uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Candace June Webb

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biology)

GRADE / DEGREE

Department of Biology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Structure and Regulation of a Novel Multigene Family Implicated in Soybean Nodule Development in the Context of Senescence

TITRE DE LA THÈSE / TITLE OF THESIS

Douglas Johnson

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

John Arnason

Brian Miki

Linda Bonen

Sharon Regan (Queen's University)

Christine Charest

Gary W. Slater

Le Doyen de la Faculté des études superieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Structure and Regulation of a Novel Multigene Family Implicated in Soybean Nodule Development in the Context of Senescence

BY

Candace June Webb

Thesis submitted to the School of Graduate Studies
and Research in partial fulfillment of the requirements
for the Ph.D. degree in Biology

University of Ottawa
Ottawa, Ontario
2006

Your file Votre référence
ISBN: 978-0-494-18613-8
Our file Notre référence
ISBN: 978-0-494-18613-8

I declare that I am the sole author of this thesis.

I authorize the University of Ottawa to lend this thesis to individuals or other institutions for the purpose of scholarly research.

I further authorize the University of Ottawa to reproduce this thesis by photocopying or other means, in total or in part, at the request of other institutions for the purpose of scholarly research.

Candace Webb

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Dr. Doug Johnson, for his support and help throughout this research and for taking a chance on a rookie molecular biologist.

I would also like to thank my committee members, Dr. Linda Bonen, Dr. Brian Miki, and Dr. John Arnason for their many helpful suggestions and advice throughout this project.

In addition, I would like to extend thanks to Andrew Ochalski for help with the light microscopy and John Livesey and Andrew Burt for their help with setting up the gas chromatography experiments.

Many people have shared good times and laughter over the years, which was much needed and appreciated. They include the now Aussies, Jennifer Georgeff and Rob Beiko – thanks for the good times!

I have had many fantastic lab mates throughout this research, but I would especially like to thank Neda Mansoorian, my partner in crime, whose unfailing cheeriness and optimism saw me through. I would also like to thank the Bonen lab, especially Jen, Sophie, and Tom, for their support and companionship throughout my time here.

Also, thank you to my running partner and friend Haude Levesque for being there to listen, give good advice, and keep me in shape!

I could not have done this without my good friends Marsha Ostrovsky and Jessica Head, who let me complain or be silly on so many occasions (about so many things!) and made it all better with keen ears, sage advice, laughter, and the indispensable beer and chicken wings. Jessica, aka The Thesis Therapist and the architect of The Task-Based

Model ™, said: "Just do it, or say 'Screw it!'". Marsha, the graduate school veteran, said: "Deeeewwwww iiiitttt!". I can't quite believe it, but I have done it. Whoo hoo!

You've all made this trip well worth it and I will miss you.

Thank you to my friend Nicole Knibb, who has been there with me through all of this from the "thinking about it" stage to the application process to the end. Without her constant support, I would not have made it this far. Many rescue missions were made to Ottawa.

I would also like to thank my family, especially my mother, for encouraging me to keep going even though the road was long and hard.

Finally, I'd like to thank Darren Reid for his unfailing love and support. Thank you for being more stubborn than me and refusing to let me get away with not doing this for myself, even when I'd grown tired and didn't believe I could. "Box!"

ABSTRACT

The soybean (*Glycine max* L. Merrill) root nodule is an organ specialized for nitrogen fixation resulting from a symbiosis between the bacterium *Bradyrhizobium japonicum* and the roots. Root nodule senescence, its last stage of development, was the focus of this study. The progression of soybean root nodule senescence was examined using nitrogenase activity. During nodule development, nitrogenase activity was first observed with the appearance of nodules and peaked just prior to flowering, which marks the onset of nodule senescence. Accordingly, nitrogenase activity declined after flowering. Although nitrate and dark treatment both caused significant declines in nitrogenase activity, darkness caused a greater decline in nitrogenase activity than nitrate.

The *SAN1A* gene, originally isolated from senescent root nodules (Chan, 1995), and two other members of the *SAN1* gene family, were characterized. These genes represent a novel family of highly conserved, tandemly repeated genes in soybean. *SAN1A* and *SAN1B* encode predicted proteins of 352 and 353 amino acids, respectively, whereas *SAN1C* encodes a truncated protein of 126 amino acids and is likely a pseudogene. Despite their occurrence in tandem, gene conversion does not appear to have occurred within the evolutionary history of these genes. During natural senescence, *SAN1A* and *SAN1B* expression patterns suggested that *SAN1A* may play a role throughout nodule development, whereas *SAN1B* is likely to be more important during senescence. Both nitrate and dark treatment caused a decrease in *SAN1A* expression and an increase in *SAN1B* expression. Although both *SAN1A* and *SAN1B* were expressed in all soybean tissues examined, their patterns of expression differed. These results suggest that *SAN1A* and *SAN1B* have divergent functions or that *SAN1B* may replace *SAN1A* under conditions

such as senescence. Preliminary results also suggest that *SAN1A* may be light regulated. Finally, a test of *SAN1C* expression in soybean tissues showed that it is expressed, despite the fact that it encodes a truncated protein.

Database searches suggested that *SAN1* genes are related to plant dioxygenases. A gene tree and BLASTP searches revealed that the *SAN1* gene products are most closely related to uncharacterized dioxygenases identified in *Arabidopsis*, rice, and maritime pine, with the closest, identifiable dioxygenases being gibberellin 20-oxidases. Thus, the *SAN1* genes may encode paralogues of known dioxygenases or novel dioxygenases. The evidence collected to date suggests that the *SAN1* genes encode dioxygenases with roles in various aspects of plant metabolism. Further characterization of this family may yield more clues about their role(s).

RÉSUMÉ

Les nodules racinaires de soja (*Glycine max* L. Merrill), qui sont le résultat d'une symbiose entre la bactérie *Bradyrhizobium japonicum* et les racines, sont des organes spécialisés pour la fixation de l'azote. L'objet de cette étude est la sénescence du nodule racinaire, dernière étape de son développement. La progression de la sénescence a été examinée en utilisant l'activité de la nitrogénase. Durant le développement des nodules, l'activité de la nitrogénase a été observée lorsque les nodules apparaissent. L'activité a atteint son maximum juste avant la floraison de la plante, marquant ainsi le commencement de la sénescence des nodules. En conséquence, l'activité de la nitrogénase a diminué après la floraison. Les deux traitements utilisés dans cette étude, l'exposition aux nitrates et à l'obscurité, ont conduit à la sénescence des nodules et au déclin significatif de leur activité nitrogénasifier. Cependant, l'obscurité a causé un plus grand déclin dans l'activité de la nitrogénase que le nitrate.

Le gène *SAN1A*, d'abord isolé chez les nodules de racines sénescentes (Chan, 1995), ainsi que deux membres de la famille des gènes *SAN1* ont été caractérisés. Ces gènes représentent une famille nouvelle en soja. Cette famille de gènes est fortement conservée et répétée en tandem. *SAN1A* et *SAN1B* codent pour des protéines pourvues respectivement de 352 et 353 acides aminés, mais *SAN1C* code pour une protéine tronquée de 126 acides aminés, et il est possiblement un pseudogène. En dépit de leur présence en tandem, la conversion génique ne semble pas s'être produite dans l'histoire évolutive de ces gènes. Les modes d'expression de *SAN1A* et *SAN1B* pendant la sénescence naturelle suggèrent que *SAN1A* peut jouer un rôle dans tout le développment des nodules. *SAN1B* peut être plus important pendant la sénescence. Le nitrate et

l'obscurité ont causé une diminution de l'expression de *SAN1A* et une augmentation de l'expression de *SAN1B*. Leurs modes d'expression étaient différents dans plusieurs tissus de soja, bien que *SAN1A* et *SAN1B* aient été exprimés en tous les tissus examinés. Ces résultats suggèrent que *SAN1A* et *SAN1B* ont des fonctions divergentes, ou *SAN1B* peut remplacer *SAN1A* dans des conditions telles que la sénescence. Les résultats préliminaires suggèrent que *SAN1A* peut être réglé aussi par la lumière. Finalement, la mesure de l'expression de *SAN1C* dans des tissus de soja a prouvé qu'il est exprimé, malgré le fait qu'il code pour une protéine tronquée.

Les recherches dans des banques de données suggèrent que les gènes *SAN1* sont liés aux dioxygénases des plantes. Un arbre phylogénétique et une recherche avec BLASTP ont indiqué que les produits des gènes *SAN1* sont étroitement liés aux dioxygénases non caractérisées et identifiées chez des espèces comme *Arabidopsis*, le riz, et le pin maritime, avec les dioxygénases les plus étroitement apparentées étant les oxydases des gibbérelline. Ainsi, les gènes *SAN1* peuvent coder pour des paralogues des dioxygénases connues ou des nouvelles dioxygénases. Les résultats assemblés jusqu'ici suggèrent que les gènes *SAN1* codent pour la famille des dioxygénases ayant des rôles dans divers aspects du métabolisme des plantes. Une caractérisation plus détaillée de cette famille pourra fournir plus d'indices sur leur rôle(s).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

LIST OF ABBREVIATIONS

2-ODDs	-	2-oxoglutarate-dependent dioxygenases/2-oxoglutarate-Fe (II) oxygenases
A	-	adenine
aa	-	amino acids
$A_{260/280}$	-	absorbance at 260 or 280 nanometres
ACC	-	1-aminocyclopropane-1-carboxylic acid
ACCox	-	1-aminocyclopropane-1-carboxylic acid oxidase
At	-	*Arabidopsis thaliana*
ATP	-	adenosine triphosphate
AVG	-	aminoethoxyvinylglycine
bp	-	base pairs
Bcl	-	B cell lymphoma
BLAST	-	Basic Local Alignment Search Tool
BSA	-	bovine serum albumin
BY-2	-	bright yellow-2
C	-	cytosine
cDNA	-	complementary DNA
$Cn_1:n_2$	-	fatty acid chain of length n_1 with n_2 double bonds
C-terminus	-	carboxy terminus
cv	-	cultivar
DAPI	-	4',6'-diamidino-2-phenylindole
dATP	-	adenine
dCTP	-	cytosine

DDSA	-	dodecenylsuccinic anhydride
dGTP	-	guanine
DMP	-	2,4,6-tri[dimethylaminomethyl] phenol
DNA	-	deoxyribonucleic acid
***DN*ase**	-	deoxyribonuclease
dpi	-	days post-inoculation
DTT	-	dithiothreitol
dTTP	-	thymine
dUTP	-	deoxyuridine triphosphate
dXTP	-	deoxyribonucleotide
EDTA	-	ethylenediamine tetraacetic acid
EGTA	-	ethyleneglycol-*bis*(β-aminoethyl)-N, N, N′, N′-tetraacetic acid
EF hand	-	helix-loop-helix calcium binding domain
EM	-	electron microscopy
ER	-	endoplasmic reticulum
EST	-	expressed sequence tag
FAA	-	formaldehyde-acetic acid-alcohol (ethanol)
F3H	-	flavanone 3β-hydroxylase
FLS	-	flavonol synthase
FNS	-	flavone synthase
f. sp.	-	forma specialis
G	-	guanine
GA2Box	-	gibberellin 2β-hydroxylase
GA3ox	-	gibberellin 3β-hydroxylase

GA20ox	-	gibberellin 20-oxidase
GC	-	gas chromatography
GFP	-	green fluorescent protein
GITC	-	guanidinium isothiocyanate
G-protein	-	guanine nucleotide binding protein
GUS	-	β-glucuronidase
H6H	-	hyoscyamine 6β-hydroxylase
HPLC	-	high performance liquid chromatography
ICE	-	interleukin-1β-converting enzyme
IPNS	-	isopenicillin N synthase
kb	-	kilobases
kbp	-	kilobase pairs
kDa	-	kilodaltons
KOAc	-	potassium acetate
L.	-	classification according to Linnaeus
Lb	-	leghemoglobin
LB	-	Luria broth
LDOX	-	leucoanthocyanidin dioxygenase
LysM	-	lysin motif
M	-	molar or moles per litre
mamp	-	milliamperes
$Mg(OAc)_2$	-	magnesium acetate
mM	-	millimolar or millimoles per litre
MNA	-	methyl nadic anhydride

MOPS	-	3-(*N*-morpholino) propanesulfonic acid
mRNA	-	messenger RNA
Mts	-	megatonnes
ng	-	nanograms
NaOAc	-	sodium acetate
NaP_i	-	sodium phosphate
NHS-biotin	-	*N*-hydroxysuccinimidyl biotin
Nod	-	nodulation
NO_x	-	nitrogen oxides
N-terminus	-	amino terminus
OPA	-	One-Phor-All
ORF	-	open reading frame
^{32}P	-	radioactive isotope of phosphorus
PAR	-	photosynthetically active radiation
PCD	-	programmed cell death
PCR	-	polymerase chain reaction
PEG	-	polyethylene glycol
pg	-	picograms
Ps	-	*Pisum sativum*
pv.	-	pathovar
PVP	-	polyvinylpyrrolidone
RACE	-	rapid amplification of cDNA ends
RAP	-	RNA arbitrarily primed
RNA	-	ribonucleic acid

RNAi	-	RNA interference
***RN*ase**	-	ribonuclease
ROS	-	reactive oxygen species
RT	-	reverse transcriptase
SAG	-	senescence-associated gene
SAN	-	senescence-associated nodulin
SDS	-	sodium dodecyl sulphate
SSC (20X)	-	3 M NaCl, 0.3 M sodium citrate, pH 7.0
sp./spp.	-	species
sym	-	symbiosis
T	-	thymine
T4	-	bacteriophage
TAAB812	-	epoxy resin developed by TAAB Laboratories Equipment Ltd.
TA-cloning	-	3′ "A" overhangs left by polymerases such as *Taq* polymerase are ligated to vectors with complementary 3′ "T" overhangs
Taq	-	*Thermus aquaticus*
TBE (10X)	-	0.89 M Tris, 0.89 M Borate, 20 mM EDTA
TdT	-	terminal deoxytransferase
TE	-	tracheary element
TE	-	10 mM Tris, 1 mM EDTA
TEMED	-	N,N,N,N-tetramethylethylenediamine
TMV	-	tobacco mosaic virus
TUNEL	-	TdT-mediated dUTP-biotin nick end labeling
U	-	unit(s)

X-gal - 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

INTRODUCTION

This thesis describes the last stage of root nodule development, senescence. Preliminary light microscopy assessments of morphological changes occurring during nodule senescence are presented. In addition, the timing of senescence during natural nodule development and senescence induced by nitrate and darkness was assessed by monitoring acetylene reduction and leghemoglobin mRNA expression. The bulk of the work consists of the description of a novel gene family isolated from soybean, the *SAN1* genes. Gene structures and features are presented, along with information about their possible functions, as determined through the use of bioinformatics and phylogenetics. A study of the expression of two of the *SAN1* genes, *SAN1A* and *SAN1B*, was carried out. Expression patterns were determined throughout natural nodule development, during senescence induced by the application of fixed nitrogen or darkness, and in various soybean tissues. In addition, the results of a test for the diurnal variation of *SAN1A* are presented. *SAN1C* expression in various tissues of soybean was investigated.

CHAPTER 1

LITERATURE REVIEW

1.1 General Introduction

Nitrogen is an important element required for plant growth since it is a key component of proteins, chlorophyll, and nucleic acids (Crawford, 1995). The most prevalent form of nitrogen in soil is nitrate, which plants take up and convert to other nitrogen-containing compounds (Vance, 1997). However, plants can also take up nitrogen in the form of ammonium or urea.

Of the nitrogen available to the biosphere, 99.95% is found in the atmosphere, dissolved in bodies of water, and in soils and sediments (Smith and Gallon, 1993; Winstanley *et al.*, 2005). Most terrestrial plants only have access to nitrogen in the soil, but the soil is nitrogen-deficient in many environments. Plants that are able to obtain nitrogen through associations with soil microorganisms are at an advantage in such environments.

1.1.1 Nitrogen Fertilizer and Problems with its Use

Between 1960 and 2000, the world's population doubled to 6 billion people and projections suggest that the world's population will reach 8 billion by 2025, and perhaps 10 billion by 2040 (Dyson, 1999; Vance, 2001, and references therein). To keep pace with this growth, food production must continue to increase dramatically. Therefore, the use of nitrogen fertilizer must also increase. In 2000, 88 Mts of nitrogen were applied to crops worldwide. Applied nitrogen will reach amounts of 120 Mts per year between 2030 and 2040 if global usage continues at a similar rate (Frink *et al.*, 1999). Since the efficiency of nitrogen uptake by crops is only about 50%, much of the applied nitrogen

remains in the soil (Smil, 1999; Socolow, 1999). This remaining nitrogen is responsible for a number of environmental and health problems.

Microbial nitrification and denitrification of the soil is a major contributor of N_2O and NO_x gases (Socolow, 1999). These gases contribute to ozone depletion, the greenhouse effect, acid rain, and particulate air pollution. In addition, bodies of water containing high concentrations of nitrates due to surface runoff of fertilizer may become eutrophic and hypoxic, leading to aquatic "dead zones", such as the one found in the Gulf of Mexico (Galloway *et al.*, 1995; Socolow, 1999). Excess NO_3^- in drinking water has been linked to methemoglobinemia in small children, a dangerous condition whereby excess nitrites in the blood can oxidize the Fe^{2+} in hemoglobin to Fe^{3+}, which cannot carry oxygen (Galloway *et al.*, 1995; Smil, 1999; Socolow, 1999).

The production of nitrogen fertilizer by the Haber-Bosch process consumes extensive amounts of fossil fuel (Galloway *et al.*, 1995; Heichel, 1987). In the United States, nitrogen fertilizer production alone accounts for 3-5% of annual natural gas consumption. Fertilizer production is therefore very susceptible to changes in energy availability and prices increase with rising energy costs.

Furthermore, although used extensively in the developed world, nitrogen fertilizer is not readily available to subsistence farmers in developing countries due to its high cost, weak infrastructure, and poor transportation (Vance, 2001).

One strategy for decreasing global dependence on nitrogen fertilizer is a wider use of legumes. Understanding how legumes acquire an ability to fix nitrogen and how this ability is affected by various environmental cues is important in terms of human health, the environment, and agronomy (Puppo *et al.*, 2005).

1.1.2 The Importance and Potential of Legumes

Legumes are important for a number of reasons. Forage and grain legumes account for 27% of the world's primary crop production and grain legumes provide 33% of human dietary protein in the form of foods such as chickpeas, peanuts, tofu, soy milk, and miso (Graham and Vance, 2003). In addition, 35% of the world's processed vegetable oil is derived from legumes, chiefly from soybean and peanut (Graham and Vance, 2003; Vance *et al.*, 2000). Soybean and peanut are also important sources of feed protein for the chicken and pork industries. Some legumes are important sources of wood (Sprent and Parsons, 2000). For example, in Brazil, the tree *Melanoxylon brauna* (bara na) is often used for construction due to its durable and attractive wood. Legumes such as *Sesbania* spp. and *Phaseolus* spp. are also important for tree fallow and alley cropping systems that are applied in subsistence farming in Africa and Asia (Buresh and Cooper, 1999). In addition, legumes have industrial and medicinal uses. Biodegradable plastics, oils, gums, dyes, and inks can be produced from legumes (Morris, 1997; Paetau *et al.*, 1994). For example, galactomannan gums from *Cyamopsis* spp. and *Sesbania* spp. are used as thickeners and in pill formulations. Isoflavones from soybean and other legumes have been shown to lower cancer risk and blood cholesterol, and more recently, legume-derived phytoestrogens have been suggested as alternatives to traditional hormone replacement therapy in post-menopausal women (Kennedy, 1995; Molteni *et al.*, 1995; Wuttke *et al.*, 2002). Finally, legumes can decrease farmers' dependence on nitrogen fertilizer. Currently, agriculturally important legumes and legumes in natural ecosystems fix 40-60 Mts and 3-5 Mts of nitrogen per year, respectively (Graham and Vance, 2003; Smil, 1999). The use of legumes, in conjunction with proper soil testing and minimal fertilizer, could reduce global nitrogen fertilizer use by up to 20 Mts per year (Smil, 1999;

Socolow, 1999). Increased use of legumes in agriculture would also reduce leaching and volatization of excess nitrogen, since biologically fixed nitrogen is not as susceptible to these physical and chemical processes.

1.1.3 Occurrence of Nitrogen Fixation and Nitrogen-Fixing Symbioses

Microorganisms called diazotrophs can fix nitrogen and make it available to plants, either as free-living prokaryotes or in symbiosis with a plant partner (Smith and Gallon, 1993).

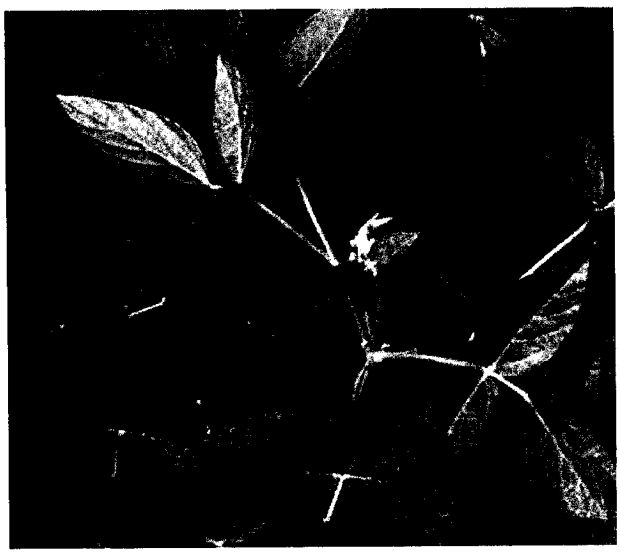
Free-living bacteria which fix nitrogen include *Clostridium* spp., *Azotobacter* spp., and cyanobacteria such as *Anabaena* spp. (Raven *et al.*, 1992). *Anabaena* spp. are often found in association with *Azolla* spp., floating water ferns common to rice paddies, and can contribute up to 50 kilograms of fixed nitrogen per hectare of rice.

Other diazotrophs, chiefly species of the Rhizobiaceae family, form symbiotic relationships with plants in which they provide fixed nitrogen in return for carbon (Van Rhijn and Vanderleyden, 1995). Each bacterial species forms symbioses with specific hosts, most often legumes, resulting in the formation of novel organs called nodules on the plant roots, where nitrogen fixation occurs (Mylona *et al.*, 1995). For example, the rhizobia *Bradyrhizobium japonicum* and *Sinorhizobium meliloti* form symbiotic relationships with *Glycine max* L. Merrill (soybean) (Figure 1A) and *Medicago sativa* (alfalfa), respectively (Cohn *et al.*, 1998). Interestingly, some legume-rhizobia symbioses result in stem nodules as well as root nodules. Stem nodules are found in *Sesbania rostrata*, a tropical legume native to West Africa, when in a symbiosis with *Azorhizobium caulinodans*, and also in several species of the tropical legume genus *Aeschynomene* when infected by various *Rhizobium* strains (Alazard, 1985; Ndoye *et al.*, 1994). In addition,

Figure 1: Legumes of Note

A. *Glycine max* L. Merrill, an important agronomic legume. B. *Lotus japonicus*, and C. *Medicago truncatula*, model legumes. Pictures were taken from http://stephenville.tamu.edu/~butler/foragesoftexas/species/soybean.html, soybean; http://www.h.chiba-u.jp/iden/; *Lotus japonicus*; http://wwwscieng.murdoch.edu.au/centres/others/ingorsi/network/mtrunc.html; *Medicago truncatula*.

A.



B.

C.

Bradyrhizobium sp. strain *Parasponia* forms nodules on the roots of non-legume trees of the genus *Parasponia*.

Species of the genus *Frankia*, a group of soil actinomycetes, can form root nodules on various actinorhizal plants such as *Alnus* spp. (Mylona *et al.*, 1995). However, unlike rhizobium-legume symbioses, the plants which form relationships with *Frankia* spp. are derived from diverse plant families, including the Betulaceae (*Alnus* spp.), the Casuarinaceae (*Casuarina* spp.), and the Rhamnaceae (*Ceanothus* spp.) (Benson and Silvester, 1993).

Since biological nitrogen fixation is controlled by genetics, improvement through breeding and biotechnology should be possible; however, this task may prove to be difficult, since a large number of plant and bacterial gene products are associated with the formation of a functional root nodule (Schultze and Kondorosi, 1998; Vance, 1997). The hope is that key insights into nitrogen fixation may be gained by research on nodulation, and its regulation, in the model legumes *Lotus japonicus* and *Medicago truncatula* (Figure 1B,C) (Cook, 1999; Handberg and Stougaard, 1992).

1.2 The Rhizobium-Legume Symbiosis

Root nodule development is a continuum of complex events, although it is arbitrarily divided into three stages - early, late, and senescent - to provide a convenient framework for study (Hirsch, 1992). The early stage consists of the events surrounding the formation of the nodule structure, the late stage is represented by a mature, nitrogen-fixing nodule, and the senescent stage is characterized by degeneration of the nodule.

1.2.1 The Early Stage of Nodule Development

1.2.1.1 Exchange of Signals between Legumes and Rhizobia

An exchange of signals between legumes and rhizobia initiates the process of nodule formation. Plant exudates of compounds such as amino acids, sugars, and flavonoids can attract rhizobia in the soil to the root by chemotaxis (Aguilar *et al.*, 1988; Barbour *et al.*, 1991; Gulash *et al.*, 1984). Once the rhizobia have gathered at the root's surface, legumes produce various types of compounds, usually flavonoids or isoflavonoids, to signal to the bacteria to prepare for nodulation (Hirsch *et al.*, 2001; Stougaard, 2000). Soybean produces the isoflavonoids daidzein and genistein (Figure 2A) to signal to *B. japonicum* (Mellor and Collinge, 1995).

Legume flavonoids are bound by constitutively expressed NODD proteins in the rhizobia (Mylona *et al.*, 1995). Since NODD proteins only recognize certain flavonoids, they contribute to the specificity of the legume-rhizobium relationship (Nap and Bisseling, 1990). *NODD* can be a single copy gene or part of a multigene family (Fisher and Long, 1992). For example, in *Azorhizobium caulinodans* ORS571, only one *NODD* locus has been found, whereas in *B. japonicum* USDA 191 and *S. meliloti*, two and three copies of *NODD* have been identified, respectively (Appelbaum *et al.*, 1988; Goethals *et al.*, 1990; Honma and Ausubel, 1987). When a flavonoid binds to NODD, a conformational change occurs that allows it to bind to elements in the promoters of other rhizobial *NOD* genes, activating their transcription (Hirsch *et al.*, 2001; Perret *et al.*, 2000). The products of these other *NOD* genes are the enzymes which catalyze the synthesis of Nod factors.

Nod factors are the signals that the bacteria return to the legume during the initiation of nodule formation (Downie and Walker, 1999; Mylona *et al.*, 1995; Stougaard, 2000). They are lipochitooligosaccharide (LCO) molecules which typically consist of a

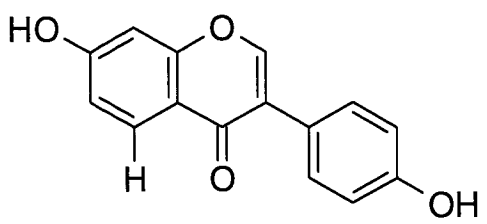
Figure 2: Communication between Legumes and Rhizobia Required for Nodule Formation

A. Daidzein and genistein are isoflavonoids produced by soybean (adapted from Hirsch *et al.*, 2001).

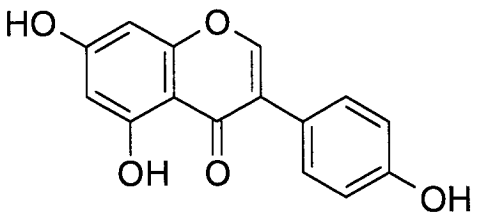
B. Structure of a typical Nod factor (from Mylona *et al.*, 1995). Nod factors consist of a backbone typically composed of three to five *N*-acetylglucosamine residues, decorated with different substituents depending on the rhizobium species. R_1 is typically a methyl group or unsubstituted. R_2 is a fatty acid moiety. R_3 is usually a carbamyl or acetyl group or the position can remain unsubstituted. R_4 varies greatly; it can be a sulfate, an acetyl, an arabinosyl, a fucosyl, or a methylfucosyl group, or the position may remain unsubstituted. R_5 is usually a glyceryl group or unsubstituted.

C. Summary of early communication between legumes and rhizobia (adapted from Stougaard, 2000). Nodulation occurs when a plant and rhizobium recognize one another's signal molecules. Nodulation occurs when a plant and rhizobium recognize one another's signal molecules. Nodules will only be formed when the plant's flavonoids are recognized by the rhizobial species and the rhizobium produces Nod factors recognizable to the legume. +++, normal nodulation; -, no nodulation.

A.

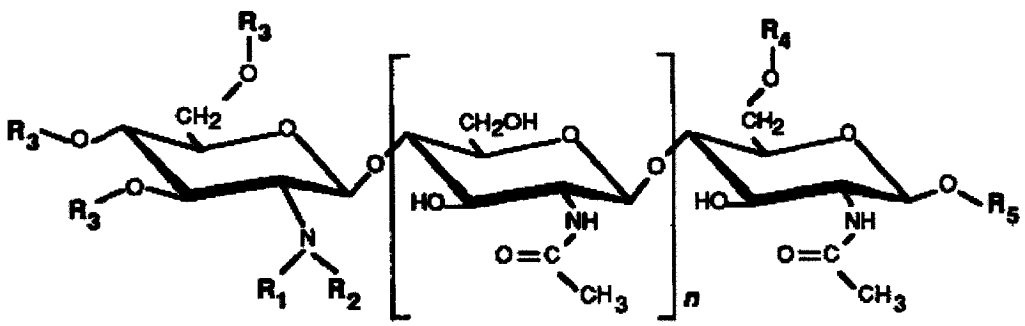


Daidzein

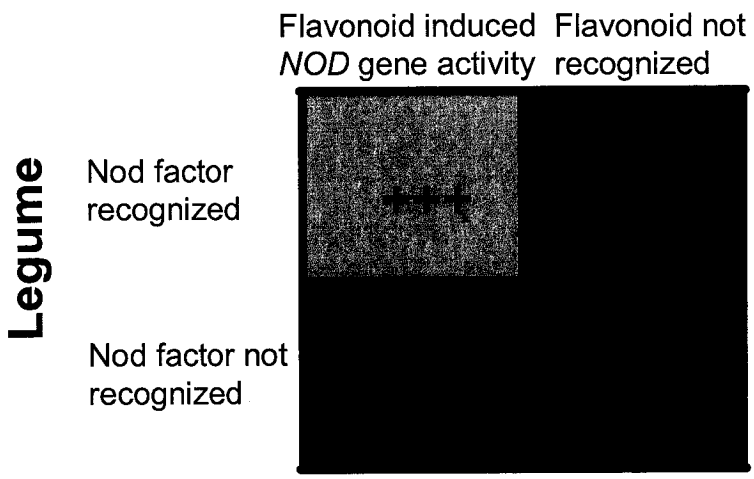


Genistein

B.



C.



three to five β-1,4-linked *N*-acetylglucosamine backbone with an added fatty acid group at the nonreducing end (Mylona *et al.*, 1995). Various other substitutions can occur on the sugar polymer, including sulfuryl, fucosyl, mannosyl, or arabinosyl groups at the reducing end and carbamoyl, glycerol, methyl, acetyl, or fucosyl additives at other positions (Stougaard, 2000). Figure 2B depicts the basic structure of a Nod factor. *B. japonicum* LCOs consist of two to three sugar moieties with a fatty acid (R_2) that is most often an unsaturated C18:1 group, but can also be a C16:0 or C16:1 group (Carlson *et al.*, 1993; Sanjuan *et al.*, 1992). The R_1 and R_5 positions are unsubstituted, R_3 is either unsubstituted or an acetyl group, and R_4 is a 2-*O*-methylfucosyl group. Nod factors with atypical lengths, different types of oligosaccharide backbones, or with novel substitutions at their terminal residues have also been described. For example, some of the Nod factors produced by *Rhizobium* sp. GRH2 have six sugar backbones, *Sinorhizobium fredii* USDA 191 synthesizes a pentameric Nod factor with a glucosyl group in place of the middle *N*-acetylglucosamine, and the reducing terminal residues of 3-*O*-S-2-*O*-methylfucosyl and 4-*O*-acetyl-2-*O*-methylfucosyl are unique to *Sinorhizobium* sp. NGR234 (Bec-Ferte *et al.*, 1996; Lopez-Lara *et al.*, 1995; Price *et al.*, 1992, 1996).

Rhizobial *NOD* genes are classified as either common or host-specific (Fisher and Long, 1992; Nap and Bisseling, 1990; Van Rhijn and Vanderleyden, 1995). The common *NOD* genes, *NODA*, *NODB*, and *NODC* are so called because each one of them is functionally interchangeable between many rhizobia (Brewin, 1991). These genes code for proteins involved in the construction of the *N*-acetylglucosamine backbone of the Nod factor. NODC is a β-glucosaminyl transferase that catalyzes the linkage of UDP-*N*-acetyl- glucosamine monomers into a chitooligosaccharide backbone (Albrecht *et al.*, 1999; Hirsch *et al.*, 2001). NODB is a chitooligosaccharide deacetylase which removes

the acetyl group from the nonreducing end of the sugar backbone to make way for the fatty acid substituent (Hirsch *et al.*, 2001; Schultze and Kondorosi, 1998). NODA, an acyltransferase, then transfers the fatty acid chain to the deacetylated nonreducing end of the sugar (Albrecht *et al.*, 1999; Cohn *et al.*, 1998; Hirsch *et al.*, 2001).

Host-specific *NOD* genes vary considerably between rhizobial species, depending on the substitutions that occur on the chitooligosaccharide backbones of the Nod factors that they produce (Fisher and Long, 1992; Nap and Bisseling, 1990; Van Rhijn and Vanderleyden, 1995). For example, the *NODF* and *NODE* genes of *S. meliloti* code for an acyl carrier protein and a β-ketoacyl synthase, respectively, which together synthesize the unsaturated fatty acid chains found on the terminal nonreducing end of the Nod factors (Demont *et al.*, 1993). *NODL*, *NODP*, *NODQ*, and *NODH* of *S. meliloti* encode homologues of an acetyl transferase, an ATP sulfurylase, an adenosine 5′ phosphosulfate kinase, and a sulfotransferase, respectively (Roche *et al.*, 1991; Spaink *et al.*, 1991). NODL acetylates the nonreducing end of the chitooligosaccharide, and NODQ, NODP, and NODH work in conjunction to bring about sulfation of the reducing end. The fucosyltransferase enzyme that carries out the addition of the 2-*O*-methylfucose group to the reducing end of *B. japonicum* Nod factors is encoded by *NODZ* (Quinto *et al.*, 1997).

The range of plants that a given rhizobium can form symbioses with is limited by its Nod factors, since legumes will only respond to the Nod factors of certain rhizobial species (Fisher and Long, 1992; Schultze and Kondorosi, 1998). This phenomenon, and the fact that rhizobial *NOD* gene transcription is only promoted when NODD recognizes a plant's flavonoids, provides the specificity of the legume-rhizobium relationship. Figure 2C summarizes the early communications that must occur between plants and rhizobia to ensure nodule formation.

1.2.1.2 Legume Responses to Nod Factors

Nod factors cause many responses in the host legume, some of which occur within seconds, and others which happen only after several days (Downie and Walker, 1999). Figure 3 summarizes the effects of Nod factors on legumes.

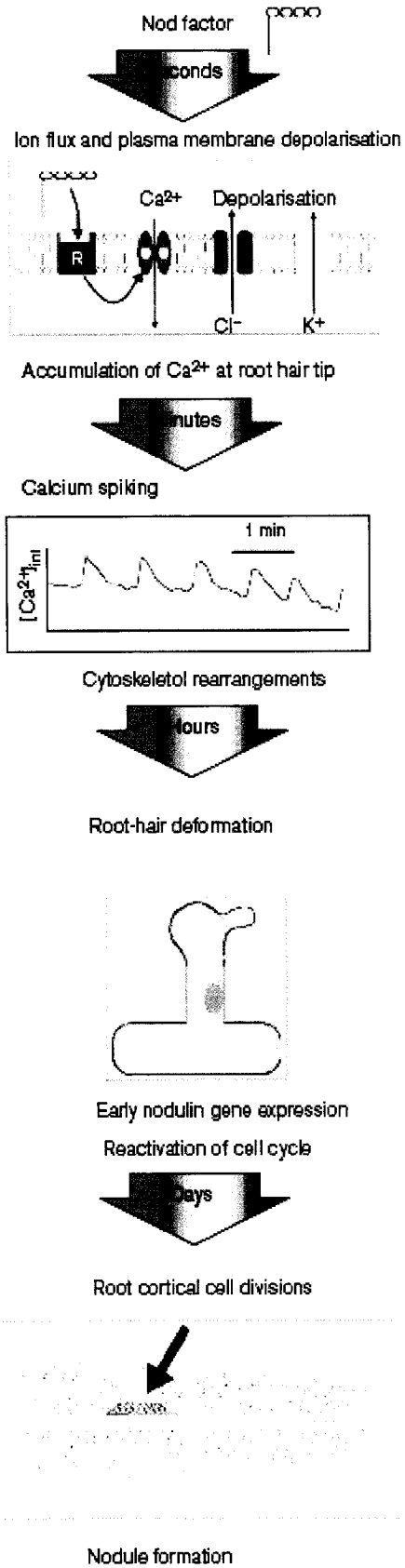
1.2.1.2.1 Ion Fluxes and Membrane Depolarization of Root Hair Cells

Early studies of alfalfa root hairs detected membrane depolarization in response to Nod factor application (Ehrhardt *et al.*, 1992; Felle *et al.*, 1995; Kurkidjian, 1995). Dissections of this response have determined which ion movements are responsible for this depolarization. Using non-invasive ion selective electrodes, Felle *et al.* (1998) showed that an influx of Ca^{2+} occurred seconds after Nod factor application to alfalfa root hairs perfused with buffer containing Ca^{2+} ions. Furthermore, they showed that a Cl^- efflux which occurred seconds after the Ca^{2+} influx led to the observed membrane depolarization, that the depolarization was counteracted by an efflux of K^+ ions, and that the membrane was repolarized through the action of H^+ pumps. In follow up studies, Felle *et al.* (1999a,b) demonstrated the influx of Ca^{2+} into the root hair cells of alfalfa directly by using invasive ion selective electrodes. This early influx of Ca^{2+} has also been detected in *Phaseolus vulgaris* and *M. truncatula* (Cardenas *et al.*, 1999; Shaw and Long, 2003).

The early Ca^{2+} influx response to Nod factors may be an early step in a G-protein signal transduction cascade. Studies by Pingret *et al.* (1998) showed that G-protein agonists were able to cause early Nod factor responses in the root hair cells of *M. truncatula*, but only if Ca^{2+} ions were included in the medium. In addition, bacterial pertussis toxin, a known G-protein antagonist, was able to abolish these responses. The importance of Ca^{2+} in the G-protein agonist-induced responses was highlighted by testing

Figure 3: Early Responses of Legumes to Nod Factors
(adapted from Downie and Walker, 1999)

The early responses to Nod factors are conserved among legumes. Within seconds of exposure to Nod factors, ion fluxes, membrane depolarization, and accumulation of Ca^{2+} at the tips of root hairs occurs. A G-protein-mediated signal transduction pathway may operate, but whether or not this is related to the initial ion fluxes or the downstream calcium spiking is unknown. Arrangements of the actin cytoskeleton occur minutes after exposure of the legume to Nod factors. Within hours, root hair deformation, the expression of early nodulins, and the reactivation of the cell cycle to initiate formation of nodule primordia occurs. Within days of Nod factor exposure, root cortical cell divisions lead to nodule formation. No cause and effect is implied by the order of events depicted in this figure.


Nod factor
conds
Ion flux and plasma membrane depolarisation
Ca2+
Depolarisation
R
Cl-
K+
Accumulation of Ca2+ at root hair tip
nutes
Calcium spiking
1 min
[Ca2+]cyt
Cytoskeletol rearrangements
lours
Root-hair deformation
Early nodulin gene expression
Reactivation of cell cycle
ays
Root cortical cell divisions
Nodule formation

various Ca^{2+} influx or release antagonists, including ruthenium red, EGTA, and La^{3+}. All of these antagonists blocked the early responses to Nod factors induced by G-protein agonists, indicating that both mobilization of intracellular Ca^{2+} stores and an influx of Ca^{2+} are required for these events. Antagonists of phospholipase C activity, an enzyme whose activation is mediated by G-proteins and which participates in initiating phosphorylation cascades in phosphoinositide second messenger pathways, also block the early Nod factor responses of root hair cells induced by G-protein agonists. Taken together, these results suggest that Nod factor signal transduction involves G-proteins and the activation of Ca^{2+} and phosphoinositide second messenger pathways.

1.2.1.2.2 Cytoplasmic Alkalinization of Root Hair Cells

Like membrane depolarization, the alkalinization of the root hair cytoplasm occurs seconds after Nod factor application (Downie and Walker, 1999). Felle *et al.* (1996) showed that when Nod factors were applied to alfalfa root hairs at a concentration of 10^{-7} M, an increase in intracellular pH by more than 0.2 units was observed within 15 seconds. This effect was concentration-dependent and persisted beyond the time of membrane repolarization until withdrawal of the Nod factor, indicating that this response may be independent of membrane depolarization. In alfalfa, sulfated Nod factors are required for nodulation to occur. Although all Nod factors tested were able to elicit intracellular alkalinization, the sulfated Nod factor caused the greatest increase in pH, suggesting that intracellular alkalinization may be one of the events involved in Nod factor signal transduction.

1.2.1.2.3 Calcium Spiking

Within minutes of Nod factor application, root hair cells display oscillations in cytosolic Ca^{2+} concentrations (Downie and Walker, 1999). Ca^{2+} oscillations or spikes

were first described by Ehrhardt *et al.* (1996). Approximately 9 minutes after the application of Nod factors to alfalfa root hairs that had been injected with a Ca^{2+}-binding dye, a regular oscillation of the Ca^{2+} with a mean period of 60 seconds was detected. The oscillations did not appear as sinusoidal waves, but as asymmetric spikes with the rising phase of the peak occurring faster than the declining phase. The response continued as long as experimental observations were made, up to 3 hours in some cases. It was observed that the calcium spikes were initiated in the perinuclear region, suggesting that the Ca^{2+} stores or channels that mediated the Ca^{2+} release were located there. When Nod factors from non-nodulating rhizobial species were applied to alfalfa, or Nod factors recognized by alfalfa were applied to root hairs of the non-legume tomato or a non-nodulating alfalfa mutant, no Ca^{2+} spiking occurred, indicating that the Ca^{2+} response was due to Nod factor recognition by the plant. Pea, *M. truncatula*, *Phaseolus* spp., *Vicia hirsuta*, *Vicia sativa*, and *L. japonicus* all exhibit Ca^{2+} oscillations in response to Nod factors (Cardenas *et al.*, 1999; Wais *et al.*, 2000; Walker *et al.*, 2000).

The role of oscillating Ca^{2+} concentrations in root hair cells responding to Nod factors remains obscure. In mammalian cells, Ca^{2+} spiking plays a role in developmental gene regulation (Meldolesi, 1998). In plants, Ca^{2+} oscillations are involved in regulating the opening of guard cell apertures in the stomata and have been associated with pollen tube growth (Franklin-Tong *et al.*, 1996; Staxen *et al.*, 1999). Ca^{2+} oscillations may play a role in a signalling pathway that leads to the expression of genes involved in nodulation.

1.2.1.2.4 Changes in Root Hair Morphology and Cytoskeletal Rearrangements

Minutes to hours after exposure to Nod factors, root hair deformation and curling occur (Downie and Walker, 1999). Rearrangements of the actin cytoskeleton appear to be

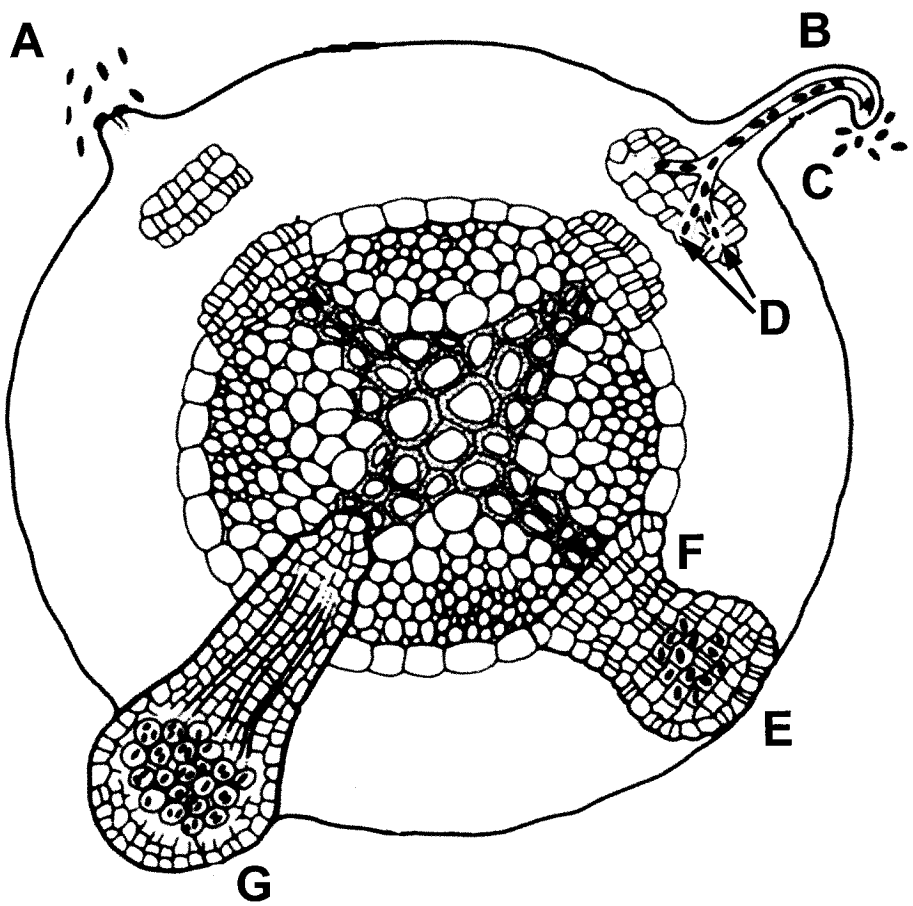
the underlying alterations that give rise to these events (D'Haeze and Holsters, 2002; Esseling and Emons, 2004).

Studies of root hair deformation in *V. sativa* indicate that it begins minutes after Nod factors are sensed by the plant and is characterized by an increase in actin filaments and gradual swelling at the tip of the root hair cell (Heidstra *et al.*, 1994). Approximately one hour later, an outgrowth of the root hair appears (De Ruijter *et al.*, 1998). At this time, the actin cytoskeleton resembles that of a growing root hair in that the filaments appear concentrated in the subapical region of the outgrowth (Miller *et al.*, 1999). It is thought that bundles of actin filaments promote polar growth by aligning and releasing Golgi vesicles into the root tip. The application of cytochalasin D, an actin-depolymerizing drug, to Nod factor-treated root hairs inhibited root hair elongation, indicating that properly aligned actin filaments were necessary for the polar growth of the root hair.

Root hair curling occurs after root hair deformation, within hours of the initial perception of Nod factors, eventually giving rise to curls called shepherd's crooks in which the rhizobia are trapped (Figure 4B,C) (Downie and Walker, 1999; Nap and Bisseling, 1990). Initially, it was thought that root hair curling required the presence of rhizobia (Catoira *et al.*, 2001), but a new assay developed by Esseling *et al.* (2003) has shown that Nod factors can induce root hair curling on their own. A small droplet of Nod factors was applied to the side of a tip of a growing root hair of *M. truncatula*. An immediate reorientation of the root hair axis of growth towards the site of application was observed and was most pronounced after 30 minutes. Even if Nod factors were only applied once at a given spot on a root hair, growth continued towards the site of application and the root hair remained oriented in that direction after growth ceased.

Figure 4: Formation of a Nodule (adapted from Taiz and Zeiger, 1991)

The steps leading to the production of a functional nodule are depicted from the perspective of a root hair cross-section. A. Rhizobia are attaching to the root hair and releasing Nod factors (in response to flavonoid compounds released by the plant host) which will stimulate cortical cell divisions. B. Nod factors have induced root hair curling or formation of "shepherd's crooks". C. Rhizobia are being trapped on the surface of the root hair. D. Infection threads are formed, allowing the bacteria to enter the cortical cells. E. Rhizobia trapped in cortical cells are dividing and differentiating in preparation for nitrogen fixation. F. Nodule formation continues by repeated cortical cell divisions. G. Emerging nodule structure.


A
B
C
D
F
E
G

Furthermore, applications of Nod factors at different spots on a single root hair resulted in reorientations in several directions at once. If several droplets of Nod factors were applied to the same side of the root hair, curling occurred, resulting in a 180° curl in some cases. These results suggest that Nod factors alone are responsible for setting up a new growth axis in the root hair tip and that root hair curling occurs due to a gradual shift of growth towards this axis.

1.2.1.2.5 Infection Thread and Nodule Primordium Formation

Once root hair curling has taken place and rhizobia are trapped within the shepherd's crooks, tubular structures called infection threads begin to form (Figure 4D) (Mylona *et al.*, 1995). Prior to their appearance, cytoplasmic bridges are observed in the outer cortical cells in line with the radial walls of the cortical cells that will eventually divide to form the nodule primordium (D'Haeze and Holsters, 2002). These bridges form preinfection threads through which the actual infection threads will eventually grow. Nod factors are sufficient to form the preinfection threads, but rhizobia must be present for subsequent infection thread formation. At the site of bacterial attachment to the root hair, hydrolysis of the plant cell wall occurs, the plasma membrane of the root hair cell invaginates, and new cell wall material is deposited around the bacteria, initiating a new infection thread (Hirsch, 1992; Mylona *et al.*, 1995). The infection thread proceeds through the cortex of the root hair and eventually branches, allowing the bacteria to travel into and between cells. Since the bacteria never make direct contact with the plant cells, and because the composition of the infection thread wall is very similar to those of the surrounding plant cells, elicitation of plant defence responses is avoided.

During infection thread formation, cortical cells of the root hair are mitotically reactivated to form the nodule primordium (Figure 4E,F) (Hirsch, 1992; Mylona *et al.*,

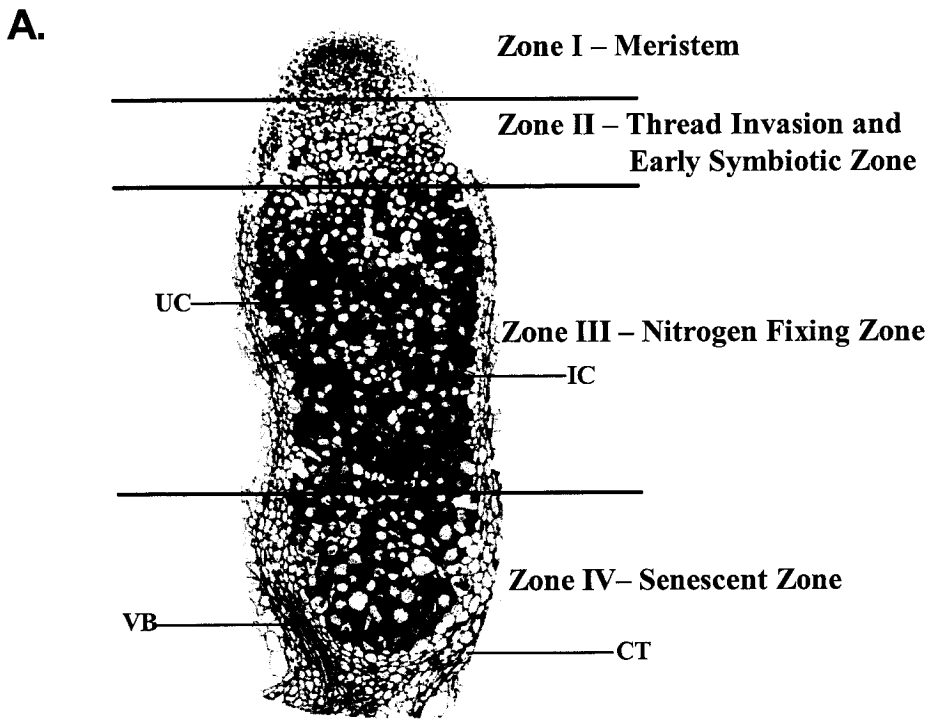
1995). At this point, a distinction between two different types of nodules can be made. Some legumes, such as pea and *M. truncatula*, produce indeterminate nodules, whereas other legumes, such as soybean and *L. japonicus*, produce determinate nodules. The location of the mitotically reactivated cortical cells determines which type of nodule is formed. Indeterminate nodules originate from inner cortical cell divisions, whereas determinate nodules arise from outer cortical cell divisions (Hirsch, 1992). In Figure 4 (E and F), cell divisions are taking place in the outer layers of the cortex indicating formation of a determinate nodule. Indeterminate nodules (Figure 5A) have a persistent meristem; therefore, the nodule continually grows in length, resulting in an elongated shape. All stages of nodule development are represented within a single nodule due to an age gradient that extends from the meristem to the nodule's point of attachment to the plant root. Four zones can be distinguished – the thread invasion zone, the early symbiotic zone, the nitrogen fixing zone, and the senescent zone. By comparison, determinate nodules (Figure 5B) lack a persistent meristem, and therefore appear spherical. Cell division stops early in development and any subsequent increases in size are due to cell enlargement. As a result, there is no age gradient in determinate nodules.

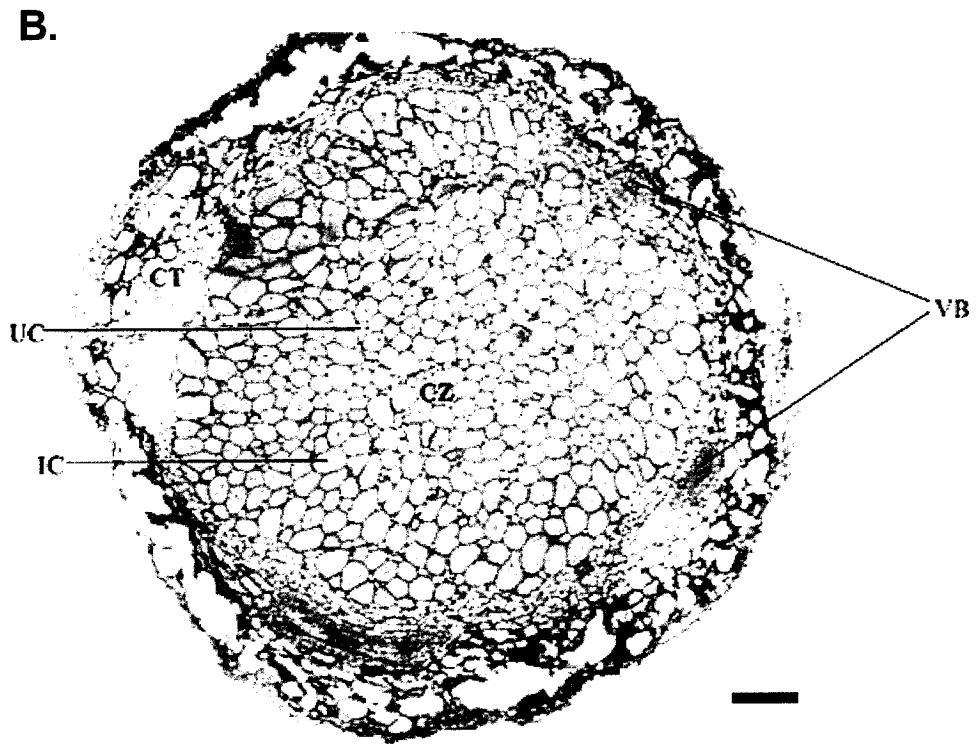
The mechanism that determines which cortical cells divide during nodule formation is not known, but the cells are mitotically reactivated before the infection threads reach them, indicating that Nod factors act at a distance (Mylona *et al.*, 1995; Schultze and Kondorosi, 1998). Since both cytokinins and inhibitors of polar auxin transport can induce the formation of nodule-like structures, it is possible that Nod factors reactivate cell division by disrupting the auxin-cytokinin balance (Cooper and Long, 1994; Hirsch *et al.*, 1989). In addition, the cortical cells which give rise to nodule primordia are located opposite the protoxylem poles (Mylona *et al.*, 1995; Schultze and

Figure 5: Structures of Indeterminate and Determinate Nodules

A. Light micrograph of a cross-section of an indeterminate, alfalfa nodule (adapted from Hirsch, 1992). The shape of the nodule is elongate due to the persistence of the meristem. Indeterminate nodules can be divided into four sections. Zone I is the meristem, where cell division occurs. Zone II is the thread invasion and early symbiotic zone, where infection threads invade the cortical cells and release the bacteria to differentiate into nitrogen fixing bacteroids. Zone III contains cells that are full of nitrogen fixing bacteroids. Zone IV contains cells in which nitrogen fixation is in decline.

B. Light micrograph of a cross-section of a determinate, soybean nodule (C. Webb, unpublished data). The section is stained with toluidine blue. Determinate nodules are round since cell division ceases once the organ is formed. Nitrogen fixation takes place in infected cells of the central zone. Bar, 50 μm. CT, cortex; CZ, central zone; IC, infected cell; UC, uninfected cell; VB, vascular bundle.


A.
Zone I – Meristem
Zone II – Thread Invasion and
Early Symbiotic Zone
UC
Zone III – Nitrogen Fixing Zone
IC
Zone IV– Senescent Zone
VB
CT


B.
CT
UC
CZ
IC
VB

Kondorosi, 1998). It is now known that the release of uridine from these poles confers susceptibility to mitosis on these cells (Smit *et al.*, 1995). Nod factors may initiate the release of uridine. Ethylene may also play a role in positioning of nodule primordia by telling which cells not to divide (Cohn *et al.*, 1998; Geurts and Bisseling, 2002). Transcripts of ACC oxidase, the enzyme catalyzing the last step in ethylene biosynthesis, have been shown to accumulate in cortical cells opposite the phloem poles, and treating roots with ethylene inhibitors allows nodule primordia to develop opposite the phloem poles (Geurts and Bisseling, 2002; Heidstra *et al.*, 1997b). Nod factors may induce a local increase in ethylene production to prevent the division of these cells.

1.2.1.2.6 The Expression of Early Nodulins (*ENOD*s)

Together with morphological and physiological changes, plant gene expression is an early response to bacterial Nod factors. Genes that are only expressed in nodules are called nodulins; those that are expressed within hours of Nod factor exposure are termed early nodulins or *ENOD*s (Cohn *et al.*, 1998; Hirsch *et al.*, 2001).

Many of the *ENOD*s that have been isolated appear to encode cell membrane or wall proteins, which highlights the importance of changes to these structures during nodule development (Schultze and Kondorosi, 1998). A well characterized *ENOD* is *PsENOD12*, which was isolated from a pea nodule cDNA library (Scheres *et al.*, 1990a). PsENOD12 is a hydroxyproline-rich glycoprotein, which suggests it may function as a cell wall protein. Furthermore, the *PsENOD12* transcript was localized to cells through which infection threads migrate, nodule primordia, and the invasion zone of mature pea nodules, suggesting that ENOD12 could be a component of infection thread walls.

Other nodulins encoding putative cell membrane and wall proteins have also been cloned, for example *ENOD2*, *ENOD5*, *RIP1*, and *ENOD12* genes from other species

(Schultze and Kondorosi, 1998). ENOD2 is thought to play a role in producing the oxygen barrier (see Section 1.2.3.2) of functional nodules (Van de Wiel *et al.*, 1990). ENOD5 is related to arabinogalactan proteins and may be a component of nodule membranes (Nap and Bisseling, 1990; Scheres *et al.*, 1990b). *RIP1* (*Rhizobium-Induced Peroxidase 1*) encodes a peroxidase that is expressed in nodule primordia (Cook *et al.*, 1995; Peng *et al.*, 1996). Since peroxidases are plant cell wall enzymes involved in lignin synthesis, RIP1 may have a role in cell wall lignification during nodule formation (Taiz and Zeiger, 1991).

An intriguing early nodulin is *ENOD40*. The *ENOD40* transcript appears in dividing cortical cells across from the protoxylem poles during the formation of nodule primordia (Charon *et al.*, 1997; Sousa *et al.*, 2001). Studies of transgenic *M. truncatula* plants overexpressing *ENOD40* suggested that its product might be responsible for regulating the dedifferentiation of cortical cells and could be a limiting factor in nodule primordia formation (Charon *et al.*, 1997, 1999). Although the *ENOD40* message is approximately 0.7 kb, only a small peptide of 12 to 13 amino acids was predicted by examination of its ORFs (Charon *et al.*, 1997). However, two peptides of 13 and 27 amino acids, translated from two overlapping ORFs in a polycistronic mRNA, are now known to be required for ENOD40's biological activity (Rohrig *et al.*, 2002, 2004; Sousa *et al.*, 2001). Work to elucidate the function of these two peptides has been carried out in soybean (Rohrig *et al.*, 2002, 2004). Both peptides bind to a 93 kDa subunit of sucrose synthase, suggesting that they could have a role in controlling sucrose use in nodules. In addition, computer analysis of the *ENOD40* mRNA has indicated that it can form stable secondary structures, a feature shared with biologically active RNAs (Girard *et al.*, 2003). Five domains of secondary structure are shared between all leguminous *ENOD40*

sequences that have been examined. This level of conservation between different species suggests that these domains may have a role in the function of the RNA. Chemical and enzymatic probing of the domains supports the proposed structure of domains 1, 3, and 5, and partially supports the proposed structure of domains 2 and 6. More recent work by Campalans *et al.* (2004) has shown that a novel RNA binding protein interacts with *ENOD40* mRNA in the nucleus and that this complex is then exported into the cytoplasm. Although the results of the *ENOD40* studies are interesting, how the peptides and/or the mRNA operate in nodule formation is not clear. Furthermore, *ENOD40* is not limited to legumes since homologues have been found in other dicots such as *Arabidopsis* and tomato, and in monocots including ryegrass, corn, and rice (Compaan *et al.*, 2003; Guzzo *et al.*, 2005; Kouchi *et al.*, 1999; Larsen, 2003; Vleghels *et al.*, 2003). The widespread occurrence of *ENOD40* in non-legumes indicates that it also plays a role in plant processes other than nodulation.

Many other *ENOD*s have been isolated, including genes encoding chitinases, lectins, carbonic anhydrases, pathogenesis-related proteins, and enzymes of the phenylpropanoid pathway, such as chalcone synthase and phenylalanine ammonia lyase (Estabrook and Sengupta-Gopalan, 1991; Goormachtig *et al.*, 1998; Schultze and Kondorosi, 1998). Overall, *ENOD*s are well conserved between legume species. For example, *ENOD12* genes have been found in broadbean, alfalfa, pea, spring vetch, and *M. truncatula* (Allison *et al.*, 1993; Perlick and Puhler, 1993; Pichon *et al.*, 1992; Scheres *et al.*, 1990a; Vijn *et al.*, 1995).

1.2.1.3 The Use of Mutants to Tease Apart Early Signalling during Nodule Formation

The use of plants mutated in early responses of legumes to Nod factors allows these events to be placed in order of their occurrence. In addition, cloning of the genes responsible for these mutations provides insight into the proteins responsible for nodule development. Naturally occurring mutants in agronomically important legumes such as pea, soybean, and alfalfa have been isolated (Esseling and Emons, 2004). However, due to the large genome sizes of pea and soybean, the polyploidy of soybean and alfalfa, and the lack of efficient transformation procedures in pea and soybean, it is difficult to analyze these mutants. Therefore, legume researchers rely on mutants of the more easily manipulated model legumes, *M. truncatula* and *L. japonicus*.

Most of the mutants that have been described are non-nodulating or ineffectively nodulating plants (Stougaard, 2001). Non-nodulating mutants may be blocked at different stages. For example, the *L. japonicus nin* mutant is blocked in inner cortical cell divisions, but root hair curling still occurs, whereas the *dmi1* and *dmi2* mutants of *M. truncatula* are blocked in nearly all early Nod factor-induced responses (Ane *et al.*, 2004; Schauser *et al.*, 1999; Wais *et al.*, 2000). The responses that are blocked have allowed the gene products corresponding to the mutated loci to be placed in order of their time of appearance. Figure 6 illustrates the order of action of selected gene products during nodulation relative to Ca^{2+} spiking.

The three *dmi* (*does not make infection*) mutants of *M. truncatula* are well studied non-nodulating mutants (Esseling and Emons, 2004; Wais *et al.*, 2000). *Dmi1* and *dmi2* mutants do not exhibit any responses to Nod factors from calcium spiking onward. *Dmi3* mutants show the same lack of responses, except that calcium spiking is maintained. *DMI1* encodes a predicted protein containing four transmembrane domains, two of which contain leucine zipper motifs and a proline-rich domain that could be

Figure 6: Gene Products Ordered Relative to Ca^{2+} Spiking
(adapted from Kistner and Parniske, 2002, Stougaard, 2001, and Wais *et al.*, 2000)

The action of the products of genes isolated from mutants can be placed in order based on where the nodulation program is blocked in the mutant plants. In this example, mutants have been ordered with respect to Ca^{2+} spiking. DMI, Does not Make Infection; HCL, Hair Curling; NFP, Nod Factor Perception; NIN, Nodule Inception; NSP, Nodulation Signalling Pathway.

Nod factors
NFP
DMI1
DMI2
CASTOR
POLLUX
Ca2+ spiking
NIN
HCL
NSP1
NSP2
DMI3
Infection thread formation, cell division, nodule organogenesis

involved in facilitating protein-protein interactions (Ane *et al.*, 2004). In addition, the central portion of DMI1 shares homology with multimeric, ligand-gated potassium channels. DMI1 may be a membrane spanning protein that participates in the formation of a receptor complex or channel. Homologues of *DMI1* are present in other legumes, including *M. sativa*, *L. japonicus*, and *Vigna*, and database searches predict similar genes in 28 monocot and dicot species. *DMI2* is the *M. truncatula* homologue of *NORK* (see Section 1.2.1.4.2.1) (Endre *et al.*, 2002). Since calcium spiking is retained in *dmi3* mutants, the *DMI3* gene product acts downstream of this event (Levy *et al.*, 2004; Mitra *et al.*, 2004). *DMI3* encodes a predicted protein with an N-terminal serine-threonine kinase domain flanked by a calmodulin-binding domain, with three calcium-binding EF hand domains towards its C-terminus (Levy *et al.*, 2004). *DMI3* is homologous to calcium and calmodulin-dependent protein kinases (CCaMKs) from trumpet lily and tobacco, and to putative CCaMKs from rice, apple, and the moss *Physcomitrella patens* (Mitra *et al.*, 2004). DMI3 may recognize and decode the calcium oscillations that occur during early stages of nodule formation (Ehrhardt *et al.*, 1996; Levy *et al.*, 2004).

Other examples of non-nodulating mutants include the *hcl* (*hair curling*), the *nfp* (*nod factor perception*), and the *nsp* (*nodulation signalling pathway*) mutants of *M. truncatula* (Ben Amor *et al.*, 2003; Catoira *et al.*, 2001; Oldroyd and Long, 2003).

The *NIN* (*Nodule Inception*) gene of *L. japonicus* encodes a transcription factor and was the first such gene involved in nodulation to be isolated (Schauser *et al.*, 1999). *Nin* mutant plants are chlorotic and do not form nodules, but resume normal growth once transferred to a nitrogen-containing medium, suggesting that the *nin* mutation is nodule-specific. *Nin* mutants also display normal early responses to Nod factors, but are blocked before the initiation of infection threads and cortical cell division. The *NIN* gene encodes

a putative protein which contains several domains similar to those of transcription factors. Interestingly, *NIN* and related *Arabidopsis* sequences share the consensus motif RWPXRK with *Chlamydomonas* minus dominance proteins (Schauser *et al.*, 1999, and references therein). This shared motif may have functional significance since the minus dominance proteins are developmental regulators of gametogenesis which, like nodulation, is induced by nitrogen limitation. NIN is likely to be a transcriptional regulator of genes involved in nodule development.

The *nsp1* and *nsp2* non-nodulating mutants of *M. truncatula* have recently been cloned (Kalo *et al.*, 2005; Smit *et al.*, 2005). *NSP1* and *NSP2* mutants exhibit calcium spiking, but not downstream Nod factor responses such as *ENOD* gene expression. Like *NIN*, both of these genes are thought to encode transcription factors. It is known that both legume-rhizobium and mycorrhizal symbioses share components of an early signalling pathway (Albrecht *et al.*, 1999). For example, the *DMI1*, *DMI2*, and *DMI3* genes of *M. truncatula* are required for the successful formation of both symbioses (Catoira *et al.*, 2000). *NSP1* and *NSP2*, however, are only required for formation of a successful legume-rhizobium symbiosis and are thought to represent the first step of the nodulation-specific signalling pathway, acting downstream of the CCaMK encoded by *DMI3*. *NSP1* and *NSP2* encode proteins with similarity to the GRAS family of putative transcription factors, which are known to be involved in other developmental processes in plants, such as gibberellin and phytochrome signalling and root development (Bolle, 2004; Kalo *et al.*, 2005; Smit *et al.*, 2005). Like other members of the GRAS family, NSP2 has a variable N-terminus containing homopolymeric stretches of amino acids characteristic of the activation domains of transcription factors, followed by a GRAS domain consisting of two leucine-rich regions and other conserved motifs of unknown function (Kalo *et al.*,

2005). *NSP1* also encodes a GRAS family protein, but its sequence differs greatly from NSP2 at the amino acid level, being merely 17% identical and 32% similar to NSP2 (Smit *et al.*, 2005). Expression studies of the two genes indicate that both *NSP1* and *NSP2* transcripts are found in roots; however, whereas inoculation of *M. truncatula* roots with *S. meliloti* causes increased expression of *NSP2*, it does not affect the expression level of *NSP1* (Kalo *et al.*, 2005; Smit *et al.*, 2005). *NSP1*- and *NSP2-GFP* constructs were prepared to study the localization of the proteins. Prior to inoculation with *S. meliloti* or Nod factor, NSP2-GFP was chiefly observed in the nuclear envelope and to a lesser extent, in the endoplasmic reticulum (Kalo *et al.*, 2005). After Nod factor application, the localization of NSP2-GFP diffused throughout the nucleus. If NSP2 acts as a transcription factor, this shift may be important to the regulation of its activity. NSP1-GFP was located in the nucleus, but its localization was unaffected by Nod factor application (Smit *et al.*, 2005). In addition, DMI3 was also localized to the nucleus (Kalo *et al.*, 2005). Since NSP1, NSP2, and DMI3 are all located in the nucleus, and DMI3 has been shown to act directly upstream of NSP1 and NSP2, it is possible that NSP1 and NSP2 are direct targets of the CCaMK encoded by *DMI3*, which may phosphorylate the NSPs allowing them to promote transcription of their targets (Kalo *et al.*, 2003; Smit *et al.*, 2005; Udvardi and Scheible, 2005).

Two recently discovered *L. japonicus* mutants, *castor* and *pollux*, are also blocked in the formation of nodules (Imaizumi-Anraku *et al.*, 2005). *CASTOR* and *POLLUX* encode plastid-localized proteins that act early in the signal transduction pathway shared between mycorrhizal and legume-rhizobium symbioses, between Nod factor perception and calcium spiking. Since the two proteins cannot replace each other functionally despite high sequence similarity, they may act in concert, possibly as a heteromeric

complex. Database searches have shown that these proteins could form structures similar to calcium-gated potassium channels; therefore, it is possible that CASTOR and POLLUX mediate ion fluxes between the plastids and the cytosol that are required for the initiation of calcium oscillations during early responses to Nod factors.

1.2.1.4 The Search for a Nod Factor Receptor

Nod factors act at very low concentrations with high specificity, implying that plant receptors mediate their perception (Cohn *et al.*, 1998; Heidstra and Bisseling, 1996). The search for these receptors has been a dominant theme of nodulation research (Parniske and Downie, 2003).

1.2.1.4.1 One Receptor or Two?

Studies to determine which structural features of Nod factors are necessary to elicit the early responses of legume root hairs led to the idea that two Nod factor perception events occur (Heidstra and Bisseling, 1996). Ardourel *et al.* (1994) proposed a two receptor model of Nod factor recognition. In this model, the first receptor would be a low stringency or "signalling" receptor to mediate events such as root hair deformation, whereas the second would be a high affinity entry or "uptake" receptor required for allowing bacterial entry into the plant. The other possibility is a single receptor model (Heidstra and Bisseling, 1996). In this case, one receptor with different binding affinities would be responsible for mediating all of the responses. Low affinity binding would be responsible for inducing early nodulation events such as root hair deformation; high affinity binding would result in bacterial entry and the full range of responses. Most research on early responses to Nod factors has supported the two receptor model.

1.2.1.4.2 Nod Factor Receptor Candidates

Early attempts at discovering Nod factor receptors led to the consideration of various candidates, including Nod factor binding proteins, carbohydrate-binding proteins called lectins, and chitinase-like proteins (D'Haeze and Holsters, 2002; Etzler *et al.*, 1999; Goormachtig *et al.*, 2001; Kim *et al.*, 2000; Minic *et al.*, 2000). However, recently isolated receptor-like kinases are believed to be the most likely candidates.

1.2.1.4.2.1 Receptor-Like Kinases as Nod Factor Receptors

A new class of receptor kinases has been identified by cloning of Nod factor perception mutants (Parniske and Downie, 2003). The discovery of these kinases, which may be the elusive Nod factor receptors, has led to the current model of Nod factor receptor operation.

