (1996) **96** (4): 655–661.

Guamét J.J., Schwartz E., Pichersky E., and Noodén L.D. **Characterization of Cytoplasmic and Nuclear Mutations Affecting Chlorophyll and Chlorophyll-Binding Proteins during Senescence in Soybean** *Plant Physiology* (1991) **96** (1): 227-231.

Guo Y., Cai Z., and Gan S. **Transcriptome of *Arabidopsis* leaf senescence** *Plant, Cell and Environment* (2004) **27**: 521–549.

Guo Y and, Gan S. **AtNAP, a NAC family transcription factor, has an important role in leaf senescence** *The Plant Journal* (2006) **46** (4): 601-612.

Hajouj T., Michelis R. and Gepstein S. **Cloning and characterization of a receptor-like protein kinase gene associated with senescence** *Plant Physiology* (2000) **124** (3): 1305-1314.

Hamilton J.P. and Buell R.C. **Advances in plant genome sequencing** *The Plant Journal* (2012) **70**: 177–190.

He Y. and Gan S. **A gene encoding an acyl hydrolase is involved in leaf senescence in *Arabidopsis*** *The Plant Cell* (2002) **14**: 805–815.

- Higginson T., Li S.F., and Parish R.W. **AtMYB103 regulates tapetum and trichome development in *Arabidopsis thaliana*** *The Plant Journal* (2003) **35**: 177-192.
- Higo K., Ugawa, Iwamoto M. and Korenaga T. **Plant cis-acting regulatory DNA elements (PLACE) database** *Nucleic Acids Research* (1999) **27** (1): 297-300.
- Hill J., Nelson E., Tilman D., Polasky S., and Tiffany D **Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels** *Proceedings of the National Academy of Sciences* (2006) **103** (30): 11206–11210.
- Hinderhofer K. and Zentgraf U. **Identification of a transcription factor specifically expressed at the onset of leaf senescence** *Planta* (2001) **213** (3): 469-473.
- Huynh L.E., Toai T.V., Streeter J. and Banowitz G. **Regulation of flooding tolerance of SAG12:ipt *Arabidopsis* plants by cytokinin** *Journal of Experimental Botany* (2005) **56**: 1397–1407.
- Hyten D.L., Choi I-Y., Song Q, Specht J.E., Carter T.E., Shoemaker R.C., Hwang E-Y., Matukumalli L.K. and Cregan P.B. **A High Density Integrated Genetic Linkage Map of Soybean and the Development of a 1536 Universal Soy Linkage Panel for Quantitative Trait Locus Mapping** *Crop Science* (2010) **50**: 960–968.
- Jaiswal P., Ni J., Yap I., Ware D., Spooner W., Youens-Clark K., Ren L., Liang C., Zhao W., Ratnapu K., Faga B., Canaran P., Fogleman M., Hebbard C., Avraham S., Schmidt S., Casstevens T.M., Buckler E.S., Stein L. and McCouch S. **Gramene: a bird's eye view of cereal genomes** *Nucleic Acids Research* (2006) **34** (Database issue): D717-D723.
- Jung C., Seo J.S., Han S.W., Koo Y.J., Kim C.H., Song S.I., Nahm B.H., Choi Y.D., and Cheong J.J. **Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*** *Plant Physiology* (2008) **146**: 623-635.
- Kim K.S., Hill C.B., Hartman C.L., Hyten D.L., Hudson M.E., Diers B.W. **Fine mapping of the soybean aphid-resistance gene Rag2 in soybean PI 200538** *Theoretical and Applied Genetics* (2010) **121**: 599–610.
- Kim M., Hyten D.L., Bent A.F. and Diers B.W. **Fine Mapping of the SCN Resistance Locus rhg1-b from PI 88788** *The Plant Genome* (2010) **3**: 81-89.
- King G.A., Davies K.M., Stewart R.J., and Borst W.M. **Similarities in gene expression during the postharvest-induced senescence of spears and the natural foliar senescence of asparagus** *Plant Physiology* (1995) **108**: 125-128.

Krul W.R. **Nucleic Acid and Protein Metabolism of Senescing and Regenerating Soybean Cotyledons** *Plant Physiology* (1974) **54**: 36-40.

Lamesch P., Berardini T.Z., Li D., Swarbreck D., Wilks C., Sasidharan R., Muller R., Dreher K., Alexander D.L., Garcia-Hernandez M., Karthikeyan A.S., Lee C.H., Nelson W.D., Ploetz L., Singh S, Wensel A. and Huala E. **The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools** *Nucleic Acids Research* (2011) 1–9.

Latchman D.S. **Transcription factors: an overview** *International Journal of Biochemistry and Cell Biology* (1997) **12** (1): 305-12.

Lee T.H., Tang H., Wang X. and Paterson A.H. **PGDD: a database of gene and genome duplication in plants** *Nucleic Acids Research* (2013) **41** (Database issue): D1152-D1158.

Li X.P., Gan R., Li P.L., Ma Y.Y., Zhang L.W., Zhang R., Wang R., and Wang N.N. **Identification and functional characterization of a leucine-rich repeat receptor-like kinase gene that is involved in regulation of soybean leaf senescence** *Plant Molecular Biology* (2006) **61**: 829–844.

Lim P.O., Lee I.C., Kim J., Kim H.J., Ryu J.S., Woo H.R. and Nam H.G. **Auxin response factor 2 (ARF2) plays a major role in regulating auxin-mediated leaf longevity** *Journal of Experimental Botany* (2010) **61** (5): 1419-1430.

Lim P.O., Kim H.J., and Nam H.G. **Leaf Senescence** *Annual Review of Plant Biology* (2007) **58**: 115–36.

Lim P.O., Woo H.R. and Nam H.G. **Molecular genetics of leaf senescence in Arabidopsis** *Trends in Plant Science* (2003) **8** (6): 272-278.

Liu L., Zhou Y., Szczerba M.W., Li X., and Lin Y. **Identification and application of a rice senescence-associated promoter** *Plant Physiology* (2010) **53** (3): 1239-1249.

Liu X., Li Z., Jiang Z., Zhao Y, Peng J., Jin J., Guo H., Luo J. **LSD: a leaf senescence database** *Nucleic Acids Research* (2011) **39** (Database issue) D1103–D1107.

Lyons E. and Freeling M. **How to usefully compare homologous plant genes and chromosomes as DNA sequences** *The Plant Journal* (2008) **53** (4): 661–673.

Martin C. and Paz-Ares J. **MYB transcription factors in plants** *Trends in Genetics* (1997) **13** (2): 67-73.

Maroof S.M.A., Tucker D.M., Skoneczka J.A., Bowman B.C., Tripathy S. and Tolin S.A. **Fine Mapping and Candidate Gene Discovery of the Soybean Mosaic Virus Resistance Gene, Rsv4** *The Plant Genome* (2010) **3** (1): 14-22.

McCabe M.S., Garratt L.C., Schepers F., Jordi W.J.R.M, Stoopen G.M., Davelaar E., Rhijn J.H.A.V, Power J.B., and Davey M.R. **Effects of PSAG12-IPT Gene Expression on Development and Senescence in Transgenic Lettuce** *Plant Physiology* (2001) **127**: 505–516.

Mehrtens F., Kranz H., Bednarek P., and Weisshaar B. **The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis** *Plant Physiology* (2005) **138**: 1083-1096.

Miao Y., Laun T., Zimmermann P. and Zentgraf U. **Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis** *Plant Molecular Biology* (2004) **55** (6): 853-867.

Noh Y-S and Amasino R.M. **Identification of a promoter region responsible for the senescence-specific expression of SAG12** *Plant Molecular Biology* (1999) **41**: 181–194.

Noodén L.D., Guiamet J.J. and John I. **Senescence mechanisms** *Physiologia Plantarum* (1997) **101**: 746-753.

Noodén L.D. and Leopold A.C., **Senescence and Aging in Plants** *Academic Press San Diego* (1988) 33-34, 506-507.

Ori N., Juarez M.T., Jackson D., Yamaguchi J., Banowetz G.M., and Hake S. **Leaf Senescence Is Delayed in Tobacco Plants Expressing the Maize Homeobox Gene *knotted1* under the Control of a Senescence-Activated Promoter** *The Plant Cell* (1999) **11**: 1073–1080.

Pabo C.O. and Sauer R.T. **Transcription factors: structural families and principles of DNA recognition** *Annual Review of Biochemistry* (1992) **61** (1): 053-95.

Park J.S., Kim J.B., Cho K.J., Cheon C.I., Sung M.K., Choung M.G., and Roh K.H. **Arabidopsis R2R3-MYB transcription factor AtMYB60 functions as a transcriptional repressor of anthocyanin biosynthesis in lettuce (*Lactuca sativa*)**. *Plant Cell Reports* (2008) **27**: 985–994.

Peleg Z. and Blumwald E. **Hormone balance and abiotic stress tolerance in crop plants** *Current Opinions in Plant Biology* (2011) **14**: 290–295.

Perry, S.E., Nichols, K.W., and Fernandez, D.E. **The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development** *Plant Cell* (1996) **8**: 1977–1989.

Ptashne, M. **How eukaryotic transcriptional activators work** *Nature* (1988) **335**: 683–689.

Qin H., Gu Q., Zhang J., Sun L., Kuppu S., Zhang Y., Burow M., Payton P., Blumwald E. and Zhang H. **Regulated expression of an isopentenyltransferase gene (IPT) in peanut significantly improves drought tolerance and increases yield under field conditions** *Plant Cell Physiology* (2011) **52** (11): 1904-1914.

Quirino B.F., Noh Y-S., Himelblau E. and Amasino R.M. **Molecular aspects of leaf senescence** *Trends in Plant Science Reviews* (2000) **5** (7): 278-282.

Reed J.W. **Roles and activities of Aux/IAA proteins in Arabidopsis** *Trends in Plant Science* (2001) **6** (9): 420-425.

Riechmann J.L. and Meyerowitz E.M. **The AP2/EREBP family of plant transcription factors** *The Journal of Biological Chemistry* (1998) **379** (6): 633-646.

Rivero R.M., Kojima M., Gepstein A., Sakakibara H., Mittler R., Gepstein S. and Blumwald E. **Delayed leaf senescence induces extreme drought tolerance in a flowering plant** *Proceedings of the National Academy of Sciences* (2007) **104** (49): 19631-19636.

Rivero R.M., Gimeno J., Van Deynze A., Walia H. and Blumwald E. **Enhanced cytokinin synthesis in tobacco plants expressing P-SARK:IPT prevents the degradation of photosynthetic protein complexes during drought** *Plant Cell Physiology* (2010) **51**: 1929–1941.

Robatzek S. and Somssich I.E. **Targets of AtWRKY6 regulation during plant senescence and pathogen defense** *Genes and Development* (2002) **16**: 1139–1149.

Rubinstein B. and Leopold A.C. **The Nature of Leaf Abscission** *Quarterly Review of Biology* (1964) **39** (4): 356-372.

Rubio-Somoza I. and Weigel D. **MicroRNA networks and developmental plasticity in plants** *Trends in Plant Science* (2011) **16** (5): 258-264.

Schmutz J., Cannon S.B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D.J., Song Q., Thelen J.J., Cheng J., Xu D., Hellsten U., May G.D., Yu Y., Sakurai T., Umezawa T., Bhattacharyya M.K., Sandhu D., Valliyodan B., Lindquist E., Peto M., Grant D., Shu S., Goodstein D., Barry K., Futrell-Griggs M., Abernathy B., Du J., Tian J., Zhu L., Gill N., Joshi T., Libault M., Sethuraman A., Zhang X-C., Shinozaki K., Nguyen H.T., Wing R.A., Cregan P., Specht J., Grimwood J., Rokhsar D., Stacey G.,

- Shoemaker R.C. and Jackson S.A. **Genome sequence of the palaeopolyploid soybean** *Nature* (2010) **463**: 178-183.
- Schreier, L. and Chandlee J.M. **Characterization of MADS-Box Genes of Soybean (*Glycine max*, *L. Merr.*) and Initial Functional Analysis in Whole Plant Senescence.** *Plant Biology* (2009)
- Simmonds N.W. **Monocarpy, Calendars and Flowering Cycles in Angosperms** *Kew Bulletin* (1980) **35** (2): 235-245.
- Smart M. C. **Tansley Review No. 64 Gene expression during leaf senescence.** *New Phytologist* (1994) **126**: 419-448.
- Stracke R., Werber M. and Weisshaar B. **The R2R3-MYB gene family in *Arabidopsis thaliana*** *Current Opinion in Plant Biology* (2001) **4**: 447-456.
- Swartzberg D., Dai N., Gan S., Amasino R., and Granot D. **Effects of Cytokinin Production under Two SAG Promoters on Senescence and Development of Tomato Plants** *Plant Biology* (2006) **8**: 579-586.
- Swartzberg D., Kirshner B., Rav-David D., Elad Y. and Granot D. **Botrytis cinerea induces senescence and is inhibited by autoregulated expression of the IPT gene** *European Journal of Plant Pathology* (2007) **120** (3): 289-297.
- Walker A.B., Sidor I.F., O'Keefe T., Cremer M. and Berlinsky D.L. **Partial Replacement of Fish Meal with Soy Protein Concentrate in Diets of Atlantic Cod** *North American Journal of Aquaculture* (2010) **72** (4): 343-353.
- Winichayakul S., Moyle R.L., Ryan D.J., Farnden K.J.F., Davies K.M., and Coupe S.A. **Distinct cis-elements in the *Asparagus officinalis* asparagine synthetase promoter respond to carbohydrate and senescence signals** *Functional Plant Biology* (2004) **31**: 573-582.
- Woo H.R., Kim J.E., Kim J., Kim J., Lee U., Song I-J., Kim J-H., Lee H-Y., Nam H.G. and Lim P.O. **The RAV1 transcription factor positively regulates leaf senescence in *Arabidopsis*** *Journal of Experimental Botany* (2010) **61** (14): 3947-3957.
- Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li1, Z., Xiaohui1, D, Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G. and Li-Jia, Q. **The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family** *Plant Molecular Biology* (2006) **60**: 107-124.

Ye J., Coulouris G., Zaretskaya I., Cutcutache I., Rozen S., and Madden T.L. **Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction** *BMC Bioinformatics* (2012) **13**: 134.

Yuan Z., Pan A., Jian Z., Xu S., Gan S., Huang H., Zhang D. **Senescence delay characterization of transgenic *Brassica chinensis* L. containing an anti-senescence chimeric gene SAG12-IPT** *Journal of Plant Physiology and Molecular Biology* (2002) **28**: 379-384.

Zhang X., Ju H.W., Chung M.S., Huang P., Ahn S.J. and Kim C.S. **The R-R-type MYB-like transcription factor, AtMYBL, is involved in promoting leaf senescence and modulates an abiotic stress response in *Arabidopsis*** *Plant Cell Physiology* (2011) **52** (1): 138-148.

Zhang Z.J., Chen J., H.T. Nguyen. **Auto-regulated expression of bacterial isopentenyltransferase gene promotes T-DNA transformation in soybean** *United States Patent Application Publication* (2008) Publication No. US 2008/0184393 A1.

Chapter I

Candidate Gene Discovery for Primary Regulatory Genes of Soybean Senescence Using a Map Based Approach

by

Robert Almeida¹ and Joel M. Chandler²

is prepared in the format specified by *Plant Physiology*

**¹MS Candidate, Department of Cell and Molecular Biology, The University of
Rhode Island, Kingston, RI 02881. Email: ralme1@my.uri.edu**

**²Professor, Department of Cell and Molecular Biology, The University of Rhode
Island, Kingston, RI 02881. Email: jmchandlee@gmail.com**

Abstract

Plant senescence is a genetically determined developmental program characterized by systematic degradative processes that involves activation of new gene activity and down-regulation of other genes that ultimately leads to cell, tissue, organ and whole plant death. Elucidating senescence regulatory pathways and participating genes will allow for the development of strategies to improve crop yields and also curtail post-harvest losses. Three genes are known to be primary regulators of senescence in soybean; namely, *g*, *D1*, and *D2*. In double and triple mutant combinations these genes confer an evergreen leaf and seed phenotype. The double mutation *d1d1d2d2* shows an inhibition of degradation of chlorophyll and chlorophyll binding protein, but photosynthesis declines and the leaves still abscise. In the triple mutant *GGd1d1d2d2* the leaves maintain the normal photosynthetic capacity, but still abscise. So, while the senescence program is not entirely blocked in the mutant background, it is altered. Studies have shown that the expression patterns of soybean senescence associated genes (SAGs) are regulated differentially by *g*, *D1* and *D2*. Due to the pivotal regulatory nature of these three genes for senescence, it is important to identify their specific nature. As such, an analysis using available soybean genome resources (SoyBase, Phytozome, COGE, etc.) was undertaken. This has resulted in the identification of a gene, Glyma01g41610.2, which encodes a putative transcription factor residing within the marker boundaries of the *D1* locus on chromosome 1 which also shares a high level of synteny with a region on chromosome 11 and includes a paralogous gene, Glyma11g03770.2, within the *D2* marker boundaries. Similarly, another gene, Glyma01g00510.1, also encodes a putative transcription factor and is located within the marker boundaries of the *g* locus on chromosome 1. These genes

were selected as candidates representing *g*, *D1*, and *D2* for RT-PCR analysis. None of the initial candidates exhibited a differential expression profile when comparing wild-type and mutant allelic versions in isogenic genetic backgrounds. The *D2* and *g* genes were then selected for Sanger sequencing to determine if sequence differences were responsible for the observed phenotypic variations. No differences in sequence were observed when comparing wild type and mutant allelic forms. However, sequence variations were observed when comparing *g* in Harosoy versus the reference genome cultivar, Williams 82. A second candidate *g* gene (Glyma01g00520.4) was selected from the defined marker boundary interval on chromosome 1 but did not demonstrate a differential pattern of expression using RT-PCR.

Introduction

Senescence is the sequence of biochemical and physiological events comprising the final stage of development for a plant tissue, organ or whole plant, from the mature, fully developed state until death. The changes that take place in senescence represent a genetically programmed sequence, with close coordination at the cell and tissue levels (Smart, 1994). During senescence, leaf cells undergo orderly changes in cell structure, metabolism, and gene expression which ultimately results in redistribution of nutrients to growing and reproductive organs before death and ultimately whole leaf abscission (Balazadeh et al., 2008). Thus, although senescence is a deleterious process for the leaf organ, it critically contributes to the fitness of the whole plant by ensuring optimal production of offspring and better chance of survival. Leaf senescence is thus an evolutionarily selected developmental process and comprises an important phase in the plant life cycle (Lim et al., 2007).

Soybean (*Glycine max L. Merr.*) is an agriculturally and economically important crop (<http://www.soystats.com>). During its annual life cycle, soybean senescence occurs primarily in the leaves which appear to be the target of the senescence-inducing influence from the seeds. This system of leaf death may be true for other species as well (Nooden, 1988) making soybean a candidate model for basic/fundamental studies.

The ability to regulate gene transcription is central both to tissue specific-gene expression and to the regulation of gene activity in response to specific stimuli (Latchman, 1997). Leaf senescence is an active process involving the differential expression of hundreds of genes and therefore it is presumed that numerous

transcription factors are involved as central elements of the regulatory network (Woo et al., 2010). Three different genes (*g*, *D1*, *D2*) are known to be involved with regulating the whole plant/leaf senescence program of soybean because mutations contribute to an evergreen (non-senescent) phenotype. Homozygous *d1d1d2d2* lines remain green and show an inhibition of chlorophyll degradation and chlorophyll-binding proteins yet they still undergo a decline in photosynthetic activity and leaf abscission (Guamet et al.1991, Canfield et al. 1995). When combined with the dominant mutant *G* (*GGd1d1d2d2*) a decline in photosynthetic activity does not take place, but the leaves still abscise. It has been suggested that *d1d2* may control a central regulatory process in the senescence program and that homozygosity at both nuclear loci is required because the two are homeologous (duplicate) loci in the ancient tetraploid soybean genome (Guamet and Giannibelli 1996). The genes responsible for these primary regulators remain unknown; however, because of their important role in the upstream regulation of the senescence program it is plausible they encode transcription factors.

Using genetic and physical maps it's possible to identify putative candidates for the transcription factor genes representing *g*, *D1*, and *D2*. By employing a map based approach the current study aimed to identify transcription factors that may ultimately represent the primary soybean regulator genes. For analysis of these three genes, near-isogenic lines are available as single, double and triple mutant combinations and provided the basis for expression analysis of these candidate transcription factor genes. To examine a functional role for these candidates, RT-PCR

was performed with candidates derived from the map-based strategy. Candidates not demonstrating differential expression were analyzed by Sanger sequencing.

Materials and Methods

Selection of Candidate Genes

The putative identity for *g*, *D1*, *D2*, the three genes responsible for regulating the soybean senescence program, was determined using genetic maps available through SoyBase. To resolve all possible candidate genes for *g*, the chromosome 1 (D1a) soybean 2003-composite map (www.soybase.org) was used to identify DNA markers that straddled a chromosome block surrounding *g* (Figure 1.1a). The position of *g* near the telomeric end of chromosome 1 (Figure 1.1a), and the lack of specification of map units in SoyBase, prevented selecting a marker on this telomeric end by this approach. To overcome this problem, the first base pair of chromosome 1 was used as a border marker to ensure all possible candidate genes for *g* were captured. The marker, SAT_332, representing a simple sequence repeat (SSR) motif of (AT)₂₅, was used to set the other marker boundary toward the centromeric end of the chromosome with an end position of 355,784 base pairs. This marker was then verified to be in the correct flanking orientation using a newer version of the soybean genetic map (consensus 4.0 map) that saturates the soybean genome with SNP markers for better resolution. This updated version of the soybean genetic map was then compared with a current sequence map in SoyBase to verify that the marker selected actually was located flanking *g* (Figure 1.1a).

Next, this information was used with the Comparative Genomics (CoGe) OrganismView tool to screen the *Glycine max* genome for genes present within the chromosomal block bordered by the marker boundaries. The start and end physical locations of the boundary markers were entered for chromosome 1 and used to extract

the most recently annotated coding sequences (CDS) (soybean v1.1) between them and entered into a spreadsheet database. Forty-two (42) gene annotations were identified between these markers. However, because some of these represented alternative transcripts, a total of 32 discrete genes were identified for screening (Figure 1.2a). All genes were screened for domains with putative transcription factor identity.

To identify *DI*, the chromosome 1 (D1a) soybean 2003-composite map in SoyBase was used to find DNA markers that could be used as border markers delineating a chromosomal interval containing *DI* (Figure 1.1b). A centromeric end marker, BARC-030807-06945, whose sequence starts at 53,063,806 base pairs, and a telomeric end marker, Sat_160, whose sequence end is at 53,236,862 base pairs, were selected as the two closest markers with available genomic sequence information flanking the *DI* gene. These markers were then verified to be in the correct flanking orientation using the newer version of the soybean genetic map (consensus 4.0 map). This updated version of the soybean genetic map was then compared with a current sequence map in SoyBase to verify that the markers selected actually were flanking *DI* (Figure 1.1b).

This information was used with the CoGe OrganismView tool screen the *Glycine max* genome for annotated genes within the border boundaries. The start and end physical locations of the boundary markers were entered for chromosome 1 and used to extract the most recently annotated CDS (soybean v1.1) between them and were subsequently entered into a spreadsheet database. Twenty-eight gene annotations were found between these border markers. However, these included 7 alternative

transcripts leaving a total of 21 discrete genes to be screened (Figure 1.2b). All genes were then analyzed for domains with putative transcription factor identity.

As previously, to identify *D2*, the chromosome 11 (B1) soybean 2003-composite map in SoyBase was used to find DNA markers that could be used as borders delineating a chromosomal interval containing *D2* (Figure 1.1c). A telomeric end marker, BARC-029533-06211, whose sequence begins at 546,754 base pairs and a centromeric end marker, Sat_272, whose sequence end is at 2,710,583 base pairs, were selected as the two closest markers flanking the *D2* gene on either side. These markers were then verified to be in the correct flanking orientation using the newer version of the soybean genetic map (consensus 4.0 map). This updated version of the soybean genetic map was then compared with a current sequence map in SoyBase to verify that the markers selected actually were flanking *D2*.

Next, this information was used with the CoGe OrganismView tool to screen the *G. max* genome for annotated genes present within the border boundaries as previously done with the *D1* gene interval. The start and end physical locations of the boundary markers were entered for chromosome 11 and used to extract the most recently annotated CDS (soybean v1.1) between them which were subsequently entered into a spreadsheet database. Four hundred twenty-three gene annotations were found between these border markers. However, these included 133 alternative transcripts leaving a total of 290 genes to be screened (Figure 1.2c). All genes were then analyzed for domains with putative transcription factor identity.

Due to the presumed nature of the *D1* and *D2* duplication, both Synmap and Synfind (CoGe) were used to determine if any transcription factor genes resided within the boundaries delineated for each gene to support the possibility of *D1* and *D2* representing a gene duplication event. To determine if chromosomes 1 and 11 shared syntenic regions, Synmap was first used to determine soybean whole genome synteny (Figure 1.3). Next, Synfind was used to determine if Glyma01g41610.2 (chromosome 1) and Glyma11g03770.2 (chromosome 11) located in the syntenic regions of these two chromosomes shared high levels of synteny at the gene level by entering these genes into the “specify feature” section and selecting the most recent annotated CDS version. This was then searched against *G. max* under the Specify Organism tab in two separate trials (*D1* search for *D2* vs. *D2* search for *D1*) with default settings (Figure 1.4).

Primer Design

Primers were designed using Primer 3 software, (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) available through the NCBI portal, and the available nucleotide sequences of selected candidate genes. Primers used for RT-PCR expression analysis were designed to span the 5' or 3' UTR and exon junction for unique *g*, *D1*, and *D2* products ranging from 100 to 300 base pairs (Table 1.1). Primers were designed to amplify a template for sequencing the *g* candidate gene by obtaining a region of 150 base pairs upstream and 200 base pairs downstream of the genomic sequence (Table 1.2). Primers were designed to amplify a template for sequencing the *D2* candidate gene by obtaining a region of 1500 base pairs upstream and 1600 base pairs downstream of the genomic sequence (Table 1.3). Sequencing

primers were generated from within the amplified template of both *g* (Table 1.4) and *D2* (Table 1.5). Default settings for the Primer3 program were used except for: 1) the “Exclusion tab” which was then selected to prevent redundant sequences from appearing in the Results field by excluding Refseq transcripts of predicted mRNA and ncRNA and 2) the “Organism” field which was selected for “*G. max.*” The 5 sets of results returned were then used to query the Phytozome soybean database using TBLASTN to ensure the sequence was located within the gene of interest and also to verify the primers were unique in the genome. Primers were also screened against Harosoy genomic DNA to verify amplification of the target genomic sequences in that strain.

RNA Extraction

Total leaf RNA was isolated from several developmental stages throughout the normal life cycle using Harosoy (Table 1.6) and isogenic lines harboring genes that affect leaf senescence (Table 1.7). Leaves were harvested immediately into liquid nitrogen and RNA was extracted the same day, using a standard phenol/chloroform extraction protocol and LiCl precipitation (Maniatis et al., 1986; Schreier and Chandlee, 2009). RNA was quantified using a spectrophotometer and assayed for quality and quantity by gel electrophoresis in formaldehyde-containing agarose gels (Figure 1.5). The soybean developmental stages V5 and R7 were used in this study. V5 (vegetative 5) is characterized by five fully expanded, green trifoliolate leaves are found on the plant and in R7 (reproductive 7) one major pod has changed to a brown color on the main stem (<http://www.ag.ndsu.edu/pubs/plantsci.htm>).

DNA Extractions

Whole genomic DNA was extracted using a standard protocol.

RT-PCR Analysis

Preliminary RT-PCR analysis was performed with total RNA from stages V5 and R7 of soybean using seven different isogenic backgrounds (Table 1.7). RT-PCR analysis was performed using 15 units of Invitrogen Superscript RIII reverse transcriptase (Life Technologies, Carlsbad CA) and buffer in conjunction with the AccessQuick RT-PCR System (Promega, Madison WI) and 1 μ g of total RNA in each reaction. Reaction controls were performed with actin-specific (*Sac3*) primers (Table 1.1) on each of the RNA samples tested. The Reverse Transcription cycle was run at 50°C for 50 min for one cycle (Eppendorf Thermocycler Model 5331). PCR was carried out as follows: an initial denaturation step at 85°C for 5 min; 40 cycles with 1 min at 94°C, 1 min at 48°C, and 2 min at 72°C; final extension step of 72°C for 7 min followed by; a hold at 4°C. This program was used for all the primer sets analyzed. The products were screened using a 2% molecular biology grade agarose (Fisher Scientific) gel in 100mM Tris-acetate and 2 mM EDTA.

g and *D2* Template Amplification

PCR analysis was performed with total DNA from Harosoy (*gD1D2*)/ L64-2489 (*Gd1d2*) for the *g* candidate and Harosoy/ L69-4266 (*gD1d2*) for the *D2* candidate. PCR analysis was performed using a Phusion High Fidelity Taq Polymerase kit (New England Biolabs, Ipswich MA) with 120 ng total DNA. Reaction controls had no template. PCR was carried out as follows for the putative *g* gene: an initial

denaturation step at 98°C for 30 sec; 30 cycles with 7 sec at 98°C, 20 sec at 62.5°C, and 3 min 30 sec at 72°C a; final extension step of 72°C for 7 min followed by; a hold at 4°C. PCR was carried out as follows for the putative *D2* gene: an initial denaturation step at 98°C for 30 sec; 30 cycles with 7 sec at 98°C, 30 sec at 71°C, and 2 min 30 sec at 72°C a; final extension step of 72°C for 7 min followed by; a hold at 4°C.

Sequence Analysis

Verification of the identity of the amplicons was performed by sequence analysis. Bands were excised from the gels and purified with the Wizard SV Gel and PCR clean-up system according to the manufacturer's instructions (Promega, Madison WI). Bands were sequenced using the facilities at the University of Rhode Island's Genomics and Sequencing Center (URIGSC).

Results and Discussion

g Gene Analysis

Analysis of domains with putative transcription factor identity for the 32 genes identified between selected markers resulted in the identification of only one gene, Glyma01g00510.1, encoding an auxin response transcription factor (ARF). This product is able to form hetero and homodimers (Okushima et al., 2005) which could explain the dominant nature of the *g* mutation. A Harosoy isolate for the single mutation of *g* was not available for direct comparison of the *g* expression profile relative to Harosoy. Therefore, a comparison was made between early (V5) and late (R7) developmental time points of wild type Harosoy (*gDID2*) vs. mutants with the available genetic backgrounds: L69-4265 (*Gd1* double mutant), L69-971 (*Gd2* double mutant), L69-4267 (*d1d2* double mutant) and L64-2489 (*Gd1d2* triple mutant) (Table 1.7) to elucidate possible differences in the expression profiles of *g* using RT-PCR. However, after functional analysis, there appeared to be no significant difference in expression of these transcripts in early and late development in any of the genetic backgrounds (Figure 1.6). This may be attributed to the lack of availability of an isolate harboring a single *g* mutation.

To determine if an alteration in the coding sequence of Glyma01g00510.1 could account for the nature of the *g* mutation, Sanger sequencing was utilized. DNA templates from Harosoy and from the L64-2489 triple mutant were used and compared to the Williams 82 cultivar sequence available through www.phytozome.net. No differences were identified between Harosoy and the L64-2489 genomic sequences.

However, four deletions totaling 21 base pairs and 24 SNP's were identified between these sequences and the Williams 82 sequence (Figure 1.7).

Due to the lack of evidence implicating Glyma01g00510.1 as the *g* gene responsible for the evergreen phenotype, another gene, Glyma01g00520.4, located between the selected markers encoding a methyl-CPG-binding domain protein 02, was examined. A search of the LSD implicated an *A. thaliana* gene with this annotation in the senescence program (Li et al., 2010). Functional analysis of this gene demonstrated no detectable difference in expression of these transcripts in early and late development in any of the genetic backgrounds examined (Figure 1.8) suggesting it was not involved in the mutant phenotype.

D1 and *D2* Gene Analysis

Marker boundary mapping of *D1* and *D2* identified a total of 22 *D1* and 290 *D2* genes. Analysis of these genes for domains with putative transcription factor identity revealed 2 within the *D1* boundaries and 41 within the *D2* boundaries. One of the two genes identified as a transcription factor, Glyma01g41610.2, encoding a soybean MYB gene located between *D1* boundaries, was used to identify a paralogous gene within a highly syntenic region on chromosome 11, Glyma11g03770.2. Glyma11g03770.2 also putatively encodes a soybean MYB gene located between the defined *D2* boundaries (Figures 1.3 and 1.4). Interestingly, the best match in *A. thaliana* for Glyma11g03770.2 using the Synfind program within CoGe was *AtMYB5* where the promoter has been demonstrated to play a role in plant senescence (Heazelwood et. al, 2011). For the putative *D1* gene, a comparison was made between

early (V5) and late (R7) developmental time points of wild type Harosoy (*gD1D2*) vs. L73-54 (*d1* single mutant), L69-4267 (*d1d2* double mutant), and L64-2489 (*gd1d2* triple mutant). For the putative *D2* gene, a comparison was made between early (V5) and late (R7) developmental time points of wild type Harosoy (*gD1D2*) vs. L69-4266 (*d2* single mutant), L69-4267 (*d1d2* double mutant), and L64-2489 (*gd1d2* triple mutant) (Table 1.7). RT-PCR was used to characterize the expression profile of both *D1* and *D2* in these genetic backgrounds (Figure 1.9 and 1.10 respectively). However, there appeared to be no significant difference in the expression of these transcripts in early and late development in any of the genetic backgrounds.

To determine if an error in the coding sequence of Glyma11g03770.2 could account for the nature of the *D2* mutation, Sanger sequencing was utilized. DNA templates from Harosoy and the L69-4266 single mutant (*gD1d2*) were used and compared to the Williams82 cultivar sequence available through www.phytozome.net (Goodstein et al., 2012). No differences were identified between the Harosoy and L64-2489 genomic sequences and the model cultivar Williams82 (Figure 1.11).

Although the results obtained in this study did not lead to the anticipated outcomes, they do open up other avenues of inquiry that warrant further investigation. The method for candidate gene identification outlined in this study represents a logical strategy that allows for selection of a gene(s) that can be screened quickly and cost effectively by trial and error compared to the map based cloning technique. Currently, an alternative putative *g* gene, Glyma01g00520.4, remains to be sequenced. There exists the possibility that an error in the coding region of this gene could adversely

affect gene functionality and therefore, the senescence developmental program (Zemach and Grafi, 2006).

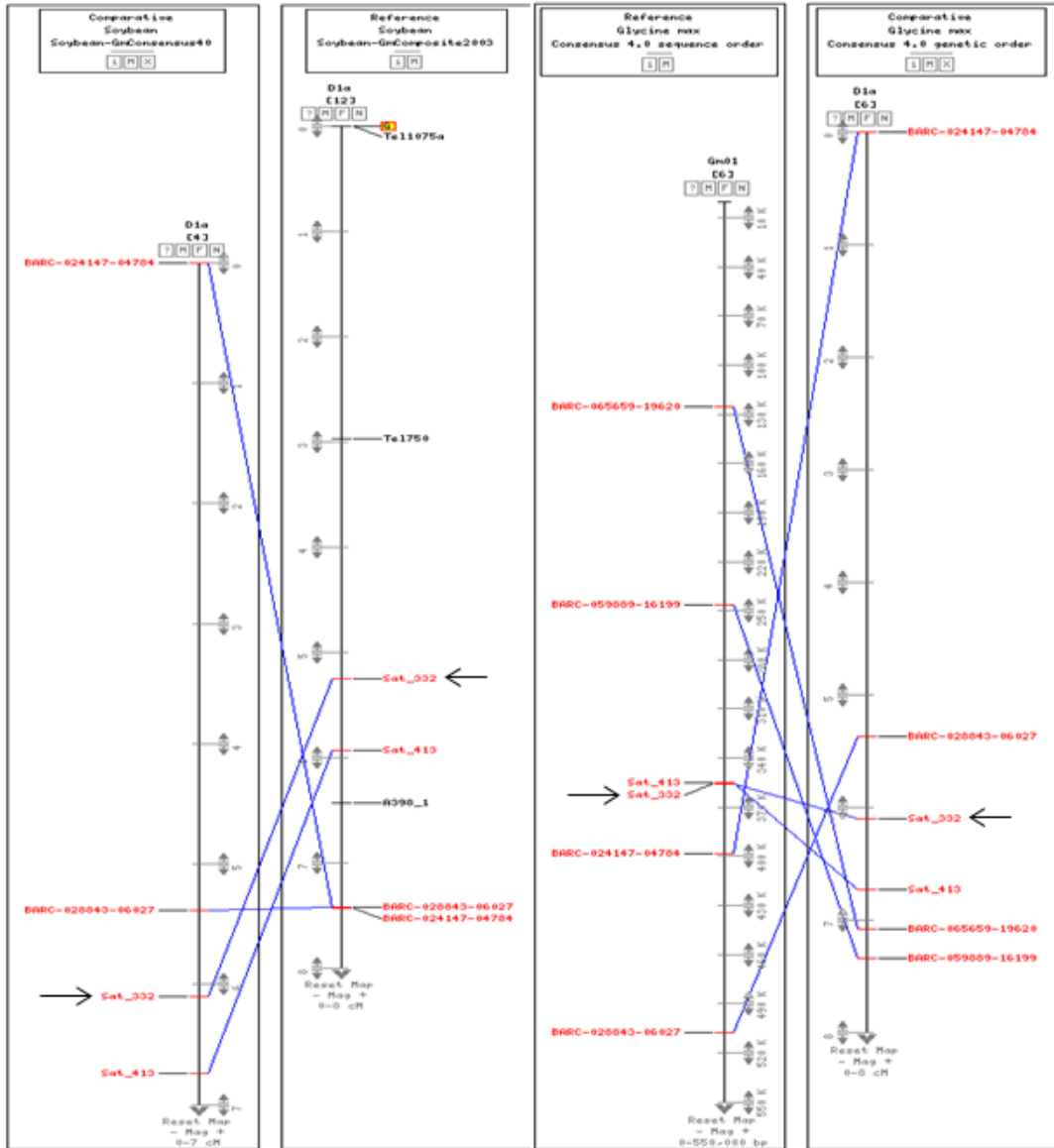
The Fast Neutron Database, a resource of the soybase.org repository may aid in future gene selection. Fast neutron radiation was used to induce deletion mutations in the soybean genome followed by cataloging plant variation for seed composition, maturity, morphology, pigmentation, and nodulation traits (Bolon et al., 2011). Future updates to the database may contribute to the identification of genes relevant to senescence.

Future advances in the development of other resources such as with soybean TILLING (Targeting Induced Local Lesions IN Genomes), which provides a high-throughput reverse genetic method to obtain an allelic series from a chemically mutagenized population (Cooper et al., 2008), and a Soybean Transposon Insertion Mutant Database which allows for functional analysis of soybean genes through transposon mutagenesis (Mathieu et al., 2009), could also offer valuable tools that could be adapted to aid in identifying the primary regulatory genes involved in senescence and their function.

1.1a)

Linkage Maps

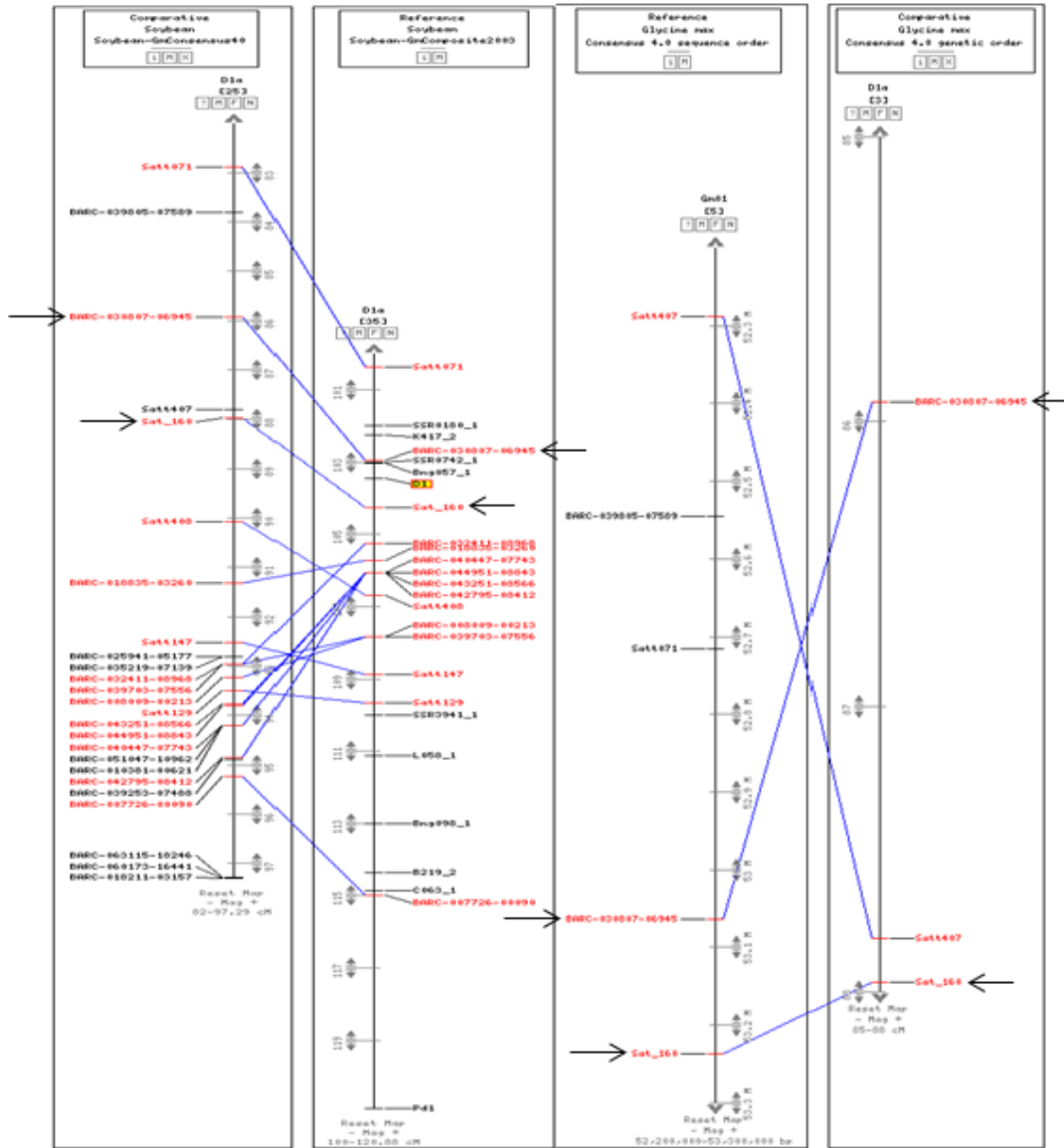
Genetic and Sequence Map Comparison



1.1b)

Linkage Maps

Genetic and Sequence Map Comparison



1.1c)

Linkage Maps

Genetic and Sequence Map Comparison

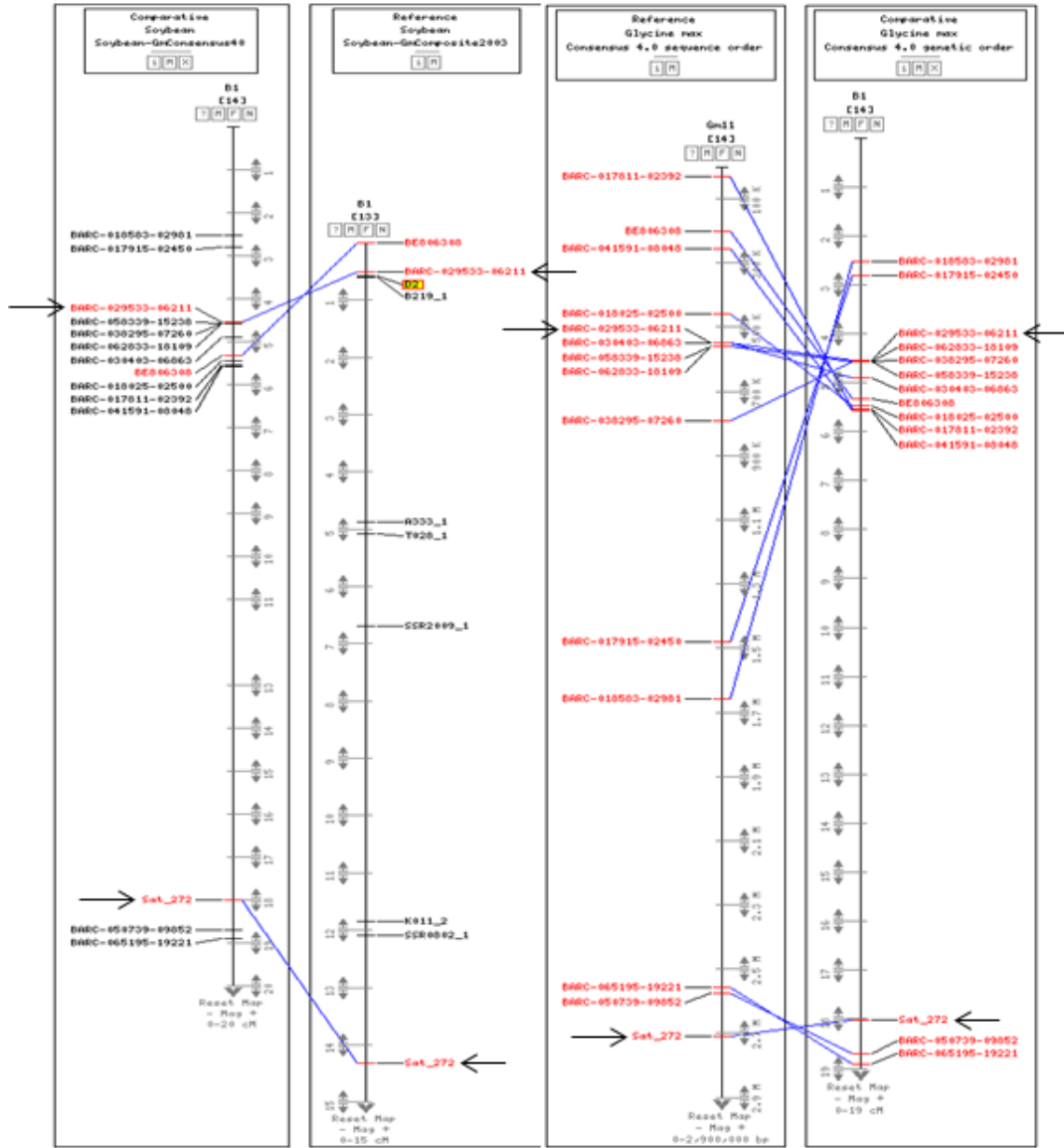
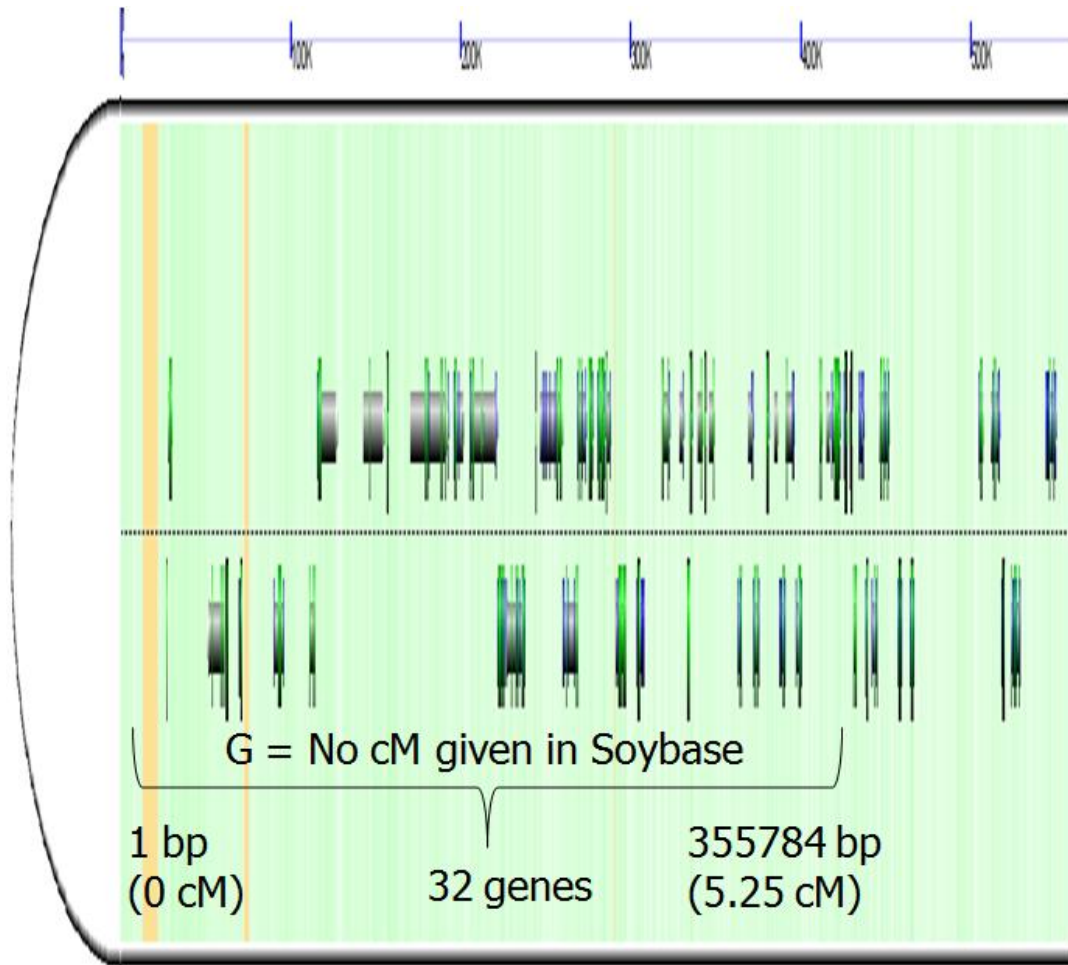