Endre *et al.* (2002) isolated a receptor kinase gene from an alfalfa non-nodulation mutant. Although resistant to nodulation, the mutant was able to form spontaneous nodules, indicating that at least some of its nodule initiation and development pathway was intact. The gene giving rise to this mutation was called the *NOdulation Receptor Kinase* or *NORK*. By synteny, the *NORK* locus was also cloned from *M. truncatula* and putative *NORK* genes were sequenced from the related legumes *Melilotus alba*, *Pisum sativum*, *V. hirsuta*, and *L. japonicus*. The NORKs share motifs reminiscent of receptors, including an N-terminal signal motif, a putative extracellular domain containing three leucine-rich repeat motifs, a transmembrane segment, and an intracellular domain with conserved elements of serine-threonine protein kinases. Since NORK is a necessary component of the nodulation signalling pathway and because it is confined to the legume family, NORK is a candidate for a Nod factor receptor. When the *NORK* results of Endre *et al.* (2002) were presented, Stracke *et al.* (2002) simultaneously reported the cloning of *SYMRK* or *SyMbiosis Receptor-like Kinase* from *L. japonicus*. *SYMRK* is the *Lotus*

orthologue of *NORK* (Szczyglowski and Amyot, 2003). Root hair deformation was increased in the *Lotus SYMRK* mutants, indicating that its protein might play a role in the negative regulation of Nod factor responses (Stracke *et al.*, 2002). In addition, expression of the nodulin leghemoglobin (Lb) was abolished. These results suggested that NORK/SYMRK has a role in signal transduction between Nod factor perception and Lb gene expression. Orthologues of *NORK/SYMRK* are the *SYM19* locus of pea and the *DMI2* locus of *M. truncatula*. Since the *SYM19* gene of pea is required for calcium spiking, NORK/SYMRK must operate between the recognition of Nod factors and the activation of calcium spiking and Lb gene expression (Stracke *et al.*, 2002; Walker *et al.*, 2000). NORK/SYMRK is thought to be a secondary receptor involved in the early stages of the nodulation program.

The identification of a novel class of receptor-like kinases with LysM motifs has added to the knowledge of how Nod factors are perceived (Limpens *et al.*, 2003; Madsen *et al.*, 2003; Radutoiu *et al.*, 2003). Two early nodulation mutants of *L. japonicus*, *nfr1* (*nod factor receptor 1*) and *nfr5*, were isolated due to their inability to display root hair deformation after treatment with *Mesorhizobium loti* or purified Nod factors (Radutoiu *et al.*, 2003). *Nfr1/symrk* and *nfr5/symrk* double mutants indicated that their gene products act upstream of SYMRK. When the early nodulin genes *NIN* and *ENOD2* were monitored in Nod factor- or rhizobia- treated roots of *nfr1* and *nfr5* mutants, low levels of their transcripts, similar to those observed in untreated wild type roots, were detected. This finding indicated that functional *NFR1* and *NFR5* genes are required for the activation of *NIN* and *ENOD2*. *NFR1* and *NFR5* have been cloned by map-based cloning (Madsen *et al.*, 2003; Radutoiu *et al.*, 2003). *NFR1* produces two alternatively spliced transcripts which give rise to two predicted proteins containing signal peptides and two

LysM motifs at the N-termini, transmembrane segments, and C-termini containing serine-threonine kinase motifs; the presence of these motifs suggest that the proteins could act as receptors with extracellular LysM and intracellular kinase domains. The *NFR5* gene encodes a protein with an N-terminal signal peptide followed by three LysM motifs, a transmembrane domain, and a C-terminal serine-threonine kinase domain (Madsen *et al.*, 2003). However, an aspartic acid residue normally conserved in domain VII of serine-threonine kinases is missing and domain VIII, which comprises the activation loop of the kinase, is highly divergent or absent. Normally, this activation loop must be phosphorylated to make the active site of the kinase available for substrates. NFR5 may only be able to function if it acts in conjunction with another protein. Studies of LysM motifs in *Escherichia coli* proteins suggest that they are able to bind *N*-acetylglucosamine molecules. Therefore, NFR1 and NFR5 have the potential to bind Nod factors through their LysM domains.

Two other LysM receptor-like kinases (*LYK*s) that are orthologues of the $SYM2^A$ locus in pea have been identified in *M. truncatula* (Geurts *et al.*, 1997; Limpens *et al.*, 2003). In the $sym2^A$ line of pea, root hair deformation and curling occur to entrap rhizobia, but if a particular Nod factor substituent is absent, the process of nodulation is halted after the initiation of infection threads. A similar phenotype was observed when *S. meliloti nodF/nodL* mutants were used to infect *M. truncatula* root hairs in which homologues of $SYM2^A$ had been silenced by RNA interference (Limpens *et al.*, 2003). These mutants also resulted in infection thread abortion. The $SYM2^A$ gene and its *M. truncatula* orthologues are thought to be good candidates for Nod factor entry or uptake receptor genes. The *M. truncatula* genes, *LYK3* and *LYK4*, are receptor-like kinase genes with LysM domains. Since RNAi knockdowns of *LYK3* and *LYK4* show reduced

infection thread formation but still retain earlier responses to Nod factors, support is provided for the two receptor model of Nod factor perception and LYK3 and LYK4 could be bacterial entry receptors.

The current model of how Nod factor perception is thought to occur is outlined in Figure 7.

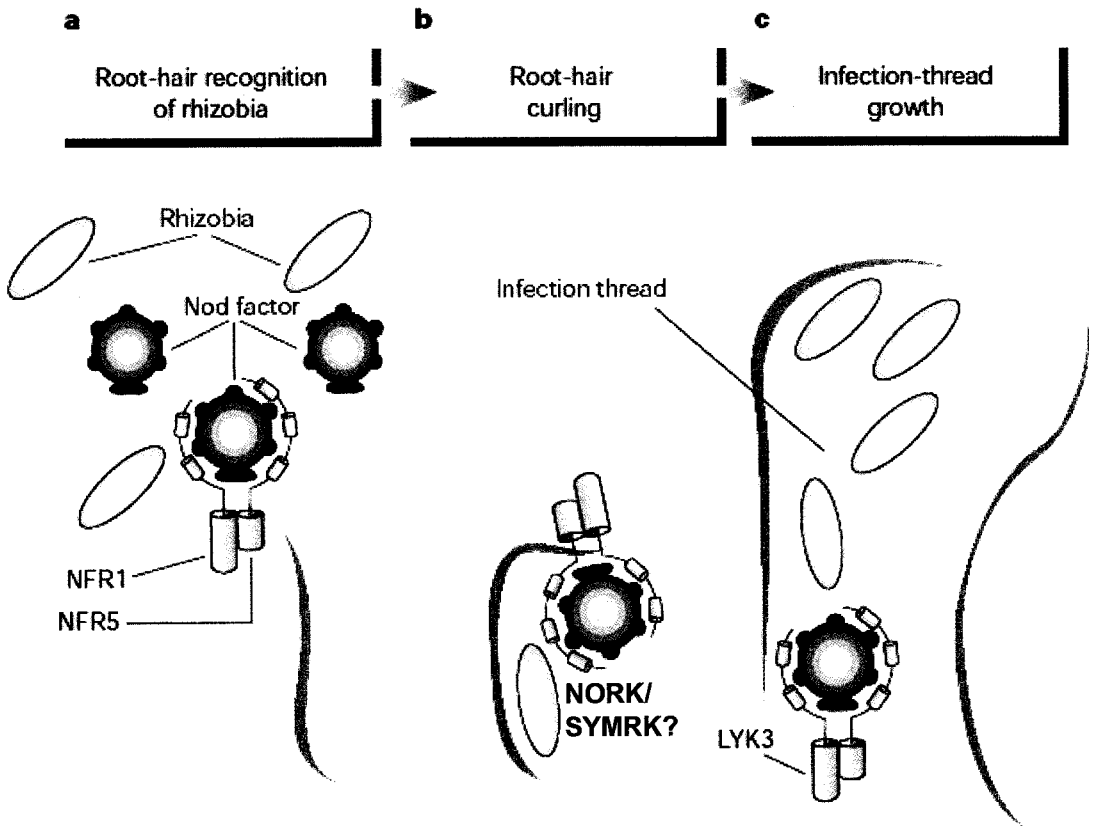
1.2.2 Regulation of Nodule Formation

Nodule formation is tightly regulated in legumes (Geurts and Bisseling, 2002; Stougaard, 2000; Szczyglowski and Amyot, 2003). Nodulation occurs only under nitrogen-limiting conditions, but even when this requirement is met, plants limit nodule number through local and global mechanisms.

In the presence of high concentrations of nitrate and/or ammonium, plants have no need for nodules and nodulation is suppressed (Malik *et al.*, 1987). In addition, when nitrate is applied to already nodulated plants, premature senescence of the nodules occurs (Brewin, 1991). The details of how nitrate regulates nodulation are not fully understood, but some aspects have been investigated. Experiments with split root systems have indicated that nitrate acts locally, and that nitrate inhibition is not due to nutritional effects because of increased assimilation of nitrogen (Carroll and Mathews, 1990; Streeter, 1988). Although nodulation is suppressed by the addition of nitrate, the perception of Nod factors still occurs, since nodulin genes can still be induced in the presence of nitrate (Heidstra *et al.*, 1997a).

The number of nodules is also thought to be limited locally by ethylene (Schultze and Kondorosi, 1998). Ethylene-insensitive *sickle* mutants of *M. truncatula* were hyperinfected by rhizobia; however, the mutants only exhibited nodule formation in the

Figure 7: Model of Nod Factor Perception (adapted from Parniske and Downie, 2003) Different receptors are thought to mediate the various events that occur during nodulation. NFR1 and NFR5 are likely to work together as the "signalling" receptor for recognition of the correct Nod factors and initiation of the nodulation program. The second receptor may be NORK/SYMRK, which would allow root hair curling and bacterial entry. LYK3 may function as another bacterial entry receptor with a role in the maintenance of infection thread formation. LYK, LysM Receptor-Like Kinase; NFR, Nod Factor Receptor; NORK, Nodulation Receptor Kinase; SYMRK, Symbiosis Receptor-Like Kinase.


a
b
c
Root-hair recognition of rhizobia
Root-hair curling
Infection-thread growth
Rhizobia
Nod factor
NFR1
NFR5
Infection thread
NORK/
SYMRK?
LYK3

susceptible area of the root, suggesting that the effect of ethylene is local (Penmetsa and Cook, 1997). Studies on wild type *M. truncatula* by Oldroyd *et al.* (2001) showed that ethylene has an inhibitory effect on early responses to Nod factors including calcium spiking, root hair deformation, *ENOD* expression, and infection thread initiation. When plants were treated with ACC, a precursor of ethylene, or AVG, an inhibitor of its biosynthesis, nodule numbers were decreased and increased, respectively. Furthermore, nodulation-deficient pea *sym5* mutants exhibit the formation of nodules when treated with inhibitors of ethylene perception, suggesting that a hypersensitivity to ethylene is responsible for their reduced nodulation phenotype (Fearn and LaRue, 1991). All of these results are consistent with a role for ethylene as a negative regulator of nodulation. Conflicting evidence comes from studies on soybean, since most cultivars do not exhibit inhibition of nodulation when exogenous ethylene is applied (Schmidt *et al.*, 1999; Xie *et al.*, 1996). However, nodulation is tightly regulated in this legume by some mechanism. It is possible that ethylene only regulates nodulation in legumes that produce indeterminate nodules such as *M. truncatula* (Stougaard, 2000).

The global regulation of nodulation, called the autoregulation of nodulation (AON), is controlled by the shoot and therefore requires long distance communication between the shoot and the root (Esseling and Emons, 2004). Grafting studies with hypernodulating soybean *nts-1* (*nitrate-tolerant symbiotic-1*) mutants demonstrated that shoot genotype controls root genotype, leading to the prediction that a compound from the nodule travels to the shoot followed by transportation of an inhibitor back to the root and nodules (Caetano-Anolles and Gresshoff, 1991). In accordance with this model, when the shoot of the hypernodulation mutant of *L. japonicus* (*har1* or *hypernodulation aberrant root 1*) was grafted on to a wild type root system, hypernodulation of the roots occurred

(Krusell *et al.*, 2002; Nishimura *et al.*, 2002). Receptor kinases appear to be involved in AON, since the *HAR1* and *NTS-1* loci encode these types of proteins. However, how they function is unknown (Krusell *et al.*, 2002; Searle *et al.*, 2003). Unfortunately, the mobile signal that mediates AON also remains undiscovered (Penmetsa *et al.*, 2003).

1.2.3 The Functional Nodule

Once the nodule is fully formed, the rhizobia begin to carry out nitrogen fixation, providing the plant with extra nitrogen to aid its growth. In exchange for nitrogen, the plant furnishes the bacteria with carbon acquired through photosynthesis.

1.2.3.1 Nodule Morphology

Nodules have characteristic structures (Figure 5 A,B). The central zone of determinate nodules and the nitrogen fixing zone of indeterminate nodules contain both infected and uninfected cells (Layzell and Atkins, 1997). The infected cells are usually larger than the uninfected cells and contain many bacteria-filled enclosures called symbiosomes, enveloped by peribacteroid membranes. At this point the rhizobia have differentiated into bacteroids, which carry out nitrogen fixation. Between one and four bacteroids reside within each symbiosome. The uninfected cells are distributed throughout the central or nitrogen fixing zone of determinate and indeterminate nodules and at least one is in contact with each infected cell. Uninfected cells are highly vacuolated compared to infected cells. Figure 8 depicts an infected cell adjacent to an uninfected cell in a determinate soybean nodule. Surrounding the central zone is the nodule cortex. The cells of the cortex are uninfected. Within the cortex, vascular bundles containing xylem and phloem connect the nodule to the root and the rest of the plant.

1.2.3.2 Nitrogen Fixation

In nodules, nitrogen fixation is carried out by the rhizobial enzyme nitrogenase

Figure 8: Inside the Central Zone of a Determinate Nodule (from Smith and Gallon, 1993)

This figure shows two uninfected cells flanked by infected cells. Note that the infected cells are much larger than the uninfected cells and that they are full of symbiosomes containing nitrogen-fixing bacteroids. Also note the large size of the vacuoles in the uninfected cells. CW, cell wall; INF, infected cell; UN, uninfected cell; Vac, vacuole.


vac
1 µm

according to the equation presented in Scheme 1:

$$N_2 + 8H^+ + 8e^- + 16\ \text{Mg-ATP} \xrightarrow{\text{nitrogenase}} 2\ NH_3 + H_2 + 16\ \text{Mg-ADP} + 16\ P_i$$

Scheme 1: Nitrogen Fixation Reaction (Mylona *et al.*, 1995)

Nitrogenase is a two component enzyme consisting of a homodimeric Fe, and a tetrameric molybdenum-iron (MoFe), protein (Mylona *et al.*, 1995). The Fe portion of the enzyme contains an Fe-S centre and two Mg-ATP binding sites, whereas the MoFe protein contains two [$MoFe_6S_8$] clusters which represent the active site (Layzell and Atkins, 1997). The Fe protein accepts electrons from either ferredoxin or flavoredoxin and passes them to the MoFe protein, where they are used to reduce N_2.

As shown in Scheme 1, a large amount of ATP is required for the reduction of N_2 to ammonium (Layzell and Atkins, 1997). ATP is also required for the synthesis of nitrogen-containing compounds derived from NH_3, as well as for normal cell growth and maintenance. Therefore, the nodule has a very high respiration rate, approximately four-fold higher than a comparable mass of root tissue, in order to provide the energy for all of these processes. Although necessary, O_2 can irreversibly denature the MoFe portion of nitrogenase (Layzell and Atkins, 1997). Therefore, the nodule has developed various strategies to protect nitrogenase from O_2. One strategy is the production of Lb, which is synthesized in large quantities just prior to the onset of nitrogen fixation (Smith and Gallon, 1993). Lb is a high affinity O_2 carrier similar to the animal protein myoglobin. As such, it controls the concentration of free O_2 in the nodule and delivers O_2 to the bacteroids. The morphology of the nodule also contributes to the limitation of free O_2 (Mylona *et al.*, 1995). Since O_2 can diffuse very rapidly through intercellular spaces, the

intercellular spaces of the innermost area of the nodule cortex are few and very small. Furthermore, if nodules are exposed to increased or decreased ambient O_2 concentrations over the longer term (days or weeks), the number of cell layers in this region can increase or decrease accordingly (Layzell and Atkins, 1997). Iannetta *et al.* (1995) suggest that legumes reduce O_2 diffusion in the short term by replacing the intercellular spaces of the inner cortex with water or a glycoprotein. Overall, nodule metabolism is limited by O_2 levels, providing a capacity to readily consume any additional O_2 (Layzell and Atkins, 1997).

1.2.3.3 Processing and Transport of Metabolites

In terms of whole plant physiology, nodules can be viewed as nitrogen sources and carbon sinks (Mylona *et al.*, 1995). The carbon source for the nodules is sucrose transported from the leaves, which is converted to usable sugars by the catabolic enzyme sucrose synthase in the infected cells of the root nodules (Layzell and Atkins, 1997; Mylona *et al.*, 1995; Smith and Gallon, 1993). Sucrose catabolites provide carbon for growth, for starch reserves in plant amyloplasts, for bacteroid poly-β-hydroxybutyrate granules, and for the production of nitrogen compounds that are exported from the nodule. In the cytosol of the infected cells, the C_4 acids malate, succinate, and fumarate, are synthesized and transported to the bacteroids in the symbiosomes. Once taken up by the bacteroids, the tricarboxylic acid cycle operates at high levels to assimilate the carbon (Smith and Gallon, 1993). Studies have shown that glucose and other disaccharides are poorly assimilated by the bacteroids, and that C_4 acid uptake mutants are unable to fix nitrogen, although they can grow, divide, and produce normal looking nodules. These results suggest that the metabolism of glucose supports the growth of bacteroids, whereas the metabolism of C_4 acids fuels nitrogen fixation.

The form in which nitrogen is exported from the nodule depends on whether the legume forms determinate or indeterminate nodules (Layzell and Atkins, 1997; Mylona *et al.*, 1995; Smith and Gallon, 1993). Legumes which form determinate nodules export the ureides allantoin and allantoic acid, whereas legumes that form indeterminate nodules export the amide asparagine (Layzell and Atkins, 1997).

In legumes which form determinate nodules, complicated metabolic conversions take place to produce ureides (Smith and Gallon, 1993). The assimilation of ammonium into glutamate and the subsequent biosynthesis of ureides are spatially separated (Mylona *et al.*, 1995). GS (glutamine synthetase) and GOGAT (glutamate synthase) are found in both infected and uninfected cells of the nodule's central zone, the purine synthesis pathway and its supporting reactions occur within the plastids of infected cells, and the enzymes involved in the purine oxidation reactions that result in ureides are found in the peroxisomes of the uninfected cells (Layzell and Atkins, 1997; Mylona *et al.*, 1995). Once glutamate has been formed in the plastids of infected cells by GS/GOGAT, purine biosynthesis resulting in the formation of inosine monophosphate (IMP) occurs (Layzell and Atkins, 1997). IMP is then exported to the cytosol, where it is oxidized to xanthine and uric acid by xanthine oxidoreductase. Uric acid is transported to the peroxisomes of the uninfected cells, where it is further oxidized to allantoin by uricase (Layzell and Atkins, 1997; Mylona *et al.*, 1995). Allantoin can be oxidized to allantoic acid by allantoinase in the ER of uninfected cells. A tubular system derived from the ER that is closely associated with the peroxisomes forms a continuous network through the uninfected cells to transport ureides to the vascular bundle in the nodule cortex (Mylona *et al.*, 1995).

In indeterminate nodules, aspartate is formed from ammonium when oxaloacetate is transaminated by aspartate aminotransferase (Smith and Gallon, 1993). It is then converted to asparagine by glutamine-dependent asparagine synthase. The location of asparagine synthase in the infected cells of the nodule is still not clear; it could be plastidic or cytosolic (Layzell and Atkins, 1997). The asparagine is then transported to the rest of the plant by transfer cells found in the vascular bundles of the nodule (Mylona *et al.*, 1995).

1.2.4 Genomics and Proteomics of Legume Nodulation

Advances in the genomics of various rhizobial and leguminous species are providing new insights into nodulation. The rhizobial species *S. meliloti*, *M. loti*, and *B. japonicum* have been studied intensively (Weidner *et al.*, 2003). *S. meliloti* and *M. loti* infect the model legumes *M. truncatula* and *L. japonicus*, respectively, and *B. japonicum* infects soybean. To date, the entire genomes of *S. meliloti* and *M.loti*, and part of the *B. japonicum* genome have been sequenced. Large scale genomics projects to study *L. japonicus* and *M. truncatula* are also being carried out. Although soybean is a polyploid characterized by a large genome, it is an important crop and therefore genome sequencing has also been initiated. Furthermore, many ESTs have been sequenced from *L. japonicus*, *M. truncatula*, and soybean. The ESTs are available in databases which allow identification of nodule-specific or nodule-enhanced genes. Websites where the above sequence information can be accessed are listed in Table I.

Transcriptome and proteome studies have been used to identify genes and proteins involved in symbiotic nitrogen fixation (Weidner *et al.*, 2003). For example, Colebatch *et al.* (2002) constructed macroarrays carrying 2304 *L. japonicus* cDNAs isolated from root nodules and identified 83 genes that were upregulated. Of the upregulated genes,

Table I: Genomic Sequence and EST Databases of Selected Legumes and Rhizobia

The websites containing genomic sequences and ESTs of legumes and their partner rhizobial species are presented. TIGR, The Institute for Genomic Research.

Species	Database	Genome Sequences or ESTs	Website
Sinorhizobium meliloti	*Sinorhizobium meliloti* strain 1021 Genome Project	Genome sequences	http://bioinfo.genopole- toulouse.prd.fr/annotation/iANT/bacteria/rhime/
Mesorhizobium loti	RhizoBase *Mesorhizobium loti* strain MAFF303099	Genome sequences	http://www.kazusa.or.jp/rhizobase/Mesorhizobium/index.html
Bradyrhizobium japonicum	*Bradyrhizobium japonicum* Genome Project	Genome sequences	http://www.genome.clemson.edu/projects/stc/b.japonicum/BJ__Ba/
Lotus japonicus	*Lotus japonicus* Genome Sequencing Project	Genome sequences	http://www.kazusa.or.jp/lotus/
	TIGR *L. japonicus* Gene Index	ESTs	http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=l_japonicus
	Lotus japonicus EST Index	ESTs	http://www.kazusa.or.jp/en/plant/lotus/EST/
Medicago truncatula	TIGR *Medicago* Genome Resources	Genome sequences	http://www.tigr.org/tdb/e2k1/mta1/

Species	Database	Genome Sequences or ESTs	Website
	Medicago truncatula Sequencing Resources	Genome sequences	http://www.medicago.org/genome/
	TIGR *Medicago* Gene Index	ESTs	http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=medicago
	Medicago truncatula Consortium Database Version 2.0	ESTs	http://www.medicago.org/MtDB2/Queries/SimilarityDB2.html
Glycine max (soybean)	SoyBase	Genome sequences	http://129.186.26.94/default.html
	TIGR Soybean Gene Index	ESTs	http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=soybean

50 had not been previously described as nodule-enhanced. Bestel-Corre *et al.* (2002) identified proteins found in *M. truncatula* root nodules by combining 2D electrophoresis and mass spectrometry. Using the EST database of *M. truncatula* to predict the fragmentation profiles of proteins, they were able to identify ~50 proteins involved in the *M. truncatula* and *S. meliloti* symbiosis. A recent innovation is a custom microarray containing the entire *S. meliloti* genome and approximately 10,000 genes of *M. truncatula*, allowing the gene expression of both symbiotic partners to be monitored in the nodule simultaneously (Barnett *et al.*, 2004).

1.3 Programmed Cell Death

Senescence, the final phase of soybean nodule development, is likely to be characterized by programmed cell death (PCD). Evidence for this suggestion has been provided by other studies on root nodule senescence (Alesandrini *et al.*, 2003; Evans *et al.*, 1999; Matamoros *et al.*, 1999) (Section 1.3.2.7.2) and senescence in other plant organs, which have been shown to exhibit features of PCD (Beers, 1997; Greenberg, 1996; Nooden *et al.*, 1997) (Section 1.3.2.7). The following discussion briefly introduces animal cell death, but focuses on plant cell death, and features that it shares with the process in animals. In addition, regulation of plant cell death and plant senescence as a form of PCD are discussed.

1.3.1 PCD in Animals

Cell death refers to the collection of processes responsible for dismantling cells. In animal cells, many types of PCD occur including apoptosis, autophagy, and less well defined types in which features of both processes have been observed in the same cell

(Assuncao Guimaraes and Linden, 2004; Leist and Jaattela, 2001; Lockshin and Zakeri, 2004). In addition, studies have shown that necrosis may be programmed and not just an uncontrolled form of death (Proskuryakov *et al.*, 2003). The emerging view of PCD is that it is a continuum of cell death styles with necrosis at one extreme and apoptosis at the other (Assuncao Guimaraes and Linden, 2004).

1.3.2 Plant PCD

Although originally characterized in animal systems, PCD is now known to be widespread in the plant kingdom (Mittler, 1998; Pennell and Lamb, 1997). Initial attempts to describe plant PCD involved searching for morphological and biochemical features similar to those of animal apoptosis. However, it is becoming apparent that plant PCD can vary just as widely as animal PCD, exhibiting hallmarks of apoptosis, autophagic cell death, necrosis, or combinations of these death styles (Assuncao Guimaraes and Linden, 2004; Levine *et al.*, 1996; Proskuryakov *et al.*, 2003; Zhivotovsky, 2002).

PCD is observed both during plant development and in the responses of plants to their environment. It plays a role during various plant reproductive processes, including gymnosperm pollen development and unisexual flower formation, anther dehiscence, the self-incompability response, and fertilization of angiosperms (Beers, 1997; Mittler, 1998; Van Doorn and Woltering, 2005). Numerous examples of PCD also occur throughout vegetative growth and development. During seed germination in monocots such as barley, PCD occurs in the seed's aleurone layer (Mittler, 1998; Pennell and Lamb, 1997). Root cap cells undergo PCD to provide a layer of dead cells to protect the growing root tip and leaf shape may be molded by PCD of selected cells, such as in the genus *Monstera* (Greenberg, 1996; Mittler, 1998; Van Doorn and Woltering, 2005). Hollow stems in

plants such as maize arise from the PCD of pith cells and PCD occurs during the formation of the conducting tissues of phloem and xylem (Beers, 1997; Greenberg, 1996; Fukuda, 1997; Mittler, 1998; Pennell and Lamb, 1997;Van Doorn and Woltering, 2005). PCD also occurs during the senescence of plant organs, including leaf and flower senescence and fruit ripening, and during whole plant senescence that occurs after fertilization (Beers, 1997; Greenberg, 1996; Nooden *et al.*, 1997).

PCD is activated by plants in response to biotic and abiotic stresses in their environment. For example, in response to hypoxia, plants can form aerenchyma, or tissue with internal air spaces, by PCD of selected cells in the cortex of the submerged tissues (Pennell and Lamb, 1997). When plants are challenged with biotic stresses such as pathogens, one of two outcomes is possible. The plant will recognize an avirulent pathogen and infection will be arrested, or it will not recognize a virulent pathogen and an infection will progress from its origin to the rest of the plant. When an avirulent pathogen is recognized, a rapid PCD response called the hypersensitive response (HR) is inititated in the cells surrounding the infection site (Greenberg, 1996; Mittler and Lam, 1996; Pennell and Lamb, 1997).

1.3.2.1 Morphological Features of Plant PCD

Morphological features similar to those of animal PCD have been described during plant PCD, although there is great deal of variation in the features observed. Similar to animal cells, cell shrinkage and condensation is usually observed, along with membrane blebbing, although the nucleus usually remains intact and condenses instead of fragmenting (McKenna *et al.*, 1998; Mittler, 1998; Pennell and Lamb, 1997). Nuclear changes are highly variable and are often different from those observed in animal cells undergoing PCD (Mittler and Lam, 1996). During PCD in cells undergoing the HR and

in prematurely degenerating tapetums of sunflower *pet1-cms* (*petiolaris 1-cytoplasmic male sterility*) mutants, plant nuclear material typically clumps throughout the nucleus rather than condensing at the periphery of the nuclear envelope (Balk and Leaver, 2001; Mittler and Lam, 1996). When PCD occurs in *Arabidopsis* protoplasts treated with UV-C radiation or in tomato protoplasts treated with the mycotoxin AAL, the nuclei fragment (Danon and Gallois, 1998; Wang *et al.*, 1996). In addition, both crescent-shaped and round nuclei appear prior to nuclear fragmentation in dying cells of UV-C-treated *Arabidopsis* protoplasts and in those of tobacco stems infected with TMV (Danon and Gallois, 1998; Mittler *et al.*, 1997). However, although chromatin condensation occurs during TE formation, neither condensation nor fragmentation of the nuclei occurs; instead, the nuclei maintain their shape even after the nucleic acids have been digested (Obara *et al.*, 2001).

Since plants do not have an immune system comparable to that of animals, they do not have macrophages (Mittler, 1998). Therefore, formation of apoptotic bodies and the engulfment of cell corpses are not usually observed during plant PCD, although similar structures have been observed during toxin-induced PCD in tomato and during HR-induced PCD in soybean cell cultures and leaves (Levine *et al.*, 1996; Wang *et al.*, 1996). Plants have rigid cell walls, which would prevent the release of cellular contents into the intercellular space in the form of apoptotic bodies. Instead, it is thought that the recycling of material from dying cells occurs by autolysis of cellular debris into low molecular weight compounds that are released by ion leakage for uptake by nearby cells (Greenberg, 1996; Panavas *et al.*, 1998; Pennell and Lamb, 1997).

1.3.2.2 Biochemical Features of Plant PCD

Various biochemical features of animal PCD have been observed in plants, including the ordered fragmentation of nuclear DNA, the activation of proteases and Ca^{2+}-dependent nucleases, the translocation of negatively charged phospholipids to the outer membrane, and the appearance of ROS (Danon *et al.*, 2000; Greenberg, 1996; Pennell and Lamb, 1997).

DNA laddering has been described in many instances of plant PCD. During plant development, DNA laddering has been observed in petal senescence of pea and petunia and during leaf senescence (Orzaez and Granell, 1997a; Xu and Hanson, 2000; Yen and Yang, 1998). It has also been observed in barley aleurone protoplasts undergoing PCD and during the HR induced in cowpea leaves by rust fungus (Ryerson and Heath, 1996; Wang *et al.*, 1996). DNA laddering occurs during induced PCD in AAL toxin-treated tomato protoplasts and leaflets, menadione-treated tobacco protoplasts, heat-treated cucumber cotyledons, UV-C-treated protoplasts and seedlings of *Arabidopsis*, and D-mannose-treated corn cell suspensions (Balk *et al.*, 1999; Danon and Gallois, 1998; Stein and Hansen, 1999; Sun *et al.*, 1999; Wang *et al.*, 1996). In addition, PCD that occurs in response to cold stress in tobacco cells, nutrient deprivation in *Arabidopsis* cell suspensions, and salt stress in barley roots is accompanied by DNA laddering (Callard *et al.*, 1996; Katsuhara, 1997; Koukalova *et al.*, 1997). However, no laddering has been observed in TE formation, and in soybean cells induced to undergo PCD by treatment with the bacterial pathogen *Pseudomonas syringae* pv. *glycinea*, 50 kb fragments of DNA were observed, but not laddering (Levine *et al.*, 1996; Obara *et al.*, 2001).

Proteolysis is associated with both animal and plant PCD and various classes of proteases have been linked to developmental and pathogen-, stress-, or toxin-induced PCD (Hoeberichts and Woltering, 2002). Aspartic proteases have been observed in PCD

of the nucellus after ovule fertilization, during TE formation, and during leaf and flower senescence (Beers, 1997; Beers *et al.*, 2000; Chen and Foolad, 1997). Cysteine endopeptidases of castor bean, housed in structures called ricinosomes, are released when PCD of the endosperm occurs during seed germination (Schmid *et al.*, 1999). Both cysteine and serine proteases have been isolated from *Zinnia* cells during TE differentiation and serine proteases are required during PCD accompanying the elicitin-induced HR in tobacco suspension cultures (Sasabe *et al.*, 2000; Ye and Varner, 1996). Furthermore, thiol and serine proteases are upregulated during barley and oat leaf senescence and PCD accompanying anther dehiscence is characterized by the accumulation of thiol proteases and an increase in ubiquitin and ubiquitin-conjugated protein levels associated with proteasome activity (Beers, 1997; Koltunow *et al.*, 1990; Li *et al.*, 1995). The proteasome may play a role in regulating TE PCD, since inhibition of the proteasome can prevent or delay TE differentiation (Woffenden *et al.*, 1998). In soybean cell cultures, cysteine proteases are induced by oxidative stress-induced PCD and the cysteine protease inhibitor cystatin blocks this cell death (Solomon *et al.*, 1999). Therefore, it is possible that proteases and endogenous protease inhibitors interact to regulate PCD, similar to caspases and IAPs in animal apoptosis (Hoeberichts and Woltering, 2002).

Nucleases have also been implicated in plant PCD. Ca^{2+}-dependent *DN*ase activity is detected during petunia petal senescence (Xu and Hanson, 2000). In addition, endonucleases appear during PCD induced by D-mannose in corn cells and *Arabidopsis* roots, and in tobacco cells undergoing PCD due to virus infection or wounding (Mittler and Lam, 1995, 1997; Stein and Hansen, 1999).

Phosphatidylserine translocation to the outer membrane of a cell is an early indicator of PCD in animals, detectable by the binding of annexin V (Danon *et al.*, 2000). The presence of this biochemical marker has not been widely studied in plant PCD, but experiments with tobacco cells induced to undergo PCD by camptothecin treatment showed that annexin V binds to the outer membrane of the cells, suggesting that phosphatidylserine translocation may be occurring (O'Brien *et al.*, 1997).

ROS are involved in animal PCD and have also been detected in various types of plant PCD (Hoeberichts and Woltering, 2002). For example, an increase in levels of ROS has been detected in naturally senescing pea leaves (Pastori and Del Rio, 1997). ROS in the form of H_2O_2 are able to trigger cell death in carrot and soybean cell cultures, and superoxide can trigger PCD in *Arabidopsis lsd1* (*lesion simulating disease 1*) or *rcd* (*runaway cell death*) mutants (Jabs *et al.*, 1996; Levine *et al.*, 1996; McCabe *et al.*, 1997). The appearance of an oxidative burst is well known in the HR, and in cultured parsley cells, the oxidative burst is one of the first responses to fungal pathogens (Jabs *et al.*, 1997; Lamb and Dixon, 1997; Levine *et al.*, 1994). In addition, the HR induced in tomato cotyledons by elicitors of the pathogen *Cladosporium fulvum* is accompanied by an oxidative burst (May *et al.*, 1996). However, an accumulation of ROS has not been observed during TE formation (Jones and Dangl, 1996).

1.3.2.3 Caspase-Like Proteins (CLPs) in Plants

Caspases play an important role in animal PCD as both regulators and effectors of apoptosis (Danon *et al.*, 2000; Hoeberichts and Woltering, 2002). Although caspase genes have not been identified in plant genomes by homology searches, caspase inhibitor experiments have indicated that caspase-like activity is present in plant cells during PCD (Lam *et al.*, 2001). For example, De Jong *et al.* (2000) used peptide caspase inhibitors to

show that caspase-like activity is present during camptothecin-, staurosporin-, and fumonisin B1-induced PCD in tomato cells and Clarke *et al.* (2000) showed that peptide caspase inhibitors blocked H_2O_2- and NO-induced cell death in *Arabidopsis* suspension cultures. Peptide inhibitors specific for caspase-1 and caspase-3 of mammals both abolished HR-induced cell death in tobacco leaves infected with *P. syringae* pv. *phaseolicola* and caspase activity was detected in tobacco leaves infected with TMV (Del Pozo and Lamb, 1998). Furthermore, transgenic tobacco plants expressing the broad range caspase inhibitor p35 exhibited partial inhibition of HR-induced cell death (Del Pozo and Lam, 2003).

The absence of identifiable caspase gene sequences in the *Arabidopsis* genome suggests that they are divergent between plants and animals (Watanabe and Lam, 2004). However, recent work using iterative homology searches have led to the identification of metacaspases in plants and fungi (Uren *et al.*, 2000). Two types of metacaspases have been discovered. Type I metacaspases, specific to the plant kingdom, are characterized by an N-terminal zinc finger and proline-rich repeat motif. The N-terminal zinc finger-proline motif is also found in LSD1, a protein thought to be involved in regulating PCD during the HR (Jabs *et al.*, 1996; Uren *et al.*, 2000). Type II metacaspases, found in both plants and fungi, have predicted proteolytic domains corresponding to the p20 and p10 subunits of caspases (Uren *et al.*, 2000). Structural modelling of the metacaspases indicates that they have significant homology to caspases in terms of tertiary structure (Woltering *et al.*, 2002). Three and six genes encode type I and II metacaspases in the *Arabidopsis* genome, respectively (Watanabe and Lam, 2004). During the HR in *Arabidopsis*, two type I metacaspases are upregulated, whereas the message levels of the two type II metacaspases that were tested were unchanged. In contrast, a type II

metacaspase was rapidly induced upon infection of tomato leaves with a fungal pathogen, although this same metacaspase was not induced in toxin-induced PCD in tomato cell suspensions (Hoeberichts *et al.*, 2003). The understanding of the biochemistry and molecular biology of these plant enzymes is in its infancy and it is not clear whether the metacaspases are equivalent to classical animal caspases in terms of their cellular targets and roles in regulation of PCD (Lam *et al.*, 2001). However, the yeast metacaspase YCA1 (yeast caspase 1) has been shown to exhibit caspase-like protease activity during H_2O_2-stimulated PCD in yeast (Madeo *et al.*, 2002).

A second group of caspase-related proteins found in plants are the legumains (Chen *et al.*, 1998). Even though their targets contain asparagine N-terminal to the cleavage sites, and not aspartate, they are thought to be evolutionarily related to caspases and their tertiary structures are very similar to those of animal caspases. Recently, the legumain vacuolar processing enzyme (VPE), originally isolated as an enzyme that processes seed storage proteins, was found to be essential for the TMV-induced HR in tobacco leaves (Hatsugai *et al.*, 2004). Since both caspase-1 inhibitors and VPE inhibitors blocked the HR, VPE was identified as an enzyme with caspase-1 activity. VPE is thought to operate in a manner similar to regulatory caspases by promoting the maturation and activation of various vacuolar proteins prior to vacuole collapse during cell death. Since *VPE* is also upregulated during leaf senescence in *Arabidopsis*, it may be a general regulator of plant PCD (Kinoshita *et al.*, 1999).

Other CLPs involved in plant PCD continue to be identified (Lam, 2005). Chichkova *et al.* (2004) found an additional, nuclear-localized CLP with caspase-1 activity in TMV-induced HR in tobacco leaves. In oat leaf cell death induced by the toxin victorin, a CLP with caspase-3-like activity which was excreted to the extracellular space

was detected (Coffeen and Wolpert, 2004). Upon further investigation, two distinct CLP activities were distinguishable, one which cleaved caspase-3-specific substrates and one which cleaved a more general caspase substrate. Both of these proteins contained subtilisin-like serine protease motifs and were therefore called "saspases" (serine-containing aspartate-specific proteases). Their role in PCD remains unknown. It is possible that multiple CLPs, localized to different compartments, could operate during plant PCD (Lam, 2005).

1.3.2.4 Mitochondrial Involvement in Plant PCD

Release of cytochrome c and other factors from the intermembrane space of the mitochondria is an early event that commits animal cells to death through the intrinsic pathway of apoptosis. Zhao *et al.* (1999) showed that a cytosolic extract of carrot cells supplemented with animal cytochrome c could induce the appearance of apoptotic-like features in mouse liver nuclei. This result implied that cytochrome c could induce PCD in both plant and animal cells. Evidence suggesting that loss of mitochondrial integrity, and subsequent release of molecules such as cytochrome c, play a role in plant PCD continues to accumulate.

Tiwari *et al.* (2002) showed that oxidative stress caused an increase in mitochondrial permeability and release of cytochrome c from the intermembrane space, resulting in ATP depletion and PCD in *Arabidopsis* cell cultures. Furthermore, it was shown that blocking this mitochondrial permeability transition prevented cell death. The opening of a mitochondrial pore has also been implicated in the control of victorin-induced PCD in oat leaves (Curtis and Wolpert, 2002). Recent research by Yao *et al.* (2004) on PCD triggered in *Arabidopsis* protoplasts by various treatments showed that mitochondrial membrane potential is lost early and can be used as a marker for PCD.

Together, these results suggest that increased mitochondrial permeability plays a role in the regulation of plant PCD.

Various studies have shown that the release of cytochrome c is correlated with plant PCD. For example, cytochrome c is released from the mitochondria prior to the appearance of any features of PCD in heat shocked cucumber cotyledons, during menadione-induced PCD in tobacco protoplasts, and during PCD in the prematurely degenerating tapetums of *pet1-cms* sunflower plants (Balk and Leaver, 2001; Balk *et al.*, 1999; Sun *et al.*, 1999). D-mannose-induced PCD in corn cells is also accompanied by release of cytochrome c from the mitochondria, although it does not appear in the cytosol until after nuclear degradation has begun; in this case, it may reflect that loss of mitochondrial integrity is a result of PCD instead of an initiating signal (Stein and Hansen, 1999). Yu *et al.* (2002) examined the role of mitochondria during TE formation in *Zinnia* mesophyll cell cultures. Cytochrome c was released from the mitochondria prior to PCD, but blocking the cell death program did not prevent the release of cytochrome c, implying that its release is not required for PCD. Interestingly, no cytochrome c release from the mitochondria was observed during PCD that occurs when petunia petals senesce (Xu and Hanson, 2000). The role of cytochrome c in plant PCD is still uncertain.

During the HR, the mitochondria remain relatively stable until late in the process of PCD and it is not clear whether increased permeability or leakage of cytochrome c plays a role in HR-induced PCD (Lam *et al.*, 2001; Mittler and Lam, 1996). However, an HR-inducing bacterial virulence factor (harpin) disrupts mitochondrial function in tobacco cells (Xie and Chen, 2000). In addition, the expression of BAX, an animal pro-apoptotic BCL-2 family protein, in tobacco cells initiates HR-like cell death (Lacomme and Santa

Cruz, 1999). Studies of alternative oxidase (AOX), an inner mitochondrial membrane (IMM) enzyme unique to plants, suggest that the plant mitochondrion may be a signal generator for HR-induced PCD by the production of ROS (Lam *et al.*, 2001). AOX may act as a type of "safety valve" for the control of ROS production (Lam *et al.*, 1999).

Although several studies point to a role for the mitochondria in plant PCD, it is not clear whether the mitochondria initiate cell death or whether they release molecules such as cytochrome c after organellar breakdown has begun.

1.3.2.5 The Vacuole and Plant PCD

The vacuole, an organelle unique to plants, may also play a role in carrying out PCD (Mittler, 1998). During many instances of plant PCD, the vacuole increases in size until late in the process when it bursts, releasing hydrolytic enzymes into the cell. Vacuolar disruption followed by rapid degeneration of cellular contents has been observed in PCD during the development of TEs, aerenchyma tissue, and phloem cells, and during leaf and flower senescence (Campbell and Drew, 1983; Fukuda, 2000; Van Doorn and Woltering, 2005; Vercher *et al.*, 1987). In PCD that occurs during embryo elimination in polyembryonic pine seeds, the cytoplasm and organelles are slowly engulfed and lysed by a high number of small vacuoles (Filonova *et al.*, 2002). This slow phase of cell death is followed by a sudden vacuolar collapse, after which point the protoplast is rapidly removed.

The role that the vacuole plays in HR-induced PCD is not clear. When tobacco leaves are infected with TMV, vacuolar collapse occurs, but the condensed cytoplasm remains intact, suggesting that the vacuolar contents are not responsible for killing the cell (Hatsugai *et al.*, 2004; Mittler *et al.*, 1997). Cells undergoing HR-induced PCD in oat leaves and soybean leaves and cells that had been infected with the avirulent pathogens

Pseudomonas coronata f. sp. *avenae* and *P. syringae* pv. *glycinea*, respectively, exhibited neither increased vacuolization nor vacuolar collapse (Greenberg and Yao, 2004; Levine *et al.*, 1996). Furthermore, when toxins such as AAL, victorin, or fumonisin B1 are infiltrated in leaves or applied to cell cultures, cell shrinkage and condensation occurs during PCD, but the vacuole remains intact (Van Doorn and Woltering, 2005).

Recently, senescence-associated vacuoles (SAVs) have been observed in *Arabidopsis* and soybean leaf senescence (Otegui *et al.*, 2005). During senescence, a collection of small, lytic vacuoles develops in the peripheral cytoplasm of chloroplast-containing cells such as mesophyll and guard cells. When *SAG12*, a senescence-associated gene encoding a cysteine protease, was fused to *GUS*, *GUS* expression was localized to SAV-containing cells in the leaves. In addition, a SAG12-GFP protein was localized to the SAVs. Therefore, SAVs may sequester these cysteine proteases until they are activated, possibly by a change in SAV pH, and released into the cytoplasm. SAVs have a different membrane composition from the central vacuole, suggesting that they do not originate from it. Although the dense appearance of the SAV resembles autophagic compartments, *Arabidopsis* mutants that cannot form autophagosomes still contain SAVs, suggesting that they are not autophagic in origin. Different vacuoles may provide unique lytic compartments for dismantling of specific cellular components during PCD.

1.3.2.6 Regulation of Plant PCD

The key molecules which control plant PCD remain largely unidentified, but attempts to find plant homologues of animal regulatory genes are beginning to yield insight into the regulation of plant PCD (Watanabe and Lam, 2004).

DAD1 (*Defender* against *Apoptotic Death 1*) was first isolated as a gene encoding a protein that could rescue cells of a temperature-sensitive hamster line from apoptosis

(Nakashima *et al.*, 1993). *DAD1* has since been identified in the rice EST database and has been cloned from various plant species including *Arabidopsis*, pea, tomato, mandarin, and apple (Apte *et al.*, 1995; Dong *et al.*, 1998; Hoeberichts and Woltering, 2001; Moriguchi *et al.*, 2000; Orzaez and Granell, 1997a; Sugimoto *et al.*, 1995). Transformation of the temperature-sensitive hamster cell line from which *DAD1* was originally isolated with *Arabidopsis DAD1* resulted in the suppression of cell death (Gallois *et al.*, 1997). Since DAD1 is highly conserved evolutionarily and in terms of its functionality, it may be a universal regulator of cell death (Hoeberichts and Woltering, 2002).

Sequence analysis of the growing databases of plant ESTs and the sequenced genomes of *Arabidopsis* and rice have failed to unambiguously identify any homologues of the animal *BCL-2* genes (Watanabe and Lam, 2004). However, immunolocalization studies in tobacco leaves and maize meristematic tissues have detected BCL-2-like epitopes in mitochondria, chloroplasts, and nuclei (Dion *et al.*, 1997). Lacomme and Santa Cruz (1999) showed that expression of the mouse pro-apoptotic *BAX* gene in tobacco induced cell death in leaves. In addition, when the anti-apoptotic gene *BCL-X_L* was expressed in tobacco, it was able to suppress cell death induced by UV, the herbicide paraquat, or TMV (Mitsuhara *et al.*, 1999). Although none of the corresponding *BCL-2* genes have been identified in plants, there is evidence that their action is conserved between animals and plants.

Homologues of *BI-1* (*Bax Inhibitor-1*), first isolated from a human cDNA library as a suppressor of yeast apoptosis, have been found in various plant species including tobacco, canola, and *Arabidopsis* (Bolduc *et al.*, 2003; Lam *et al.*, 2001; Shaham *et al.*, 1998; Xu and Reed, 1998). Like animal *BI-1*s, plant *BI-1*s can also inhibit PCD. For

example, BAX-induced apoptosis is inhibited in human kidney cells transformed with tobacco or canola *BI-1* and in yeast transformed with *Arabidopsis BI-1* (Bolduc *et al.*, 2003; Kawai *et al.*, 1999). Furthermore, antisense suppression of tobacco *BI-1* in BY-2 tobacco cells correlates with increased stress-induced PCD (Bolduc and Brisson, 2002). *AtBI-1* is also upregulated in response to wounding and pathogen challenge in *Arabidopsis* plants, suggesting a role in various types of stress (Sanchez *et al.*, 2000). Interestingly, expression of *AtBI-1* in a human fibrosarcoma cell line was able to induce, rather than inhibit, apoptosis (Yu *et al.*, 2002). Since the induction of apoptosis could be repressed by coexpression of the caspase inhibitor XIAP, it is thought that *BI-1* may be able to function as a pro-survival or pro-death protein depending on the cellular context. All known *BI-1* proteins contain 5-7 transmembrane α-helices, suggesting they may form pores similar to those thought to be formed by the BCL-2 family in the mitochondria (Lam, 2004; Lam *et al.*, 2001).

Three *BI* genes containing C-terminal *BI-1* domains fused to variable N-terminal domains have been identified in the *Arabidopsis* genome (Lam *et al.*, 2001). Use of the N-terminus of *AtBI-2* to search for related sequences in plant gene databases resulted in the identification of a new gene family called *ABR*s (*AtBI-2-RELATED PROTEINS*). There are 13 of these genes in *Arabidopsis*, each member encoding a protein of approximately 650 amino acids, containing structural hallmarks of channel-forming proteins. These proteins may interact with BI-1 and could be the functional equivalents of BCL-2 proteins in plants.

Plant-specific genes that encode regulators of HR-induced PCD have been isolated. *LSD1* was isolated as the gene responsible for the mutation in *Arabidopsis* plants exhibiting the *rcd* (*runaway cell death*) phenotype, in which cell death cannot be

stopped after an initial stimulus sets it in motion (Jabs *et al.*, 1996). It has therefore been defined as a negative regulator of cell death. A relative of *LSD1*, *LOL1* (*LSD-Like-1*) has also been characterized in *Arabidopsis* (Lam, 2004). Both LOL1 and LSD1 have conserved zinc finger motifs and are putative transcription factors (Epple *et al.*, 2003). LOL1 is thought to be a positive regulator of cell death since suppression of this gene in *lsd1* mutants suppresses rcd, whereas increased expression in wild type plants results in accelerated HR-induced cell death. It may be that LSD1 and LOL1 are plant cell death regulators that function antagonistically to establish a threshold for the activation of HR-induced PCD. It will be interesting to see whether they also regulate other types of plant PCD (Lam, 2004). *LOL2* is a third member of this gene family, but it has not been characterized. *Mlo* (*mycoplasma-like organism*) loss of function mutants, originally described in barley, display spontaneous lesion-mimic phenotypes of the initiation class in which PCD is initiated spontaneously in the absence of any inductive signals (Lam, 2004). The MLO proteins are thought to be negative regulators of cell death.

Plant hormones play critical roles in the regulation of plant development; therefore, it is not surprising that they have also been implicated in some forms of PCD. Ethylene, a well known stress-responsive and senescence hormone, has been shown to have a role in the regulation of some forms of PCD. For example, it is involved in senescence-associated PCD during pea carpel senescence, in victorin-induced senescence in oats (Navarre and Wolpert, 1999; Orzaez and Granell, 1997) and in fumonisin B1-induced PCD in *Arabidopsis* protoplasts (Asai *et al.*, 2000). During seed germination in barley, cell death in the aleurone layers is promoted by gibberellins and blocked by abscisic acid (Wang *et al.*, 1996). Gibberellins have also been implicated in PCD of epidermal cells that cover adventitious root primordia in deepwater rice; however, they

act synergistically with ethylene and cannot promote PCD on their own (Steffens and Sauter, 2005). Abscisic acid has also been found to mediate or protect against cell death in maize endosperm development and during the development of barley anthers (Wang *et al.*, 1999; Young and Gallie, 2000). Jasmonic acid, a hormone that is known to be involved in plant defence and stress responses, appears to have a role in the negative regulation of PCD. For example, the application of methyl jasmonate to O_3-sensitive *Arabidopsis* plants inhibits the spread of O_3-induced PCD (Rao and Davis, 2001). Furthermore, *jar1* mutants, which are insensitive to jasmonic acid, exhibited enhanced propagation of PCD following O_3 exposure (Overmyer *et al.*, 2000). Cytokinins influence many aspects of plant development, including seed germination, vascular tissue development, and in the regulation of leaf senescence (Ferreira and Kieber, 2005). However, they have also been shown to induce apoptosis in carrot and *Arabidopsis* cell cultures, indicating that they, too, are involved in the regulation of PCD (Carimi *et al.*, 2003). Salicylic acid is a molecule involved in plant disease resistance, but it is thought that it may also play a role in the regulation of other forms of PCD (Hoeberichts and Woltering, 2002). For example, salicylic acid signalling is required for fumonisin B1-induced cell death in *Arabidopsis* protoplasts (Asai *et al.*, 2000). A current model of how plant PCD occurs suggests that hormones may affect the process by either enhancing or attenuating ROS (Hoeberichts and Woltering, 2002). However, as more information becomes available on cross-talk between hormone response pathways, this model will undoubtedly be refined.

1.3.2.7 Plant Senescence

Senescence, the last stage of plant development, refers to both the aging of specific tissues and organs as a plant matures, and to monocarpic or whole plant senescence that occurs after flowering (Greenberg, 1996; Nooden *et al.*, 1997).

Plant senescence is thought to be a specialized form of PCD. DNA laddering and other hallmarks of PCD were shown to coincide with petal senescence in petunia (Xu and Hanson, 2000). During the last stages of leaf senescence, when the leaves are almost completely yellow, chromatin condensation and DNA laddering are observed (Yoshida, 2003). In tomato leaves undergoing senescence, cysteine proteases are upregulated (Drake *et al.*, 1996). In addition, the accumulation of ROS has been demonstrated in senescing pea leaves (Pastori and Del Rio, 1997).

Senescence usually involves the breakdown of cellular material and its recapture for use in other organs (Pennell and Lamb, 1997). For example, during leaf senescence, cellular components such as proteins, lipids, and amino acids are degraded and the nutrients are mobilized from the leaves for reuse in other parts of the plant (Lim *et al.*, 2003; Quirino *et al.*, 2000). In addition, although most cellular proteins show substantial decreases in their levels, degradative enzymes such as cysteine, aspartic, and serine proteases, as well as metalloproteases, all increase in activity, concentration, or expression (Buchanan-Wollaston and Ainsworth, 1997; Drake *et al.*, 1996; Smart *et al.*, 1995; Weaver *et al.*, 1998). The efficient recycling of nutrients requires that senescence occur more slowly than other types of PCD, which are generally associated with acute, rapid cellular breakdown. Therefore, cells in senescing organs undergo a gradual, orderly disassembly, with many of the organelles remaining intact until very late in the process (Nooden *et al.*, 1997).

1.3.2.7.1 Genes Upregulated during Plant Senescence

Many studies have supported the idea that the onset of plant senescence involves the expression of a large number of genes whose protein products are required to carry out senescence-related processes (Gepstein, 2004). To date, a large number of genes involved in senescence have been isolated from various plants. In addition, the development of microarray technology has enabled researchers to monitor the fate many genes simultaneously during plant senescence (Gepstein *et al.*, 2003; Guo *et al.*, 2004). Table II gives examples of the types of genes that have been isolated. Many of the genes encode enzymes which are involved in the degradation of proteins, nucleic acids, lipids, and polysaccharides (Gepstein, 2004). This fact is not surprising, given that senescence is an organized dismantling of cells resulting in a breakdown of cellular components (Nooden *et al.*, 1997; Pennell and Lamb, 1997). It is apparent that *SAG*s are also upregulated during plant responses to stress and pathogens (Lim *et al.*, 2003; Nooden *et al.*, 1997; Quirino *et al.*, 2000). Therefore, some features are likely to be shared between senescence and these processes at the molecular biological level.

1.3.2.7.2 Root Nodule Senescence

In root nodules, senescence has not been well characterized and it is not certain whether it represents a case of PCD, although some features of PCD have been observed during the process. For example, ROS appear in naturally senescing soybean nodules and an ordered breakdown of the cytoplasm and membranes reminiscent of PCD occurs in pea and bean nodules senescing in response to dark or nitrate treatment (Alesandrini *et al.*, 2003; Evans *et al.*, 1999; Matamoros *et al.*, 1999). In addition, internucleosomal DNA fragmentation has been observed in senescing soybean nodules (Alesandrini *et al.*, 2003). Cysteine proteases have also been implicated in pea and soybean nodule senescence, although CLPs have not been detected in nodules (Alesandrini *et al.*, 2003; Kardailsky

Table II: Plant Senescence-Associated Genes

Examples of plant senescence-associated genes, organized into various functional categories, are presented. Senescence refers to naturally occurring senescence unless otherwise specified. *APG*; autophagy; *At*, Arabidopsis thaliana; *bZIP*, basic region leucine zipper; *LE*, extracellular ribonuclease; *LSC*, leaf senescence cDNA; *LX*, intracellular ribonuclease; MAP, mitogen-activated protein; *MPK*, MAP kinase; PR, pathogenesis-related; MYB, myeloblastosis oncogene; *SAG*, senescence-associated gene; *SARK*, senescence-associated receptor-like kinase; *SENU*, senescence-upregulated; *SFP*, sugar-porter family protein; *SR*, senescence-related; *tbz*, tobacco bZIP; WRKY, amino acids in conserved domain of this family of transcription factors; *Zm*, *Zea mays*.

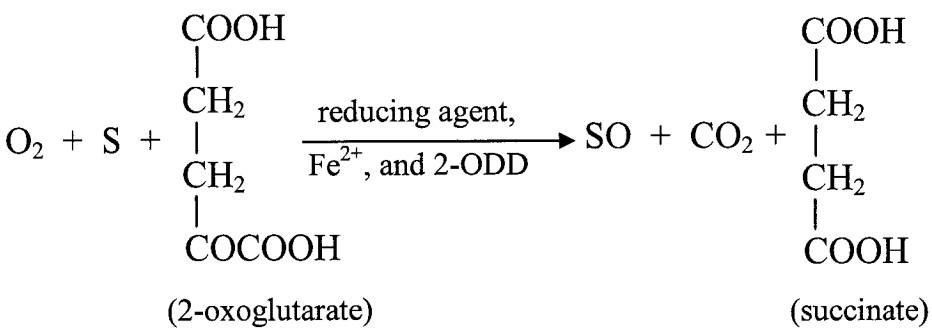
Functional Category	Gene Name	Plant	Type of Gene Product	Plant Processes	Reference
Macromolecule Degradation	*SAG101*	*Arabidopsis*	acyl hydrolase	leaf senescence	He and Gan, 2002
	SAG12	*Arabidopsis*	cysteine protease	leaf senescence	Lohman *et al.*, 1994
	LX, LE	tomato	ribonuclease	leaf senescence phosphate starvation wound response (*LE*)	Lers *et al.*, 1998
Nutrient Recycling	*LSC460*	canola	glutamine synthetase	leaf senescence	Buchanan-Wollaston and Ainsworth, 1997
	AtAPG7	*Arabidopsis*	*APG8/12*-activating enzyme	leaf senescence	Doelling *et al.*, 2002
	SFP1	*Arabidopsis*	monosaccharide transporter	leaf senescence	Quirino *et al.*, 2001
Defence and Stress Response Mechanisms	*SENU4*	tomato	PR protein P6	leaf senescence	John *et al.*, 1997
	LSC210	canola	metallothionein	leaf senescence	Buchanan-Wollaston and Ainsworth, 1997
	SR8	carnation	glutathione-*S*-transferase	petal senescence	Meyer *et al.*, 1991
Transcriptional Regulation	*tbzF*	tobacco	bZIP transcription factor	petal senescence leaf senescence	Yang *et al.*, 2001

Functional Category	Gene Name	Plant	Type of Gene Product	Plant Processes	Reference
	AtWRKY6	*Arabidopsis*	WRKY transcription factor	leaf senescence pathogen response	Robatzek and Somssich, 2001
	AtMYB6	*Arabidopsis*	MYB family transcription factor	dark-induced leaf senescence	Lin and Wu, 2004
Signal Transduction	*ZmMPK5*	maize	MAP kinase	leaf senescence cold stress	Berberich *et al.*, 1999
	SARK	bean	receptor-like kinase	leaf senescence	Hajouj *et al.*, 2000

and Brewin, 1996; Puppo *et al.*, 2005). Furthermore, increased activity of acid, serine, and thiol proteases has been detected in senescing alfalfa, soybean, and French bean nodules, respectively (Pfeiffer *et al.*, 1983b; Pladys and Rigaud, 1985; Pladys and Vance, 1993). A recent study comparing the transcriptomes of non-senescent nodules in *Medicago truncatula* has identified gene tags suggesting that the hormones ethylene and jasmonic acid play a positive role in nodule senescence, whereas gibberellins play a negative role (Van de Velde *et al.*, 2006). However, to date, little has been known about gene expression in senescing root nodules.

1.4 Dioxygenases

Dioxygenases can be loosely defined as enzymes which add two atoms of molecular oxygen to a given substrate (DeCarolis and DeLuca, 1994). Two-oxoglutarate-dependent dioxygenases (also called 2-oxoacid-dependent dioxygenases or 2-ODDs) are characterized by their involvement in the type of reaction outlined in Scheme 2. S is the substrate and SO is the oxygenated substrate (Prescott, 1993). In many cases, 2-ODDs carry out hydroxylation reactions and therefore SO would be a hydroxylated product. Ferrous iron and a reducing agent, often ascorbate, are required for maximum activity.



Scheme 2: Catalytic Oxidation of Substrate by 2-ODDs (Prescott, 1993)

2-ODDs are ubiquitous since they are found in microorganisms, animals, and plants (De Carolis and De Luca, 1994; Prescott, 1993). They play many roles in both the primary and secondary metabolism of organisms. Examples of well known 2-ODDs include mammalian prolyl-4-hydroxylase, involved in the post-translational modification of collagen, mammalian ε-*N*-trimethyl-L-lysine hydroxylase and γ-butyrobetaine hydroxylase, two enzymes involved in the synthesis of L-carnitine, a molecule involved in the translocation of fatty acids across the mitochondrial membrane, and isopenicillin N synthase, a protein isolated from various microorganisms which is involved in the synthesis of the β-lactam antibiotic cephalosporin (De Carolis and De Luca, 1994).

1.4.1 Plant 2-ODDs

Several plant 2-ODDs involved in a wide variety of processes have also been described. For example, the flavonoid biosynthetic proteins flavanone 3β-hydroxylase, flavone synthase, flavonol synthase, and leucoanthocyanidin dioxygenase have been isolated and characterized in a wide variety of plant species (Prescott and John, 1996). Many 2-ODDs play important roles in the synthesis of plant hormones. The synthesis of gibberellins, hormones which control many aspects of plant growth and development, is partially carried out by 2-ODDs (Hedden, 1999; Hedden and Phillips, 2000). For example, gibberellin 7-oxidase, gibberellin 20-oxidase, and gibberellin 3β-hydroxylase carry out steps in the synthesis of active gibberellins, whereas gibberellin 2β-hydroxylase produces physiologically inactive gibberellins. ACC oxidase catalyzes the last step in the biosynthesis of the hormone ethylene (Barry *et al.*, 1996; Prescott and John, 1996). Enzymes of importance in alkaloid biosynthesis have also been isolated (Prescott and John, 1996). Examples are hyoscyamine 6β-hydroxylase, which is involved in the

formation of scopolamine, a tropane alkaloid, and desacetoxyvindoline 4-hydroxylase, which is involved in the formation of vinblastine and vincristine (De Carolis and De Luca, 1993; Hashimoto and Yamada, 1987). Scopolamine is an anticholinergic agent and vinblastine and vincristine are well known anticancer agents (Duflos *et al.*, 2002; Yun *et al.*, 1992). Other genes thought to encode 2-ODDs, many of unknown function, have been, and continue to be, isolated. For example, in a recent study of pine drought-responsive genes, an unknown gene whose predicted protein contains motifs common to 2-ODDs was isolated (Dubos and Plomion, 2003). In addition, the sequenced *Arabidopsis* genome has revealed the presence of many genes thought to encode 2-ODDs.

1.5 The Soybean Nodule as a Model for Plant Senescence

To gain a greater understanding of the processes which occur during the last stage of nodule development, and plant senescence in general, the soybean nodule was chosen as a model system. The root nodule proves to be advantageous for studying plant senescence in soybean for several reasons. First, senescence is relatively easy to monitor. It can be followed by the acetylene reduction assay, which measures nitrogenase activity, since nitrogenase activity declines in senescing soybean nodules (Chan, 1995; this study). In addition, soybean nodules are determinate and once they have formed, no further cell division takes place (Hirsch, 1992). Therefore, each nodule provides a developmentally homogeneous population of cells for study. Furthermore, soybean nodule senescence can be induced by the addition of fixed nitrogen or by placing the plant in the dark, providing a population of nodules synchronized in terms of development (Becana and Sprent, 1987). Studying soybean nodule senescence will add to the growing knowledge about plant

senescence and nodule development and how these processes are regulated. Understanding how, when, and why nodules senesce may give clues about how to sustain or improve nitrogen-fixing symbioses.

1.6 Objectives of Research

Since plant senescence displays attributes of PCD, it was hypothesized that soybean nodule senescence would also share many features of PCD, including morphological changes and changes in gene expression. To determine whether any characteristics of plant PCD are recognizable in senescing soybean nodules, light microscopy experiments were initiated to observe any changes in morphology.

Root nodule senescence can be monitored by measuring levels of nitrogenase activity using the acetylene reduction assay or by observing levels of Lb transcripts. These techniques were employed to determine when nodule senescence began during soybean development and to verify that senescence occurred when nitrate or darkness was applied to plants. Any differences in the rate of nodule senescence between nitrate and dark treatments were also observed.

A cDNA named *SAN1A* (*Senescence-Associated Nodulin 1A*) previously isolated by Chan (1995) by screening of a cDNA library constructed from mRNA of senescing soybean nodules. The predicted protein of *SAN1A* contains motifs conserved among plant dioxygenases, and therefore its gene product is thought to encode an enzyme belonging to this superfamily. In addition, Southern analysis had indicated that *SAN1A* belongs to a small, multigene family (Chan, 1995). During the course of this research, two other genes of this family, *SAN1B* and the probable pseudogene *SAN1C*, were isolated. The major

portion of this work focused on determining whether, and how, the expression profiles of *SAN1A* and *SAN1B* differed during natural and induced senescence and across a range of soybean plant tissues as would be expected for duplicated genes that have diverged in regulation. Sequence analysis of the gene family was carried out in order to aid in determination of possible functions for the SAN1 proteins. Functional motifs, gene structure, and the evolutionary relationship of the gene family to other plant dioxygenases were determined. Potential regulatory sequences of the genes were also identified.

CHAPTER 2

MATERIALS AND METHODS

2.1 Growth and Harvest of Soybean (*Glycine max* L. Merrill) Tissues

Seeds of the soybean cultivar Maple Arrow were obtained from Dr. Elroy Cober, Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Ottawa. Prior to germination, seeds were surface sterilized by treatment with 25% bleach twice for 5 minutes, followed by two 5 minute rinses in sterile, distilled water. The seeds were then dispersed onto sterilized trays lined with moistened filter paper (VWR Brand Blotting Paper, VWR Canlab, Montreal, QC) and germinated in the dark for 5-7 days at room temperature (~22°C). During the germination period, the filter paper was moistened with sterile, distilled water once per day. Seeds that did not germinate or that displayed evidence of fungal infection were discarded.

Bradyrhizobium japonicum strain 61A76 was prepared for inoculation by growth in AIEHM liquid medium (0.1% yeast extract, 3.9 mM Hepes, 0.88 mM Na_2HPO_4, 1X salts (1.76 mM Na_2SO_4, 5.6 mM NH_4Cl, 0.7 mM $MgSO_4 \cdot 7H_2O$, 0.019 mM $FeCl_3$, 0.088 mM $CaCl_2 \cdot 2H_2O$), 0.15% arabinose). The culture was grown with shaking at 30°C for 3-4 days until it reached the late log phase and then stored at 4°C until the seedlings were ready for inoculation.

One seedling was placed in each 15 cm pot filled with vermiculite moistened with distilled water, and inoculated with 4 mL of *B. japonicum* culture. The plants were placed into a CMP 3244 growth cabinet (Controlled Environments Limited, Winnipeg, MB) and grown under the following conditions: 16 h, 25°C, 325 μmol m^{-2} s^{-1} PAR day/8 h, 20°C night; relative humidity of 50%.

Up to 12 dpi, plants were watered once daily with 1/8-strength nitrogen-free Hoagland's solution (Hoagland and Arnon, 1950) supplemented with 0.1 mM NH_4NO_3. After 12 dpi, the plants were watered once daily with 1/4-strength nitrogen-free Hoagland's solution, again supplemented with 0.1 mM NH_4NO_3. From 28 dpi onward, the plants were watered once daily with 1/4-strength nitrogen-free Hoagland's solution. Additional nitrogen was no longer required since nodules were fully functional by this stage of development.

Plant tissues harvested for RNA preparation were hand-picked, frozen in liquid nitrogen, and stored at -80°C until needed. Nodules were picked at 7, 10, 15, 20, 24, 28, 32, 36, 40, 45, 50, 55, 60, 65, and 70 dpi for studies of natural senescence. At 7 and 10 dpi, no nodules were visible; therefore, entire root systems were harvested instead of nodules. By 15 dpi, small nodules were visible. For tissue distribution studies, 6 day old seedlings, cotyledons from 10 dpi plants, inoculated roots from 24 dpi plants, stems from 24 dpi plants, young leaves from 24 dpi plants, old leaves from 40 dpi plants, senescing leaves from 65 dpi plants, flowers from 36 dpi plants, nodules from both 24 dpi and 40 dpi plants, and pods from 65 dpi plants were harvested. Young leaves, found at the top of the plant, were light green and not fully expanded, while old leaves, found at lower levels, were dark green and fully expanded, without any signs of yellowing. Senescing leaves, also found at lower levels of the plant, showed obvious signs of yellowing and browning. Whole root systems, stems, and all flowers and pods were harvested from each plant.

For induced senescence and light regulation studies, 0 h corresponded to 8:00 am, 21 dpi with *B. japonicum*. Nodules were harvested from four to five plants every 4 hours for 48 hours for light regulation studies. To induce senescence by nitrate, approximately 1 L of 20 mM NH_4NO_3 was applied to the plants daily (Chan, 1995). To induce

senescence by darkness, the lights were turned off. Following both treatments, nodules were picked from five plants at 0 h, 12 h, 24 h, 36 h, 48 h, 72 h, and 96 h post-treatment.

2.2 Light Microscopy

Root nodules were fixed with 10 mL FAA at 4°C overnight and then transferred to 100 mM NaP_i buffer, pH 7 for storage at 4°C. Following fixation, root nodules were dehydrated under vacuum in a series of 50%, 60%, 70%, 80%, 95%, and 100% ethanol solutions and left overnight in 100% ethanol at room temperature.

Nodules to be embedded in epoxy resin were first immersed in propylene oxide three times for one hour each. They were then placed in one part resin, two parts propylene oxide for one hour. The resin solution consisted of a 2:1:1 ratio by weight of TAAB 812, MNA, and DDSA, respectively (all components except the catalyst) (Marivac Canada Inc., St. Laurent, QC). The nodules were moved to a one part resin, one part propylene oxide solution for one hour and then to a two parts resin, one part propylene oxide solution for another hour. The nodules were then left in fresh resin solution without catalyst overnight. The next day new resin was prepared with ~12 drops of DMP-30 catalyst per 20 g of resin and the nodules incubated in this mixture for one hour. All incubations were carried out at room temperature. Finally, a batch of resin with catalyst was prepared and poured into bullet tubes containing one nodule each. The tubes were incubated overnight at 60°C to allow the resin to solidify.

Resin-embedded root nodules were sectioned on an LKB Ultratome III (LKB Bromma, Uppsala, Sweden) employing a diamond knife with an attached boat. Sections were placed on Colorfrost or Superfrost slides (VWR Canlab, Montreal, QC) in drops of

double-distilled water and heat-fixed to the slides at the lowest temperature setting on a Corning PC-35 hot plate (Corning Incorporated, Corning, NY).

The soybean root nodule sections were stained with a solution of 1% (w/v) toluidine blue O in boric acid, pH 8.0 for three minutes on the lowest temperature setting of the hot plate. Slides were rinsed with double-distilled water and allowed to dry. Cytoseal-60 mounting medium (Electron Microscopy Sciences, Hatfield, PA) was applied to the samples to affix the cover slips to the slides and preserve the stain. Samples were visualized with a Zeiss Axioskop microscope (Carl Zeiss, Inc., Thornwood, NY). Photographs were taken on Kodak MAX film, ISO 400 or 800 (Eastman Kodak Company, Rochester, NY). Negatives were scanned with a SprintScan 35 scanner (Polaroid Canada, Inc., Mississauga, ON) and images enhanced by Adobe Photoshop Version 5 (Adobe Systems Canada, Ottawa, ON) for presentation.

Nodule cells were observed at 0 h, 12 h, 24 h, 36 h, and 48 h during nitrate-induced senescence and at 21 dpi (pre-senescent), 32 dpi, 40 dpi, 50 dpi, and 60 dpi during natural senescence.

2.3 Acetylene Reduction Assays

The nitrogenase (EC 1.18.6.1) activity of soybean nodules was measured by the acetylene reduction assay (Hardy *et al.*, 1968). Root systems were severed at the base of the stem, and placed into a 40 mL jar (Wheaton Scientific, Millville, NJ) that was then sealed with a 13 X 20 mm rubber septum stopper (Wheaton Scientific, Millville, NJ). To initiate the assay, 3 mL of air were removed from the jar by a 5 mL sample lock syringe (Hamilton Company, Reno, NV) and replaced with 3 mL of acetylene gas, prepared by

the reaction of calcium carbide with water. To monitor ethylene production, 5 μL samples were removed with a 50 μL sample lock syringe (Hamilton Company, Reno, NV) at 0, 15, and 30 minutes and injected onto a Varian Star 3400 gas chromatograph (Varian Instrument Group, Walnut Creek, CA) outfitted with an HP-PLOTU column 15 m in length, with an internal diameter of 0.53 mm, and a 20 μm film (Agilent Technologies Canada Inc., Mississauga, ON). The injected samples were subjected to 50°C for 2.5 minutes, followed by ramping to 120°C, followed by 1.5 minutes at 120°C. Samples of 0%, 0.2%, 0.4%, 0.6%, 0.8%, and 1% (v/v) ethylene in nitrogen were run as calibration standards each day that root systems were assayed. Upon completion of the assays, nodules were removed from the root systems, dried overnight at 120°C, and weighed. Nitrogenase activity was calculated from the slope of the curve (ethylene production vs. time) and expressed as nanomoles of ethylene produced per minute per gram of dry weight of root nodule.

To monitor how nitrogenase activity changed throughout nodule development, the root systems of three soybean plants were tested per time point for 7, 10, 15, 20, 24, 28, 32, 36, 40, 45, 50, 55, 60, 65, and 70 dpi. To determine how nitrogenase activity changed during nitrate- and dark-induced senescence, three root systems were assayed at each of the following time points: 0 h, 12 h, 24 h, 36 h, 48 h, 72 h, and 96 h. Data were presented as the means ± standard error for each time point tested.