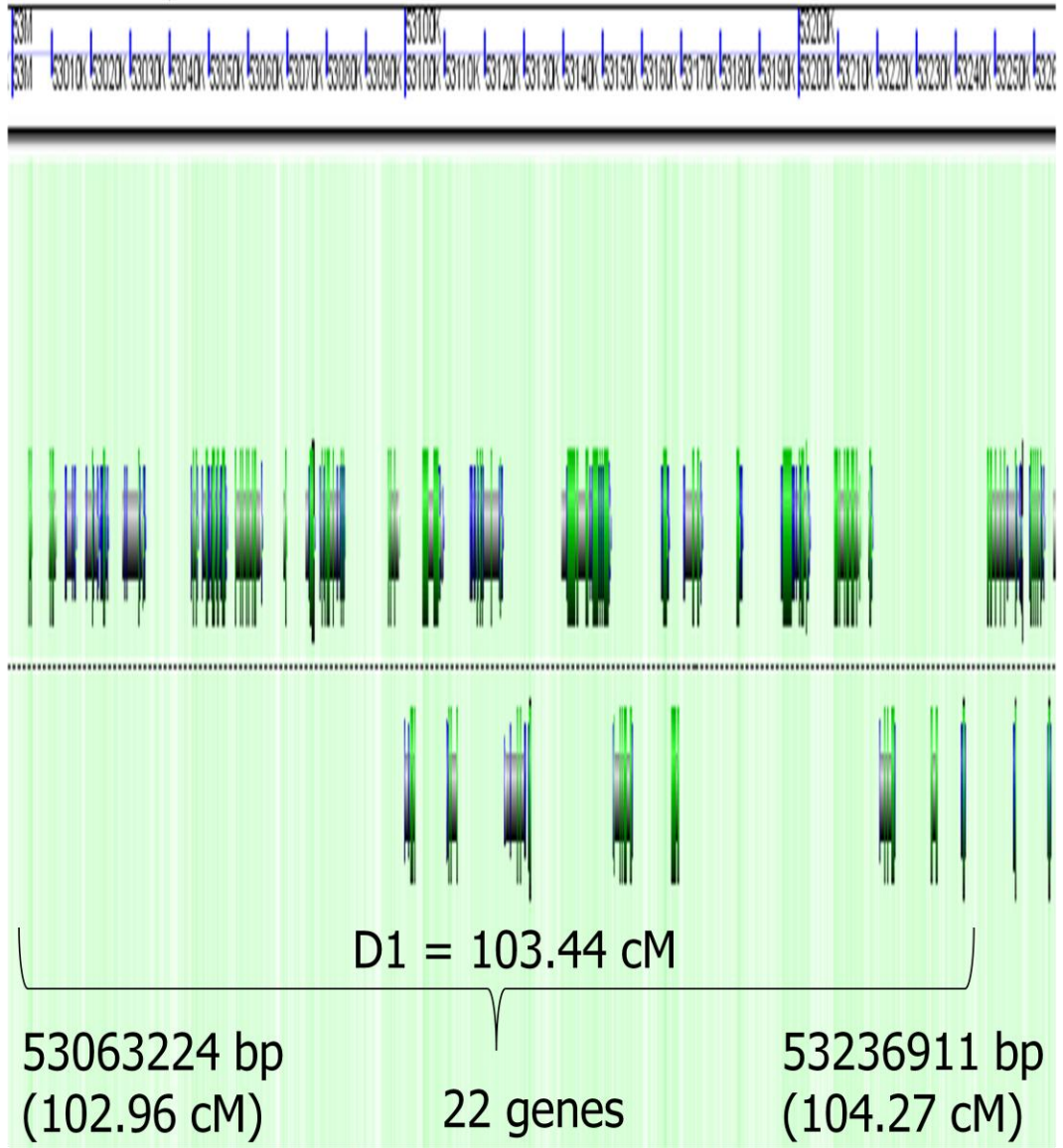
Figure 1.1a-c. Linkage and sequence map comparisons

a) The first base pair of chromosome 1 is used along with the centromeric end marker SAT_332 (arrows) to define a chromosomal block that contains the *g* gene on the 2003 composite and consensus 4.0 linkage maps (left). Right panel shows the consensus 4.0 linkage map vs. the physical map on chromosome 1. b) The arrows highlight a telomeric end marker BARC-030807-06945 and a centromeric end marker Sat_160 that flank the *D1* gene on both the 2003 composite and consensus 4.0 linkage maps (left). Right panel shows the consensus 4.0 linkage map vs. the physical map on chromosome 1. c) The arrows highlight a telomeric end marker BARC-029533-06211 and a centromeric end marker Sat_272 that flank the *D2* gene on both the 2003 composite and consensus 4.0 linkage maps (left). Right shows the consensus 4.0 linkage map vs. the physical map on chromosome 11.

1.2a)



1.2b)



1.2c)

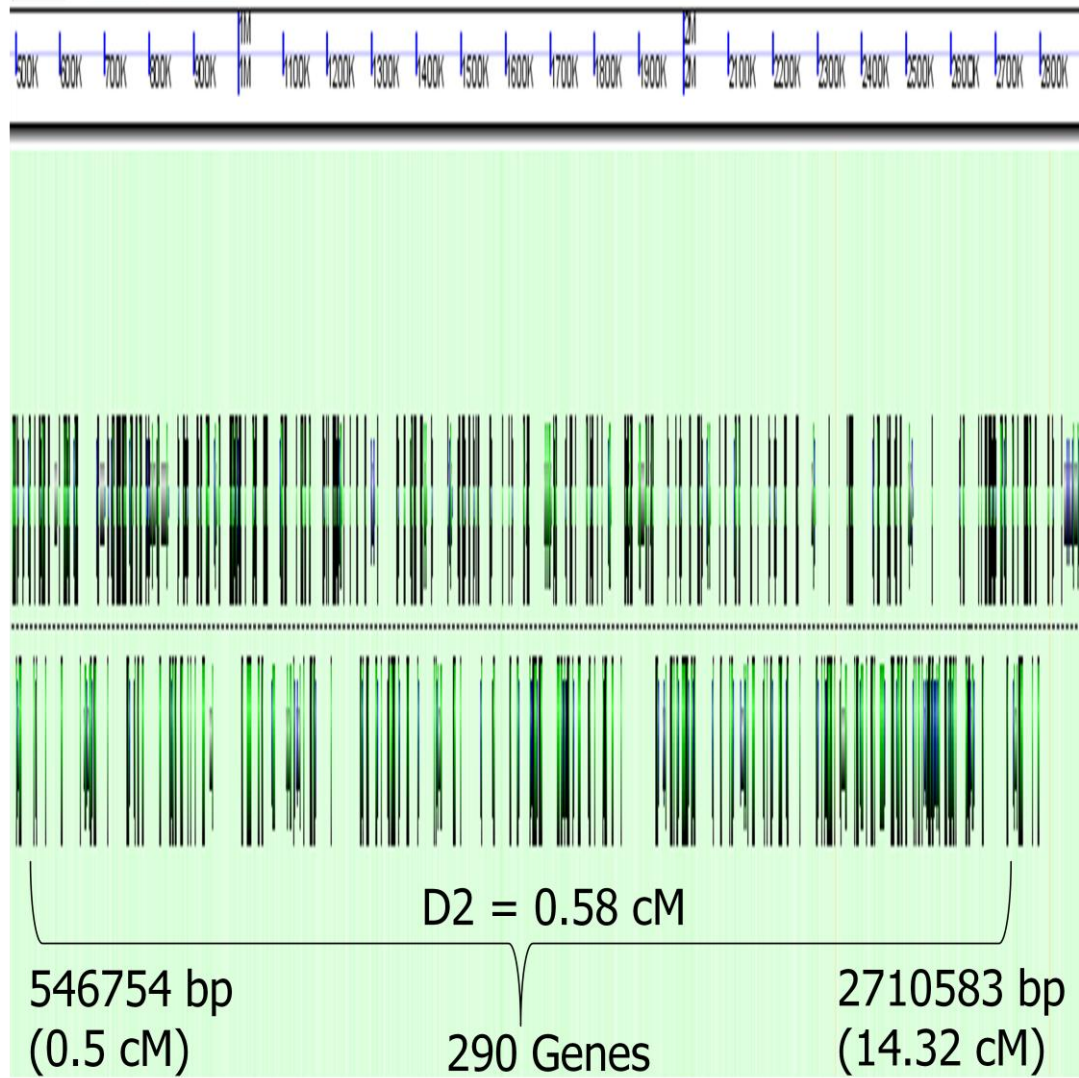


Figure 1.2a-c. Physical maps

a) A physical map of chromosome 1 showing the 32 annotated genes residing between the region defined by the first base of chromosome 1 and the SAT_332 marker selected for analysis using the linkage maps. b) A physical map of chromosome 1 showing the 22 annotated genes residing between the region defined by the telomeric end marker BARC-030807-06945 and the centromeric end marker Sat_160 selected for analysis using the linkage maps. c) A physical map of chromosome 11 showing the 290 annotated genes residing between the region defined by the telomeric end marker BARC-029533-06211 and the centromeric end marker Sat_272 selected for analysis using the linkage maps.

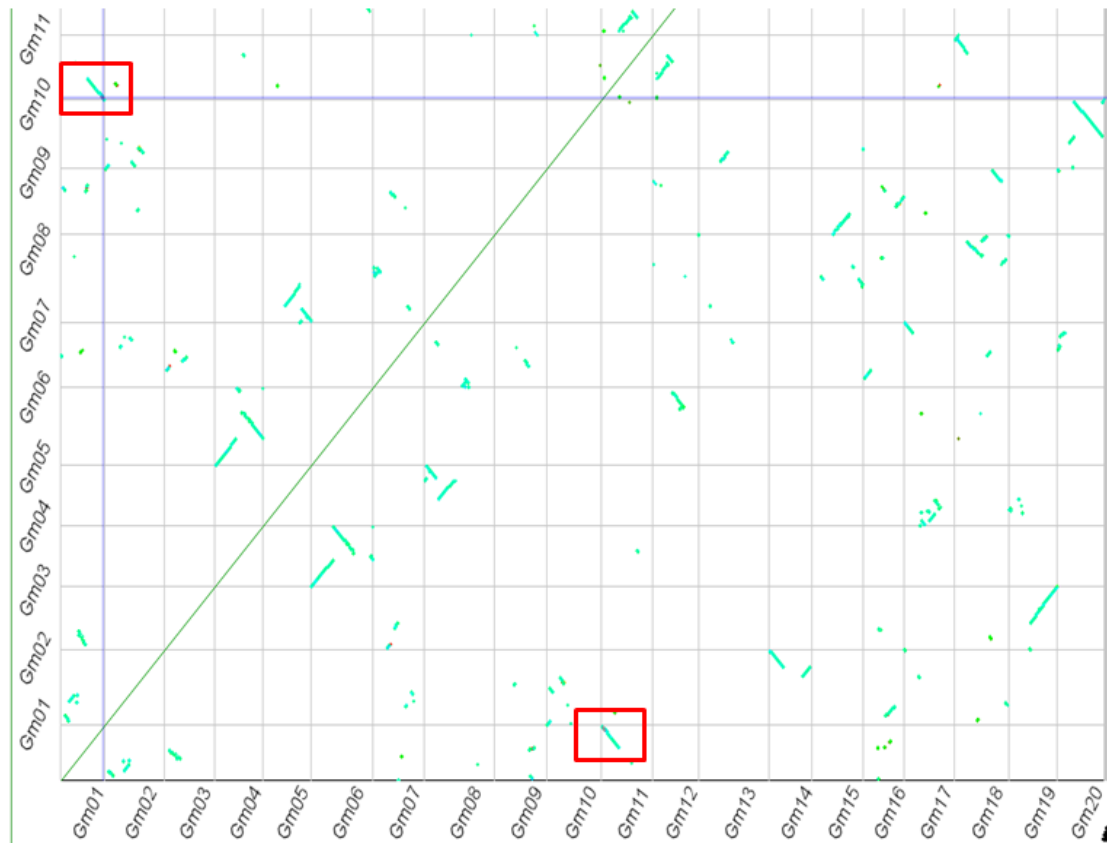


Figure 1.3. Soybean whole genome syntenic plot

Whole genome synteny alignment of soybean vs. soybean genes. The boxed regions highlight duplicated regions of chromosome 1 and 11 where both D1 and D2, respectively, reside. Within Synmap, *G. max* was selected for both organisms 1 and 2. The “Analysis Option” tab was selected and default settings used with the exceptions of “Syntenic Depth” and “CodeML” choices. These advanced analytical tools identify orthologous syntenic regions by the relative evolutionary distance of syntenic gene pairs using synonymous mutation rates and the algorithm “Quota Align” for screening syntenic regions to enforce a specific mapping of syntenic regions between genomes. The “Quota Align” option was selected for Syntenic Depth using a ratio of 1 to 1 coverage depth. The “CodeML” option was then changed from the default selection of “none” for calculating syntenic CDS pairs and color dots to “Synonymous (Ks)” substitution rates. The “Display Options” tab was also selected and default settings used except in the cases of the “Sorting Chromosomes” and “Dotplot Axes Metric” tabs. The chromosomes were sorted by name and the dotplot metric selected for analysis was genes before generating the Synmap.



Figure 1.4. Soybean syntenic gene alignment

Syntenic region of Glyma01g41610.2 encoding a soybean MYB gene located between D1 boundaries (top pink box) on chromosome 1. SynFind identifies a syntenic region on chromosome 11, Glyma11g03770.2, which also encodes a soybean MYB gene located between D2 boundaries (bottom pink box).

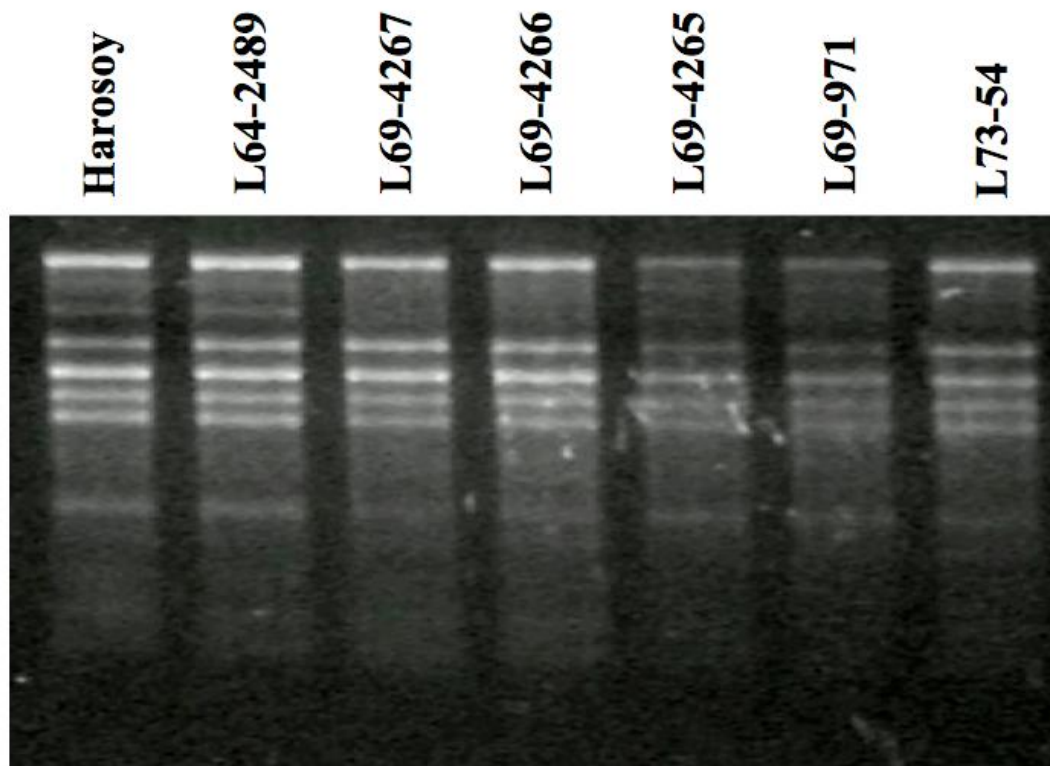
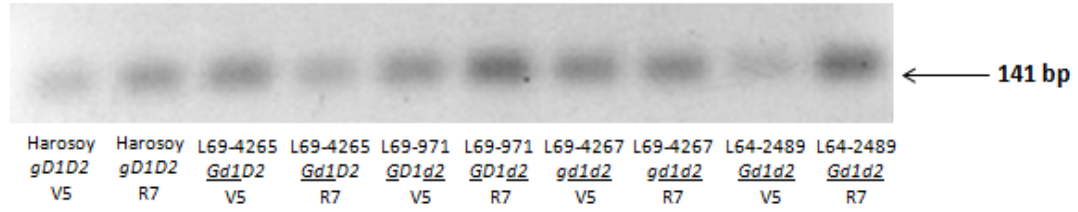


Figure 1.5. Gel electrophoresis of RNA samples extracted from seven isogenic lines of soybean (Schreier and Chandlee, 2009). Shown is a representative agarose gel of RNA extracted from leaf tissue (Table 1). All RNA samples used in this study exhibited similar quality.

a)



b)



Figure 1.6. RT-PCR analysis of a putative *g* gene Glyma01g00510.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and five isogenic lines [WT Harosoy, L69-4265, L69-971, L69-4265, L64-2489] with *g* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

Query 1 TAATTTGATAACTTTTTATGTTTAAATAAGACTGGTGAAGAATAAAAAGAAAGGAGGTGA 60
 Sbjct 1 TAATTTGATAACTTTTTATGTTTAAATAAGACTGGTGAAGAATAAAAAGAAAGGAGGTGA 60

Query 61 GAATGATGGTGTAACGGTGAGAGCGTGGCGTCTGGTCCATAAAGGGAGGTAAACTGA 120
 Sbjct 61 GAATGATGGTGTAACGGTGAGAGCGTGGCGTCTGGTCCATAAAGGGAGGTAAACTGA 120

Query 121 TTGCAGTATGAAGAAGAAGAGCAGCATAAAGCGGAGCTGTGGCACGCTTGTGCGGGGCC 180
 Sbjct 121 TTGCAGTATGAAGAAGAAGAGCAGCATAAAGCGGAGCTGTGGCACGCTTGTGCGGGGCC 180

Query 181 TTTGGTTAAGCTTCCTCCCTCCGGCAGCAGCAGTCATCTACTTCCCTCAAGGTCACAGCGA 240
 Sbjct 181 TTTGGTTAAGCTTCCTCCCTCCGGCAGCAGCAGTCATCTACTTCCCTCAAGGTCACAGCGA 240

Query 241 ACAAGTCTCTGCTTCTCTCAACAGAGATGTCCACTCTCAAATCCCCAACTATCCCAATCT 300
 Sbjct 241 ACAAGTCTCTGCTTCTCTCAACAGAGATGTCCACTCTCAAATCCCCAACTATCCCAATCT 300

Query 301 CCCCTCCAAGCTACTCTGTCTCCTCCACTCTCACTTTGCATGCTGATCCCCAACTGA 360
 Sbjct 301 CCCCTCCAAGCTACTCTGTCTCCTCTACTCTCACTTTGCATGCTGATCCCCAACTGA 360

Query 361 TCAAGTCTACGCTCAGATTACTCTTCAACCCCTCCCTTCTGTACGCTTCTAACCATCCA 420
 Sbjct 361 TCAAGTCTACGCTCAGATTACTCTTCAACCCCTCCCTTCTGTACGCTTCTAACCATCCA 420

Query 421 TTCCTTCTCTCCACTTCCCATTCCTCTCTGTTGTTTATCACATTCATTTAGTTTGACA 480
 Sbjct 421 TTCCTTCTCTCCACTTCCCATTCCTCTCTGTTGTTTATCACATTCATTTAGTTTGACA 480

Query 481 AGGATGCTCTCTTGAGATCTGATCTCGTCTCGAATCCACCAAGCCACCACCTGACTTCT 540
 Sbjct 481 AGGATGCTCTCTTGAGATCTGATCTCGTCTCGAATCCACCAAGCCACCACCTGACTTCT 540

Query 541 TCTGCAAACAACCTACCGCCAGTGATACTAGCACTCACGGTGGTTTCTCCGTCCCTCGCC 600
 Sbjct 541 TCTGCAAACATCTTACCGCCAGTGATACTAGCACTCACGGTGGTTTCTCCGTCCCTCGCC 600