2.4 Isolation of DNA

2.4.1 Isolation of Genomic DNA

Genomic DNA was extracted from soybean leaves according to protocols outlined by Hattori *et al.* (1987) and Taylor and Powell (1985). The DNA was provided by D. Johnson and N. Mansoorian.

2.4.2 Isolation of Lambda Phage DNA

Titration, growth, isolation of phage, and purification of phage DNA were carried out as described in Maniatis *et al.* (1982). To assess whether contaminating RNA was present in the DNA preparation, agarose gel electrophoresis (Section 2.8.1) was performed. If RNA was detected, *RN*ase was added to a concentration of 100 μg/mL to the phage DNA dissolved in 0.5X OPA buffer (5 mM Tris-acetate, pH 7.5, 5 mM $Mg(OAc)_2$, 25 mM KOAc) and the mixture incubated at 37°C for 60 minutes. The *RN*ase was removed by sequential extractions with phenol, phenol-chloroform, and chloroform (Section 2.4.6). To remove any remaining low molecular weight RNA, 500 μL of 20% PEG 8000 and 1 M NaCl were added to the DNA preparation and it was left overnight at 4°C. Recovery of the DNA was achieved by centrifugation at 7000 rpm for 10 minutes. The supernatant was discarded and the DNA pellet washed twice with 80% ethanol for 10 minutes at room temperature, each wash followed by centrifugation at 7000 rpm for 10 minutes. The DNA pellet was suspended in 500 μL of 1X TE buffer.

2.4.3 Isolation of Plasmid DNA

Ten mL of an overnight culture of DH5αF′ or DH5α cells transformed with plasmid were transferred to a 17 X 100 polypropylene tube and centrifuged at 4500 rpm for 5 minutes at 4°C to pellet the bacterial cells. Plasmid DNA isolation was carried out using the Wizard Minipreps DNA Purification System (Promega Corporation, Madison,

WI) according to the manufacturer's protocol, except that DNA was eluted from the resin with 0.01X TE, pH 8.0, heated to 65°C.

2.4.4 Isolation and Purification of DNA Fragments from Agarose Gels

Restriction fragments or PCR products were isolated from agarose gels for cloning or for re-amplication. DNA was removed from gels by cutting out the area of interest with a scalpel. The gel slice was then placed in a Costar Spin-X Centrifuge Tube Filter spin column with a cellulose acetate filter of 0.45 μm (Corning Incorporated, Corning, NY) and left in a -20°C freezer overnight.

The following day, the gel slice was thawed and the spin column was centrifuged at 14000 rpm for 15 minutes at room temperature. The column flow-through was transferred to a 1.7 mL Eppendorf tube and the volume measured. To concentrate and purify the DNA, either the QIAquick Gel Extraction kit or the QIAquick PCR Purification kit (Qiagen Inc., Mississauga, ON) was used, or a phenol-chloroform extraction was performed (Section 2.4.6). Extractions using the Qiagen kits were carried out according to the manufacturer's protocol. Recovered DNA was stored at 4°C.

2.4.5 Isolation of Plasmid DNA Using the "Easyprep" Boiling Method

To isolate plasmids for screening of clones by restriction digest or colony PCR, the "easyprep" boiling method was used (Berghammer and Auer, 1993).

2.4.6 Phenol-Chloroform Purification of DNA

To purify DNA, 0.5 volumes of buffered phenol (50mM Tris-HCl, pH ~8.0) were added to a solution of DNA and the mixture vortexed. The mixture was left undisturbed at room temperature for 10 minutes, vortexed again, and then centrifuged at 14000 rpm for 3 minutes at room temperature. The aqueous phase was removed to a new tube and 0.5 volumes of 1:1 phenol:chloroform were added. This mixture was also vortexed,

followed by centrifugation at 14000 rpm for 3 minutes at room temperature. The aqueous phase was again removed to a fresh tube and extracted by adding 0.5 volumes of chloroform and vortexing. The resulting suspension was centrifuged at 14000 rpm for 3 minutes. The final aqueous phase was removed to a new tube. To this solution, 0.1 volumes of 3 M NaOAc, pH 5.5 and 2.5 volumes of 95% ethanol were added, and the mixture left overnight at room temperature to allow DNA precipitation.

DNA was recovered by centrifuging the mixture for 10 minutes at room temperature at 14000 rpm. The supernatant was discarded, 1 mL of 80% ethanol added to the pellet, and the mixture placed on a rotary shaker for 10 minutes. The tube was then centrifuged for 5 minutes at room temperature at 14000 rpm. The wash process was repeated and the pellet dried for 5 minutes under vacuum or for 10 minutes at 37°C. The pellet was suspended in an appropriate volume of 0.01X TE, pH 8.0.

2.4.7 Assessment of Quantity, Purity, and Integrity of DNA Samples

The concentration and purity of DNA samples were assessed by UV-spectrophotometry using the GeneQuant RNA/DNA Calculator (Pharmacia Biotech, Amersham Biosciences, Baie d'Urfe, QC). Quality of DNA was also verified by performing a restriction digest (Section 2.5). If the ratio of A_{260}/A_{280} was between 1.75 and 2.0, the preparation was considered pure. To assess the integrity of the DNA, agarose gel electrophoresis (Section 2.8.1) was employed.

2.5 Restriction Digests of DNA

Restriction digest reactions were typically composed of 100 ng -1 μg of DNA, 1X or 2X OPA buffer (1X OPA: 10 mM Tris-acetate, pH 7.5, 10 mM $Mg(OAc)_2$, 50 mM

KOAc; 2X OPA: 20 mM Tris-acetate, pH 7.5, 20 mM $Mg(OAc)_2$, 100 mM KOAc), depending on the enzyme, and 10-20 U of the restriction enzyme or enzymes (Pharmacia Biotech, Amersham Biosciences, Baie d'Urfe, QC). Reactions were carried out at 37°C for 1-2 hours or overnight. Results of digestion were assessed by agarose gel electrophoresis (Section 2.8.1).

2.6 Cloning of DNA

2.6.1 Ligations

2.6.1.1 Standard Ligations

DNA fragments generated by restriction digest were ligated into pUC8, pUC19 (Pharmacia Biotech, Amersham Biosciences, Baie d'Urfe, QC), or pGEM4Z (Promega Corporation, Madison, WI) vectors. Vectors were prepared for ligation by restriction with the appropriate enzyme or enzymes (Section 2.5), but the amount of DNA was scaled up to ensure a final concentration of 50-100 ng/μL. Enzymes were inactivated by heating the mixture to 65°C or 85°C. Following enzyme inactivation, the vector was treated with calf intestinal alkaline phosphatase (Promega Corporation, Madison, WI) to prevent self-ligation. Phosphatase treatment was carried out following the manufacturer's protocol.

Intermolecular ligations were carried out by combining insert and vector in a molar ratio of 1:1 to 10:1 insert:vector, 1X T4 DNA Ligase buffer (5X is 250 mM Tris-HCl, pH 7.6, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25%(w/v) PEG 8000), and 1 U of T4 DNA ligase (Invitrogen Canada Inc., Burlington, ON). The volume was brought to 20 μL with sterile, distilled water and the mixture incubated overnight at 15°C. Intramolecular (self) ligations were carried out by combining 1 μL of digested plasmid,

1X T4 DNA Ligase buffer, and 1 U of T4 DNA ligase. Sterile, distilled water was added to bring the final volume to 20 μL and the reaction mixture was incubated overnight at 15°C.

2.6.1.2 Ligation of PCR Products

PCR products generated by *Taq* polymerase were cloned using a TA-cloning procedure. Products were ligated into the pGEM-T or pGEM-T Easy vector (Promega Corporation, Madison, WI) according to the manufacturer's instructions.

2.6.2 Transformation of Bacteria with Plasmid

Transformations were generally carried out by the Hanahan procedure with the DH5αF′ strain of *Escherichia coli* (Hanahan, 1983) or with subcloning efficiency competent DH5α cells (Invitrogen Canada Inc., Burlington, ON) according to the manufacturer's protocol.

Dilutions of transformed cells were plated on LB agar plates supplemented with 50 μg/mL ampicillin and 0.002% X-gal. The plates were incubated at 37°C overnight. A typical positive control consisted of cells transformed with 10 ng of pUC19 DNA; a negative control consisted of untreated cells.

2.7 Isolation of Total RNA

Isolation of total RNA from soybean tissues was carried out by the GITC–phenol–chloroform method (Chomczynski and Sacchi, 1987), including a high salt precipitation step outlined by Chomczynski and Mackey (1995). For root and seedling extractions, volumes of extraction components were reduced by half per gram of fresh tissue to maximize yield. All glassware, plastics, and solutions used for extraction were *RN*ase-

free. To measure the concentration and purity of RNA samples, UV-visible spectrophotometry using a GeneQuant RNA/DNA Calculator (Pharmacia Biotech, Amersham Biosciences, Baie d'Urfe, QC) was employed. If the ratio of A_{260}/A_{280} was between 1.8 and 2.0, the preparation was considered pure. To assess the integrity of the isolated RNA, agarose gel electrophoresis (Section 2.8.2) was employed. For long-term storage, RNA samples were kept at -80°C or aliquots were precipitated in 0.1 volumes of 3 M NaOAc, pH 5.5 and 2.5 volumes of 95% ethanol and stored at -20°C or -80°C.

2.8 Agarose Gel Electrophoresis

2.8.1 DNA Gel Electrophoresis

DNA fragments generated by restriction digests or PCR were separated by electrophoresis in agarose gels prepared in 0.5X TBE buffer. Prior to loading a sample, 10X DNA loading buffer (25% Ficoll 400, 0.1% SDS, 1X TE, pH 8.0, 0.001% bromophenol blue, 0.001% xylene cyanol) was added to a final concentration of 1X. Agarose concentrations ranged from 0.8-2% (w/v) in 0.5X TBE, depending on the sizes of fragments to be resolved. Ethidium bromide was added to the gel to a final concentration of 0.5 μg/mL for visualization of the DNA. DNA ladders used as size markers included λ*Hin*dIII (Pharmacia Biotech, Amersham Biosciences, Baie d'Urfe, QC) and pφX/HaeIII (prepared by D. Johnson) for higher molecular weight molecules (>1300 bp) and lower molecular weight molecules (<1300 bp), respectively. The separated DNA was visualized on a UV transilluminator (Ultra-Violet Products Inc., San Gabriel, CA). Photographs of the gels were obtained by the ChemiDoc system (Bio-Rad Laboratories, Hercules, CA) or by a Polaroid MP-4 Land Camera (Polaroid Canada Inc., Mississauga, ON).

2.8.2 RNA Gel Electrophoresis

To carry out electrophoretic separation of RNA, 1.5% (w/v) agarose gels were prepared in 1X MOPS buffer (200 mM MOPS, 50 mM NaOAc, 8.6 mM EDTA, pH 7) and 10% formaldehyde. All glassware, plastics, and solutions used to carry out RNA gel electrophoresis were *RN*ase-free. Between 2.5 and 20 μg of total RNA were loaded into each lane of the gel. Per μL of RNA sample, 2.25 μL of RNA loading buffer (50% (w/v) formamide, 1X MOPS, 5% (w/v) glycerol, 17.5% (w/v) formaldehyde, 0.6 mM EDTA, 2.5 mg/mL bromophenol blue, 5 μg/mL ethidium bromide) were added. Prior to loading, the samples were heated at 65°C for 20 minutes and then chilled on ice. Running buffer (1X MOPS) was prepared in sterile, distilled water. Once the RNA samples were loaded, the RNA was run into the gel at 100 volts until the samples had cleared the wells. The voltage was then reduced to allow migration of RNA through the gel at a rate of approximately 1 cm/h. To visualize the RNA, the gel was examined on a UV transilluminator (Ultra-Violet Products Inc., San Gabriel, CA) and a photograph taken with either the ChemiDoc system (Bio-Rad Laboratories, Hercules, CA) or a Polaroid MP-4 Land Camera (Polaroid Canada Inc., Mississauga, ON).

2.9 Transfer of DNA and RNA from Gel to Membrane

2.9.1 Transfer of DNA to Membrane

After electrophoretic separation (Section 2.8.1), DNA could be transferred to a membrane for subsequent hybridization (Section 2.12.1). Gels containing DNA to be transferred were treated twice with denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 20 minutes each. The gel was then rinsed with distilled water and treated twice with

neutralization buffer (3 M NaOAc, pH 5.0) for 20 minutes each. For small gels (25-50 mL), treatment times were shortened to 10 minutes. If large fragments (>5 kb) of DNA were to be transferred, depurination in 0.25 M HCl was carried out twice for 5 minutes each prior to the denaturation step. DNA was transferred to Hybond-N (Amersham Biosciences UK Ltd., Buckinghamshire, UK) membrane by the capillary method using 20X SSC as the transfer buffer.

The next day, the transfer apparatus was disassembled and the DNA UV cross-linked to the membrane for 5 minutes using an intermediate wavelength UV lamp (Model XX-15 Blak-Ray, Ultra-Violet Products, Inc., San Gabriel, CA). To confirm that transfer had occurred, the gel was stained with a 0.5 μg/mL ethidium bromide solution and visualized on a UV transilluminator (Ultra-Violet Products Inc., San Gabriel, CA). Membranes were stored at 4°C.

2.9.2 Transfer of RNA to Membrane

If RNA was to be transferred to a membrane after electrophoretic separation (Section 2.8.2), the gel was washed twice for 20 minutes with distilled water to remove formaldehyde and ethidium bromide, and then washed for 20 minutes with 2X SSC. The method of capillary transfer described in Section 2.9.1 was employed. For RNA transfers, GeneScreen+ (NEN Life Science Products, Boston, MA) membrane was used. The membrane was briefly immersed in distilled water and then in 2X SSC for 10 minutes prior to use.

2.10 Dot Blots

Dot blots were included with Southern and Northern hybridizations to determine the detection limit of each labelled probe. Serial dilutions of probe DNA were prepared in 20X SSC to obtain concentrations of 10 ng/μL, 1 ng/μL, 100 pg/μL, and 10 pg/μL. Hybond-N membrane (Amersham Biosciences UK Ltd., Buckinghamshire, UK) was divided into a series of 1 X 4 grids. One μL of each DNA dilution was applied to each box in a single grid. Once the solutions had dried, the membrane was treated with denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 5 minutes, followed by a 5 minute treatment with neutralization solution (3 M NaOAc, pH 5.0). A final 5 minute treatment with 2X SSC was also carried out. The DNA was UV cross-linked to the membrane for 5 minutes using an intermediate wavelength UV lamp (Model XX-15 Blak-Ray, Ultra-Violet Products, Inc., San Gabriel, CA). The dot blots were stored at 4°C.

2.11 Radioactive Probe Preparation

DNA to be used for probe preparation was obtained by PCR amplification of cloned cDNA (Section 2.13), followed by purification using the QIAquick PCR Purification kit (Qiagen Inc., Mississauga, ON). Lb and 18S probes were amplified from the clones pCT-2 and 18S soy, respectively, using M13 universal primers and the conditions outlined for colony PCR (Section 2.13.1). The full-length *SAN1A* cDNA probe was amplified from the *SAN1A* CDS clone with the *SAN1A* CDS-specific primers (Table III). PCR conditions consisted of 30 cycles of 30 seconds at 94°C, 30 seconds at 60°C, followed by 1.5 minutes at 72°C. An initial denaturation at 94°C for 5 minutes and final extension at 72°C for 10 minutes were also included in the program. Probes for relative RT-PCR experiments were prepared from clones of the *SAN1A*, *SAN1B*, and 18S products

Table III: List of PCR Primers

Primers used during the course of this work are listed with respect to the protocol in which they were employed. The sequences of the primers and their annealing sites are also included.

CDS, coding sequence; MCS, multiple cloning site.

Experimental Procedure	Name	Sequence (5′-3′)	Annealing Site
Colony PCR	M13 Universal F M13 Universal R	CGCCAGGGTTTTCCAGTCACGAC AGCGGATAACAATTTCACACAGGA	Adjacent to MCS of pUC and pGEM vectors.
Cycle Sequencing	Universal F (IRD 800) Universal R (IRD 700)	CACGACGTTGTAAAAACGA GGATAACAATTTCACACAGG	Adjacent to MCS of pUC and pGEM vectors.
5′ RACE			
Gene-Specific Primers	*SAN1A*-5RGSP1 *SAN1A*-5RGSP2 *SAN1A*-5RGSP3	CCGTTTACGTCTCACT GTCCTGTGCAAGAATGGTCAAG GACGTTCTTGGTGTGTTCTGTG	(−) 717 bp on *SAN1A* CDS. (−) 678 bp on *SAN1A* CDS. (−) 321 bp on *SAN1A* CDS.
PCR of dCTP-tailed cDNA	Abridged Anchor Primer	GGCCACGCGTCGACTAGTAC$(G)_{16}$	C-tail of first strand cDNA.
Nested PCR	Abridged Universal Anchor Primer	GGCCACGCGTCGACTAGTAC	Complementary sequence of abridged anchor primer.
3′ RACE			
Gene-Specific Primers	*SAN1A*-3RGSP1 *SAN1A*-3RGSP2	CCTGCACACGACACCAAAGTAA AGGCCATACAATTGGGGCAAGT	(+) 892 bp on *SAN1A* CDS. (+) 958 bp on *SAN1A* CDS.
First Strand cDNA Synthesis	Adapter Primer	GGCCACGCGTCGACTAGTAC$(T)_{17}$	PolyA tail of mRNA.

Experimental Procedure	Name	Sequence (5′-3′)	Annealing Site
Nested PCR Anchor Primer	Abridged Universal	GGCCACGCGTCGACTAGTAC	Complementary sequence of adapter primer.
***SAN1A* Full-Length Probe Preparation**	*SAN1A*CDSF	TAGGATCCGCCATGGGAGAGGTTGAC	(+) 11 bp upstream of *SAN1A* start codon.
	*SAN1A*CDSR	TCTCGAGGCGCTGGAATCCTGCTAA	(−) 1070 bp from *SAN1A* start codon in 3′ UTR.
Extension of *SAN1A* Genomic DNA Sequence	T7	GCAACTCGTGAAAGGTAGGC	T7 area of left arm of EMBLλ3 SP6/T7 vector.
	SAN1A-1F	AGTGAAACCAACCCCAAATGCTTA	Coding region of *SAN1A* (+) 1142 bp from start codon.
	SAN1A-2F	CAGGTGCCAGTGCAATACCATATT	Coding region and second intron of *SAN1A* (+) 1191 bp from start codon.
RAP-PCR	A3	AATCTAGAGCTCTCCTGG	Not applicable.
	B3	CATACACGCGTATACTGG	Not applicable.
	C3	CCATGCGCATGCATGAGA	Not applicable.
Relative RT-PCR			
SAN1A Primers	*SAN1A*F	CCACTCTTCAGGGTTGTAACAC	(+) 460 bp on *SAN1A* CDS.
	*SAN1A*R	GCTCCTCCAAAGGCTTTACTTT	(−) 907 bp on *SAN1A* CDS.
SAN1B Primers	*SAN1B*F	GACTTGAAGTGAGACCTAAAG	(+) 710 bp on *SAN1B* CDS.
	*SAN1B*R	TCAACGCTGCTACATACCAATG	(−) 1123 bp from *SAN1B*

Experimental Procedure	Name	Sequence (5′-3′)	Annealing Site
18S Primers	QuantumRNA Universal 18S Internal Standards (Ambion Inc., Austin, TX)	Sequence unknown (proprietary).	Unknown site on soybean 18S.
***RT-PCR-SAN1C* Tissue Distribution**	*SAN1C*F *SAN1C*R	CAACCTGAAGACATTCCTGTGA GGTTGGTTTCACTCTTGATCAG	(+) 64 bp from *SAN1C* start codon. (−) 732 bp from *SAN1C* start codon.

using the PCR conditions outlined in Section 2.14.5, Step 5, except that the number of cycles was increased to 30 to maximize the amount of product obtained. *SAN1C* probe was prepared from the *SAN1C* clone using the PCR conditions described in Section 2.14.2.1.

Preparation of radioactive probes was carried out by the multiprime labelling method. Between 25 and 50 ng of DNA to be labelled were combined with sterile, distilled water to a volume of 12 μL in a 1.7 mL Eppendorf tube. This mixture was boiled for 5 minutes and then chilled on ice. To this solution, 6 μL of 5X Random Primer Extension buffer (250 mM Tris-HCl, pH 7.4, 25 mM $MgCl_2$, 50 mM DTT, 2 mg/mL BSA, 12.5 U/mL random hexamers), 6 μL of 5X dXTP Mix (25 mM Tris-HCl, pH 7.4, 100 μM dXTP stock), 0.5 U of the large fragment of DNA polymerase I (Invitrogen Canada Inc., Burlington, ON), and 50 μCi of α-^{32}P dCTP (Amersham Biosciences Corp., Piscataway, NJ) were added. The reaction was incubated for 45 minutes at room temperature before being terminated by the addition of 70 μL of Stop Buffer (50 mM Tris-HCl, 10 mM EDTA, 100 mM NaCl, 0.1% SDS, pH 8.0).

To separate unincorporated nucleotides from the reaction products, the solution was passed through a 1 mL column of Sephadex G-50 (Pharmacia Biotech, Amersham Biosciences, Baie d'Urfe, QC) and purified probe collected in a 1.7 mL Eppendorf tube. Success of the labelling reaction was assessed by monitoring the probe with an Eberline Model E-140 Geiger counter (Eberline Services, Lionville, PA). Before adding the prepared probe to a hybridization reaction, it was denatured by boiling for 5 minutes followed by chilling on ice.

2.12 Southern and Northern Hybridization

2.12.1 Southern Hybridization

The standard solution for Southern hybridization experiments consisted of 6X SSC, 50 mM NaP_i, pH 7, 5 mM EDTA, 5X Denhardt's solution (0.1% Ficoll 400, 0.1% PVP, 0.1% BSA), 0.2 mg/mL herring sperm DNA, and 0.2% SDS. Stringency was adjusted if required. Pre-hybridization of the membrane was carried out for 1-4 hours with 4 mL of solution per 100 cm^2 of membrane. Denatured probe was added directly to this solution and hybridization allowed to occur overnight at 65°C.

Membranes were washed on the following day. One wash of 20 minutes with 2X SSC, 0.1% SDS at 65°C and two washes of 20 minutes each with 0.1X SSC, 0.1% SDS at 65°C were carried out. For each wash, 50 mL of solution was used. The washed membranes were exposed to X-OMAT AR film (Eastman Kodak Company, Rochester, NY), B-Plus film (Wolf X-Ray Corporation, West Hempstead, NY), or to a phosphor screen (Bio-Rad Laboratories, Hercules, CA) for varying amounts of time.

2.12.2 Northern Hybridization

Prior to Northern hybridization, the membrane was pre-washed with 50 mL of 0.1X SSC, 0.1% SDS for 20 minutes at 65°C. After discarding this solution, hybridization solution consisting of 10% dextran sulfate, 0.1% SDS, 6X SSC, 100 μg/mL herring sperm DNA, and 5X Denhardt's solution (0.1% Ficoll 400, 0.1% PVP, 0.1% BSA), adjusted to pH 7, was added in the amount of 10 mL of solution per 100 cm^2 of membrane. Dextran sulfate was omitted from the hybridization solution if the target RNA was that of a highly expressed gene, such as Lb or 18S. Pre-hybridization was carried out

for 3-4 hours at 65°C prior to adding denatured probe to the solution. Hybridization was carried out overnight at 65°C.

The next day, the membranes were washed. Two 15 minute washes in 2X SSC at 65°C were carried out followed by two 30 minute washes in 2X SSC, 0.1% SDS at 65°C. A final 15 minute wash in 0.1X SSC, 0.1% SDS at 65°C was performed. For each wash, 50 mL of solution were employed. The membranes were exposed to X-OMAT AR film (Eastman Kodak Company, Rochester, NY), B-Plus film (Wolf X-Ray Corporation, West Hempstead, NY), or to a phosphor screen (Bio-Rad Laboratories, Hercules, CA) for varying amounts of time.

2.13 Polymerase Chain Reaction

PCR was generally carried out in a reaction volume of 20 µL. A typical reaction was composed of 1X PCR buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl), 0.2 µM each of forward and reverse primers, 1-2.5 mM of $MgCl_2$, 200 µM each of dGTP, dCTP, dATP, and dTTP, 0.5 U of *Taq* polymerase (Invitrogen Canada Inc., Burlington, ON), sterile, distilled water to a final volume of 18 or 19 µL, and 1-2 µL of the DNA to be amplified. PCR was run on an MJ Research PTC-100 (MJ Research Inc., Waltham, MA), a Perkin-Elmer GeneAmp PCR System 9600 (PerkinElmer Inc., Boston, MA), Eppendorf Mastercycler, or Mastercycler ep Gradient (Eppendorf, Hamburg, Germany) thermocycler. An initial denaturation of 5 minutes at 94°C was included in each PCR program. A cycle consisting of a 94°C denaturation step for 30 seconds, an annealing step between 30 seconds and 1 minute at a primer-appropriate temperature, and an extension step at 72°C between 30 seconds and 3 minutes was repeated between 20 and

40 times depending on the abundance of the target being amplified. The length of the extension step depended on the expected length of the target; generally, *Taq* polymerase adds nucleotides to a template at a rate of approximately 35-100 nucleotides per second (Wittwer and Garling, 1991). A final extension at 72°C for 5-10 minutes was also carried out. For all PCRs, a positive control containing the target sequence and a negative control with no added DNA template were performed. The outcome of PCR was assessed by gel electrophoresis (Section 2.8.1). Primer design was aided by primer analysis features of DNAMAN Version 4.22 (Lynnon BioSoft, Vaudreuil, QC).

2.13.1 Colony PCR

Colony PCR was carried out following transformation to screen clones for inserts. Bacteria were picked from patches of clones using a sterile toothpick, placed in 50 µL of sterile, distilled water, and mixed. The suspension was then placed at 65°C for 20 minutes. Following lysis, the suspensions were centrifuged at 14000 rpm for 10 minutes at room temperature. One µL of the supernatant was used for PCR amplification. For colony PCR, the reaction composition was the same as that described in Section 2.13, except that the forward and reverse primers were universal M13 primers (Table III), $MgCl_2$ was used at a concentration of 2 mM, and water was added to a volume of 19 µL. The PCR program consisted of 30 cycles of 30 seconds at 94°C, 30 seconds at 59°C, followed by 1.5 minutes at 72°C. An initial denaturation at 94°C for 5 minutes and final extension at 72°C for 10 minutes were also included in the program.

2.14 RT-PCR and Related Techniques

RT-PCR and the related techniques of 5′ and 3′ RACE, RAP-PCR, and relative RT-PCR were carried out during the course of this research. All glassware, plastics, and solutions used to carry out RT reactions were *RN*ase-free.

2.14.1 *DN*ase Treatment of RNA Samples

To remove genomic DNA from RNA samples prior to RT-PCR experiments, *DN*ase I treatment (Invitrogen Canada Inc., Burlington, ON) was carried out according to manufacturer's instructions.

2.14.2 Standard RT-PCR

For each reaction, 1-5 µg of total RNA were combined with 0.5 µg of oligod$T_{(13-18)}$ primer (Invitrogen Canada Inc., Burlington, ON) and DEPC water to a volume of 13 µL. This mixture was incubated at 70°C for 10 minutes and then chilled on ice for 5 minutes. The contents were collected by briefly centrifuging the tubes. To each of the tubes, 2 µL of 10X PCR buffer, 1 µL of 50 mM $MgCl_2$, 1 µL of 10 mM dXTPs, and 2 µL of 0.1 M DTT (all products supplied by Invitrogen Canada Inc., Burlington, ON) were added. The contents were mixed and then briefly centrifuged. Finally, 200 U of SuperScript II *RN*ase H^- RT (Invitrogen Canada Inc., Burlington, ON) were added to the reaction mixture. The reaction was carried out at 42°C for 60 minutes and then terminated by incubation at 70°C for 10 minutes. Reactions without added RT were carried out as negative controls to test for the presence of genomic DNA. Without these controls, false positives could result following subsequent PCR amplification, since both the genomic DNA and cDNA could be amplified. The prepared cDNA was stored at -20°C. PCR was carried out as described in Section 2.13. For each 20 µL PCR, 1-2 µL of the RT reaction were amplified.

2.14.2.1 Tissue Distribution of *SAN1C* mRNA

RT-PCR was used to detect *SAN1C* mRNA in various soybean tissues. Standard conditions were employed, except that random hexamers were used to prepare cDNA (Section 2.14.5, Step 2). Although the data collected were not quantitative, Northern analysis (Section 2.12.2) of 18S was carried out to normalize the RNA concentrations of the samples prior to RT. PCR was carried out on 2 μL of cDNA. The primers used to amplify *SAN1C* can be found in Table III. PCR conditions consisted of 30 cycles of 30 seconds at 94°C, 30 seconds at 55°C, followed by 1.5 minutes at 72°C. An initial denaturation at 94°C for 5 minutes and final extension at 72°C for 10 minutes were also included in the program. After PCR amplification, the products were analyzed by gel electrophoresis (Section 2.8.1) and the products transferred to Hybond-N membrane (Amersham Biosciences UK Ltd., Buckinghamshire, UK) for Southern hybridization with the *SAN1C* probe. Standard hybridization conditions were used (Section 2.12.1). The results were obtained by exposure to a phosphor screen (Bio-Rad Laboratories, Hercules, CA).

2.14.3 5′ and 3′ RACE

All buffers, enzymes, and standard primers used in 5′ and 3′ RACE were obtained from Invitrogen Canada Inc., Burlington, ON, unless otherwise stated.

5′ RACE

Step 1: First Strand cDNA Synthesis

For first strand cDNA synthesis, 100-500 ng of total RNA were combined with 2.5 pmol of a gene-specific primer (GSP1) (Table III), an antisense primer designed to anneal ~300 bp from the 5′ end of the mRNA, and the reaction prepared and carried out according to manufacturer's instructions. Negative controls without added RT were

included to ensure that the products did not represent genomic DNA. To remove any residual RNA, 4 U of *RN*ase H were added and the mixture incubated at 37°C for 20 minutes. The cDNA was stored at -20°C until needed.

Step 2: Purification of cDNA

The cDNA was purified from the RT reaction components using the QiaQuick PCR Purification System (Qiagen Inc., Mississauga, ON). The manufacturer's protocol was modified slightly. One additional wash with 375 µL PE buffer was carried out followed by a wash with 375 µL of 70% ethanol. In addition, elution was carried out with sterile, distilled water, pH 7 instead of EB buffer.

Step 3: Terminal Deoxynucleotidyl Transferase (TdT) Tailing of cDNA

The 3′ end of the cDNA was tailed by the addition of dCTP to provide a binding site for the abridged anchor primer (Table III) for subsequent PCR amplification. The tailing reaction was carried out according to the manufacturer's instructions.

Step 4: PCR of dCTP-Tailed cDNA

PCRs were set up as described in Section 2.13, except that the primers used were the abridged anchor primer and a second gene-specific primer (GSP2) (Table III), which was nested with respect to GSP1. $MgCl_2$ was added to a concentration of 1.5 mM and 2 µL of dC-tailed cDNA were added to each reaction. The PCR program consisted of 30 cycles of 30 seconds at 94°C, 30 seconds at 60°C, followed by 1 minute at 72°C. An initial denaturation step at 94°C for 5 minutes and a final extension at 72°C for 5 minutes were included in the program. A second nested PCR was carried out to increase the likelihood that the RACE product represented the intended target. Primers employed were a third gene-specific primer (GSP3) (Table III), nested with respect to GSP2, and the abridged universal amplification primer (AUAP) (Table III). The first PCR product was

diluted 100-fold and 2 μL were amplified. The PCR conditions and program were the same as for the first round of amplification.

3' RACE

Step 1: First Strand cDNA Synthesis

First strand cDNA synthesis was carried out on up to 5 μg of total RNA with an oligodT adapter primer (Table III) according to manufacturer's instructions. Negative controls without added RT were included to ensure that the products did not represent genomic DNA. After the reaction was completed, 4 U of *RN*ase H were added to the tube and the mixture incubated at 37°C for 20 minutes. The cDNA was stored at -20°C.

Step 2: PCR of cDNA

PCR was carried out as described in Section 2.13, except that the primers employed were the AUAP and a gene-specific primer (GSP1) (Table III). $MgCl_2$ was added to a concentration of 1.5 mM. To each PCR, 2 μL of cDNA were added. The PCR program consisted of 30 cycles of 30 seconds at 94°C, 30 seconds at 60°C, followed by 1 minute at 72°C. An initial denaturation step at 94°C for 5 minutes and a final extension at 72°C for 5 minutes were included in the program. A nested PCR was also carried out to increase the likelihood that the RACE products represented the intended target. The primers employed were a second gene-specific primer (GSP2) (Table III), nested with respect to GSP1, and the AUAP. The primary PCR product was diluted 100-fold and 2 μL were amplified. The PCR conditions and program were the same as those used in the first round of amplification.

Analysis of 5' and 3' RACE Products

All RACE products were analyzed by gel electrophoresis (Section 2.8.1) and cloned by the TA-cloning method (Section 2.6.1.2). Transformation was carried out by

the Hanahan method (Section 2.6.2). The largest clones, determined by colony PCR (Section 2.13.1), were selected for plasmid preparation (Section 2.4.3) and cycle sequencing (Section 2.15) to determine the sequence of the DNA amplified by RACE.

2.14.4 Differential Display RT-PCR

A modified version of differential display RT-PCR, RAP-PCR (Stratagene, La Jolla, CA), was employed to isolate additional senescence-associated genes from soybean nodules. The RAP-PCR primers used in these experiments were A3, B3, and C3 (Table III).

Step 1: RT Reactions

One μg of total RNA, 1 μL of 25 mM RAP-PCR primer (Stratagene, La Jolla, CA), and DEPC H_2O were combined to a final volume of 9.5 μL. This mixture was incubated at 70°C for 10 minutes, chilled on ice, and then briefly centrifuged to collect the contents. To this mixture, 4 μL of 5X First Strand buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM $MgCl_2$), 2 μL of 0.1M DTT, and 2.5 μL of 10 mM dXTP were added. The mixture was then incubated at 42°C for 1 to 5 minutes, 200 U of SuperScript II *RN*ase H^- RT (Invitrogen Canada Inc., Burlington, ON) added, and the resulting reaction mixture incubated for 50 minutes at 42°C. Termination of the reaction was carried out by incubation at 70°C for 10 minutes. Negative controls without added RT were also carried out to test for the presence of genomic DNA. Reaction products were stored at -20°C.

Step 2: PCR

PCR was carried out in a volume of 50 μL. For each reaction, 29.15 μL of sterile, distilled H_2O, 5 μL of 10X PCR buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 1.5 μL of 50 mM $MgCl_2$, 0.25 μL of 10 mM dXTPs, 2 μL of the 25 μM RAP-PCR primer

employed in the RT reaction, and 0.2 µL of *Taq* polymerase (Invitrogen Canada Inc., Burlington, ON) were combined. To this mixture, 10 µL of 10X diluted cDNA and 1.2 µCi of [α-^{32}P] dCTP (Amersham Biosciences Corp., Piscataway, NJ) were added. The PCR was carried out as follows: denaturation for 5 minutes at 94°C, one low stringency cycle consisting of denaturation for 1 minute at 94°C, annealing for 5 minutes at 36°C, and a 5 minute extension at 72°C, 40 high stringency cycles consisting of denaturation for 1 minute at 94°C, annealing for 2 minutes at 55°C, and a 2 minute extension at 72°C, and a final extension for 10 minutes at 72°C.

Step 3: Sequencing Gels

The RAP-PCR products were separated on a 4% acrylamide gel prepared by combining 33.6 g of urea, 8 mL of a 40% acrylamide stock (19:1 acrylamide:bisacrylamide), 43 mL of distilled water, and 4 mL 10X TBE. This mixture was filtered through blotting paper (VWR Brand Blotting Paper, VWR Canlab, Montreal, QC) before adding 56 µL of TEMED and 560 µL 10% ammonium persulfate. The acrylamide solution was poured between two clean glass plates and polymerization allowed to occur overnight. The next day, the gel was pre-run for 2 hours at 20 mamp. Five µL of stop buffer (80% deionized formamide, 50 mM Tris-HCl, pH 8.3, 1mM EDTA, 0.1% (w/v) xylene cyanol, 0.1% bromophenol blue) were combined with 5 µL of the PCR products and the mixture heated at 80°C for 10 minutes prior to loading. Three µL of each sample were loaded in duplicate and the gel run at 1.8 kV for approximately 6 hours. After the run was completed, the gel was transferred to filter paper (VWR Brand Blotting Paper, VWR Canlab, Montreal, QC) and dried on a gel dryer (Drygel Sr. Slab Gel Dryer Model SE1160, Hoeffer Scientific Instruments, San Francisco, CA). The dried

gel was exposed to BioMax MR film (Eastman Kodak Company, Rochester, NY) for various amounts of time.

Step 4: Isolation and Cloning of Bands

Bands of interest were cut from the dried gel, immersed in 50 µL of 1X TE, pH 8.0 in a 1.7 mL Eppendorf tube, and placed at 65°C for one hour, followed by an overnight incubation at room temperature. The following day, the mixture was centrifuged for 10 minutes at 14000 rpm at room temperature and the supernatant removed to a fresh tube. PCR reamplification of the DNA was carried out with the appropriate RAP-PCR primer. The PCR conditions were the same as those used for the original reaction, except that the low stringency cycle was eliminated. The resulting DNA was cloned using the TA-cloning method (Section 2.6.1.2) and transformation was carried out as described in Section 6.2. Colony PCR (Section 2.13.1) and gel electrophoresis (Section 2.8.1) were used to determine which clones contained inserts of the expected size. Clones of the appropriate size were analyzed by sequencing (Section 2.15).

2.14.5 Relative RT-PCR

To monitor the expression of genes whose message levels were below the detection limit of Northern analysis, relative RT-PCR (Ambion Inc., Austin, TX) was employed.

Step 1: Primer Design

Primers were designed to amplify small regions of the genes *SAN1A* and *SAN1B*. Since relative RT-PCR is a multiplex PCR method, it was important to target gene regions which would give product sizes similar to, but still distinguishable from, the 18S control. The 18S product was 317 bp; therefore, primers were designed to ensure that the *SAN1A*

or *SAN1B* products were no more than 200 bp larger or smaller than that of the 18S product (QuantumRNA Internal Standards Kit, Ambion Inc., Austin, TX). *SAN1A* and *SAN1B* primer sequences are listed in Table III. The specificity of the primers for each of the target genes was verified by PCR with each primer set on each of the *SAN1A*, *SAN1B*, and *SAN1C* sequences. PCR with the *SAN1A* primers only resulted in a product when *SAN1A* was the target and PCR with the *SAN1B* primers only resulted in a product when *SAN1B* was the target (data not shown). Therefore, the *SAN1A* and *SAN1B* primers were considered to be specific for their respective targets.

Step 2: cDNA Preparation for Relative RT-PCR Experiments

Reverse transcription was carried out as described in Section 2.14.2, except that 0.5 μg of random hexamers were used instead of oligod$T_{(13\text{-}18)}$ primer. All RNA was *DN*ase-treated and 2.5 μg of total RNA were used in each RT reaction.

Step 3: Determination of the Linear Range of Amplification of SAN1A and SAN1B

To allow relative quantitation, it was necessary to determine the linear range of amplification of the target sequences. The linear range refers to the cycle numbers over which the accumulation of products occurs at a constant rate (QuantumRNA Internal Standards Kit, Ambion Inc., Austin, TX). In other words, if amount of product is plotted versus cycle number, the linear range is the region of the curve which is a straight line. In order to determine the linear range, cDNA from a time point at which the target mRNA of interest is abundant was employed.

A cocktail of PCR reagents plus cDNA was prepared and split into a sufficient number of tubes to allow an aliquot to be removed every two cycles between 10 and 40 cycles. Each aliquot of the reaction was composed of 2 μL each of 2 μM forward and

reverse primers, 1X PCR buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl), 2 µL of 2 mM dXTPs, 1-2.5 mM of $MgCl_2$, 0.5 U of *Taq* polymerase (Invitrogen Canada Inc., Burlington, ON), 2 µL of cDNA, and sterile, distilled H_2O to a final volume of 20 µL. PCR conditions consisted of an initial denaturation at 94°C for 5 minutes, followed by 40 cycles of a 30 second denaturation step at 94°C, a 30 second annealing step carried out at a primer-appropriate temperature, and a 1 minute extension at 72°C. Once the PCR was running, one aliquot was removed after every two cycles beginning at cycle 10.

The PCR products were run on a 1.25 % nondenaturing agarose gel (Section 2.8.1) and transferred to Hybond-N membrane for Southern hybridization with the target gene product. Standard hybridization conditions were used (Section 2.12.1). The results were obtained by exposure to a phosphor screen (Bio-Rad Laboratories, Hercules, CA) and the intensity of the signal in each lane quantified using Quantity One software (Bio-Rad Laboratories, Hercules, CA). The log of signal intensity was plotted versus the number of cycles using Excel XP (Microsoft Corporation, Redmond, WA). The number of cycles subsequently used for amplification of the target was the number of cycles in the middle of the linear portion of the curve. Linear range determination was carried out twice for each target gene.

Step 4: Determination of the Optimal Ratio of 18S Primers:Competimers

Since the 18S primers compete for resources with the target in the multiplex reaction, it is necessary to ensure that the 18S signal is attenuated to a level similar to that of the target (QuantumRNA Internal Standards Kit, Ambion Inc., Austin, TX), otherwise 18S will outcompete the target for PCR reagents, leading to a loss of exponential amplification of the target, which would make quantitation unattainable. Generally,

Ambion recommends 18S primer:competimer ratios of 4:6, 3:7, 2:8, and 1:9 for abundant transcripts, moderately expressed transcripts, rare transcripts, and extremely rare transcripts, respectively. In some cases, intermediate ratios could be required.

A series of PCRs were set up to determine the ratio required for each target. Reactions in total volumes of 20 µL containing 2 µL each of 2 µM forward and reverse target primers, 1X PCR buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl), 2 µL of 10 mM dXTPs, 1-2.5 mM $MgCl_2$, 0.5 U of *Taq* polymerase (Invitrogen Canada Inc., Burlington, ON), and water to a final volume of 16.4 µL were prepared. To each reaction, 1.6 µL of 18S primers:competimers (Ambion Inc., Austin, TX) in ratios of 1:9, 2:8, 3:7, and 1.5:8.5, and 2 µL of cDNA were added. Conditions for the PCR consisted of an initial denaturation step for 5 minutes at 94°C, followed by the previously determined optimal number of cycles, each consisting of a 30 second denaturation step at 94°C, a 30 second annealing step at a temperature compatible with the target's primers, and a 1 minute extension at 72°C. A final extension for 10 minutes at 72°C was also carried out. A positive control for RT, positive PCR controls for both the target and 18S, and a negative PCR control were also performed. The PCR products were run on a 1% nondenaturing agarose gel (Section 2.8.1) and transferred to Hybond-N. Southern hybridization of the membrane was performed with both the target and 18S probes under standard conditions (Section 2.12.1). The membranes were exposed to phosphor screens (Bio-Rad Laboratories, Hercules, CA) and the relative intensity of the target signal was compared to that of 18S using Quantity One software (Bio-Rad Laboratories, Hercules, CA). The ratio of 18S primers:competimers for which similar intensities of the target and 18S were

obtained was the ratio chosen for subsequent experiments. Each ratio determination experiment was performed twice.

Step 5: Employing Relative RT-PCR for Determination of SAN1A and SAN1B Expression Patterns

To determine the expression pattern of *SAN1A*, a ratio of 18S primers:competimers of 1.5:8.5 was employed. PCR was carried out in reaction volumes of 20 µL each. Each reaction was composed of 1X PCR buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl), 0.2 µM each of the *SAN1A* forward and reverse primers, 1.6 µL of 1.5:8.5 18S primers:competimers (Ambion Inc., Austin, TX), 1.5 mM $MgCl_2$, 200 µM each of dGTP, dCTP, dATP, and dTTP, 0.5 U of *Taq* polymerase (Invitrogen Canada Inc., Burlington, ON), water to a final volume of 18 µL, and 2 µL of the cDNA sample to be amplified. PCR conditions consisted of an initial denaturation step of 94°C for 5 minutes, followed by 21 cycles of a 30 second denaturation step at 94°C, a 30 second annealing step at 57°C, and a 1 minute extension at 72°C. A final extension of 10 minutes at 72°C was also performed.

For determination of the expression pattern of *SAN1B*, a 1:9 ratio of 18S primers:competimers was employed; however, to compensate for the extremely low expression of *SAN1B*, less of the primer-competimer mix was used in each reaction. PCR was carried out in reaction volumes of 20 µL each. Each reaction was composed of 1X PCR buffer, 0.2 µM each of the *SAN1B* forward and reverse primers, 0.6 µL of 10X diluted 1:9 18S primers:competimers (Ambion Inc., Austin, TX), 2 mM $MgCl_2$, 200 µM each of dGTP, dCTP, dATP, and dTTP, 0.5 U *Taq* polymerase (Invitrogen Canada Inc., Burlington, ON), water to a final volume of 18 µL, and 2 µL of the cDNA sample to be amplified. PCR conditions consisted of an initial denaturation step of 94°C for 5 minutes,

followed by 25 cycles of a 30 second denaturation step at 94°C, a 30 second annealing step at 57°C, and a 1 minute extension at 72°C. A final extension of 10 minutes at 72°C was also performed.

All PCR products were run on 1% nondenaturing agarose gels (Section 2.8.1) and transferred to Hybond-N for Southern hybridization (Section 2.12.1). Probes were prepared from *SAN1A* or *SAN1B* and 18S PCR products and hybridizations were carried out under standard conditions. The membranes were exposed to phosphor screens (Bio-Rad Laboratories, Hercules, CA) and the intensity of either the *SAN1A* or *SAN1B* signal was compared to the intensity of the 18S signal using Quantity One software (Bio-Rad Laboratories, Hercules, CA).

2.15 DNA Sequencing

Cycle sequencing was carried out by the Li-Cor method using the Thermo Sequenase Primer Cycle Sequencing kit (Amersham Biosciences, Baie d'Urfe, QC) with fluorescently labelled universal forward and reverse Li-Cor primers (Table III) according to manufacturer's instructions. Completed sequencing reactions were then sent to Canadian Molecular Research Services Ltd. (Orleans, ON) for reading of the sequence.

Alternatively, DNA samples to be sequenced were sent to Barry Panchuk at the DNA Technologies Unit of the Plant Biotechnology Institute (Saskatoon, SK).

2.16 Sequence Analysis and Bioinformatics

Sequence analysis was carried out using a variety of computer programs. Sequence alignments were carried out with DNAMAN Version 4.22 (Lynnon BioSoft,

Vaudreuil, QC), CLUSTAL W Version 1.8, or CLUSTAL X (Thompson *et al.*, 1994, 1997). Editing of sequence alignments was carried out with BioEdit Version 7.0.5.2 (Hall, 1999).

In silico restriction analyses were performed with Webcutter 2.0 (http://rna.lundberg.gu.se/cutter2/), WatCut (http://watcut.uwaterloo.ca), or DNAMAN Version 4.22 (Lynnon BioSoft, Vaudreuil, QC).

Sequence similarity searches were carried out using the various BLAST programs offered by NCBI (http://www.ncbi.nlm.nih.gov), including the conserved domain search function (Altschul *et al.*, 1997; Marchler-Bauer and Bryant, 2004). Searches of soybean and other plant EST databases were carried out using the BLAST programs at The Institute for Genomic Research (TIGR) site (http://www.tigr.org). Searches of the *Lotus japonicus* genome were carried out at the *Lotus japonicus* Genome Sequencing Project web page (http://www.kazusa.or.jp/lotus/). *Medicago truncatula* genomic sequences were accessed through The TIGR *Medicago truncatula* database (http://www.tigr.org/tdb/e2k1/mta1) or through the *Medicago truncatula* Sequencing Resources web page (http://www.medicago.org/genome/). *Arabidopsis* genomic sequences were accessed through the TIGR *Arabidopsis thaliana* Database (http://www.tigr.org/tdb/e2k1/ath1) and rice genomic sequences through the TIGR Rice Database (http://www.tigr.org/tdb/e2k1/osa1).

Searches for protein motifs of the predicted proteins of the *SAN1* genes were carried out by scanning the PROSITE database (http://us.expasy.org/prosite/) for profiles and patterns (Hulo *et al.*, 2004).

Promoter regions of the *SAN1* gene family and intergenic sequences of the *SAN1* and *Arabidopsis* At3g19000 and At3g19010 genomic regions were aligned using

PipMaker (Schwartz *et al.*, 2000). Conserved sequences were identified by comparison with sequences of known plant promoter elements in the PLACE (http://www.dna.affrc.go.jp/PLACE) (Higo *et al.*, 1999) and PlantCARE (http://intra.psb.ugent.be:8080/PlantCARE) (Lescot *et al.*, 2002) databases.

Phylogenetic analysis of the *SAN1* gene family was carried out using a variety of programs. A gene tree based on protein sequences was constructed using the neighbour-joining method and bootstrapping was performed 10,000 times using programs contained in PHYLIP Version 3.63 (Felsenstein, 1989). The resulting tree was drawn by TreeView Version 1.6.6 (Page, 1996).

Testing for gene conversion between the *SAN1* genes was carried out with GENECONV Version 1.81 (Sawyer, 1989).

2.17 Statistics

Statistical analyses were carried out using SYSTAT Version 10 (SPSS Inc., Chicago, IL). One-way ANOVAs were carried out in order to determine whether mean nitrogenase activity, Lb expression, and *SAN1A* or *SAN1B* expression differed during natural nodule development or in response to treatment with nitrate or darkness, since ANOVA tests for the effects of discrete independent variables, such as developmental time points or time of treatment with nitrate or dark. Tukey's test (HSD) at 5% level of significance was used to determine critical differences between means. Tukey's test is the most commonly used post hoc test. In addition, it reduces the probability of type I error and is especially appropriate when making comparisons between six or more means (Aspelmeier, 2002).

CHAPTER 3

RESULTS: PART I

Characterization of Soybean Nodule Senescence

To provide a framework for the study of gene expression during soybean nodule senescence, it was necessary to characterize, to a limited extent, the process of senescence itself. The morphology of senescing nodule cells was examined by light microscopy to determine whether changes similar to those previously described as occurring during PCD could be observed. The progression of soybean root nodule senescence was examined using both a physiological and a molecular marker. Nitrogenase activity was used as the physiological marker of senescence and Lb mRNA expression, as measured by Northern analysis, was employed as the molecular marker of the process, since both are known to decline during root nodule senescence (Becana and Sprent, 1987; Chan, 1995; Matamoros *et al.*, 1999; Sutton, 1983). Both assays were performed to assess when, and how rapidly, nodule senescence occurs during the course of natural development in soybean. Additionally, these assays were carried out on root systems of soybean plants in which nodule senescence had been induced by nitrate application or darkness to ensure that both treatments caused senescence, and to compare the rates of senescence between the two types of induction. The efficacy of Lb mRNA expression as a marker of nodule senescence was also compared with that of nitrogenase activity.

3.1.1 Morphological Characteristics of Soybean Nodule Senescence

Although nodule senescence has not been studied extensively, several researchers have discovered in senescing nodules features corresponding to those of PCD. For example, reactive oxygen species appear in naturally senescing soybean nodules (Evans *et*

al., 1999). In addition, studies by Matamoros *et al.* (1999) showed that time-dependent cytoplasmic and membrane degradation occurs in cells of *Pisum sativum* (pea) and *Phaseolus vulgaris* (bean) nodules senescing in response to dark or nitrate treatments.

Plant PCD is characterized by certain morphological changes, including cell shrinkage and condensation, membrane blebbing, and condensation and/or fragmentation of the nucleus (Levine *et al.*, 1996; Mittler, 1998; Pennell and Lamb, 1997). In addition, the vacuole may progressively increase in size until the vacuolar membrane ruptures. Although morphological features of PCD, such as changes to the nuclei and cytoplasmic and membrane alterations can be observed by light or electron microscopy, initial assessments of PCD are typically carried out by light microscopy (Loo and Rillema, 1998; Ziegler and Groscurth, 2004).

3.1.1.1 Experimental Techniques and Design

Preliminary light microscopy studies were carried out to determine which morphological features of PCD, if any, could be detected in soybean nodules undergoing senescence, and to identify any other unique changes that might be occurring. A series of naturally senescing nodules and one of nodules undergoing induced senescence were observed. Two series of plants were observed: naturally senescing nodules and nodules undergoing induced senescence. The morphology of cells in the senescing soybean nodules was observed in both toluidine blue O-stained epoxy resin-embedded nodule sections. Toluidine blue O was chosen to highlight overall morphology, since it is a general stain, interacting indiscriminately with the nucleus and the components of the cytoplasm (Berlyn and Miksche, 1976). One field of one section, representative of features observed in several fields of several different nodule sections, was chosen for

presentation for each time point examined during natural or induced nodule senescence. Different fields of view were presented for each stain for each time point examined.

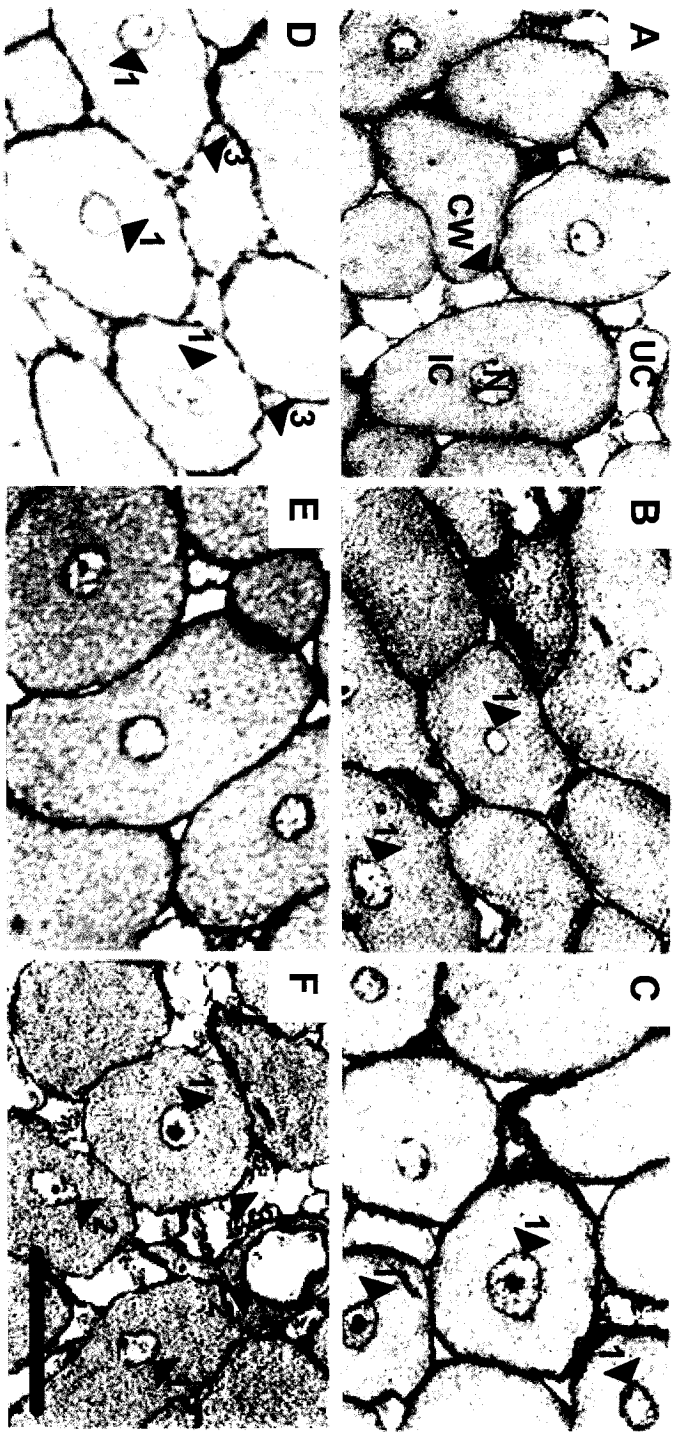
3.1.1.2 Morphological Features of Cells in Nodules Undergoing Natural Senescence

Morphological changes occurring during natural senescence were observed at 21 dpi (prior to senescence), 32 dpi, 40 dpi, 50 dpi, and 60 dpi. By 32 dpi, senescence was underway (Chan 1995; this study). The largest nodules were selected for microscopy at each time point since they were likely to represent events arising from the initial infections with *Bradyrhizobium japonicum* at day zero, while smaller nodules were avoided since they were more likely to represent secondary infections by the rhizobia still present in the rhizosphere. Observations focused on the central zone of the root nodule since events characteristic of senescence are more readily detectable in this region (Matamoros *et al.*, 1999). The infected cells are larger and appear granular due to the presence of bacteroids. Organelles of the uninfected cells are not readily visible due to the large size of the vacuoles that push the cytoplasmic contents to their margins.

In 21 dpi nodules, most nuclei of the infected cells appeared intact and spherically shaped. Bacteroids seemed to be evenly distributed throughout the cytoplasm (Figure 9A). In addition, the cell walls were clearly defined and there did not appear to be visible spaces between any of the uninfected and infected cells. The nuclei were intact and spherical. By 32 dpi, some nuclei appeared misshapen, but the cell walls remained unchanged in appearance and the bacteroids were evenly distributed within the infected cells (Figure 9B). At 40 dpi, no changes to bacteroid distribution were observed, although a slight crinkling of the cell walls was evident (Figure 9C). A greater proportion of nuclei were misshapen by this time point. Between 40 and 50 dpi, few changes occurred; more nuclei were misshapen and slightly more crinkling of cell walls was noticeable, although this

Figure 9: Bright Field Micrographs of Soybean Nodule Cross-Sections from Various Time Points during Naturally Occurring Senescence

Sections were resin-embedded, sliced to a thickness of 2 μm, and stained with toluidine blue O. All photos obtained with a 40X dry objective. (A) 21 dpi. (B) 32 dpi. (C) 40 dpi. (D), (E) 50 dpi. (F) 60 dpi. (1) misshapen nucleus; (2) advanced nuclear deformation; (3) cell wall crinkling. Bar, 50 μm. UC, uninfected cell; IC, infected cell; N, nucleus; CW, cell wall.


A
CW
UC
IC
N
B
C
D
1
3
E
F

was not apparent in all fields observed at 50 dpi (Figure 9 D,E). By 60 dpi, more dramatic changes in the cells could be observed (Figure 9F). Cell walls were very crinkled and nuclei were very misshapen, although the bacteroids still appeared to be evenly distributed throughout the infected cells.

Overall, the major changes observed in nodule cells during naturally occurring soybean nodule senescence were crinkling of the cell walls and deformation of the nuclei, the extent of each appearing to increase with time.

3.1.1.3 Morphological Features of Cells in Nodules Undergoing Induced Senescence

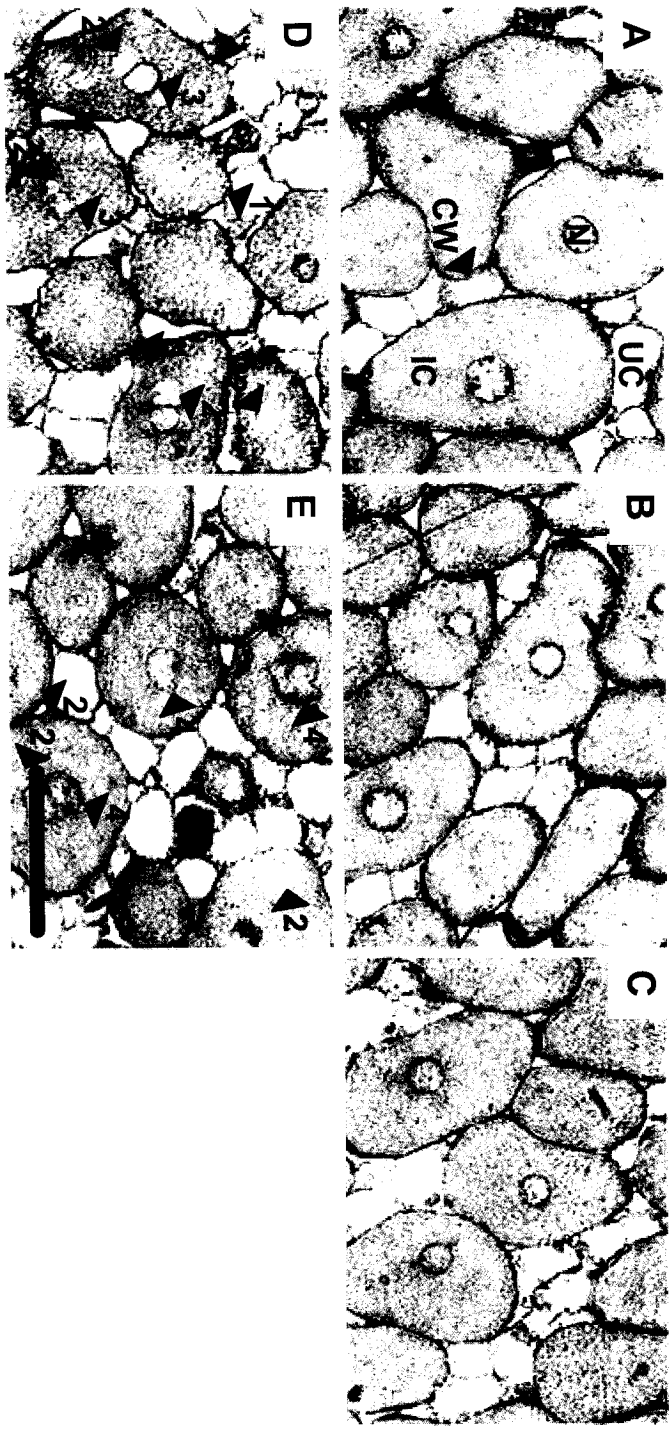
Morphological features of nodule cells undergoing senescence induced by treatment with NH_4NO_3 were also examined. The induction of senescence was initiated at 21 dpi since nodules of this age are mature, but not yet senescent (Chan, 1995; this study). Nodules were harvested and prepared for microscopy at 0, 12, 24, 36, and 48 hours after the initial addition of 20 mM NH_4NO_3.

Characteristics of nodules at 21 dpi were the same as those described in the preceding section (Figure 10A). Few changes were visible between 0 and 24 hours after induction of senescence (Figure 10 A-C). By 36 hours after the initial addition of NH_4NO_3, toluidine blue O-stained samples exhibited many differences from the preceding three time points. At this time point, most of the cell walls in the field appeared crinkled, misshapen nuclei were readily apparent, and the infected cells appeared to have zones that were cleared of bacteroids (Figure 10D). After 48 hours of nitrate treatment, still more nuclei appeared misshapen and the degree of deformation was more severe (Figure 10E). Interestingly, the crinkled appearance of the cell walls could no longer be observed.

Prior to 36 hours after induction of senescence by nitrate application, few changes could be observed. However, by this time point, zones cleared of bacteroids, misshapen

Figure 10: Bright Field Micrographs of Soybean Nodule Cross-Sections from Various Time Points after the Induction of Nodule Senescence by the Addition of Nitrate

Sections were resin-embedded, sliced to a thickness of 2 μm, and stained with toluidine blue O. All photos obtained with a 40X dry objective. After induction of senescence at: (A) 0 hours, (B) 12 hours, (C) 24 hours, (D) 36 hours, (E) 48 hours. (1) cell wall crinkling; (2) cleared zone; (3) misshapen nucleus; (4) advanced nuclear deformation. Bar, 50 μm. UC, uninfected cell; IC, infected cell; N, nucleus; CW, cell wall.


A
CW
N
IC
UC
B
C
D
E
2
2
2
2

nuclei, and crinkling of the cell walls could be readily observed. The cleared zones were likely to be indicative of some type of cytoplasmic disruption of the infected cells. Their appearance could have been due to an increase in vacuole size that is often observed during plant PCD (Mittler, 1998), but may also have been due to the formation of spaces that are often observed during necrotic cell death (Ziegler and Groscurth, 2004) or the destruction of bacteroids. Light microscopy did not provide enough resolution to determine which of these possibilities is correct.

It is possible to say that at least one feature of PCD was observed during nodule senescence induced by NH_4NO_3 application, that of nuclear deformation. Cytoplasmic condensation, as evidenced by cell wall crinkling, and increased vacuolation could also have been occurring, but this would require verification by electron microscopy, which would provide enhanced resolution.

3.1.1.4 Summary of Light Microscopic Observations of Root Nodule Senescence

Cells of both naturally senescing soybean nodules nitrate-treated nodules exhibited crinkling of the cell walls and deformation of the nuclei. However, the crinkling of the cell walls was more pronounced in the nodule cells undergoing induced senescence than in the cells of the nodules undergoing natural senescence. Interestingly, cleared zones were only observed in the cells of nodules induced to senesce by nitrate application.

3.1.2 Monitoring the Progression of Soybean Nodule Senescence

3.1.2.1 Physiological Monitoring of Soybean Nodule Senescence – Acetylene Reduction Assays

The rate of nitrogen fixation can be monitored by acetylene reduction assays as described by Hardy *et al.* (1968). During this assay, acetylene substitutes for nitrogen as a substrate for nitrogenase in the nodule's bacteroids; however, instead of ammonium,

ethylene gas is produced. This conversion can be monitored by gas chromatography (GC).

Rates of nitrogen fixation can be used to monitor the progression of soybean nodule senescence since an overall decline in its rate occurs during the senescence of root nodules (Chan, 1995; Matamoros *et al.*, 1999; Puppo *et al.*, 2005). Nodule senescence can be allowed to occur naturally or it can be initiated by the application of nitrate or by placing plants in the dark. Nodule senescence occurs naturally during whole plant senescence that is initiated around the time of flowering (Chan, 1995; Sutton, 1983).

3.1.2.1.1 Separation of Ethylene and Acetylene – Determination of Retention Times

Before using the acetylene reduction assay to monitor nitrogen fixation in the nodules, determination of the conditions that would best separate the ethylene and acetylene with the HP-PLOTU column (Agilent Technologies Canada Inc., Mississauga, ON) was necessary. Good separation was achieved at 50°C and the retention times of ethylene and acetylene were found to be ~1.6 and ~2.3 minutes, respectively. Figure 11C is a chromatogram showing peaks of ethylene and acetylene.

3.1.2.1.2 Preparation of Ethylene Standards

To determine the amount of ethylene produced during an acetylene reduction assay, it is necessary to relate ethylene peak area detected by GC to a known amount of ethylene. This determination was carried out by use of ethylene standards ranging from 0.2% to 1% (v/v) that were prepared in nitrogen. Plants with nodules that were senescing (> 40 dpi) and plants with mature, non-senescent nodules (~20 dpi) were used to estimate of the range of standards required. Standards were run each time a set of samples was assayed. Figure 11A depicts the relationship of peak area to the percentage of ethylene for a sample set of standards.

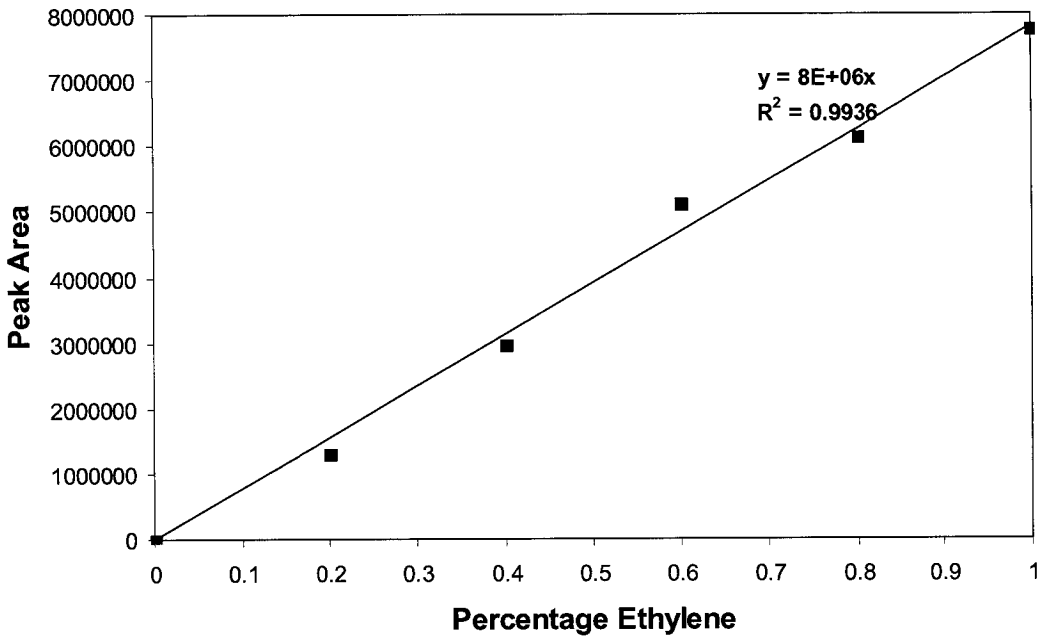
Figure 11: Preparations for the Acetylene Reduction Assay

A. Ethylene standards were prepared in the following concentrations: 0, 0.2, 0.4, 0.6, 0.8, and 1% (v/v) ethylene in nitrogen. 10 μL of each standard was loaded on the GC. The peak areas were recorded and plotted versus the percentage of ethylene, allowing peak areas of ethylene to be compared to known amounts of ethylene for determination of the rate of nitrogen fixation. Percentage ethylene was converted to nanomoles of ethylene based on the density of the gas and its molecular weight.

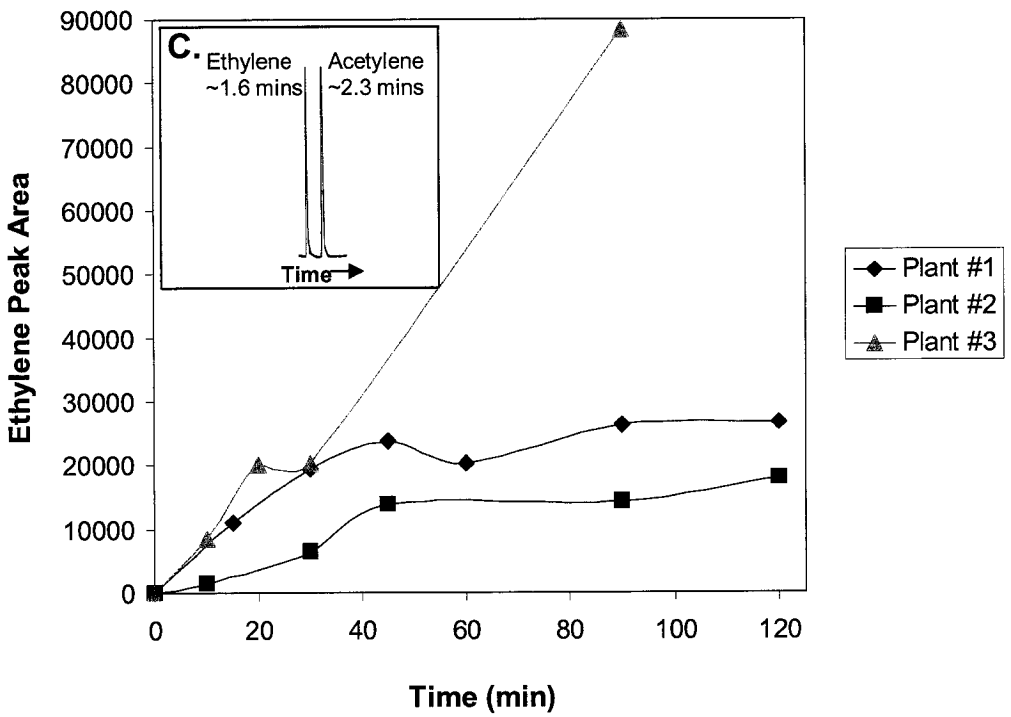
B. The production of ethylene by nodulated soybean root systems was monitored over a time period of 120 minutes to determine the linear range of the acetylene reduction assay. Three plants - of ages 38 dpi, 47 dpi, and 47 dpi for plants #1, #2, and #3, respectively - were tested. The results of the linear range tests are depicted.

C. Retention times of ethylene and acetylene were determined to be ~1.6 and ~2.3 minutes, respectively, at a temperature of 50°C on a Varian Star 3400 gas chromatograph (Varian Instrument Group, Walnut Creek, CA) outfitted with an HP-PLOTU column 15 m in length, with an internal diameter of 0.53 mm and a 20 μm film (Agilent Technologies Canada Inc., Mississauga, ON).

A.


8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
Peak Area
y = 8E+06x
R2 = 0.9936
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Percentage Ethylene

B.


90000
80000
70000
60000
50000
40000
30000
20000
10000
0
Ethylene Peak Area
C.
Ethylene
~1.6 mins
Acetylene
~2.3 mins
Time
0
20
40
60
80
100
120
Time (min)
Plant #1
Plant #2
Plant #3

3.1.2.1.3 Linear Range of the Acetylene Reduction Assay

The linear range of the acetylene reduction assay was determined to ensure that samples were taken at times that would reflect the true reaction rate. It was expected that the production of ethylene would level off after a certain amount of time due to saturation of the enzyme or due to O_2 depletion in the sample container (Hardy *et al.*, 1968). In addition, ethylene is a plant hormone that has been implicated in promoting and accelerating plant senescence (Beers, 1997). Therefore, if it accumulates to high enough levels, the hormone itself might cause the process of senescence to become autocatalytic, causing the rate of acetylene reduction to decrease more rapidly, affecting the linear range of the assay. Since only ~2% of the acetylene is converted to ethylene in the course of a typical assay (Hardy *et al.*, 1968), the decline in acetylene is not expected to be limiting and therefore acetylene is considered to be constant.