Query 601 GTGCCGCCGAGAAGATTTCCCTCCTCTGTATCTCCATTCACCCCGTTTCCCTTTTCTA 660
 Sbjct 601 GTGCCGCCGAGAAGATTTCCCTCCTCTGTATCTCCATTCACCCCGTTTCCCTTTTCTA 659

Query 661 ACCACTTCACTCATTCTTCACTCTTTTCACTAATTCTTTAATCCCCGCTTCTGTAGGAT 720
 Sbjct 660 ACCCTTCACTCATTCTTCACTCTTTTCACTAATTCTTTAATCCCCCTTCTGTAGGAT 719

Query 721 TATTCTATGCAACCTCCTGCTCAAGAACTCGTCGCCAGGGATTTGCACGACACCGTTTGG 780
 Sbjct 720 TATTCTATGCAACCTCCTGCTCAAGAACTCGTCGCCAGGGATTTGCACGACACCGTTTGG 779

Query 781 AAATTCGCCACATATAACCGTGGTACGTTTTCAGACTCCAAGGAAATGTAAGGTTGATCT 840
 Sbjct 780 ACATTCGCCACATATAACCGTGGTACGTTTTCAGACTCCAAGGAAATGTAAGGTTGATCT 839

Query 841 AGTAGTGAAGGGTGAAGGAAAGGGGAAGATGTCGCGGTTTGAATCCTTCCCTAACAAA 900
 Sbjct 840 AGTAGTGAAGGGTGAAGGAAAGGGGAAGATGTCGCGGTTTGAATCCTTCCCTAACAAA 899

Query 901 CAATACTAACAACTAATATTTGTCCATCCAAAATAAAAACGTTTTCAGACTCCAACAAT 960
 Sbjct 900 CAATACTAACAACTAATATTTGTCCATCCAAAATAAAAACGTTTTCAGACTCCAACAAT 959

Query 2821 TGGGAACAGTCCTGCCGCCGAGCAACAGTTGGGTGATATCACTCAACAAATCGAGGCAGA 2880
 Sbjct 2820 TGGGAACAGTCCTGCCGCCGAGCAACAGTTGGGTGATATCACTCAACAAACCGAGGCAGA 2879
 Query 2881 ACTTGGCAAATCAAACAATCCCTCAGGGTCAAGTTCAGGCCAACTCGTGCATCCCCAGA 2940
 Sbjct 2880 ACTTGGCAAATCAAACAATCCCTCAGGGTCAAGTTCAGGCCAACTCGTGCATCCCCAGA 2939
 Query 2941 ACATTGTCCAAACCAACAATATTCTTCAACAGCAGCAACCATCCAGTCAAACCATCAAC 3000
 Sbjct 2940 ACATTGTCCAAACCAACAATATTCTTCAACAGCAGCAACCATCCAGTCAAACCATCAAC 2999
 Query 3001 TACATAGAAGCCTCTCTCAGAACCCATCACAGCAGCAGCAGACAATTATTGGTCAGA 3060
 Sbjct 3000 TACATAGAAGCCTCTCTCAGAACCCATCACAGCAGCAGCAGACAATTATTGGTCAGA 3059
 Query 3061 ATCAACACCAAATTTGATACAGTCTCTATGCCTGATCATGTTCAACAATTACAGATGT 3120
 Sbjct 3060 ATCAACACCAAATTTGATACAGTCTCTATGCCTGATCATGTTCAACAATTACAGATGT 3119
 Query 3121 CTGACGATCAGATTCAGCTGCAGTTGTTACAGAAGCTTCAACAACAAAAACAAACTCT 3180
 Sbjct 3120 CTGACGATCAGATTCAGCTGCAGTTGTTACAGAAGCTTCAACAACAAAAACAAACTCT 3179
 Query 3181 TGGCCCAACAAACTGCATTGCAGCATTCTACTCAACTTACTCAAATCCAGGATCGGCAAA 3240
 Sbjct 3180 TGGCCCAACAAACTGCATTGCAGCATTCTACTCAACTTACTCAAATCCAGGATCGGCAAA 3239
 Query 3241 GACAGCTTTTAGACAAAACGCATAAATTGTCTAGAGCACTAACACCTGGTCAAGTACGGG 3300
 Sbjct 3240 GACAGCTTTTAGATAAAACGCATAAATTGTCTAGAGCACTAACACCTGGTCAAGTACGGG 3299
 Query 3301 AAATTCCTCCTATTTTTTCAGAATTCACTCCCCAAGGCTAATTCTATCTCAAATCCGATTA 3360
 Sbjct 3300 AAATTCCTCCTATTTTTTCAGAATTCACTCCCCAAGGCTAATTCTATCTCAAATCCGATTA 3359
 Query 3361 CAAAGGCCAATTGCCAGAGCAATATTCAATTCTATCAGCAGCCCAAGCTTCAACAGCAGC 3420
 Sbjct 3360 CAAAGGCCAATTGCCAGAGCAATATTCAATTCTCAGCAGCCCAAGCTTCAACAGCAGC 3419
 Query 3421 AACCTGGCTTGCTTTCTGAAATGCCTGGCCATACGGCACTTCACCCTACCACTACAACCA 3480
 Sbjct 3420 AACCTGGCTTGCTTTCTGAAATGCCTGGCCATATGGCACTTCACCCTACCACTACAACCA 3479
 Query 3481 ACCAACTTCTGCTGCTGGCAGTAGTATACTGACTGGAGCAGGTGGTGCAGGTGGTGCAG 3540
 Sbjct 3480 ACCAACTTCTGCTGCTGGCAGTAGTATACTGACTGGAGCAGG-----TGGTGCAG 3530
 Query 3541 GGCAGTCTGTAATTACTGATGAAGTTCTATCTTGCTCCACCTCACCTTCTGCAAATAACT 3600
 Sbjct 3531 GGCAGTCTGTAATTACTGATGAAGTTCTATCTTGCTCCACCTCACCTTCTGCAAATAACT 3590
 Query 3601 GTACCAATGCACTTCCACAATTGATAAATTCACGATTCCAGAGAAGCACACTGGTAGGCG 3660
 Sbjc 3591 GTACCAATGCACTTCCACAATTGATAAATTCACGATTCCAGAGAAGCACACTGGTAGGCG 3650
 Query 3661 ATGACATGGCCAGTCTGCTGCCACAATCTTGAAGTTCAAGTGCCTTAGAAACCACGTCAT 3720
 Sbjct 3651 ATGACATGGCCAGTCTGCTGCCACAATCTTGAATTCAGTGCCTTAGAAACCACGTCAT 3710

Query 3721 CAAATGCAAACATGTTGAAAGATTTACAGCCAAAGTCTGAAGTTAAGCCATCTCTGAATA 3780
 |||
 Sbjct 3711 CAAATGCAAACATGTTGAAAGATTTACAGCCAAAGTCTGAAGTTAAGCCATCTCTGAATA 3770

Query 3781 TTTCCAAAATTCAGAATCAAGGGCATTGCTCCTCAGACATACTTGAATGGTAATGCTG 3840
 |||
 Sbjct 3771 TTTCCAAAATTCAGAATCAAGGGCATTGCTCCTCAGACATACTTGAATGGTAATGCTG 3830

Query 3841 CCCATACAGATTGTTTGGACACATCATCTTCTACAACCTTCAGTTTGCCTTTCTCAGAGTG 3900
 |||
 Sbjct 3831 CCCATACAGATTGTTTGGACACATCATCTTCTACAACCTTCAGTTTGCCTTTCTCAGAGTG 3890

Query 3901 ACGCTCATATGCATCAGAATAACAATCCATTATCTTACAATCCACAGTCTTTGTTGTTTA 3960
 |||
 Sbjct 3891 ACGCTCATATGCATCAGAATAACAATCCATTATCTTACAATCCACAGTCTTTGTTGTTTA 3950

Query 3961 GAGACAATAATCAAGATGGGGAAGTTCAGGCAGATGCTAGGAGCAATATTCCTTATGCCA 4020
 |||
 Sbjct 3951 GAGACAATAATCAAGATGGGGAAGTTCAGGCAGATGCTAGGAGCAATATTCCTTATGCCA 4010

Query 4021 ATAACATTGATAGCCAAATGGGAATGCCACTGAATCCAGATTCACTCTCAACCAAAGGCA 4080
 |||
 Sbjct 4011 ATAACATTGATAGCCAAATGGGAATGCCACTGAATCCAGATTCACTCTCAACCAAAGGCA 4070

Query 4081 CATTGAGGTTGGGGAAGGATTTGTCTAATAATTTCTCTTCAGAAGGCATGCTTGTAATT 4140
 |||
 Sbjct 4071 CATTGAGGTTGGGGAAGGATTTGTCTAATAATTTCTCTTCAGAAGGCATGCTTGTAATT 4130

Query 4141 ATGAAATTAACAGAGATGCCAGCAGGAACCTTCGTCTTCAATGGTTTCACAGACATTTG 4200
 |||
 Sbjct 4131 ATGAAATTAACAGAGATGCCAGCAGGAACCTTCGTCTTCAATGGTTTCACAGACATTTG 4190

Query 4201 GTGTACCTGATATGGCCTTCAATTCAATTGATTCCACAATAGATGATAGTAACCTTCTTGA 4260
 |||
 Sbjct 4191 GTGTACCTGATATGGCCTTCAATTCAATTGATTCCACAATAGATGATAGTAACCTTCTTGA 4250

Query 4261 ATAGTGGTCCATGGGCTCCACCACCAGCACCACCCTGCCACCTCTGCCACCTCTGCCAC 4320
 |||
 Sbjct 4251 ATAGTGGTCCATGGGCTCCACCACCAGCACCACCCTGCCACCTCTGCCACCTCTGCCAC 4301

Query 4321 CAGCACAGTTTCAGCGGATGAGGACATATACCAAGGTTTATTGATGTTATTGTTATTCA 4380
 |||
 Sbjct 4302 CAGCACAGTTTCAGCGGATGAGGACATATACCAAGGTTTATTGATGTTATTGTTATTCA 4361

Query 4381 ATAATAAATCTGTCTTGATGGTAGATGTAAGCTGTTTCATGGGTGTGTGTGCTTTTCTTT 4440
 |||
 Sbjct 4362 ATAATAAATCTGTCTTGATGGTAGATGTAAGCTGTTTCATGGGTGTGTGTGCTTTTCTTT 4421

Query 4441 TATTCTAAATTAGCTTTTTTCTTTTGGCCCTTTAGGTATATAAACGTGGAGCTGTGGGAA 4500
 |||
 Sbjct 4422 TATTCTAAATTAGCTTTTTTCTTTTGGCCCTTTAGGTATATAAACGTGGAGCTGTGGGAA 4481

Query 4501 GATCCATAGACATAACACGGTATTTCAGGTTATGAGGAGCTTAAACAGGATCTAGCTCGTA 4560
 |||
 Sbjct 4482 GATCCATAGACATAACACGGTATTTCAGGTTATGAGGAGCTTAAACAGGATCTAGCTCGTA 4541

Query 4561 GGTTTGGCATAGAGGGACAGCTGGAGGATCGGCAAAGGATAGGTTGGAAACTTGTCTATG 4620
 |||
 Sbjct 4542 GGTTTGGCATAGAGGGACAGCTGGAGGATCGGCAAAGGATAGGTTGGAAACTTGTCTATG 4601

Query 4621 TGGATCATGAGAGTGATGTTCTACTATTGGGAGATGACCCTTGGGAGTAAGTCTGACCAT 4680
 |||
 Sbjct 4602 TGGATCATGAGAGTGATGTTCTACTATTGGGAGATGACCCTTGGGAGTAAGTCTGACCAT 4661


```

Query 5581 GAAGGTGGAGAAGCGAGCTGCTCCTAGTGCTACAAATCAATTAAGTTACTATAAACTGG 5640
          |||
Sbjct 5560 GAAGGTGGAGAAGCGAGCTGCTCCTAGTGCTACAAATCAATTAAGTTACTATAAACTGG 5619

Query 5641 TACTGAATCGAATGTTTGGAGCACCCGGACAAAAC TAGACGTGTTTCAAGCCATAGATTT 5700
          |||
Sbjct 5620 TACTGAATCGAATGTTTGGAGCACCCGGACAAAAC TAGACGTGTTTCAAGCCATAGATTT 5679

Query 5701 GTTGTGTGTGTAATGTTGAAAAGAATTGTTCTTGGCATGGTCAGAAATATTTAATTCTAT 5760
          |||
Sbjct 5680 GTTGTGTGTGTAATGTTGAAAAGAATTGTTCTTGGCATGGTCAGAAATATTTAATTCTAT 5739

Query 5761 GTATTGTAGAAAAGAAAGGTTTAATATTGAGCGGTGATTTTGAATGTATAAAGTCATAGGC 5820
          |||
Sbjct 5740 GTATTGTAGAAAAGAAAGGTTTAATATTGAGCGGTGATTTTGAATGTATAAAGTCATAGGC 5799

Query 5821 AGCTAGATTTTGC GGAAGCTTTCACGATCTTCCATCCTTGGAGGCACCTTAGTAATTTTT 5880
          |||
Sbjct 5800 AGCTAGATTTTGC GGAAGCTTTCACGATCTTCCATCCTTGGAGGCACCTTAGTAATTTTT 5859

Query 5881 AGCCTAGAGGACTGTGAAGTAGAGAAACTTCAAATGTTGTAAGGTCATGTAGACACTGTT 5940
          |||
Sbjct 5860 AGCCTAGAGGACTGTGAAGTAGAGAAACTTCAAATGTTGTAAGGTCATGTAGACACTGTT 5919

Query 5941 GCATTGTTTGCTTGAAGCATAAGTACTAAAAATGGTGGAGAATTAATTGACTGTACAGGA 6000
          |||
Sbjct 5920 GCATTGTTTGCTTGAAGCATAAGTACTAAAAATGGTGGAGAATTAATTGACTGTACAGGA 5979

Query 6001 GATCGATTACTAGTATTTGACACATGAATCTAATGTTTCATCATCGACTCTTATTCTAATT 6060
          |||
Sbjct 5980 GATCGATTACTAGTATTTGACACATGAATCTAATGTTTCATCATCGACTCTTATTCTAATT 6039

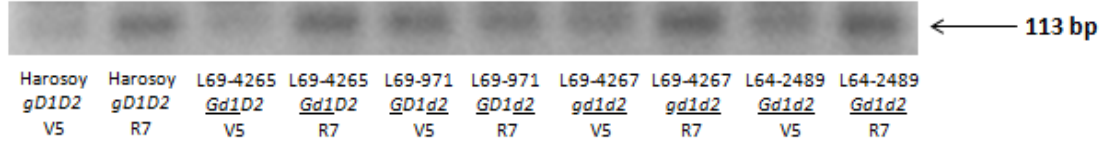
Query 6061 CAGTAATTGTTACGA 6075
          |||
Sbjct 6040 CAGTAATTGTTACGA 6054

```

Key: 5'UTR Exon 3'UTR SNP Deletion

Figure 1.7. Genomic sequencing results for a candidate *g* gene Glyma01g00510.1. Nucleotide sequences for wild type HAR (*gD1D2*) and L64-2489 (*Gd1d2*) mutant are identical (Query) and are aligned to the model soybean cultivar Williams 82 (Sbjct).

a)



b)

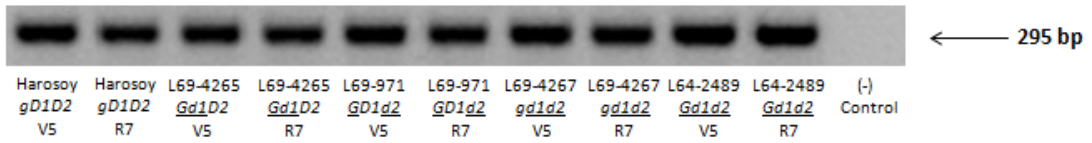


Figure 1.8. RT-PCR analysis of an alternate putative *g* gene Glyma01g00520.4 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and five isogenic lines [WT Harosoy, L69-4265, L69-971, L69-4265, L64-2489] with *g* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

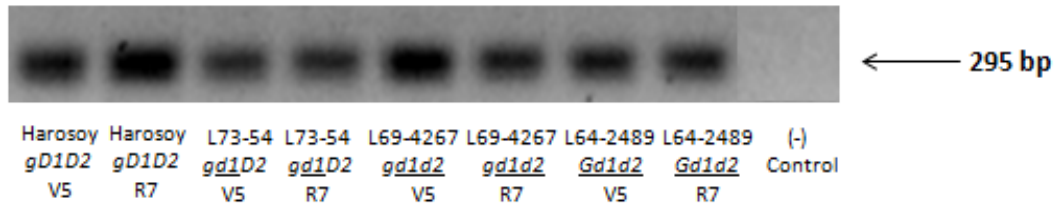


Figure 1.9. RT-PCR analysis of a putative *DI* gene Glyma01g41610.2 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and four isogenic lines [WT Harosoy, L73-54, L69-4267, L64-2489] with *DI* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)



Figure 1.10. RT-PCR analysis of a putative *D2* gene Glyma11g03770.2 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and four isogenic lines [WT Harosoy, L69-4266, L69-4267, L64-2489] with *D2* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

1	CCTGAAGCA	CTCTTCATT	TTAAATTACC	CCTGAATGAA	TGACCAGCAA	TGATGTGCTC
61	ATTGACCTTT	CAAGCCCTAC	AACTACAATG	AATGACTGTT	CCCGCCAAGT	CACCATGTTA
121	TAGTACTACA	CGTGCTGACT	CCCCATAAAA	TAATGCCATA	ACCAAACCTCT	TTATCAACAT
181	CATAGTTGAA	GTGTCTCTAT	CCCTCTCGTA	TATGCATGGC	ACCCCCGAGG	AATGTTGAAA
241	CTGCTAAGAA	AACAAACAAC	AGAGGAGCAT	GGACCGCGGA	GGAAGATCAG	AAGCTAGCCC
301	AGTGCATTGA	AATTCACGGT	GCAAAGAAAAT	GGAAAACTGT	TGCAATTTAAA	TCAGGTTATA
361	TATAATTAGC	TTGTAATGCT	AATTTCCAAT	ATACATATGC	ATTAATTCTT	CTTGACATGC
421	GTGTTTAAAT	TCTATATGGT	TTGATAGGTC	TAAACAGATG	TGGAAAAAGT	TGCAGGCTGA
481	GATGGTTAAA	CTATTTGAGA	CCAAATATCA	AGAGGGGCAA	CATATCAGTT	GAGGAAGAGG
541	ATTTGATTAT	TAGGCTTCAC	AAACTGCTGG	GAAACAGGTT	TTAATTTGTT	GCATTCAAAA
601	CTAAGATAAG	AATATTACTT	CCTCTTTCAT	ATTATAATTA	TCGTGTAAGA	AAGATAAATT
661	ATTTTTAAAA	TAAGTTTTAT	TTTAAATTTT	TAATGTAATA	TTAATTATAT	TTTTTTACTT
721	TATTTTATTT	TATTTATTTA	TTTATTTAT	ATATGGATAA	AATAATTTAA	GACGATTATT
781	TTTTTTTTTG	AACGAAGAAA	ATATTTTAT	ATAGTTTATG	AACATGTGTT	GTGAATATTA
841	CTATACATAT	AATGTCCATA	TTTTTTGTGTT	AATTAATTTAA	GGTGGTCATT	GATAGCTAAG
901	AGACTTCCAG	GACGAACAGA	CAATGAAAAT	AAGAACTACT	GGAAACTCTG	TTTGTGCAAA
961	AAAGTTAATC	ACAAAAAGGT	GAAACCTGAA	ACTTCAACGG	CACAAAACAAC	GCATACGACT
1021	CAGAATACCG	AGGAAAGAGC	AGCGGTAGAG	AATAAAGAGG	GCACTGATGA	TGGGAGTGGG
1081	GACTCAGAAG	TTAACTTTGA	TGTGAACGAA	TTCTTCAACT	TCTCCATTGA	ATCTGGACCC
1141	TTTGCCTTGG	ATTGGATGAT	CTAAAACCTG	ATGAGGAATT	GATAATTACG	TGTCTCTCTT
1201	AATTTATTAG	TGTGAGCTTG	TTTTTATATA	CTTGTTATGT	GGTCATTGTT	AAGTGAGTC
1261	CTAGCAATAG	CAATACATGT	GTGAGTTCCA	TTAAATAAAT	TGAACTCTCA	TCATTTGATT
1321	TAGTGGAACT	TACATATATA	TAATTAAGTG	TAACAATAAT	AATGTGTTCA	TGAGAGTACT
1381	GAAT					

Key: 5'UTR Exon 3'UTR

Figure 1.11. Genomic sequencing results for candidate *D2* gene Glyma11g03770.2. Nucleotide sequences for Williams 82, wild type Harosoy (*gDID2*) and a single Harosoy mutant L69-4266 (*gDId2*) all yielded identical results.

Table 1.1: RT-PCR primers for four candidate primary regulatory genes and control in soybean

Name	Gene Model	Transcript Size	Product Size	Forward Primer	Reverse Primer
<i>g</i>	Glyma01g00510.1	4123 bp	141 bp	GTGGCCTCCA AATCAAGCC	CGCCTTTCGGA ACTAGGACA
<i>g</i> (alternate)	Glyma01g00520.4	1944 bp	113 bp	CCAGAAGGGT GTGGTGGAAG	GTTGCTTGCCT CTTGCAGG
<i>D1</i>	Glyma01g41610.2	741 bp	110 bp	ATCGCACAAAG GTGAAGCCT	ACTCCCATCAT CAATGCGCT
<i>D2</i>	Glyma11g03770.2	987 bp	149 bp	AATGACTGTTC CCGCCAAGT	TTCAACATTCC TCGGGGGTG
Actin (<i>Sac3</i>) (soybean)*	/	1128 bp	295 bp	GCGAGAAATTG TCCGTGACA	TTGAGCCACCA CTAAGAACA

*control primer set

Table 1.2: Primers used for amplifying *g* target sequence

Name	Product Size	Forward Primer	Reverse Primer
Gforward	6239	TCATCACTAAATT AAACGGGCCT	/
Greverse		/	TGCACTTTGTTCT CACTATAGCAC

Table 1.3: Primers used for amplifying *D2* target sequence

Name	Product Size	Forward Primer	Reverse Primer
D2-1F	4384 bp	GAGAGAGGAG GAAGTCGTGG	/
D2-1R		/	TGCAACGCT TCTTCACAT GC

Table 1.4: Primers used for sequencing *g* template

Name	Forward Primer	Reverse Primer
GForward	TCATCACTAAAT TAAACGGGCCT	/
Greverse	/	TGCACTTGT TCTACTATA
GT1F	TCATCACTAAAT TAAACGGGCCT	/
GT1R	/	CTGCAGCAG GATTGAAATG ATTAA
GT2F	GGGGATGCTTGG CTGGAAAT	/
GT2R	/	TGAGGAGCA AAATGCCCTT
GT3F	GGGCATTTTGCT CCTCAGAC	/
GT3R	/	TGCACTTGT TCTACTATA GCA
gt4pp9f	TCATGAAGTAAA CTTGGGTGGTIG	/
gt5pp3f	TCTTGCTCCACC TCACCTTC	/
gt6pp3f	AGGATCTAGCTC GTAGGTTTGG	/
GT7-1	/	AGTATCCCA CCGTCAGAG

Table 1.5: Primers used for sequencing *D2* template

Name	Forward Primer	Reverse Primer
D2ALT2-F	TGAGACCAG TAGACCACAA ACA	/
D2ALT2-R	/	GTCATTGGA TACATGTTA GGAGCA
PP2-F	AGGCTTCACA AACTGCTGG	/

Table 1.6: RNA samples used for expression analysis

	V5*	Microarray # / Year	R7**	Microarray # / Year
	[] ng/uL		[] ng/uL	
GENOTYPES				
Harosoy	3053	Microarray 1 / 2009	1069	Microarray 6 pt. 2/ 2008
L73-54	2831	Microarray 1/ 2008	1030	Microarray 6 pt. 1/ 2008
L69-4266	3585	Microarray 1/ 2008	1713	Microarray 5/ 2008
L69-4265	6036	Microarray 1/ 2008	994	Microarray 6 pt. 1/ 2008
L67-971	2725	Microarray 1/ 2008	496	Microarray 6 pt. 2/ 2008
L69-4267	4407	Microarray 1/ 2008	558	Microarray 6 pt. 1/ 2008
L64-2489	3288	Microarray 1 / 2009	1073	Microarray 6 pt. 2/ 2008

*V5- The developmental stage at which five fully expanded, green trifoliolate leaves are on the plant (www.ag.ndsu.edu/pubs/plantsci.htm).

**R7- The developmental stage at which one major pod has changed to a brown color on the main stem (www.ag.ndsu.edu/pubs/plantsci.htm).

Microarray's provided by (Schreier and Chandlee, 2009).

Table 1.7: Soybean “Evergreen” mutants available as Harosoy isolines

Line	Phenotype			Genotype	Gene Combination
	Leaves	Seed Coats	Pods		
Harosoy	WT	WT	WT	<i>ggD1D1D2D2</i>	--
L73-54*	WT	WT	WT	<i>ggd1d1D2D2</i>	<i>D1</i>
L69-4266*	WT	WT	WT	<i>ggD1D1d2d2</i>	<i>D2</i>
L69-4265*	WT ¹	Stay Green	WT ¹	<i>GGd1d1D2D2</i>	<i>Gd1</i>
L69-971*	WT ¹	Stay Green	WT ¹	<i>GGD1D1d2d2</i>	<i>Gd2</i>
L69-4267*	Stay Green ²	WT	Stay Green ³	<i>ggd1d1d2d2</i>	<i>d1d2</i>
L64-2489*	Stay Green ²	Stay Green	Stay Green ³	<i>GGd1d1d2d2</i>	<i>Gd1d2</i>

* Harosoy Isolines

WT= Wild type (yellow leaves at maturity)

¹slightly delayed senescence relative to Harosoy

²abscission proceeds normally

³ pods turn green to brown with no yellow, transitional phase

American Soybean Association copyright ©2014 www.soystats.com

Balazadeh S., Riano-Pachon D.M. and Mueller-Roeber B. **Transcription factors regulating leaf senescence in *Arabidopsis thaliana*** *Plant Biology* (2008) **10**: 63–75.

Bolon Y.T., Haun W.J., Xu W.W., Grant D., Stacey M.G., Nelson R.T., Gerhardt D.J., Jeddloh J.A., Stacey G., Muehlbauer G.J., Orf J.H., Naeve S.L., Stupar R.M. and Vance C.P. **Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean** *Plant Physiology* (2011) **156** (1): 240-53.

Canfield M.R., Guiamét J.J., and Noodén L.D. **Alteration of Soybean Seedling Development in Darkness and Light by the Stay-green Mutation *cytG* and *Gd1d2*** *Annals of Botany* (1995) **75** (2): 143-150.

Cooper J.L., Till B.J., Laport R.G., Darlow M.C., Kleffner J.M., Jamai A., El-Mellouki T., Liu S., Ritchie R., Nielsen N., Bilyeu K.D., Meksem K., Comai L. and Henikoff S. **TILLING to detect induced mutations in soybean** *BMC Plant Biology* (2008) **24**: 8-9.

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

Guiamet J.J. and Giannibelli C.M. **Nuclear and cytoplasmic “stay-green” mutations of soybean alter the loss of leaf soluble proteins during senescence** *Physiologia Plantarum* (1996) **96** (4): 655–661.

Guiamét J.J., Schwartz E., Pichersky E., and Noodén L.D. **Characterization of Cytoplasmic and Nuclear Mutations Affecting Chlorophyll and Chlorophyll-Binding Proteins during Senescence in Soybean** *Plant Physiology* (1991) **96** (1): 227-231.

Latchman D.S. **Transcription factors: an overview** *International Journal of Biochemistry and Cell Biology* (1997) **12** (1): 305-12.

Liu X., Li Z., Jiang Z., Zhao Y, Peng J., Jin J., Guo H., Luo J. **LSD: a leaf senescence database** *Nucleic Acids Research* (2011) **39** (Database issue) D1103–D1107.

Lim P.O., Kim H.J., and Nam H.G. **Leaf Senescence** *Annual Review of Plant Biology* (2007) **58**: 115–36.

Mathieu M., Winters E.K., Kong F., Wan J., Wang S., Eckert H., Luth D., Paz M., Donovan C., Zhang Z., Somers D., Wang K., Nguyen H., Shoemaker R.C., Stacey G.

and Clemente T. **Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository** *Planta* (2009) **229** (2): 279-89.

Noodén L.D. and Leopold A.C., **Senescence and Aging in Plants** *Academic Press San Diego* (1988) 33-34, 506-507.

Okushima Y., Overvoorde P.J., Arima K., Alonso J.M., Chan A., Chang C., Ecker J.R., Hughes B., Lui A., Nguyen D., Onodera C., Quach H., Smith A., Yu G., and Theologis A. **Functional Genomic Analysis of the Auxin Response Factor Gene Family Members in *Arabidopsis thaliana*: Unique and Overlapping Functions of ARF7 and ARF19** *The Plant Cell* (2005) **17**: 444–463.

Schreier, L. and Chandlee J.M. **Characterization of MADS-Box Genes of Soybean (*Glycine max*, *L. Merr.*) and Initial Functional Analysis in Whole Plant Senescence.** *Plant Biology* (2009)

Smart M. C. **Tansley Review No. 64 Gene expression during leaf senescence.** *New Phytologist* (1994) **126**: 419-448.

Woo H.R., Kim J.E., Kim J., Kim J., Lee U., Song I-J., Kim J-H., Lee H-Y., Nam H.G. and Lim P.O. **The RAV1 transcription factor positively regulates leaf senescence in *Arabidopsis*** *Journal of Experimental Botany* (2010) **61** (14): 3947–3957.

Zemach A. and Grafi G. **Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation** *TRENDS in Plant Science* (2007) **12** (2): 80-85.

Chapter II

Evaluating the Role of a Senescence Response Element (SRE) in Senescence Associated Genes (SAGs) of Soybean

by

Robert Almeida¹ and Joel M. Chandlee²

is prepared in the format specified by *Plant Physiology*

**¹MS Candidate, Department of Cell and Molecular Biology, The University of
Rhode Island, Kingston, RI 02881. Email: ralme1@my.uri.edu**

**²Professor, Department of Cell and Molecular Biology, The University of Rhode
Island, Kingston, RI 02881. Email: jmchandlee@gmail.com**

Abstract

Soybean [*Glycine max* (L.) Merr.] is an important agricultural crop that serves as a major source of oil and protein for consumption worldwide. Soybean whole plant senescence is characterized by a systematic degradative process that involves up-regulation of new gene activity and down-regulation of other genes that ultimately lead to cell, tissue and organ death. Genes up-regulated during senescence are collectively referred to as senescence associated genes (SAGs). Some SAGs have been shown to contain a unique senescence response element (SRE) within their promoters that confers a senescence-specific pattern of expression. This is best exemplified by the *SAG12* (cysteine protease) gene of *Arabidopsis*. The 33 base pair SRE for the *SAG12* gene has been shown to harbor a well-conserved 7 base pair sequence that is also found in SAGs from other plant species. To determine if a related SRE could be responsible for regulating soybean SAGs, a genome-wide study was performed using publically available databases. Genes selected for analysis in soybean were previously identified as SAGs in other plant species. This search led to the identification of several soybean genes that harbor this SRE. These genes were bioinformatically analyzed using various structural criteria to identify the best potential soybean ortholog for each gene type. Structural criteria included measures of alignment similarity with the *Arabidopsis* SRE, proximity of the SRE to the transcription start site and other features of gene architecture, polypeptide sequence identity, and phylogenetic and syntenic relationships. Genes meeting the defined structural criteria underwent evaluation for functional roles in soybean senescence through RT-PCR analysis using a suite of isogenic lines exhibiting normal as well as

delayed senescence phenotypes. The selected mutants represent different combinations (single, double and triple mutants) of genes *G*, *d1* and *d2* that give rise to evergreen leaves and green seed phenotypes. None of the genes selected for functional analysis demonstrated evidence of differential expression among the selected isolines. However, many genes harboring an SRE have yet to be investigated.

Introduction

Leaf senescence is a highly regulated developmental process that ends with the programmed death of leaf cells (Swidzinski *et al.*, 2002). During leaf senescence, cellular components such as proteins, lipids, and nucleic acids are degraded, and the released nutrients are mobilized from the leaves for re-use in other parts of the plant (Lim *et al.*, 2003; Noodén and Guiamet, 1991; Quirino *et al.*, 2000). Understanding the molecular mechanics of senescence has direct application to the improvement of crop plants by reduction of spoilage and increased yields. Efforts to inhibit or delay the effects of senescence have been made using fusion proteins produced from gene constructs that incorporate the *Arabidopsis thaliana* *SAG12* gene (encoding a cysteine protease) senescence response element (SRE) (Gan and Amasino, 1995; Noh and Amasino, 1999). The *SAG12* SRE is 33 base pairs in length and is found within an essential promoter element located -472 to -784 upstream of the transcription start site (Noh and Amasino, 1999). The SRE is highly conserved in the orthologous *SAG12* gene of rapeseed (Noh and Amasino, 1999) and the asparagine synthetase (*AS*) gene of asparagus (Winichayakul *et al.*, 2004) two distantly related taxa. The SRE of the *AS* gene was aligned to the *SAG12* SRE of both *Arabidopsis* and rapeseed and a highly conserved seven base pair region was identified within the well conserved 33 base pair region (Figure 2.1).

The goal of this study was to examine a collection of soybean genes encoding an *AS*, transcription factors (*MYB*, *MADS* and *MYB60*), proteases (*SAG12* and *SAG2*), a lipase (*SAG101*), and a transmembrane receptor kinase (*SARK*) for the presence of a conserved SRE promoter element. These genes were chosen because of experimental

evidence implicating them in the senescence program of various plant species. These genes were used to select soybean orthologs for a comprehensive bioinformatic analysis to identify an associated, potentially functional, SRE element. Candidate genes meeting defined structural criteria were analyzed functionally to elucidate a role in the soybean senescence pathway.

Three different genes (*g*, *D1*, *D2*) are thought to be involved with regulating the whole plant/leaf senescence program of soybean because non-functional mutations in any one of them contributes to an evergreen (non-senescent) phenotype. The fact that mutations in these genes are known to alter the progression of the normal senescence program suggests that they function as major regulators of at least portions of the overall senescence pathway (Zeng, 1999). Homozygous *d1d1d2d2* lines remain green and show an inhibition of chlorophyll degradation and chlorophyll-binding proteins yet they still undergo a decline in photosynthetic activity and leaf abscission (Guiamet et al. 1991, Canfield et al. 1995). When combined with the dominant mutation, *G* (*GGd1d1d2d2*), a decline in photosynthetic activity fails to take place but the leaves still abscise. It has been suggested that *d1d2* may control a central regulatory process in the senescence program (Guiamet and Giannibelli 1996). Near-isogenic lines are available as single, double and triple mutant combinations for these three genes (Guiamet and Giannibelli 1996). This material can serve as the basis for expression analysis of candidate genes during senescence. As patterns of altered expression are characterized, a better understanding of the genetic regulation of the senescence pathways will unfold. Seven near-isogenic lines of soybean (Harosoy, L64-2489, L69-4266, L69-4265, L69-971, L69-4267 and, L73-54; Table 2.1) were

analyzed for the expression patterns of candidate genes throughout development using RT-PCR. The senescence program of soybean is poorly understood and further research into this area will contribute to a greater understanding of the developmental program for this economically important agricultural crop.

Results and Discussion

A total of 174 genes obtained through bioinformatic and structural analysis (Materials and Methods) were screened for 9 gene types representing 6 gene families of which 38 demonstrated evidence of an SRE (Table 2.2). Of these 38, 8 genes were selected for functional analysis using RT-PCR (Table 2.3) because of how well they met defined criteria in the Materials and Methods section compared to other genes that showed evidence of an SRE within the same gene family. Comparisons were made using early green leaf tissue (V5) and late yellow/evergreen leaf tissue (R7-12) of both wild type Harosoy (*gDID2*) and a Harosoy triple mutant (*Gdid2*) to delineate developmental differences in expression between these isolines. A total of 9 genes representing the SAG family [SAG12 (7) and SAG2 (2)] showed evidence of an SRE but fared poorly in the other criteria of structural analysis (gene architecture, polypeptide alignment, phylogenetic and syntenic analysis) and therefore were not examined functionally. The genes SAG101 (Glyma13g04540.1, Figure 2.2), TMYB (Glyma06g4555.1, Figure 2.3), 11MADS28 (Glyma11g36890.1, Figure 2.4), 18MADS28 (Glyma18g00801.1, Figure 2.5), SARK (Glyma12g36090.1, Figure 2.6) AS (Glyma18g06840.4, Figure 2.7) and MYB60 (Glyma19g29750.2, Figure 2.8) showed evidence for an SRE and were supported by other structural component criteria (see Materials and Methods). Functional analysis, however, did not yield evidence of a differential expression profile for any of the genes examined.

TMYB (Glyma06g4555.1) contains 68 base pairs identical to a 70 base pair sequence previously identified by microarray analysis as down-regulated 5-fold in expression. This 70 base pair sequence was selected for further analysis because it

shares a high sequence identity with a gene annotated as a MYB transcription factor previously implicated in the senescence program. Another gene, Glyma10g11200.1, annotated as a peroxidase harboring the defined SRE but matching the 70 base pair sequence exactly was also examined. Functional analysis, however, did not yield evidence of a differential expression profile (Figure 2.9).

MADS28, another gene previously identified and determined to be involved in the senescence program, yielded conflicting results. While this study determined that 11*MADS28*, [which harbors an SRE (Glyma11g36890.1) and not a non-SRE containing duplicate version 18*MADS28* (Glyma18g00801.1)], is the copy with a likely role in the senescence program, a discrepancy exists between Schreier and Chandlee, 2009 and the current study. In the previous study a differential expression pattern between wild type Harosoy (*gD1D2*) and the triple mutant (*Gdl1d2*) showed a perceptible upregulation of the transcript in the wild type isolate at the R7 developmental phase. However, in the current study, upregulation was evident in the triple mutant isolate of the R7 developmental phase (Figure 2.4). A possible explanation for this observation may be the primers used for analysis. The previous study used primers that spanned the fifth and sixth exon junction and the current study used primers that spanned the sixth exon and 3' UTR junction.

While none of the genes examined functionally in this study exhibited differential expression, the method used to derive the candidates has been well developed to serve as an efficient tool for gene mining. Rather than relying on costly and time consuming methods such as microarray analysis, this approach allows for a rapid evaluation of large numbers of orthologs from various species. It incorporates

data from various platforms and allows for the integration of information even with continually updating databases. There remain many genes from various gene families in soybean that have been identified as potentially involved in the senescence program using this method but await further functional analysis.

Materials and Methods

Selection of Candidate Genes for Functional Analysis

The strategy to identify candidate soybean orthologs harboring an SRE involved following guidelines for analysis of 8 specific structural criteria. Genes selected for functional analysis were required to: 1) exhibit an E-value less than 1.0×10^{-10} in a TBLASTN search; 2) exhibit sequence identity greater than or equal to 48.5% to the SRE 33 base pair consensus; 3) exhibit sequence identity greater than or equal to 71.4% to the SRE 7 base pair consensus; and 4) exhibit a proximity of the SRE sequence upstream to the transcription start site of the gene similar to the model *SAG12* gene. Gene architecture, polypeptide sequence identity, phylogenetic and syntenic relationships were also evaluated, but were given less emphasis than the previous four criteria for determining orthology (Figure 2.10). A total of 9 gene types representing 6 gene families were screened for soybean orthologs or homologs (Table 2.2).

Identification of Candidate Soybean Ortholog Genes Using *Arabidopsis thaliana* Gene Models

A comprehensive literature and database search of Pubmed, Google Scholar, the Leaf Senescence Database (LSD) and the *Arabidopsis* Information Resource (TAIR) for genes involved in leaf senescence was undertaken to identify key genes of *A. thaliana* that are implicated experimentally in the senescence program. From this list, several were selected and their *A. thaliana* peptide sequences were determined [*SARK* (AT4G30520.1), *SAG12* (AT5G45890.1), *SAG2* (AT5G60360.1), *SAG101*

(AT5G14930.2), and *MYB60* (AT1G08810.1)] using TAIR. These sequences were used to perform a global TBLASTN (protein to translated nucleotide query) alignment with the soybean genome within www.phytozome.net to identify putative orthologous soybean genes. Alignments were returned as high scoring pairs (HSP's) and the top 10 soybean results for each gene type containing a region of 3000 base pairs upstream of the gene representing the promoter of each with E-values less than 1.0×10^{-10} were selected for further structural analysis. To generate additional quality candidates and to ensure the best possible soybean orthologs were represented, the Smith and Waterman algorithm (local alignment) was also employed. Within the [phytozome.net](http://www.phytozome.net) gene page for each *A. thaliana* gene listed, the "Peptide Homologs" tab was selected. On the resulting page, a filter was selected that allowed for the local alignment of only orthologous soybean peptide sequences. The top 10 soybean results for each gene type with a region of 3000 base pairs upstream of the gene, representing the promoter of each, was selected with no score precluding them from further structural analysis (Figure 2.11). This procedure yielded a total of 20 soybean ortholog candidates for each gene derived from *A. thaliana* using the two alignment algorithms.

Identification of Candidate Soybean Ortholog Genes Using Alternate Plant Gene Models

A literature and database search of Pubmed, Google Scholar, the Leaf Senescence Database (LSD) and previous research also yielded genes implicated in the senescence program derived from plant species other than *A. thaliana*. These peptide sequences included asparagus AS (accession number X99552.1), soybean

SARK (Glyma13g34100.1), soybean *TMYB* (Glyma06g45554.1), and soybean *MADS28* (Glyma11g36890.1 and Glyma18g00800.1). This set of genes required individually unique methods to obtain ortholog/homolog soybean candidate genes for alignment analysis. The method used for each selected gene is outlined below.

Asparagus Asparagine Synthetase (*AS*)

The promoter region of the *AS* gene implicated in senescence was derived from the partially sequenced asparagus genome database. An ortholog to asparagus *AS* was first identified in *A. thaliana* because of the well annotated nature of the genome. This was achieved using the peptide sequence of the protein encoded by the asparagus *AS* gene and performing a TBLASTN against the *A. thaliana* genome to discover the ortholog with the best HSP overall (AT3G47340.1). After the ortholog was confirmed to have an asparagine synthetase annotation in TAIR, the AT3G47340.1 amino acid sequence was used to screen the soybean genome for orthologs by TBLASTN alignment in the Phytozome database. The top 10 HSP results with a region of 3000 base pairs upstream of the gene representing the promoter of each with E-values less than 1.0 e-10 were selected for further structural analysis. To generate additional quality candidates and to ensure the best possible soybean orthologs were represented, the Smith and Waterman algorithm (local alignment) was also employed. Within the phytozome.net gene page for the *A. thaliana AS* gene AT3G47340.1, the “Peptide Homologs” tab was selected. On the resulting page, a filter was selected that allowed for the local alignment of only orthologous soybean peptide sequences. The top 10 soybean results with a region of 3000 base pairs upstream of the gene representing the

promoter of each was selected with no score precluding them from further structural analysis. This procedure yielded a total of 20 soybean ortholog candidates for the *A. thaliana AS* gene using the two alignment algorithms.

Soybean *SARK*

Soybean *SARK* (Glyma13g34100.1) has been implicated in the senescence program by Li et al., (2006) but does not contain evidence of an SRE. However, this gene was used to search for other potential soybean homologs that could contain an SRE. The amino acid sequence of Glyma13g34100.1 was used to screen the soybean genome for homologs by TBLASTN alignment in the Phytozome database. The top 10 HSP soybean results with a region of 3000 base pairs upstream of the gene representing the promoter of each with E-values less than 1.0 e-10 were selected for further structural analysis. To generate additional quality candidates and to ensure the best possible soybean homologs were represented, the Smith and Waterman algorithm (local alignment) was also employed. Within the phytozome.net gene page for the soybean *SARK* gene Glyma13g34100.1, the “Peptide Homologs” tab was selected. On the resulting page, a filter was selected that allowed for the local alignment of only homologous soybean peptide sequences. The top 10 soybean results with a region of 3000 base pairs upstream of the gene representing the promoter of each was selected with no score precluding them from further structural analysis. This procedure yielded a total of 20 soybean homolog candidates for the soybean *SARK* gene using the two alignment algorithms.

Soybean *TMYB*

In previous work by (Schreier and Chandlee, 2009) and colleagues using microarray analysis, *TMYB* was identified as being differentially expressed in senescing leaf tissue in soybean. A 70 base pair sequence representing this MYB was first used to perform a BLASTN (nucleotide query to nucleotide) against the soybean genome using the Phytozome database with E-value results below 1.0×10^{-10} being the cutoff for acceptable homolog candidate genes. The top HSP results produced through this global alignment algorithm were obtained for a total of 10 homolog soybean candidates with a region of 3000 base pairs upstream of the gene representing the promoter of each selected for further structural analysis. In an attempt to obtain a higher number of quality candidates, the 70 base pair sequence was used to screen the soybean genome at the National Center for Biotechnology Information (NCBI). The search yielded only one result, a predicted myb-like protein, MYB4-Like. The translated 229 amino acid peptide sequence of MYB4-Like was used to screen the soybean genome with TBLASTN (protein translated nucleotide query) in the Phytozome database. The top HSP results were obtained for a total of 10 homolog soybean candidates with a region of 3000 base pairs upstream of the gene representing the promoter of each selected for further structural analysis. To utilize the Smith and Waterman algorithm for identifying additional candidates, further steps were needed. In the microarray analysis *TMYB* was identified based on a soybean nucleotide sequence of limited size (70 nucleotides). The Smith and Waterman algorithm could not be used with such a small sequence because no single soybean gene could be clearly identified as representing this sequence. To overcome this issue, the top overall

MYB sequence obtained from the TBLASTN search (Glyma06g45554.1) was used to gather candidates with the Smith and Waterman method. Within the phytozome.net gene page for Glyma06g45554.1, the “Peptide Homologs” tab was selected. On the resulting page, a filter was selected that allowed for the local alignment of only homologous soybean peptide sequences. The top 10 results with a region of 3000 base pairs upstream of the gene representing the promoter of each was selected with no score precluding them from further analysis. The three alignment strategies used to identify *TMYB* homologs yielded a total of 25 candidate soybean genes (the lower total number was due to overlap of the same sequences identified by the different algorithms).