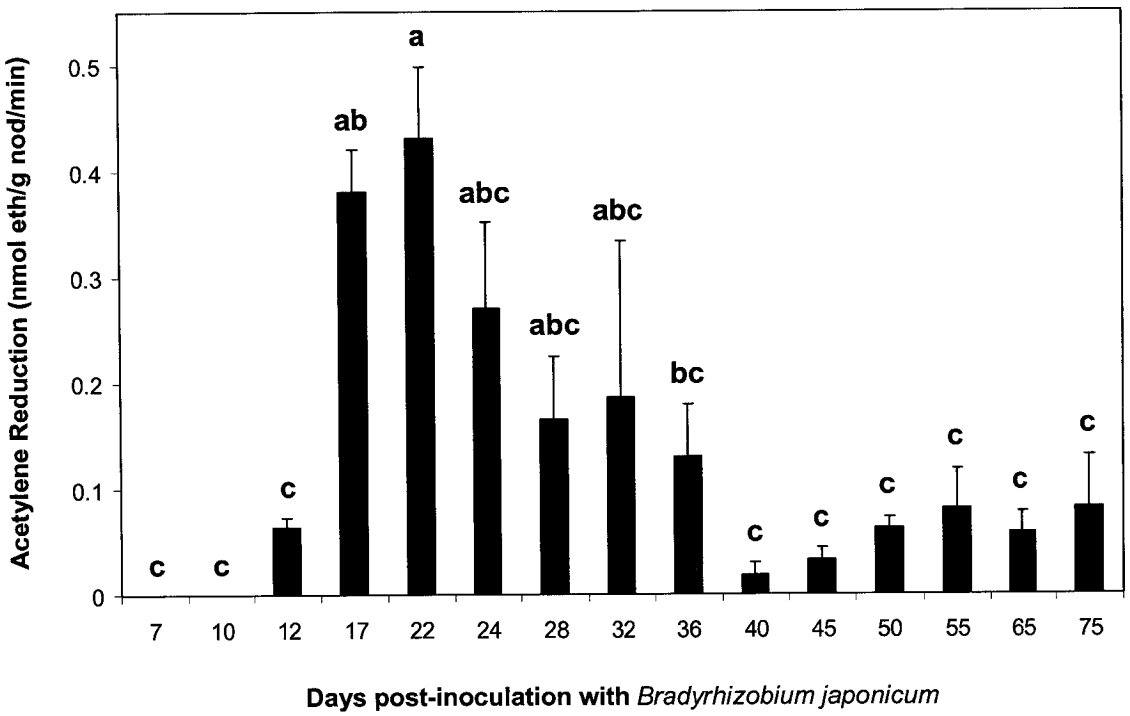
Three plants were tested to determine the linear range of the acetylene reduction assay over a total time period of 120 minutes (Figure 11B). For plants #1 and #2, the reaction was found to be in the linear range up to 35 minutes. Plant #3 gave an anomalous result, since acetylene reduction did not level off; however, it is not clear why this occurred. The linear range was at least 35 minutes for all three plants tested. Therefore, for subsequent assays, samples were removed between 0 and 30 minutes after assay initiation in order to ensure that measurements were taken in the linear range, thereby reflecting the correct rate of nitrogen fixation.

3.1.2.1.4 Nitrogen Fixation in Nodules throughout Soybean Development

Nitrogen fixation was monitored throughout natural soybean development between 7 and 75 dpi (Figure 12). At 7 and 10 dpi, no nodules were visible. Therefore,

Figure 12: Nitrogen Fixation during Natural Soybean Nodule Development

Acetylene reduction assays were performed on root systems of soybean plants at 7, 10, 12, 17, 22, 24, 28, 32, 36, 40, 45, 50, 55, 65, and 75 dpi with *Bradyrhizobium japonicum*. Results are presented as mean ± SE of the rate of acetylene reduction versus dpi with *B. japonicum*. Three plants from one set were assayed for each time point. Means with different letter(s) are statistically different (Tukey's test, $p<0.05$). eth, ethylene; g nod, grams of nodules (dry weight).


Acetylene Reduction (nmol eth/g nod/min)
0
0.1
0.2
0.3
0.4
0.5
c
c
c
ab
a
abc
abc
bc
c
c
c
c
c
c
7
10
12
17
22
24
28
32
36
40
45
50
55
65
75
Days post-inoculation with Bradyrhizobium japonicum

as expected, no nitrogen fixation was detected. By 12 dpi, tiny, pink nodules were visible as bumps on the roots and nitrogen fixation was occurring. Nitrogen fixation continued to increase until it peaked at 22 dpi, after which point it began to decline, as expected. Under these growth conditions, flowering, a state that precedes nodule senescence, occurred between 22 and 24 dpi. Between 24 and 36 dpi, levels of nitrogenase activity were similar. By 40 dpi, nitrogen fixation had declined considerably. After 40 dpi, there was a slight increase to 75 dpi.

Although the trends on the graphs point to increases and declines, very few statistically significant differences in the rates of nitrogen fixation were found. For example, plants examined at both 24 and 32 dpi exhibited large variations in their rates of nitrogen fixation, displayed as large standard errors on the graph. In a natural population of soybean plants, the nodule population is not necessarily uniform. Some plants may become infected before other plants and the nodules on an individual plant may be at various stages of development due to differences in the timing of infection. This problem of variation in nitrogen fixation values between plants could be partially remedied by examining more plants at each time point or by removing the soybean plants to a rhizobia-free growth medium after the first crop of nodules has been formed. Due to space and time constraints, this was not feasible for this study.

3.1.2.1.5 Nitrogen Fixation in Nodules Induced to Senesce by Nitrate and Darkness

Applying nitrate to and placing plants in the dark cause root nodules to senesce (Brewin, 1991; Matamoros *et al.*, 1999; Pfeiffer *et al.*, 1983a). To test that nitrate and darkness induced soybean root nodule senescence, acetylene reduction assays were performed to monitor the rate of nitrogen fixation after initiation of the treatments. In

addition, a comparison between the rates of nodule senescence induced by nitrate or darkness was possible.

3.1.2.1.5.1 Nitrogen Fixation in Nodules of Untreated Soybean Plants

Initially, untreated plants were assayed at 0, 24, 48, 72, and 96 hours (21 to 25 dpi) to ensure that no decline in nitrogen fixation was occurring that could interfere with the interpretation of the results when nitrate or darkness was applied. Zero hours corresponded to 8:00 am for a 21 dpi plant (control or treated). There were no significant differences between the time points (one-way ANOVA, $p=0.236$), indicating that no significant decline in nitrogen fixation occurs between 0 and 96 hours (21 to 25 dpi) in nodules of untreated plants (data not shown). Therefore, any declines in nitrogen fixation observed in nitrate- and dark-treated plants could be attributed to the treatments and not to natural senescence of the plants' root nodules.

3.1.2.1.5.2 Nitrogen Fixation in Nodules of Nitrate-Treated Plants

Soybean plants that had been treated daily with 20 mM NH_4NO_3 were assayed for their nitrogen fixation ability at 0, 12, 24, 36, 48, 72, and 96 hours after the beginning of treatment (Figure 13A). It is clear that the application of nitrate caused a decline in the rate of nitrogen fixation between 0 and 96 hours (one-way ANOVA, $p=0$). Tukey's test ($p<0.05$) indicated that the rate of nitrogen fixation showed a significant decline by 24 hours after the initiation of treatment, when the rate appeared to drop to approximately half of that at zero hours. Between 24 and 96 hours, the mean rate of nitrogen fixation did not change significantly, although the trend of the graph shows a slight continual decline in nitrogen fixation. By 96 hours, the rate of nitrogen fixation had declined to approximately one third of the initial value. Therefore, nitrate application was successful at inducing soybean nodule senescence.

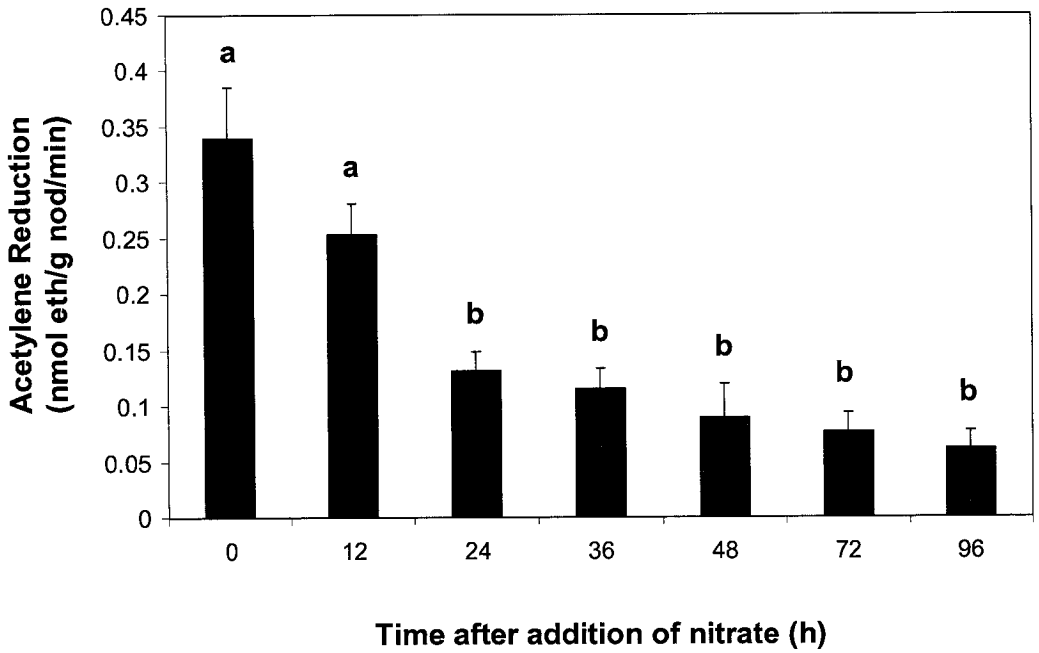
Figure 13: Nitrogen Fixation during Induced Senescence of Nodules

Acetylene reduction assays were performed on root systems of soybean plants at 0, 12, 24, 36, 48, 72, and 96 hours after the initiation of senescence-inducing treatments. Zero hours corresponded to 8:00 am, 21 dpi with *Bradyrhizobium japonicum*. Results are presented as mean ± SE of the rate of acetylene reduction versus hours elapsed after the initiation of treatment. Three plants from one set were assayed for each time point. Means which are statistically different from one another are marked with different letters (Tukey's test, $p<0.05$). eth, ethylene; g nod, grams of nodules (dry weight).

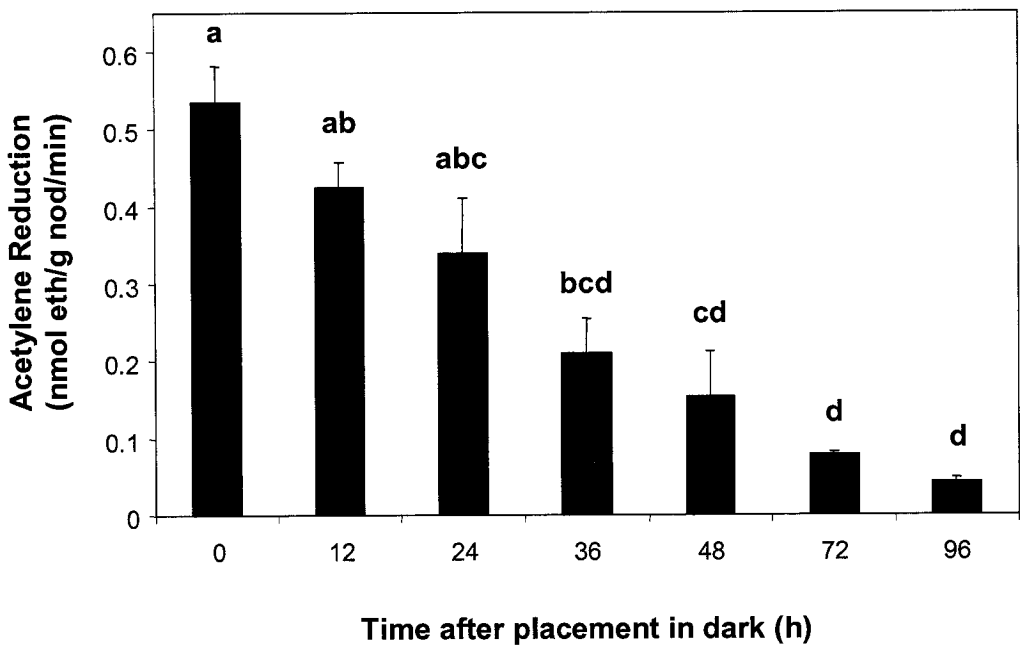
A. Nitrogen fixation during nodule senescence induced by the application of 20 mM NH_4NO_3.

B. Nitrogen fixation during nodule senescence induced by darkness.

A.


Acetylene Reduction (nmol eth/g nod/min)
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
a
a
b
b
b
b
b
0
12
24
36
48
72
96
Time after addition of nitrate (h)

B.


Acetylene Reduction (nmol eth/g nod/min)
0.6
0.5
0.4
0.3
0.2
0.1
0
a
ab
abc
bcd
cd
d
d
0
12
24
36
48
72
96
Time after placement in dark (h)

3.1.2.1.5.3 Nitrogen Fixation in Nodules of Dark-Treated Plants

The nodules of soybean plants that had been placed in the dark were also assayed for their ability to fix nitrogen at 0, 12, 24, 36, 48, 72, and 96 hours after the beginning of treatment (Figure 13B). Like nitrate, placing plants in the dark has the ability to initiate nodule senescence, exhibited by the decline in the rate of nitrogen fixation between 0 and 96 hours (one-way ANOVA, $p=0$). Tukey's test ($p<0.05$) indicated that by 36 hours, the mean rate of nitrogen fixation had declined to less than half of the initial value of nitrogen fixation. Dark treatment seemed to cause a less sharp decline in nitrogen fixation than nitrate application between 12 and 24 hours, but the graph displays a more continual decline to 96 hours. In addition, the overall reduction in the rate of nitrogen fixation appeared to be greater by dark treatment than by nitrate, since its final rate (96 h) was approximately 10% of the initial value.

3.1.2.2 Molecular Biological Monitoring of Soybean Nodule Senescence – Northern Analysis of Lb

As nitrogen fixation declines during soybean nodule senescence, the mRNA and protein levels of Lb, a key molecule involved in carrying oxygen to the bacteroids, also declines (Chan, 1995; Gallusci *et al.*, 1991; Matamoros *et al.*, 1999; Pfeiffer *et al.*, 1983a). Therefore, since Lb message levels are regulated at the level of transcription, it should be able to act as a molecular marker for senescence as an alternative to, or in conjunction with, acetylene reduction assays (Verma *et al.*, 1981). Northern analysis of Lb message levels was carried out on naturally senescing nodules and on nodules induced to senesce by the application of nitrate or by placement of plants in the dark.

3.1.2.2.1 Lb Message Levels in Nodules during Natural Soybean Development

Lb expression was monitored in nodules throughout natural soybean development between 7 and 75 dpi with *Bradyrhizobium japonicum* (Figure 14 A and B). At 7 and 10 dpi, no nodules were visible. Therefore, as expected, Lb expression was barely detectable. By 15 dpi, small, pink nodules were visible and Lb expression was readily detectable. From 15 to 75 dpi, Lb expression did not vary significantly (Tukey's test, $p<0.05$). However, trends indicate that there was increased expression of Lb at 28 and 55 dpi (Figure 14 A and B). After 55 dpi, a decline in Lb expression seemed to occur. Unfortunately, Lb mRNA levels did not mirror the declines observable in levels of nitrogenase activity during natural development of soybean nodules (see Figure 12). Therefore, it would be best to monitor natural nodule senescence using the acetylene reduction assay.

During natural development, the population of nodules on a given plant may be varied in terms of age, giving rise to large variations in Lb expression between sets of plants at a given time point. Re-infection by *Bradyrhizobium* may occur, causing increased variation in the stages of development represented in the nodule population. In addition, although nodule senescence may be occurring, Lb decline may occur late in the process of senescence. Although Lb is thought to be regulated at the level of transcription (Verma *et al.*, 1981), RNA turnover may be low and reduced transcription may not be easily detectable.

3.1.2.2.2 Lb Message Levels in Nodules of Untreated Soybean Plants

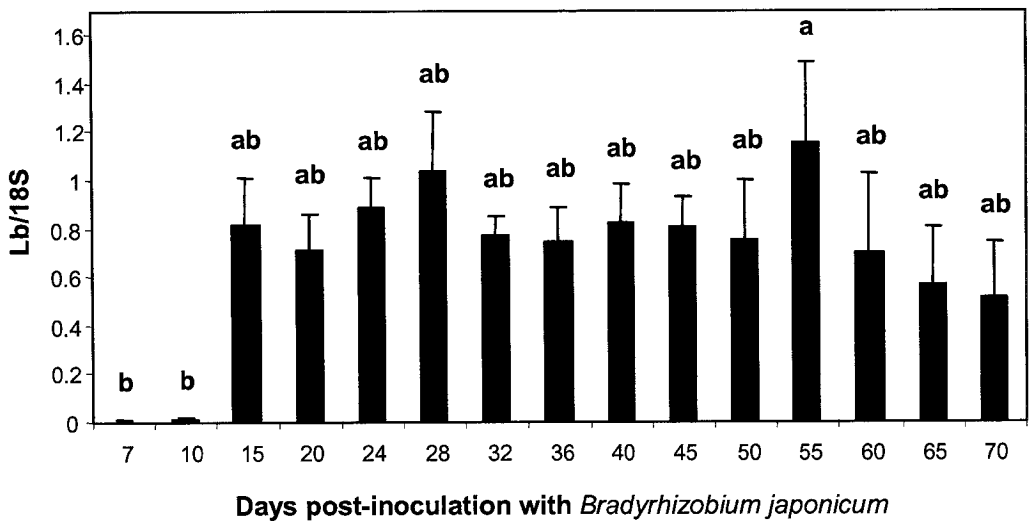
Initially, Northern analysis of Lb was carried out on one set of untreated plants at 0, 24, 48, 72, and 96 hours (21 to 25 dpi) to ensure that no natural decline in Lb expression was occurring during this particular time period in the life of the plants. Zero hours corresponded to 8:00 am for a 21 dpi plant (control or treated). There did not

Figure 14: Lb Expression in Nodules during Natural Senescence

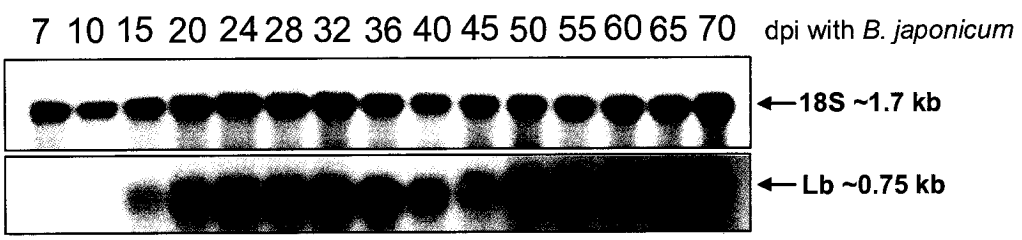
A. Northern blot analysis to monitor Lb expression was carried out on RNA extracted from root nodules harvested at 7, 10, 12, 17, 22, 24, 28, 32, 36, 40, 45, 50, 55, 65, and 75 dpi with *Bradyrhizobium japonicum*. Results are presented as mean ± SE of the Lb/18S signal versus dpi. Each time point represents the mean Lb/18S signals of nodules collected from three different sets of plants. For each set of plants, nodules were collected from 4-5 plants for each time point. Means that are statistically different from one another are marked with different letters (Tukey's test, $p<0.05$).

B. Northern blot of Lb and 18S carried out on nodules throughout nodule development for one set of soybean plants. 2.5 μg of total nodule RNA per lane was electrophoresed on a 1.5% agarose gel containing formaldehyde and transferred to GeneScreen+ membrane (NEN Life Science Products, Boston, MA). For detection of Lb mRNA, hybridization was carried out consecutively with ^{32}P-labelled Lb and 18S cDNA fragments under standard conditions (Section 2.12.2).

A.


Lb/18S
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
b b ab ab ab ab ab ab ab ab a ab ab ab
7 10 15 20 24 28 32 36 40 45 50 55 60 65 70
Days post-inoculation with Bradyrhizobium japonicum

B.


7 10 15 20 24 28 32 36 40 45 50 55 60 65 70 dpi with B. japonicum
←18S ~1.7 kb
← Lb ~0.75 kb

appear to be a decline in Lb expression between 0 and 96 hours in nodules of untreated soybean plants (data not shown). Therefore, any declines in Lb expression detected in nitrate- or dark-treated plants should have been a result of treatment and not from natural processes occurring in the plants' root nodules.

3.1.2.2.3 Lb Expression in Nodules of Nitrate-Treated Plants

Lb expression was monitored in nodules of soybean plants at 0, 12, 24, 36, 48, 72, and 96 hours after the initiation of treatment with 20 mM NH_4NO_3 (Figure 15 A,B). Although there was no statistically significant decline in the Lb/18S signal between 0 and 96 hours (one-way ANOVA, p=0.636), the trend of the graph indicates that expression of Lb was reduced by nitrate treatment (Figure 15A). The Northern blot also shows that a slow decline in Lb expression occurred due to nitrate treatment (Figure 15B). Nitrogenase activity and Lb message levels both declined in nodules of nitrate-treated soybean plants; however, measurement of nitrogenase activity appears to be a more sensitive method of monitoring senescence. Therefore, it would be best to monitor nitrate-induced senescence by measurement of nitrogenase activity instead of by Northern analysis of Lb mRNA.

3.1.2.2.4 Lb Expression in Nodules of Dark-Treated Plants

Northern analysis was also carried out to monitor Lb expression in nodules of soybean plants that had been placed in the dark. Nodule RNA was extracted from plants after 0, 12, 24, 36, 48, 72, and 96 hours of treatment (Figure 15 C,D). Like nitrate, placing plants in the dark has the ability to initiate nodule senescence. As with nitrate treatment, the reduction in Lb expression was not significant between 0 and 96 hours (one-way ANOVA, p=0.531), but a trend of decline in Lb expression was observed on the graph (Figure 15C) and Northern blot (Figure 15D). In addition, the Northern blot

Figure 15: Lb Expression in Nodules Undergoing Induced Senescence

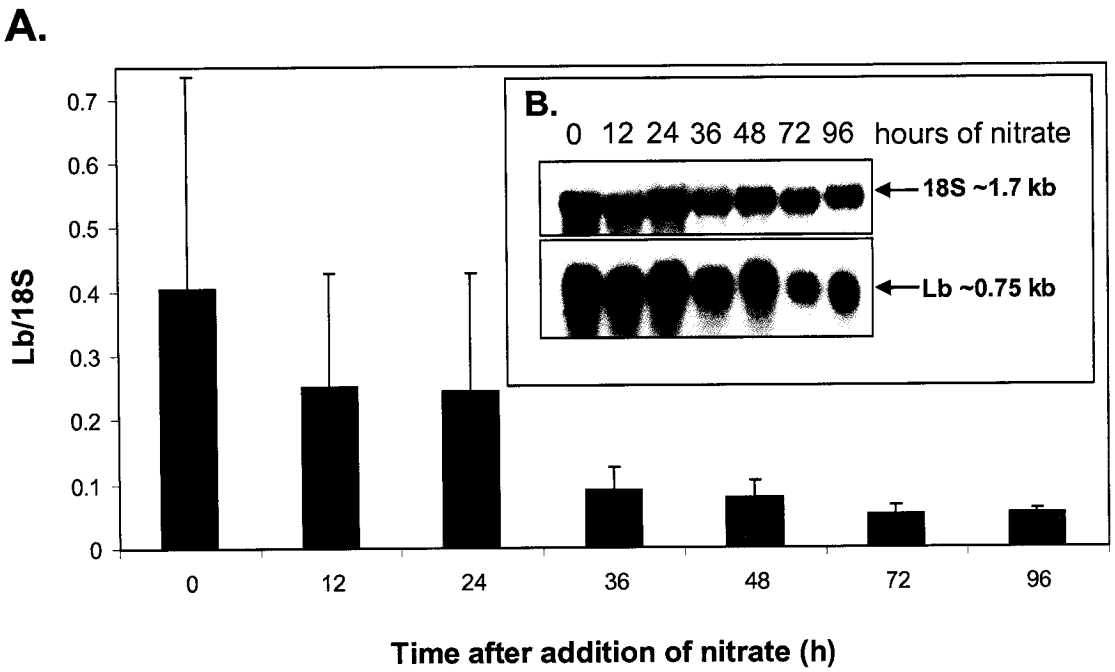
Northern blot analysis to monitor Lb expression was carried out on RNA extracted from root nodules harvested at 0, 12, 24, 36, 48, 72, and 96 hours after initiation of senescence-inducing treatments. Zero hours corresponded to 8:00 am, 21 dpi with *Bradyrhizobium japonicum*. Results are presented as the mean ± SE of the Lb/18S signal versus hours after initiation of nitrate treatment. Each time point represents the mean Lb/18S signals of nodules collected from three different sets of plants. For each set of plants, nodules were collected from 4-5 plants for each time point.

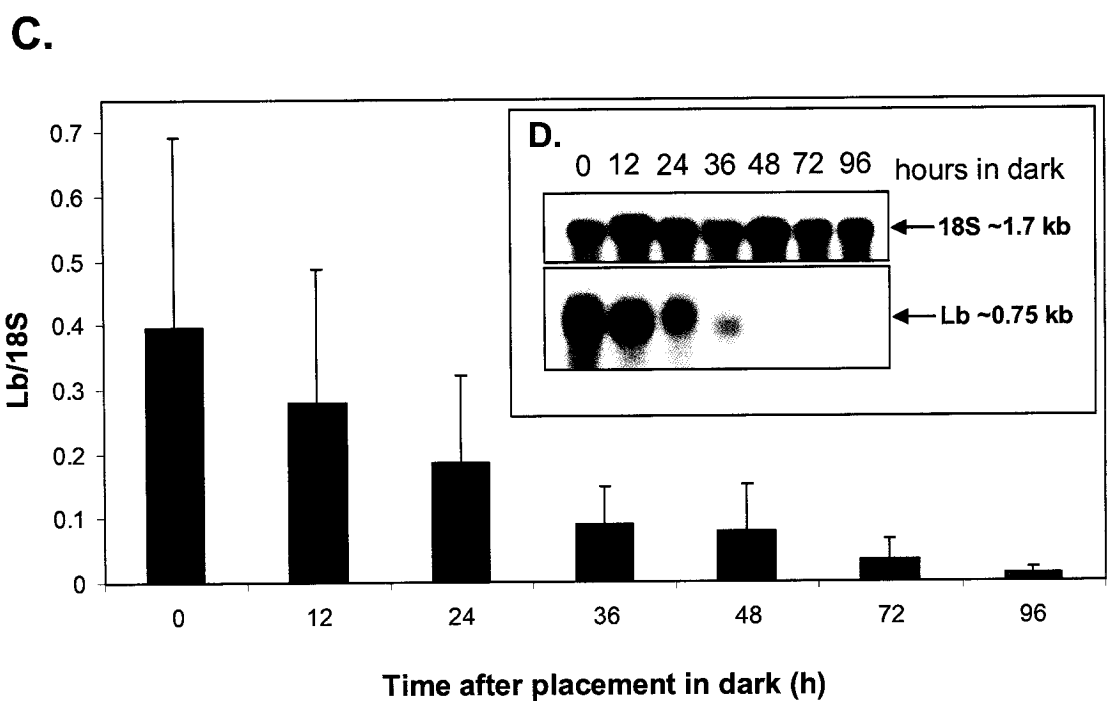
A. Lb expression during nodule senescence induced by the application of 20 mM NH_4NO_3. The decline in Lb expression between 0 and 96 hours was not statistically significant (one-way ANOVA, p=0.636), but the trend of the graph indicates a slight decline.

B. A sample Northern blot of Lb and 18S carried out on nodules of plants treated with nitrate between 0 and 96 hours. 2.5 μg of total nodule RNA per lane was electrophoresed on a 1.5% agarose gel containing formaldehyde and transferred to GeneScreen+ membrane (NEN Life Science Products, Boston, MA). For detection of Lb mRNA, hybridization was carried out consecutively with ^{32}P-labelled Lb and 18S cDNA fragment under standard conditions (Section 2.12.2).

C. Lb expression during nodule senescence induced by darkness. The decline in Lb expression between 0 and 96 hours was not statistically significant (one-way ANOVA, p=0.531), but the trend of the graph indicates a decline.

D. A sample Northern blot of Lb and 18S carried out on nodules of plants placed in the darkness between 0 and 96 hours. See B (this figure caption) for details.


A.
B.
0 12 24 36 48 72 96 hours of nitrate
18S ~1.7 kb
Lb ~0.75 kb
Lb/18S
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 12 24 36 48 72 96
Time after addition of nitrate (h)


C.
D.
0 12 24 36 48 72 96 hours in dark
18S ~1.7 kb
Lb ~0.75 kb
Lb/18S
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 12 24 36 48 72 96
Time after placement in dark (h)

shown (Figure 15D) indicates a more dramatic decline in Lb mRNA occurred during dark treatment than during nitrate treatment. As was found for nitrate treatment, measurement of nitrogenase activity is a more sensitive method of detecting dark-induced senescence than Northern analysis of Lb mRNA.

3.1.2.3 Summary of Acetylene Reduction Assay and Lb Northern Analysis Results for Soybean Nodule Senescence

Acetylene reduction assays of naturally senescing nodules indicated a peak in nitrogen fixation occurred at 22 dpi, followed by a decline to 75 dpi. Northern analysis of Lb expression in naturally senescing nodules did not indicate changes in message accumulation after 15 dpi. The acetylene reduction assay therefore appears to be a better way to monitor naturally occurring nodule senescence than Lb mRNA levels.

When soybean plants were treated with nitrate, nodule senescence was readily detectable by the acetylene reduction assay after 24 hours of treatment. However, the decline of the Lb message was less dramatic than the decline in nitrogenase activity during nitrate treatment. Placing soybean plants in the dark resulted in a decline in both nitrogenase activity, which was significantly reduced by 36 hours, and Lb mRNA. Although both types of measurements are capable of detecting soybean nodule senescence induced by nitrate or darkness, measurement of nitrogenase activity was found to be the more sensitive method of detection.

Induction of senescence by nitrate and dark treatments was carried out to produce a more homogeneous population of cells for gene expression studies. The results of the acetylene reduction assays suggest that increased homogeneity in the population of nodule cells was achieved by these treatments, since the standard errors in nitrogenase activity were smaller in nodules undergoing dark and nitrate treatment compared with those that

were senescing naturally. No significant declines in Lb expression were observed during either natural or induced senescence; therefore, it was not possible to make any conclusions from these results about whether nitrate or dark treatment had a homogenizing effect on the nodule cell population.

CHAPTER 3

RESULTS: PART II

Characterization of the *SAN1* Gene Family

A 935 bp cDNA clone, called *SAN1*G1, of the *SAN1A* (*Senescence-Associated Nodulin 1A*) gene was initially isolated by Chan (1995) by the differential screening of a cDNA library constructed from mRNA isolated from the nodule tissue of the soybean cv Maple Arrow 40 dpi with *Bradyrhizobium japonicum* 61A76. To verify that *SAN1A* was a senescence-associated gene, preliminary Northern analysis was carried out, revealing that its message increased between 22 dpi and 40 dpi (Chan, 1995). Comparison of the size of the *SAN1A* mRNA determined by Northern analysis (1.4 kb) with that of *SAN1*G1 (935 bp) suggested that a partial cDNA clone had been isolated.

Southern analysis indicated that *SAN1A* was likely to belong to a small, multigene family since four fragments were detectable in genomic DNA cut with either *Eco*RI or *Hind*III (Chan, 1995). Several DNA clones hybridizing to *SAN1A* were isolated for further analyses by screening a soybean genomic library (cv Resnik, Clontech Laboratories, Mountain View, CA) with *SAN1*G1.

3.2.1 The Complete Coding Sequence of *SAN1A*

To obtain cDNA sequence containing the entire coding sequence of *SAN1A*, 5′ and 3′ RACE were carried out. Figure 16 shows the binding sites of the RACE primers on the original *SAN1A* DNA clone, *SAN1*G1 (A), and the final *SAN1A* coding sequence obtained (B). Nested PCR was carried out to acquire the final 5′ and 3′ RACE products (see Section 2.14.3). The consensus sequence of three 5′ and 3′ RACE clones was determined

Figure 16: Extension of *SAN1A* cDNA Sequence by RACE

A. The binding sites of the primers used for 5′ and 3′ RACE are shown on the originally isolated *SAN1A* cDNA fragment. The 5′ RACE nested primers are shown in red text; the 3′ RACE nested primers are shown in blue text.

B. The *SAN1A* cDNA obtained by RACE is depicted. The start and stop codons are underlined and are shown in red. The potential polyadenylation sequence is underlined and shown in blue. The *SAN1A* CDS forward and reverse primers used to produce the full length *SAN1A* probe are also shown.

A.

```
1    TGAGGAATTTTTCATCAAAGATCAAACTAGCTTTATTCGACTCAACCACTATCCTCCAT

61   GCCCTTACCCTGACCTTGCTCTTGGCGTCGGTCGACACAAGGACCCCGGTGCCTTGACCA
     SAN1A-5RGSP2                                  SAN1A-5RGSP1
121  TTCTTGCACAGGACGAAGTTGGAGGACTTGAAGTGAGACGTAAACGGGATCAAGAGTGGA

181  TTAGAGTGAAACCAACCCCAAATGCTTATATTATCAACATTGGTGATACTGTTCAGGTTT

241  GGAGCAATGATGCATATGAGAGTGTGGATCACAGAGTGGTGGTCAACTCTGAGAAGGAAA
                                   SAN1A-3RGSP1
301  GGGTTTCCATTCCGTTCTTCTTCTTCCCTGCACACGACACCAAAGTGAAGCCTTTGGAGG
                                   SAN1A-3RGSP2
361  AGCTGATAAATGAACAAAACCCTTCAAAATATAGGCCATACAATTGGGGCAAGTTTCTTG

421  TCCACAGAGGGAACAGCAATTTCAAGAAGCAAAATGAGGAGAACATTCAAATTCATCATT

481  ACAAGATAGCTTAAAGGGTAGTATATTAGCAGGATTCCAGCGCCATATAGCTTATGCATA

541  TAAATTGGGTATGATATCATGATATGTAGCAGCGTTGATCCAAGACAACGTTTCTGATAT

601  AAATGACTGAGTCAAATTTACTTCCTAGTGTGCCTTAATGTATGCAAGTATCACCCAATA

661  AAACAACTAAATTATTA
```

B.

```
                                                SAN1ACDSF
1      ACACGTGCAGCAAATTGCTTACCCAATCACCAGCCATGGGAGAGGTTGACCCAGCTTTCA
61     TTCAAGAACCACAACACAGGCCAAATCTCTCCACCATTCAAGCAGAAGGAATTCCCATAA
121    TTGATCTCTCTCCAATAACCAACCACACAGTTTCAGACCCTTCTGCAATTGAAAGCTTGG
181    TGAAGGAGATAGGACGTGCATGCCAGGAGTGGGGCTTCTTCCAAGTAACAAACCATGGAG
241    TGCCACTCACTCTAAGGCAAAACATTGAGAAAGCCTCAAAACTGTTCTTTGCTCAGACTC
301    TGGAGGAGAAGAGAAAGGTTAGCAGAAATGAGAGCTCTCCAATGGGTTATTATGACACAG
361    AACACACCAAGAACGTCAGGGACTGGAAAGAAGTCTTTGATTTTCTAGCCAAAGACCTCA
421    CTTTCATTCCTCTCACTTCTGATGAACATGATGATCGAGTCAATCAATGGACTAATCAAT
481    CACCTCAATACCCTCCACTCTTCAGGGTTGTAACACAAGAGTATATTCAGGAGATGGAAA
541    AGCTGTCCTTTAAGCTGTTGGAGCTTATAGCTTTGAGCTTAGGCCTTGAAGCAAAGAGGT
601    TTGAGGAATTTTTCATCAAAGATCAAACTAGCTTTATTCGACTCAACCACTATCCTCCAT
661    GCCCTTACCCTGACCTTGCTCTTGGCGTCGGTCGACACAAGGACCCCGGTGCCTTGACCA
721    TTCTTGCACAGGACGAAGTTGGAGGACTTGAAGTGAGACGTAAACGGGATCAAGAGTGGA
781    TTAGAGTGAAACCAACCCCAAATGCTTATATTATCAACATTGGTGATACTGTTCAGGTTT
841    GGAGCAATGATGCATATGAGAGTGTGGATCACAGAGTGGTGGTCAACTCTGAGAAGGAAA
901    GGCTTTCCATTCCGTTCTTCTTCTTCCCTGCACACGACACCAAAGTAAAGCCTTTGGAGG
961    AGCTGATAAATGAGCAAAACCCTTCAAAATATAGGCCATACAATTGGGGCAAGTTTCTTG
1021   TCCACAGAGGGAACAGCAATTTCAAGAAGCAAAATGAGGAGAACATTCAAATTCATCATT
                                                SAN1ACDSR
1081   ACAAGATAGCTTAAAGGGTAGTATATTAGCAGGATTCCAGCGCCATATAGCTTATGCATA
1141   TAAATTGGGTATGATATCATGATATGTAGCAGCGTTGATCCAAGACAACGTTTCTGATAT
1201   AAATGACTGAGTCAAATTTACTTCCTAGTGTGCCTTAATGTATGCAAGTATCACCCAATA
1261   AAACAACTAAATTATTACC
```

to ensure that the correct coding sequence was achieved. The resulting assembled cDNA sequence was 1279 bp, suggesting that a nearly full length cDNA had been obtained, since the full length mRNA size had previously been estimated to be 1.4 kb by Northern analysis (Chan, 1995). The cDNA sequence contained an open reading frame of 1059 bp. Start and stop codons and a potential polyadenylation signal are highlighted in Figure 16B. This cDNA clone, called *SAN1A* CDS, was used as a probe in subsequent hybridization experiments.

3.2.2 Strategy for Obtaining Sequences of *SAN1* Genes and cDNAs from cvs Resnik and Maple Arrow

A genomic DNA clone named λ*SAN1* was isolated by screening of the soybean genomic DNA library with the *SAN1A* cDNA fragment *SAN1*G1. Since preliminary restriction and Southern analyses with the full length *SAN1A* CDS probe suggested that there were two or three *SAN1*-like genes contained within the 12.5 kb λ*SAN1* genomic DNA clone (data not shown), sequencing efforts were focused on this clone. Sequencing was initiated by subcloning fragments of λ*SAN1*.

Digestion with *SstI* generated two fragments of approximately 11 kb and 1.3 kb, while digestion with *SalI* revealed three fragments of approximately 3.1, 5.1 and 4.2 kb. Sequencing of λ*SAN1* employed two strategies. In the first approach, the two *SstI* fragments were cloned into plasmid vectors and sequenced. The entire 11 kb fragment was sequenced by primer walking through both strands (NRC-PBI, DNA Technologies Unit, Saskatoon, SK), while the smaller 1.3 kb fragment was subcloned and sequenced separately (D. Johnson). To verify the sequence at the joint between the two fragments, a small *Kpn*I-*Eco*RV fragment that overlaps both the 11 kb and 1.3 kb *Sst*I fragments was cloned and sequenced (D. Johnson). The second method followed a traditional "top-

down" approach. The three *SalI* fragments were cloned into plasmid vectors, and further subcloning generated five clones, CW1 to CW5, covering the approximately 12.5 kb of the original genomic clone (Figure 17). Each subclone was sequenced separately by primer walking (NRC-PBI, DNA Technologies Unit, Saskatoon, SK). The sequences obtained by the two different approaches were 100% identical, confirming that the correct sequence of the genomic region contained in the λ*SAN1* DNA clone had been determined.

Since the *SAN1A* gene was truncated in the λ*SAN1* DNA clone, a PCR approach was taken to extend the genomic sequence to include the entire gene. A primer specific to the T7 region of the EMBLλ3 SP6/T7 vector (Chapter 2, Table III) was used in combination with two nested *SAN1A*-specific primers (Chapter 2, Table III; Appendix 1, Figure 1) to amplify potential *SAN1A* gene products from the soybean library (D. Johnson). The resulting PCR products were cloned and the largest one sequenced. The region of *SAN1A* genomic DNA sequence extended by PCR was 100% identical to that of the original *SAN1A* cDNA. The strategy for the assembly of the various sequences of the *SAN1* genomic region is depicted in Figure 17.

Following completion of the λ*SAN1* sequence, the numbers and positions of the *SAN1* sequences were determined. Three genes arranged in tandem were identified and named *SAN1A*, *SAN1B*, and *SAN1C* based on their order of isolation, although their actual arrangement on the clone (5′ to 3′) is *SAN1B*, *SAN1C*, and *SAN1A* (Figure 18). In addition, the sequencing of the λ*SAN1* clone revealed the presence of an unrelated gene, called *MSG*, upstream of *SAN1B* (Figure 18). The *MSG* gene encodes a putative major latex protein homologue (Stromvik *et al.*, 1999). This finding should aid in the positioning of the *SAN1* genes with respect to other genes in the soybean genome.

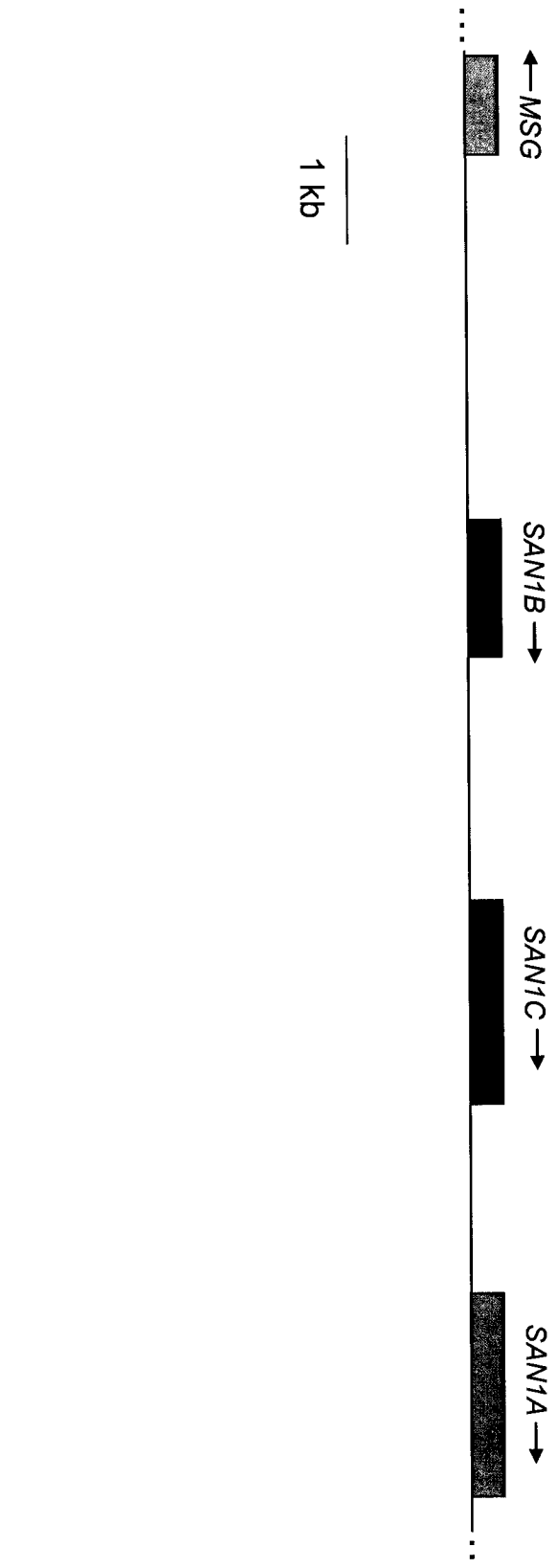
Figure 17: Strategy for the Assembly of the Final Sequence of the Soybean Genomic Region Containing the *SAN1* Genes

The final assembled genomic region is labelled the Assembled *SAN1* Genomic Region (15,012bp). The 12 kb *Sst*I fragment, the 2.2 kb Extended *Sst*I 1.3 kb Fragment, and the 2.1 kb fragment representing the PCR extended *SAN1A* sequence are indicated above the line. The positions of subclones CW1 to CW5, prepared as part of a "top-down" approach to obtaining the sequence, are indicated below the line. CW1 contains a 4.2 *Sal*I fragment; CW2 and CW3 are a *Pst*I fragment and a *Pst*I deletion of a *Sal* 5.1 subclone, respectively; CW4 and CW5 are a *Hind*III fragment and a *Hind*III deletion of a *Sal* 3.1 subclone, respectively.

SANIA_PCR_Extension
Extended_SstI_Fragment_1.3kb
SstI_Fragment_12_kb
Assembled_SAN1_Genomic_Region
CW1
CW5
CW3
CW4
CW2
1
15012

Figure 18: Arrangement of the *SAN1* Genes in the Soybean Genome

SAN1A, *SAN1B*, and *SAN1C* are arranged in tandem in the soybean genome. An unrelated gene, *MSG*, resides upstream of *SAN1B*. Arrows represent the direction of transcription of the genes.


MSG
SAN1B
SAN1C
SAN1A
1 kb

Since only genomic DNA sequence was obtained for cv Resnik, cDNA sequences for each of the *SAN1* genes were derived by alignment of their genomic sequences with the coding sequence of the *SAN1A* cDNA isolated from cv Maple Arrow. DNAMAN Version 4.22 (Lynnon BioSoft Inc., Vaudreuil, QC) was used to carry out the alignments. Since each of *SAN1B* and *SAN1C* from cv Resnik is at least 84% identical with the *SAN1A* coding sequences (Section 3.2.5, Table V), introns could be inferred from gaps in the alignments between the *SAN1A* coding sequence from cv Maple Arrow and the *SAN1* genes from cv Resnik. The putative exon-intron boundaries contained typical nuclear intron splice sites of plants, indicating that the gaps in the alignments are indeed indicative of the presence of introns (Lorkovic *et al.*, 2000).

All *SAN1* gene sequences and their derived coding sequences from the soybean cv Resnik can be found in Appendix 1 (Figures 2, 3, and 4).

3.2.3 Number of *SAN1* Genes in the Soybean Genome

Southern blot analysis of the soybean cv Maple Arrow genome was attempted in order to determine whether the *SAN1* gene family consists only of the three genes described, or whether there are more members of this family. However, the results were not conclusive (data not shown) and the actual number of genes in the family remains unknown.

In another attempt to address the question of gene number in soybean, an *in silico* approach was taken. BLASTN searches of the GSS (genome survey sequences) database of soybean maintained by Soybase (http://soybase.org/blast/GSSBlast.html) with the genomic DNA sequences of each of the *SAN1* genes revealed only one significant match from soybean cv Forrest (Acc. CG822679). The match was 97% identical to the *SAN1C* sequence over 295 nucleotides and 84-89% identical to shorter regions of *SAN1A* and

SAN1B, suggesting that *SAN1C* had been identified. No other significant matches to any of the *SAN1* genes were obtained. The *in silico* approach to determining the number of *SAN1* sequences present in soybean was not successful at this time, likely due to the fact that soybean genome sequencing remains incomplete.

3.2.4 Structure of the *SAN1* Genes

The *SAN1* genes of cv Resnik are each comprised of three exons and two introns (Figure 19A). The sizes of the exons appear to be conserved between the three genes. Exon #1 ranges from 470 to 471 bp, exon #2 is between 321 and 334 bp, and exon #3 ranges from 255 to 261 bp (Figure 19B). However, the sizes of the introns differ markedly between the genes. The first intron of *SAN1A* (393 bp) is similar in size to that of *SAN1C* (358 bp), but the first intron of *SAN1B* is much smaller at only 117 bp. The size of the second intron is well conserved between *SAN1A* and *SAN1B* at 258 bp and 257 bp, respectively. However, the second intron of *SAN1C* is much larger than those of *SAN1A* and *SAN1B* at 445 bp.

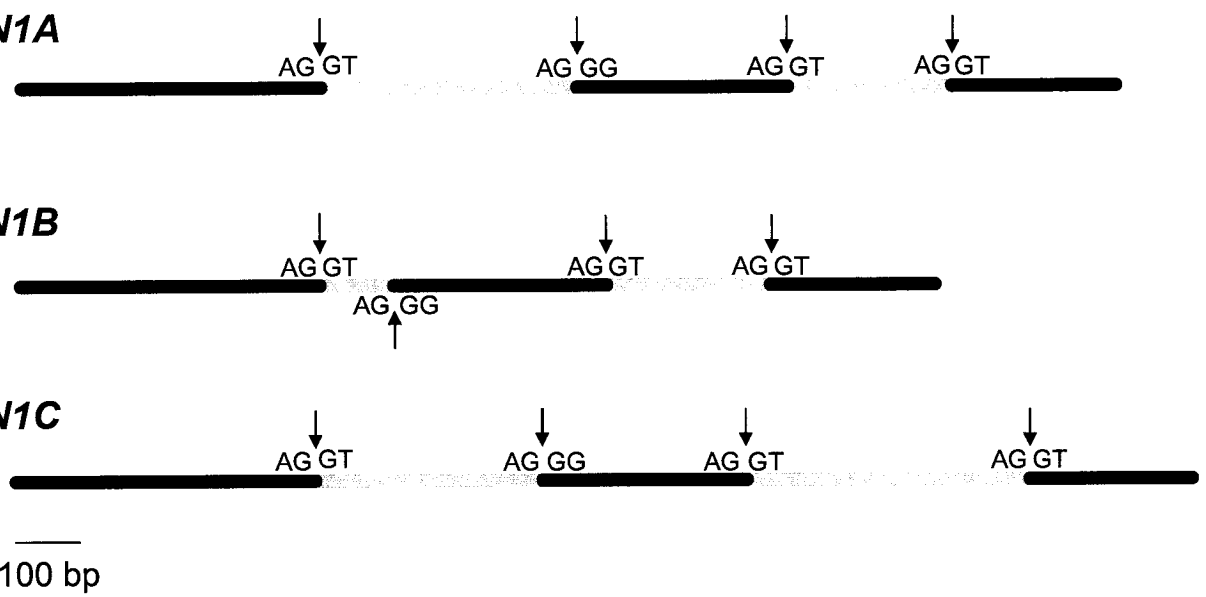
The conserved bases expected for dicot plant nuclear intron splice sites are present in the three *SAN1* genes (Lorkovic *et al.*, 2000). All six introns have the conserved 5′-GT…AG-3′ motif and are AT-rich, as expected for plant introns; however, the known plant branchpoint consensus sequence, CURAY, similar to those found in vertebrate introns, was not found in all of the introns of the *SAN1* genes (Table IV). It was only found in the first intron of *SAN1A* and the second introns of *SAN1B* and *SAN1C*. The reason for their absence is not certain, but various U-rich sequences found throughout the entire intron are known to be required for proper splicing and the CURAY sequence might not be necessary in all introns (Lorkovic *et al.*, 2000). Polypyrimidine tract sequences ($U_4Y_3U_3$) as defined by Brown *et al.* (2002) were not found in any of the *SAN1*

Figure 19: Structures of the *SAN1* Genes

A. The exon-intron structures of the *SAN1* genes from cv Resnik are depicted. The black bars represent exons and the green bars represent introns. The exon-intron and intron-exon splice junctions are denoted by arrows. The splice site sequences are also given.

B. A table displaying the sizes of the exons and introns of the *SAN1* genes from cv Resnik.

A.



B.

Exon/Intron Number	Size (bp)		
	SAN1A	*SAN1B*	*SAN1C*
Exon #1	470	470	471
Intron #1	393	117	358
Exon #2	331	334	321
Intron #2	258	257	445
Exon #3	258	261	255

Table IV: Features of the Introns of the *SAN1* Genes

Features of the introns of the *SAN1* genes, such as AT-richness, and the presence of motifs, such as conserved 5′-GT and 3′-AG dinucleotides and branchpoint sequences, are presented.

SAN1 Gene	Intron Number	Percent A+T	5′-GT & 3′-AG?	Putative Branchpoint Sequence(s)
SAN1A	1	62	Yes.	CTAAT (71 bp upstream of 3′-AG) CTAAT (115 bp upstream of 3′-AG)
	2	68	Yes.	No known.
SAN1B	1	74	Yes.	No known.
	2	68	Yes.	CTGAT (80 bp upstream of 3′-AG) CTAAT (171 bp upstream of 3′-AG)
SAN1C	1	69	Yes.	No known.
	2	65	Yes.	CTAAT (80 bp upstream of 3′-AG) CTAAC (138 bp upstream of 3′-AG) CTAAT (178 bp upstream of 3′-AG) CTAAC (363 bp upstream of 3′-AG)

introns. That polypyrimidine tracts were not identified is not surprising, since plants do not generally have long polypyrimidine tracts such as those found in the introns of vertebrates (Brown *et al.*, 2002; Lorkovic *et al.*, 2000).

The exon sequences flanking the *SAN1* gene intron sequences are also conserved. At each of the 5′ splice sites for the first and second introns of all three *SAN1* genes, the -2 and -1 positions are A and G, respectively, which are commonly found in these positions (position -2, A_{62}; position -1, G_{79}) (Lorkovic *et al.*, 2000). The conserved exon sequences at the 3′ splice sites in plants are more variable (position +1, G_{57}, position +2 T_{44}). For the second exon of all the *SAN1* genes, +1 and +2 positions are GG, and for the third exon of all the *SAN1* genes, +1 and +2 positions are GT respectively. Taken together, the features of the introns of the *SAN1* multigene family represent those of "typical" plant nuclear introns.

Since the *SAN1* cDNA sequences of soybean cv Resnik were inferred by comparison of the genomic sequences with the *SAN1A* cDNA cloned from soybean cv Maple Arrow, it was not possible to determine the lengths of the 5′ and 3′ UTRs. However, the entire 184 bp 3′ UTR had been obtained for *SAN1A* from cv Maple Arrow (Figure 16B) by 3′ RACE. A potential polyadenylation signal, often termed the near-upstream element (NUE) in plants, AATAAA, was identified 24 bp upstream of the polyadenylation site. NUEs determine the site of polyadenylation and usually occur between 10 and 30 nucleotides upstream of the actual polyadenylation site (Rothnie, 1996; Wu *et al.*, 1995); therefore, the AATAAA sequence may serve as a polyadenylation signal during *SAN1A* mRNA processing. The polyA tail of *SAN1A* is preceded by the sequence CA, which is expected, since plant polyadenylation sites are generally marked by YA (Rothnie, 1996; Wu *et al.*, 1995). In addition to the NUE and polyadenylation

sites, plants require a sequence upstream of the NUE, called the far-upstream element or FUE, for efficient polyadenylation (Li and Hunt, 1997; Rothnie, 1996; Wu *et al.*, 1995). However, none of these sequences were identifiable in *SAN1A* from cv Maple Arrow. Although determination of the extent of the 3′ UTRs of the *SAN1* genes of cv Resnik was not achievable, it was possible to identify whether they share the putative polyadenylation signals found in *SAN1A* from cv Maple Arrow by alignment of the *SAN1A* cDNA sequence from cv Maple Arrow with the *SAN1* genes of cv Resnik. DNAMAN Version 4.22 (Lynnon BioSoft, Vaudreuil, QC) was used to carry out the alignments. In all three *SAN1* genes from cv Resnik, the potential polyadenylation signal AATAAA was conserved with that of *SAN1A* from Maple Arrow, although the CA sequence prior to the polyA tail in *SAN1A* from Maple Arrow was not. The 3′ UTRs of the *SAN1* genes may also harbour sequences involved in interactions with RNA-binding factors or with the translation machinery (Mignone *et al.*, 2002). A scan of the UTR database (UTRdb) (Mignone *et al.*, 2005) indicates that none of the genes share these types of sequences. The results of a more comprehensive search for plant regulatory elements in the 3′ UTRs are presented in Section 3.4.3.

3.2.5 Sequence Conservation between the Soybean *SAN1* Genes at the DNA and Amino Acid Levels

Comparisons were made between the DNA sequences of the *SAN1A*, *SAN1B*, and *SAN1C* genes from cv Resnik. Sequences were aligned and the identity between them calculated using DNAMAN Version 4.22 (Lynnon BioSoft, Vaudreuil, QC). Comparisons were made between pairs of *SAN1* gene sequences from start codon to stop codon, between the coding sequences alone, and between each of the introns (Table V).

Comparison of each pair of *SAN1* genes revealed that over the entire gene

Table V: Conservation of *SAN1* DNA Sequences

Each pair of *SAN1* sequences isolated from soybean cultivar Resnik were compared over their entire gene sequences between the start codon and stop codon, over their predicted coding sequences, and over each of their intron sequences. Sequence identity is expressed as percent identity.

Genes Compared	Region Compared	Percent Identity
SAN1A	Entire gene	67
SAN1B	Coding sequence	87
	Intron #1	14
	Intron #2	51
SAN1A	Entire gene	64
SAN1C	Coding sequence	85
	Intron #1	33
	Intron #2	26
SAN1B	Entire gene	61
SAN1C	Coding sequence	84
	Intron #1	17
	Intron #2	28

sequence and the coding regions, *SAN1A* and *SAN1B* appear to be slightly more similar than either of them is to *SAN1C*. However, a neighbour-joining tree including *SAN1A*, *SAN1B*, and *SAN1C* coding sequences suggests that *SAN1B* and *SAN1C* are most closely related, suggesting they could have arisen from a more recent duplication event and that *SAN1A* may have originated earlier than either of *SAN1B* and *SAN1C* (data not shown). A comparison of the sequences of intron #1 showed that *SAN1A* and *SAN1C* share more identical nucleotides than either of them does with *SAN1B*. Intron #2 is much more highly conserved between *SAN1A* and *SAN1B* than between either of *SAN1A* and *SAN1C* or *SAN1B* and *SAN1C*. The results of the intron sequence comparisons suggest that gene conversion may have occurred in the evolutionary history of the *SAN1* genes (see Section 3.4.5.2). Although they are AT-rich, most of the intron comparisons indicate little sequence similarity. In some cases, this may be due to limitations in alignment software when comparing introns of different sizes, for example the first intron of *SAN1B* is much smaller than that of *SAN1A*. In addition, a lack of functional constraint on introns compared to coding regions may be responsible for this observation (Graur and Li, 2000). However, it is interesting that the second introns of *SAN1A* and *SAN1B* exhibit such a high level of sequence conservation; this observation may suggest that they contain regulatory elements important for the expression of these genes. Table VIII (Section 3.4.3) shows that there are sequences corresponding to those of regulatory elements shared by the second introns of *SAN1A* and *SAN1B*. Another possibility for the high level of sequence conservation between the second introns of *SAN1A* and *SAN1B* is alternative splicing to give rise to additional proteins with functional differences. However, translation of the *SAN1A* and *SAN1B* cDNA sequences with the second introns suggests that this is not the case due to the presence of many stop codons within the introns.

The predicted coding sequences of the *SAN1* genes from cv Resnik, obtained by alignment of their genomic DNA sequences with the coding sequence of *SAN1A* from cv Maple Arrow, and the coding sequence of *SAN1A* from cv Maple Arrow, were translated using DNAMAN Version 4.22. *SAN1A* and *SAN1B* are predicted to produce proteins of 352 and 353 amino acids, respectively (Figure 20 A,B). However, translation of *SAN1C* suggests that it produces a truncated protein of 126 amino acids since it contains a stop codon 379 bp downstream of the start codon (Figure 20C). The presence of the stop codon is not likely to be due to sequence error, since the gene was sequenced by two independent methods (Section 3.2.2), and both methods produced identical sequences. The presence of several more in-frame stop codons after this initial one, suggests that read through during translation is unlikely. Therefore, the longest uninterrupted ORF of 379 bp gives rise to a truncated SAN1C protein of 126 amino acids, if it is translated at all, and *SAN1C* could represent a pseudogene. Interestingly, between the stop codon and the putative polyadenylation sequence (Section 3.2.4), a motif corresponding to that of an internal ribosome binding site was identified by scanning the UTRdb (Mignone *et al.*, 2005). The existence of this motif suggests that translation of another ORF might be possible. There are several other ORFs, but BLASTP searches of the translated ORFs did not reveal any similarities to known proteins, although this does not rule out the possibility of a novel protein being produced.

Comparison of the predicted protein sequences of *SAN1A* and *SAN1B* reveals an overall identity of 83% and similarity of 89% (Figure 21A). If the first 126 amino acids of all three SAN1 predicted proteins are compared, an identity of 88% and a similarity of 95% are revealed (Figure 21B). However, over the first 126 amino acids, the SAN1A and SAN1B proteins are 86% identical and 89% similar, whereas each of SAN1A and SAN1B

Figure 20: Predicted Protein Sequences of *SAN1* Genes of Soybean cv Resnik

A. Translation of *SAN1A* shown with its coding sequence.

B. Translation of *SAN1B* shown with its coding sequence.

C. Translation of *SAN1C* shown with its coding sequence.

The predicted amino acid sequences are highlighted in yellow.

A.

```
1          ATGGGAGAGGTTGACCCAGCTTTCATTCAAGAACCACAACACAGGCCAAATCTCTCCACC
1           M  G  E  V  D  P  A  F  I  Q  E  P  Q  H  R  P  N  L  S  T

61         ATTCAAGCAGAAGGAATTCCCATAATTGATCTCTCTCCAATAACCAACCACACAGTTTCA
21          I  Q  A  E  G  I  P  I  I  D  L  S  P  I  T  N  H  T  V  S

121        GACCCTTCTGCAATTGAAAGCTTGGTGAAGGAGATAGGACGTGCATGCCAGGAGTGGGGC
41          D  P  S  A  I  E  S  L  V  K  E  I  G  R  A  C  Q  E  W  G

181        TTCTTCCAAGTAACAAACCATGGAGTGCCACTCACTCTAAGGCAAAACATTGAGAAAGCC
61          F  F  Q  V  T  N  H  G  V  P  L  T  L  R  Q  N  I  E  K  A

241        TCAAAACTGTTCTTTGCTCAGACTCTGGAGGAGAAGAGAAAGGTTAGCAGAAATGAGAGC
81          S  K  L  F  F  A  Q  T  L  E  E  K  R  K  V  S  R  N  E  S

301        TCTCCAATGGGTTATTATGACACAGAACACACCAAGAACGTCAGGGACTGGAAAGAAGTC
101         S  P  M  G  Y  Y  D  T  E  H  T  K  N  V  R  D  W  K  E  V

361        TTTGATTTTCTAGCCAAAGACCCCACTTTCATTCCTCTCACTTCTGATGAACATGATGAT
121         F  D  F  L  A  K  D  P  T  F  I  P  L  T  S  D  E  H  D  D

421        CGAGTCAATCAATGGACTAATCAATCACCTCAATACCCTCCACTCTTCAGGGTTGTAACA
141         R  V  N  Q  W  T  N  Q  S  P  Q  Y  P  P  L  F  R  V  V  T

481        CAAGAGTATATTCAGGAGATGGAAAAGCTGTCCTTTAAGCTGTTGGAGCTTATAGCTTTG
161         Q  E  Y  I  Q  E  M  E  K  L  S  F  K  L  L  E  L  I  A  L

541        AGCTTAGGCCTTGAAGCAAAGAGGTTTGAGGAATTTTTCATCAAAGATCAAACTAGCTTT
181         S  L  G  L  E  A  K  R  F  E  E  F  F  I  K  D  Q  T  S  F

601        ATTCGACTCAACCACTATCCTCCATGCCCTTACCCTGACCTTGCTCTTGGCGTCGGTCGA
201         I  R  L  N  H  Y  P  P  C  P  Y  P  D  L  A  L  G  V  G  R

661        CACAAGGACCCCGGTGCCTTGACCATTCTTGCACAGGACGAAGTTGGAGGACTTGAAGTG
221         H  K  D  P  G  A  L  T  I  L  A  Q  D  E  V  G  G  L  E  V

721        AGACGTAAACGGGATCAAGAGTGGATTAGAGTGAAACCAACCCCAAATGCTTATATTATC
241         R  R  K  R  D  Q  E  W  I  R  V  K  P  T  P  N  A  Y  I  I

781        AACATTGGTGATACTGTTCAGGTTTGGAGCAATGATGCATATGAGAGTGTGGATCACAGA
261         N  I  G  D  T  V  Q  V  W  S  N  D  A  Y  E  S  V  D  H  R

841        GTGGTGGTCAACTCTGAGAAGGAAAGGCTTTCCATTCCGTTCTTCTTCTTCCCTGCACAC
281         V  V  V  N  S  E  K  E  R  L  S  I  P  F  F  F  F  P  A  H

901        GACACCAAAGTAAAGCCTTTGGAGGAGCTGATAAATGAGCAAAACCCTTCAAAATATAGG
301         D  T  K  V  K  P  L  E  E  L  I  N  E  Q  N  P  S  K  Y  R

961        CCATACAATTGGGGCAAGTTTCTTGTCCACAGAGGGAACAGCAATTTCAAGAAGCAAAAT
321         P  Y  N  W  G  K  F  L  V  H  R  G  N  S  N  F  K  K  Q  N

1021       GAGGAGAACATTCAAATTCATCATTACAAGATAGCTTAA
341         E  E  N  I  Q  I  H  H  Y  K  I  A  *
```

B.

```
1        ATGGAAAAGGTTGACCAAGCTTTCATCCAATACCCACAACACAGGCCAAATCTCACTATC
1         M  E  K  V  D  Q  A  F  I  Q  Y  P  Q  H  R  P  N  L  T  I

61       ATCCAACCTGAACACATTCCCATAATAGACCTCTCTCCAATAACCAACCACACAGTTTCA
21        I  Q  P  E  H  I  P  I  I  D  L  S  P  I  T  N  H  T  V  S

121      CATCCATCTCCCATTGAAGGCTTGGTGAAGGAGATAGGAAATGCATGCAAAGAGTGGGGA
41        H  P  S  P  I  E  G  L  V  K  E  I  G  N  A  C  K  E  W  G

181      TTCTTCCAAGTTATCAACCATGGGGTGTCCCTCACTCTAAGGCAAAACATTGAGAAAGCC
61        F  F  Q  V  I  N  H  G  V  S  L  T  L  R  Q  N  I  E  K  A

241      TCAAAATTGTTCTTTGCTCAGAGTTTAGAAGAGAAGAGGAAGGTTAGCAGAGATGAGAGT
81        S  K  L  F  F  A  Q  S  L  E  E  K  R  K  V  S  R  D  E  S

301      TCTCCCATGGGTTATTATGACACAGAGCACACAAAAAACATTAGGGACTGGAAAGAAGTG
101       S  P  M  G  Y  Y  D  T  E  H  T  K  N  I  R  D  W  K  E  V

361      TTTGATTTTCTAGCCAAAGACCCTACTTTTGTTCCTCTCACTTCGGATGAACATGATAAT
121       F  D  F  L  A  K  D  P  T  F  V  P  L  T  S  D  E  H  D  N

421      CGTCTCACTCAATGGACTAATCCATCCCCTCAATACCCTCCACACTTCAGGGATATAATT
141       R  L  T  Q  W  T  N  P  S  P  Q  Y  P  P  H  F  R  D  I  I

481      AAAGAGTATGTTGAAGAGATGGAAAAGCTGTCCTTTAAGTTGATGGAGCTTATAGCTTTG
161       K  E  Y  V  E  E  M  E  K  L  S  F  K  L  M  E  L  I  A  L

541      AATTTAGGCCTTGAGGCAAAGAGGTTTGAGGAATTTTTCATCAAAGATCAGACTAGCTTT
181       N  L  G  L  E  A  K  R  F  E  E  F  F  I  K  D  Q  T  S  F

601      CTTAGACTCAACTACTATCCTCCATGCCCTTACCCTCACCTTGCCCTTGGTGTTGGTCGA
201       L  R  L  N  Y  Y  P  P  C  P  Y  P  H  L  A  L  G  V  G  R

661      CATAAGGACTCTGGTCCCTTAACCATTCTTGCACAAGATGAAGTTGGAGGACTTGAAGTG
221       H  K  D  S  G  P  L  T  I  L  A  Q  D  E  V  G  G  L  E  V

721      AGACCTAAAGCAGATCAAGACTGGATAAGAGTGAAACCTATTCCAAATGCTTATATTATC
241       R  P  K  A  D  Q  D  W  I  R  V  K  P  I  P  N  A  Y  I  I

781      AACCTTGGTGATATGATTCAGGTTTGGAGCAATGATGCCTATGAGAGTGTGGAACACAGA
261       N  L  G  D  M  I  Q  V  W  S  N  D  A  Y  E  S  V  E  H  R

841      GTGGTGGTCAACTCTGAGAAAGCAAGGTTTTCCATTCCATTCTTCCTCTTCCCTGCACAT
281       V  V  V  N  S  E  K  A  R  F  S  I  P  F  F  L  F  P  A  H

901      GACACTGTAGTCAAGCCTTTGGAGGAGCTAATAAATGAGCAAAACCCTTCAAAATTTAGG
301       D  T  V  V  K  P  L  E  E  L  I  N  E  Q  N  P  S  K  F  R

961      CCATACAAATGGGGCAAGTTTCTTGTCCACAGACTGGATAGCAATTTCAAGAAACAAAAT
321       P  Y  K  W  G  K  F  L  V  H  R  L  D  S  N  F  K  K  Q  N

1021     GCGGAGAATGTTCAAACTTATCATTATAAGATAAGAGAGTAG
341       A  E  N  V  Q  T  Y  H  Y  K  I  R  E  *
```

C.

```
1       ATGGGAGAGGTCGACCCAGCTTTCATCCAATACCTACAACACAGGCCAAAGTTCTCTACC
1        M  G  E  V  D  P  A  F  I  Q  Y  L  Q  H  R  P  K  F  S  T

61      ATCCAACCTGAAGACATTCCTGTGATAGATATCTCCCCAATAACCAACCACACAGCTTCA
21       I  Q  P  E  D  I  P  V  I  D  I  S  P  I  T  N  H  T  A  S

121     GATCCATGTTCTATTGGAGGCTTAGTGAAGGAGATAAGAAATGCAAGCAAGGAGTGGGGT
41       D  P  C  S  I  G  G  L  V  K  E  I  R  N  A  S  K  E  W  G

181     TTCTTCCAATTATCAAACCATGCTGTGCCCCTCACTCTAAGGCAAAACATTGAGAAAGCC
61       F  F  Q  L  S  N  H  A  V  P  L  T  L  R  Q  N  I  E  K  A

241     TCAAGATTGTTCTTTGCACAGAATTTGGAGGAGAAGAGGAAGGTTAGTAAAGATGAGAAC
81       S  R  L  F  F  A  Q  N  L  E  E  K  R  K  V  S  K  D  E  N

301     AATACCACCGGTTACCATGACAGAGAGCACCCCAAGAACATTAGGGACTGGAAAGAAGTG
101      N  T  T  G  Y  H  D  R  E  H  P  K  N  I  R  D  W  K  E  V

361     TTTGATTTCCTAGCCAAATAATCCACTTTCATTCCTCTCACTTCTGATGAACATGATGAT
121      F  D  F  L  A  K  *  S  T  F  I  P  L  T  S  D  E  H  D  D

421     CGACTCACTCAGTTTGACTAATCAATCTCCTTAATACCTTCAAAATTTCAGGGTTATAAT
141      R  L  T  Q  F  D  *  S  I  S  L  I  P  S  K  F  Q  G  Y  N

481     ACAAGAGTATATTCAAGAGATGGAAAAGCTGTGCTTTAAGCTGTTGGAGCTTATAGCTTT
161      T  R  V  Y  S  R  D  G  K  A  V  L  *  A  V  G  A  Y  S  F

541     AAGCTTAGGTGTTGAAGCAAAGAGGTTTGAAGAACTTTTCATCAAAGATCAAACTAGTTC
181      K  L  R  C  *  S  K  E  V  *  R  T  F  H  Q  R  S  N  *  F

601     TATTCTACTCAACTGCTATCCTCCATGCCCTTACCCTCACCTTGCTCATGGCGTTGGTTG
201      Y  S  T  Q  L  L  S  S  M  P  L  P  S  P  C  S  W  R  W  L

661     ACACAAGGACCCTGGTGCCTTGACCATTCTTGCACAAGATGAGGTTGGAGGACTTGGAGT
221      T  Q  G  P  W  C  L  D  H  S  C  T  R  *  G  W  R  T  W  S

721     GAAACGTAAAGCTGATCAAGAGTGAAACCAACCCCAGATGCTTATACTATCAACTTTGGT
241      E  T  *  S  *  S  R  V  K  P  T  P  D  A  Y  T  I  N  F  G

781     GACATTATTCAGGTTTGGAGCAATGATGCATATGAGAGTGTGGAACACAGAGTGGTGGTC
261      D  I  I  Q  V  W  S  N  D  A  Y  E  S  V  E  H  R  V  V  V

841     AACTCTGAGAAAGAAAGGTTCTCCATTCCATTCTTCTACCTTTCGCATGAAACTGAGGTC
281      N  S  E  K  E  R  F  S  I  P  F  F  Y  L  S  H  E  T  E  V

901     AAGCCTTTGGAGAAGCTGATAAATGAGCAAAACCCTTGAAAATATAGGCCATACAAATGG
301      K  P  L  E  K  L  I  N  E  Q  N  P  *  K  Y  R  P  Y  K  W

961     GGCAAGTTTCTCACCCATAGAAAGAATACCAATTTCAAGAAACAAAATGTGGAGAACCGT
321      G  K  F  L  T  H  R  K  N  T  N  F  K  K  Q  N  V  E  N  R

1021    CAAATTTATCATTATAAGCTTGCTTAA
341      Q  I  Y  H  Y  K  L  A  *
```

Figure 21: Comparison of the Predicted Amino Acid Sequences of the *SAN1* Genes

A. Alignment of SAN1A and SAN1B predicted protein sequences of soybean cv Resnik. Identical and similar amino acids are indicated by asterisks and periods, respectively.

B. Alignment of the first 126 amino acids of the SAN1A and SAN1B predicted proteins with the SAN1C truncated protein sequence of soybean cv Resnik. Identical and similar amino acids are indicated by asterisks and periods, respectively.