Soybean *MADS28*

Previous research indicated that a *MADS*-box gene (*MADS28*) was differentially regulated in wild type and triple mutant isogenic varieties of soybean (Schreier and Chandlee, 2009). However, because of the duplicated nature of the soybean genome, two genes were possible candidates representing *MADS28* (Glyma11g36890.1 and Glyma18g00800.1). These sequences were verified to ensure current annotations for both sequences were present in phytozome.net. While the Glyma11g36890.1 gene remained the same, Glyma1800800.1 had been updated to Glyma1800801.1. The nucleotide sequences of the genes and 3000 base pairs upstream of the transcription start site were obtained for structural analysis.

Ortholog Alignment with the consensus 33 base pair SRE and the consensus 7 Base Pair Conserved Promoter Element

The ortholog and homolog soybean candidate genes produced from both global and local alignments were then subjected to nucleotide alignments using CLUSTALW and Geneious alignment programs to obtain the best possible nucleotide sequence comparison in a 3000 base pair promoter region upstream of each gene (Figure 2.12). The default settings used for alignments with CLUSTALW were as follows: slow alignment type, IUD DNA Weight Matrix, gap open default penalty of 10, gap extension default penalty of 0.20, gap distance default penalty of 5, no end gaps, no iterations and number of 1. The default settings used for Geneious alignments were as follows: cost matrix: 65% similarity (IUB) (5.0/4.5), gap open penalty, 12; gap extension penalty, 3; alignment type, global alignment with free end gaps; automatically determine sequence's direction (build guide tree via alignment not used) and refinement iterations, 2.

Alignments were performed using the entire gene sequence and 3000 base pair upstream region to confirm the 33 base pair SRE and that the 7 base pair conserved sequence within the SRE was indeed in the region upstream of the transcription start site. Alignments were omitted from further structural analysis if they did not have an SRE alignment of greater than or equal to 48.5% (16 of 33), a 7 base pair alignment of greater than or equal to 71.4% (5 of 7), and were not located in the upstream region. Gaps in the 33 base pair alignments found in the promoter were deleted and inspected to determine if they met the defined alignment criteria and discarded if they did not. This analysis was performed for all genes in this study regardless of identification

method. Soybean *SARK* is the only gene whose homologs did not produce candidates meeting these criteria.

Gene Architecture

Comparative gene architecture was examined to determine if conservation of exon/intron structures existed between the gene used for the search and the putative soybean ortholog. The publically available database Gramene (<http://gramene.org>) was used to search for the gene of interest and the ortholog transcript structures so that a comparison could be made between the two. While no quantitative criteria were established to define of an evolutionary relationship, this analysis contributed to ultimately selecting genes for functional testing. Similarities in length of genes, number of exons/ introns, and nucleotide alignments of over 50% in exons (performed using Geneious alignments with sequences obtained through global and local alignments) were used with other evidence aid in supporting a claim of orthology with the genes examined in this study (Figure 2.13). The sequence for the *TMYB* was generated from within the soybean genome but because individual one soybean gene could definitively be identified as representing this sequence, a comparison of gene structures could not be achieved.

Polypeptide Sequence Alignment

To further corroborate orthologous relationships, polypeptide sequence identity was examined. As with gene architecture, no quantitative criteria were established to define a relationship but rather the alignments were used to substantiate the previous

evidence for an orthologous relationship. Genes were examined to ensure that domains aligned properly with known sequences and a percentage of identical sites was used to assist in selecting genes for functional testing (Figure 2.14). Alignments were performed using the alignment tool within the Geneious program with peptide sequences obtained using BLASTN and TBLASTN. Default amino acid alignment parameters were as follows: Cost Matrix: Blosum62, gap open penalty of 12, gap extension penalty of 3; and alignment type: global alignment with free end gaps. This procedure was followed for all genes examined. Peptide alignments were performed only for homologs of Glyma06g45554.1 found using the Smith and Waterman method for *TMYB*.

Phylogenetic Analysis

A phylogenetic approach was undertaken to provide additional insight into any relationships between genes of different species. Soybean genes harboring an SRE within nodes in the same clade as the ortholog of each gene type were considered stronger candidates. To construct trees, the peptide sequence of the model for ortholog/homolog discovery of each gene type was used to screen the Phytozome database using TBLASTN against the *Arabidopsis*, rapeseed, rice, common bean, sorghum, and corn genomes. The best overall HSP result returned for each species was selected and the amino acid sequences imported to the Geneious program for phylogenetic evaluation. The peptide sequences for these genes and all soybean genes identified during global and local alignments were then aligned followed by evaluation of the phylogenetic tree (Figure 2.15). The default settings of Geneious Tree Builder

were as follows: cost matrix, Blosum62, gap open penalty of 12, gap extension penalty of 3; alignment type, global alignment with free end gaps; genetic distance model, Jukes-Cantor; tree build method, Neighbor-joining; and no outgroup. Tree nodes were then examined to see how orthologous soybean genes selected for harboring an SRE clustered with genes presumed to be orthologs or homologs.

Syntenic Analysis

The final criteria used to aid in the selection of orthologs in soybean for functional analysis was synteny. The Synfind portal of CoGe was used to screen *A. thaliana* orthologs and soybean homologs against the soybean genome. *A. thaliana* orthologs and soybean homologs were entered in the name field of the Specify Feature section and the appropriate coding sequence (CDS) selected. In the organism name field under the Specify Organisms section, the genome to be queried, *Glycine max*, was entered and the appropriate annotated version selected. Default settings were used for the Configure Parameters tab before running the analysis (*A. thaliana Col.* vs. soybean). Genes were determined to share a “chromosomal neighborhood” if 4 genes existed with similar structures and available annotations were present. More genes in the same syntenic region strengthened the likelihood of an orthologous relationship for this analysis. No results from syntenic analysis decreased the likelihood of an orthologous relationship, but no gene was eliminated based solely on this finding. Running the analysis in reverse (soybean vs. *A. thaliana*) did not necessarily guarantee the same results. This can be attributed to the difference in levels of amino acid identity between the two genes in question. As with the gene architecture, polypeptide

sequence identity and phylogenetic analysis, and results (or lack thereof) of syntenic analysis (Figure 2.16) were combined with the other analyses to determine the best candidate(s) for functional analysis.

Primer Design

Gene specific primers for the selected candidate soybean orthologs were designed using Primer 3 software (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) available through the NCBI portal. Primers were designed to span either the 5' or 3' UTR and exon junction for unique products ranging from 100 to 300 base pairs. Default settings were used except for the Exclusion tab which was then checked and Organism field which was changed to "*Glycine max*". If no 5' or 3' UTR was available, exon regions outside of conserved domains were used to generate unique products. The 5 sets of results returned were then used to query the Phytozome soybean database using TBLASTN to ensure the sequence was within the gene of interest and to verify the primers were unique in the genome (Table 2.4)

RNA Extraction

Total leaf RNA was isolated from several developmental stages throughout the normal life cycle using Harosoy and isogenic lines harboring genes that affect leaf senescence (Table 2.1). Leaves were harvested immediately into liquid nitrogen and RNA was extracted the same day, using a standard phenol/chloroform extraction protocol and LiCl precipitation (Maniatis et al., 1986; Schreier and Chandlee, 2009). RNA was quantified using a spectrophotometer and assayed for quality and quantity

by electrophoresis in formaldehyde-containing agarose gels (Figure 2.17). The soybean developmental stages V5 and R7 were used in this study. V5 (vegetative 5) is characterized by five fully expanded, green trifoliolate leaves are found on the plant and in R7 (reproductive 7) one major pod has changed to a brown color on the main stem (Table 2.5) (<http://www.ag.ndsu.edu/pubs/plantsci.htm>).

RT-PCR analysis

Preliminary RT-PCR analysis was performed with total RNA from the V5 and R7 developmental stages of soybean leaves from the seven different isogenic backgrounds, with the described experimental primer sets (Tables 2.6 and 2.4). RT-PCR analysis was performed using the AccessQuick RT-PCR System (Promega, Madison WI) with 1 μ g of total RNA in each reaction. Reaction controls were performed with actin-specific (*Sac3*) primers (Table 2.4) on each of the RNA samples tested. The Reverse Transcription cycle was run at 45° C for 45 min for one cycle (Eppendorf thermocycler model 5331). The PCR was carried out as follows: an initial denaturation step at 94° C for 4 min; 40 cycles with 1 min at 94° C, 1 min at 48° C, and 2 min at 72° C; final extension step of 72° C for 7 min; and finally, a hold at 4° C to complete the program. This program was used for all the primer sets analyzed. The products were screened using a 2% molecular biology grade agarose (Fischer Scientific) gel in 100 mM Tris-acetate and 2 mM EDTA.

Sequencing

Verification of the identity of the amplicons was performed through sequence analysis. Bands were excised from the gel and purified with the Wizard SV Gel and

PCR clean-up system according to the manufacturer's instructions (Promega). Bands were sequenced using the facilities at the University of Rhode Island's Genomics and Sequencing Center (URIGSC).

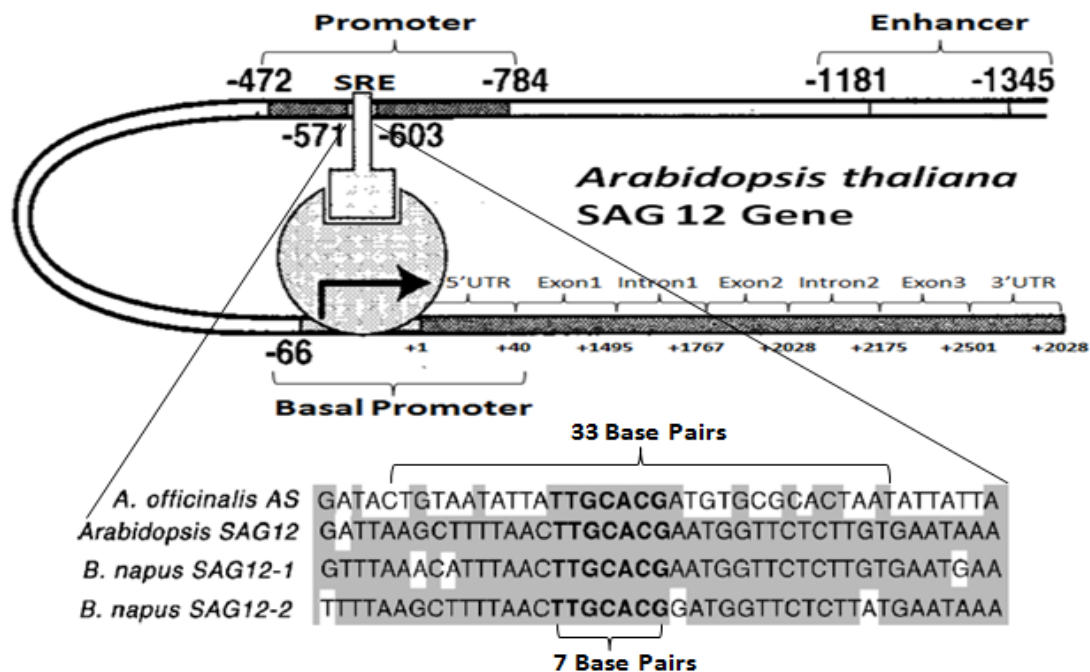
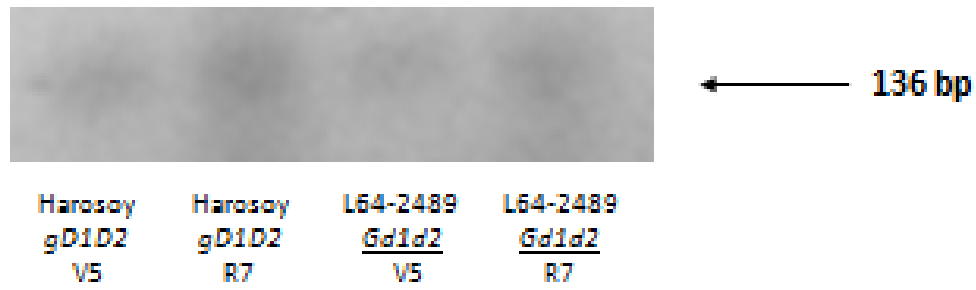


Figure 2.1. Conserved promoter sequences used as selection criteria for soybean SAG's.

A highly senescence-specific promoter has been identified in the *Arabidopsis thaliana* SAG12 gene. A subsequent study identified a conserved 7 base pair sequence within the 33 base pair Senescence Response Element (SRE). These two sequences serve as the basis for selection of candidate genes.

a)



b)

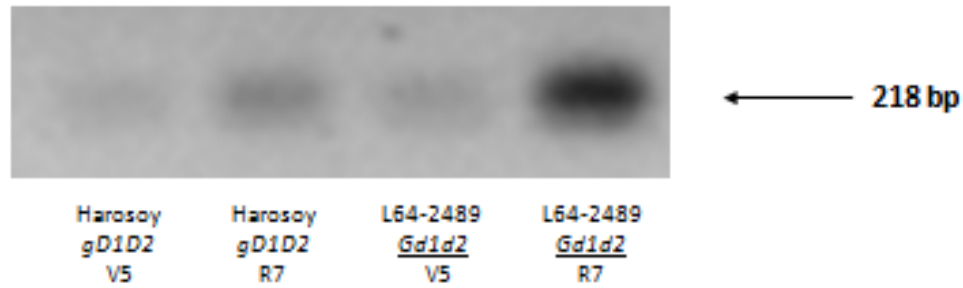


Figure 2.2. RT-PCR analysis of a putative *SAG101* gene Glyma13g04540.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with *SAG101* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

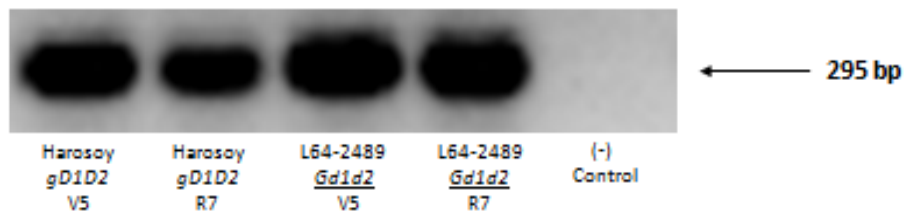
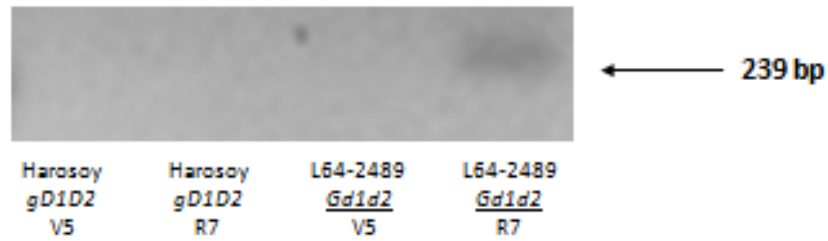


Figure 2.3. RT-PCR analysis of a putative *TMYB* gene Glyma06g4555.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with *TMYB* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1ug of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

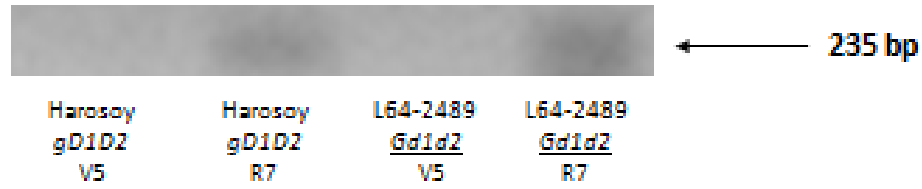


Figure 2.4. RT-PCR analysis of an alternate putative *IIMADS28* gene Glyma11g36890.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with alternate *IIMADS28* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

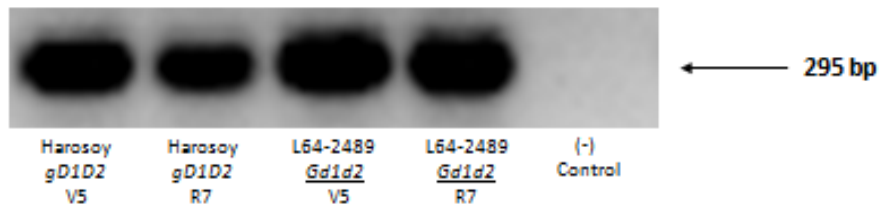
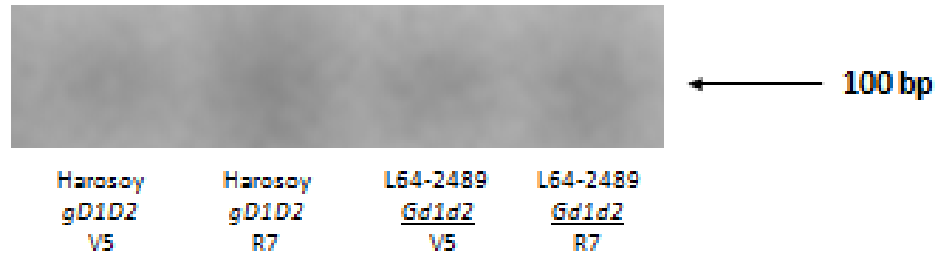


Figure 2.5. RT-PCR analysis of an alternate putative *18MADS28* gene Glyma18g00801.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with alternate *18MADS28* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

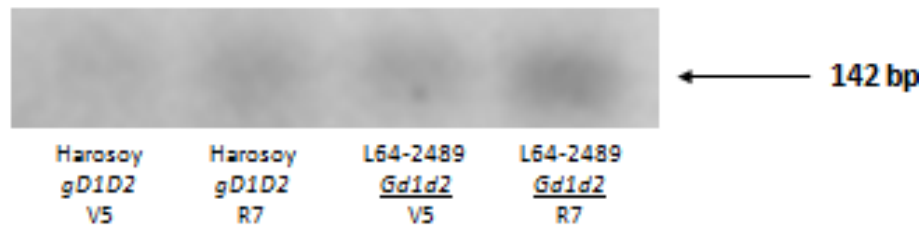


Figure 2.6. RT-PCR analysis of an alternate putative *SARK* gene Glyma12g36090.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with alternate *SARK* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1ug of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

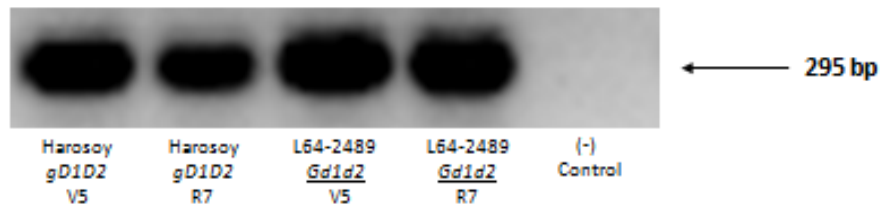
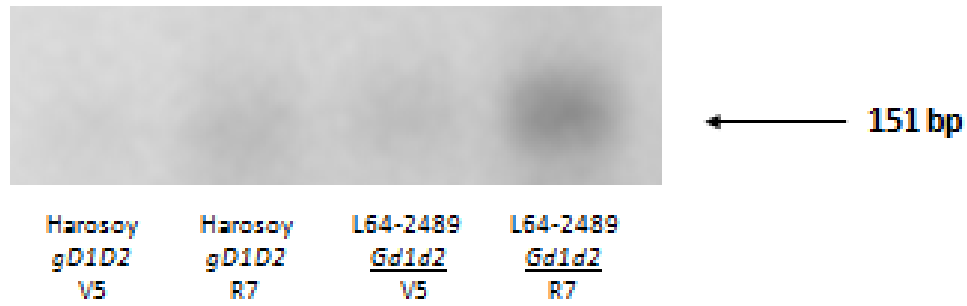


Figure 2.7. RT-PCR analysis of a putative *AS* gene Glyma18g06840.4 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with *AS* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1ug of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

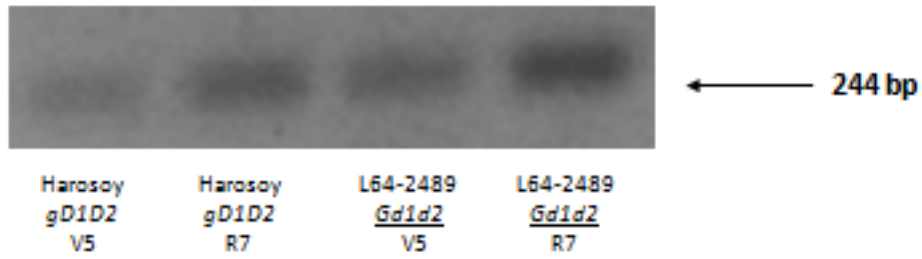


Figure 2.8. RT-PCR analysis of a putative *MYB60* gene Glyma19g29750.2 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with *MYB60* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

a)



b)

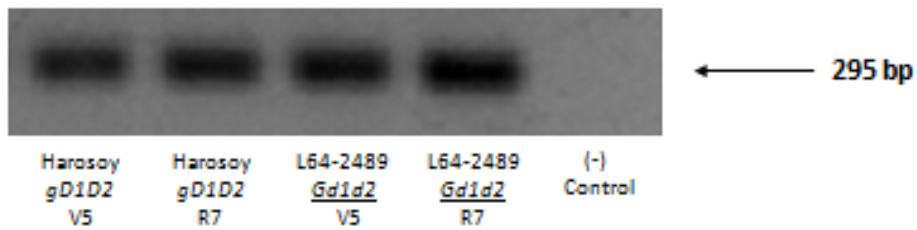


Figure 2.9. RT-PCR analysis of an alternate putative *TMYB* gene Glyma10g11200.1 comparing expression patterns between senescing and non-senescing soybean leaf tissue.

a) RT-PCR analysis of two developmental stages V5 and R7 and two isogenic lines [WT Harosoy and L64-2489] with alternate *TMYB* specific primers. Underlined represents the mutant allele.

b) RT-PCR reaction controls: 1 μ g of RNA with actin-specific (*Sac3*) primers for each RNA sample tested. (-) Reaction components without RNA.

Criteria	
A.	Blast search results e-value less than 1.0 e-10
B.	≥ 48.5% sequence identity at SRE 33 base pair level
C.	≥ 71.4% sequence identity at consensus 7 base pair level
D.	Proximity of the SRE upstream to genes transcription start site
E.	Gene architecture
F.	Polypeptide sequence identity
G.	Phylogenetic relationships
H.	Syntenic relationships

Figure 2.10. Structural criteria guidelines for gene selection

Genes selected for functional analysis were required to exhibit A-D. E-H were also evaluated but with less emphasis than the previous four criteria.

Smith & Waterman Method (Local Alignment)		TBLASTN (Global Alignment)	
Genes	E-Value	Genes	E-Value
1. Glyma19g29750.2	7.4 e-55 ★	1. Glyma12g01960.1	2.4 e-40
2. Glyma03g00890.1	3.5 e-52	2. Glyma11g01150.1	1.5 e-37
3. Glyma05g01080.1	1.5 e-50	3. Glyma13g05370.1	4.3 e-34
4. Glyma04g33720.1	3.5 e-50	4. Glyma01g06192.1	2.7 e-34
5. Glyma17g10820.1	7.0 e-50	5. Glyma16g13440.2	2.6 e-30
6. Glyma06g20800.2	8.3 e-50	6. Glyma18g49360.1	3.7 e-29
7. Glyma18g49360.1	2.5 e-48	7. Glyma04g11040.1	8.9 e-28
8. Glyma13g05370.1	1.3 e-46	8. Glyma06g00630.1	4.1 e-27
9. Glyma09g37340.1	1.3 e-46	9. Glyma14g24500.2	4.7 e-27
10. Glyma19g02600.1	5.3 e-40	10. Glyma15g15400.2	2.6 e-27

Figure 2.11. A. *Arabidopsis thaliana MYB60* vs. Soybean TBLASTN results

The *Arabidopsis thaliana MYB60* gene peptide sequence (At1g08810.1) was used to search the soybean genome for orthologs. Both global and local alignment algorithms were used to generate results that could be screened for the presence of an SRE and 7 base pair element. (A.) All values are below e-10 meeting the criteria. Arrows indicate overlap using the different algorithms. Star highlights gene used for further analysis.