A.

```
SAN1Atrans   MGEVDPAFIQEPQHRPNLSTIQAEGIPIIDLSPITNHTVSDPSAIESLVKEIGRACQEWG
SAN1Btrans   MEKVDQAFIQYPQHRPNLTIIQPEHIPIIDLSPITNHTVSHPSPIEGLVKEIGNACKEWG
             *  ** **** *******. ** * *************** ** ** ****** **.***

SAN1Atrans   FFQVTNHGVPLTLRQNIEKASKLFFAQTLEEKRKVSRNESSPMGYYDTEHTKNVRDWKEV
SAN1Btrans   FFQVINHGVSLTLRQNIEKASKLFFAQSLEEKRKVSRDESSPMGYYDTEHTKNIRDWKEV
             **** **** *****************.********* ***************.******

SAN1Atrans   FDFLAKDPTFIPLTSDEHDDRVNQWTNQSPQYPPLFRVVTQEYIQEMEKLSFKLLELIAL
SAN1Btrans   FDFLAKDPTFVPLTSDEHDNRLTQWTNPSPQYPPHFRDIIKEYVEEMEKLSFKLMELIAL
             **********.******** *..**** ****** ** . .**..*********.*****

SAN1Atrans   SLGLEAKRFEEFFIKDQTSFIRLNHYPPCPYPDLALGVGRHKDPGALTILAQDEVGGLEV
SAN1Btrans   NLGLEAKRFEEFFIKDQTSFLRLNYYPPCPYPHLALGVGRHKDSGPLTILAQDEVGGLEV
              ******************.*** ******* ********** * **************

SAN1Atrans   RRKRDQEWIRVKPTPNAYIINIGDTVQVWSNDAYESVDHRVVVNSEKERLSIPFFFFPAH
SAN1Btrans   RPKADQDWIRVKPIPNAYIINLGDMIQVWSNDAYESVEHRVVVNSEKARFSIPFFLFPAH
             * * **.****** *******.** .***********.********* * ***** ****

SAN1Atrans   DTKVKPLEELINEQNPSKYRPYNWGKFLVHRGNSNFKKQNEENIQIHHYKIA-
SAN1Btrans   DTVVKPLEELINEQNPSKFRPYKWGKFLVHRLDSNFKKQNAENVQTYHYKIRE
             ** ***************.*** ********  ******* **.*  ****
```

B.

```
SAN1Atrans   MGEVDPAFIQEPQHRPNLSTIQAEGIPIIDLSPITNHTVSDPSAIESLVKEIGRACQEWG
SAN1Btrans   MEKVDQAFIQYPQHRPNLTIIQPEHIPIIDLSPITNHTVSHPSPIEGLVKEIGNACKEWG
SAN1Ctrans   MGEVDPAFIQYLQHRPKFSTIQPEDIPVIDISPITNHTASDPCSIGGLVKEIRNASKEWG
             *  ** ****  ****  . ** * **.**.******* * *. *  *****  *..***

SAN1Atrans   FFQVTNHGVPLTLRQNIEKASKLFFAQTLEEKRKVSRNESSPMGYYDTEHTKNVRDWKEV
SAN1Btrans   FFQVINHGVSLTLRQNIEKASKLFFAQSLEEKRKVSRDESSPMGYYDTEHTKNIRDWKEV
SAN1Ctrans   FFQLSNHAVPLTLRQNIEKASRLFFAQNLEEKRKVSKDENNTTGYHDREHPKNIRDWKEV
             ***. ** * ***********.***** ********. *    ** * ** **.******

SAN1Atrans   FDFLAK
SAN1Btrans   FDFLAK
SAN1Ctrans   FDFLAK
             ******
```

Figure 22: Linear Range of *SAN1A* and *SAN1B* Amplification by RT-PCR

The linear ranges of *SAN1A* (A) and *SAN1B* (B) amplification were determined by plotting log (cpm) versus the number of cycles. The dashed lines show the number of cycles that is thought to represent the middle of the linear range of amplification. Each graph displays the results of two trials for each gene. For *SAN1A* and *SAN1B*, 21 cycles and 25 cycles were chosen for subsequent RT-PCR experiments, respectively.

cpm, counts per minute.

are only 75% identical to SAN1C. SAN1A and SAN1B are 84% and 82% similar to SAN1C, respectively. This data indicates that SAN1A and SAN1B appear to be more closely related to each other than either of them is to SAN1C.

When the *SAN1A* coding sequences of cvs Maple Arrow and Resnik were compared, it was found that they differ by two nucleotides (C↔T and A↔T) changes that give rise to two non-conservative differences in the predicted proteins (proline↔leucine and aspartate↔valine).

3.2.6 Summary of *SAN1* Gene Characteristics

In this study, three *SAN1* genes, *SAN1A*, *SAN1B*, and *SAN1C*, were isolated from the soybean cv Resnik. The genes form a multigene family and occur in tandem in the genome. Analysis of the genomic Southern blot data and *in silico* approaches has not allowed determination of the exact number of genes in the family, but there are at least three. In addition, the *SAN1* genes reside close to the *MSG* locus in the soybean genome.

Each of the *SAN1* genes consists of three exons and two introns. The sizes and positions of the exons are well conserved between the genes, but the intron sizes are not. The conserved bases expected for dicot plant nuclear intron splice sites are present in the introns of the three *SAN1* genes and the introns are generally AT-rich. Consensus branchpoint sequences were observed in some of the introns, but polypyrimidine tracts were not. Exon sequences that flank the intron sequences of the *SAN1* genes are also conserved. Therefore, the *SAN1* genes appear to contain typical plant nuclear introns. When the 3′ UTRs were examined for typical polyadenylation signals and sites, a potential NUE of sequence AATAAA was observed in all three *SAN1* genes of cv Resnik and *SAN1A* of Maple Arrow.

Overall DNA and amino acid sequence identity suggests that *SAN1A* and *SAN1B* are more closely related to one another than either of them is to *SAN1C*, although phylogenetic analysis suggests that *SAN1B* and *SAN1C* may be more closely related than either one is to *SAN1A*. *SAN1A* and *SAN1B* encode putative proteins of 352 and 353 amino acids, respectively. The longest ORF of *SAN1C* could only encode a protein of 126 amino acids, if it is translated at all, and it could represent a pseudogene. Two nucleotides differ between the sequences of the *SAN1A* genes of cvs Maple Arrow and Resnik which give rise to two amino acid differences in the predicted proteins.

CHAPTER 3

RESULTS: PART III

Expression of the *SAN1* Gene Family

The expression of the members of the soybean *SAN1* gene family was investigated using relative RT-PCR. Since the senescent phase of nodule development was the focus of this research, the expression of the *SAN1A* and *SAN1B* genes was investigated in nodules undergoing senescence. The levels and patterns of their expression were determined in naturally senescing nodules and in those in which senescence had been induced by nitrate and dark treatment. In addition, *SAN1A* expression was investigated throughout two light and dark cycles to determine whether its expression levels were influenced by light. The expression of *SAN1A*, *SAN1B*, and the pseudogene *SAN1C* was also monitored in the various tissues of soybean. Determination of changes in *SAN1C* expression in senescent nodule tissue was not pursued, since it encodes a truncated protein which is not likely to be functional.

3.3.1 Determination of *SAN1A* and *SAN1B* Message Levels

3.3.1.1 Rationale for Use of Relative RT-PCR

Initially, Northern analysis was chosen to monitor expression levels; however, low and irreproducible levels of expression were monitored for *SAN1A* and therefore a more sensitive approach was required. Relative RT-PCR (Ambion Inc., Austin, TX), a semi-quantitative, multiplex approach to determining expression levels, was chosen. By amplifying the target gene of interest, in this case *SAN1A* or *SAN1B*, along with an 18S rRNA control in the same reaction, the relative amounts of the target can be determined as a fraction of 18S rRNA (QuantumRNA Internal Standards Kit). 18S rRNA was chosen as

the control over housekeeping genes such as β-actin or NADPH oxidase because the rRNA levels do not vary. As an alternative to 18S rRNA, ribosomal protein genes could also have been employed. A more detailed description of the steps necessary to validate this method was provided in Section 2.14.5, while the following three sections present the results of the validation steps.

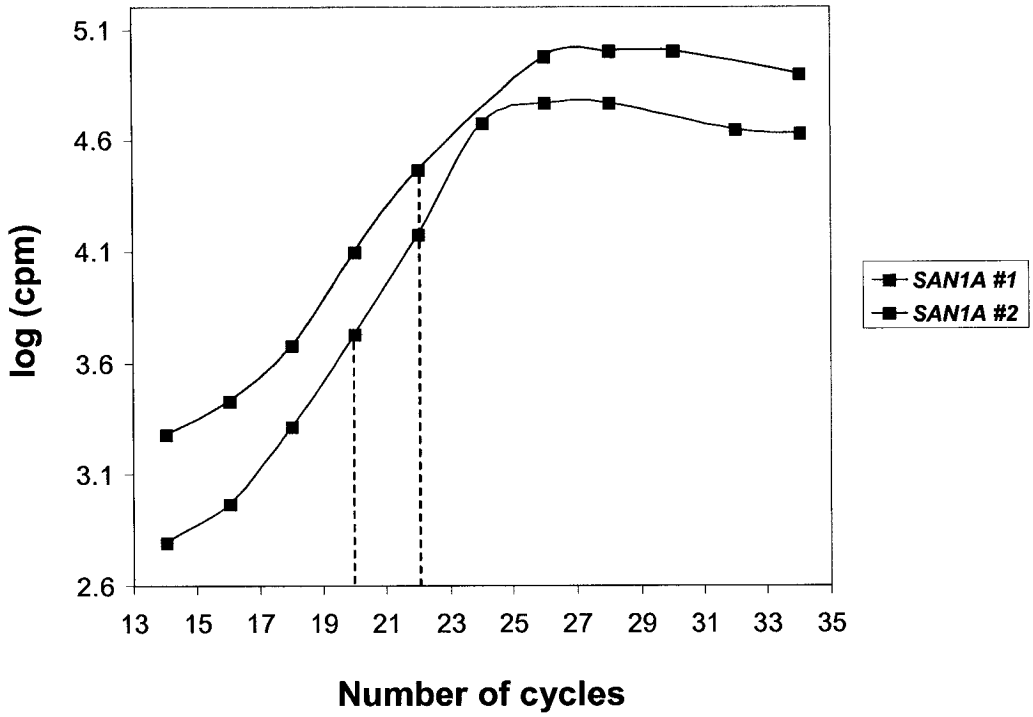
3.3.1.2 Verification of Identities of RT-PCR Products

The RT-PCR products amplified by both gene-specific and the proprietary 18S primers (Ambion Inc., Austin, TX) were cloned and sequenced to verify that the correct targets had been amplified. The *SAN1A* and *SAN1B* RT-PCR product clones were 100% identical to their predicted cDNA sequences, confirming that the correct products had been obtained (data not shown). The 18S rRNA RT-PCR product was 99% identical to the soybean 18S sequence (GenBank, accession X02623.1), implying that the correct product had been obtained (data not shown).

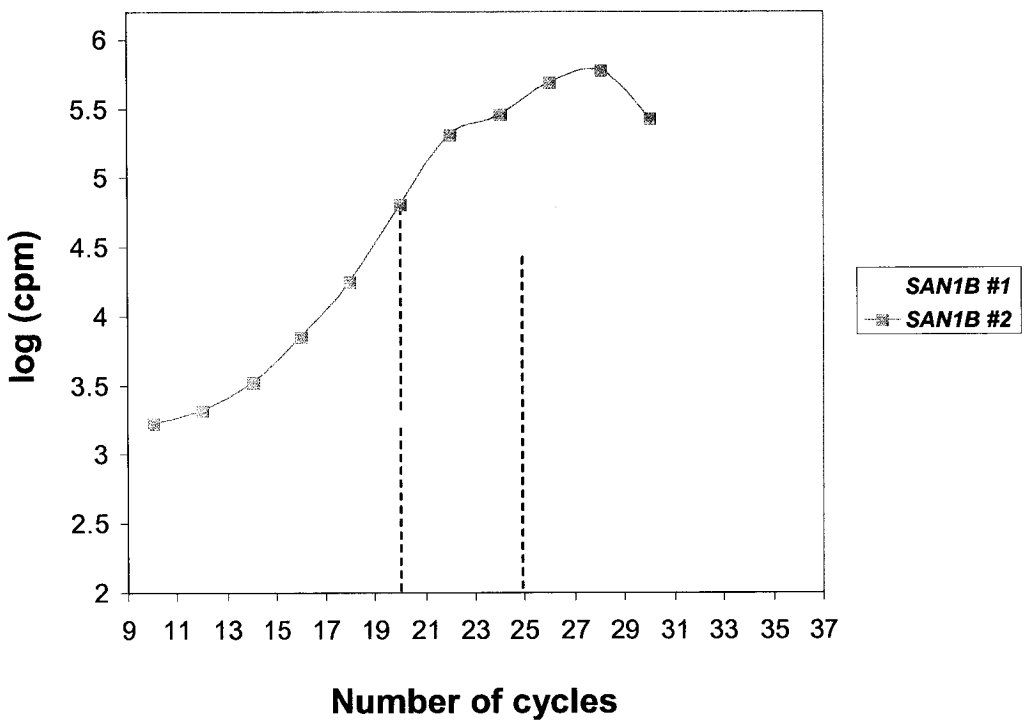
3.3.1.3 Linear Range of Amplification of *SAN1A* and *SAN1B*

In order to ensure that relative RT-PCR results were semi-quantitative, the linear ranges of amplification of the *SAN1A* and *SAN1B* target sequences were determined using RNA from tissues with median levels of target mRNA. For *SAN1A* and *SAN1B*, duplicate determinations with RNA from 24 dpi nodules were performed (Figure 22 A and B). For *SAN1A*, the average middle of the linear range occurs around 21 cycles (replicates of 22 and 20 cycles) and this value was used in subsequent experiments. For *SAN1B*, the first trial gave a linear range mid-point of 25, whereas the second one gave a mid-point of only 20 cycles. Since the first trial produced a more reliable curve with less scattered points, 25 cycles was chosen as the number of cycles to be used in subsequent *SAN1B* relative

A.


5.1
4.6
4.1
3.6
3.1
2.6
log (cpm)
13 15 17 19 21 23 25 27 29 31 33 35
Number of cycles
SAN1A #1
SAN1A #2

B.


6
5.5
5
4.5
4
3.5
3
2.5
2
log (cpm)
9 11 13 15 17 19 21 23 25 27 29 31 33 35 37
Number of cycles
SAN1B #1
SAN1B #2

RT-PCR experiments.

In addition, the range over which the amplification of *SAN1B* was linear was greater than that for *SAN1A*. *SAN1B* amplification was found to be linear over 12 (cycles 14 to 26) or 16 cycles (cycles 16 to 32), whereas *SAN1A* amplification was linear over 8 cycles (cycles 16 to 24 and cycles 18 to 26). This observation suggests that *SAN1A* is expressed at higher levels in the 24 dpi nodules than *SAN1B*.

3.3.1.4 Determination of Ratios of 18S Primers to Competimers

Since 18S is so highly expressed in all tissues, it was necessary to attenuate its accumulation by PCR so that it does not outcompete the target gene for available reagents in the multiplex reaction. In order to achieve this goal, Ambion Inc. (Austin, TX) has designed proprietary competimers which bind to the target 18S sequences, but which do not allow amplification from those sites. The ideal ratio of primers to competimers for a given target gene is one for which the resultant ratio of target to 18S is close to one (QuantumRNA Internal Standards Kit).

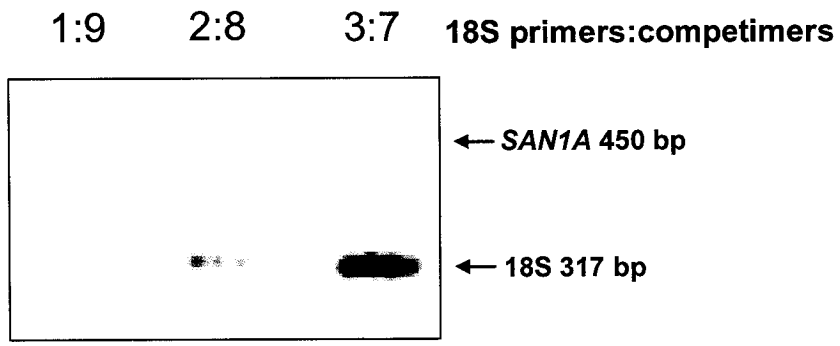
In order to determine experimentally the ratio of 18S primers to competimers to be used for *SAN1A* and *SAN1B* expression studies, tests with 1:9, 2:8, and 3:7 ratios were carried out (Figures 23A and 24A, respectively). *SAN1A* RT-PCR products were not visible on the gel for any of the ratios tested and the 18S band was strong when the 3:7 ratio was employed, fainter when the 2:8 ratio was used, and no longer visible at a ratio of 1:9. This result suggested that the correct ratio of 18S primers to competimers was likely to lie between 1:9 and 2:8. Like *SAN1A*, *SAN1B* DNA was not detected for any of the ratios tested. However, strong 18S bands were observed for all three ratios, suggesting that the ratio at which comparable levels of *SAN1B* and 18S would be achieved would be lower than 1:9.

Figure 23: Determination of 18S Primer: Competimer Ratio for *SAN1A* Expression Analyses by Relative RT-PCR

A. 1% agarose gel of multiplex RT-PCR products of *SAN1A* and 18S prepared from RNA extracted from 21 dpi nodules using 1:9, 2:8, and 3:7 ratios of 18S primers to competimers.

B. Relative RT-PCR to monitor *SAN1A* and 18S was carried out using ratios of 18S primers to competimers of 1:9, 1.5:8.5, and 2:8 on RNA extracted from root nodules harvested at 0, 12, 24, 36, 48, 72, and 96 hours after initiation of treatment with 20 mM NH_4NO_3. Results are presented as the *SAN1A*/18S signal obtained for each ratio of 18S primers to competimers versus hours of nitrate treatment.

A.


1:9
2:8
3:7
18S primers:competimers
SAN1A 450 bp
18S 317 bp

B.

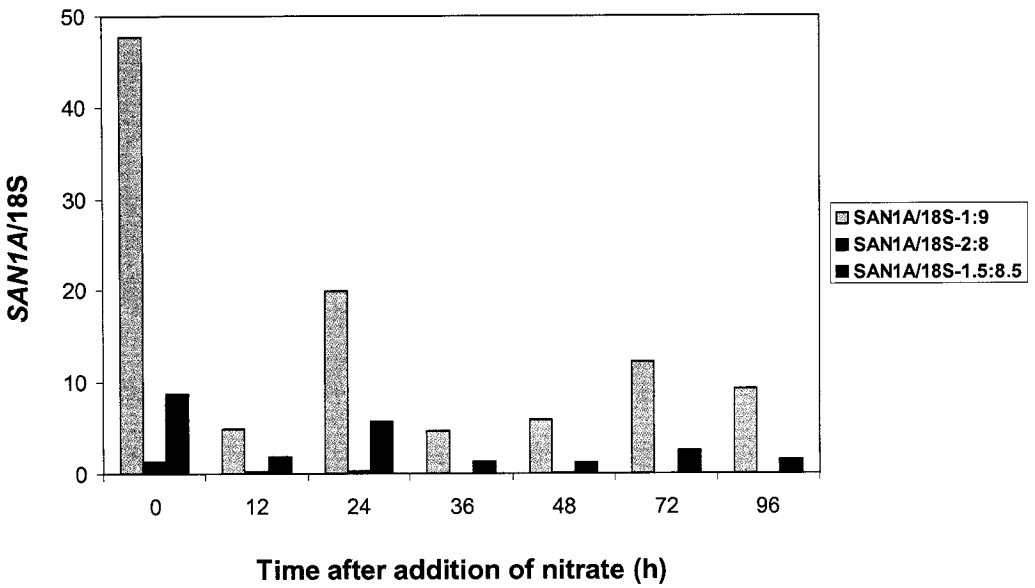
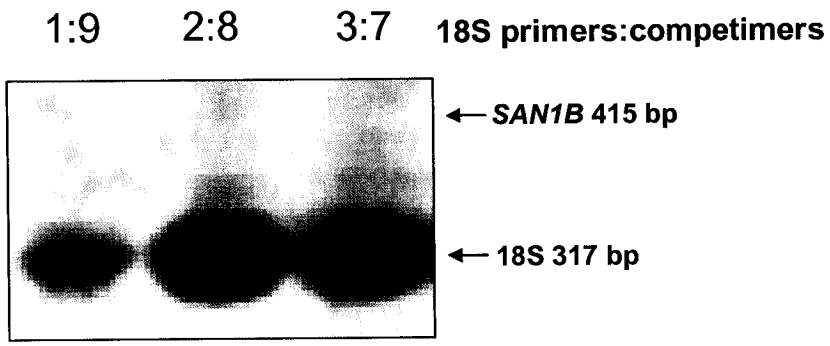

50
40
30
20
10
0
SAN1A/18S
0
12
24
36
48
72
96
Time after addition of nitrate (h)
SAN1A/18S-1:9
SAN1A/18S-2:8
SAN1A/18S-1.5:8.5

Figure 24: Determination of 18S Primer: Competimer Ratio for *SAN1B* Expression Analyses by Relative RT-PCR

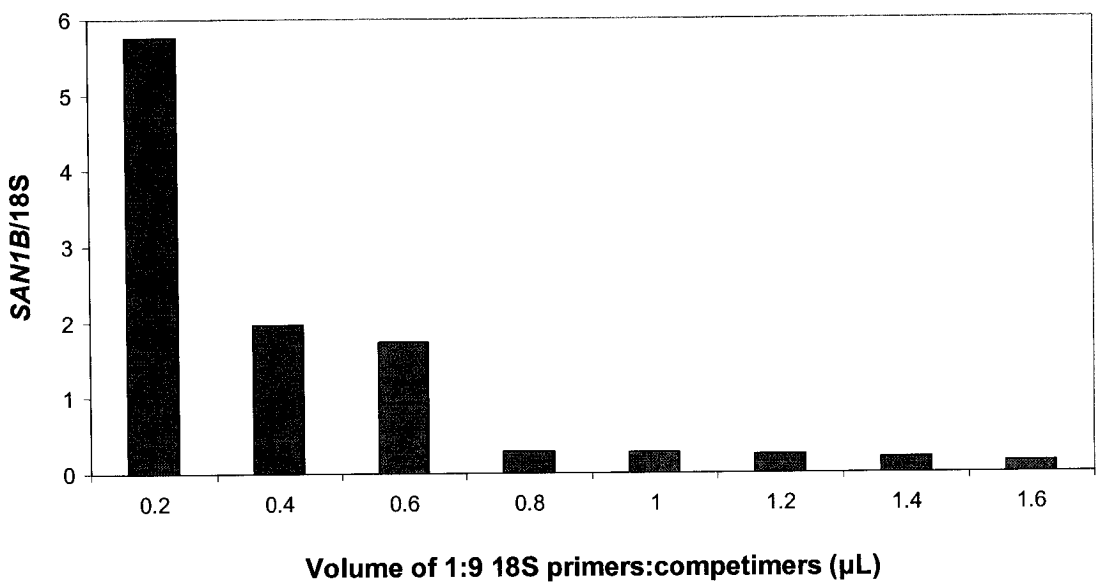
A. 1% agarose gel of multiplex RT-PCR products of *SAN1B* and 18S prepared from RNA extracted from 21 dpi nodules using 1:9, 2:8, and 3:7 ratios of 18S primers to competimers.

B. Relative RT-PCR to monitor *SAN1B* and 18S was carried out using different amounts of 1:9 18S primers to competimers on RNA extracted from root nodules harvested from 21 dpi root nodules. Results are presented as the *SAN1B*/18S signal obtained for each amount of 1:9 18S primers to competimers.

A.


1:9
2:8
3:7
18S primers:competimers
SAN1B 415 bp
18S 317 bp

B.


SAN1B/18S
0
1
2
3
4
5
6
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
Volume of 1:9 18S primers:competimers (µL)

Since the initial ratio tests implied that the correct ratio of 18S primers to competimers for determination of *SAN1A* levels was likely to be between the ratios of 2:8 and 1:9, the ratios of 2:8, 1:9, and an intermediate ratio of 1.5:8.5 were tested on a set of RNA extracted from nodules undergoing nitrate-induced senescence (Figure 23B). The 1.5:8.5 ratio gave results in which the *SAN1A*/18S signal ranged between 1.2 and 8.8, whereas the 1:9 ratio produced a range of 4.7 and 47.7 and the 2:8 ratio produced a range of 0.01 and 1.34. The 1.5:8.5 18S primer to competimer ratio resulted in *SAN1A*/18S ratios that were between one and ten for the set of nitrate-induced RNA samples, whereas the 1:9 ratio resulted in *SAN1A*/18S ratios that were mostly smaller than one and the 2:8 ratio resulted in *SAN1A*/18S ratios that were much higher than one. Therefore, the 1.5:8.5 ratio of 18S primers to competimers was chosen for subsequent determinations of *SAN1A* expression patterns by relative RT-PCR.

A similar experiment was carried for *SAN1B*. Since the correct ratio of 18S primers to competimers appeared to be lower than 1:9 (Figure 24A), it was necessary to either lower the ratio further or to decrease the volume of 1:9 18S primers and competimers added to the 20 μL PCR reaction to reduce the 18S signal. The latter method was chosen and eight different volumes were tested with one cDNA sample prepared from 21 dpi nodule RNA (Figure 24B). Since 0.6 μL of 18S 1:9 primers to competimers gave a ratio of *SAN1B*/18S signal of 1.74, one at which *SAN1B* and 18S are amplified at comparable levels, this volume of 1:9 18S primers to competimers was used in subsequent experiments.

To recapitulate, the ratio of 18S primers to competimers used for the concomitant amplification of *SAN1A* and 18S was 1.5:8.5. For multiplex PCR of *SAN1B* and 18S, a ratio of 1:9 was employed and the overall amount of the primer mix was reduced from

1.6 μL to 0.6 μL in a 20 μL PCR reaction. Like those of the linear range determinations for *SAN1A* and *SAN1B*, these results imply that the expression levels of *SAN1A* are higher than that of *SAN1B*.

3.3.1.5 *SAN1A* and *SAN1B* Expression

SAN1A and *SAN1B* expression was monitored using relative RT-PCR in naturally senescing nodules and in nodules induced to senesce by nitrate and dark treatments. In addition, the expression of *SAN1A* during light-dark cycling was also carried out. Finally, the tissue distribution of the *SAN1A* and *SAN1B* messages were also determined.

3.3.1.5.1 *SAN1A* and *SAN1B* Message Levels in Nodules during Natural Soybean Development

SAN1A and *SAN1B* expression was monitored in naturally senescing soybean nodules between 7 and 70 dpi with *Bradyrhizobium japonicum*. (Figures 25 A,B and 26 A,B). At 7 and 10 dpi, no nodules were visible and therefore RNA was prepared from inoculated roots. At 15 dpi, small pink nodules began to appear. Therefore, between 15 and 70 dpi with *B. japonicum*, expression was monitored in nodules.

For both *SAN1A* and *SAN1B*, no significant differences in expression were observed between 7 and 70 dpi with *B. japonicum* (one-way ANOVA, $p=0.696$ for *SAN1A*; one-way ANOVA, $p=0.086$ for *SAN1B*) (Figures 25A and 26A, respectively). However, the trends that are observed on the graphs and Southern blots imply that *SAN1A* and *SAN1B* behaved differently in terms of their expression throughout nodule development.

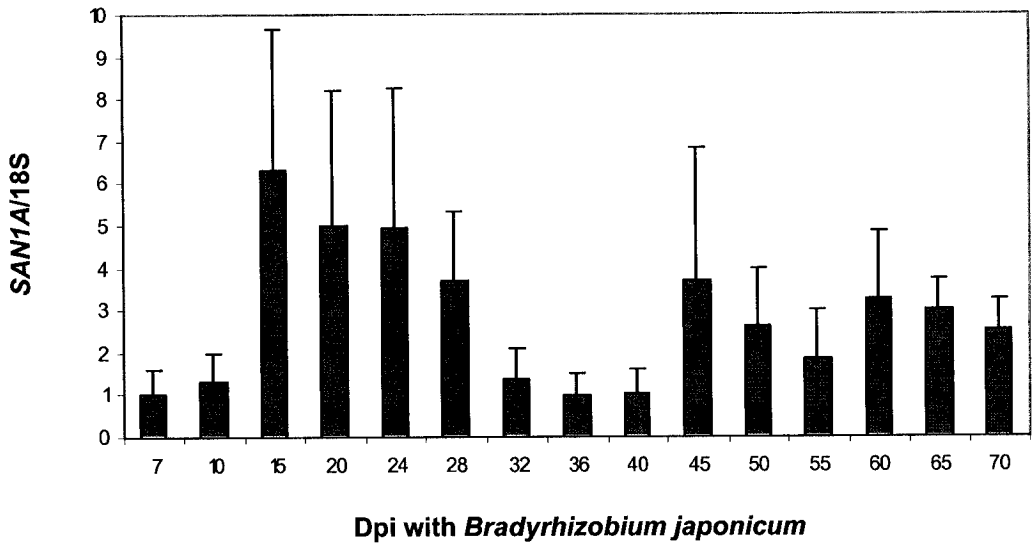
In general, *SAN1A* expression appeared to be low in the inoculated roots, but increased once nodules have been formed at 15 dpi (Figure 25A). Message levels remained relatively high between 15 and 28 dpi, but then exhibited lower levels between

Figure 25: *SAN1A* Expression in Nodules during Natural Development

A. Relative RT-PCR was used to monitor *SAN1A* and 18S expression in RNA extracted from root tissue harvested at 7 and 10 dpi with *Bradyrhizobium japonicum*, and from root nodules harvested at 12, 17, 22, 24, 28, 32, 36, 40, 45, 50, 55, 65, and 70 dpi with *B. japonicum*. Results are presented as the mean ± SE of the *SAN1A*/18S signal versus dpi with *B. japonicum*. Each time point represents the mean *SAN1A*/18S signals of nodules collected from three different sets of plants. For each set of plants, nodules were collected from 4-5 plants for each time point.

B. Southern blots of the relative RT-PCR products hybridized sequentially with ^{32}P-labelled 458 bp *SAN1A* cDNA and 317 bp 18S cDNA probes under standard conditions (Chapter 2, Sections 11 and 12.1). Twenty μL of PCR product was electrophoresed on a 1% agarose gel and transferred to Hybond-N membrane (Amersham Biosciences UK Ltd., Buckinghamshire, UK). The pictures represent exposures to a phosphor screen (Bio-Rad Laboratories, Hercules, CA).

A.


SAN1A/18S
0
1
2
3
4
5
6
7
8
9
10
7
10
15
20
24
28
32
36
40
45
50
55
60
65
70
Dpi with Bradyrhizobium japonicum

B.

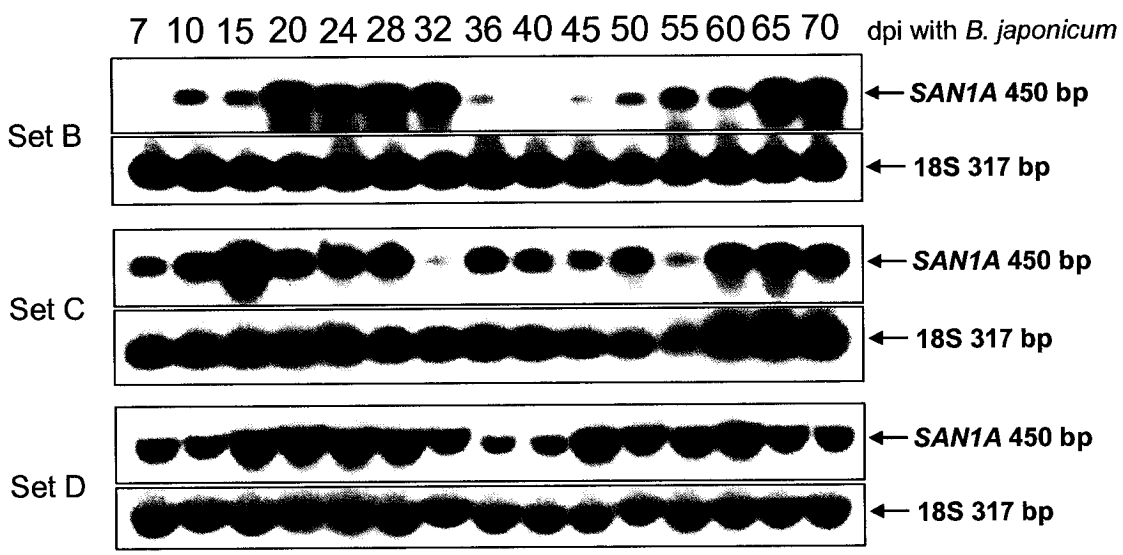
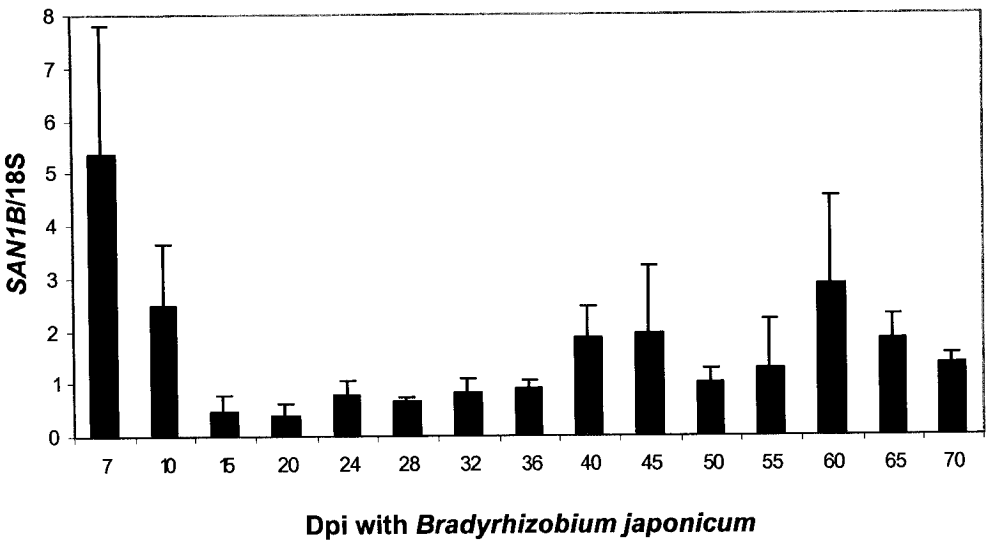

7 10 15 20 24 28 32 36 40 45 50 55 60 65 70 dpi with B. japonicum
Set B
SAN1A 450 bp
18S 317 bp
Set C
SAN1A 450 bp
18S 317 bp
Set D
SAN1A 450 bp
18S 317 bp

Figure 26: *SAN1B* Expression in Nodules during Natural Development

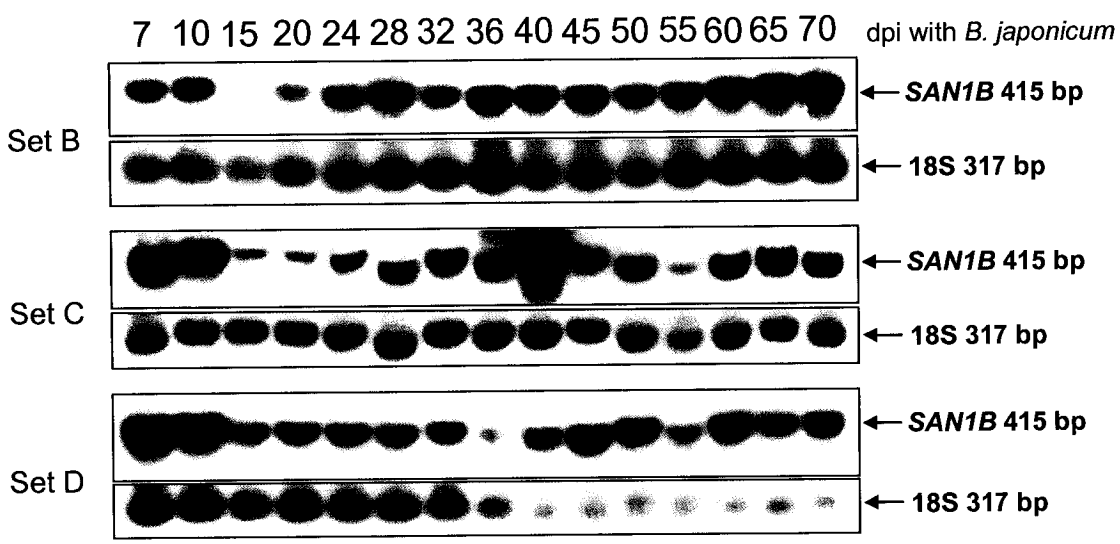
A. Relative RT-PCR was used to monitor *SAN1B* and 18S expression in RNA extracted from root and nodule tissue of soybean during natural development. Refer to the caption of Figure 25 (A) for details.

B. Southern blots of the relative RT-PCR products hybridized sequentially with ^{32}P-labelled 415 bp *SAN1B* and 317 bp 18S cDNA probes under standard conditions (Chapter 2, Sections 11 and 12.1). Twenty μL of PCR product was electrophoresed on a 1% agarose gel and transferred to Hybond-N membrane (Amersham Biosciences UK Ltd., Buckinghamshire, UK). The pictures represent exposures to a phosphor screen (Bio-Rad Laboratories, Hercules, CA).

A.


SAN1B/18S
0
1
2
3
4
5
6
7
8
7
10
15
20
24
28
32
36
40
45
50
55
60
65
70
Dpi with Bradyrhizobium japonicum

B.


7 10 15 20 24 28 32 36 40 45 50 55 60 65 70
dpi with B. japonicum
Set B
SAN1B 415 bp
18S 317 bp
Set C
SAN1B 415 bp
18S 317 bp
Set D
SAN1B 415 bp
18S 317 bp

32 and 40 dpi. By 45 dpi, the message of *SAN1A* had increased again and remained at similar levels until 70 dpi. *SAN1A* appeared to be most highly expressed early after nodule formation and then again during senescence after 40 dpi. The individual results for each set of plants tested are also shown (Figure 25B). In all three sets, the expression of *SAN1A* was low in *Bradyrhizobium*-inoculated roots prior to the appearance of nodules, higher when nodules first began fixing nitrogen (15 or 20 dpi) until after the onset of senescence at 24 dpi (between 15/20 dpi and 28/32 dpi), and low again until later in senescence, when expression levels rose a second time. The second rise in levels of *SAN1A* expression varied between the three sets. For example, it occurred at 65, 50, and 45 dpi in Sets B, C, and D, respectively. In addition, one set (B) appeared to lag behind the other two (C,D) in development since the first large increase in *SAN1A* expression was not observed until 20 dpi, rather than at 15 dpi. Large amounts of variability in gene expression between naturally grown sets of plants were also observed when Lb mRNA was monitored throughout the development of these three sets of soybean plants (Section 3.1.2.2.1).

By contrast, *SAN1B* expression appeared to be higher in inoculated roots than in pre-senescent nodules (Figure 26A). Low levels of expression were observed between 15 and 36 dpi, but by 40 dpi, message levels had increased again. At 50 dpi, a drop in *SAN1B* messenger RNA levels was noted, but this was followed by an increase to 60 dpi. After 60 dpi, a slight decline in message levels occurred. It appears that *SAN1B* was expressed at higher levels in inoculated roots than in nodules and that it was more highly expressed in senescing nodules than in younger nodules. Figure 26B shows the results for each individual set of plants tested for *SAN1B* expression. All three sets showed similar patterns, although the timing was slightly different in each set. Expression of *SAN1B* was

much higher in inoculated roots at 7 and 10 dpi than it was in young nodules at 15 dpi. After 15 dpi, gradual increases in *SAN1B* expression were observed until 40 (B,C) or 45 dpi (D). After this point in development, slight decreases followed by second increases in expression were observed. Another decrease in expression was observed in Set D between 60 and 70 dpi. This same decline was not observed in Sets B and C. In addition, the highest levels of *SAN1B* expression were obtained at 70 dpi in Set B, whereas the highest levels were obtained at 7 dpi in Sets C and D. Overall, the patterns of the *SAN1B* gene expression between the three natural sets appeared similar, but there was clearly variability in gene expression between the three sets, making it difficult to observe significant differences in message levels.

That there was no overall statistically significant changes in expression of *SAN1A* and *SAN1B* throughout nodule development was not surprising given that the population of nodules on a given plant may be varied in terms of age, giving rise to large variations in gene expression between sets of plants at a given time point. Re-infection by *Bradyrhizobium* may occur, causing increased variation in the stages of development represented in the nodule population. These results were similar to the results obtained when Lb expression was monitored in naturally senescing nodules (Section 3.1.2.2.1), suggesting that changes in gene expression may be difficult to detect in a naturally senescing population of root nodules.

Provided that regulation of these genes occurs chiefly at the level of transcription, these results indicate that both SAN1A and SAN1B may have roles in senescence, but that SAN1A is more likely than SAN1B to have a role throughout all the stages of nodule development.

3.3.1.5.2 *SAN1A* and *SAN1B* Expression in Nodules of Nitrate-Treated Plants

SAN1A and *SAN1B* expression was monitored in nodules of soybean plants at 0, 12, 24, 36, 48, 72, and 96 hours after the initiation of treatment with 20 mM NH_4NO_3. Zero hours corresponded to 8:00 am for a 21 dpi plant (Figure 27 A,B,C,D).

Both *SAN1A* and *SAN1B* expression changed in a significant manner during the 96 hours of nitrate treatment. Tukey's test ($p<0.05$) indicated that *SAN1A* expression declined significantly between 0 and 12 hours of nitrate treatment (Figure 27A), suggesting that nitrate treatment, either directly or indirectly, represses *SAN1A* expression. However, there appeared to be a slight recovery of the message levels by 24 hours, followed by another sharp decline by 36 hours (Figure 27 A and B). Another slight increase was observed at 48 hours, but between 36 and 96 hours of treatment, levels of *SAN1A* message were statistically similar. Overall, nitrate treatment caused *SAN1A* expression to decrease to approximately 10% of its original level by 96 hours. The decline and recovery of the *SAN1A* mRNA may indicate light regulation of *SAN1A* since the pattern of expression seemed to correlate with the different stages in the light-dark cycle (Figure 27 A) under which the plants were grown. This hypothesis was tested as described in the next section.

SAN1B behaved very differently than *SAN1A* in terms of its expression during nitrate treatment (Figure 27 C and D). During the 96 hours of nitrate treatment, there was a clear, although slow increase of *SAN1B* expression. By 96 hours, a significant change in *SAN1B* message levels from 0 hours had occurred. *SAN1B* expression appeared to show a positive correlation with nitrate-induced senescence and by 96 hours had increased to approximately four times the original mRNA level.

SAN1A and *SAN1B* responded differently to nitrate application in terms of their expression. Nitrate appeared to have a negative effect on *SAN1A* message levels and a

Figure 27: *SAN1A* and *SAN1B* Expression in Nodules Undergoing Nitrate-Induced Senescence

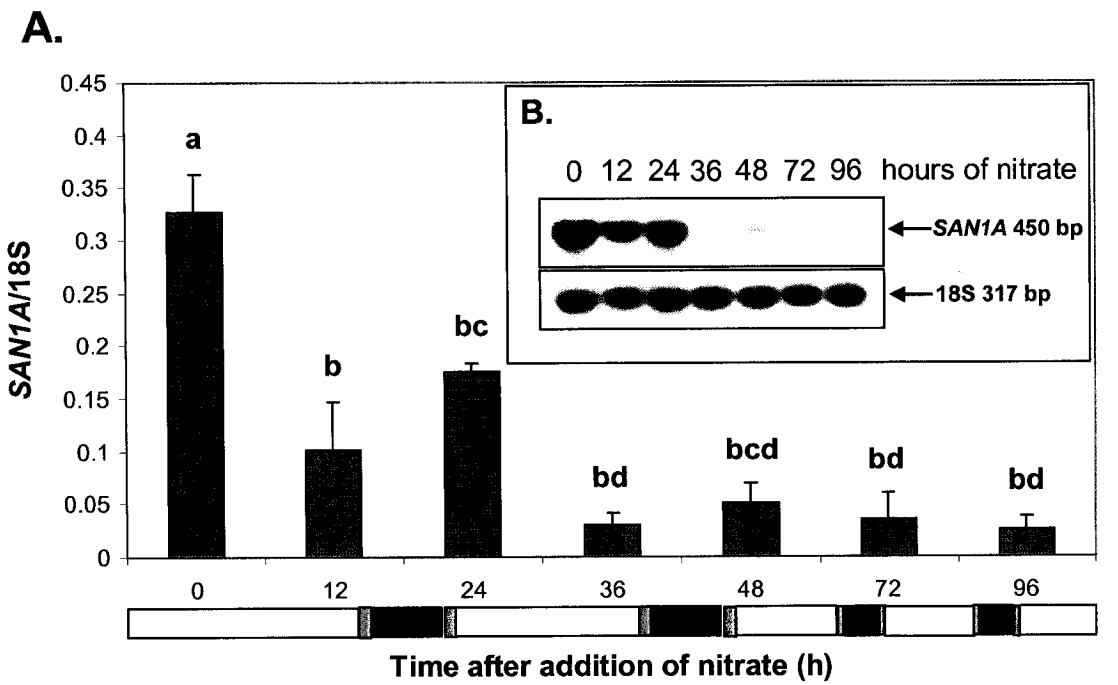
Relative RT-PCR was used to monitor *SAN1A* and 18S and *SAN1B* and 18S expression in RNA extracted from root nodules harvested at 0, 12, 24, 36, 48, 72, and 96 hours after initiation of treatment with 20 mM NH_4NO_3. Zero hours corresponded to 8:00 am, 21 dpi with *Bradyrhizobium japonicum*. Results are presented as the mean ± SE of the *SAN1A*/18S or *SAN1B*/18S signal versus hours of nitrate treatment. Each time point represents the mean *SAN1A*/18S signals of nodules collected from three different sets of plants. For each set of plants, nodules were collected from 4-5 plants for each time point. Means which are statistically different from one another are marked with different letters (Tukey's test, $p<0.05$).

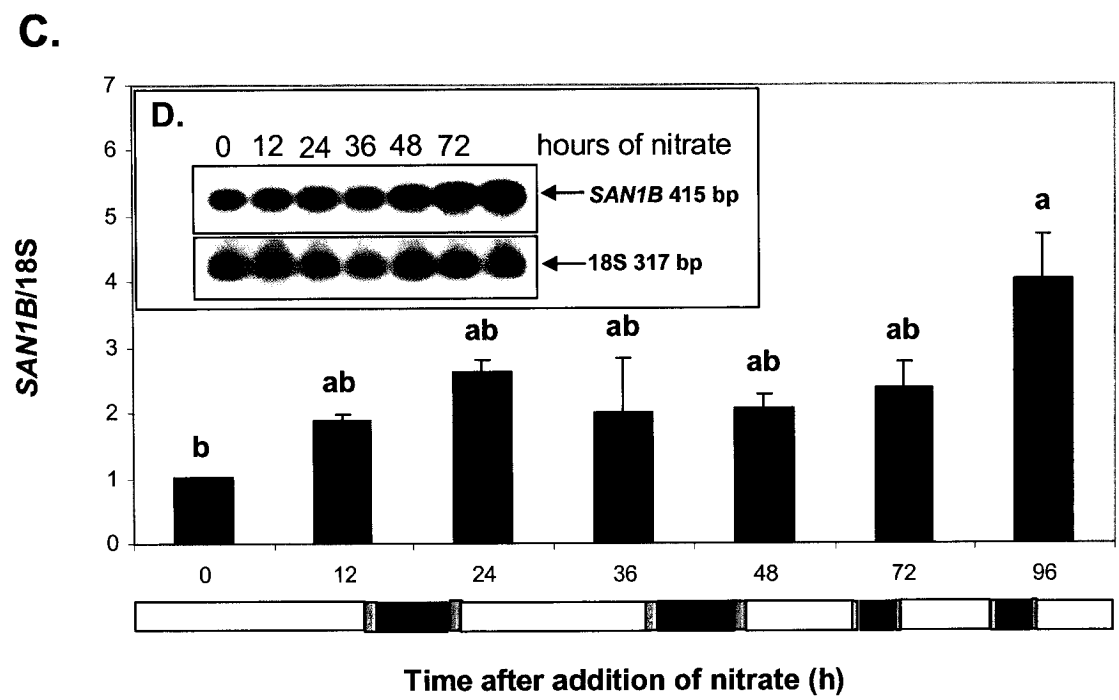
A. *SAN1A*/18S versus hours of nitrate treatment. The bar beneath the hours on the graph depicts the cycle of light and dark applied to the soybean plants. White sections represent full light, gray sections represent half light (dawn or dusk), and black sections represent darkness.

B. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 25 (B) for details. The pictures displayed here represent one hour exposures.

C. *SAN1B*/18S versus hours of nitrate treatment.

D. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 26 (B) for details. The pictures displayed here represent one hour exposures.


A.
B.
0 12 24 36 48 72 96 hours of nitrate
SAN1A 450 bp
18S 317 bp
SAN1A/18S
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
a
b
bc
bd
bcd
bd
bd
0
12
24
36
48
72
96
Time after addition of nitrate (h)


C.
D.
0 12 24 36 48 72 hours of nitrate
SAN1B 415 bp
18S 317 bp
SAN1B/18S
7
6
5
4
3
2
1
0
b
ab
ab
ab
ab
a
0
12
24
36
48
72
96
Time after addition of nitrate (h)

positive one on *SAN1B* message levels.

3.3.1.5.3 Light Regulation of *SAN1A*?

The unusual expression pattern of *SAN1A* during nitrate treatment of nodules suggested that it might be regulated by light. The message levels exhibited declines followed by recoveries over 12 hour intervals until 48 hours after nitrate application was initiated (Figure 27 A and B). Since these 12 hour intervals represented different points in the light-dark cycle, it was hypothesized that these declines and recoveries might represent a circadian rhythm dampened by nitrate treatment. To test this hypothesis, *SAN1A* expression was monitored more frequently, with samples taken every 4 hours over a 44 hour period, in the absence of nitrate (Figure 28 A,B).

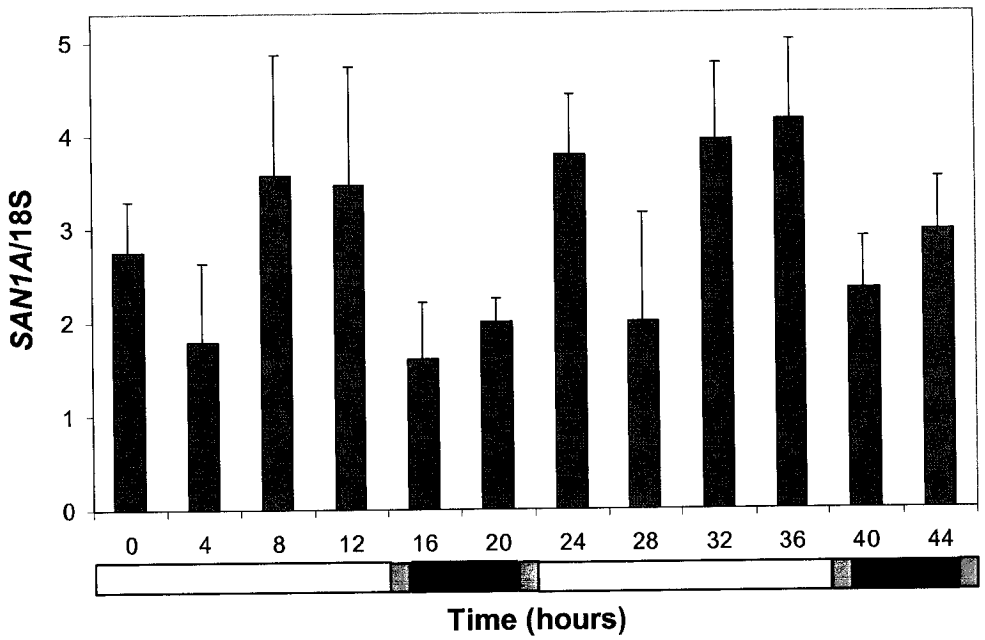
No statistically significant changes in *SAN1A* expression were observed between 0 and 44 hours (one-way ANOVA, p=0.367), suggesting that *SAN1A* may not exhibit strong diurnal variation. However, the trend of the graph (Figure 28A) suggests that *SAN1A* expression may be affected by light and dark to a limited extent, since it seemed to be highest in the latter portion of the light period and lowest during the dark portion of the light-dark cycle. Intermediate levels appeared to occur during the first part of the light period. Although the time points examined in the nitrate-treated nodules all fell within the light portion of the light-dark cycle, they fell on different points within it. For example, 0 and 24 hours fell near the beginning of the light period, whereas 12 and 36 hours occurred near the end of it. This result suggested that the unusual pattern of *SAN1A* expression in nitrate-treated nodules might be due to the combined effects of diurnal variation and nitrate. This possibility is explored further in the Discussion, Section 4.2.6. *SAN1A* expression may be affected by light and dark cycles, but further experiments would be required to verify this result.

Figure 28: *SAN1A* Expression in Nodules throughout Light-Dark Cycling

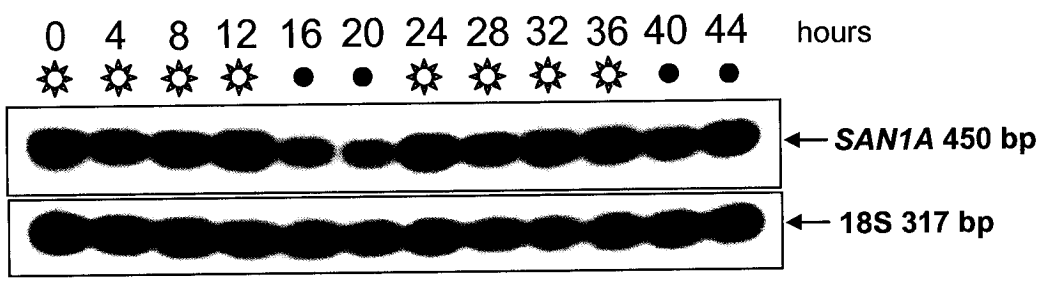
A. Relative RT-PCR was used to monitor *SAN1A* and 18S expression in RNA extracted from nodules harvested at 0, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, and 44 hours with *Bradyrhizobium japonicum*. Zero hours corresponded to 8:00 am, 21 dpi with *B. japonicum*. The bar beneath the hours on the graph depicts the cycle of light and dark applied to the soybean plants. White sections represent full light, gray sections represent half light (dawn or dusk), and black sections represent darkness. Results are presented as the mean ± SE of the *SAN1A*/18S signal versus hours. Each time point represents the mean *SAN1A*/18S signals of nodules collected from three different sets of plants.

B. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 25 (B) for details. The *SAN1A* and the 18S pictures represent 10 minute exposures. ☼, light; ●, darkness.

A.


SAN1A/18S
5
4
3
2
1
0
0 4 8 12 16 20 24 28 32 36 40 44
Time (hours)

B.


0 4 8 12 16 20 24 28 32 36 40 44 hours
SAN1A 450 bp
18S 317 bp

3.3.1.5.4 *SAN1A* and *SAN1B* Expression in Nodules of Dark-Treated Plants

SAN1A and *SAN1B* mRNA levels were monitored in nodules of soybean plants after the transfer of the plants to the dark (Figure 29 A,B,C,D). Zero hours corresponded to 8:00 am for a 21 dpi plant. The expression of both *SAN1A* and *SAN1B* showed significant changes between 0 and 96 hours of darkness. As with nitrate treatment, changes in *SAN1A* and *SAN1B* expression differed from each other.

SAN1A expression was significantly reduced between 0 and 12 hours of darkness (Figure 29A), suggesting that dark-induced senescence of nodules repressed the expression of *SAN1A* (Tukey's test, $p<0.05$). The expression levels remained low and statistically the same between 12 and 96 hours of dark treatment. However, the trend on the graph suggests that a further slight decline of *SAN1A* expression occurred between 12 and 36 hours, followed by a slight recovery in its expression between 36 and 96 hours after dark treatment (Figure 29B). Overall, dark induction of nodule senescence caused a swift reduction of *SAN1A* expression to between 1 and 15% of its original level in untreated 21 dpi nodules.

As with nitrate treatment, *SAN1B* behaved very differently from *SAN1A* in nodules of dark-treated soybean plants (Figure 29C). Between 0 and 36 hours, there was no change in *SAN1B* expression; however, between 36 and 96 hours, there was an increase in *SAN1B* expression, resulting in a significant increase in mRNA level by 72 hours followed by a continual rise up to 96 hours (Tukey's test, $p<0.05$). *SAN1B* expression appeared to show a positive correlation with dark-induced senescence in nodules such that the final level of *SAN1B* mRNA was approximately five times its original amount prior to treatment.

The expression of *SAN1A* and *SAN1B* in nodules following dark treatment and

Figure 29: *SAN1A* and *SAN1B* Expression in Nodules Undergoing Dark-Induced Senescence

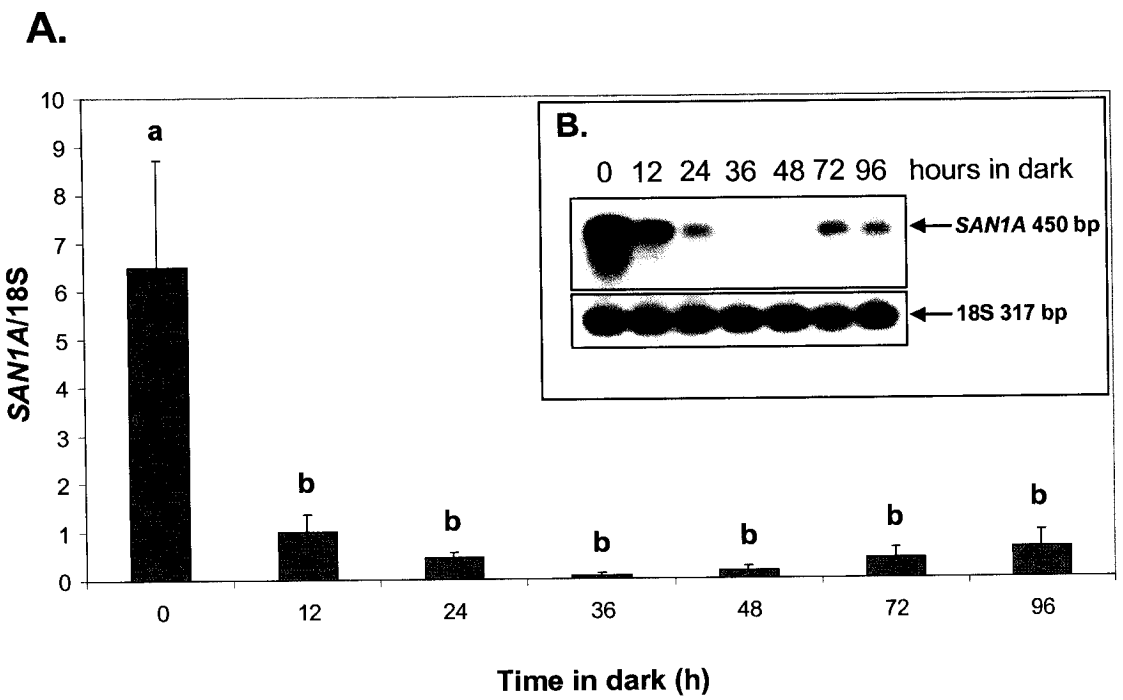
Relative RT-PCR was used to monitor *SAN1A* and 18S and *SAN1B* and 18S expression in RNA extracted from root nodules harvested at 0, 12, 24, 36, 48, 72, and 96 hours after placement of plants in the dark. Zero hours corresponded to 8:00 am, 21 dpi with *Bradyrhizobium japonicum*. Results are presented as the mean ± SE of the *SAN1A*/18S or *SAN1B*/18S signal versus hours of darkness. Each time point represents the mean *SAN1A*/18S or *SAN1B*/18S signals of nodules collected from three different sets of plants. For each set of plants, nodules were collected from 4-5 plants for each time point. Means which are statistically different from one another are marked with different letters (Tukey's test, $p<0.05$).

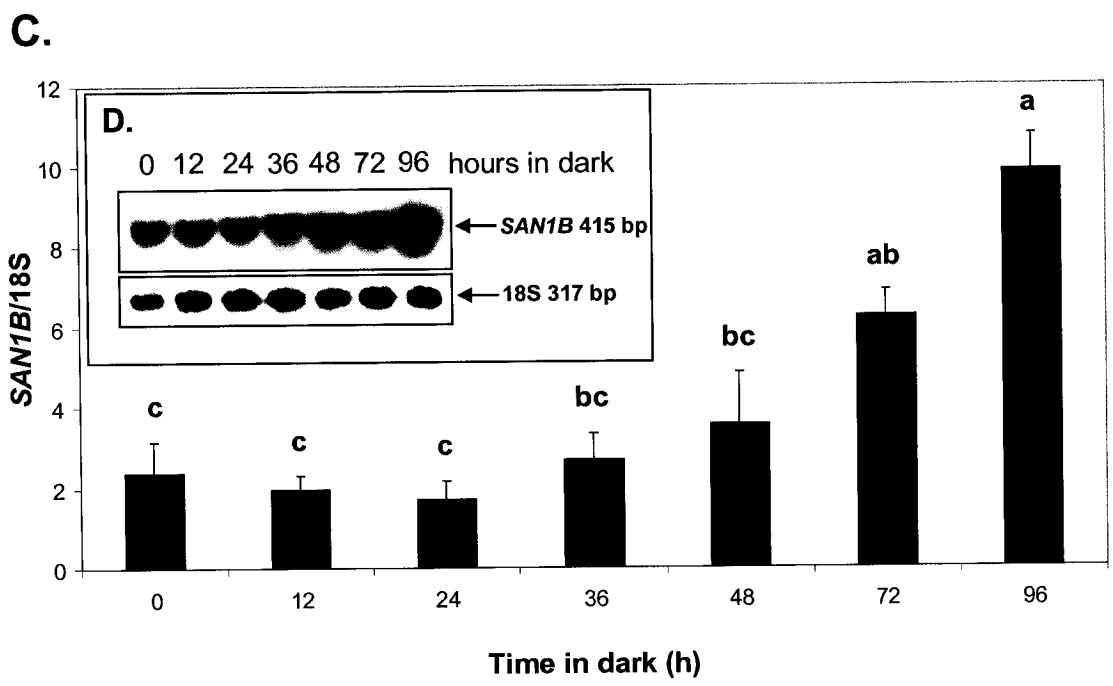
A. *SAN1A*/18S versus hours of darkness.

B. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 25 (B) for details. The *SAN1A* and 18S pictures represent overnight exposures.

C. *SAN1B*/18S versus hours of darkness.

D. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 26 (B) for details. The *SAN1B* and 18S pictures represent one hour and ten minute exposures, respectively.


A.
B.
0 12 24 36 48 72 96 hours in dark
SAN1A 450 bp
18S 317 bp
SAN1A/18S
Time in dark (h)
a
b
b
b
b
b
b
0
12
24
36
48
72
96


C.
D.
0 12 24 36 48 72 96 hours in dark
SAN1B 415 bp
18S 317 bp
SAN1B/18S
Time in dark (h)
c
c
c
bc
bc
ab
a
0
12
24
36
48
72
96

nitrate treatment was not identical. Overall, the addition of nitrate or dark treatment resulted in a decline in *SAN1A* mRNA levels in nodules and an increase in *SAN1B* mRNA levels.

3.3.1.5.5 *SAN1A* and *SAN1B* Expression in Various Soybean Tissues

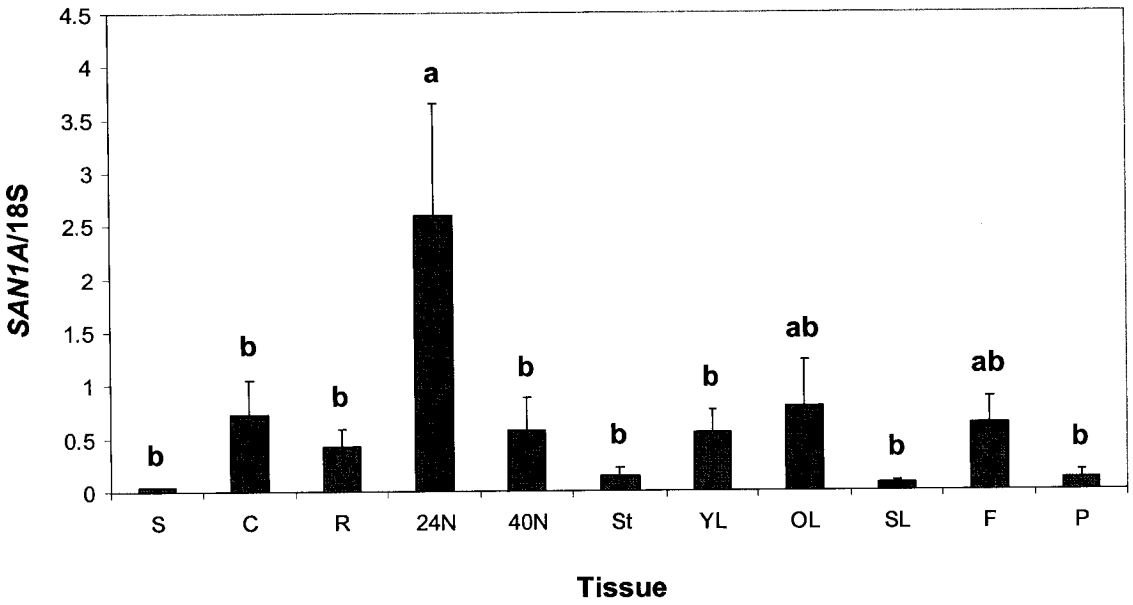
SAN1A and *SAN1B* expression was monitored in several different tissues of soybean (Figures 30 A,B and 31 A,B). While both *SAN1A* and *SAN1B* were expressed in all the soybean tissues examined, their expression levels among the tissues varied. Tukey's test ($p<0.05$) revealed that *SAN1A* expression was low and not significantly different in seedlings, cotyledons, roots, 40 dpi nodules, stems, young leaves, old leaves, senescing leaves, flowers, and pods (Figure 30A). Its expression levels were highest in 24 dpi nodules. This highest level of expression was not significantly different from that in old leaves and flowers, but was significantly different from that in all of the other tissues examined. The trend observed on the graph (Figure 30A) shows that the lowest levels of *SAN1A* expression were found in seedlings, senescing leaves, and pods. Intermediate levels of expression were found in cotyledons, roots, 40 dpi nodules, young leaves, old leaves, and flowers. Interestingly, in both nodules and leaves, *SAN1A* expression appeared to be lower in senescent tissues. Although the expression level of *SAN1A* was not significantly lower in senescing leaves compared with young and old leaves, the trends apparent on the graph (Figure 30A) and the Southern blot (Figure 30B) suggest that it was lower in senescing leaves. When taken with the previous results obtained for *SAN1A* expression in naturally senescing nodules and those induced to senesce by nitrate or dark treatment, this suggests that *SAN1A* expression may be repressed in senescing tissues in general. In addition, the results of the tissue distribution of *SAN1A* expression imply that *SAN1A* should be termed a nodule-enhanced gene.

Figure 30: *SAN1A* Expression in Various Tissues of Soybean

A. Relative RT-PCR was used to monitor *SAN1A* and 18S expression in RNA extracted from 6 day old seedlings (S), cotyledons (C) from 10 dpi plants, inoculated roots (R) from 24 dpi plants, nodules from 24 (24N) and 40 (40N) dpi soybean, stems (St) from 24 dpi plants, young leaves (YL) from 24 dpi plants, old leaves (OL) from 40 dpi plants, senescing leaves (SL) from 65 dpi plants, flowers (F) from 36 dpi plants, and pods (P) from 65 dpi plants. Results are presented as the mean ± SE of the *SAN1A*/18S signal versus soybean tissue. Each time point represents the mean *SAN1A*/18S signals of tissues collected from three different sets of plants. For each set of plants, tissues were collected from 4-5 plants for each time point. Means which are statistically different are marked with different letters (Tukey's test, $p<0.05$).

B. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 25 (B) for details. The *SAN1A* and 18S pictures represent overnight exposures.

A.


SAN1A/18S
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
a
b
b
b
b
b
b
ab
b
ab
b
S
C
R
24N
40N
St
YL
OL
SL
F
P
Tissue

B.

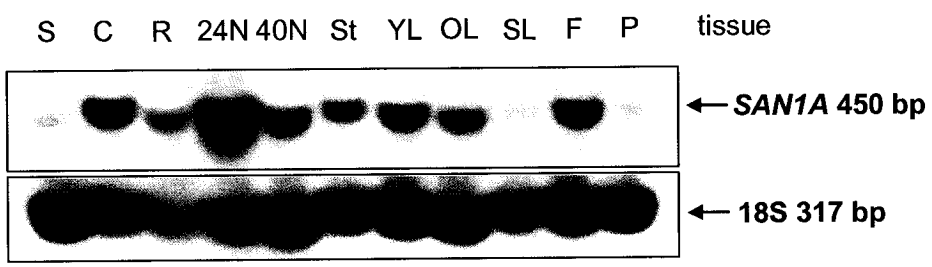
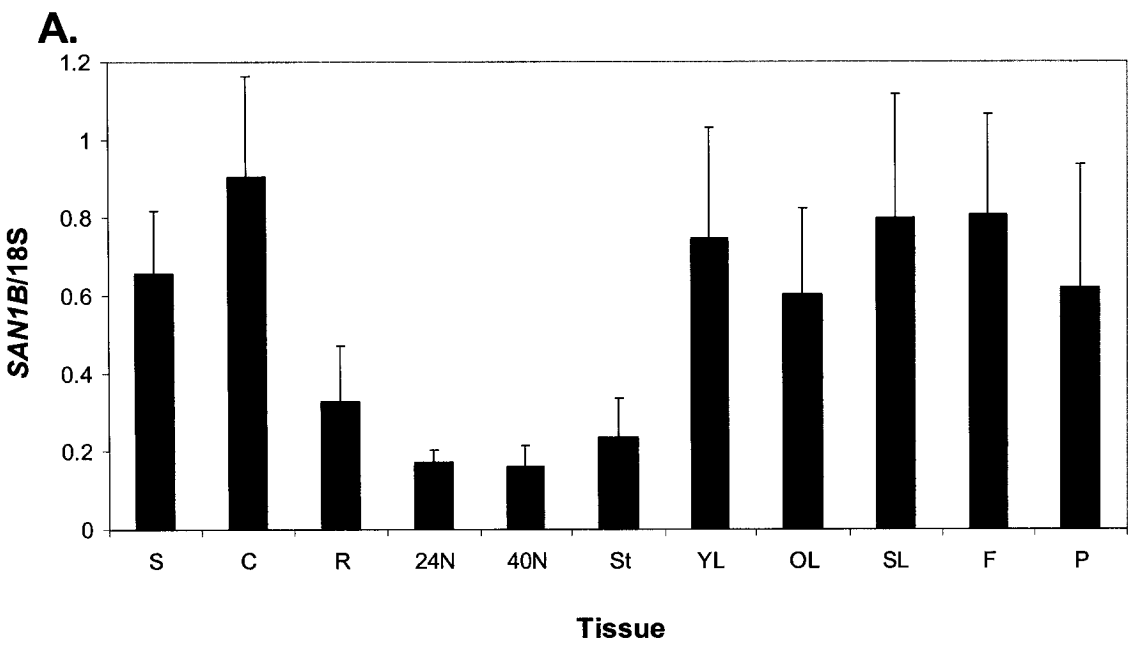
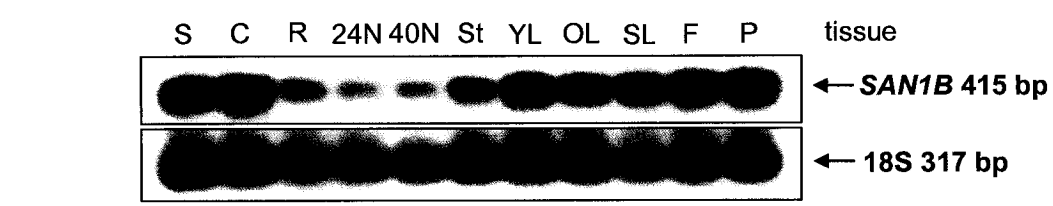

S C R 24N 40N St YL OL SL F P tissue
SAN1A 450 bp
18S 317 bp

Figure 31: *SAN1B* Expression in Various Tissues of Soybean

A. Relative RT-PCR was used to monitor *SAN1B* and 18S expression in RNA extracted from various soybean tissues. Refer to the caption of Figure 30 (A) for details.

B. A Southern blot of the relative RT-PCR products. Refer to the caption in Figure 26 (B) for details. The *SAN1B* and 18S pictures represent two day and overnight exposures, respectively.


A.
1.2
1
0.8
0.6
0.4
0.2
0
SAN1B/18S
S
C
R
24N
40N
St
YL
OL
SL
F
P
Tissue


B.
S C R 24N 40N St YL OL SL F P
tissue
SAN1B 415 bp
18S 317 bp

SAN1B expression did not vary significantly between the soybean tissues examined (one-way ANOVA, p= 0.180) (Figure 31A). However, the trend of the graph and the Southern blot (Figure 31B) suggest that the levels of *SAN1B* message were lowest in roots, nodules, and stems, and higher and comparable in seedlings, cotyledons, young, old, and senescing leaves, flowers, and pods. Unlike *SAN1A*, *SAN1B* cannot be said to be nodule-enhanced. The tissue distribution did not suggest that *SAN1B* expression is enhanced in naturally senescing tissues as it was in nodules of plants treated with nitrate or darkness since its message did not appear to be increased in either of senescing nodules or leaves.

SAN1A and *SAN1B* were both expressed in all tissues examined, but *SAN1A* appeared to be nodule-enhanced and may decline in both senescing nodules and leaves, whereas *SAN1B* did not appear to be enhanced in nodules or in naturally senescing tissues.

3.3.2 Tissue Distribution of *SAN1C*

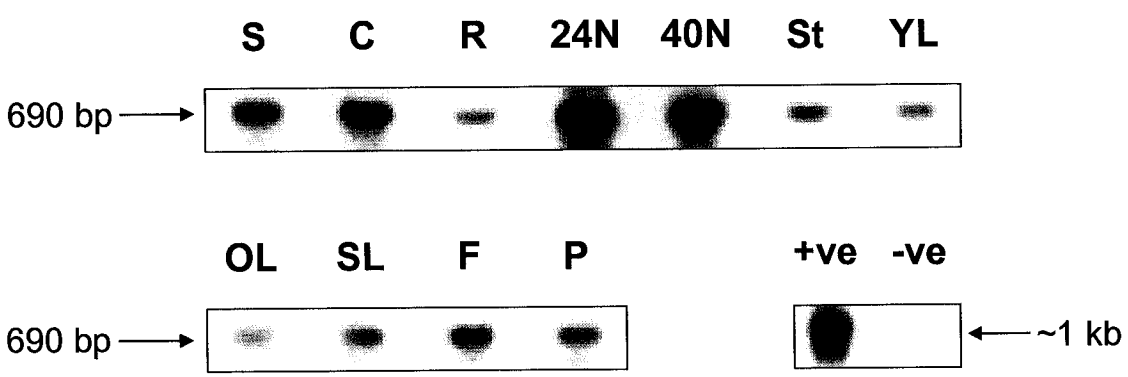
The *SAN1C* gene encodes a truncated protein (see Section 3.2.5); therefore, semi-quantitative relative RT-PCR was not carried out for this gene as for *SAN1A* and *SAN1B*. However, standard, non-quantitative RT-PCR experiments with *SAN1C*-specific primers (Table III) were carried out to determine whether its message was detectable in a variety of soybean tissues. The RT-PCR product obtained by use of the *SAN1C* primers was cloned and sequenced to ensure that the correct sequence was being amplified. The PCR product clone was 100% identical to the target region of *SAN1C*, confirming that the correct sequence had been obtained (data not shown).

The results of the *SAN1C* tissue distribution are displayed in Figure 32. The experiment was performed three times with similar results. It was clear that *SAN1C* is

Figure 32: Tissue Distribution of *SAN1C* mRNA

A Southern blot of *SAN1C* RT-PCR products in various tissues of soybean. The experiment was carried out as described in Chapter 2, Section 14.2.1 and repeated three times with the same result. The result of a 1 hour exposure is depicted.

24N, 24 dpi nodules; 40N, 40 dpi nodules; C, cotyledons; F, flowers; OL, old leaves; P, pods; R, roots; S, seedlings; SL, senescing leaves; St, stems; YL, young leaves.


S
C
R
24N
40N
St
YL
690 bp
OL
SL
F
P
+ve
-ve
690 bp
~1 kb

still expressed, and that its message could be found in all of the soybean tissues that were investigated. Although the results were not quantitative, it appeared that the highest levels of its message were likely to be found in nodules, followed by cotyledons and seedlings. Since *SAN1C* encodes a truncated protein, its expression patterns were not investigated further; however, the tissue distribution of the expression of this probable pseudogene is interesting (see Section 4.2.9).

3.3.3 Summary of *SAN1* Gene Family Expression Results

The tissue distributions of all three *SAN1* genes examined showed that all of the genes are expressed in each of the soybean tissues examined. *SAN1A* and *SAN1C* expression appeared to be nodule-enhanced, whereas *SAN1B* expression was not.

Monitoring of *SAN1A* and *SAN1B* expression throughout the natural development of soybean nodules showed that *SAN1A* message levels declined at the onset of senescence and then recovered, whereas *SAN1B* message levels remained low until late in senescence. In addition, the tissue distribution results suggested that *SAN1A* may also decline in naturally senescing tissues besides nodules, since its expression appeared to be lower in senescing leaves than young or old leaves. *SAN1B* expression was not significantly different between non-senescent and senescent leaves and nodules.

SAN1A and *SAN1B* expression behaved very differently in nodules induced to senesce by treatment of soybean with nitrate or dark. *SAN1A* mRNA showed a rapid decline in nodules during both treatments, whereas *SAN1B* mRNA showed a gradual increase during these treatments. *SAN1A* is likely to be repressed during natural and induced senescence, whereas *SAN1B* expression is likely to be promoted during induced senescence, but remain unaffected during natural senescence, at least until late in the process.

CHAPTER 3

RESULTS: PART IV

Bioinformatics and Evolution of the *SAN1* Gene Family

The functions of the *SAN1* genes are unknown; however, in an attempt to infer possible functions, bioinformatics tools were employed including BLASTP searches of the *nr* database (GenBank) and a protein sequence motif searches. TBLASTN searches of selected plant genomic and EST databases with the SAN1 protein sequences were also performed to look for the presence and expression of homologues in other plant species. Consensus sequences of the upstream regions, introns, and UTRs of *SAN1A* and *SAN1B* genes were scanned with the plant-specific regulatory sequence identification programs, PLACE and PlantCARE. The evolutionary relationships between the *SAN1* genes and their plant gene relatives were also investigated. Finally, a search for gene conversion events among the three *SAN1* genes was performed.

3.4.1 BLASTP Searches Reveal Similarities between SAN1s and Dioxygenases

BLASTP searches were carried out using the default settings, except that the low complexity filter was turned off. The results with the predicted protein sequences of *SAN1A*, *SAN1B*, and *SAN1C* are summarized in Table VI. The top matches for each SAN1 sequence were accessions of putative plant 2-ODDs; however, their ranking differed, reflecting differences in the SAN1 protein sequences as described in Section 3.2.5. Due to redundancy in the information from the *nr* database, the matches to *Arabidopsis* genes represented just two loci, At3g19000 and At3g19010. The truncated protein encoded by *SAN1C* gave the same BLASTP results as those of SAN1A and

Table VI: Results of BLASTP Search with the Predicted Protein Sequences of the *SAN1* Genes

The top eight results of a BLASTP search on June 20, 2006 using each of the predicted amino acid sequences of the *SAN1* genes from the soybean cultivar Resnik are presented. The matches are displayed with their accession numbers, species of origin, and probable function. The match number precedes the accession number in parentheses. Matches are in order from the top one to the eighth, but hits representing the same sequence are grouped together. Levels of identity and/or similarity with a given SAN1 sequence and the E-values of the matches are also presented. For the top match, the levels of identity and similarity are both presented, whereas only the levels of identity are shown for the other seven matches. +ves, number of similar amino acids; 2OG-Fe(II) oxygenase, 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily; IDs, number of identical amino acids.

Input Sequence	Accession Number(s)	Species	Description	Level of Similarity	E-value
SAN1A 1-352 aa	(1)AAN15625,BAB01696, AAM20659,NP_566623 (5) NP_974337	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein (At3g19000)	IDs: 207/356=58% +ves: 258/356=72% IDs: 159/290=54%	1e-116 5e-82
	(2)AAT77035	*Oryza sativa*	putative oxidoreductase	IDs: 192/347=55%	4e-111
	(3)BAB01697,AAO22576, NP_566624 (4)AAM65669 (6)NP_850613	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein (At3g19010)	IDs: 188/353=52% IDs: 187/353=52% IDs: 141/270=52%	6e-109 1e-108 5e-77
	(7)XP_480999,BAD05850	*Oryza sativa*	putative IPNS and related dioxygenases	IDs: 143/347=41%	6e-70
	(8)CAC83305	*Pinus pinaster*	putative oxylase protein	IDs: 118/211=55%	4e-58
SAN1B 1-353 aa	(1)AAN15625 ,BAB01696, AAM20659,NP_566623 (5)NP_974337	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein (At3g19000)	IDs: 197/358=55% +ves: 252/358=70% IDs: 149/276=53%	2e-111 1e-79
	(2)BAB01697,AAO22576, NP_566624 (3)AAM65669 (6)NP_850613	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein (At3g19010)	IDs: 181/354=51% IDs: 180/354=50% IDs: 135/270=50%	2e-106 5e-106 3e-75
	(4)AAT77035	*Oryza sativa*	putative oxidoreductase	IDs: 178/347=51%	7e-103

Input Sequence	Accession Number(s)	Species	Description	Level of Similarity	E-value
	(7)XP_480999,BAD05850	*Oryza sativa*	putative IPNS and related dioxygenases	IDs: 133/347=38%	5e-64
	(8)CAC83305	*Pinus pinaster*	putative oxylase protein	IDs: 114/211=54%	1e-56
SAN1C 1-126 aa	(1)NP_850613	*Arabidopsis thaliana*	oxidoreductase, 2OG-Fe(II) oxygenase family protein (At3g19010)	IDs: 60/128=46% +ves: 83/128=64%	3e-27
	(2)AAM65669			IDs: 60/128=46%	3e-27
	(3)BAB01697,AAO22576, NP_566624			IDs: 60/128=46%	3e-27
	(4)AAN15625 ,BAB01696, AAM20659,NP_566623	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein (At3g19000)	IDs: 64/129=49%	7e-27
	(5)NP_974337			IDs: 64/129=49%	7e-27
	(6)AAT77035	*Oryza sativa*	putative oxidoreductase	IDs: 50/123=50%	1e-21
	(7)XP_480999,BAD05850	*Oryza sativa*	putative IPNS and related related dioxygenases	IDs: 46/121=38%	3e-18
	(8)CAC83305	*Pinus pinaster*	putative oxylase protein	IDs: 46/121=38%	5e-14

SAN1B. If it were functional, one might have expected to find other shortened dioxygenase sequences in the *nr* database at NCBI. Since none were found, *SAN1C* is thought to be a pseudogene.

Although diverse, plant 2-ODDs share common motifs (Prescott and John, 1996) as discussed below. Despite these shared motifs, a given plant dioxygenase may not exhibit high overall amino acid sequence identity between species, e.g., F3Hs, LDOXs, and GA20oxs are ~74%, 47% and 53% identical, respectively, between species. This lack of sequence similarity can make it difficult to classify unknown 2-ODDs by sequence analysis alone.

As indicated in Table VI, the best matches were not to annotated accessions of diverse plant dioxygenases (GA20oxs, LDOXs, F3Hs, FLSs, ACCoxs, GA3oxs, GA2oxs, and H6Hs), but to unknown proteins containing the conserved amino acid residues found in dioxygenases from evolutionarily divergent species such as *Arabidopsis* (dicot), *Oryza sativa* (rice, a monocot), and *Pinus pinaster* (maritime pine, a gymnosperm). The best matches were 58%, 55%, and 46% identical to the SAN1A, SAN1B, and SAN1C predicted proteins, respectively. This result suggests that the *SAN1* genes either encode novel dioxygenases related to the unknown accessions from *Arabidopsis*, rice, and maritime pine, or are divergent types (paralogues) of previously characterized dioxygenases. Plant 2-ODDs range in size from 310 to 400 aa with ACCoxs at the shorter end of this range (312 to 325 aa) and the others described above having 325 to 400 aa. The predicted sizes of SAN1A and SAN1B, 352 and 353 aa, respectively, are within the normal rage for plant 2-ODDs that are not ACCoxs.

The positions of the introns of plant 2-ODD genes are highly conserved (Prescott and John, 1996). Either two or three introns distributed among four positions may be

found, but the intron at position four is conserved between all plant dioxygenases described to date. SAN1A and SAN1B have two introns each, at conserved positions two and four. This arrangement is typically found in gibberellin 20-oxidase genes (Xu *et al.*, 1995).

Plant dioxygenases are most closely related to the IPNSs of microorganisms (De Carolis and De Luca, 1994), with both being members of a larger superfamily of 2-oxoglutarate-Fe (II) oxygenases. Thus if *SAN1* genes encode dioxygenases, the SAN1 conceptual proteins should contain the same conserved amino acid residues as other dioxygenases. Residues that SAN1 proteins share with other plant dioxygenases, IPNSs, and with enzymes belonging to the larger 2-ODD superfamily are shown in the alignment in Figure 33. The alignment highlights the degree of sequence similarity of the SAN1 predicted proteins with IPNS-related proteins and 2-ODDs. Four of the five domains (I, III, IV, and V) that are highly conserved between plant dioxygenases and the IPNSs (Britsch *et al.*, 1993) are found in the SAN1A and SAN1B proteins (truncated SAN1C only has domain I). Functions have not been assigned to Domains I and III, but Domain IV contains a histidine and an aspartic acid residue involved in binding of Fe(II) and Domain V contains a histidine involved in Fe(II)-binding and the arginine residue responsible for binding of 2-oxoglutarate within the highly conserved R-X-S motif of plant dioxygenases (Wellmann *et al.*, 2002). A leucine-rich domain common to other plant dioxygenases, in addition to the four canonical IPNS-related dioxygenase domains, has also been identified in the SAN1A and SAN1B predicted proteins. In addition, four other amino acids, two glycines, one proline, and a histidine are conserved among all plant dioxygenases, although functions have not yet been assigned to these amino acids (Wellmann *et al.*, 2002). SAN1A and SAN1B contain all four of these conserved amino

Figure 33: Amino Acid Sequence Conservation between SAN1 Predicted Proteins and Other Dioxygenases

Amino acid residues conserved between the SAN1 predicted proteins and IPNS-related dioxygenases as determined by the NCBI Conserved Domain Search are displayed (Marchler-Bauer and Bryant, 2004). Amino acids which are identical between dioxygenases and the SAN1 sequences are shown in pink and similar amino acids are green. Underlined sequences correspond to domains conserved among plant dioxygenases by Britsch *et al.* (1993). Domain II, represented by W/I/V-P-D, is missing from the SAN1 protein sequences. The double underlined region represents a leucine-rich domain. The squiggly underline highlights the conserved R-X-S motif (Lukacin and Britsch, 1997). Asterisks highlight amino acids of known or probable functional importance (Wellmann *et al.*, 2002); blue asterisks are those of unknown function conserved among all dioxygenases; red asterisks are the amino acids involved in binding Fe(II); the green asterisk is the conserved arginine believed to bind 2-oxoglutarate. The regions conserved between the SAN1 predicted proteins and the 2-oxoglutarate-Fe(II)-dependent oxygenase superfamily are highlighted in gray.

```
SAN1A      1   MGEVDPAFIQEPQHRPNLSTIQAEGIPIIDLSPITNHTVSDPSAIESLVK
SAN1B      1   MEKVDQAFIQYPQHRPNLTIIQPEHIPIIDLSPITNHTVSHPSPIEGLVK
SAN1C      1   MGEVDPAFIQYLQHRPKFSTIQPEDIPVIDISPITNHTASDPCSIGGLVK

                        *      *
SAN1A     51   EIGRACQEWGFFQVTNHGVPLTLRQNIEKASKLFFAQTLEEKRKVSRNES
SAN1B     51   EIGNACKEWGFFQVINHGVSLTLRQNIEKASKLFFAQSLEEKRKVSRDES
SAN1C     51   EIRNASKEWGFFQLSNHAVPLTLRQNIEKASRLFFAQNLEEKRKVSKDEN
                          I

SAN1A    101   SPMGYYDTEHTKNVRDWKEVFDFLAKDPTFIPLTSDEHDDRVNQWTNQSP
SAN1B    101   SPMGYYDTEHTKNIRDWKEVFDFLAKDPTFVPLTSDEHDNRLTQWTNPSP
SAN1C    101   NTTGYHDREHPKNIRDWKEVFDFLAK........................

SAN1A    151   QYPPLFRVVTQEYIQEMEKLSFKLLELIALSLGLEAKRFEEFFIKDQTSF
SAN1B    151   QYPPHFRDIIKEYVEEMEKLSFKLMELIALNLGLEAKRFEEFFIKDQTSF
SAN1C          ..................................................

                  *             * *
SAN1A    201   IRLNHYPPCPYPDLALGVGRHKDPGALTILAQDEVGGLEVRRKRDQEWIR
SAN1B    201   LRLNYYPPCPYPHLALGVGRHKDSGPLTILAQDEVGGLEVRPKADQDWIR
SAN1C          ..................................................
                   III                    IV
                     *               *         *
SAN1A    251   VKPTPNAYIINIGDTVQVWSNDAYESVDHRVVVNSEKERLSIPFFFFPAH
SAN1B    251   VKPIPNAYIINLGDMIQVWSNDAYESVEHRVVVNSEKARFSIPFFLFPAH
SAN1C          ..................................................
                                                    V

SAN1A    301   DTKVKPLEELINEQNPSKYRPYNWGKFLVHRGNSNFKKQNEENIQIHHYK
SAN1B    301   DTVVKPLEELINEQNPSKFRPYKWGKFLVHRLDSNFKKQNAENVQTYHYK
SAN1C          ..................................................

SAN1A    351   IA.
SAN1B    351   IRE
SAN1C          ...
```

acids, whereas SAN1C only contains one glycine and the histidine. The portion of the alignment in Figure 33 that is highlighted in gray shows the amino acids conserved between the SAN1 predicted proteins and enzymes of the larger 2-oxoglutarate-Fe (II) oxygenase superfamily. The conserved amino acids are confined to the carboxy-terminus as is the case between unrelated 2-oxoglutarate-Fe (II) oxygenases in general. These observations strengthen the inference that proteins encoded by the *SAN1* genes are dioxygenases.

3.4.2 Motifs Found in the SAN1 Predicted Proteins

PROSITE (http://us.expasy.org/prosite/) was used to scan the SAN1 predicted proteins for profiles and patterns shared with other proteins (Hulo *et al.*, 2004) in addition to the conserved domains discussed above (Table VII). Only commonly occurring profiles were detected including sites for N-glycosylation, casein kinase II phosphorylation, N-myristoylation, protein kinase C phosphorylation, cAMP- and cGMP-dependent protein kinase phosphorylation, and tyrosine kinase phosphorylation. The number of these sites varied between the three proteins with the truncated protein of SAN1C having the fewest (Table VII). Since the recognition sites within the motifs are short, they represent patterns with a high probability of occurrence (PROSITE, http://us.expasy.org/prosite/); therefore, it cannot be said with any certainty that the proteins are really modified as predicted. Direct experimental evidence is required to verify whether the SAN1 proteins are actually modified.

3.4.3 Regulatory Sequences of the *SAN1* Genes

Expression data for the *SAN1A* and *SAN1B* genes presented in Part III of Results suggested that they may share regulatory elements, as they are both expressed in all soybean tissues. However, their differential expression during dark and nitrate treatments

Table VII: Profiles Found in the SAN1 Predicted Protein Sequences

Protein profiles were identified in each of the SAN1 predicted proteins by scanning the PROSITE database on June 20, 2006 (http://us.expasy.org/prosite/) with each amino acid sequence. The only profiles identified, besides those shared with dioxygenases, were commonly found motifs. The table presents these motifs, their position in the SAN1 amino acid sequence, their descriptions, and their consensus sequences. Standard amino acid abbreviations are used in the consensus sequences and X represents any amino acid. { } indicates that the amino acid cannot be found at that position in a consensus sequence.

PROSITE Profile Description	Consensus Sequence of Profile	Positions and Sequences in: SAN1A	SAN1B	SAN1C
N-glycosylation site	N-X{P}-S/T-X{P}	17-20: NLST 36-39: NHTV 98-101: NESS	17-20: NLTI 36-39: NHTV	36-39: NHTA 54-57: NASK 100-103: NNTT 101-104: NTTG
Casein kinase II phosphorylation site; S/T is the phosphorylation site	S/T-X-X-D/E	38-41: TvsD 43-46: SaiE 88-91: TleE 96-99: SrnE 134-137: TsdE 285-288: SekE	43-46: SpiE 88-91: SleE 96-99: SrdE 134-137: TsdE	38-41: TasD 96-99: SkdE
N-myristoylation site	G-X{EDRKHPFYW}-X -X-S/T/A/G/C/N-X{P}	68-73: GVplTL	68-73: GVslTL	None identified.
Protein kinase C phosphorylation site; S/T is the phosphorylation site	S/T-X-R/K	72-74: TlR 171-173: SfK 285-287: SeK	72-74: TlR 171-173: SfK 285-287: SeK	72-74: TlR
cAMP- and cGMP-dependent protein kinase phosphorylation site; S/T is the phosphorylation site	R/K-R/K-X-S/T	93-96: RKvS	93-96: RKvS	93-96: RKvS
Tyrosine sulfation site	None determined.	267-281: qvwsndaYesvdhrv	267-281: qvwsndaYesvdhrv	None identified.

in root nodules, and during the stages of natural root nodule senescence suggests the presence of regulatory elements specific to each of *SAN1A* and *SAN1B*. These specific cis-acting sequences would be expected in addition to the common eukaryotic regulatory elements such as CAAT boxes, TATA boxes, and polyadenylation signals shared by most genes (Lehninger *et al.*, 1993). Conserved sequences are likely to be located in the typical promoter region of the genes, upstream of the transcriptional start site (Rombauts *et al.*, 2003), and are likely to be short. For these analyses, the 5′ upstream region was taken as ~1700 bp for each gene, the length of sequence between the predicted 3′ UTR of the preceding gene and the start codon of the gene of interest.

Two approaches were used. PipMaker (Schwartz *et al.*, 2000) displays the outcome of a BLASTZ alignment as "pips" or percent identity plots that demonstrate a level of identity ≥50% when gap-free segments of the alignment of two or more genomic regions are displayed. Typically, this program is used to compare homologous genomic regions between two species in order to find regulatory elements, e.g., enhancers in the intergenic regions of *Dlx* homeobox genes from human, mouse, zebrafish and pufferfish (Ghanem *et al.*, 2003). In this case the entire *SAN1A* and *SAN1B* gene regions were compared. PipMaker finds regions that are conserved between genomic sequences, but does not identify regulatory motifs. Therefore, DNAMAN was used to separately align regions with potential regulatory elements, including the 5′ upstream regions, introns, and 3′ UTRs, and the consensus sequences were scanned for potential regulatory elements in the PLACE (http://www.dna.affrc.go.jp/PLACE) (Higo *et al.*, 1999) and PlantCARE (http://intra.psb.ugent.be:8080/PlantCARE) (Lescot *et al.*, 2002) databases (Table VIII).

PipMaker identified high regions of sequence similarity between the exons of *SAN1A* and *SAN1B*, as expected, and short regions of conserved sequence within the

Table VIII: Regulatory Elements Shared between *SAN1A* and *SAN1B*

The 5′ upstream regions (~1720 bp), each of the two introns, and the 3′ UTRs of *SAN1A* and *SAN1B* were aligned using DNAMAN Version 4.22 (Lynnon BioSoft, Vaudreuil, QC) and the resulting consensus sequences scanned for potential regulatory elements using the PLACE and PlantCARE databases. ARR, *Arabidopsis thaliana* response regulators; Dof, DNA-binding with one finger; ERF3, ethylene-responsive factor 3; *legA*, legumin A; N, any nucleotide; PEPC, phosphoenylpyruvate carboxylase; RE, regulatory element; *rolD*, root locus D, a plant oncogene; SKN-1, a transcription factor originally isolated from *Caenorhabditis elegans* implicated in determining the fate of the ventral blastomere; WRKY, amino acids in conserved domain of this family of transcription factors; Y, cytosine or thymine.

Region Compared	Shared REs	Strand and Location within Region	Signal Sequence	Description
5′ upstream	CACTFTPPCA1	(+) 452	YACT	key component of mesophyll expression module 1 in a *Flaveria trinervia PEPC* gene
	ARR1AT	(+) 1196	NGATT	ARR-1 binding element; ARR is a cytokinin responsive regulator in *Arabidopsis*
	DOFCOREZM	(−) 1519 (−) 1657	AAAG	core site for binding of Dof proteins in maize
Intron 1	EBOXBNNAPA	(+) 242 (−) 242	CANNTG	E-box of *napA* seed storage protein gene of *Brassica napus*
	MYB2CONSEN-SUSAT	(+) 242	YAACKG	MYB recognition site found in promoters of dehydration-responsive and ABA signalling genes of *Arabidopsis*
	MYCCONSEN-SUSAT	(+) 242 (−) 242	CANNTG	MYC recognition site in genes involved in abiotic stress and ABA signalling of *Arabidopsis*
	MYBCORE	(−) 242	CNGTTR	binding site for MYB proteins; AtMYB2 is involved in regulation of water stress in *Arabidopsis*; petunia MYB is involved in regulation of flavonoid synthesis
	GTGANTG10	(+) 410	GTGA	motif found in promoter of tobacco late pollen gene *g10*

Region Compared	Shared REs	Strand and Location within Region	Signal Sequence	Description
Intron 2	ARR1AT	(+) 278	NGATT	ARR-1 binding element; ARR is a response regulator in *Arabidopsis*
	CAATBOX1	(−) 280	CAAT	CAAT consensus sequence from *legA* gene of pea
	CCAATBOX1	(−) 280	CCAAT	common sequence found in 5′ non-coding regions of eukaryotic genes
3′ UTR	TATA-Box	(+) 4	TATA	core promoter element usually found -30 upstream of transcription start site in eukaryotic genes
	ROOTMOTIF-TAPOX1	(+) 5	ATATT	motif found in *rolD* promoter; promotes root-specific gene expression
	SKN-1-like-motif	(−) 105	GTCAT	Skn-1-like motif required for endosperm-specific expression in rice; requires cooperation of AACA, GCN4, and ACGT motifs
	WBOXHVISO1	(+) 106	TGACT	W-box element found in barley *isoamylase1* promoter
	WBOXNTERF3	(+) 106	TGACY	W-box found in tobacco ERF3 (transcriptional repressor) gene; may be involved in wound-activation of ERF3
	WRKY71OS	(+) 106	TGAC	core of TGAC-containing W-box; binding site of rice WRKY71, a transcriptional repressor of gibberellin signalling

Region Compared	Shared REs	Strand and Location within Region	Signal Sequence	Description
	POLASIG1	(+) 168	AATAAA	plant polyadenylation signal

the second introns and the 3′ UTRs; however, no areas of conserved sequence were detected in the 5′ upstream regions or first introns (data not shown). These results suggest that few regulatory elements are conserved between the two *SAN1* genes - a result consistent with the different expression patterns of *SAN1A* and *SAN1B*.

Despite the low overall level of conservation within non-exonic regions of *SAN1A* and *SAN1B*, the scan of the consensus sequences did identify shared potential regulatory elements (Table VIII). For example, their 5′ upstream regions contain DOFCOREZM, the core site for DOF protein binding in maize. DOF proteins are plant-specific transcription factors (Yanagisawa and Schmidt, 1999) that are involved in regulating processes such as carbon fixation and metabolism in maize (Yanagisawa, 2000; Yanagisawa and Sheen, 1998), endosperm-specific expression in maize (Vicente-Carbajosa *et al.*, 1997), and stress response in *Arabidopsis* (Zhang *et al.*, 1995). They also share CACTFTPPCA1, which directs mesophyll-specific expression of C4PEPC in *Flaveria trinervia* (Gowik *et al.*, 2004), and ARR1AT, a cytokinin-responsive regulator binding element found in *Arabidopsis* and rice (Ross *et al.*, 2004; Sakai *et al.*, 2000). No CAAT or TATA boxes were found to be conserved between *SAN1A* and *SAN1B*; however, when the upstream region of each gene was scanned separately, several potential CAAT and TATA box sequences were identified. In addition, potential regulatory sequences are shared by *SAN1A* and *SAN1B* in their first introns, their second introns, and their 3′ UTRs (Table VIII).

While some of these identifications are likely coincidental, e.g., the TATA box in the 3′ UTR, this process has identified many potential sites of regulation. However, experiments would have to be carried out to determine whether any of these sites truly regulate the expression of these genes. *SAN1A* and *SAN1B* are also differentially

expressed and were therefore also individually scanned for regulatory motifs. The number of "hits" was too many to be presented, but regulatory elements involved in responsiveness to various signals, including light, hormones, stress, senescence, and sugar, and those involved in tissue-specific expression were identified. Each gene has many results compared to those shared by both *SAN1A* and *SAN1B*, as would be expected by their differing expression patterns. Divergence of function and regulation often occurs in duplicated genes (Graur and Li, 2000).

3.4.4 Evolutionary History of the *SAN1* Genes

Although the specific functions of the SAN1 proteins are not known, the proteins share amino acid signatures of proteins belonging to the 2-oxoglutarate-Fe(II) oxygenase superfamily, more specifically to the IPNS-related plant dioxygenases. A gene tree of plant dioxygenases was constructed to determine which dioxygenases were most closely related to the *SAN1* genes. In addition, tandemly repeated genes often undergo gene conversion; this process was also investigated in the *SAN1* genes (Graur and Li, 2000).

3.4.4.1 Gene Tree of Plant Dioxygenases

A tree containing the major families of plant dioxygenases, along with the SAN1A and SAN1B sequences from cv Resnik was constructed. Only protein sequences or predicted protein sequences were used as the DNA sequence identity among plant dioxygenases is generally low, in part due to degeneracy of the genetic code.

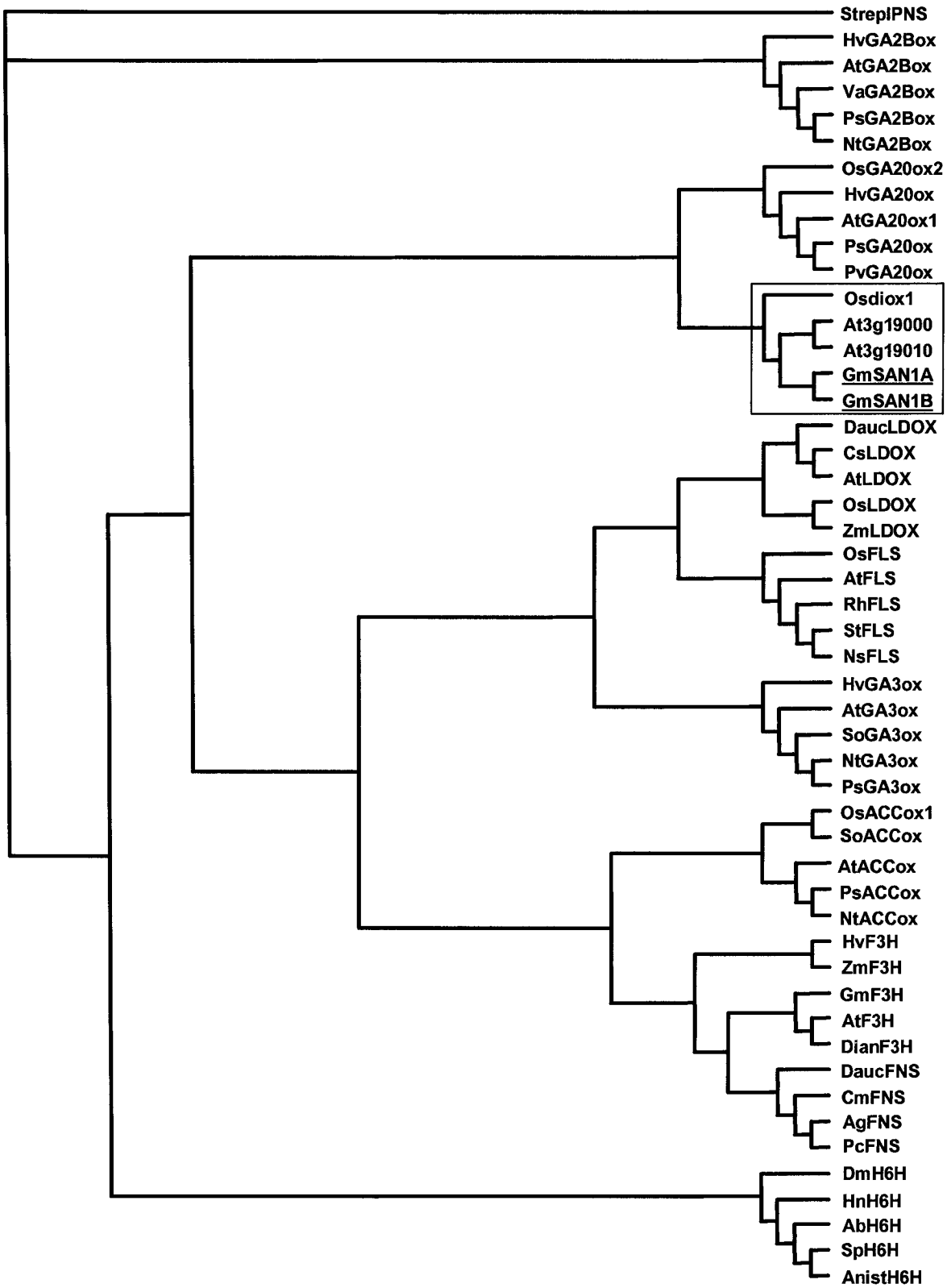
Dioxygenases were selected to represent each of the major families of these enzymes in plants. The families used in the construction of the tree were the SAN1 sequences and the related unknown dioxygenases identified in BLASTP searches (see Section 3.4.1), including GA20oxs, GA3oxs, GA2Boxs, F3Hs, FLSs, FNSs, LDOXs, H6Hs, and ACCoxs. Five members of each family were chosen randomly, except that at

least one monocot sequence and one legume sequence was selected for inclusion in each family, if the sequences were available. The outgroup was IPNS from *Streptomyces cattleya*. IPNSs are the related to the plant dioxygenases, but are only found in fungi and bacteria (De Carolis and De Luca, 1994). See Appendix 2, Table I for a list of sequences.

Following alignment of the sequences (Appendix 2, Figure 1), a neighbour-joining tree was constructed (Figure 34). Trees were also constructed using protein parsimony and maximum likelihood methods; however, these are not shown as they gave the same topology. The SAN1 sequences and related dioxygenases from *Arabidopsis* and rice identified by BLASTP (Section 3.4.1) form a separate group with the GA20oxs. SAN1A and SAN1B are paired together and are more closely related to the unknown dioxygenases from *Arabidopsis* (dicot) than the unknown dioxygenase from rice (monocot). The topology also suggests that the gene duplication events which gave rise to the *SAN1* genes and the *Arabidopsis* genes may have occurred after *Arabidopsis* and soybean split into species. If they had arisen prior to the split in the lineages, one might expect to see orthologues from each of the species paired together. In addition to placing SAN1 with the GA20oxs, tree analysis was used to explore the clustering of other plant dioxygenases characterized to date. LDOXs, FLSs, and GA3oxs tree together, whereas ACCoxs, F3Hs, and FNSs form a separate group. Gebhardt *et al.* (2005) recently published a neighbour-joining tree of the flavonoid biosynthetic genes that shows similar clustering. Unfortunately, trees with other plant dioxygenases are unavailable for comparison. Interestingly, enzymes that operate in the same pathway, such as those of the gibberellin synthetic and degradation pathways, are not more closely related to each other than to the other plant dioxygenases and are dispersed throughout the tree. This phenomenon is also apparent for the enzymes that play roles in flavonoid synthesis. The

Figure 34: Neighbour-Joining Tree of Plant Dioxygenases

A neighbour-joining tree based on protein sequences representative of the major families of plant dioxygenases is shown. The tree was constructed using the neighbour-joining method and bootstrapping was performed 10,000 times using programs contained in PHYLIP Version 3.63 (Felsenstein, 1989). The tree was drawn using TreeView Version 1.6.6 (Page, 1996). All nodes are supported by bootstrap values of 100%; therefore, none of the nodes are labelled. The outgroup is StrepIPNS (Accession Q53932). The SAN1 sequences are labelled as GmSAN1A and GmSAN1B and are presented in underlined red font.


StreptPNS
HvGA2Box
AtGA2Box
VaGA2Box
PsGA2Box
NtGA2Box
OsGA20ox2
HvGA20ox
AtGA20ox1
PsGA20ox
PvGA20ox
Osdiox1
At3g19000
At3g19010
GmSAN1A
GmSAN1B
DaucLDOX
CsLDOX
AtLDOX
OsLDOX
ZmLDOX
OsFLS
AtFLS
RhFLS
StFLS
NsFLS
HvGA3ox
AtGA3ox
SoGA3ox
NtGA3ox
PsGA3ox
OsACCox1
SoACCox
AtACCox
PsACCox
NtACCox
HvF3H
ZmF3H
GmF3H
AtF3H
DianF3H
DaucFNS
CmFNS
AgFNS
PcFNS
DmH6H
HnH6H
AbH6H
SpH6H
AnistH6H

enzymes belonging to a single pathway do not seem to have evolved directly from one another.

When monocot sequences were included, the monocots generally fell outside of the dicot sequences, if there was one sequence, or treed together if two were included. However, OsACCox (rice ACCox) is paired with SoACCox (spinach ACCox). This result may have occurred due to the fact that the ACCox genes occur as multigene families (Prescott and John, 1996). Since the sequences for each plant dioxygenase family were selected randomly without regard for orthologues or paralogues, it is possible that the rice and spinach that the rice and spinach ACCoxs represent one member of the ACCox family, whereas the other sequences are representative of another member of the family.

3.4.4.2 Testing for Gene Conversion between the *SAN1* Genes in Soybean

Tandemly repeated genes are likely candidates for gene conversion (Graur and Li, 2000). The genomic DNA sequences of *SAN1A*, *SAN1B*, and *SAN1C* (Appendix 1), with an additional ~1700 bp of upstream sequence, were employed in a test for gene conversion events using GENECONV Version 1.81 (Sawyer, 1989). The P-value produced by the gene conversion test was 0.4794 (non-significant), suggesting that either gene conversion has not occurred between the three *SAN1* genes or that the program is not able to detect it.

3.4.5 Summary of Bioinformatics and Evolution Results

Analysis of the BLASTP results suggest that *SAN1* gene products are similar to, and share key residues with, plant dioxygenases and isopenicillin N synthases which are part of the 2-oxoglutarate-Fe(II) oxygenase superfamily. However, the *SAN1* genes appear to encode novel dioxygenases or divergent proteins encoded by paralogues of previously

described dioxygenases. A neighbour-joining tree constructed from the protein sequences of plant dioxygenases revealed that they are most closely related to the gibberellin 20-oxidases.

Motifs identified in the SAN1 putative protein include sites for glycosylation, phosphorylation, myristoylation, and sulfation. However, large proteins commonly contain such sites and their function must be tested experimentally.

An examination of the 5′ upstream regions, introns, and 3′ UTRs for potential regulatory sequences revealed little conservation between the *SAN1A* and *SAN1B* genes, as expected by their different expression profiles. Diversification of function is a common fate of duplicated genes (Graur and Li, 2000).

Finally, testing for gene conversion between the *SAN1* genes revealed that it is unlikely that it has occurred in the evolutionary history of this gene family.

CHAPTER 4

DISCUSSION

4.1 Characterization of Soybean Nodule Senescence

4.1.1. Features of PCD Observed during Soybean Nodule Senescence

Light microscopic studies of senescing soybean nodules were undertaken in order to determine whether features resembling those of PCD, such as nuclear deformation, cleared zones which could be indicative of vacuolation, or "crinkling" of the cell walls which could be indicative of cellular condensation, could be observed during soybean root nodule senescence. In naturally senescing nodules, nuclear deformation (by 32 dpi) and crinkling of the cell walls (by 40 dpi), were observed, but cleared zones were not. In soybean cells undergoing senescence induced by nitrate, all three features were observed.

During natural senescence, the population of nodules is likely to be asynchronous with each nodule developing at its own pace; therefore, nodules may not exhibit features of PCD at predictable time points. For example, at 50 dpi, the crinkling of the cell walls and nuclear deformation could be observed in some fields, but not others. Nodules from plants of the same age may not show the same features of PCD, and nodules from different time points may show the same features of PCD. Thus, the features observed during this study, while suggestive of PCD, are not helpful for determining the timing of senescence. These observations support those of Puppo *et al.* (2005), who showed that adjacent infected cells within a nodule may not senesce at the same rate, indicated by differing cytoplasmic density between adjacent non-senescent and senescent cells, due to changes in the symbiosomes and the appearance of vesicles in the senescent cells. Cleared zones were not observed in naturally senescing nodules, but a more extensive

study that includes later time points or more fields of 60 dpi nodules may reveal their presence.

Cells undergoing induced senescence demonstrated nuclear deformation by 24 h and cleared zones by 36 h. Crinkling of the cell walls was observed by 36 h following application, but was no longer observable at 48 h. Cell wall crinkling may indicate cellular condensation, but other processes such as osmotic stress may be implicated, e.g., the vacuole may have burst, releasing its contents, making the cell hypotonic so that it lost fluid and appeared condensed. If osmotic balance was regained, the cell could regain its normal appearance before it was further dismantled.

While crinkling of cell walls may represent the initiation of cellular condensation or shrinkage (Mittler, 1998; Pennell and Lamb, 1997), there are few morphological studies linking this feature to senescence. Microscopic studies of naturally senescing nodules of soybean (Puppo *et al.*, 2005) or nitrate- or dark-induced senescence in pea and bean nodules (Matamoros *et al.*, 1999) did not reveal changes in cell walls; however, a study of root cells of soybean *root necrosis* (*rn*) mutants by Kosslak *et al.*(1997) showed convolution of the cell walls. Soybean *rn* mutants have disease symptoms in the absence of pathogens, including characteristics of cell death. Crinkling of the cell wall could also result from changes in osmotic pressure, perhaps due to the leakage of small molecules (ions) that occurs during petal senescence and in necrotic cell death (Panavas *et al.*, 1998; Ziegler and Groscurth, 2004). The role of cell wall crinkling remains unclear. Is it a prelude to further cell shrinkage or simply a result of changes in osmotic potential? To resolve this problem, more observations during both induced and natural senescence would be helpful, but electron microscopy of the same samples would allow

determination of whether the plasma membrane of the cells had begun to detach from the cell wall or not.

Nuclear deformation is a clear hallmark of PCD. In animal apoptosis, condensation and fragmentation of the nucleus are typically observed (Assuncao and Guimaraes and Linden, 2004; Leist and Jaattela, 2001; McKenna *et al.*, 1998; Ziegler and Groscurth, 2004), whereas in plants, condensation is more often observed (Mittler, 1998; Pennell and Lamb, 1997). Thus, the nuclear deformation observed in senescing soybean nodule cells in this study appears to be a common feature of plant PCD. The micrographs presented in this study suggest that nuclear deformation occurs in both naturally senescing and nitrate-treated soybean nodule cells. In the natural senescence series at 50 dpi, one nucleus appears to be fragmenting, while another appears to be condensing. A more comprehensive study during natural and induced senescence and the use of electron microscopy may allow a greater understanding of whether the nuclei fragment or condense or do both, and whether chromatin condensation occurs. However, determining the fate of the nucleus may not be sufficient for categorizing nodule PCD as apoptotic, autophagic, or some combination of these types, since nuclear changes occur during all types of cell death (Van Doorn and Woltering, 2005; Ziegler and Groscurth, 2004). The studies of nodule senescence discussed above (Matamoros *et al.*, 1999; Puppo *et al.*, 2005) focused on bacteroids and organelles, and make no mention of the nuclei, perhaps because nuclei cannot be viewed in every infected nodule cell in a given field (this study).

However, Alesandrini *et al.* (2003) found TUNEL-positive nuclei and chromatin condensation in senescing soybean nodules on five week-old plants. Only one TUNEL-positive nucleus per infected cell was observed, suggesting that the nuclei had not fragmented during senescence. Initially, TUNEL-positive nuclei were localized to the

periphery of the central zone but spread towards the centre by 7.5 weeks, suggesting that nodule senescence does not occur at the same time in all cells. By contrast, Groten *et al.* (2005) observed neither DNA laddering nor TUNEL-positive nuclei in naturally senescing pea nodules. Pea nodules are indeterminate, whereas soybean nodules are determinate. Differences in cell death mechanisms may exist between the two types or the growth conditions may have been sufficiently different to cause disparate outcomes. Events in the nucleus have not yet been useful for identifying the type of cell death, apoptosis or another type, that occurs in root nodules. It would be useful to look for DNA laddering in soybean root nodules during senescence since laddering distinguishes apoptotic cell death from other types, whereas the appearance of TUNEL-positive nuclei does not necessarily infer the type of cell death.

Cleared zones ("patches") were observed in nodules undergoing nitrate-induced, but not natural, senescence. Although this observation is intriguing, the relatively low magnifications employed make a precise interpretation difficult. Reduced cytoplasmic density accompanied by the appearance of vesicles was described in naturally senescing soybean nodules (Puppo *et al.*, 2005). Their patchy distribution was similar to that observed in the current study. Matamoros *et al.* (1999) observed cytoplasmic disruption under the light microscope, but not vesicles, in senescing pea and bean nodules, which could also give a patchy appearance. In other plant systems exhibiting PCD, or at least a form of PCD, vacuoles and vesicles have been observed. However, what do they represent? Are they indicative of generalized cellular disruption, an increasing central vacuole, shrinkage of the cells from their walls, autophagy, or are they vacuoles with a different origin? The role of vacuoles or vesicles in senescence and PCD is still not clear. These several different possibilities make interpretation of the data difficult.

One hypothesis about nodule symbiosis asserts that the bacteria must hide from plant defences (Mellor and Collinge, 1995). If this is the case, senescence may activate the previously suppressed defence response and nodules may be expected to exhibit features of PCD and the hypersensitive response (HR). Cell shrinkage and nuclear deformation could be reminiscent of apoptosis, whereas the cleared zones could represent vacuoles that appear during autophagy. It is likely that plant PCD does not fit neatly into a single category, but exhibits features of both apoptosis and autophagy, as has been suggested in studies of somatic embryogenesis in Norway spruce (Filonova *et al.*, 2000).

While nuclear deformation is the one feature of PCD that has been established in both naturally occurring and induced senescence of soybean nodules, cellular condensation, represented by cell wall crinkling, could also be occurring, and vacuolation is a possible feature of induced senescence. During naturally occurring senescence, nodule growth is likely asynchronous; therefore, features of PCD may not be easily detectable at predictable time points. In contrast, induction of senescence by nitrate creates a population of nodules that senesce as a group and may allow features of PCD to be more readily observable.

Various avenues of research are available to explore and expand upon these ideas:

i. a study of DNA laddering would be useful to ascertain whether this hallmark of apoptosis is present or not;

ii. an in-depth EM study of the progression of both natural and induced senescence in soybean root nodules would be useful to answer questions about the origin of the cleared zones - are they due to the formation of autophagosomes, vesicles or vacuoles, cellular disruption, or increased vacuolization? For example, autophagosomes would be discernible from other vesicles or vacuoles since they have a double membrane (Cuervo,

2004; Shintano and Klionsky, 2004). Do the vesicles bud from the central vacuole or another organelle such as the Golgi body, or do they arise in a different manner? Are they similar to the senescence-associated vacuoles described by Otegui *et al.* (2005) that harbour senescence-associated cysteine proteases, e.g., *SAG12*? EM would allow in-depth studies of the cell wall crinkling, the phenomenon of nuclear deformation and its relation to condensation or fragmentation, and changes to other organelles suggested by Puppo *et al.* (2005). A detailed examination of how each organelle changes has not been undertaken. Quantification of the traits observed over various fields prepared from several nodules at each time point might allow ordering of the changes in time.

4.1.2 Nitrogenase Activity and Lb mRNA Levels during Nodule Senescence

Nitrogenase activity and Lb mRNA were monitored during nodule senescence. The results of the acetylene reduction assays indicated that in naturally senescing soybean nodules, nitrogen fixation peaked at 22 dpi, declined to 40 dpi, and showed a slight, non-significant increase between 40 and 75 dpi. Nitrogen fixation in the root nodules was significantly reduced after 24 and 36 hours of nitrate and dark treatment of the soybean plants, respectively. Lb expression did not show a significant decline over the life of the soybean nodule under natural growth conditions, although nitrate and dark treatment produced declines in its message levels in the nodules over the 96 hour treatment period.

4.1.2.1 Nitrogenase Activity and Lb mRNA throughout Natural Development in Soybean Nodules

Only a few studies measuring the levels of nitrogenase activity in soybean nodules, or those of other legumes, throughout the plants' life spans are available for comparison (Gottlob-McHugh, 1989; Groten *et al.*, 2005; Hardy *et al.*, 1968; Nash and Schulman, 1976), as most studies have focused on the earlier stages. Of the soybean

studies, all used different cultivars and only the first two were carried out in the controlled environments of growth chambers. The studies of Nash and Schulman (1976) and Gottlob-McHugh (1989) are described here, but since the work of Hardy *et al.* (1968) was a field study, and the work of Groten *et al.* (2005) was carried out on pea, which forms indeterminate nodules, they will not be described in detail.

Nash and Schulman (1976) monitored nitrogenase activity from its first appearance at 12 dpi to 80 dpi (after pod filling). Nitrogen fixation rapidly increased between the first detection of nodules (12 dpi) and its peak level (35 dpi) and then exhibited a steady decline after 35 dpi until 80 dpi. In their study, the first flowers did not appear until 35 dpi, whereas in the current study they appeared earlier, between 22 and 24 dpi. The difference may be attributable to three factors: (i) different soybean cultivars (cv Kanrich vs cv Maple Arrow), since cultivars differ in terms of development and nitrogen fixation patterns (Hardy *et al.*, 1968); (ii) different regimes used for their growth, since Nash and Schulman (1976) used a 12 hour light period, a daytime temperature of 27°C, a nighttime temperature of 21°C, and a relative humidity of 50%, whereas the plants in the current study were cultivated under 16 hours of light per day with a daytime temperature of 25°C, a night time temperature of 20°C, and a relative humidity of 50%; (iii) different watering solutions.

In the study of Nash and Schulman (1976), nitrogenase activity had declined to ~66% of its peak value by 55 dpi (20 days post-peak) and 20% by 70 dpi (35 days post-peak). The current study also found the post-peak decline to be more gradual than the initial increase since a significant decline in nitrogenase activity was not observed until 40 dpi (18 days after the peak). At 40 dpi, nitrogenase activity had declined to approximately 4% of the peak value, suggesting that the decline of nitrogenase activity

was more rapid than that of the plants in the Nash and Schulman (1976) study. However, by 55 dpi, a recovery of nitrogenase activity was observed, a result similar to that of Klucas and Arp (1977), who observed an increase late in senescence between 50 and 57 dpi in cvs Calland and Beeson. No late stage recovery was observed by Nash and Schulman (1976), perhaps due to differences in cultivars, growth conditions, watering solutions, or measurement fluctuations and/or differences in re-inoculation rates.

Gottlob-McHugh (1989) found a similar pattern of nitrogenase activity in naturally developing nodules, although the events occurred at earlier times than in the current study. Nitrogenase activity was initially detectable at 10 dpi followed by a rapid increase to a maximum at 16-18 dpi, after which there was a slow decline between 18 and 30 dpi, such that by 36 dpi the remaining activity was ~58% of the peak value. In the current study, nitrogenase activity did not appear until 12 dpi, peaked at 22 dpi when flowering was observed, and declined slowly afterwards such that at 36 dpi nitrogenase activity had fallen to ~35% of the highest value. In addition, peak nitrogenase activity was ~1.5 nmol eth/min/g nod in the Gottlob-McHugh work, whereas it was ~0.42 nmol eth/min/g nod in the current study. Although the standard errors in the measurements preclude a detailed comparison of the two sets of experiments, the patterns are roughly comparable, especially given the differences in the watering solutions.

There is surprisingly little data available about changes in Lb mRNA throughout root nodule development. It is clear from the current study that steady state Lb mRNA does not decline in nodules of naturally developing soybeans once senescence begins, even though nitrogenase activity does.

Gottlob-McHugh (1989) observed Lb mRNA levels from the first appearance of nodules through early senescence. Lb mRNA was detectable in the first nodules at 10 dpi

and increased steadily until 24 dpi, and had declined slightly by 34 dpi. Since nitrogenase activity peaked at 16 dpi in Gottlob-McHugh's (1989) study, it is clear that Lb mRNA levels do not peak until after nitrogenase activity. In the current study, Lb mRNA is detectable by 15 dpi, but the small nodules visible at 12 dpi were not monitored for Lb since they did not provide sufficient biomass to extract the amounts of RNA required for Northern analysis. However, since the 12 dpi nodules were pink, it is likely that Lb mRNA would have been detectable. A Lb RT-PCR experiment could confirm this hypothesis. Little variation in Lb mRNA levels was observable between 15 and 70 dpi, and no peak was observed around the time of flowering (24 dpi), but a slight increase was observable between 20 and 28 dpi, indicating a lag between the peak in nitrogenase activity (22 dpi) and Lb message levels. This result is similar to that found by Gottlob-McHugh (1989). In the current study, a slight decline was observed after 28 dpi, but levels of Lb mRNA remained very similar between 32 and 50 dpi. At 55 dpi, a slight increase was observed, followed by a slight decline to 75 dpi. Gottlob-McHugh (1989) also only observed a slight decline in Lb mRNA after its peak. In general, the results observed in the current study are similar to those obtained by the previous study by Gottlob-McHugh (1989).

Nash and Schulman (1976) also observed changes in Lb throughout soybean nodule development, but they observed protein levels instead of mRNA levels. However, their data is discussed here since Lb is thought to be regulated at the level of transcription (Verma *et al.*, 1981), and therefore protein levels are likely to be indicative of mRNA levels. Nash and Schulman (1976) observed that Lb protein accumulation lags behind the appearance of nitrogenase activity prior to flowering and its decline occurs later than the decline in nitrogenase activity. They detected Lb protein by 12 dpi, whereas nitrogenase

was first detected 10 dpi, although clearly the different sensitivities of the detection methods could pose problems. They also showed that Lb protein levels peaked along with nitrogenase activity at flowering around 35 dpi, but after flowering, nitrogenase activity declined, whereas Lb levels remained nearly constant until after 70 dpi. Like the results of Nash and Schulman (1976), once Lb mRNA was detectable at 15 dpi in the current study, it showed little variation until 70 dpi, but did not peak at flowering (24 dpi). Similar to the results of the current study, Nash and Schulman (1976) did not observe any significant declines in Lb protein after flowering.

Lb expression does not correlate well with nitrogenase activity in the root nodules of naturally developing soybean plants and is not a useful marker for natural senescence. It has long been thought that Lb protein binds oxygen to bacteroids to furnish them with enough oxygen for respiration at the same time preventing the oxygen from destroying the nitrogenase enzyme (Layzell and Atkins, 1997; Smith and Gallon, 1993). However, it was not clear whether the presence of Lb was required for nitrogenase to be produced or whether it was synthesized as a consequence of nodule development (Bergersen and Goodchild, 1973; Hardy *et al.*, 1971; Johnson and Hume, 1973; Nash and Schulman, 1976; Smith, 1949). A recent study by Ott *et al.* (2005) employed an RNAi knockout approach in *Lotus japonicus* to abolish the expression of all symbiotic Lbs. Their results showed that neither nitrogenase protein, nor nitrogen fixation was detectable unless Lb was produced, showing that the presence of Lb protein is required for the establishment of nitrogen fixation in root nodules. It is not certain why this is the case, but Lb may have other roles in nodules besides carrying oxygen to the bacteroids, since its expression has also been detected in the symbiosis between plants and mycorrhizae fungus, in which

nitrogen fixation does not occur (Vieweg *et al.*, 2004). Its role in the plant-fungus symbiosis is not known.

The fact that Lb is required for the appearance of nitrogenase suggests that it should be present by the time nitrogenase activity is detectable. In the current study, nitrogenase was detectable by 12 dpi when nodules were first seen on the roots. Although Lb mRNA was not monitored in 12 dpi nodules, they were pink, suggesting that Lb was present and therefore, Lb mRNA would have been detectable by this time point. As mentioned above, RT-PCR employing Lb-specific primers could confirm this hypothesis. Similar to the results of Nash and Schulman (1976), it was observed that although nitrogenase activity declined, Lb mRNA did not decline with it, but remained at similar levels throughout nodule senescence. As suggested by Downie (2005), Lb may have another role in the nodule besides that of an oxygen carrier, although this secondary role remains unknown. In addition, Nash and Schulman (1976) found that the larger the nodule, the greater the amount of Lb per gram fresh weight of nodule tissue and per milligrams of protein, despite declining nitrogenase activity. As nodules increase in size, an increase in oxygen diffusion would be required. Therefore, more Lb could be required. As the soybean plants age, a greater proportion of the nodules are larger. In this study, RNA for Lb Northern analysis was prepared from a mixture of larger nodule sizes, and therefore, declines in Lb expression might not occur with nitrogenase activity since at the later time points, larger nodules containing higher levels of Lb could have constituted a significant proportion of the nodule population.

4.1.2.2 Nitrogenase Activity and Lb mRNA during Induced Senescence in Soybean Nodules

Nitrogenase activity and Lb mRNA levels were monitored during nitrate- and dark-induced senescence. Nitrogenase activity declined significantly by 24 and 36 hours after the initiation of nitrate and dark treatment, respectively. Lb mRNA did not significantly decline in nodules during either nitrate or dark treatment of the soybean plants, although a trend of decline was observed in both cases. In addition, the dark treatment appeared to cause a slightly greater decline in Lb mRNA levels than nitrate treatment.

Gordon *et al.* (2002) compared nitrogenase activity in soybean plants treated with 10mM KCl (control) or 10 mM KNO_3 over a 48 hour period. In their study, nitrogenase activity only began to differ at 18 hours and from that time the activity of the nitrate-treated plants remained lower than that of the controls. In the current study, significantly lower nitrogenase activity was observed by 24 hours of nitrate treatment. Since measurement of nitrogenase activity was not continuous in the current study, a significant drop in activity may have occurred earlier than 24 hours. In both studies, nitrogenase activity varied little between 24, 36, and 48 hours after the initiation of nitrate treatment.

The results obtained in this study are also similar to those obtained by Matamoros *et al.* (1999) on bean nodules treated with 10 mM KNO_3. They, too, found that a significant decrease in nitrogenase activity occurred by 24 hours of treatment, but it had only declined to ~70% of its initial value instead of 50% of this value, as was found in the current study. However, they also detected significant declines between 24 and 48 hours and again between 48 and 96 hours, whereas in the current study, no further decline was observed between 24 and 96 hours. Differences in results were expected due to interspecies variation between common bean and soybean.

Lb mRNA has been monitored by Northern analysis in nitrate-treated soybean nodules by Gordon *et al.* (1997) and Chan (1995). In agreement with the current work, slight declines in Lb message were detected in both studies. Although quantitative data was not collected, the declines were evident by 48 hours (Chan, 1995) and six days (Gordon *et al.*, 1997) of treatment. In the current study, statistically significant declines in Lb message were not observed, but by 36 hours, evidence of a slight decline in the message levels was apparent by viewing the Northern blot. Chan (1995) monitored Lb message at 24 and 48 hours after the initiation of nitrate treatment, but not at 36 hours; it is possible that a decline in Lb message would have been observed by that time point. Since the same cultivar was used for both this study and Chan's (1995) work, similar results were expected for the two studies. Congruent with the results obtained in the current study, Gordon *et al.* (1997) observed only a slight decrease in Lb mRNA by six days of treatment with 10 mM KNO_3. Unfortunately, earlier time points were not monitored; therefore, the time that the decline first occurred could not be determined. In general, the two previous studies are in agreement with the current one in that only slight declines in Lb mRNA levels are observed in response to nitrate.

More commonly, Lb protein has been monitored in nodules undergoing senescence. Matamoros *et al.* (1999) showed that Lb protein levels in bean nodules until 96 hours (four days) of treatment with 10 mM KNO_3. The slow decline in Lb protein mirrors the results by Chan (1995) and Gordon *et al.* (1997) for Lb mRNA.

Based on the results of the current study and the work of Gordon *et al.* (1997, 2002), Chan (1995), and Matamoros *et al.* (1999), Lb does not appear to be as sensitive a marker for nitrate-induced senescence as nitrogenase activity. Furthermore, a study carried out in lucerne nodules showed that much smaller nitrate concentrations (ie. ≤ 5

mM) were required to effect a decrease in nitrogenase activity than to achieve a decrease in Lb (Becana and Sprent, 1987). Concentrations of nitrate between 10 and 20 mM were required to cause detectable declines in Lb protein. It is likely that this difference in sensitivity to nitrate between nitrogenase and Lb is also found in other nodule types, such as those of soybean and common bean.

In a study by Pfeiffer *et al.* (1983a), acetylene reduction by soybean plants was nearly abolished after four days of dark treatment, with only 0.6% of the original pre-treatment nitrogenase activity remaining. The present study produced similar results in that dark treatment causes a dramatic reduction in the nitrogenase activity of soybean nodules. However, in the current work, approximately 10% of the pre-treatment activity remained after four days of treatment. That more nitrogenase activity remains after four days of dark treatment in the current study than in that of Pfeiffer *et al.* (1983a) may be attributable to cultivar differences. In addition, the photoperiod employed to grow the plants was one hour shorter in the Pfeiffer *et al.* (1983a) study than in the current one. However, both the current results and those of Pfeiffer *et al.* (1983a) imply that nitrogenase activity is more sensitive to darkness than nitrate since the activity declines more dramatically after four days in the dark than after four days of nitrate application.

Matamoros *et al.* (1999) measured nitrogenase activity in pea nodules of plants placed in the dark. Their study found that by 24 hours, the activity had been reduced to 7% of its initial value and was abolished by 48 hours, also suggesting that nitrogenase activity is more dramatically decreased by dark than nitrate. The more rapid decline in nitrogenase activity observed in the pea plants compared with the soybean plants of this study may be due to differences in nodule type, since pea plants produce indeterminate nodules, whereas soybean produces determinate ones, or species differences.

Although data on the effect of darkness on the Lb message is not available, levels of the protein have been monitored in response to this treatment. Pfeiffer *et al.* (1983a) found that Lb protein in soybean nodules was reduced by 25% by four days of darkness (96 hours) and that it was no longer detectable by 8 days of treatment. In the current study, the Lb message had dropped to approximately 2.5% of the level in the untreated plants by 96 hours, suggesting that the mRNA levels drop more rapidly than the protein ones, as would be expected given the relative stability of the Lb protein (Coventry and Dilworth, 1976).

During induced senescence, nitrogenase activity appears to decline more rapidly than Lb transcription, making it a more sensitive marker for nitrate- and dark-induced senescence in soybean nodules. The lesser response of Lb to the senescence-stimulating treatments may be due to the stability of the protein, low RNA turnover, or an additional, as yet uncharacterized role of this protein in nodules (Coventry and Dilworth, 1976; Downie, 2005).

It is clear that nitrogen and darkness both cause nodule senescence. However, darkness is thought to cause a decline in nodule function through reduced levels of photosynthate (Gordon *et al.*, 1993), whereas it is not certain how nitrate induces nodule senescence. Nodule function may be reduced by nitrate through carbohydrate-deprivation (Gordon *et al.*, 1997; Vessey and Waterer, 1992), feedback inhibition by products of nitrogen assimilation such as glutamine (Neo and Layzell, 1997), or decreased O_2 diffusion into nodules, which would reduce the rate of respiration of the bacteroids (Gordon *et al.*, 2002). A study by Fujikake *et al.* (2003) provides more evidence for the carbohydrate-deprivation hypothesis, since sucrose supply is decreased in nodules during nitrate-induced senescence, but questions remain about whether this is a result of nitrate

application or a secondary effect due to reduced use of the carbohydrate by the senescing nodules.

4.2 The *SAN1* Gene Family

One of the main objectives of this thesis was the isolation and characterization of genes that are differentially expressed during senescence. The three *SAN1* genes isolated during the course of this work are found in tandem in the order *SAN1B*, *SAN1C*, and *SAN1A*. Within ~4 kb upstream of the *SAN1* genes, a single gene, called *MSG*, encoding a major latex protein (Stromvik *et al.*, 1999), was identified. The position of *MSG* in the soybean genome could reveal the positions of the *SAN1* genes, but unfortunately, it has not yet been mapped (Soybase, http://soybase.org/). No other genes were identified.

4.2.1 The *SAN1* Genes Encode Unknown IPNS-Related Dioxygenases that are Close Relatives of the GA20Oxs

The results of BLASTP searches and examination of the motifs of the SAN1 proteins suggest that they belong to the IPNS-related dioxygenase family, which includes all known plant dioxygenases and the IPNS genes of microorganisms (De Carolis and De Luca, 1994). The IPNS-related dioxygenases belong to a gene superfamily known as the 2-oxoglutarate-Fe (II) oxygenases. Sequence conservation between plant dioxygenases is often low, making it difficult to identify a given dioxygenase by sequence analysis alone (Prescott and John, 1996). For example, the three GA20oxs of *Arabidopsis* are only 64% identical. Between plant species, the situation is even worse, since in these cases GA20oxs can only be expected to share approximately 53% identity. In general, amino acid sequences of plant dioxygenases of unrelated functions are 27-32% identical; therefore, any identity higher than this is considered indicative of a closer evolutionary

relationship. The SAN1 predicted proteins are most closely related to unknown dioxygenases of the divergent species *Arabidopsis*, rice, and maritime pine, which are in turn most closely related to the GA20oxs.

Another way of determining the relationships between dioxygenases is provided by the high degree of conservation between intron number and position (Prescott and John, 1996). Plant dioxygenases may have up to three introns at four variable positions in the gene. For example, the genes encoding ACCoxs and H6Hs have three introns at positions 1, 3, and 4, whereas GA20oxs and F3Hs have two introns at positions three and four and two and four, respectively. Length of the encoded proteins and whether the dioxygenase is encoded by one or more genes may also provide clues as to function for unknown dioxygenases.

Since the SAN1 predicted proteins are encoded by a multigene family, they are more likely to be related to ACCoxs, GA20oxs, GA3oxs, and GA2oxs than to the single gene encoded F3Hs, FNSs, FLSs, LDOXs, or H6Hs. *SAN1A* and *SAN1B*, the two *SAN1* genes which encode full length dioxygenases encode proteins of 352 and 353 amino acids, respectively. A scan of the plant dioxygenses in the protein database at NCBI (http://ncbi.nlm.nih.gov) revealed that the lengths of characterized plant dioxygenases range between approximately 310 and 400 amino acids in length. ACCoxs tend to be at the shorter end of this range, around 312 to 325 amino acids. Other plant dioxygenases, including LDOXs, F3Hs, FLSs, GA20oxs, GA3oxs, GA2oxs, and the H6Hs range in size from approximately 325 to 400 amino acids. SAN1A and SAN1B fall within the middle of the latter range for dioxygenase sizes; therefore, their sizes do not indicate which of the latter group of dioxygenases they are most closely related to.

The protein sequences of *SAN1A* and *SAN1B* are 83% identical and 89% similar. This level of identity is more typical of the ACCox multigene family, since within species identities are >90% and between species identities are found to be ~70% (Prescott and John, 1996). Interestingly, when TBLASTN searches of the genome databases of *M. truncatula* were carried out, more than one gene sequence >70% identical to the *SAN1A* and *SAN1B* genes was found. Unfortunately, potential homologues in *L. japonicus* were not found. The putative *Arabidopsis* homologues of the *SAN1* genes identified by a BLASTP search of the *nr* database and by TBLASTN searches of the *Arabidopsis* genome at TIGR (http://www.tigr.org/tdb/e2k1/ath1) were around 58% identical to the *SAN1* gene products in soybean. In addition, the putative rice and maritime pine homologues also shared between 38% and 55% identity with the SAN1 predicted proteins. These levels of identity between distantly related species appear to be slightly higher than expected for GA20oxs, but lower than expected for ACC oxidases. However, since the BLASTP results show that the SAN1 predicted proteins share greater than 50% identity with known GA20oxs, it is likely that they have a closer evolutionary relationship to these plant dioxygenases than with others. Unfortunately, it is not possible to classify the *SAN1* genes based on their levels of identity with homologues in other plant species. The *SAN1* genes have two introns each, similar to the genes of F3Hs andGA20oxs. The two introns are found at conserved positions three and four; this suggests a closer relationship of the *SAN1* genes to those encoding GA20oxs rather than F3Hs.

Since sequence and gene structure analyses were not able to determine what the *SAN1* genes encode, a gene tree was constructed to help determine their degree of evolutionary relatedness to the other enzymes in the IPNS-related dioxygenase family. The tree again revealed that the *SAN1* genes are most closely related to the GA20oxs, but

the *SAN1*-related unknown dioxygenases from *Arabidopsis*, rice, and maritime pine identified by the BLASTP searches treed together with the SAN1 proteins in a separate clade from the GA20oxs, despite the divergence of the species in which they had been identified. This result does not rule out the possibility that the *SAN1* genes encode GA20oxs, but it does suggest that they could encode a new subset of this gene family or novel dioxygenases. The evolutionary relationships determined by the neighbour-joining tree in this study do not agree with those determined by chemical analyses of various plant species. For example, chemical analyses of flavonoids across various plant species have suggested that F3Hs, FLSs, and FNSs were the first plant dioxygenases to have arisen, followed by the enzymes of the gibberellin synthesis and degradation pathway. The ACCoxs and LDOXs are thought to have evolved next and the H6Hs are thought to be the most recently evolved. However, the tree in this study implies that the GA2oxs are the oldest dioxygenases, followed by the H6Hs. The rest of the dioxygenases appear to be more recent. Differences between the molecular phylogenetic analysis of the dioxygenases and the chemical analyses are not likely to be due to the methodology of tree construction since protein parsimony and maximum likelihood trees gave similar results (data not shown). Studies of the molecular evolution of plant dioxygenases are still rare. However, one recent study of the molecular evolution of the flavonoid dioxygenases from the Apiaceae family indicates that LDOXs and FLSs are more closely related to each other than either of them is to the FNSs and F3Hs (Gebhardt *et al.*, 2005). This result was also observed in the tree constructed for the current study. The phylogeny of the plant dioxygenases is still not clear and further studies will be required to tease apart the evolutionary history of these enzymes. In general, the divergent sequences of

homologous plant dioxygenases makes it difficult to analyze the evolutionary relationships between them by any of the commonly used phylogenetic methods.

To determine the true identity of the dioxygenases encoded by the *SAN1* genes, it would be necessary to perform biochemical assays for their activities. For example, the genes of *SAN1A* and *SAN1B* could be cloned into expression vectors, the proteins produced in *Escherichia coli*, and then tested for their substrate specificities. FLS from parsley, GA20ox from pea, GA3ox from *Arabidopsis*, and F3H from petunia have all been successfully expressed in *E. coli* (Leonard *et al.*, 2005; Lester *et al.*, 2005; Lukacin *et al.*, 2000). Unfortunately, full length cDNA sequences for both *SAN1A* and *SAN1B* were not available at the outset of this project, and once they had been obtained, budgetary and time constraints did not allow for these experiments to be carried out. In addition, discussions with Dr. Peter Hedden at the International Society Plant Molecular Biology Congress 2000 (Québec) indicated that it could be difficult to carry out these expression experiments. Assaying for activity could also prove difficult since it is has not been possible to determine the probable functions of the *SAN1*-encoded dioxygenases. Prior to these biochemical tests, a metabolomics approach to determining possible functions of the *SAN1* gene products could be carried out. For example, extracts of both pre-senescent and senescent root nodules could be prepared and their metabolite profiles compared. Unidentified compounds that show changes between non-senescent and senescent nodule extracts could be candidates for further investigation. Their identification might lead to clues as to the function of the *SAN1* gene products. Genetic approaches to abolish, attenuate, or increase the expression of the *SAN1* genes could also be used to narrow down the number of possible dioxygenase enzyme activities to test for. Sense constructs could be prepared for overexpression and antisense constructs or an

RNAi approach could be used to knockout or knockdown the expression of the genes. Phenotypes of plants overexpressing, not expressing, or expressing the *SAN1* genes at low levels could be observed. For example, if the *SAN1* genes encode GA20oxs, overexpressing the genes might cause taller than average plants, whereas knockdowns or knockouts of the genes might result in dwarf plants (Hedden and Phillips, 2000); these phenotypes are readily observable. In addition, manipulation of gibberellin levels and comparing the resultant phenotypes to transgenic plants with altered levels of *SAN1* gene expression might provide clues as to whether the *SAN1* genes encode GA20oxs. However, since changes in phenotype might not be obvious, chemical profiles of non-transgenic plant extracts versus transgenic plant extracts might allow the identification of changes in the physiology of the plant resulting from over- or underexpression of the genes.

Unfortunately, sequence analysis, bioinformatics, and evolutionary approaches have not revealed the function of proteins encoded by the *SAN1* genes, only that they are plant dioxygenases and part of the larger superfamily of 2-oxoglutarate Fe(II) oxygenases. However, the number and position of the introns of the *SAN1* genes, the results of a BLASTP search, and the gene tree of the plant dioxygenases suggest that they are most closely related to the GA20oxs. The BLASTP search also revealed that the most closely related dioxygenases to the SAN1 predicted proteins were those of unknown dioxygenases found in divergent species such as the angiosperms *Arabidopsis* and rice and the gymnosperm, maritime pine. This result suggests that the *SAN1* genes are likely to encode novel dioxygenases that may have evolved from the GA20oxs. It is also possible that they could encode divergent types or as yet undescribed paralogues of characterized dioxygenases.

4.2.2 Occurrence of *SAN1*-Like Sequences in Other Legume Genomes

Future studies on the role(s) of SAN1 proteins would be enhanced if their manipulation by reverse genetics was possible. While soybean can be transformed, the process can take over 15 months from infection or bombardment to seed (D. Johnson, personal communication). Thus, a model system would be immensely helpful if it contained identifiable *SAN1* gene homologues. As previous BLASTP results (Section 3.4.1) had identified potential homologues in *Arabidopsis* and rice, searches were limited to the model legumes *L. japonicus* and *M. truncatula* (Table IX). Protein sequences were used due to expected divergence in sequences between species.

L. japonicus TAC and BAC end sequences (http://www.kazusa.or.jp/lotus/) and *M. truncatula* BAC end sequences (http://www.medicago.org/genome/) were scanned. Only matches of low significance were obtained for both *Lotus* databases while *Medicago* revealed some potential *SAN1* homologues; however, these lack extensive annotation. A scan of the *L. japonicus* Gene Index with the SAN1 predicted protein sequences did not yield any highly significant matches. When the *M. truncatula* Gene Index was scanned with the three SAN1 protein sequences, the top match was the same for each of SAN1A, SAN1B, and SAN1C, although other significant matches were also found for each SAN1 sequence (not shown). Given that the model legumes are closely related to soybean, it is likely that *SAN1* homologues are to be found in *M. truncatula* and *L. japonicus*; however, at this time, only sequences from *M. truncatula* have been identified. Thus, while none of the results could be identified as *SAN1* homologues with any degree of confidence, as the databases improve such homologues may be identified.

Once completed genomic data becomes available for soybean, *Lotus*, and *Medicago*, the total number of *SAN1* genes in the legume genomes will be revealed. In

Table IX: The Occurrence and Expression of Potential *SAN1* Homologues in Other Legumes

The sequenced portions of the legume genomes of *Lotus japonicus* and *Medicago truncatula* were scanned with the SAN1 protein sequences. The scans were carried out using the TBLASTN algorithm on June 20, 2006. The TIGR Gene Indices of *Lotus japonicus* and *Medicago truncatula* were scanned with each of the SAN1 sequences using the TBLASTN algorithm on June 20, 2006 to determine if homologous expressed sequences could be found in these species. The database scanned, the matches accompanied by their identifiers and their description, and the significance of the matches are presented. For each scan, only the top match is displayed. +ves, number of similar amino acids; 2OG, 2-oxoglutarate; BAC, bacterial artificial chromosome; GB, GenBank database; IDs, number of identical amino acids/nucleotides; LDOX, leucoanthocyanidin dioxygenase; TAC, transformation-competent artificial chromosome; TC, tentative contig.