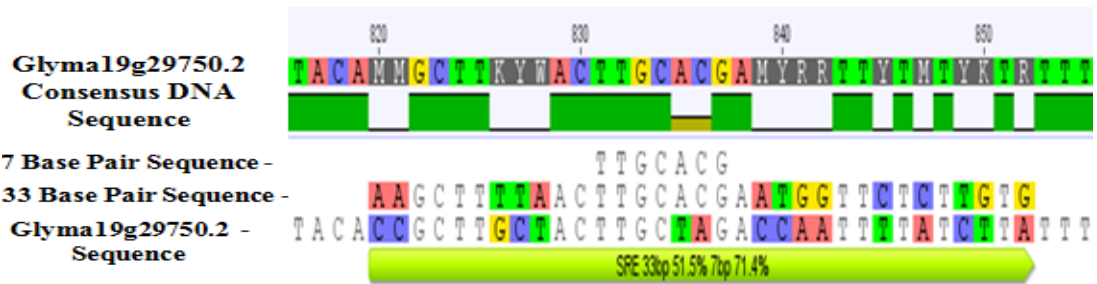


Figure 2.12. B. SRE Alignment C. Conserved 7 Base Pair Sequence Alignment D. Proximity of SRE to Gene Transcription Start Site

ClustalW and Geneious alignments were performed to visualize sequence identities of Glyma19G29750.2 found through the Smith and Waterman search and the SRE and conserved 7 bp sequences. (B.) In this instance there is a greater than 48.5% match (51.5%) at the SRE level and (C.) 71.4% match with the conserved 7 base pair element. (D.) This promoter element was located upstream of the Glyma19G29750.2 gene.

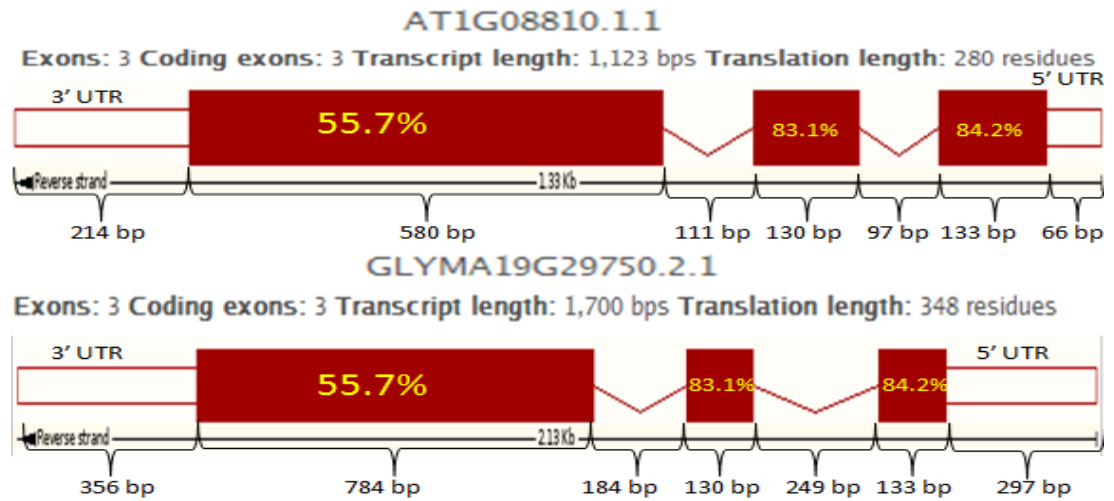


Figure 2.13. E. Gene architecture

Structures of both the *Arabidopsis thaliana* and soybean ortholog *MYB60* genes.



Figure 2.14. F. Polypeptide sequence alignment

MYB60 At1G08810.1 shares 184 identical sites (52.1%) with Glyma19G29750.2.

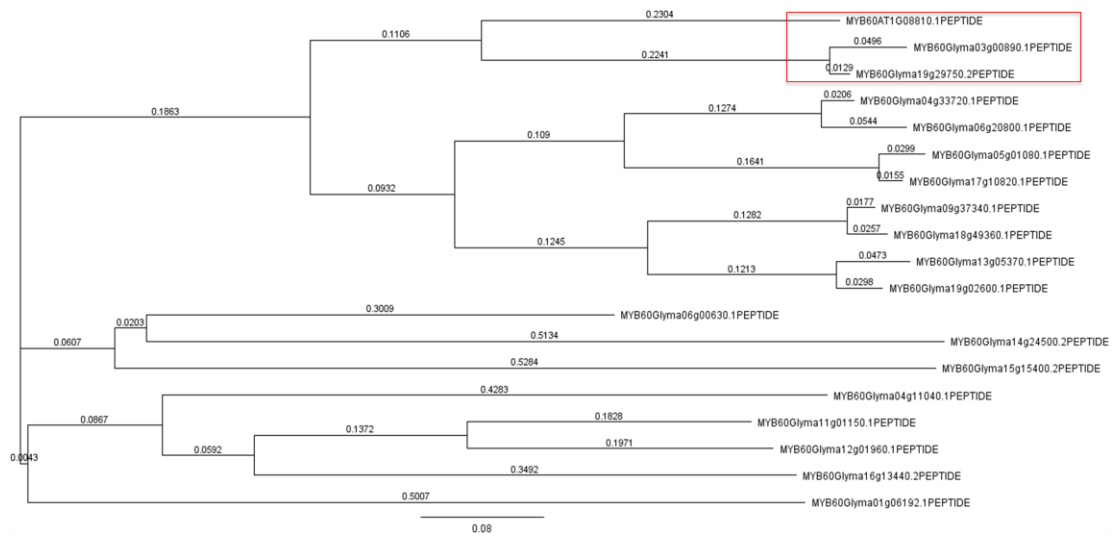


Figure 2.15. G. Phylogenetic analysis

Jukes-Cantor Neighbor Joining Phylogenetic Tree made using the *Arabidopsis thaliana* MYB60 peptide sequence and soybean sequences found using TBLASTN and Smith and Waterman alignments. Highlighted is the node that contains Glyma19g29750.2 and its duplicate which cluster with *AtMYB60* suggesting an evolutionary relationship.

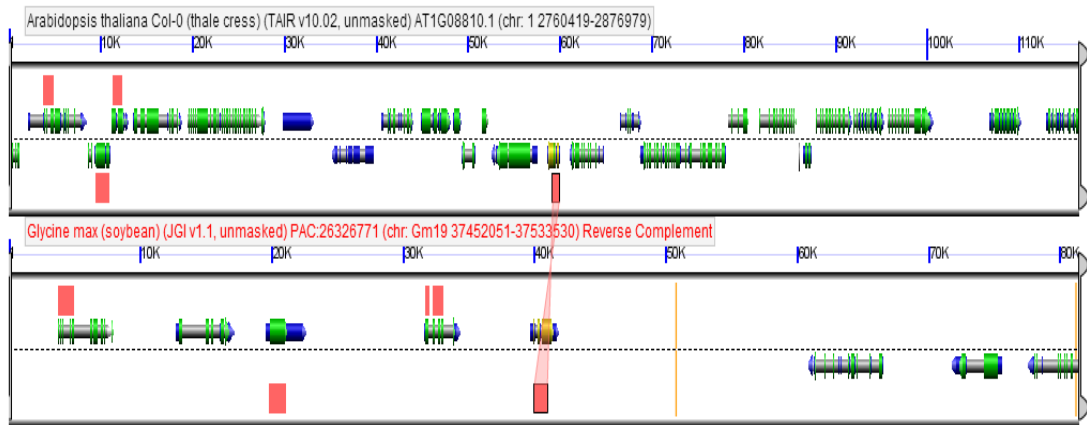


Figure 2.16. H. Syntenic analysis

Highlighted pink boxes represent the syntenic relationship observed between *AtMYB60* (At1g08810.1) and its ortholog in soybean (Glyma19g29750.2). Three other genes within a 60K region are also observed with similar functional annotations. This evidence implies a close evolutionary relationship between the two chromosomal “neighborhoods” where the gene resides.

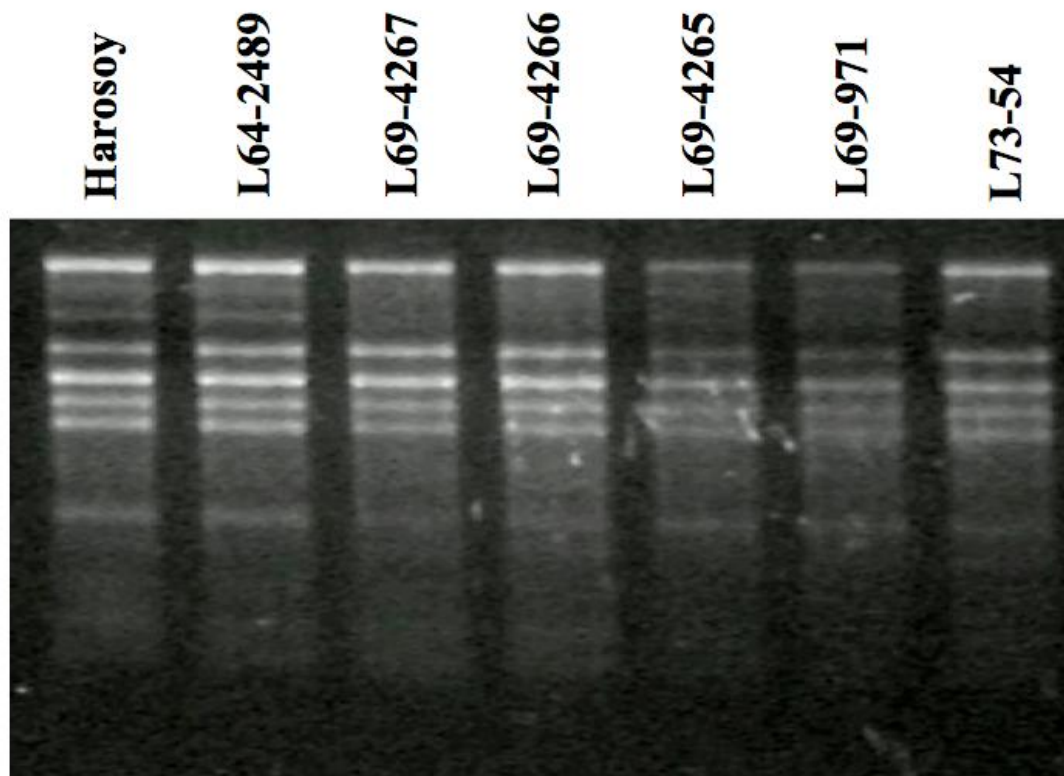


Figure 2.17: Gel electrophoresis of RNA samples extracted from all seven isogenic lines of soybean (Schreier and Chandlee, 2009). The gel shown is a representative agarose gel of RNA extracted from leaf tissue of seven isogenic lines of soybean (Table 1). All RNA's used in this study exhibited similar quality.

Table 2.1: Soybean “Evergreen” mutants available as Harosoy isolines

Line	Phenotype			Genotype	Gene Combination
	Leaves	Seed Coats	Pods		
Harosoy	WT	WT	WT	<i>ggD1D1D2D2</i>	--
L73-54*	WT	WT	WT	<i>ggd1d1D2D2</i>	<i>D1</i>
L69-4266*	WT	WT	WT	<i>ggD1D1d2d2</i>	<i>D2</i>
L69-4265*	WT ¹	Stay Green	WT ¹	<i>GGd1d1D2D2</i>	<i>Gd1</i>
L69-971*	WT ¹	Stay Green	WT ¹	<i>GGD1D1d2d2</i>	<i>Gd2</i>
L69-4267*	Stay Green ²	WT	Stay Green ³	<i>ggd1d1d2d2</i>	<i>d1d2</i>
L64-2489*	Stay Green ²	Stay Green	Stay Green ³	<i>GGd1d1d2d2</i>	<i>Gd1d2</i>

* Harosoy isolines

WT= Wild type (yellow leaves at maturity)

¹slightly delayed senescence relative to Harosoy

²abscission proceeds normally

³Pods turn green to brown with no yellow, transitional phase

RNA's extracted from various soybean backgrounds. These lines represent wild type (Harosoy) and various allelic combinations of single, double and triple mutants of *g*, *D1* and *D2*.

Table 2.2: Selected candidate gene families

Gene of Interest	Function
SIMYB	Senescence-Specific Transcription Factor
AoAS	Nitrogen Remobilization
AtSARK	Senescence-Associated Receptor Kinase (Transmembrane Protein, Leucine Rich)
AtSAG12	Senescence-Specific Cysteine Protease
AtSAG 2	Senescence-associated Cysteine Protease
AtSAG101	Acyl Hydrolase
AtMYB60	Anthocyanin Biosynthesis Transcription Factor ★
GmSEP1.1/.2	MADS-box Transcription Factor
GmSEP 3.3/.4	MADS-box Transcription Factor
GmSOC.1/.2	MADS-box Transcription Factor
GmAG.1/.2	MADS-box Transcription Factor

A database and literature search for genes associated with senescence in other plant species was conducted. A list was compiled and screened for soybean orthologs for which 3000 base pairs upstream were then evaluated for the presence of an SRE and conserved element. The *Arabidopsis thaliana MYB60* gene (At1g08810.1) (starred) will serve as an example of a typical gene screening strategy in this report.

Table 2.3: List of all genes meeting structural criteria from various families

Gene	Soybean (v1.1)	%ID 33 bp	%ID 7 bp	Proximity	Gene Architecture	Gene length	Alignment Method
TMYB	Glyma05g02550.2	54.5	71.4	(-1601,-1569)	3 exons	(-3000,+1948)	Geneious
	Glyma06g45547.1	51.5	71.4	(-1548,-1516)	3 exons	(-3000,+1512)	Geneious
	Glyma06g45554.1	51.5	85.7	(-645,-613)	3 exons	(-3000,+2392)	Geneious
	Glyma06g45570.2	51.5	100	(-60,-28)	3 exons	(-3000,+1847)	Geneious
	Glyma10g11200.1*	57.6	85.7	(-803,-771)	3 exons	(-3000,+2388)	Geneious
	Glyma11g01150.1	54.5	71.4	(-2100,-1978)	3 exons	(-3000,+1793)	Geneious
	Glyma12g11330.2	48.5	85.7	(-159,-127)	3 exons	(-3000,+2783)	Geneious
	Glyma12g32541.1	60.6	85.7	(-802,-770)	4 exons	(-3000,+2649)	Geneious
	Glyma19g34740.2	57.6	85.7	(-1141,-1109)	3 exons	(-3000,+2249)	Geneious
Soybean SARK	Glyma12g36090.1	63.6	85.7	(-1805,-1773)	22 exons	(-3000,+15882)	ClustalW
	Glyma13g34100.1	N/A	N/A	N/A	24 exons	(-3000,+12992)	Geneious/ClustalW
Arabidopsis thaliana SARK AT4G30520.1	Glyma14g01720.1	57.6	71.4	(-1606,-1574)	1 exon	(-3000,+5876)	Geneious
	AT4G30520.1	51.5	57.1	(+1448,+1480)	11 exons	(-3000,+3308)	Geneious
Asparagine Synthetase AT3G47340.1	Glyma11g27480.1	57.6	85.7	(-2342,-2310)	13 exons	(-3000,+4901)	Geneious
	Glyma11g38130.1	54.5	85.7	(-1336,-1304)	14 exons	(-3000,+5337)	Geneious
	Glyma18g06840.4	57.6	71.4	(-833,-801)	13 exons	(-3000,+5134)	ClustalW
	AT3G47340.1	54.5	57.1	(-2517,-2485)	10 exons	(-3000,+3360)	ClustalW
SAG12 AT5G45890.1	Glyma010s00210.2	51.5	71.4	(-70,-38)	2 exons	(-3000,+4105)	Geneious
	Glyma06g43530.1	57.6	71.4	(-2889,-2857)	2 exons	(-3000,+1113)	Geneious
	Glyma12g15120.2	54.5	71.4	(-1537,-1505)	2 exons	(-3000,+4105)	Geneious
	Glyma12g15130.1	51.5	85.7	(-421,-389)	2 exons	(-3000,+4108)	Geneious
	Glyma12g15750.2	51.5	85.7	(-1486,-1454)	2 exons	(-3000,+4522)	Geneious
	Glyma12g15760.1	57.6	71.4	(-1034,-1002)	2 exons	(-3000,+4226)	Geneious
	Glyma12g15780.1	54.5	71.4	(-554,-522)	2 exons	(-3000,+4226)	Geneious
	AT5G45890.1	100	100	(-659,-635)	3 exons	(-3000,+4672)	ClustalW
SAG2 AT5G60360.1	Glyma04g01640.1	54.5	71.4	(-991,-959)	4 exons	(-3000,+4803)	Geneious
	Glyma06g43530.1	57.6	71.4	(-2889,-2857)	2 exons	(-3000,+4341)	Geneious
	AT5G60360.1	N/A	N/A	(-1032,-1000)	8 exons	(-3000,+5472)	Geneious
SAG101 AT5G14930.2	Glyma06g19890.1	57.6	71.4	(-428,-396)	3 exons	(-3000,+3728)	Geneious
	Glyma13g04540.1	48.5	85.7	(-2339,-2307)	4 exons	(-3000,+3386)	Geneious
	Glyma13g04561.1	51.5	85.7	(-2817,-2785)	4 exons	(-3000,+2252)	Geneious
	Glyma19g01605.1	54.5	85.7	(-1478,-1446)	4 exons	(-3000,+7116)	Geneious
	AT5G14930.2	45.5	28.6	(-165,-133)	3 exons	(-3000,+2315)	Geneious
MYB60 AT1G08810.1	Glyma03g00890.1	54.5	71.4	(-2449,-2417)	3 exons	(-3000,+1139)	Geneious
	Glyma04g33720.1	57.6	71.4	(-1214,-1182)	3 exons	(-3000,+2040)	Geneious
	Glyma05g01080.1	57.6	71.4	(-2314,-2282)	3 exons	(-3000,+1945)	ClustalW
	Glyma11g01150.1	54.5	71.4	(-2010,-1978)	3 exons	(-3000,+1793)	Geneious
	Glyma12g01960.1	51.5	100	(-2389,-2357)	3 exons	(-3000,+1773)	Geneious
	Glyma16g13440.2	57.6	71.4	(-1467,-1435)	3 exons	(-3000,+1500)	ClustalW
	Glyma18g49360.1	60.6	85.7	(-2542,-2510)	3 exons	(-3000,+2219)	Geneious
	Glyma19g29750.2	51.5	71.4	(-2180,-2148)	3 exons	(-3000,+1919)	ClustalW
	AT1G08810.1	42.5	71.4	(-559,-527)	3 exons	(-3000,+1331)	Geneious
MADS MADS 28/SEP1	Glyma11g36890.1	54.5	71.4	(-209,-177)	6 exons	(-3000,+4537)	Geneious
	Glyma18g00801.1**	N/A	N/A	N/A	6 exons	(-3000,+4745)	Geneious /ClustalW
	Glyma13g06730.1	57.6	71.4	(-1534,-1502)	8 exons	(-3000,+8571)	Geneious

Highlighted represent genes selected for functional analysis.

*Alternate TMYB not found by this study. Gene was identified through BLASTN of soybean database using the 70 base pair sequence identified by microarray (Schreier and Chandlee, 2009)

**Gene did not meet structural criteria detailed in this study. Gene was examined to determine whether Glyma11g36890.1 or Glyma18g00801.1 harboring an SRE was responsible for previously observed differential expression.

Table 2.4: PCR primers for eight candidate senescence associated genes (SAG's) selected for developmental expression analysis in soybean leaves

Name	Gene Model	Transcript Size	Product Size	Forward Primer	Reverse Primer
SAG101	Glyma13g04540.1	2424 bp	136 bp	TGAACGTCCTGTG GTTGAA	CCACAACAAGGGA GGTA
TMYB	Glyma06g45554.1	989 bp	218 bp	CATCCACAAGCACA GTGGTC	CAAGCACAGAGATGT CAGTTAAG
TMYB (alternate)	Glyma10g11200.1	1230 bp	244 bp	AATGCTCTTTGGT GTGCG	TGGGTAGAAATCGGC GTAGC
11MADS28	Glyma11g36890.1	1140 bp	239 bp	GTCAGAAGGCCAA GCATGA	AGCTAGCTAGGTGCT GTTTATGA
18MADS28	Glyma18g00801.1	1008 bp	235 bp	TGAGTGTGGACATA CTCTGAGC	CCGCTGCAAACTCG TACTG
SARK	Glyma12g36090.1	3632 bp	100 bp	TTCTGATACCGACA CACCGC	AGGGAAGCCAGGAAC ACAAA
AS	Glyma18g06840.4	2165 bp	142 bp	GTAGTTGCCCTAGA ACCAGCA	GGATGGCACAACACT TAAGAC
MYB60	Glyma19g29750.2	1700 bp	151 bp	GAGAACAGTGGTGC TCTCC	CACTTCCAACTTG AAGGGA
Actin (<i>Sac3</i>) (soybean)*	/	1128 bp	295 bp	GCGAGAAATTGTCC GTGACA	TTGAGCCACCTAA GAACA

*control primer set

Table 2.5: RNA samples used for expression analysis

	V5*	Microarray # / Year	R7**	Microarray # / Year
	[] ng/uL		[] ng/uL	
GENOTYPES				
Harosoy	3053	Microarray 1 / 2009	1069	Microarray 6 pt. 2/ 2008
L64-2489	3288	Microarray 1 / 2009	1073	Microarray 6 pt. 2/ 2008

*V5- The developmental stage at which five fully expanded, green trifoliolate leaves are on the plant (www.ag.ndsu.edu/pubs/plantsci.htm).

**R7- The developmental stage at which one major pod has changed to a brown color on the main stem (www.ag.ndsu.edu/pubs/plantsci.htm).

Canfield MR, Guiamet JJ, Nooden LD: **Alteration of soybean seedling development in darkness and light by the stay green mutation *cytG* and *Gd1d2***. *Annals of Botany* (1995) **75**: 143-150.

Gan S. and Amasino R.M. **Inhibition of Leaf Senescence by Autoregulated Production of Cytokinin** *Science* (1995) **270** (5244): 1986-1988.

Guiamet JJ and Gianibelli MC: **Nuclear and cytoplasmic stay green mutations of soybean alter the loss of leaf soluble proteins during senescence**. *Physiologia Plantarum* (1996) **96**: 655-661.

Guiamet J.J., Shwartz E., Pichersky E. and Nooden L.D. **Characterization of cytoplasmic and nuclear mutations affecting chlorophyll and chlorophyll binding proteins during senescence in soybean**. *Plant Physiology* (1991) **96**: 227-231.

Lim P.O., Woo H.R. and Nam H.G. **Molecular genetics of leaf senescence in *Arabidopsis*** *Trends in Plant Science*, (2003) **8** (6): 272-278.

Noh Y-S and Amasino R.M. **Identification of a promoter region responsible for the senescence-specific expression of SAG12** *Plant Molecular Biology* (1999) **41**: 181–194.

Quirino B.F., Noh Y-S., Himelblau E. and Amasino R.M. **Molecular aspects of leaf senescence** *Trends in Plant Science Reviews* (2000) **5** (7): 278-282.