Database Scanned	*SAN1* Sequence	Match No. and Identifier	Description	Identities and/or Similarities	E-value
Lotus japonicus TAC ends	SAN1A	(1)AG242743 (GenBank) LjT11f04_not	Not annotated.	IDs: 50/161=31% +ves: 76/161=47%	2e-16
	SAN1B	(1)AG242743 (GenBank) LjT11f04_not	Not annotated.	IDs: 50/160=31% +ves: 74/160=46%	3e-16
	SAN1C	(1)AG240792 (GenBank) LjT07e07_not	Not annotated.	IDs: 31/109=28% +ves: 48/109=43%	8e-06
Lotus japonicus BAC ends	SAN1A	(1)AG222284 (GenBank) LjB06013_r	Not annotated.	IDs: 41/70=58% +ves: 50/70=70%	1e-17
	SAN1B	(1)AG234196 (GenBank) LjB26i11_f	Not annotated.	IDs: 41/112=36% +ves: 66/112=58%	3e-17
	SAN1C	(1)AG219728 (GenBank) LjB01p10_r	Not annotated.	IDs: 14/47=29% +ves: 29/47=60%	7e-04
Medicago truncatula BAC ends	SAN1A	(1)mte1-68M12FM1	Not annotated.	IDs: 87/124=70% +ves: 106/124=85% IDs: 77/102=75% +ves: 87/102=85%	1.6e-86

Database Scanned	*SAN1* Sequence	Match No. and Identifier	Description	Identities and/or Similarities	E-value
Medicago truncatula BAC ends	SAN1B	(1)mte1-68M12FM1	Not annotated.	IDs: 89/124=71% +ves: 106/124=85% IDs: 74/97=76% +ves: 88/97=90%	1.1e-87
	SAN1C	(1)mte1-68M12FM1	Not annotated.	IDs: 7/8=87% +ves: 8/8=100% IDs: 57/92=61% +ves: 74/92=80%	5.1e-30
Lotus japonicus Gene Index	SAN1A	(1)BI417790	Singleton.	IDs: 101/147=68% +ves: 120/147=81%	3.4e-56
	SAN1B	(1)BI4417790	Singleton.	IDs: 102/153=66% +ves: 120/153=78%	5.7e-54
	SAN1C	(1)AV411854	Singleton.	IDs: 54/118=45% +ves: 74/118=62%	1.4e-24
Medicago truncatula Gene Index	SAN1A	(1)TC103144	Contig contains 6 ESTs.	IDs: 276/352=78% +ves: 317/352=90%	3.2e-155
	SAN1B	(1)TC103144	Contig contains 6 ESTs.	IDs: 259/351=73% +ves: 307/351=87%	1.2e-146
	SAN1C	(1)TC103144	Contig contains 6 ESTs.	IDs: 83/126=65% +ves: 104/126=82%	9.7e-44

addition, the evolutionary relationships of the *SAN1* genes to their homologues in the other legumes and to other plant dioxygenase genes may become clearer. Transcriptomics may allow the correlation of *SAN1A* and *SAN1B* expression with other genes expressed in root nodules, possibly revealing clues to their function and their role in nodule physiology.

4.2.3 The Sequence of *SAN1A* is Different between cvs Resnik and Maple Arrow

Comparison of the deduced cDNA sequences of *SAN1A* from cvs Resnik and Maple Arrow revealed two single nucleotide changes giving rise to two amino acid changes- proline to leucine and aspartic acid to valine. These differences were not thought to be due to sequence error, since as described in Sections 3.2.1 and 3.2.2, two independent sequences of cv Resnik were 100% identical and the cv Maple Arrow gene was cloned and sequenced three times. Due to functional constraints, synonymous changes between orthologues are expected (Graur and Li, 2000); however, cvs Resnik and Maple Arrow do not share common ancestors for at least two previous generations back (Mansoorian, 2004). These nonsynonymous changes do not occur within any of the five canonical plant dioxygenase domains, nor do they affect residues responsible for 2-oxoglutarate or Fe(II) binding in the protein sequence (Figure 33). Proline and leucine both have aliphatic side chains and are therefore hydrophobic; however, proline may introduce changes in protein structure due to its having an aliphatic ring rather than a chain (Lehninger *et al.*, 1993). The second amino acid change occurs within the leucine-rich domain that may be responsible for protein-protein or protein-membrane interactions (Britsch *et al.*, 1993). Although the change does not affect any of the leucine residues themselves, aspartic acid has a negatively charged side chain at physiological pH, whereas

valine has an aliphatic one, which could have an affect on protein structure. The significance of these changes is not known.

4.2.4 Relative RT-PCR was Useful for Determining the Expression Patterns of the *SAN1A* and *SAN1B* Genes

An RT-PCR-based approach for determining the expression patterns of *SAN1A* and *SAN1B* was chosen since *SAN1A* expression proved to be too low for Northern analysis. Relative RT-PCR was chosen since it is a cost effective, yet sensitive and semi-quantitative method of determining gene expression. Several controls were necessary to validate the methodology (Sections 2.14.5, 3.3.1.3, and 3.3.1.4). Primers designed to amplify regions of *SAN1A* and *SAN1B* gave products 100% identical to their intended targets, while the Ambion proprietary 18S primers gave a product 99% identical to the soybean 18S sequence (GenBank X02623.1). Sequencing errors or incorporation of errors when RT or *Taq* polymerase was used or cultivar differences may account for the 18S sequence variation.

Relative RT-PCR was able to detect significant differences in steady state mRNA levels of *SAN1A* and *SAN1B* in nodules induced to senesce by nitrate and dark treatment and in the tissue distribution experiments. However, although trends could be detected in *SAN1A* and *SAN1B* levels in naturally senescing nodules or in *SAN1A* levels during light and dark cycling, significant differences between expression levels were not observed. Northern analyses of Lb mRNA levels in naturally senescing nodules suggested that there was much variability in untreated soybean nodules, making it difficult to detect significant differences in gene expression.

Since each multiplex PCR reaction is optimized individually so that the target and 18S are amplified to similar levels, it was not possible to quantitatively compare levels of

SAN1A and *SAN1B* mRNA. This would require optimization with primers designed to amplify *SAN1A* and *SAN1B* with 18S in a single reaction. However, with qPCR, a quantitative method that determines the actual amounts of each mRNA in a given tissue under a given condition, *SAN1A* and *SAN1B* levels could be measured with ease.

4.2.5 *SAN1A* and *SAN1B* Exhibit Radically Different Patterns of Expression

When the expression patterns of *SAN1A* and *SAN1B* were compared in naturally senescing nodules, nodules induced to senesce by nitrate or dark treatments, and in the various tissues of soybean, the expression of *SAN1A* and *SAN1B* mRNA differed markedly. Since *SAN1A* and *SAN1B* occur in tandem, a reasonable hypothesis is that their function and regulation have diverged following gene duplication (Graur and Li, 2000).

SAN1A expression resembles that of a late nodulin and is roughly correlated with nodule development and senescence, as measured by the acetylene reduction assay. Its expression was low in roots prior to the appearance of nodules, increased with the onset of nitrogen fixation to reach a maximum at ~22-24 dpi when nitrogenase activity peaked, and declined once senescence had begun. *SAN1A* expression increased again during late senescence, whereas nitrogenase activity only showed a slight increase during the late stages of senescence. In contrast, *SAN1B* expression does not appear to show any correlation with nitrogenase activity. Its message levels were higher in roots before nodules appeared, low in actively nitrogen fixing nodules, and higher again in senescing nodules from 40 dpi onwards. The tissue distribution results also confirmed these observations since *SAN1A* expression was higher in 24 dpi than in 40 dpi nodules and *SAN1B* expression was higher in roots that in nodules. Assuming that the levels of protein correlate with the levels of mRNA, these results suggest that *SAN1A* has a functional role related to nodule development whereas *SAN1B* is more likely to have a role during the

latter stages of nodule senescence. Both genes are expressed in other tissues and must play a role in these tissues as well.

Applying high levels of fixed nitrogen to root nodules causes generalized senescence to occur, affecting all of the processes in the nodule, including gene expression (Becana and Sprent, 1987). Nitrate treatment had opposite effects on the expression of *SAN1A* and *SAN1B*; *SAN1A* mRNA levels decreased while *SAN1B* mRNA levels increased. The decline in *SAN1A* expression paralleled the changes in nitrogenase activity; however, the decline was not continuous as was observed for nitrogenase activity (Section 3.1.2.1.5.2), but showed decreases followed by recoveries every 12 hours until 48 hours, after which point levels continued to be reduced. Since this is suggestive of light regulation, it was hypothesized that these declines and recoveries represent diurnal variation in *SAN1A* gene expression dampened by nitrate treatment. Preliminary results of a test for diurnal variation of *SAN1A* were reported in Section 3.3.1.5.3 and are discussed in Section 4.2.6. Gordon *et al.* (2002) had previously observed diurnal variation in nitrogenase activity dampened by nitrate treatment in nodules. However, diurnal variation in nitrogenase activity was not observed in the current study in nitrate-treated or untreated nodules. This failure to observe the diurnal variation of nitrogenase activity may relate to the sampling times, which may not have been ideal for detection of its variation. Plants were only tested every 12 hours until 48 hours of treatment, after which time assays were conducted every 24 hours until 96 hours of nitrate treatment. In contrast to the general decline observed for *SAN1A* expression, *SAN1B* mRNA levels exhibited a slow and steady increase between 0 and 96 hours of nitrate treatment and there were no indications of light-regulated changes in expression. As suggested for natural senescence, the SAN1B protein is likely to have a role in the process of senescence.

Dark treatment causes rapid declines in the mRNA of key nodule proteins such as Lb, sucrose synthase, and glutamine synthetase in root nodules (Gordon *et al.*, 1993). Dark treatment affected *SAN1A* and *SAN1B* expression. Between 0 and 96 hours of dark treatment, *SAN1A* mRNA rapidly declined while *SAN1B* expression increased steadily. Darkness causes changes similar to those observed in natural senescence, suggesting that these processes may be related (Pfeiffer *et al.*, 1983a). Therefore, it might be expected that *SAN1A* expression would decrease and *SAN1B* expression would increase during dark treatment, as seen for their expression during natural senescence. Again, the results imply that the SAN1A protein may have a role in nodule function, whereas SAN1B would be more likely to be involved in nodule senescence. Since *SAN1A* appears to be expressed chiefly in non-senescing nodules, and *SAN1B* is expressed in senescing nodules, it is possible that *SAN1B* replaces *SAN1A* during senescence.

Tissue distributions of *SAN1A* and *SAN1B* expression again highlighted differences in their regulation. Although both genes were expressed in all tissues tested, their patterns differed markedly. Transcript levels of *SAN1A* were relatively low in seedlings, senescing leaves, flowers, and pods, with slightly higher levels being found in cotyledons, roots, 40 dpi nodules, young and old leaves, and the highest levels being found in 24 dpi nodules. The expression of *SAN1B* was relatively low in 24 and 40 dpi nodules, slightly higher in roots and stems, and relatively high all other tissues. These results suggest that *SAN1A* is generally nodule-enhanced, but that *SAN1B* is not.

The results of the expression studies indicate that the regulation of *SAN1A* and *SAN1B* have diverged since duplication. In addition, the broad tissue distribution of their expression suggests that they were recruited from other plant pathways for use in nodulation, which is typical of genes involved in the symbiosis (Silverstein *et al.*, 2006).

It also appears that *SAN1A* may replace *SAN1B* during nodule development, since *SAN1A* is expressed at higher levels in non-senescent nodules, whereas *SAN1B* is expressed at higher levels in senescent nodules.

Although the predicted coding sequence of *SAN1C* indicates that it forms a truncated protein, a non-quantitative RT-PCR test was carried out to determine whether it was expressed. Interestingly, the distribution of *SAN1C* mRNA among tissues was found to be similar to the distribution of *SAN1A* mRNA. This observation, along with sequence and intron comparisons presented in Sections 3.2.4 and 3.2.5 suggest that *SAN1A* and *SAN1C* may be more closely related to one another than either one of them is to *SAN1B*. The duplication that gave rise to these genes may be more recent than that which gave rise to *SAN1B* and the progenitor of *SAN1A* and *SAN1C*.

A scan of the Soybean Gene Index at TIGR (http://www.tigr.org) identified seven ESTs with >80% identity to *SAN1A* and *SAN1B* and one that was 100% identical to *SAN1A*. These ESTs were expressed in a variety of tissues, including seeds, cotyledons, leaves, roots, and flowers, and in response to salt stress, wounding, salicylic acid application, and pathogens. Specifically, TC205378, the probable homologue of *SAN1A*, was expressed in germinated seeds, in whole seedlings, and in response to pathogens. These results extend the tissues and conditions where *SAN1* homologues are expressed in soybean. If we include in this analysis scans of the rice and *Arabidopsis* Gene Indices (http://www.tigr.org), then putative *SAN1* genes from these species are also expressed in a wide variety of tissues including roots, leaves, flowers, and seeds, and in response to various abiotic and biotic stresses in their respective species (dark, drought). One could speculate that if the regulatory systems of the *SAN1* genes and their homologues are conserved, the advantages of *Arabidopsis* as a model system could be exploited.

4.2.6 *SAN1A* Expression May be Regulated by Light

Several genes expressed in nodules, including nitrogenase (of bacterial origin) and asparagine synthetase are responsive to light (Gordon *et al.*, 2002; Osuna *et al.*, 1999). When *SAN1A* expression in nodules was monitored during nitrate treatment, a pattern of expression suggestive of light regulation was observed; therefore, it was hypothesized that the declines and recoveries in its mRNA levels might represent diurnal variations in gene expression (Sections 3.3.1.5.3 and 4.2.6). A test to investigate whether *SAN1A* expression did exhibit diurnal variation in the absence of nitrate was carried out. Levels of *SAN1A* mRNA were monitored more frequently, every 4 hours for 44 hours, starting at 21 dpi. The results suggested that there could be diurnal variation, since message levels appeared to increase from early in the day until late in the day and to decline in the dark (Figure 32). However, these results were not statistically significant and must be treated as preliminary. A quantitative technique such as qPCR would be useful in providing definitive evidence.

Most plant dioxygenases are also known to be regulated by light, e.g., GA20ox and ACCox mRNA exhibit light-dark cycling (Jackson *et al.*, 2000; Machackova *et al.*, 1997); GA3oxs and GA2oxs show increased and decreased expression, respectively, when exposed to red light (Nakaminami *et al.*, 2003; Toyomasu *et al.*, 1993); enzymes involved in flavonoid and anthocyanin synthesis, including FLS, FNS, F3H, and LDOX are all regulated by light (Gong *et al.*, 1997; Hartmann *et al.*, 2005; Lukacin *et al.*, 2001). Since *SAN1A* is evolutionarily related to these enzymes, it is possible that it is regulated by light. If this is the case, the advantages of *Arabidopsis* as a model system could be exploited to further investigate the possibility of light regulation.

4.2.7 Regulatory Elements of *SAN1A* and *SAN1B*

Due to their differing expression patterns, *SAN1A* and *SAN1B* were not expected share many regulatory elements except those necessary for basal transcription (TATA boxes, etc.) and tissue-specific sequences. They would not be expected to share nitrate and dark responsive elements, since they respond differently to these treatments. *SAN1A* might also be expected to have light responsive elements (Section 4.2.6). The upstream regions, introns, and 3′ UTR regions of *SAN1A* and *SAN1B* were scanned for shared plant regulatory elements. The identification of regulatory elements *in silico* is a preliminary step - motifs are often short (4-6 nt), and therefore may arise by chance and may not be functional. Experiments that monitor the expression of genes containing mutated regulatory element motifs would be necessary to identify the actual elements responsible for directing expression of the *SAN1* genes. Table VIII (Section 3.4.3) provides an extensive list of the motifs identified by this approach, but here, only a subset of these motifs whose functions could relate mainly to nodule development, nitrate and dark responses, and light regulation will be discussed.

SAN1A and *SAN1B* share the regulatory element motifs of DOFCOREZM and ARR1AT in their upstream regions. DOFCOREZM is the core binding site for DOF proteins in maize, which are plant-specific transcription factors involved in various processes (Yanagisawa and Schmidt, 1999). An *Arabidopsis* DOF protein, OBP1, interacts with bZIP transcription factors that respond to stress, enhancing their binding to the DNA (Zhang *et al.*, 1995). Given that *SAN1A* and *SAN1B* expression responds to dark and nitrate stress in nodules, it may not be surprising that they share regulatory elements that bind transcription factors controlling expression in response to stress. ARR1AT is an element that is directly responsive to cytokinins and which has so far been found in

Arabidopsis and rice (Ross *et al.*, 2004; Sakai *et al.*, 2000, 2001). In nodules, cytokinins have a role in processes of early nodule development, such as root hair deformation and curling, and induce the expression of genes such as *ENOD2*, *ENOD12*, and *ENOD40* (Bladergroen and Spaink, 1998; Cooper and Long, 1994; Fang and Hirsch, 1998; Hirsch *et al.*, 1989). Since the highest levels of *SAN1A* expression are found in newly formed nodules, it is possible that *SAN1A* is upregulated in response to increased cytokinins during early nodulation. *SAN1B* expression could be repressed by cytokinins and therefore its levels are only able to increase later in nodule development when cytokinin levels decline.

Potential regulatory elements were also shared between the introns of *SAN1A* and *SAN1B*. In the first intron, these elements included the MYB and MYC motifs. The MYB and MYC consensus sequences are typically found in genes that respond to abiotic stresses such as water deficit (Abe *et al.*, 2003; Urao *et al.*, 1993). Both *SAN1A* and *SAN1B* respond to dark and nitrate stress in nodules. However, these elements are typically found upstream of the coding sequences and not in the introns. A MYB protein isolated from *Petunia hybrida* has also been shown to activate transcription of the chalcone synthase genes, which encode key proteins in flavonoid synthesis (Solano *et al.*, 1995). The *SAN1* genes are related to genes encoding other proteins involved in flavonoid synthesis, such as FLS and F3H. If as part of the phenylpropanoid pathway, FLS and F3H share regulatory elements with chalcone synthase, they may also share regulatory elements with the *SAN1* genes. In the second introns of the *SAN1* genes, only three sequences corresponding to regulatory elements were identified. These sequences included the previously described ARR1AT motif. ARR1AT was discussed above, but it has been identified in upstream regions and not introns (Ross *et al.*, 2004; Sakai *et al.*,

2000). That regulatory elements were identified in the introns was not surprising since they have been found in the introns of various genes in plants. For example, the vernalization genes of barley are thought to be regulated by sequences in the first intron and the rice tubulin genes contain regulatory elements required for high levels of expression in their first intron (Fiume *et al.*, 2004; Fu *et al.*, 2005). In addition, the *ZWICHEL* gene of *Arabidopsis*, which encodes a Ca^{2+}-calmodulin-regulated kinesin, is partially regulated by the last two introns of the upstream *β-HYDROXYBUTYRYL-COA HYDROLASE 1* gene (Reddy and Reddy, 2004). Although many of the potential regulatory elements shared by the introns of the *SAN1A* and *SAN1B* genes were originally identified in the upstream regions of their respective genes, it is possible that they can reside in introns.

Lastly, a search was carried out for potential regulatory elements shared between the 3′ UTRs of the *SAN1A* and *SAN1B* genes. Several potential regulatory motifs were found including POLASIG1 and WBOXNTERF3. The POLASIG1 motif, AATAAA, was previously identified in the 3′ UTRs of *SAN1A* and *SAN1B* and is likely to be functional in directing the polyadenylation of their messenger RNAs (Wu *et al.*, 1995; Rothnie, 1996) (Section 3.2.4). WBOXNTERF3 was isolated from the wound-inducible tobacco gene *ERF3*, which is a repressor of transcription when wounding occurs (Nishiuchi *et al.*, 2004). *SAN1A* and *SAN1B* expression is affected by dark and nitrate stress in nodules, and the EST data for closely related sequences in soybean show that these genes may be generally responsive to stress. Wounding may be another stress which affects their expression levels.

When the upstream regions of each of *SAN1A* and *SAN1B* were independently scanned for regulatory elements, many more motifs were found for each gene. *SAN1A*

contained a total of 374 motifs, comprising 77 different types, whereas *SAN1B* contained 430 motifs belonging to 82 different types. Both the *SAN1A* and *SAN1B* genes contained regulatory elements involved in various responses, including responsiveness to light, hormones, abiotic and biotic stresses, senescence, and sugar. In addition, the *SAN1* genes contained regulatory elements that promote gene expression in a wide variety of tissues, including leaves, shoots, roots, nodules, seeds, cotyledons, and flowers. Elements that promote flavonoid biosynthesis were also identified. These results suggest that the *SAN1A* and *SAN1B* genes respond to a variety of stresses and are expressed in many tissue types, in agreement with the expression studies presented in this work.

A similar analysis of putative *Arabidopis* homologues found in the At3g19000 and At3g19010 genes identified motifs shared with both *SAN1* genes. Although the use of *Arabidopis* as a model system would be very helpful, the significance of such conserved sequences in putative *Arabidopsis* homologues is uncertain.

A scan of the *SAN1A* gene's upstream region identified many potential light responsive regulatory element motifs (see Section 4.2.6), including three -10PEHVPSBD motifs, one BOXIINTPATPB motif, one CCA1ATLHCB21 motif, two CIRCADIANLELHC motifs, twelve GATA BOX motifs, eleven GT1CONSENSUS motifs, one IBOX, three IBOXCORE sequences, three INRNTPSADB motifs, and one TBOXATGAPB sequence (Chan *et al.*, 2001; Gidoni *et al.*, 1989; Giuliano *et al.*, 1988; Kapoor and Sugiura, 1999; Nakamura *et al.*, 2002; Piechulla *et al.*, 1998; Terzaghi and Cashmore, 1995; Thum *et al.*, 2001; Wang *et al.*, 1997). Although there is no evidence that *SAN1B* is light regulated, it also contains light responsive motifs and thus both genes may be diurnally regulated.

4.2.8 A Model Describing Potential Roles for the *SAN1* Gene Products

Although the functions of *SAN1A* and *SAN1B* in soybean metabolism are unknown, a model based upon the data presented in this thesis will be described as model building can lead to hypotheses that can later be subject to experimental testing. The model is based upon the following evidence:

i. *SAN1A* and *SAN1B* have similar, but not identical, sequences;

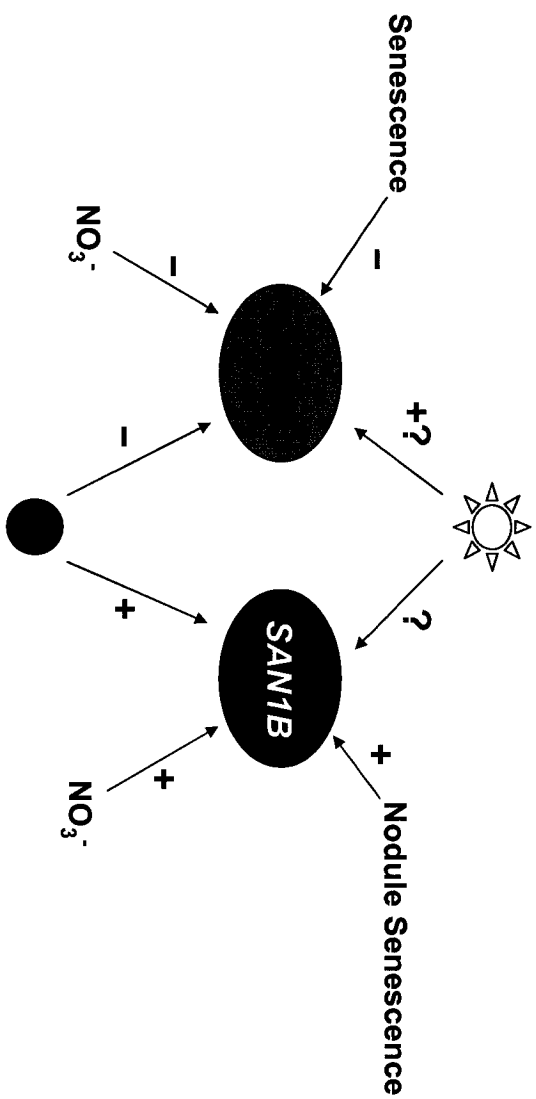
ii. the *SAN1* genes encode unknown IPNS-related dioxygenases most closely related to the GA20oxs (Hedden and Phillips, 2000). These enzymes catalyze the synthesis of active gibberellins, which play many roles in plant growth and development and in the mediation between environmental signals, such as light and photoperiod, and induced physiological responses such as flowering and stem elongation;

iii. *SAN1A* and *SAN1B* have divergent roles, are differentially regulated, and may be regulated by different signals.

The model presented in Figure 35 summarizes the potential forms of regulation, the genes' responses, and potential functions of the *SAN1A* and *SAN1B* gene products. *SAN1A* and *SAN1B* expression in nodules are both affected by dark and nitrate. Nitrate and dark increased the expression of *SAN1B*, but decreased the expression of *SAN1A*. During natural development, *SAN1A* was expressed in actively nitrogen fixing nodules, but its expression declined during senescence. In addition, it appeared to be expressed at lower levels in senescing leaves than in non-senescent leaves. Therefore, *SAN1A* may be generally downregulated by plant senescence signals. *SAN1B* appeared to be upregulated during natural nodule senescence, but not in senescing leaves. Since senescence appears to affect the expression of *SAN1A* and *SAN1B* in opposing manners, it is possible that *SAN1B* may replace *SAN1A* during senescence. *SAN1A* is likely to be upregulated in the

Figure 35: A Working Model Describing Potential Roles for the *SAN1* Gene Products

The potential mechanisms of regulation of the *SAN1A* and *SAN1B* genes are outlined. Arrows indicate regulatory input. The sun represents light and the black circle represents darkness. +, positive regulation; -, negative regulation; ?, unknown regulation. *SAN1A* and *SAN1B* are regulated by dark and nitrate in nodules and may also be regulated by senescence signals, and light. The functions of the *SAN1A* and *SAN1B* gene products are unknown, but they are most closely related to plant dioxygenases involved in flavonoid, anthocyanin, and hormone biosynthesis and degradation.


Senescence
−
NO_3^-
−
−
+?
+
?
SAN1B
+
NO_3^-
+
Nodule Senescence

light and downregulated in the dark; no information is available on the light responsiveness of *SAN1B*, although like *SAN1A*, *SAN1B* contains motifs corresponding to light responsive regulatory elements.

If the *SAN1* genes encode GA20oxs and not novel dioxygenases, then this model should also be applicable to these enzymes. GA20oxs exhibit light-dark cycling, such as that observed in *SAN1A* expression (Jackson *et al.*, 2000). Like GA20oxs, the *SAN1* genes are expressed in all tissues examined (Hedden and Phillips, 2000). However, gibberellins, and therefore GA20oxs expression, are at their highest levels in seeds; however, it is not certain whether this is true if the *SAN1* genes. In addition, recent studies have indicated a role for gibberellins in the process of nodulation. For example, Ferguson *et al.* (2005) showed that gibberellins are required for and have a direct role in nodule formation in pea. Lievens *et al.* (2005) showed that a GA20ox gene is uparegulated during nodulation of *Sesbania rostrata*. Therefore, this model could apply to GA20oxs in several aspects, although not in all, suggesting that the *SAN1* genes are more likely to encode novel dioxygenases.

Based on this model, a number of future experiments suggest themselves. qPCR experiments could be carried out to determine whether *SAN1A* is really light regulated as well as whether *SAN1B* is also regulated by light. In addition, whether *SAN1A* and *SAN1B* are downregulated and upregulated, respectively, during natural and dark-induced senescence in other tissues could be carried out using relative RT-PCR. To assess potential roles of the *SAN1* gene products in soybean metabolism, knockouts, knockdowns, and overexpression experiments could be carried out. Any alterations in morphological and physiological features of the transformants could be observed. For example, if the *SAN1* genes encode gibberellin 20-oxidases, a knockdown would be

expected to result in dwarf plants. Another approach to determination of the possible roles of the *SAN1* genes in plant metabolism would be the expression of the genes in *E. coli* followed by assays for the activity of known plant dioxygenases. If no known dioxygenase activities were observed, the *SAN1* genes would be expected to encode novel dioxygenases. In this case, comparison of the chemical profiles of nodule extracts in plants with altered levels of *SAN1* expression to wild type plants by methods such as HPLC or GC might yield clues as to the functions of their gene products.

REFERENCES

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and R.A. Yamaguchi-Shinozaki (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signalling. The Plant Cell **15**: 63-78.

Aguilar, J.M., Ashby, A.M., Richards, J.M., Loake, G.J., Watson, M.D., and C.H. Shaw (1988) Chemotaxis of *Rhizobium leguminosarum* biovar *phaseoli* Towards Flavonoid Inducers of the Symbiotic Nodulation Genes. Journal of General Microbiology **134**: 2741-2746.

Alazard, D. (1985) Stem and Root Nodulation in *Aeschynomene* spp. Applied and Environmental Microbiology **50**: 732-734.

Albrecht, C., Geurts, R., and T. Bisseling (1999) Legume Nodulation and Mycorrhizae Formation: Two Extremes of Host Specificity Meet. The EMBO Journal **18**: 281-288.

Alesandrini, F., Mathis, R., Van de Sype, G., Herouart, D., and A. Puppo (2003) Possible Roles for a Cysteine Protease and Hydrogen Peroxide in Soybean Nodule Development and Senescence. New Phytologist **158**: 131-138.

Allison L.A., Kiss, G.B., Bauer, P., Poiret, M., Pierre, M., Savoure, A., Kondorosi, E., and A. Kondorosi (1993) Identification of Two Alfalfa Early Nodulin Genes with Homology to Members of the Pea *ENOD12* Gene Family. Plant Molecular Biology **21**: 375-380.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and D.J. Lipman (1997) Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. Nucleic Acids Research **25**: 3389-3402.

Ane, J.-M., Kiss, G.B., Riely, B.K., Penmetsa, R.V., Oldroyd, G.E.D., Ayax, C., Levy, J., Debelle, F., Baek, J.-M., Kalo, P., Rosenberg, C., Roe, B.A., Long, S.R., Denarie, J., and D.R. Cook (2004) *Medicago truncatula DMI1* Required for Bacterial and Fungal Symbioses in Legumes. Science **303**: 1364-1367.

Appelbaum, E.R., Thompson, D.V., Idler, K., and N. Chartrain (1988) *Rhizobium japonicum* USDA 191 Has Two *NODD* Genes that Differ in Primary Structure and Function. Journal of Bacteriology **170**: 12-20.

Apte, S.S., Mattei, M.G., Seldin, M.F., and B.R. Olsen (1995) The Highly Conserved *Defender Against the Death 1* (*DAD1*) Gene Maps to Human Chromosome 14q11-q12 and Mouse Chromosome 14 and has Plant and Nematode Homologs. FEBS Letters **363**: 304-306.

Ardourel, M., Demont, N., Debelle, F., Maillet, F., DeBilly, F., Prome, J.-C., Denarie, J., and G. Truchet (1994) *Rhizobium meliloti* Lipooligosaccharide Nodulation Factors: Different Structural Requirements for Bacterial Entry into Target Root Hair Cells and Induction of Plant Symbiotic Developmental Responses. The Plant Cell **6**: 1357-1374.

Asai, T., Stone, J.M., Heard, J.E., Kovtun, Y., Yorgey, P., Sheen, J., and F.M. Ausubel (2000) Fumonisin B1-Induced Cell Death in *Arabidopsis* Protoplasts Requires Jasmonate-, Ethylene-, and Salicylate-Dependent Signaling Pathways. The Plant Cell **12**: 1823-1836.

Aspelmeier, J. (2002) Multiple Comparison Handout. (http://www.radford.edu/~jaspelme/611/Multiple%20Comparisons%20Handout.doc)

Assuncao Guimaraes, C. and R. Linden (2004) Programmed Cell Death: Apoptosis and Alternative Deathstyles. European Journal of Biochemistry **271**: 1638-1650.

Balk, J. and C.J. Leaver (2001) The *pet1-cms* Mitochondrial Mutation in Sunflower is Associated with Premature Programmed Cell Death and Cytochrome c Release. The Plant Cell **13**: 1803-1818.

Balk, J., Leaver, C.J., and P.F. McCabe (1999) Translocation of Cytochrome c from the Mitochondria to the Cytosol Occurs During Heat-Induced Programmed Cell Death in Cucumber Plants. FEBS Letters **463**: 151-154.

Barbour, W.M., Hatterman, D.R., and G. Stacey (1991) Chemotaxis of *Bradyrhizobium japonicum* to Soybean Exudates. Applied and Environmental Microbiology **57**: 2635-2639.

Barnett, M.J., Toman, C.J., Fisher, R.F., and S.R. Long (2004) A Dual-Genome Symbiosis Chip for Coordinate Study of Signal Exchange and Development in a Prokaryote-Host Interaction. Proceedings of the National Academy of Sciences USA **101**: 16636-16641.

Barry, C.S., Blume, B., Bouzayen, M., Cooper, W., Hamilton, A.J., and D. Grierson (1996) Differential Expression of the 1-Aminocyclopropane-1-Carboxylate Oxidase Gene Family in Tomato. The Plant Journal **9**: 525-535.

Becana, M. and J.I. Sprent (1987) Nitrogen Fixation and Nitrate Reduction in the Root Nodules of Legumes. Physiologia Plantarum **70**: 757-765.

Bec-Ferte, M.-P., Krishnan, H.B., Savagnac, A., Pueppke, S.G., and J.-C. Prome (1996) *Rhizobium fredii* Synthesizes an Array of Lipooligosaccharides, including a Novel Compound with Glucose Inserted into the Backbone of the Molecule. FEBS Letters **393**: 273-279.

Beers, E. (1997) Programmed Cell Death during Plant Growth and Development. Cell Death and Differentiation **4**: 649-661.

Beers, E., Woffenden, B.J., and C. Zhao (2000) Plant Proteolytic Enzymes: Possible Roles during Programmed Cell Death. Plant Molecular Biology **44**: 399-415.

Ben Amor, B., Shaw, S.L., Oldroyd, G.E.D., Maillet, F., Penmetsa, R.V., Cook, D., Long, S.R., Denarie, J., and C. Gough (2003) The *NFP* Locus of *Medicago truncatula* Controls an Early Step of Nod Factor Signal Transduction Upstream of a Rapid Calcium Flux and Root Hair Deformation. The Plant Journal **34**: 495-506.

Benson, D.R. and W.B. Silvester (1993) Biology of *Frankia* Strains, Actinomycete Symbionts of Actinorhizal Plants. Microbiological Reviews **57**: 293-319.

Bergersen, F.J. and D.J. Goodchild (1973) Cellular Location and Concentration of Leghaemoglobin in Soybean Root Nodules. Australian Journal of Biological Sciences **26**: 741-756.

Berghammer, H. and B. Auer (1993) "Easypreps": Fast and Easy Plasmid Minipreparation for Analysis of Recombinant Clones in *E. coli*. Biotechniques **14**: 524-528.

Berlyn, G.P. and J.P. Miksche (1976) Botanical Microtechnique and Cytochemistry. Iowa State University Press, Ames, USA

Bestel-Corre, G., Dumas-Gaudot, E., Poinsot, V., Dieu, M., Dierick, J.F., Van, T.D., Remacle, J., Gianinazzi-Pearson, V., and S. Gianinazzi (2002) Proteome Analysis and Identification of Symbiosis-Related Proteins from *Medicago truncatula* Gaertn. by Two-Dimensional Electrophoresis and Mass Spectrometry. Electrophoresis **23**: 122-137.

Bladergroen, M.R. and H.P. Spaink (1998) Genes and Signal Molecules Involved in the Rhizobia-Leguminoseae Symbiosis. Current Opinion in Plant Biology **1**: 353-359.

Bolduc, N. and L.F. Brisson (2002) Antisense Down Regulation of *NtBI-1* in Tobacco BY-2 Cells Induces Accelerated Cell Death upon Carbon Starvation. FEBS Letters **532**: 111-114.

Bolduc, N., Ouellet, M., Pitre, F., and L.F. Brisson (2003) Molecular Characterization of Two Plant *BI-1* Homologues which Suppress Bax-Induced Apoptosis in Human 293 Cells. Planta **216**: 377-386.

Bolle, C. (2004) The Role of GRAS Proteins in Plant Signal Transduction and Development. Planta **218**: 683-692.

Brewin, N.J. (1991) Development of the Legume Root Nodule. Annual Review of Cell Biology **7**: 191-226.

Britsch, L., Dedio, J., Saedler, H., and G. Forkmann (1993) Molecular Characterization of Flavanone 3β-Hydroxylases. Consensus Sequence, Comparison with Related Enzymes, and the Role of Conserved Histidine Residues. The European Journal of Biochemistry **217**: 745-754.

Brown, J.W.S., Simpson, C.G., Thow, G., Clark, G.P., Jennings, S.N., Medina-Escobar, N., Haupt, S., Chapman, S.C., and K.J. Oparka (2002) Splicing Signals and Factors in Plant Intron Removal. Biochemical Society Transactions **30**: 146-149.

Buchanan-Wollaston, V. and C. Ainsworth (1997) Leaf Senescence in *Brassica napus*: Cloning of Senescence-Related Genes by Subtractive Hybridisation. Plant Molecular Biology **33**: 821-834.

Buresh, R.J. and P.J.M. Cooper (1999) The Science and Practice of Short-Term Improved Fallows: Symposium Synthesis and Recommendations. Agroforestry Systems **47**: 345-356.

Caetano-Anolles, G. and P.M. Gresshoff (1991) Plant Genetic Control of Nodulation. Annual Review of Microbiology **45**: 345-382.

Callard, D., Axelos, M., and L. Mazzolini (1996) Novel Molecular Markers for Late Phases of the Growth Cycle of *Arabidopsis thaliana* Cell-Suspension Cultures are Expressed during Organ Senescence. Plant Physiology **112**: 705-715.

Campbell, R. and M.C. Drew (1983) Electron Microscopy of Gas Space (Aerenchyma) Formation in Adventitious Roots of *Zea mays* L. Subjected to Oxygen Shortage. Planta **157**: 350-357.

Campalans, A., Kondorosi, A., and M. Crespi (2004) *Enod40*, a Short Open Reading Frame-Containing mRNA, Induces Cytoplasmic Localization of a Nuclear RNA Binding Protein in *Medicago truncatula*. The Plant Cell **16**: 1047-1059.

Cardenas, L., Feijo, J.A., Kunkel, J.G., Sanchez, F., Holdaway-Clarke, T., Hepler, P.K., and C. Quinto (1999) *Rhizobium* Nod Factors Induce Increases in Intracellular Free Calcium Influxes in Bean Root Hairs. The Plant Journal **19**: 347-352.

Carimi, F., Zottini, M., Formentin, E., Terzi, M., and F. Schiavo (2003) Cytokinins: New Apoptotic Inducers in Plants. Planta **216**: 413-421.

Carlson, R.W., Sanjuan, J., Bhat, U.R., Glushka, J., Spaink, H.P., Wijfjes, A.H.M., Van Brussel, A.A.N., Stokkermans, T.J.W., Peters, N.K., and G. Stacey (1993) The Structures and Biological Activities of the Lipo-Oligosaccharide Nodulation Signals Produced by Type I and Type II Strains of *Bradyrhizobium japonicum*. Journal of Biological Chemistry **268**: 18372-18381.

Carroll, B.J. and A. Mathews (1990) Nitrate Inhibition of Nodulation in Legumes. *In* P.M. Gresshoff, ed. Molecular Biology of Symbiotic Nitrogen Fixation. CRC Press, Boca Raton, USA, pp.159-180.

Catoira, R., Galera, C., de Billy, F. Penmetsa, R.V., Journet, E.-P., Maillet, F., Rosenberg, F., Cook, D., Gough, C. and J. Denarie (2000) Four Genes of *Medicago truncatula* Controlling Components of a Nod Factor Transduction Pathway. The Plant Cell **12**: 1647-1666.

Catoira, R., Timmers, A.C.J., Maillet, F., Galera, C., Penmetsa, R.V., Cook, D., Denarie, J. and C. Gough (2001) The *HCL* Gene of *Medicago truncatula* Controls *Rhizobium*-Induced Root Hair Curling. Development **128**: 1507-1518.

Chan, C.C.Y. (1995) The Isolation and Characterization of a Senescence-Associated Nodulin cDNA. M.Sc. Thesis, Department of Biology, University of Ottawa, Ottawa, ON

Chan, C.S., Guo, L., and M.C. Shih (2001) Promoter Analysis of the Nuclear Gene Encoding the Chloroplast Glyceraldehyde-3-Phosphate Dehydrogenase B Subunit in *Arabidopsis thaliana*. Plant Molecular Biology **46**: 131-141.

Charon, C., Johansson, C., Kondorosi, E., Kondorosi, A., and M. Crespi (1997) *ENOD40* Induces Dedifferentiation and Division of Root Cortical Cells in Legumes. Proceedings of the National Academy of Sciences USA **94**: 8901-8906.

Charon, C., Sousa, C., Crespi, M., and A. Kondorosi (1999) Alteration of *ENOD40* Expression Modifies *Medicago truncatula* Root Nodule Development Induced by *Sinorhizobium meliloti*. The Plant Cell **11**: 1953-1965.

Chen, F. and M.R. Foolad (1997) Molecular Organization of a Gene in Barley which Encodes a Protein Similar to Aspartic Protease and its Specific Expression in Nucellar Cells during Degeneration. Plant Molecular Biology **35**: 821-831.

Chen, J.M., Rawlings, N.D., Stevens, R.A., and A.J. Barrett (1998) Identification of the Active Site of Legumain Links it to Caspases, Clostripain and Gingipains in a New Clan of Cysteine Endopeptidases. FEBS Letters **441**: 361-365.

Chichkova, N.V., Kim, S.H., Titova, E.S., Kalkum, M., Morozov, V.S., Rubtsov, Y.P., Kalinina, N.O., Taliansky, M.E., and A.B. Vartapetian (2004) A Plant Caspase-Like Protease Activated during the Hypersensitive Response. The Plant Cell **16**: 157-171.

Chomczynski, P. and K. Mackey (1995) Modification of the Tri Reagent Procedure for Isolation of RNA from Polysaccharide- and Proteoglycan-Rich Sources. Biotechniques **19**: 924-945.

Chomczynski, P. and N. Sacchi (1987) Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction. Analytical Biochemistry **162**: 156-159.

Clarke, A., Desikan, R., Hurst, R.D., Hancock, J.T., and S.J. Neill (2000) NO Way Back: Nitric Oxide and Programmed Cell Death in *Arabidopsis thaliana* Suspension Cultures. The Plant Journal **24**: 667-677.

Coffeen, W.C. and T.J. Wolpert (2004) Purification and Characterization of Serine Proteases that Exhibit Caspase-Like Activity and are Associated with Programmed Cell Death in *Avena sativa*. The Plant Cell **16**: 857-873.

Cohn, J., Day, R.B., and G. Stacey (1998) Legume Nodule Organogenesis. Trends in Plant Science **3**: 105-110.

Colebatch, G., Kloska, S., Trevaskis, B., Freund, S., Altmann, T., and M.K. Udvardi (2002) Novel Aspects of Symbiotic Nitrogen Fixation Uncovered by Transcript Profiling with cDNA Arrays. Molecular Plant-Microbe Interactions **15**: 411-420.

Compaan, B., Ruttink, T., Albrecht, C., Meeley, R., Bisseling, T., and H. Franssen (2003) Identification and Characterization of a *Zea mays* Line Carrying a Transposon-Tagged *ENOD40*. Biochimica et Biophysica Acta **1629**: 84-91.

Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, E., and K. VandenBosch (1995) Transient Induction of a Peroxidase Gene in *Medicago truncatula* Precedes Infection by *Rhizobium meliloti*. The Plant Cell **7**: 43-55.

Cook, D.R. (1999) *Medicago truncatula* – A Model in the Making! Current Opinion in Plant Biology **2**: 301-304.

Cooper, J.B. and S.R. Long (1994) Morphogenetic Rescue of *Rhizobium meliloti* Nodulation Mutants by *Trans*-Zeatin Secretion. The Plant Cell **6**: 215-225.

Coventry D.R. and M.J. Dilworth (1976) Synthesis and Turnover of Leghaemoglobin in Lupin Root Nodules. Biochimica et Biophysica Acta **447**: 1-10.

Crawford, N.M. (1995) Nitrate: Nutrient and Signal for Plant Growth. The Plant Cell **7**: 859-868.

Cuervo, A.M. (2004) Autophagy: In Sickness and in Health. Trends in Cell Biology **14**: 70-77.

Curtis, M.J. and T.J. Wolpert (2002) The Oat Mitochondrial Permeability Transition and its Implication in Victorin Binding and Induced Cell Death. The Plant Journal **29**: 295-312.

Danon, A., Delorme, V., Mailhac, N., and P. Gallois (2000) Plant Programmed Cell Death: a Common Way to Die. Plant Physiology and Biochemistry **38**: 647-655.

Danon, A. and P. Gallois (1998) UV-C Radiation Induces Apoptotic-Like Changes in *Arabidopsis thaliana*. FEBS Letters **437**: 131-136.

De Carolis, E. and V. De Luca (1994) 2-Oxoglutarate-Dependent Dioxygenase and Related Enzymes: Biochemical Characterization. Phytochemistry **36**: 1093-1107.

De Carolis, E. and V. De Luca (1993) Purification, Characterization, and Kinetic Analysis of a 2-Oxoglutarate-Dependent Dioxygenase involved in Vindoline Biosynthesis from *Catharanthus roseus*. Journal of Biological Chemistry **268**: 5504-5511.

De Jong, A.J., Hoeberichts, F.A., Yakimova, E.T., Maximova, E., and E.J. Woltering (2000) Chemical-Induced Apoptotic Cell Death in Tomato Cells: Involvement of Caspase-Like Proteases. Planta **211**: 656-662.

Del Pozo, O. and E. Lam (2003) Expression of the Baculovirus p35 Protein in Tobacco Affects Cell Death Progression and Compromises N Gene-Mediated Disease Resistance Response to Tobacco Mosaic Virus. Molecular Plant-Microbe Interactions **16**: 485-494.

Del Pozo, O. and E. Lam (1998) Caspases and Programmed Cell Death in the Hypersensitive Response of Plants to Pathogens. Current Biology **8**: 1129-1132.

Demont, N., Debelle, F., Aurelle, H., Denarie, J., and J.-C. Prome (1993) Role of the *Rhizobium meliloti NODF* and *NODE* Genes in the Biosynthesis of Lipooligosaccharidic Nodulation Factors. Journal of Biological Chemistry **268**: 20134-20142.

De Ruijter, N.C.A., Rook, M.B., Bisseling, T., and A.M.C. Emons (1998) Lipochito-Oligosaccharides Reinitiate Root Hair Tip Growth in *Vicia sativa* with High $[Ca^{2+}]_c$ and Spectrin-Like Antigen at the Tip. The Plant Journal **13**: 341-350.

D'Haeze, W. and M. Holsters (2002) Nod Factor Structures, Responses, and Perception During Initiation of Nodule Development. Glycobiology **12**: 79R-105R.

Dion, M., Chamberland, H., St.-Michel, C., Plante, M., Darveau, A., Lafontaine, J.G., and L.F. Brisson (1997) Detection of a Homologue of BCL-2 in Plant Cells. Biochemistry and Cell Biology **75**: 457-461.

Dong, Y.-H., Zhan, X.-C., Kvarnheden, A., Atkinson, R.G., Morris, B.A., and R.C. Gardner (1998) Expression of a cDNA from Apple Encoding a Homologue of *DAD1*, an Inhibitor of Programmed Cell Death. Plant Science **139**: 165-174.

Downie, J.A. (2005) Legume Haemoglobins: Symbiotic Nitrogen Fixation Needs Bloody Nodules. Current Biology **15**: R196-R198.

Downie, J.A. and S.A. Walker (1999) Plant Responses to Nodulation Factors. Current Opinion in Plant Biology **2**: 483-487.

Drake, R., John, I., Farrell, A., Cooper, W., Schuch, W., and D. Grierson (1996) Isolation and Analysis of cDNAs Encoding Tomato Cysteine Proteases Expressed during Leaf Senescence. Plant Molecular Biology **30**: 755-767.

Dubos, C. and C. Plomion (2003) Identification of Water-Deficit Responsive Genes in Maritime Pine (*Pinus pinaster* Ait.) Roots. Plant Molecular Biology **51**: 249-262.

Duflos, A., Kruczynski, A., and J.M. Barret (2002) Novel Aspects of Natural and Modified *Vinca* Alkaloids. Current Medicinal Chemistry, Anti-Cancer Agents **2**: 55-70.

Dyson, T. (1999) World Food Trends and Prospects to 2025. Proceedings of the National Academy of Sciences USA **96**: 5929-5936.

Ehrhardt, D.W., Atkinson, E.M., and S.R. Long (1992) Depolarization of Alfalfa Root Hair Membrane Potential by *Rhizobium meliloti* Nod Factors. Science **256**: 998-1000.

Ehrhardt, D.W., Wais, R., and S.R. Long (1996) Calcium Spiking in Plant Root Hairs Responding to Rhizobium Nodulation Signals. Cell **85**: 673-681.

Endre, G., Kereszt, A., Kevel, Z., Mihacea, S., Kalo, P., and G.B. Kiss (2002) A Receptor Kinase Gene Regulating Symbiotic Nodule Development. Nature **417**: 962-966.

Epple, P., Mack, A.A., Morris, V.R., and J.L. Dangl (2003) Antagonistic Control of Oxidative Stress-Induced Cell Death in *Arabidopsis* by Two Related, Plant-Specific Zinc Finger Proteins. Proceedings of the National Academy of Sciences USA **100**: 6831-6836.

Esseling, J.J. and A.M.C. Emons (2004) Dissection of Nod Factor Signalling in Legumes: Cell Biology, Mutants, and Pharmacological Approaches. Journal of Microscopy **214**: 104-113.

Esseling, J.J., Lhuissier, E.G.P., and A.M.C. Emons (2003) Nod Factor-Induced Root Hair Curling: Continuous Polar Growth Towards the Point of Nod Factor Application. Plant Physiology **132**: 1982-1988.

Estabrook, E.M. and C. Sengupta-Gopalan (1991) Differential Expression of Phenylalanine Ammonia-Lyase and Chalcone Synthase During Soybean Nodule Development. The Plant Cell **3**: 299-308.

Etzler, M.E., Kalsi, G., Ewing, N.N., Roberts, N.J., Day, R.B., and J.B. Murphy (1999) A Nod Factor Binding Lectin with Apyrase Activity from Legume Roots. Proceedings of the National Academy of Sciences USA **96**: 5856-5861.

Evans, P.J., Gallesi, D., Mathieu, C., Hernandez, M.J., DeFelipe, M., Halliwell, B., and A. Puppo (1999) Oxidative Stress Occurs During Soybean Nodule Senescence. Planta **208**: 73-79.

Fang, Y. and A.M. Hirsch (1998) Studying Early Nodulin Gene *ENOD40* Expression and Induction by Nodulation and Cytokinin in Transgenic Alfalfa. Plant Physiology **116**: 53-68.

Fearn, J.C. and T.A. LaRue (1991) Ethylene Inhibitors Restore Nodulation to *sym5* Mutants of *Pisum sativum* L. cv. Sparkle. Plant Physiology **96**: 239-244.

Felle, H.H., Kondorosi, E., Kondorosi, A., and M. Schultze (1999a) Elevation of the Cytosolic Free [Ca^{2+}] is Indispensable for the Transduction of the Nod Factor Signal in Alfalfa. Plant Physiology **121**: 273-279.

Felle, H.H., Kondorosi, E., Kondorosi, A., and M. Schultze (1999b) Nod Factors Modulate the Concentration of Cytosolic Free Calcium Differentially in Growing and Non-Growing Root Hairs of *Medicago sativa* L. Planta **209**: 207-212.

Felle, H.H., Kondorosi, E., Kondorosi, A., and M. Schultze (1998) The Role of Ion Fluxes in Nod Factor Signalling in *Medicago sativa*. The Plant Journal **13**: 455-463.

Felle, H.H., Kondorosi, E., Kondorosi, A., and M. Schultze (1996) Rapid Alkanilization in Alfalfa Root Hairs in Response to Rhizobial Lipochitooligosaccharide Signals. The Plant Journal **10**: 295-301.

Felle, H.H., Kondorosi, E., Kondorosi, A., and M. Schultze (1995) Nod Signal-Induced Plasma Membrane Potential Changes in Alfalfa Root Hairs are Differentially Sensitive to Structural Modifications of the Lipochitooligosaccharide. The Plant Journal **7**: 939-947.

Felsenstein, J. (1989) PHYLIP – Phylogeny Inference Package (Version 3.2) Cladistics **5**: 164-166.

Ferguson, B.J., Ross, J.J., and J.B. Reid (2005) Nodulation Phenotypes of Gibberellin and Brassinosteroid Mutants of Pea. Plant Physiology **138**: 2396-2405.

Ferreira, F.J. and J.J. Kieber (2005) Cytokinin Signaling. Current Opinion in Plant Biology **8**: 518-525.

Filonova, L.H., Von Arnold, S., Daniel, G., and P.V. Bozhkov (2002) Programmed Cell Death Eliminates All But One Embryo in a Polyembryonic Plant Seed. Cell Death and Differentiation **9**: 1057-1062.

Fisher, R.F. and S.R. Long (1992) *Rhizobium*-Plant Signal Exchange. Nature **357**: 655-660.

Fiume, E., Christou, P., Giani, S., and D. Breviario (2004) Introns are Key Regulatory Elements of Rice Tubulin Expression. Planta **218**: 693-703.

Franklin-Tong, V.E., Drobak, B.K., Allan, A.C., Watkins, P.A.C., and A.J. Trewavas (1996) Growth of Pollen Tubes of *Papaver rhoeas* is Regulated by a Slow-Moving Calcium Wave Propagated by Inositol 1,4,5-Trisphosphate. The Plant Cell **8**: 1305-1321.

Frink, C.R., Waggoner, P.E., and J.H. Ausubel (1999) Nitrogen Fertilizer: Retrospect and Prospect. Proceedings of the National Academy of Sciences USA **96**: 1175-1180.

Fu, D., Szucs, P., Yan, L., Helguera, M., Skinner, J.S., Von Zitzewitz, J., Hayes, P.M., and J. Dubcovsky (2005) Large Deletions within the First Intron in *VRN-1* are Associated with Spring Growth Habit in Barley and Wheat. Molecular Genetics and Genomics **273**: 54-65.

Fujikake, H., Yamazaki, A., Ohtake, N., Sueyoshi, K., Matsuhashi, S., Ito, T., Mizuniwa, C., Kume, T., Hashimoto, S., Ishioka, N.-S., Watanabe, S., Osa, A., Sekine, T., Uchida, H., Tsuji, A., and T. Ohyama (2003) Quick and Reversible Inhibition of Soybean Root Nodule Growth by Nitrate Involves a Decrease in Sucrose Supply to Nodules. Journal of Experimental Botany **54**: 1379-1388.

Fukuda, H. (2000) Programmed Cell Death of Tracheary Elements as a Paradigm in Plants. Plant Molecular Biology **44**: 245-253.

Fukuda, H. (1997) Tracheary Element Differentiation. The Plant Cell **9**: 1147-1156.

Gallois, P., Makishima, T., Hecht, V., Despres, B., Laudie, M., Nishimoto, T., and R. Cooke (1997) An *Arabidopsis thaliana* cDNA Complementing a Hamster Apoptosis Suppressor Mutant. The Plant Journal **11**: 1325-1331.

Galloway, J.N., Schlesinger, W.H., Levy II, H., Michaels, A., and J.L. Schnoor (1995) Nitrogen Fixation: Anthropogenic Enhancement-Environmental Response. Global Biogeochemical Cycles **9**: 235-252.

Gallusci, P., Dedieu, A., Journet, E.P., Huguet, T., and D.G. Barker (1991) Synchronous Expression of Leghaemoglobin Genes in *Medicago truncatula* during Nitrogen-Fixing Root Nodule Development and Response to Exogenously Supplied Nitrate. Plant Molecular Biology **17**: 335-349.

Gebhardt, Y., Witte, S., Forkmann, G., Lukacin, R., Matern, U., and S. Martens (2005) Molecular Evolution of Flavonoid Dioxygenases in the Family Apiaceae. Phytochemistry **66**: 1273-1284.

Gepstein, S. (2004) Leaf Senescence – Not Just a "Wear and Tear" Phenomenon. Genome Biology **5**: 212.1-212.3.

Gepstein, S., Sabehi, G., Carp, M.J., Hajouj, T., Nesher, M.F., Yariv, I., Dor, C., and M. Bassani (2003) Large-Scale Identification of Leaf Senescence-Associated Genes. The Plant Journal **36**: 629-642.

Geurts, R. and T. Bisseling (2002) *Rhizobium* Nod Factor Perception and Signalling. The Plant Cell **Supplement 2002**: S239-S249.

Geurts, R., Heidstra, R., Hadri, A.E., Downie, J.A., Franssen, H., Van Kammen, A., and T. Bisseling (1997) SYM2 of Pea Is Involved in a Nodulation Factor-Perception Mechanism That Controls the Infection Process in the Epidermis. Plant Physiology **115**: 351-359.

Ghanem, N., Jarinova, O., Amores, A., Long, Q., Hatch, G., Park, B.K., Rubenstein, J.L.R., and M. Ekker (2003) Regulatory Roles of Conserved Intergenic Domains in Vertebrate *DLX* Bigene Clusters. Genome Research **13**: 533-543.

Gidoni, D., Brosio, P., Bond-Nutter, D., Bedbrook, J., and P. Dunsmuir (1989) Novel Cis-Acting Elements in Petunia *CAB* Gene Promoters. Molecular and General Genetics **215**: 337-344.

Girard, G., Roussis, A., Gultyaev, A.P., Pleij, C.W., and H.P. Spaink (2003) Structural Motifs in the RNA Encoded by the Early Nodulation Gene *ENOD40* of Soybean. Nucleic Acids Research **31**: 5003-5015.

Giuliano, G., Pichersky, E., Malik, V.S., Timko, M.P., Scolnik, P.A., and A.R. Cashmore (1988) An Evolutionarily Conserved Protein Binding Sequence Upstream of a Plant Light-Regulated Gene. Proceedings of the National Academy of Sciences **85**: 7089-7093.

Goethals, K., Van den Eeede, G., Van Montagu, M., and M. Holsters (1990) Identification and Characterization of a Functional *NODD* Gene in *Azorhizobium caulinodans* ORS571. Journal of Bacteriology **172**: 2658-2666.

Gong, Z., Yamazaki, M., Sugiyama, M., Tanaka, Y., and K. Saito (1997) Cloning and Molecular Analysis of Structural Genes Involved in Anthocyanin Biosynthesis and Expressed in a Forma-Specific Manner in *Perilla frutescans*. Plant Molecular Biology **35**: 915-927.

Goormachtig, S., Lievens, S., Van de Velde, W., Van Montagu, M., and M. Holsters (1998) *SrCHI13*, a Novel Early Nodulin from *Sesbania rostrata*, is Related to Acidic Class III Chitinases. The Plant Cell **10**: 905-915.

Goormachtig, S., Van de Velde, W., Lievens, S., Verplancke, C., Herman, S., De Keyser, A., and M. Holsters (2001) *SrCHI24*, a Chitinase Homolog Lacking an Essential Glutamic Acid Residue for Hydrolytic Activity, is Induced During Nodule Development on *Sesbania rostrata*. Plant Physiology **127**: 78-89.

Gordon, A.J., Minchin, F.R., Skot, L., and C.L. James (1997) Stress-Induced Declines in Soybean N_2 Fixation Are Related to Nodule Sucrose Synthase Activity. Plant Physiology **114**: 937-946.

Gordon, A.J., Ougham, H.J. and C.L. James (1993) Changes in Levels of Gene Transcripts and their Corresponding Proteins in Nodules of Soybean Plants Subjected to Dark-Induced Stress. Journal of Experimental Botany **44**: 1453-1460.

Gordon, A.J., Skot, C.L.J., and F.R. Minchin (2002) Short-Term Metabolic Responses of Soybean Nodules to Nitrate. Journal of Experimental Botany **53**: 423-428.

Gottlob-McHugh, S.G. (1989) Investigation of Nodulin Genes Expressed during Soybean Root Nodule Development. Ph.D. Thesis, Department of Biology, University of Ottawa, Ottawa, ON

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., and P. Westhoff (2004) Cis-Regulatory Elements for Mesophyll-Specific Gene Expression in the C4 Plant *Flaveria trinervia*, the Promoter of the C4 Phosphoenolpyruvate Carboxylase Gene. The Plant Cell **16**: 1077-1090.

Graham, P.H. and C.P. Vance (2003) Legumes: Importance and Constraints to Greater Use. Plant Physiology **131**: 872-877.

Graur, D. and W.-H. Li (2000) Fundamentals of Molecular Evolution, 2nd ed. Sinauer Associates, Inc., Sunderland, USA, pp. 264, 309.

Greenberg, J.T. (1996) Programmed Cell Death: a Way of Life for Plants. Proceedings of the National Academy of Sciences USA **93**: 12094-12097.

Greenberg, J.T. and N. Yao (2004) The Role and Regulation of Programmed Cell Death in Plant-Pathogen Interactions. Cellular Microbiology **6**: 201-211.

Groten, K., Vanacker, H., Dutilleul, C., Bastian, F., Bernard, S., Carzaniga, R., and C.H. Foyer (2005) The Roles of Redox Processes in Pea Nodule Development and Senescence. Plant, Cell, and Environment **28**: 1293-1304.

Gulash, M., Ames, P., Larosiliere, R.C., and K. Bergman (1984) Rhizobia are Attracted to Localized Sites on Legume Roots. Applied and Environmental Microbiology **48**: 149-152.

Guo, G., Cai, Z., and S. Gan (2004) Transcriptome of *Arabidopsis* Leaf Senescence. Plant, Cell and Environment **27**: 521-549.

Guzzo, F., Portaluppi, P., Grisi, R., Barone, S., Zampieri, S., Franssen, H., and M. Levi (2005) Reduction of Cell Size Induced by *ENOD40* in *Arabidopsis thaliana*. Journal of Experimental Botany **56**: 507-513.

Hall, T.A. (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series **41**: 95-98.

Hanahan, D. (1983) Studies on Transformation of *Escherichia coli* with Plasmids. Journal of Molecular Biology **166**: 557-580.

Handberg, K. and J. Stougaard (1992) *Lotus japonicus*, an Autogamous, Diploid Legume Species for Classical and Molecular Genetics. The Plant Journal **2**: 487-496.

Hardy, R.W.F., Burns, R.C., Hebert, R.R., Holsten, R.D., and E.K. Jackson (1971) Biological Nitrogen Fixation: A Key to World Protein. International Biological Program Meeting on Nitrogen Fixation, Wageningen, Netherlands, Plant Soil Special Volume: 561-590.

Hardy, R.W.F., Holsten, R.D., Jackson, E.K., and R.C. Burns (1968) The Acetylene-Ethylene Assay for N_2 Fixation: Laboratory and Field Evaluation. Plant Physiology **43**: 1185–1207.

Hartmann, U., Sagasser, M., Mehrtans, F., Stracke, R., and B. Weisshaar (2005) Differential Combinatorial Interactions of Cis-Acting Elements Recognized by R2R3-MYB, BZIP, and BHLH Factors Control Light-Responsive and Tissue-Specific Activation of Phenylpropanoid Biosynthesis Genes. Plant Molecular Biology **57**: 155-171.

Hashimoto, T. and Y. Yamada (1987) Purification and Characterization of Hyoscyamine 6β-Hydroxylase from Root Cultures of *Hyoscyamus niger* L. European Journal of Biochemistry **164**: 277-285.

Hatsugai, N., Kuroyanagi, M., Yamada, K., Meshi, T., Tsuda, S., Kondo, M., Nishimura, M., and I. Hara-Nishimura (2004) A Plant Vacuolar Protease, VPE, Mediates Virus-Induced Hypersensitive Cell Death. Science **305**: 855-858.

Hattori, J., Gottlob-McHugh, S.G., and D.A. Johnson (1987) The Isolation of High-Molecular-Weight DNA from Plants. Analytical Biochemistry **165**: 70-74.

Hedden, P. (1999) Recent Advances in Gibberellin Synthesis. Journal of Experimental Botany **50**: 553-563.

Hedden, P. and A.L. Phillips (2000) Gibberellin Metabolism: New Insights Revealed by the Genes. Trends in Plant Science **5**: 523-530.

Heichel, G.H. (1987) Legume Nitrogen: Symbiotic Fixation and Recovery by Subsequent Crops. *In* Z.R. Helsel, ed. Energy in Plant Nutrition and Pest Control. Elsevier Science, Amsterdam, The Netherlands, pp. 63-80.

Heidstra, R. and T. Bisseling (1996) Nod Factor-Induced Host Responses and Mechanisms of Nod Factor Perception. New Phytologist **133**: 25-43.

Heidstra, R., Geurts, R., Franssen, H., Spaink, H.P., Van Kammen, A., and T. Bisseling (1994) Root Hair Deformation Activity of Nodulation Factors and their Fate on *Vicia sativa*. Plant Physiology **105**: 787-797.

Heidstra, R., Nilsen, G., Martinez-Abarca, F., Van Kammen, A., and T. Bisseling (1997a) Nod Factor-Induced Expression of Leghemoglobin to Study the Mechanism of NH_4NO_3 Inhibition on Root Hair Deformation. Molecular Plant-Microbe Interactions **10**: 215-220.

Heidstra, R., Yang, W.-C., Yalcin, Y., Peck, S., Emons, A.M., Van Kammen, A., and T. Bisseling (1997b) Ethylene Provides Positional Information on Cortical Cell Division but is not Involved in Nod Factor-Induced Root Hair Tip Growth in *Rhizobium*-Legume Interaction. Development **124**: 1781-1787.

Higo, K., Ugawa, Y., Iwamoto, M., and T. Korenaga (1999) Plant Cis-Acting Regulatory DNA Elements (PLACE) Database: 1999. Nucleic Acids Research **27**: 297-300.

Hirsch, A.M. (1992) Developmental Biology of Legume Nodulation. New Phytologist **122**: 211-237.

Hirsch, A.M., Bhuvaneswari, T.V., Torrey, J.G., and T. Bisseling (1989) Early Nodulin Genes are Induced in Alfalfa Root Outgrowths Elicited by Auxin Transport Inhibitors. Proceedings of the National Academy of Sciences USA **86**: 1244-1248.

Hirsch, A.M., Lum, M.R., and J.A. Downie (2001) What Makes the Rhizobia-Legume Symbiosis So Special? Plant Physiology **127**: 1484-1492.

Hoagland, D.R. and D.I. Arnon (1950) The Water Culture Method for Growing Plants without Soil, Circular 347. University of California Agricultural Experiment Station, Berkeley, USA, pp. 1-39.

Hoeberichts, F.A., ten Have, A., and E.J. Woltering (2003) A Tomato Metacaspase Gene is Upregulated during Programmed Cell Death in *Botrytis cinerea*-Infected Leaves. Planta **217**: 517-522.

Hoeberichts, F.A. and E.J. Woltering (2002) Multiple Mediators of Plant Programmed Cell Death: Interplay of Conserved Cell Death Mechanisms and Plant-Specific Regulators. BioEssays **25**: 47-57.

Hoeberichts, F.A. and E.J. Woltering (2001) Cloning and Analysis of a *Defender against Apoptotic cell Death* (*DAD1*) Homologue from Tomato. Journal of Plant Physiology **158**: 125-128.

Honma, M.A. and F.M. Ausubel (1987) *Rhizobium meliloti* has Three Functional Copies of the *NODD* Symbiotic Regulatory Gene. Proceedings of the National Academy of Sciences USA **84**: 8558-8562.

Hulo, N., Sigrist, C.J., Le Saux, V., Langendijk-Genevaux, P.S., Bardoli, L., Gattiker, A., De Castro, E., Bucher, P., and A. Bairoch (2004) Recent Improvements in the PROSITE Database. Nucleic Acids Research **32 (Database Issue)**: D134-D137.

Iannetta, P.P.M., James, E.K., Sprent, J.I., and F.R. Minchin (1995) Time-Course of Changes involved in the Operation of the Oxygen Diffusion Barrier in White Lupin Nodules. Journal of Experimental Botany **46**: 565-575.

Imaizumi-Anraku, H., Takeda, N., Charpentier, M., Perry, J., Miwa, H., Umehara, Y., Kouchi, H., Murakami, Y., Mulder, L., Vickers, K., Pike, J., Downie, J.A., Wang, T., Sato, S., Asamizu, E., Tabata, S., Yoshikawa, M., Murooka, Y., Wu, G.-J., Kawaguchi, M., Kawasaki, S., Parniske, M., and M. Hayashi (2005) Plastid Proteins Crucial for Symbiotic Fungal and Bacterial Entry into Roots. Nature **433**: 527-531.

Jabs, T., Dietrich, R.A., and J.L. Dangl (1996) Initiation of Runaway Cell Death in an *Arabidopsis* Mutant by Extracellular Superoxide. Science **273**: 1853-1856.

Jabs, T., Tschope, M., Colling, C., Hahlbrock, K., and D. Scheel (1997) Elicitor-Stimulated Ion Fluxes and O_2^- from the Oxidative Burst are Essential Components in Triggering Defense Gene Activation and Phytoalexin Synthesis in Parsley. Proceedings of the National Academy of Sciences USA **94**: 4800-4805.

Jackson, S.D., James, P.E., Carrera, E., Prat, S., and B. Thomas (2000) Regulation of Transcript Levels of a Potato Gibberellin 20-Oxidase Gene by Light and Phytochrome B. Plant Physiology **124**: 423-430.

Johnson, H.S. and D.J. Hume (1973) Comparisons of Nitrogen Fixation Estimates in Soybean by Nodule Weight, Leghemoglobin Content, and Acetylene Reduction. Canadian Journal of Microbiology **19**: 1165-1168.

Jones, A.M. and J.L. Dangl (1996) Logjam at the Styx: Programmed Cell Death in Plants. Trends in Plant Science **1**: 114-119.

Kalo, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J., Kiss, G.B., Downie, J.A., and G.E.D. Oldroyd (2005) Nodulation Signalling in Legumes Requires NSP2, a Member of the GRAS Family of Transcriptional Regulators. Science **308**: 1786-1789.

Kapoor, S. and M. Sugiura (1999) Identification of Two Essential Sequence Elements in the Nonconsensus Type II PatpB-290 Plastid Promoter by Using Plastid Transcription Extracts from Cultured Tobacco BY-2 Cells. The Plant Cell **11**: 1799-1810.

Kardailsky, I.V. and N.J. Brewin (1996) Expression of Cysteine Protease Genes in Pea Nodule Development and Senescence. Molecular Plant-Microbe Interactions **9**: 689-695.

Katsuhara, M. (1997) Apoptosis-Like Cell Death in Barley Roots under Salt Stress. Plant and Cell Phyiology **38**: 1091-1093.

Kawai, M., Pan, L., Reed, J.C., and H. Uchimiya (1999) Evolutionally Conserved Plant Homologue of the *BAX INHIBITOR-1* (*BI-1*) Gene Capable of Suppressing Bax-Induced Cell Death in Yeast(1). FEBS Letters **464**: 143-147.

Kennedy, A.R. (1995) The Evidence for Soybean Products as Cancer Preventative Agents. Journal of Nutrition **125**: S733-S743.

Kim, Y.S., Lee, J.H., Yoon, G.M., Cho,, H.S., Park, S.-W., Suh, M.C., Choi, D., Ha, H.J., Liu, J.R., and H.-S. Pai (2000) CHRK1, a Chitinase-Related Receptor-Like Kinase in Tobacco. Plant Physiology **123**: 905-915.

Kinoshita, T., Yamada, K., Hiraiwa, N., Kondo, M., Nishimura, M., and I. Hara-Nishimura (1999) Vacuolar Processing Enzyme is Up-Regulated in the Lytic Vacuoles of Vegetative Tissues during Senescence and Under Various Stressed Conditions. The Plant Journal **19**: 43-53.

Klucas, R.V. and D. Arp (1977) Physiological and Biochemical Studies on Senescing Tap Root Nodules of Soybeans. Canadian Journal of Microbiology **23**: 1426-1432.

Koltunow, A.M., Truettner, J., Cox, K.H., Wallroth, M., and R.B. Goldberg (1990) Different Temporal and Spatial Gene Expression Patterns Occur during Anther Development. The Plant Cell **2**: 1201-1224.

Kosslak, R.M., Chamberlin, M.A., Palmer, R.G., and B.A. Bowen (1997) Programmed Cell Death in the Root Cortex of Soybean *root necrosis* Mutants. The Plant Journal **11**: 729-745.

Kouchi, H., Takane, K., So, R.B., Ladha, J.K., and P.M. Reddy (1999) Rice *ENOD40*: Isolation and Expression Analysis in Rice and Transgenic Soybean Root Nodules. The Plant Journal **18**: 121-129.

Koukalova, B., Kovarik, A., Fajkus, J., and J. Siroky (1997) Chromatin Fragmentation Associated with Apoptotic Changes in Tobacco Cells Exposed to Cold Stress. FEBS Letters **414**: 289-292.

Krusell, L., Madsen, L.H., Sato, S., Aubert, G., Genua, A., Szczyglowski, K., Duc, G., Kaneko, T., Tabata, S., De Bruijn, F., Pajuelo, E., Sandal, N., and J. Stougaard (2002) Shoot Control of Root Development and Nodulation is Regulated by a Receptor-Like Kinase. Nature **420**: 422-426.

Kurkidjian, A.C. (1995) Role of the Differentiation of Root Epidermal Cells in Nod Factor (from *Rhizobium meliloti*)-Induced Root Hair Depolarization of *Medicago sativa*. Plant Physiology **107**: 783-790.

Lacomme, C. and S. Santa Cruz (1999) BAX-Induced Cell Death in Tobacco is Similar to the Hypersensitive Response. Proceedings of the National Academy of Sciences USA **96**: 7956-7961.

Lam, E. (2005) Vacuolar Proteases Livening Up Programmed Cell Death. Trends in Cell Biology **15**: 124-127.

Lam, E. (2004) Controlled Cell Death, Plant Survival and Development. Nature Reviews: Molecular Cell Biology **5**: 305-315.

Lam, E., Kato, N., and M. Lawton (2001) Programmed Cell Death, Mitochondria and the Hypersensitive Response. Nature **411**: 848-853.

Lam, E., Pontier, D., and O. Del Pozo (1999) Die and Let Live – Programmed Cell Death in Plants. Current Opinion in Plant Biology **2**: 502-507.