Schreier, L. and Chandlee J.M. **Characterization of MADS-Box Genes of Soybean (*Glycine max*, *L. Merr.*) and Initial Functional Analysis in Whole Plant Senescence**. *Plant Biology* (2009)

Swidzinski J.A., Sweetlove L.J. and Leaver C.J. **A custom microarray analysis of gene expression during programmed cell death in *Arabidopsis thaliana*** *Plant Journal* (2002) **30** (4): 431-46.

Winichayakul S., Moyle R.L., Ryan D.J., Farnden K.J.F., Davies K.M., and Coupe S.A. **Distinct cis-elements in the *Asparagus officinalis* asparagine synthetase promoter respond to carbohydrate and senescence signals** *Functional Plant Biology* (2004) **31**: 573–582.

Zeng P. **Developmental regulation of gene expression during senescence in soybean** *Doctoral Dissertation* (1999) University of Rhode Island

Bibliography

Abe H., Urao T., Ito T., Seki M., Shinozaki K., and Yamaguchi-Shinozaki K. ***Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling** *Plant Cell* (2003) **15**: 63-78.

The American Soybean Association copyright ©2014 www.soystats.com accessed June 15, 2013

Balazadeh S., Riano-Pachon D.M. and Mueller-Roeber B. **Transcription factors regulating leaf senescence in *Arabidopsis thaliana*** *Plant Biology* (2008) **10**: 63–75.

Becker, A. and Theissen, G. **The major clades of MADS-box genes and their role in the development and evolution of flowering plants** *Molecular Phylogenetics and Evolution* (2003) **29**: 464–489.

Bolon Y.T., Haun W.J., Xu W.W., Grant D., Stacey M.G., Nelson R.T., Gerhardt D.J., Jeddeloh J.A., Stacey G., Muehlbauer G.J., Orf J.H., Naeve S.L., Stupar R.M. and Vance C.P. **Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean** *Plant Physiology* (2011) **156** (1): 240-53.

Berg J.M., Tymoczko J.L., Stryer L., and Gatto G.J. **Biochemistry 7th ed.** *W. H. Freeman and Company* (2012) 941-943.

Brivanlou A.H. and Darnell J.E. Jr. **Signal transduction and the control of gene expression** *Science* (2002) **295** (5556): 813-81.

Buchanan-Wollaston V. **The molecular biology of leaf senescence** *Journal of Experimental Botany* (1997) **48** (307): 181-199.

Canfield M.R., Guiamét J.J., and Noodén L.D. **Alteration of Soybean Seedling Development in Darkness and Light by the Stay-green Mutation cytG and Gd1d2** *Annals of Botany* (1995) **75** (2): 143-150.

Chandlee J.M. **Current molecular understanding of the genetically programmed process of leaf senescence** *Physiologia Plantarum* (2001) **113**: 1–8.

Cheema J. and Dicks J. **Computational approaches and software tools for genetic linkage map estimation in plants** *Briefings in Bioinformatics* (2009) **10** (6): 595-608.

Chen L-F. O., Hwang J-Y., Charng Y_Y., Sun C-W. and Yang S-F. **Transformation of broccoli (*Brassica oleracea* var. *italica*) with isopentenyltransferase gene via *Agrobacterium tumefaciens* for post-harvest yellowing retardation** *Molecular Breeding* (2001) **7**: 243–257.

College of Life Sciences National Laboratory of Protein Engineering and Plant Genetic Engineering and Center for Bioinformatics (2012) Leaf Senescence Database 2012 <http://www.eplantsenescence.org/> accessed April 27, 2012

Cooper J.L., Till B.J., Laport R.G., Darlow M.C., Kleffner J.M., Jamai A., El-Mellouki T., Liu S., Ritchie R., Nielsen N., Bilyeu K.D., Meksem K., Comai L. and Henikoff S. **TILLING to detect induced mutations in soybean** *BMC Plant Biology* (2008) **24**: 8-9.

Davies KM. and King GA. **Isolation and characterization of a cDNA clone for harvest induced asparagine synthetase from *Asparagus officinalis* L.** *Plant Physiology* (1993) **102**: 1337-1340.

Du H., Yang S.S., Liang Z., Feng B.R., Liu L., Huang Y.B., and Tang Y.X. **Genome-wide analysis of the MYB transcription factor superfamily in soybean** *BMC Plant Biology* (2012) **12** (106): 1-22.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. and Lepiniec, L. **MYB transcription factors in *Arabidopsis*** *Trends in Plant Science* (2010) **15** (10): 573-581.

Ellis C.M., Nagpal P., Young J.C., Hagen G., Guilfoyle T.J. and Reed J.W. **AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*** *Development* (2005) **132** (20): 4563-4574.

Fernandez, D., Heck, G.R., Perry, S.E., Patterson, S.E., Bleecker, A.B., and Fang, S.C. **The embryo MADS domain factor AGL15 acts postembryonically: inhibition of perianth senescence and abscission via constitutive expression** *Plant Cell* (2000) **12**: 183–197.

Gan S. **Mitotic and Postmitotic Senescence in Plants** *Science of Aging Knowledge Environment* (2003) **38** (7): 1-10.

Gan S. and Amasino R.M. **Inhibition of Leaf Senescence by Autoregulated Production of Cytokinin** *Science* (1995) **270** (5244): 1986-1988.

Gan S. and Amasino R.M. **Making Sense of Senescence** *Plant Physiology* (1997) **113**: 313-319.

Gatlin III D.M., Barrows F.T., Brown P, Dabrowski D., Gaylord T.G, Hardy R.W., Herman E., Hu G., — Krogdahl A., Nelson R., Overturf K., Rust M., Sealey W., Skonberg D., Souza E.J., Stone D., Wilson R., and Wurtele E. **Expanding the**

utilization of sustainable plant products in aquafeeds: a review *Aquaculture Research* (2007) **38**: 551-579.

Gaufichona L., Reisdorf-Crena M., Rothstein SJ., Chardona F., and Suzukia A. **Biological functions of asparagine synthetase in plants** *Plant Science* (2010) **179**: 141–153.

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

Gramzow, L. and Theissen, G. **A hitchhiker’s guide to the MADS world of plants** *Genome Biology* (2010) **11**: 1-11.

Geneious version (6.1.5) created by Biomatters. available from <http://www.geneious.com/>

Grant D., Nelson R.T., Cannon S.B., and Shoemaker R.C. **SoyBase, the USDA-ARS soybean genetics and genomics database** *Nucleic Acids Research* (2010) **38** (Database issue): D843–D846.

Grbic V. **SAG2 and SAG12 protein expression in senescing *Arabidopsis* plants** *Physiologia Plantarum* (2003) **119**: 263–269.

Gregersen P.L., Culetic A., Boschian L. and Krupinska K. **Plant senescence and crop productivity** *Plant Molecular Biology* (2013) [Epub ahead of print]: 1-20.

Griffiths A.J.F., Miller J.H, Suzuki D.T., Lewontin R.C. and Gelbart W.M. **An Introduction to Genetic Analysis, 7th ed.** New York: W. H. Freeman (2000) Linkage maps.

Guamet J.J. and Giannibelli C.M. **Nuclear and cytoplasmic “stay-green” mutations of soybean alter the loss of leaf soluble proteins during senescence** *Physiologia Plantarum* (1996) **96** (4): 655–661.

Guamét J.J., Schwartz E., Pichersky E., and Noodén L.D. **Characterization of Cytoplasmic and Nuclear Mutations Affecting Chlorophyll and Chlorophyll-Binding Proteins during Senescence in Soybean** *Plant Physiology* (1991) **96** (1): 227-231.

Guo Y., Cai Z., and Gan S. **Transcriptome of *Arabidopsis* leaf senescence** *Plant, Cell and Environment* (2004) **27**: 521–549.

Guo Y and, Gan S. **AtNAP, a NAC family transcription factor, has an important role in leaf senescence** *The Plant Journal* (2006) **46** (4): 601-612.

- Hajouj T., Michelis R. and Gepstein S. **Cloning and characterization of a receptor-like protein kinase gene associated with senescence** *Plant Physiology* (2000) **124** (3): 1305-1314.
- Hamilton J.P. and Buell R.C. **Advances in plant genome sequencing** *The Plant Journal* (2012) **70**: 177–190.
- He Y. and Gan S. **A gene encoding an acyl hydrolase is involved in leaf senescence in *Arabidopsis*** *The Plant Cell* (2002) **14**: 805–815.
- Higginson T., Li S.F., and Parish R.W. **AtMYB103 regulates tapetum and trichome development in *Arabidopsis thaliana*** *The Plant Journal* (2003) **35**: 177-192.
- Higo K., Ugawa, Iwamoto M. and Korenaga T. **Plant cis-acting regulatory DNA elements (PLACE) database** *Nucleic Acids Research* (1999) **27** (1): 297-300.
- Hill J., Nelson E., Tilman D., Polasky S., and Tiffany D **Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels** *Proceedings of the National Academy of Sciences* (2006) **103** (30): 11206–11210.
- Hinderhofer K. and Zentgraf U. **Identification of a transcription factor specifically expressed at the onset of leaf senescence** *Planta* (2001) **213** (3): 469-473.
- Huynh L.E., Toai T.V., Streeter J. and Banowetz G. **Regulation of flooding tolerance of SAG12:ipt *Arabidopsis* plants by cytokinin** *Journal of Experimental Botany* (2005) **56**: 1397–1407.
- Hyten D.L., Choi I-Y., Song Q, Specht J.E., Carter T.E., Shoemaker R.C., Hwang E-Y., Matukumalli L.K. and Cregan P.B. **A High Density Integrated Genetic Linkage Map of Soybean and the Development of a 1536 Universal Soy Linkage Panel for Quantitative Trait Locus Mapping** *Crop Science* (2010) **50**: 960–968.
- Jaiswal P., Ni J., Yap I., Ware D., Spooner W., Youens-Clark K., Ren L., Liang C., Zhao W., Ratnapu K., Faga B., Canaran P., Fogleman M., Hebbard C., Avraham S., Schmidt S., Casstevens T.M., Buckler E.S., Stein L. and McCouch S. **Gramene: a bird's eye view of cereal genomes** *Nucleic Acids Research* (2006) **34** (Database issue): D717-D723.
- Jung C., Seo J.S., Han S.W., Koo Y.J., Kim C.H., Song S.I., Nahm B.H., Choi Y.D., and Cheong J.J. **Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*** *Plant Physiology* (2008) **146**: 623-635.

- Kim K.S., Hill C.B., Hartman C.L., Hyten D.L., Hudson M.E., Diers B.W. **Fine mapping of the soybean aphid-resistance gene Rag2 in soybean PI 200538** *Theoretical and Applied Genetics* (2010) **121**: 599–610.
- Kim M., Hyten D.L., Bent A.F. and Diers B.W. **Fine Mapping of the SCN Resistance Locus rhg1-b from PI 88788** *The Plant Genome* (2010) **3**: 81-89.
- King G.A., Davies K.M., Stewart R.J., and Borst W.M. **Similarities in gene expression during the postharvest-induced senescence of spears and the natural foliar senescence of asparagus** *Plant Physiology* (1995) **108**: 125-128.
- Krul W.R. **Nucleic Acid and Protein Metabolism of Senescing and Regenerating Soybean Cotyledons'** *Plant Physiology* (1974) **54**: 36-40.
- Lamesch P., Berardini T.Z., Li D., Swarbreck D., Wilks C., Sasidharan R., Muller R., Dreher K., Alexander D.L., Garcia-Hernandez M., Karthikeyan A.S., Lee C.H., Nelson W.D., Ploetz L., Singh S, Wensel A. and Huala E. **The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools** *Nucleic Acids Research* (2011) 1–9.
- Latchman D.S. **Transcription factors: an overview** *International Journal of Biochemistry and Cell Biology* (1997) **12** (1): 305-12.
- Lee T.H., Tang H., Wang X. and Paterson A.H. **PGDD: a database of gene and genome duplication in plants** *Nucleic Acids Res.* (2013) **41** (Database issue): D1152-D1158.
- Li XP., Gan R., Li PL., Ma YY., Zhang LW., Zhang R., Wang R., and Wang NN. **Identification and functional characterization of a leucine-rich repeat receptor-like kinase gene that is involved in regulation of soybean leaf senescence** *Plant Molecular Biology* (2006) **61**: 829–844.
- Lim P.O., Lee I.C., Kim J., Kim H.J., Ryu J.S., Woo H.R. and Nam HG. **Auxin response factor 2 (ARF2) plays a major role in regulating auxin-mediated leaf longevity** *Journal Experimental Botany* (2010) **61** (5): 1419-1430.
- Lim P.O., Kim H.J., and Nam H.G. **Leaf Senescence** *Annual Review of Plant Biology* (2007) **58**: 115–36.
- Lim P.O., Woo H.R. and Nam H.G. **Molecular genetics of leaf senescence in Arabidopsis** *Trends in Plant Science* (2003) **8** (6): 272-278.
- Liu L., Zhou Y., Szczerba M.W., Li X., and Lin Y. **Identification and application of a rice senescence-associated promoter** *Plant Physiology* (2010) **53** (3): 1239-1249.

Liu X., Li Z., Jiang Z., Zhao Y, Peng J., Jin J., Guo H., Luo J. **LSD: a leaf senescence database** *Nucleic Acids Research* (2011) **39** (Database issue) D1103–D1107.

Lyons E. and Freeling M. **How to usefully compare homologous plant genes and chromosomes as DNA sequences** *The Plant Journal* (2008) **53** (4): 661–673.

Mathieu M., Winters E.K., Kong F., Wan J., Wang S., Eckert H., Luth D., Paz M., Donovan C., Zhang Z., Somers D., Wang K., Nguyen H., Shoemaker R.C., Stacey G. and Clemente T. **Establishment of a soybean (*Glycine max* Merr. L) transposon-based mutagenesis repository** *Planta* (2009) **229** (2): 279-89.

Martin C. and Paz-Ares J. **MYB transcription factors in plants** *Trends in Genetics* (1997) **13** (2): 67-73.

Maroof S.M.A., Tucker D.M., Skoneczka J.A., Bowman B.C., Tripathy S. and Tolin S.A. **Fine Mapping and Candidate Gene Discovery of the Soybean Mosaic Virus Resistance Gene, *Rsv4*** *The Plant Genome* (2010) **3** (1): 14-22.

McCabe M.S., Garratt L.C., Schepers F., Jordi W.J.R.M, Stoopen G.M., Davelaar E., Rhijn J.H.A.V, Power J.B., and Davey M.R. **Effects of *PSAG12-IPT* Gene Expression on Development and Senescence in Transgenic Lettuce** *Plant Physiology* (2001) **127**: 505–516.

Mehrtens F., Kranz H., Bednarek P., and Weisshaar B. **The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis** *Plant Physiology* (2005) **138**: 1083-1096.

Miao Y., Laun T., Zimmermann P. and Zentgraf U. **Targets of the WRKY53 transcription factor and its role during leaf senescence in *Arabidopsis*** *Plant Molecular Biology* (2004) **55** (6): 853-867.

Noh Y-S and Amasino R.M. **Identification of a promoter region responsible for the senescence-specific expression of *SAG12*** *Plant Molecular Biology* (1999) **41**: 181–194.

Noodén L.D., Guiamet J.J. and John I. **Senescence mechanisms** *Physiologia Plantarum* (1997) **101**: 746-753.

Noodén L.D. and Leopold A.C., **Senescence and Aging in Plants** *Academic Press San Diego* (1988) 33-34, 506-507.

North Dakota State University Soybean Production Field Guide
<http://www.ag.ndsu.edu/pubs/plantsci/rowcrops/a1172a.pdf> accessed June 15, 2013

Okushima Y., Overvoorde P.J., Arima K., Alonso J.M., Chan A., Chang C., Ecker J.R., Hughes B., Lui A., Nguyen D., Onodera C., Quach H., Smith A., Yu G., and

Theologis A. **Functional Genomic Analysis of the Auxin Response Factor Gene Family Members in *Arabidopsis thaliana*: Unique and Overlapping Functions of ARF7 and ARF19** *The Plant Cell* (2005) **17**: 444–463.

Ori N., Juarez M.T., Jackson D., Yamaguchi J., Banowetz G.M., and Hake S. **Leaf Senescence Is Delayed in Tobacco Plants Expressing the Maize Homeobox Gene *knotted1* under the Control of a Senescence-Activated Promoter** *The Plant Cell* (1999) **11**: 1073–1080.

Pabo C.O. and Sauer R.T. **Transcription factors: structural families and principles of DNA recognition** *Annual Review of Biochemistry* (1992) **61** (1): 053-95.

Park J.S., Kim J.B., Cho K.J., Cheon C.I., Sung M.K., Choung M.G., and Roh K.H. ***Arabidopsis* R2R3-MYB transcription factor AtMYB60 functions as a transcriptional repressor of anthocyanin biosynthesis in lettuce (*Lactuca sativa*)**. *Plant Cell Reports* (2008) **27**: 985–994.

Peleg Z. and Blumwald E. **Hormone balance and abiotic stress tolerance in crop plants** *Current Opinions in Plant Biology* (2011) **14**: 290–295.

Perry, S.E., Nichols, K.W., and Fernandez, D.E. **The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development** *Plant Cell* (1996) **8**: 1977–1989.

Ptashne, M. **How eukaryotic transcriptional activators work** *Nature* (1988) **335**: 683–689.

Qin H., Gu Q., Zhang J., Sun L., Kuppu S., Zhang Y., Burow M., Payton P., Blumwald E. and Zhang H. **Regulated expression of an isopentenyltransferase gene (IPT) in peanut significantly improves drought tolerance and increases yield under field conditions** *Plant Cell Physiology* (2011) **52** (11): 1904-1914.

Quirino B.F., Noh Y-S., Himelblau E. and Amasino R.M. **Molecular aspects of leaf senescence** *Trends in Plant Science Reviews* (2000) **5** (7): 278-282.

Reed J.W. **Roles and activities of Aux/IAA proteins in *Arabidopsis*** *Trends in Plant Science* (2001) **6** (9): 420-425.

Riechmann J.L. and Meyerowitz E.M. **The AP2/EREBP family of plant transcription factors** *The Journal of Biological Chemistry* (1998) **379** (6): 633-646.

Rivero R.M., Kojima M., Gepstein A., Sakakibara H., Mittler R., Gepstein S. and Blumwald E. **Delayed leaf senescence induces extreme drought tolerance in a flowering plant** *Proceedings of the National Academy of Sciences* (2007) **104** (49): 19631-19636.

Rivero R.M., Gimeno J., Van Deynze A., Walia H. and Blumwald E. **Enhanced cytokinin synthesis in tobacco plants expressing P-SARK:IPT prevents the degradation of photosynthetic protein complexes during drought** *Plant Cell Physiology* (2010) **51**: 1929-1941.

Robatzek S. and Somssich I.E. **Targets of AtWRKY6 regulation during plant senescence and pathogen defense** *Genes and Development* (2002) **16**: 1139–1149.

Rubinstein B. and Leopold A.C. **The Nature of Leaf Abscission** *Quarterly Review of Biology* (1964) **39** (4): 356-372.

Rubio-Somoza I. and Weigel D. **MicroRNA networks and developmental plasticity in plants** *Trends in Plant Science* (2011) **16** (5): 258-264.

Schmutz J., Cannon S.B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D.J., Song Q. , Thelen J.J., Cheng J., Xu D., Hellsten U., May G.D., Yu Y., Sakurai T., Umezawa T., Bhattacharyya M.K., Sandhu D., Valliyodan B., Lindquist E., Peto M., Grant D., Shu S., Goodstein D., Barry K., Futrell-Griggs M., Abernathy B., Du J., Tian J., Zhu L., Gill N., Joshi T., Libault M., Sethuraman A., Zhang X-C., Shinozaki K., Nguyen H.T., Wing R.A., Cregan P., Specht J., Grimwood J., Rokhsar D., Stacey G., Shoemaker R.C. and Jackson S.A. **Genome sequence of the palaeopolyploid soybean** *Nature* (2010) **463**: 178-183.

Schreier, L. and Chandlee J.M. **Characterization of MADS-Box Genes of Soybean (*Glycine max*, *L. Merr.*) and Initial Functional Analysis in Whole Plant Senescence.** *Plant Biology* (2009).

Simmonds N.W. **Monocarpy, Calendars and Flowering Cycles in Angosperms** *Kew Bulletin* (1980) **35** (2): 235-245.

Smart M. C. **Tansley Review No. 64 Gene expression during leaf senescence.** *New Phytologist* (1994) **126**: 419-448.

Stracke R., Werber M. and Weisshaar B. **The R2R3-MYB gene family in *Arabidopsis thaliana*** *Current Opinion in Plant Biology* (2001) **4**: 447–456.

Swartzberg D., Dai N., Gan S., Amasino R., and Granot D. **Effects of Cytokinin Production under Two *SAG* Promoters on Senescence and Development of Tomato Plants** *Plant Biology* (2006) **8**: 579–586.

Swartzberg D., Kirshner B., Rav-David D., Elad Y. and Granot D. **Botrytis cinerea induces senescence and is inhibited by autoregulated expression of the IPT gene** *European Journal of Plant Pathology* (2007) **120** (3): 289-297.

Swidzinski J.A., Sweetlove L.J. and Leaver C.J. **A custom microarray analysis of gene expression during programmed cell death in *Arabidopsis thaliana*** *Plant Journal* (2002) **30** (4): 431-46.

Walker A.B., Sidor I.F., O'Keefe T., Cremer M. and Berlinsky D.L. **Partial Replacement of Fish Meal with Soy Protein Concentrate in Diets of Atlantic Cod** *North American Journal of Aquaculture* (2010) **72** (4): 343-353.

Winichayakul S., Moyle R.L., Ryan D.J., Farnden K.J.F., Davies K.M., and Coupe S.A. **Distinct cis-elements in the *Asparagus officinalis* asparagine synthetase promoter respond to carbohydrate and senescence signals** *Functional Plant Biology* (2004) **31**: 573–582.

Woo H.R., Kim J.E., Kim J., Kim J., Lee U., Song I-J., Kim J-H., Lee H-Y., Nam H.G. and Lim P.O. **The RAV1 transcription factor positively regulates leaf senescence in *Arabidopsis*** *Journal of Experimental Botany* (2010) **61** (14): 3947–3957.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li1, Z., Xiaohui1, D, Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G. and Li-Jia, Q. **The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family** *Plant Molecular Biology* (2006) **60**: 107–124.

Ye J., Coulouris G., Zaretskaya I., Cutcutache I., Rozen S., and Madden T.L. **Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction** *BMC Bioinformatics* (2012) **13**: 134.

Yuan Z., Pan A., Jian Z., Xu S., Gan S., Huang H., Zhang D. **Senescence delay characterization of transgenic *Brassica chinensis* L. containing an anti-senescence chimeric gene SAG12-IPT** *Journal of Plant Physiology and Molecular Biology* (2002) **28**: 379-384.

Zemach A. and Grafi G. **Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation** *TRENDS in Plant Science* (2007) **12** (2): 80-85.

Zeng P. **Developmental regulation of gene expression during senescence in soybean** *Doctoral Dissertation* (1999) University of Rhode Island

Zhang X., Ju H.W., Chung M.S., Huang P., Ahn S.J. and Kim C.S. **The R-R-type MYB-like transcription factor, AtMYBL, is involved in promoting leaf senescence and modulates an abiotic stress response in *Arabidopsis*** *Plant Cell Physiology* (2011) **52** (1): 138-148.

Zhang Z.J., Chen J., H.T. Nguyen. **Auto-regulated expression of bacterial isopentenyltransferase gene promotes T-DNA transformation in soybean** *United States Patent Application Publication* (2008) Publication No. US 2008/0184393 A1.