Lamb, C. and R.A. Dixon (1997) The Oxidative Burst in Plant Disease Resistance. Annual Review of Plant Physiology and Plant Molecular Biology **48**: 251-275.

Larsen, K. (2003) Molecular Cloning and Characterization of a cDNA Encoding a Ryegrass (*Lolium perenne*) *ENOD40* Homologue. Journal of Plant Physiology **160**: 675-687.

Layzell, D.B. and C.A. Atkins (1997) The Physiology and Biochemistry of Legume N_2 Fixation. *In* D.T. Dennis and D.H. Turpin, eds. Plant Metabolism: Physiology, Biochemistry, and Molecular Biology. Longman Publishers, Toronto, Canada, pp. 495-505.

Lehninger, A.L., Nelson, D.L., and M.M. Cox (1993) Principles of Biochemistry, 2nd ed. Worth Publishers, Inc., New York, USA, pp. 870-872; 975.

Leist, M. and M. Jaattela (2001) Four Deaths and a Funeral: From Caspases to Alternative Mechanisms. Nature Reviews Molecular Cell Biology **2**: 1-10.

Leonard, E., Chemler, J., Hong Lim, K., and M.A.G. Koffas (2005) Expression of a Soluble Flavone Synthase Allows the Biosynthesis of Phytoestrogen Derivatives in *Escherichia coli*. Applied Microbiology and Biotechnology **Online First (July 15)**: 1-7.

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P., and S. Rombauts (2002) PlantCARE, a Database of Plant *Cis*-Acting Regulatory Elements and a Portal to Tools for *In Silico* Analysis of Promoter Sequences. Nucleic Acids Research **30**: 325-327.

Lester, D.R., Phillips, A., Hedden, P., and I. Andersson (2005) Purification and Kinetic Studies of Recombinant Gibberellin Dioxygenases. BMC Plant Biology **5**: 19-27.

Levine, A., Pennell, R.I., Alvarez, M.E., Palmer, R., and C. Lamb (1996) Calcium-Mediated Apoptosis in a Plant Hypersensitive Disease Resistance Response. Current Biology **6**: 427-437.

Levine, A., Tenhaken, R., Dixon, R., and C. Lamb (1994) H_2O_2 from the Oxidative Burst Orchestrates the Plant Hypersensitive Disease Resistance Response. Cell **79**: 583-593.

Levy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ane, J.-M., Lauber, E., Bisseling, T., Denarie, J., Rosenberg, C., and F. Debelle (2004) A Putative Ca^{2+} and Calmodulin-Dependent Protein Kinase Required for Bacterial and Fungal Symbioses. Science **303**: 1361-1364.

Li, Q. and A.G. Hunt (1997) The Polyadenylation of RNA in Plants. Plant Physiology **115**: 321-325.

Li, Y.Q., Southworth, D., Linskens, H.F., Mulcahy, D.L., and M. Cresti (1995) Localization of Ubiquitin in Anthers and Pistils of *Nicotiana*. Sexual Plant Reproduction **8**: 123-128.

Lievens, S., Goormachtig, S., Den Herder, J., Capoen, W., Mathis, R., Hedden, P., and M. Holsters (2005) Gibberellins are Involved in Nodulation of *Sesbania rostrata*. Plant Physiology **139**: 1366-1379.

Lim, P.O., Woo, H.R., and H.G. Nam (2003) Molecular Genetics of Leaf Senescence in *Arabidopsis*. Trends in Plant Sciences **8**: 272-278.

Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and R. Geurts (2003) LysM Domain Receptor Kinases Regulating Nod Factor-Induced Infection. Science **302**: 630-633.

Lockshin, R.A. and Z. Zakeri (2004) Apoptosis, Autophagy, and More. The International Journal of Biochemistry and Cell Biology **36**: 2405-2419.

Loo, D.T. and J.R. Rillema (1998) Measurement of Cell Death. Methods in Cell Biology **57**: 251-264.

Lopez-Lara, I.M., Van den Berg, J.D.J., Thomas-Oates, J.E., Glushka, J., Lugtenberg, B.J.J., and H.P. Spaink (1995) Structural Identification of the Lipo-Chitin Oligosaccharide Nodulation Signals of *Rhizobium loti*. Molecular Microbiology **15**: 627-638.

Lorkovic, Z.J., Wieczorek Kirk, D.A., Lambermon, M.H.L., and W. Filipowicz (2000) Pre-mRNA Splicing in Higher Plants. Trends in Plant Science **5**: 160-167.

Lukacin, R., Groning, I., Schiltz, E., Britsch, L., and U. Matern (2000) Purification of Recombinant Flavanone 3β-Hydroxylase from *Petunia hybrida* and Assignment of the Primary Site of Proteolytic Degradation. Archives of Biochemistry and Biophysics **375**: 364-370.

Lukacin, R., Matern, U., Junghanns, K.T., Heskamp, M.L., Britsch, L., Forkmann, G., and S. Martens (2001) Purification and Antigenicity of Flavone Synthase I from Irradiated Parsley Cells. Archives of Biochemistry and Biophysics **393**: 177-183.

Machackova, I., Chauvaux, N., Dewitte, W., and H. Van Onckelen (1997) Diurnal Fluctuations in Ethylene Formation in *Chenopodium rubrum*. Plant Physiology **113**: 981-985.

Madeo, F., Herker, E., Maldener, C., Wissing, S., Lachelt, S., Herlan, M., Fehr, M., Lauber, K., Sigrist, S.J., Wesselborg, S., and K.U. Frohlich (2002) A Caspase-Related Protease Regulates Apoptosis in Yeast. Molecular Cell **9**: 1-20.

Madsen, E.B., Madsen, L.H., Radutoiu, S., Olbryt, M., Rakwaiska, M., Szczyglowski, K., Sate, S., Kaneko, T., Tabata, S., Sandal, N., and J. Stougaard (2003) A Receptor Kinase Gene of the LysM Type is Involved in Legume Perception of Rhizobial Signals. Nature **425**: 637-640.

Malik, N.S.A., Calvert, S.E., and W.D. Bauer (1987) Nitrate Induced Regulation of Nodule Formation in Soybean. Plant Physiology **84**: 266-271.

Maniatis, T., Fritsch, E.F., and J. Sambrook (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, USA

Mansoorian, N. (2005) Analysis of a Novel Myb-Like Gene from Soybean. M.Sc. Thesis, Department of Biology, University of Ottawa, Ottawa, ON

Marchler-Bauer, A., and S.H. Bryant (2004) CD-Search: Protein Annotations on the Fly. Nucleic Acids Research **32 (Web Server Issue)**: W327-W331.

Matamoros, M.A., Baird, L.M., Escuredo, P.R., Dalton, D.A., Minchin, F.R., Iturbe-Ormaetxe, I., Rubio, M.C., Moran, J.F., Gordon, A.J., and M. Becana (1999) Stress-Induced Legume Root Nodule Senescence. Physiological, Biochemical, and Structural Alterations. Plant Physiology **121**: 97-111.

May, M.J., Hammond-Kosack, K.E., and J.D.G. Jones (1996) Involvement of Reactive Oxygen Species, Glutathione Metabolism, and Lipid Peroxidation in the *Cf*-Gene-Dependent Defense Response of Tomato Cotyledons Induced by Race-Specific Elicitors of *Cladosporium fulvum*. Plant Physiology **110**: 1367-1379.

McCabe, P.F., Levine, A., Meijer, P.J., Tapon, N.A., and R.I. Pennell (1997) A Programmed Cell Death Pathway Activated in Carrot Cells Cultured at Low Cell Density. The Plant Journal **12**: 267-280.

McKenna, S.L., McGowan, A.J., and T.G. Cotter (1998) Molecular Mechanisms of Programmed Cell Death. *In* T. Scheper, ed. Advances in Biochemical Engineering Biotechnology **62**. Springer-Verlag, New York, USA, pp. 1-31.

Meldolesi, J. (1998) Calcium Signalling – Oscillation, Activation, Expression. Nature **392**: 863-865.

Mellor, R.B. and D.B. Collinge (1995) A Simple Model Based on Known Plant Defence Reactions is Sufficient to Explain Most Aspects of Nodulation. Journal of Experimental Botany **46**: 1-18.

Mignone, F., Gissi, C., Liuni, S., and G. Pesole (2002) Untranslated Regions of mRNAs. Genome Biology 3: REVIEWS0004.

Mignone, F., Grillo, G., Licciulli, F., Iacono, M., Liuni, S., Kersey, P.J., Duarte, J., Saccone, C., and G. Pesole (2005) UTRdb and UTRsite: A Collection of Sequences and Regulatory Motifs of the Untranslated Regions of Eukaryotic mRNAs. Nucleic Acids Research **33**: D141-D146.

Miller, D.D, De Ruijter, N.C.A., Bisseling, T., and A.M.C. Emons (1999) The Role of Actin in Root Hair Morphogenesis: Studies with Lipochito-Oligosaccharide as a Growth Stimulator and Cytochalasin as an Actin Perturbing Drug. The Plant Journal **17**: 141-154.

Minic, Z., LeProust-LeCoester, L., LaPorte, J., DeKouchkovsky, J., and S.C. Brown (2000) Proteins Isolated from Lucerne Roots by Affinity Chromatography with Sugars Analogous to Nod Factor Moieties. Biochemical Journal **345**: 255-262.

Mitra, R.M., Gleason, C.A., Edwards, A., Hadfield, J., Downie, J.A., Oldroyd, G.E.D., and S.R. Long (2004) A Ca^{2+}/Calmodulin-Dependent Protein Kinase Required for Symbiotic Nodule Development: Gene Identification by Transcript-Based Cloning. Proceedings of the National Academy of Sciences USA **101**: 4701-4705.

Mitsuhara, I., Malik, K.A., Miura, M., and Y. Ohashi (1999) Animal Cell-Death Suppressors BCL-X(L) and CED-9 Inhibit Cell Death in Tobacco Plants. Current Biology **9**: 775-778.

Mittler, R. (1998) Cell Death in Plants. *In* When Cells Die, R.A. Lockshin and Z. Zakeri, eds. Wiley-Liss, Inc., Toronto, Canada, pp. 147-174.

Mittler, R. and E. Lam (1997) Characterization of Nuclease Activities and DNA Fragmentation Induced upon Hypersensitive Response Cell Death and Mechanical Stress. Plant Molecular Biology **34**: 209-221.

Mittler, R. and E. Lam (1996) Sacrifice in the Face of Foes: Pathogen-Induced Programmed Cell Death in Plants. Trends in Microbiology **4**: 10-15.

Mittler, R. and E. Lam (1995) Identification, Characterization, and Purification of a Tobacco Endonuclease Activity Induced upon Hypersensitive Response Cell Death. The Plant Cell **7**: 1951-1962.

Mittler, R., Simon, L., and E. Lam (1997) Pathogen-Induced Programmed Cell Death in Tobacco. Journal of Cell Science **110**: 1333-1344.

Molteni, A., Brizio-Molteni, L., and V. Persky (1995) *In vitro* Hormonal Effects of Soybean Isoflavones. Journal of Nutrition **125**: S751-S756.

Moriguchi, T., Komatsu, A., Kita, M., Akimitsu, K., Endo-Inagaki, T., and M. Omura (2000) Molecular Cloning of a Homologue of *DAD-1* Gene in Citrus: Distinctive Expression during Fruit Development. Biochimica et Biophysica Acta **1490**: 198-202.

Morris, J.B. (1997) Special Purpose Legume Genetic Resources Conserved for Agricultural, Industrial, and Pharmaceutical Use. Economic Botany **51**: 251-263.

Mylona, P., Pawlawski, K., and T. Bisseling (1995) Symbiotic Nitrogen Fixation. The Plant Cell **7**: 869-885.

Nakaminami, K., Sawada, Y., Suzuki, M., Kenmoku, H., Kawaide, H., Mitsuhashi, W., Sassa, T., Inoue, Y., Kamiya, Y., and T. Toyomasu (2003) Deactivation of Gibberellin by 2-Oxidation during Germination of Photoblastic Lettuce Seeds. Bioscience, Biotechnology, and Biochemistry **67**: 1551-1558.

Nakamura, M., Tsunoda, T., and J. Obokata (2002) Photosynthesis Nuclear Genes Generally Lack TATA-Boxes: A Tobacco Photosystem I Gene Responds to Light through an Initiator. The Plant Journal **29**: 1-10.

Nakashima, T., Sekiguchi, T., Kuraoka, A., Fukushima, K., Shibata, Y., Komiyama, S., and T. Nishimoto (1993) Molecular Cloning of a Human cDNA Encoding a Novel Protein, DAD1, whose Defect Causes Apoptotic Cell Death in Hamster BHK21 Cells. Molecular and Cellular Biology **13**: 6367-6374.

Nap, J.-P. and T. Bisseling (1990) Developmental Biology of a Plant-Prokaryote Symbiosis: The Legume Root Nodule. Science **250**: 948-954.

Nash, D.T. and H.M. Schulman (1976) Leghemoglobins and Nitrogenase Activity during Soybean Root Nodule Development. Canadian Journal of Botany **54**: 2790-2797.

Navarre, D.A. and T.J. Wolpert (1999) Victorin Induction of an Apoptotic/Senescence-Like Response in Oats. The Plant Cell **11**: 237-249.

Ndoye, I., DeBilly, F., Vasse, J., Dreyfus, B., and G. Truchet (1994) Root Nodulation of *Sesbania rostrata*. Journal of Bacteriology **176**: 1060-1068.

Neo, H.H. and D.B. Layzell (1997) Phloem Glutamine and the Regulation of O_2 Diffusion in Legume Nodules. Physiologia Plantarum **113**: 259-267.

Nishimura, R., Hayashi, M., Wu, G.-J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., Kawasaki, S., Akao, S., Ohmori, M., Nagasawa, M., Harada, K., and M. Kawaguchi (2002) *HAR1* Mediates Systemic Regulation of Symbiotic Organ Development. Nature **420**: 426-429.

Nishiuchi, T., Shinshi, H., and K. Suzuki (2004) Rapid and Transient Activation of Transcription of the *ERF3* Gene by Wounding in Tobacco Leaves: Possible Involvement of NtWRKYs and Autorepression. Journal of Biological Chemistry **279**: 55355-55361.

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

Obara, K., Kuriyama, H., and H. Fukuda (2001) Direct Evidence of Active and Rapid Nuclear Degradation Triggered by Vacuole Rupture during Programmed Cell Death in *Zinnia*. Plant Physiology **125**: 615-626.

O'Brien, I.E., Reutelingsperger, C.P., and K.M. Holdaway (1997) Annexin-V and TUNEL Use in Monitoring the Progression of Apoptosis in Plants. Cytometry **29**: 28-33.

Oldroyd, G.E.D., Engstrom, E.M., and S.R. Long (2001) Ethylene Inhibits the Nod Factor Signal Transduction Pathway of *Medicago truncatula*. The Plant Cell **13**: 1835-1849.

Oldroyd, G.E.D. and S.R. Long (2003) Identification and Characterization of *Nodulation-Signalling Pathway 2*, a Gene of *Medicago truncatula* involved in Nod Factor Signalling. Plant Physiology **131**: 1027-1032.

Orzaez, D. and A. Granell (1997a) The Plant Homologue of the *Defender Against Apoptotic Death* Gene is Down-Regulated during Senescence of Flower Petals. FEBS Letters **404**: 275-278.

Orzaez, D. and A. Granell (1997b) DNA Fragmentation is Regulated by Ethylene during Carpel Senescence in Pisum sativum. The Plant Journal **11**: 137-144.

Osuna, D., Galvez, G., Pineda, M., and M. Aguilar (1999) RT-PCR Cloning, Characterization and mRNA Expression Analysis of a cDNA Encoding a Type II Asparagine Synthetase in Common Bean. Biochimica et Biophysica Acta **1445**: 75-85.

Otegui, M.S., Noh, Y.-S., Martinez, D.E., Petroff, M.G.V., Staehelin, A., Amasino, R.M., and J.J. Guiamet (2005) Senescence-Associated Vacuoles with Intense Proteolytic Activity Develop in Leaves of *Arabidopsis* and Soybean. The Plant Journal **41**: 831-844.

Ott, T., Van Dongen, J.T., Gunther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Bock, V., Czechowski, T., Geigenberger, P., and M.K. Udvardi (2005) Symbiotic Leghemoglobins are Crucial for Nitrogen Fixation in Legume Root Nodules but Not for General Plant Growth and Development. Current Biology **15**: 531-535.

Overmyer, K., Tuominen, H., Kettunen, R., Betz, C., Langebartels, C., Sandermann, Jr., H., and J. Kangasjarvi (2001) Ozone-Sensitive *Arabidopsis rcd1* Mutant Reveals Opposite Roles for Ethylene and Jasmonate Signaling Pathways in Regulating Superoxide-Dependent Cell Death. The Plant Cell **12**: 1849-1862.

Paetau, I., Chen, C.Z., and J.L. Jane (1994) Biodegradable Plastic Made from Soybean Products: 1. Effect of Preparation and Processing on Mechanical Properties and Water Absorption. Industrial Engineering and Chemical Research **33**: 1821-1827.

Page, R.D.M. (1996) TREEVIEW: An Application to Display Phylogenetic Trees on Personal Computers. Computer Applications in the Biosciences **12**: 357-358.

Panavas, T., Walker, E.L., and R. Rubinstein (1998) Possible Involvement of Abscisic Acid in Senescence of Daylily Petals. Journal of Experimental Botany **49**: 1987-1997.

Parniske, M. and J.A. Downie (2003) Locks, Keys, and Symbioses. Nature **425**: 569-570.

Pastori, G.M. and L.A. Del Rio (1997) Natural Senescence of Pea Leaves (an Activated Oxygen-Mediated Function for Peroxisomes). Plant Physiology **113**: 411-418.

Peng, H.M., Dreyer, D.A., VandenBosch, K.A., and D. Cook (1996) Gene Structure and Differential Regulation of the *Rhizobium*-Induced Peroxidase gene *RIP1*. Plant Physiology **112**: 1437-1446.

Penmetsa, R.V. and D.R. Cook (1997) A Legume Ethylene-Insensitive Mutant Hyperinfected by its Rhizobial Symbiont. Science **275**: 527-530.

Penmetsa, R.V., Frugoli, J.A., Smith, L., Long, S.R., and D.R. Cook (2003) Dual Genetic Pathways Controlling Nodule Number in *Medicago truncatula*. Plant Physiology **131**: 998-1008.

Pennell, R.I. and C. Lamb (1997) Programmed Cell Death in Plants. The Plant Cell **9**: 1157-1168.

Perlick, A.M. and A. Puhler (1993) A Survey of Transcripts Expressed Specifically in Root Nodules of Broadbean (*Vicia faba* L.). Plant Molecular Biology **22**: 957-970.

Perret, X., Stahelin, C., and W.J. Broughton (2000) Molecular Basis of Symbiotic Promiscuity. Microbiology and Molecular Biology Reviews **64**: 180-201.

Pfeiffer, N.E., Malik, S.A., and F.W. Wagner (1983a) Reversible Dark-Induced Senescence of Soybean Root Nodules. Plant Physiology **71**: 393-399.

Pfeiffer, N.E., Torres, C.M., and F.W. Wagner (1983b) Proteolytic Activity in Soybean Root Nodules. Plant Physiology **71**: 797-802.

Pichon, M., Journet, E.P., Dedieu, A., de Billy, F., Truchet, G., and D.G. Barker (1992) *Rhizobium meliloti* Elicits Transient Expression of the Early Nodulin Gene *ENOD12* in the Differentiating Root Epidermis of Transgenic Alfalfa. The Plant Cell **4**: 1199-1211.

Piechulla, B., Merforth, N., and B. Rudolph (1998) Identification of Tomato *LHC* Promoter Regions Necessary for Circadian Expression. Plant Molecular Biology **38**: 655-662.

Pingret, J.-L., Journet, E.-P., and D.G. Barker (1998) *Rhizobium* Nod Factor Signaling: Evidence for a G Protein-Mediated Transduction Mechanism. The Plant Cell **10**: 659-671.

Pladys, D. and J. Rigaud (1985) Senescence in French Bean Nodules: Occurrence of Different Proteolytic Activities. Physiologia Plantarum **63**: 43-48.

Pladys, D. and C.P. Vance (1993) Proteolysis During Development and Senescence of Effective and Plant Gene-Controlled Ineffective Alfalfa Nodules. Plant Physiology **103**: 379-384.

Prescott, A.G. (1993) A Dilemma of Dioxygenases (or Where Biochemistry and Molecular Biology Fail to Meet). Journal of Experimental Botany **44**: 849-861.

Prescott, A.G. and P. John (1996) Dioxygenases: Molecular Structure and Role in Plant Metabolism. Annual Review of Plant Physiology and Plant Molecular Biology **47**: 245-271.

Price, N.P.J., Relic, B., Talmont, F., Lewin, A., Prome, D., Pueppke, S.G., Maillet, F., Denarie, J., Prome, J.-C., and W.J. Broughton (1992) Broad-Host-Range *Rhizobium* Species Strains NGR234 Secretes a Family of Carbamoylated, and Fucosylated, Nodulation Signals that are *O*-Acetylated or Sulphated. Molecular Microbiology **6**: 3575-3584.

Price, N.P.J., Talmont, F., Wieruszeski, J.-M., Prome, D., and J.-C. Prome (1996) Structural Determination of Symbiotic Nodulation Factors from the Broad Host-Range *Rhizobium* Species NGR234. Carbohydrate Research **289**: 115-136.

Proskuryakov, S.Y., Konoplyannikov, A.G., and V.L. Gabai (2003) Necrosis: A Specific Form of Programmed Cell Death? Experimental Cell Research **283**: 1-16.

Puppo, A., Groten, K., Bastian, F., Carzaniga, R., Soussi, M., Lucas, M.M., De Felipe, M.R., Harrison, J., Vanacker, H., and C.H. Foyer (2005) Legume Nodule Senescence: Roles for Redox and Hormone Signalling in the Orchestration of the Natural Aging Process. New Phytologist **165**: 683-701.

Quinto, C., Wijfjes, A.H.M., Bloemberg, G.V., Blok-Tip, L., Lopez-Lara, I.M., Lugtenberg, B.J.J., Thomas-Oates, J.E., and H.P. Spaink (1997) Bacterial Nodulation Protein NODZ is a Chitin Oligosaccharide Fucosyltransferase Which Can Also Recognize Substrates of Animal Origin. Proceedings of the National Academy of Sciences USA **94**: 4336-4341.

Quirino, B.F., Noh, Y.S., Himelblau, E., and R.M. Amasino (2000) Molecular Aspects of Leaf Senescence. Trends in Plant Science **5**: 278-282.

Rao, M.V. and K.R. Davis (2001) The Physiology of Ozone Induced Cell Death. Planta **213**: 682-690.

Radutoiu, S., Madsen, L.H., Madsen, E.B., Felle, H.H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and J. Stougaard (2003) Plant Recognition of Symbiotic Bacteria Requires Two LysM Receptor-Like Kinases. Nature **425**: 585-592.

Raven, P.H., Evert, R.F., and S.E. Eichhorn (1992) Biology of Plants, 5th ed. Worth Publishers, New York, USA, pp. 197-198, 607.

Reddy, V.S. and A.S.N. Reddy (2004) Developmental and Cell-Specific Expression of *ZWICHEL* is Regulated by the Intron and Exon Sequences of its Upstream Protein-Coding Gene. Plant Molecular Biology **54**: 273-293.

Roche, P., Debelle, E., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Denarie, J., and J.-C. Prome (1991) Molecular Basis of Symbiotic Host Specificity in *Rhizobium meliloti*: *NODH* and *NODPQ* Genes Encode the Sulfation of Lipo-Oligosaccharide Signals. Cell **67**: 1131-1143.

Rohrig, H., John, M., and J. Schmidt (2004) Modification of Soybean Sucrose Synthase by S-Thiolation with *ENOD40* Peptide A. Biochemical and Biophysical Research Communications 325: 864-870.

Rohrig, H., Schmidt, J., Miklashevichs, E., Schell, J., and M. John (2002) Soybean *ENOD40* Encodes Two Peptides that Bind to Sucrose Synthase. Proceedings of the National Academy of Sciences USA **99**: 1915-1920.

Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouze, P., and Y. van de Peer (2003) Computational Approaches to Identify Promoters and Cis-Regulatory Elements in Plant Genomes. Plant Physiology **132**: 1162-1176.

Ross, E.J., Stone, J.M., Elowsky, C.G., Arredondo-Peter, R., Klucas, R.V., and G. Sarath (2004) Activation of the *Oryza sativa* Non-Symbiotic Haemoglobin-2 Promoter by the Cytokinin-Regulated Transcription Factor, ARR1. Journal of Experimental Botany **55**: 1721-1731.

Rothnie, H.M. (1996) Plant mRNA 3'-End Formation. Plant Molecular Biology **32**: 43-61.

Ryerson, D.E. and M.C. Heath (1996) Cleavage of Nuclear DNA into Oligonucleosomal Fragments during Cell Death Induced by Fungal Infection or by Abiotic Treatments. The Plant Cell **8**: 393-402.

Sakai, H., Aoyama, T., and A. Oka (2000) *Arabidopsis* ARR1 and ARR2 Response Regulators Operate as Transcriptional Activators. The Plant Journal **24**: 703-711.

Sakai, H., Honma, T., Aoyama, T., Sato, S., Kato, T., Tabata, S., and A. Oka (2001) ARR1, a Transcription Factor for Genes Immediately Responsive to Cytokinins. Science **294**: 1519-1521.

Sanchez, P., De Torres Zabala, M., and M. Grant (2000) *AtBI-1*, a Plant Homologue of *BAX INHIBITOR-1*, Suppresses Bax-Induced Cell Death in Yeast and is Rapidly Upregulated during Wounding and Pathogen Challenge. The Plant Journal **21**: 393-399.

Sanjuan, J., Carlson, R.W., Spaink, H.P., Bhat, R., Mark Barbour, W., Glushka, J., and G. Stacey (1992) A 2-*O*-Methylfucose Moiety is Present in the Lipopolysaccharide Nodulation Signal of *Bradyrhizobium japonicum*. Proceedings of the National Academy of Sciences USA **89**: 8789-8793.

Sasabe, M., Takeuchi, K., Kamoun, S., Ichinose, Y., Govers, F., Toyoda, K., Shiraishi, T., and T. Yamada (2000) Independent Pathways Leading to Apoptotic Cell Death, Oxidative Burst and Defense Gene Expression in Response to Elicitin in Tobacco Cell Suspension Culture. European Journal of Biochemistry **267**: 5005-5013.

Sawyer, S.A. (1989) Statistical Tests for Detecting Gene Conversion. Molecular Biology and Evolution **6**: 526-538.

Schauser, L., Roussis, A., Stiller, J., and J. Stougaard (1999) A Plant Regulator Controlling Development of Symbiotic Root Nodules. Nature **411**: 191-195.

Scheres, B., Van de Wiel, C., Zalensky, A., Horvath, B., Spaink, H., Van Eck, H., Zwartkruis, F., Wolters, A.-M., Gloudemans, T., Van Kammen, A., and T. Bisseling (1990a) The *ENOD12* Gene Product is Involved in the Infection Process During the Pea-*Rhizobium* Interaction. Cell **60**: 281-294.

Scheres, B., Van Engelen, F., Van der Knaap, E., Van de Wiel, C., Van Kammen, A., and T. Bisseling (1990b) Sequential Induction of Nodulin Gene Expression in the Developing Pea Nodule. The Plant Cell **2**: 687-700.

Schmid, M., Simpson, D., and C. Gietl (1999) Programmed Cell Death in Castor Bean Endosperm is Associated with the Accumulation and Release of a Cysteine Endopeptidase from Ricinosomes. Proceedings of the National Academy of Sciences USA **96**: 14159-14164.

Schmidt, J.S., Harper, J.E., Hoffman, T.K., and A.F. Bent (1999) Regulation of Soybean Nodulation Independent of Ethylene Signaling. Plant Physiology **119**: 951-959.

Schultze, M. and A. Kondorosi (1998) Regulation of Symbiotic Root Nodule Development. Annual Review of Genetics **32**: 33-57.

Schwartz, S., Zhang, Z.,. Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., and W. Miller (2000) PipMaker – A Web Server for Aligning Two Genomic DNA Sequences. Genome Research **10**: 577-586.

Searle, I.R., Men, A.E., Laniya, T.S., Buzas, D.M., Iturbe-Ormaetxe, I., Carroll, B.J., and P.M. Gresshoff (2003) Long-Distance Signaling in Nodulation Directed by a CLAVATA1-Like Receptor Kinase. Science **299**: 109-112.

Shaham, S., Shuman, M.A., and I. Herskowitz (1998) Death-Defying Yeast Identify Novel Apoptosis Genes. Cell **92**: 425-427.

Shaw, S.L. and S.R. Long (2003) Nod Factor Elicits Two Separable Calcium Responses in *Medicago truncatula* Root Hair Cells. Plant Physiology **131**: 976-984.

Silverstein, K.A.T., Graham, M.A., and K.A. VandenBosch (2006) Novel Paralogous Gene Families with Potential Function in Legume Nodules and Seeds. Current Opinion in Plant Biology **9**: 142-146.

Smart, C.M., Hosken, S.E., Thomas, H., Greaves, J.A., Blair, B.G., and W. Schuch (1995) The Timing of Maize Leaf Senescence and Characterization of Senescence-Related cDNAs. Physiologia Plantarum **93**: 673-682.

Smil, V. (1999) Nitrogen in Crop Production: an Account of Global Flows. Global Biogeochemical Cycles **13**: 647-662.

Smit, G., De Koster, C.C., Schripsema, J., Spaink, H.P., Van Brussel, A.A., and J.W. Kijne (1995) Uridine, a Cell Division Factor in Pea Roots. Plant Molecular Biology **29**: 869-873.

Smit, P., Raedts, J., Portyanko, V., Debelle, F., Gough, C., Bisseling, T., and R. Geurts (2005) NSP1 of the GRAS Protein Family is Essential for Rhizobial Nod Factor-Induced Transcription. Science **308**: 1789-1791.

Smith, R.J. and J.R. Gallon (1993) Nitrogen Fixation. *In* P.J. Lea and R.C. Leegood, eds. Plant Biochemistry and Molecular Biology. John Wiley & Sons Ltd., Toronto, Canada, pp. 129-153.

Socolow, R.H. (1999) Nitrogen Management and the Future of Food: Lessons from the Management of Energy and Carbon. Proceedings of the National Academy of Sciences USA **96**: 6001-6008.

Solano, R., Nieto, C., Avila, J., Canas, L., Diaz, I., and J. Paz-Ares (1995) Dual DNA Binding Specificity of a Petal Epidermis-Specific MYB Transcription Factor (MYB.Ph3) from *Petunia hybrida*. The EMBO Journal **14**: 1773-1784.

Solomon, M., Belenghi, B., Delledonne, M., Menachem, E., and A. Levine (1999) The Involvement of Cysteine Proteases and Protease Inhibitor Genes in the Regulation of Programmed Cell Death in Plants. The Plant Cell **11**: 431-443.

Sousa, C., Johansson, C., Charon, C., Manyani, H., Sautier, C., Kondorosi, A., and M. Crespi (2001) Translational and Structural Requirements of the Early Nodulin Gene *ENOD40*, a Short-Open Reading Frame-Containing RNA, for Elicitation of a Cell-Specific Growth Response in the Alfalfa Root Cortex. Molecular and Cellular Biology **21**: 354-366.

Spaink, H.P., Sheeley, D.M., Van Brussel, A.A.N., Glushka, J., York, W.S., Tak, T., Geiger, O., Kennedy, E.P., Reinhold, V.N., and B.J.J. Lugtenberg (1991) A Novel Highly Unsaturated Fatty Acid Moiety of Lipo-Oligosaccharide Signals Determines Host Specificity of *Rhizobium*. Nature **354**: 125-130.

Sprent, J.I. and R. Parsons (2000) Nitrogen Fixation in Legume and Non-Legume Trees. Field Crops Research **65**: 183-196.

Staxen, I., Pical, C., Montgomery, L.T., Gray, J.E., Hetherington, A.M., and M.R. McAinsh (1999) Abscisic Acid Induces Oscillations in Guard-Cell Cytosolic Free Calcium that Involve Phosphoinositide-Specific Phospholipase C. Proceedings of the National Academy of Sciences USA **96**: 1779-1784.

Steffens, B. and M. Sauter (2005) Epidermal Cell Death in Rice is Regulated by Ethylene, Gibberellin, and Abscisic Acid. Plant Physiology **139**: 713-721.

Stein, J.C. and G. Hansen (1999) Mannose Induces an Endonuclease Responsible for DNA Laddering in Plant Cells. Plant Physiology **121**: 71-79.

Stougaard, J. (2001) Genetics and Genomics of Root Symbiosis. Current Opinion in Plant Biology **4**: 328-335.

Stougaard, J. (2000) Regulators and Regulation of Legume Root Nodule Development. Plant Physiology **124**: 531-540.

Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tahata, S., Sandal, N., Stougaard, J., Szczyglowski, K., and M. Parniske (2002) A Plant Receptor-Like Kinase Required for Both Bacterial and Fungal Symbiosis. Nature **417**: 959-962.

Streeter, J. (1988) Inhibition of Legume Nodule Formation and N_2 Fixation by Nitrate. CRC Critical Reviews of Plant Science **7**: 1-23.

Stromvik, M.V., Sundararaman, V.P., and L.O. Vodkin (1999) A Novel Promoter from Soybean that is Active in a Complex Developmental Pattern With and Without its Proximal 650 Base Pairs. Plant Molecular Biology **41**: 217-231.

Sugimoto, A., Hozak, R.R., Nakashima, T., Nishimoto, T., and J.H. Rothman (1995) DAD-1, an Endogenous Programmed Cell Death Suppressor in *Caenorhabditis elegans* and Vertebrates. The EMBO Journal **14**: 4434-4441.

Sun, Y.-L., Zhao, Y., Hong, X., and Z.-H. Zhai (1999) Cytochrome c Release and Caspase Activation during Menadione-Induced Apoptosis in Plants. FEBS Letters **462**: 317-321.

Sutton, W.D. (1983) Nodule Development and Senescence. *In* W.J. Broughton, ed. Nitrogen Fixation, Vol. 1 (Legumes). Clarendon Press, Oxford, England, pp. 142-212.

Szczyglowski, K. and L. Amyot (2003) Symbiosis, Inventiveness by Recruitment? Plant Physiology **131**: 935-940.

Taiz, L. and E. Zeiger (1991) Plant Physiology. The Benjamin/Cummings Publishing Company, Inc., Don Mills, Canada, pp. 331-333.

Taylor, B. and A. Powell (1985) Isolation of Plant DNA and RNA. BRL Focus **4**: 4-6.

Terzaghi, W.B. and A.R. Cashmore (1995) Light-Regulated Transcription. Annual Review of Plant Physiology and Plant Molecular Biology **46**: 445-474.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and D.G. Higgins (1997) The CLUSTAL_X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools. Nucleic Acids Research **25**: 4876-4882.

Thompson, J.D., Higgins, D.G., and T.J. Gibson (1994) CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties, and Weight Matrix Choice. Nucleic Acids Research **22**: 4673-4680.

Thum, K.E., Kim, M., Morishige, D.T., Eibl, C., Koop, H.U., and J.E. Mullet (2001) Analysis of Barley Chloroplast *PSBD* Light-Responsive Promoter Elements in Transplastomic Tobacco. Plant Molecular Biology **47**: 353-366.

Tiwari, B.S., Belenghi, B., and A. Levine (2002) Oxidative Stress Increased Respiration and Generation of Reactive Oxygen Species, Resulting in ATP Depletion, Opening of Mitochondrial Permeability Transition, and Programmed Cell Death. Plant Physiology **128**: 1271-1281.

Toyomasu, T., Tsuji, H., Yamane, H., Nakayama, M., Yamaguchi, I., Murofushi, N., Takahashi, N., and Y. Inoue (1993) Light Effects on Endogenous Levels of Gibberellins in Photoblastic Lettuce Seeds. Journal of Plant Growth Regulators **12**: 85-90.

Urao, T., Yamaguchi-Shinozaki, K., Urao, S., and K. Shinozaki (1993) An *Arabidopsis* MYB Homolog is Induced by Dehydration Stress and its Gene Product Binds to the Conserved MYB Recognition Sequence. The Plant Cell **5**: 1529-1539.

Uren, A.G., O'Rourke, K., Aravind, L.A., Pisabarro, M.T., Seshagiri, S., Koonin, E.V., and V.M. Dixit (2000) Identification of Paracaspases and Metacaspases: Two Ancient Families of Caspase-Like Proteins, One of which Plays a Key Role in MALT Lymphoma. Molecular Cell **6**: 961-967.

Vance, C.P. (2001) Symbiotic Nitrogen Fixation and Phosphorus Acquisition. Plant Nutrition in a World of Declining Renewable Resources. Plant Physiology **127**: 390-397.

Vance, C.P. (1997) The Molecular Biology of N Metabolism. *In* D.T. Dennis and D.H. Turpin, eds. Plant Metabolism: Physiology, Biochemistry, and Molecular Biology. Longman Publishers, Toronto, Canada, pp. 449-477.

Vance, C.P., Graham, P.H., and D.L. Allan (2000) Biological Nitrogen Fixation: Phosphorus a Critical Future Need? *In* F.O. Pederosa, M. Hungria, M.G. Yates, and W.E. Newton, eds. Nitrogen Fixation from Molecules to Crop Productivity. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 509-518.

Van de Velde, W., Perez Guerra, J.C., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M., and S. Goormachtig (2006) Aging in Legume Symbiosis. A Molecular View on Nodule Senescence in *Medicago truncatula*. Plant Physiology **141**: 711-720.

Van de Wiel, C., Scheres, B., Franssen, H., Van Lierop, M.J., Van Lammeren, A., Van Kammen, A., and T. Bisseling (1990) The Early Nodulin Transcript *ENOD2* is Located in the Nodule Parenchyma (Inner Cortex) of Pea and Soybean Root Nodules. The EMBO Journal **9**: 1-9.

Van Doorn, W.G. and E.J. Woltering (2005) Many Ways to Exit? Cell Death Categories in Plants. Trends in Plant Science **10**: 117-122.

Van Rhijn, P. and J. Vanderleyden (1995) The *Rhizobium*-Plant Symbiosis. Microbiological Reviews **59**: 124-142.

Vercher, Y., Molowny, A., and J. Carbonell (1987) Gibberellic Acid Effects on the Ultrastructure of Endocarp Cells of Unpollinated Ovaries of *Pisum sativum*. Physiologia Plantarum **71**: 302-308.

Verma, D.P., Haugland, R., Brisson, N., Legocki, R.P., and L. Lacroix (1981) Regulation of the Expression of Leghaemoglobin Genes in Effective and Ineffective Root Nodules of Soybean. Biochimica et Biophysica Acta **653**: 98-107.

Vessey, J.K. and J. Waterer (1992) In Search of the Mechanism of Nitrate Inhibition of Nitrogenase Activity in Legume Nodules: Recent Developments. Physiologia Plantarum **84**: 171-176.

Vicente-Carbajosa, J., Moose, S.P., Parsons, R., and R.J. Schmidt (1997) A Maize Zinc-Finger Protein Binds the Prolamin Box in *ZEIN* Gene Promoters and Interacts with the Basic Leucine Zipper Transcriptional Activator OPAQUE2. Proceedings of the National Academy of Sciences USA **94:** 7685-7690.

Vieweg, M.F., Fruhling, M., Quandt, H.J., Heim, U., Baumlein, H., Puhler, A., Kuster, H., and A.M. Perlick (2004) The Promoter of the *Vicia faba* L. Leghemoglobin Gene *VfLB29* is Specifically Activated in the Infected Cells of Mycorrhizal Roots from Different Legume and Non-Legume Plants. Molecular Plant Microbe Interactions **17**: 62-69.

Vijn, I., Yang, W.C., Pallisgard, N., Ostergaard, J.E., van Kammen, A., and T. Bisseling (1995) *VsENOD5*, *VsENOD12* and *VsENOD40* Expression during *Rhizobium*-Induced Nodule Formation on *Vicia sativa* Roots. Plant Molecular Biology **28**: 1111-1119.

Vleghels, I., Hontelez, J., Ribeiro, A., Fransz, P., Bisseling, T., and H. Franssen (2003) Expression of *ENOD40* during Tomato Plant Development. Planta **218**: 42-49.

Wais, R.J., Galera, C., Oldroyd, G., Catoira, R., Penmetsa, R.V., Cook, D., Gough, C., Denarie, J., and S.R. Long (2000) Genetic Analysis of Calcium Spiking Responses in Nodulation Mutants of *Medicago truncatula*. Proceedings of the National Academy of Sciences USA **97**: 13407-13412.

Walker, S.A., Viprey, V., and J.A. Downie (2000) Dissection of Nodulation Signaling Using Pea Mutants Defective for Calcium Spiking Induced by Nod Factors and Chitin Oligomers. Proceedings of the National Academy of Sciences USA **97**: 13413-13418.

Wang, H., Li, J., Bostock, R.M., and D.G. Gilchrist (1996) Apoptosis: a Functional Paradigm for Programmed Plant Cell Death Induced by a Host-Selective Phytotoxin and Invoked during Development. The Plant Cell **8**: 375-391.

Wang, M., Hoekstra, S., Van Bergen, S., Lamers, G.E., Oppedijk, B.J., Van der Heijden, M.W., De Priester, W., and R.A. Schilperoort (1999) Apoptosis in Developing Anthers and the Role of ABA in this Process during Androgenesis in *Hordeum vulgare* L. Plant Molecular Biology **39**: 489-501.

Wang, M., Oppedijk, B.J., Lu, X., Van Duijn, B., and R.A. Schilperoort (1996) Apoptosis in Barley Aleurone during Germination and its Inhibition by Abscisic Acid. Plant Molecular Biology **32**: 1125-1134.

Wang, Z.-Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M.S., and E.M. Tobin (1997) A MYB-Related Transcription Factor is Involved in the Phytochrome Regulation of an *Arabidopsis LHCB* Gene. The Plant Cell **9**: 491-507.

Watanabe, N. and E. Lam (2004) Recent Advance in the Study of Caspase-Like Proteases and BAX INHIBITOR-1 in Plants: Their Possible Roles as Regulator of Programmed Cell Death. Molecular Plant Pathology **5**: 65-70.

Weaver, L.M., Gan, S., Quirino, B., and R.M. Amasino (1998) A Comparison of the Expression Patterns of Several Senescence-Associated Genes in Response to Stress and Hormone Treatment. Plant Molecular Biology **37**: 455-469.

Weidner, S., Puhler, A., and H. Kuster (2003) Genomics Insights into Symbiotic Nitrogen Fixation. Current Opinion in Biotechnology **14**: 200-205.

Wellmann, F., Lukacin, R., Moriguchi, T., Britsch, L., Schiltz, E., and U. Matern (2002) Functional Expression and Mutational Analysis of Flavonol Synthase from *Citrus unshiu*. European Journal of Biochemistry **269**: 4134-4142.

Winstanley, D., Demissie, M., and S.E. Hollinger (2005) Nitrogen Properties, Nitrogen Cycles Project, Illinois State Water Survey, Illinois Council on Food and Agricultural Research Strategic Research Initiative on Water Quality, Project OOSi-040-5A-MBL (http://www.sws.uiuc.edu/nitro/nitrodesc.asp)

Wittwer, C.T. and D.J. Garling (1991) Rapid Cycle DNA Amplification: Time and Temperature Optimization. Biotechniques **10**: 76-83.

Woffenden, B.J., Freeman, T.B., and E.P. Beers (1998) Proteasome Inhibitors Prevent Tracheary Element Differentiation in *Zinnia* Mesophyll Cell Cultures. Plant Physiology **118**: 419-430.

Woltering, E.J., Van der Bent, A., and F.A. Hoeberichts (2002) Do Plant Caspases Exist? Plant Physiology **130**: 1764-1769.

Wu, L., Ueda, T., and J. Messing (1995) The Formation of mRNA 3′-Ends in Plants. The Plant Journal **8**: 323-329.

Wuttke, W., Jarry, H., Westphalen, S., Christoffel, V., and D. Seidlova-Wuttke (2002) Phytoestrogens for Hormone Replacement Therapy? The Journal of Steroid Biochemistry and Molecular Biology **83**: 133-147.

Xie, Z.-P., Staehelin, C., Wiemken, A., and T. Boller (1996) Ethylene Responsiveness of Soybean Cultivars Characterized by Leaf Senescence, Chitinase Induction and Nodulation. Journal of Plant Physiology **149**: 690-694.

Xie, Z. and Z. Chen (2000) Harpin-Induced Hypersensitive Cell Death is Associated with Altered Mitochondrial Functions in Tobacco Cells. Molecular Plant-Microbe Interactions **13**: 183-190.

Xu, Q. and J.C. Reed (1998) BAX INHIBITOR-1, a Mammalian Apoptosis Suppressor Identified by Functional Screening in Yeast. Molecular Cell **1**: 337-346.

Xu, Y. and M. Hanson (2000) Programmed Cell Death during Pollination-Induced Petal Senescence in Petunia. Plant Physiology **122**: 1323-1333.

Xu, Y.-L., Li, L., Wu, K., Peeters, A.J.M., Gage, D.A., and J.A.D. Zeevaart (1995) The *GA5* Locus of Arabidopsis thaliana Encodes a Multifunctional Gibberellin 20-Oxidase: Molecular Cloning and Functional Expression. Proceedings of the National Academy of Sciences **92**: 6640-6644.

Yanagisawa, S. (2000) DOF1 and DOF2 Transcription Factors are Associated with Expression of Multiple Genes Involved in Carbon Metabolism in Maize. The Plant Journal **21**: 281-288.

Yanagisawa, S. and R.J. Schmidt (1999) Diversity and Similarity Among Recognition Sequences of DOF Transcription Factors. The Plant Journal **17**: 209-214.

Yanagisawa, S., and J. Sheen (1998) Involvement of Maize DOF Zinc Finger Proteins in Tissue-Specific and Light-Regulated Gene Expression. The Plant Cell **10**: 75-89.

Yao, N., Eisfelder, B.J., Marvin, J., and J.T. Greenberg (2004) The Mitochondrion – an Organelle Commonly Involved in Programmed Cell Death in *Arabidopsis thaliana*. The Plant Journal **40**: 596-610.

Ye, Z.H. and J.E. Varner (1996) Induction of Cysteine and Serine Proteases during Xylogenesis in *Zinnia elegans*. Plant Molecular Biology **30**: 1233-1246.

Yen, C.-H. and C.-H. Yang (1998) Evidence for Programmed Cell Death during Leaf Senescence in Plants. Plant and Cell Physiology **39**: 922-927.

Yoshida, S. (2003) Molecular Regulation of Leaf Senescence. Current Opinion in Plant Biology **6**: 79-84.

Young, T.E. and D.R. Gallie (2000) Regulation of Programmed Cell Death in Maize Endosperm by Abscisic Acid. Plant Molecular Biology **42**: 397-414.

Yu, L.-H., Kawai-Yamada, M., Naito, M., Watanabe, K., Reed, J.C., and H. Uchimiya (2002) Induction of Mammalian Cell Death by a Plant BAX INHIBITOR. FEBS Letters **512**: 308-312.

Yu, X.-H., Perdue, T.D., Heimer, Y.M., and A.M. Jones (2002) Mitochondrial Involvement in Tracheary Element Programmed Cell Death. Cell Death and Differentiation **9**: 189-198.

Yun, D.J., Hashimoto, T., and Y. Yamada (1992) Metabolic Engineering of Medicinal Plants: Transgenic *Atropa belladonna* with an Improved Alkaloid Composition. Proceedings of the National Academy of Sciences USA **89**: 11799-11803.

Zhang, B., Chen, W., Foley, R.C., Buttner, M., and K.B. Singh (1995) Interactions between Distinct Types of DNA Binding Proteins Enhance Binding to *OCS* Element Promoter Sequences. The Plant Cell **7**: 2241-2252.

Zhao, Y., Jiang, Z.-F., Sun, Y.-L., and Z.-H. Zhai (1999) Apoptosis of Mouse Liver Nuclei Induced in the Cytosol of Carrot Cells. FEBS Letters **448**: 197-200.

Zhivotovsky, B. (2002) From the Nematode and Mammals Back to the Pine Tree: On the Diversity and Evolution of Programmed Cell Death. Cell Death and Differentiation **9**: 867-869.

Ziegler, U. and P. Groscurth (2004) Morphological Features of Cell Death. News in Physiological Sciences **19**: 124-128.

APPENDIX 1:

SAN1 Gene Sequences

Figure 1: Primers Used in Extension of *SAN1A* Gene Sequence

The locations of the nested primers on the Extended *Sst*I Fragment 1.3 kb for extension of the *SAN1A* gene sequence are shown. The start codon of the *SAN1A* gene is underlined and shown in red text.

```
1       GGGCAACTTCATCAAGGCGAGAACGATGCTTTACCAAGGTCTGCCTGATCCAAGAGAGGC
61      CGAGACTTGGGTTTATTTCAAACACTTATGTGGACAAAACACCTCAATTTGACAAATGTT
121     ATCTTCGAATTTGATTATAAAGTAGTGGTAGATGGAGTTAAAAATGTTATTAGAGACTCC
181     CTAGATTTTCACTGTATTATATATAAGTGTAGAGTCCTTCTTTCTTCTATTCTAAACTCA
241     GTAGTGAGTTTCAATAGGAGACTAATTAGTGGAAAATAGTCAAGTCCTACATCAACTGTC
301     TTAATTTTTGGAGTGCAACTTATATACCTGTGGGACAATTTCACTTAATACCAATTAGTT
361     TTAAGATGAAATCTAATACTAACAAATCATGTTGCTCATGTCTTCTTGGCAGCAGGATAT
421     CATGCTAGCATCCAAATTTTTTATTATATTCTAAATTGTATATTTCCATACAATATGAAT
481     GACGTGAAGTAATTATGTTTCCTTTCAAAAAAAAAAAAAGACATCTGCATGGTAGCATAA
541     CAGAAACCAGTCCAGGTTCAATGCACATAAATCAAAATATTTCTTTATTGTACGTGAACG
601     GAATCTTCAACTTCAAAATTGTATTTGCACCACACAAAACTTTCTCTTGTTCCTTAAATA
661     AGAGAACAACCCAAAACCTTTCCAAAATACACGTGCAGCAAATTGCTTACCCAATCACCA
721     GCC**ATG**GGAGAGGTTGACCCAGCTTTCATTCAAGAACCACAACACAGGCCAAATCTCTCC
781     ACCATTCAAGCAGAAGGAATTCCCATAATTGATCTCTCTCCAATAACCAACCACACAGTT
841     TCAGACCCTTCTGCAATTGAAAGCTTGGTGAAGGAGATAGGACGTGCATGCCAGGAGTGG
901     GGCTTCTTCCAAGTAACAAACCATGGAGTGCCACTCACTCTAAGGCAAAACATTGAGAAA
961     GCCTCAAAACTGTTCTTTGCTCAGACTCTGGAGGAGAAGAGAAAGGTTAGCAGAAATGAG
1021    AGCTCTCCAATGGGTTATTATGACACAGAACACACCAAGAACGTCAGGGACTGGAAAGAA
1081    GTCTTTGATTTTCTAGCCAAAGACCCCACTTTCATTCCTCTCACTTCTGATGAACATGAT
1141    GATCGAGTCAATCAATGGACTAATCAATCACCTCAATACCCTCCACTCTTCAGGTAATTC
1201    ATCTGTACTTTGCAAGGTTTTAAATACCGGTCGTGTTACGATTTTGTCGCGTTTGATGAT
1261    ATCACGGAAAATCACAAACAAATGTGGTTGAAGCGGTCACATTGTGATCACATACAGTTA
1321    AAAAACTTTGATGTTGTGGCCCAAATCACGGTTTGGTTTAACTGTCTATGACTGCAATGT
1381    GACCAATTGCAATTGAGACCGCATCAACTGCATTTGTCCGCGATTTTTCACGATATCAAC
1441    AAACACGAAGAAACCGCAATGTGACTGCAACTAATATTTAAAACCTTTAAAACAAAAGTG
1501    GATTGTGTTCAAAACTAATTGTTCAGCCCACAGTCATGACAGCTGGACAGTTTTTATAAT
1561    TCTATTGATGATGTGAATACGAGCAGGGTTGTAACACAAGAGTATATTCAGGAGATGGAA
1621    AAGCTGTCCTTTAAGCTGTTGGAGCTTATAGCTTTGAGCTTAGGCCTTGAAGCAAAGAGG
1681    TTTGAGGAATTTTTCATCAAAGATCAAACTAGCTTTATTCGACTCAACCACTATCCTCCA
1741    TGCCCTTACCCTGACCTTGCTCTTGGCGTCGGTCGACACAAGGACCCCGGTGCCTTGACC
1801    ATTCTTGCACAGGACGAAGTTGGAGGACTTGAAGTGAGACGTAAACGGGATCAAGAGTGG
                    SAN1A-1F
1861    ATTAGAGTGAAACCAACCCCAAATGCTTATATTATCAACATTGGTGATACTGTTCAGGTG
              SAN1A-2F
1921    CCAGTGCAATACCATATTGAAAACTGAAAATTTTAAATTTCACTTACACACTAGTTGTAA
1981    CACACTCTATCATTGGTTTAAATTAATTGAAAACTATAAAATCATGAGAGAGACATTAAA
2041    TAAGATGTGAGAACAACAACTTTCTTGTAATTTCCCATAAATTTAAACCAATGATAGAGT
2101    GTGTTAGAGAGTGTGTTGCTAGCATTTCTCTAAGGAAGAAAAACTTTCACTTGAGGATTG
2161    GGGTTGCTTGTGCAGGTTTGGAGCAATGATGCATATGAGAGTGTGGATC
```

Figure 2: Sequence of *SAN1A* Gene and Coding Sequence from Cultivar Resnik

A. *SAN1A* gene sequence, consisting of coding sequence and introns.

B. *SAN1A* coding sequence.

A.

```
1       ATGGGAGAGGTTGACCCAGCTTTCATTCAAGAACCACAACACAGGCCAAATCTCTCCACC
61      ATTCAAGCAGAAGGAATTCCCATAATTGATCTCTCTCCAATAACCAACCACACAGTTTCA
121     GACCCTTCTGCAATTGAAAGCTTGGTGAAGGAGATAGGACGTGCATGCCAGGAGTGGGGC
181     TTCTTCCAAGTAACAAACCATGGAGTGCCACTCACTCTAAGGCAAAACATTGAGAAAGCC
241     TCAAAACTGTTCTTTGCTCAGACTCTGGAGGAGAAGAGAAAGGTTAGCAGAAATGAGAGC
301     TCTCCAATGGGTTATTATGACACAGAACACACCAAGAACGTCAGGGACTGGAAAGAAGTC
361     TTTGATTTTCTAGCCAAAGACCCCACTTTCATTCCTCTCACTTCTGATGAACATGATGAT
421     CGAGTCAATCAATGGACTAATCAATCACCTCAATACCCTCCACTCTTCAGGTAATTCATC
481     TGTACTTTGCAAGGTTTTAAATACCGGTCGTGTTACGATTTTGTCGCGTTTGATGATATC
541     ACGGAAAATCACAAACAAATGTGGTTGAAGCGGTCACATTGTGATCACATACAGTTAAAA
601     AACTTTGATGTTGTGGCCCAAATCACGGTTTGGTTTAACTGTCTATGACTGCAATGTGAC
661     CAATTGCAATTGAGACCGCATCAACTGCATTTGTCCGCGATTTTTCACGATATCAACAAA
721     CACGAAGAAACCGCAATGTGACTGCAACTAATATTTAAAACCTTTAAAACAAAAGTGGAT
781     TGTGTTCAAAACTAATTGTTCAGCCCACAGTCATGACAGCTGGACAGTTTTTATAATTCT
841     ATTGATGATGTGAATACGAGCAGGGTTGTAACACAAGAGTATATTCAGGAGATGGAAAAG
901     CTGTCCTTTAAGCTGTTGGAGCTTATAGCTTTGAGCTTAGGCCTTGAAGCAAAGAGGTTT
961     GAGGAATTTTTCATCAAAGATCAAACTAGCTTTATTCGACTCAACCACTATCCTCCATGC
1021    CCTTACCCTGACCTTGCTCTTGGCGTCGGTCGACACAAGGACCCCGGTGCCTTGACCATT
1081    CTTGCACAGGACGAAGTTGGAGGACTTGAAGTGAGACGTAAACGGGATCAAGAGTGGATT
1141    AGAGTGAAACCAACCCCAAATGCTTATATTATCAACATTGGTGATACTGTTCAGGTGCCA
1201    GTGCAATACCATATTGAAAACTGAAAATTTTAAATTTCACTTACACACTAGTTGTAACAC
1261    ACTCTATCATTGGTTTAAATTAATTGAAAACTATAAAATCATGAGAGAGACATTAAATAA
1321    GATGTGAGAACAACAACTTTCTTGTAATTTCCCATAAATTTAAACCAATGATAGAGTGTG
1381    TTAGAGAGTGTGTTGCTAGCATTTCTCTAAGGAAGAAAAACTTTCACTTGAGGATTGGGG
1441    TTGCTTGTGCAGGTTTGGAGCAATGATGCATATGAGAGTGTGGATCACAGAGTGGTGGTC
1501    AACTCTGAGAAGGAAAGGCTTTCCATTCCGTTCTTCTTCTTCCCTGCACACGACACCAAA
1561    GTAAAGCCTTTGGAGGAGCTGATAAATGAGCAAAACCCTTCAAAATATAGGCCATACAAT
1621    TGGGGCAAGTTTCTTGTCCACAGAGGGAACAGCAATTTCAAGAAGCAAAATGAGGAGAAC
1681    ATTCAAATTCATCATTACAAGATAGCTTAA
```

B.

```
1       ATGGGAGAGGTTGACCCAGCTTTCATTCAAGAACCACAACACAGGCCAAATCTCTCCACC
61      ATTCAAGCAGAAGGAATTCCCATAATTGATCTCTCTCCAATAACCAACCACACAGTTTCA
121     GACCCTTCTGCAATTGAAAGCTTGGTGAAGGAGATAGGACGTGCATGCCAGGAGTGGGGC
181     TTCTTCCAAGTAACAAACCATGGAGTGCCACTCACTCTAAGGCAAAACATTGAGAAAGCC
241     TCAAAACTGTTCTTTGCTCAGACTCTGGAGGAGAAGAGAAAGGTTAGCAGAAATGAGAGC
301     TCTCCAATGGGTTATTATGACACAGAACACACCAAGAACGTCAGGGACTGGAAAGAAGTC
361     TTTGATTTTCTAGCCAAAGACCCCACTTTCATTCCTCTCACTTCTGATGAACATGATGAT
421     CGAGTCAATCAATGGACTAATCAATCACCTCAATACCCTCCACTCTTCAGGGTTGTAACA
481     CAAGAGTATATTCAGGAGATGGAAAAGCTGTCCTTTAAGCTGTTGGAGCTTATAGCTTTG
541     AGCTTAGGCCTTGAAGCAAAGAGGTTTGAGGAATTTTTCATCAAAGATCAAACTAGCTTT
601     ATTCGACTCAACCACTATCCTCCATGCCCTTACCCTGACCTTGCTCTTGGCGTCGGTCGA
661     CACAAGGACCCCGGTGCCTTGACCATTCTTGCACAGGACGAAGTTGGAGGACTTGAAGTG
721     AGACGTAAACGGGATCAAGAGTGGATTAGAGTGAAACCAACCCCAAATGCTTATATTATC
781     AACATTGGTGATACTGTTCAGGTTTGGAGCAATGATGCATATGAGAGTGTGGATCACAGA
841     GTGGTGGTCAACTCTGAGAAGGAAAGGCTTTCCATTCCGTTCTTCTTCTTCCCTGCACAC
901     GACACCAAAGTAAAGCCTTTGGAGGAGCTGATAAATGAGCAAAACCCTTCAAAATATAGG
961     CCATACAATTGGGGCAAGTTTCTTGTCCACAGAGGGAACAGCAATTTCAAGAAGCAAAAT
1021    GAGGAGAACATTCAAATTCATCATTACAAGATAGCTTAA
```

Figure 3: Sequence of *SAN1B* Gene and Coding Sequence from Cultivar Resnik

A. *SAN1B* gene sequence, consisting of coding sequence and introns.

B. *SAN1B* coding sequence.

Table I: BLAST Results for RAP-PCR Clones Isolated from Senescing Nodule Tissues

BLASTX and BLASTN results from searches carried out on June 20, 2006 are displayed for clones obtained from senescing nodule tissue by RAP-PCR. Only significant matches are displayed with their accession numbers, species of origin, and probable function. Levels of identity and similarity with the clones and the E-values of the matches are also presented. +ves, number of similar amino acids; IDs, number of identical amino acids.

A.

```
1       ATGGAAAAGGTTGACCAAGCTTTCATCCAATACCCACAACACAGGCCAAATCTCACTATC
61      ATCCAACCTGAACACATTCCCATAATAGACCTCTCTCCAATAACCAACCACACAGTTTCA
121     CATCCATCTCCCATTGAAGGCTTGGTGAAGGAGATAGGAAATGCATGCAAAGAGTGGGGA
181     TTCTTCCAAGTTATCAACCATGGGGTGTCCCTCACTCTAAGGCAAAACATTGAGAAAGCC
241     TCAAAATTGTTCTTTGCTCAGAGTTTAGAAGAGAAGAGGAAGGTTAGCAGAGATGAGAGT
301     TCTCCCATGGGTTATTATGACACAGAGCACACAAAAAACATTAGGGACTGGAAAGAAGTG
361     TTTGATTTTCTAGCCAAAGACCCTACTTTTGTTCCTCTCACTTCGGATGAACATGATAAT
421     CGTCTCACTCAATGGACTAATCCATCCCCTCAATACCCTCCACACTTCAGGTATTTATTA
481     GGGATCGAAAACATATTTGATCCTTGTTTTATTTTTTCTTGTAATTCAACTGTTAATATG
541     AACAATTAGTTAATGGGTTGTGTATAAAATGATGTGAATATTGGTAGGGATATAATTAAA
601     GAGTATGTTGAAGAGATGGAAAAGCTGTCCTTTAAGTTGATGGAGCTTATAGCTTTGAAT
661     TTAGGCCTTGAGGCAAAGAGGTTTGAGGAATTTTTCATCAAAGATCAGACTAGCTTTCTT
721     AGACTCAACTACTATCCTCCATGCCCTTACCCTCACCTTGCCCTTGGTGTTGGTCGACAT
781     AAGGACTCTGGTCCCTTAACCATTCTTGCACAAGATGAAGTTGGAGGACTTGAAGTGAGA
841     CCTAAAGCAGATCAAGACTGGATAAGAGTGAAACCTATTCCAAATGCTTATATTATCAAC
901     CTTGGTGATATGATTCAGGTACCATTTATAGATATTTGAAAGAGATGTGGTTTATTCACT
961     TCATTAATGTACTGCAGATACAAAATATATACAAGTGAACTAGCTAATCACTTCCTACAA
1021    TTATGGGATAACACAACAAAATACATGACTTGATAAGGTTGTTGCAGCTAGCAGATAATG
1081    AGAATTTAAATGCACTGATTCGAATTCAAATTTGTGCTTCTACACCATTACCATTATAAA
1141    GACTGAAAAATTGATGATTGGGGTTGTTTATGCAGGTTTGGAGCAATGATGCCTATGAGA
1201    GTGTGGAACACAGAGTGGTGGTCAACTCTGAGAAAGCAAGGTTTTCCATTCCATTCTTCC
1261    TCTTCCCTGCACATGACACTGTAGTCAAGCCTTTGGAGGAGCTAATAAATGAGCAAAACC
1321    CTTCAAAATTTAGGCCATACAAATGGGGCAAGTTTCTTGTCCACAGACTGGATAGCAATT
1381    TCAAGAAACAAAATGCGGAGAATGTTCAAACTTATCATTATAAGATAAGAGAGTAG
```

B.

```
1       ATGGAAAAGGTTGACCAAGCTTTCATCCAATACCCACAACACAGGCCAAATCTCACTATC
61      ATCCAACCTGAACACATTCCCATAATAGACCTCTCTCCAATAACCAACCACACAGTTTCA
121     CATCCATCTCCCATTGAAGGCTTGGTGAAGGAGATAGGAAATGCATGCAAAGAGTGGGGA
181     TTCTTCCAAGTTATCAACCATGGGGTGTCCCTCACTCTAAGGCAAAACATTGAGAAAGCC
241     TCAAAATTGTTCTTTGCTCAGAGTTTAGAAGAGAAGAGGAAGGTTAGCAGAGATGAGAGT
301     TCTCCCATGGGTTATTATGACACAGAGCACACAAAAAACATTAGGGACTGGAAAGAAGTG
361     TTTGATTTTCTAGCCAAAGACCCTACTTTTGTTCCTCTCACTTCGGATGAACATGATAAT
421     CGTCTCACTCAATGGACTAATCCATCCCCTCAATACCCTCCACACTTCAGGGATATAATT
481     AAAGAGTATGTTGAAGAGATGGAAAAGCTGTCCTTTAAGTTGATGGAGCTTATAGCTTTG
541     AATTTAGGCCTTGAGGCAAAGAGGTTTGAGGAATTTTTCATCAAAGATCAGACTAGCTTT
601     CTTAGACTCAACTACTATCCTCCATGCCCTTACCCTCACCTTGCCCTTGGTGTTGGTCGA
661     CATAAGGACTCTGGTCCCTTAACCATTCTTGCACAAGATGAAGTTGGAGGACTTGAAGTG
721     AGACCTAAAGCAGATCAAGACTGGATAAGAGTGAAACCTATTCCAAATGCTTATATTATC
781     AACCTTGGTGATATGATTCAGGTTTGGAGCAATGATGCCTATGAGAGTGTGGAACACAGA
841     GTGGTGGTCAACTCTGAGAAAGCAAGGTTTTCCATTCCATTCTTCCTCTTCCCTGCACAT
901     GACACTGTAGTCAAGCCTTTGGAGGAGCTAATAAATGAGCAAAACCCTTCAAAATTTAGG
961     CCATACAAATGGGGCAAGTTTCTTGTCCACAGACTGGATAGCAATTTCAAGAAACAAAAT
1021    GCGGAGAATGTTCAAACTTATCATTATAAGATAAGAGAGTAG
```

Figure 4: Sequence of *SAN1C* Gene and Coding Sequence from Cultivar Resnik

A. *SAN1C* gene sequence, consisting of coding sequence and introns.

B. *SAN1C* coding sequence.

A.

```
1       ATGGGAGAGGTCGACCCAGCTTTCATCCAATACCTACAACACAGGCCAAAGTTCTCTACC
61      ATCCAACCTGAAGACATTCCTGTGATAGATATCTCCCCAATAACCAACCACACAGCTTCA
121     GATCCATGTTCTATTGGAGGCTTAGTGAAGGAGATAAGAAATGCAAGCAAGGAGTGGGGT
181     TTCTTCCAATTATCAAACCATGCTGTGCCCCTCACTCTAAGGCAAAACATTGAGAAAGCC
241     TCAAGATTGTTCTTTGCACAGAATTTGGAGGAGAAGAGGAAGGTTAGTAAAGATGAGAAC
301     AATACCACCGGTTACCATGACAGAGAGCACCCCAAGAACATTAGGGACTGGAAAGAAGTG
361     TTTGATTTCCTAGCCAAATAATCCACTTTCATTCCTCTCACTTCTGATGAACATGATGAT
421     CGACTCACTCAGTTTGACTAATCAATCTCCTTAATACCTTCAAAATTTCAGGTAACTCAT
481     CTTCACTTTTGGAATAGTTATAGACATAAACAGCAAATTTATGGTGAAAACAGATGTATC
541     TTTTCTTTGTATATTTTAAACTACGCTACAATATATTTTTAGTCCTTGTACTTTGATCAA
601     TTTAGTGAGTCCGTGAACTTTAGAAAAACAGTGATTTTTTTGTTCTTCTTCACAGACGGT
661     TTTCATTTTCACCATAAGGAGGGACGAAAACCAATATTTTTTGAAAAGTTTAGGGACCAA
721     CTTGATCAAAGTTGAAGTTCAGATACTAAAACAAGTTTTACCGAAAGCACAGAACTGAAA
781     ACATATTTCACTCTTTAAACTATGTTCTAAATGATGTCCATATGAGCAGGGTTATAATAC
841     AAGAGTATATTCAAGAGATGGAAAAGCTGTGCTTTAAGCTGTTGGAGCTTATAGCTTTAA
901     GCTTAGGTGTTGAAGCAAAGAGGTTTGAAGAACTTTTCATCAAAGATCAAACTAGTTCTA
961     TTCTACTCAACTGCTATCCTCCATGCCCTTACCCTCACCTTGCTCATGGCGTTGGTTGAC
1021    ACAAGGACCCTGGTGCCTTGACCATTCTTGCACAAGATGAGGTTGGAGGACTTGGAGTGA
1081    AACGTAAAGCTGATCAAGAGTGAAACCAACCCCAGATGCTTATACTATCAACTTTGGTGA
1141    CATTATTCAGGTACAACTTCAGTAGCAATATTAAGACTGGAAACTTGAAATTTTATTGAC
1201    TCGTAATCTAGTTATAAGAAACAGCCACCACCTAACTCACTCAAGAAAGTTAACTTTAAG
1261    GGAAGAGGACTGCACATGTCATATAAGGACAACCATATTGTAGTTTCTGTTGTTGGTTCG
1321    TCTGGAATTGGATTTGGATTCTTTAAGCAAAGTCTTAAGCAAAGTCTCAAGTTTGAGTAT
1381    TGTGGATGAAAAAACGTGTCTAAAAGAGGAGACCTACTAATAGTGGTCTGCCCAGCTTCC
1441    CAGGAGAGATTAGTCACTAACAAAAGTATAAATACTTCACACCAATGTCATGATGACCCT
1501    AAAAAAAACATATTCTAATCTTTGAGTTTATGAACTATTTATTCCCGTAATTAGTGGAGA
1561    GAATTTTTATTTGAGGATTGGAGTTGTTTGTGCAGGTTTGGAGCAATGATGCATATGAGA
1621    GTGTGGAACACAGAGTGGTGGTCAACTCTGAGAAAGAAAGGTTCTCCATTCCATTCTTCT
1681    ACCTTTCGCATGAAACTGAGGTCAAGCCTTTGGAGAAGCTGATAAATGAGCAAAACCCTT
1741    GAAAATATAGGCCATACAAATGGGGCAAGTTTCTCACCCATAGAAAGAATACCAATTTCA
1801    AGAAACAAAATGTGGAGAACCGTCAAATTTATCATTATAAGCTTGCTTAA
```

B.

```
1       ATGGGAGAGGTCGACCCAGCTTTCATCCAATACCTACAACACAGGCCAAAGTTCTCTACC
61      ATCCAACCTGAAGACATTCCTGTGATAGATATCTCCCCAATAACCAACCACACAGCTTCA
121     GATCCATGTTCTATTGGAGGCTTAGTGAAGGAGATAAGAAATGCAAGCAAGGAGTGGGGT
181     TTCTTCCAATTATCAAACCATGCTGTGCCCCTCACTCTAAGGCAAAACATTGAGAAAGCC
241     TCAAGATTGTTCTTTGCACAGAATTTGGAGGAGAAGAGGAAGGTTAGTAAAGATGAGAAC
301     AATACCACCGGTTACCATGACAGAGAGCACCCCAAGAACATTAGGGACTGGAAAGAAGTG
361     TTTGATTTCCTAGCCAAATAATCCACTTTCATTCCTCTCACTTCTGATGAACATGATGAT
421     CGACTCACTCAGTTTGACTAATCAATCTCCTTAATACCTTCAAAATTTCAGGGTTATAAT
481     ACAAGAGTATATTCAAGAGATGGAAAAGCTGTGCTTTAAGCTGTTGGAGCTTATAGCTTT
541     AAGCTTAGGTGTTGAAGCAAAGAGGTTTGAAGAACTTTTCATCAAAGATCAAACTAGTTC
601     TATTCTACTCAACTGCTATCCTCCATGCCCTTACCCTCACCTTGCTCATGGCGTTGGTTG
661     ACACAAGGACCCTGGTGCCTTGACCATTCTTGCACAAGATGAGGTTGGAGGACTTGGAGT
721     GAAACGTAAAGCTGATCAAGAGTGAAACCAACCCCAGATGCTTATACTATCAACTTTGGT
781     GACATTATTCAGGTTTGGAGCAATGATGCATATGAGAGTGTGGAACACAGAGTGGTGGTC
841     AACTCTGAGAAAGAAAGGTTCTCCATTCCATTCTTCTACCTTTCGCATGAAACTGAGGTC
901     AAGCCTTTGGAGAAGCTGATAAATGAGCAAAACCCTTGAAAATATAGGCCATACAAATGG
961     GGCAAGTTTCTCACCCATAGAAAGAATACCAATTTCAAGAAACAAAATGTGGAGAACCGT
1021    CAAATTTATCATTATAAGCTTGCTTAA
```

APPENDIX 2:

PLANT DIOXYGENASE SEQUENCES USED IN THE CONSTRUCTION OF THE GENE TREE

Table I: Protein Sequences Used in the Construction of the Plant Dioxygenase Gene Tree

The protein sequences of the plant dioxygenases used in the construction of the gene tree are listed, along with their GenBank accession numbers, the species that they were isolated from, and the dioxygenase family to which they belong. The names of the sequences are the same names that appear on the gene tree. GB, GenBank.

Sequence Name	GB Accession	Species	Description
GmSAN1A	N/A	*Glycine max* (soybean)	senescence-associated nodulin 1A; unknown dioxygenase
GmSAN1B	N/A	*Glycine max* (soybean)	senescence-associated nodulin 1B; unknown dioxygenase
At3g19000	AAN15625	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein
At3g19010	BAB01697	*Arabidopsis thaliana*	unknown oxidoreductase; 2OG-Fe(II) oxygenase family protein
Osdiox1	AAT77035	*Oryza sativa* (rice)	putative oxidoreductase
AtGA20ox1	Q39110	*Arabidopsis thaliana*	gibberellin 20-oxidase 1
HvGA20ox	AAT94365	*Hordeum vulgaris* (barley)	gibberellin 20-oxidase
OsGA20ox2	Q8RVF5	*Oryza sativa* (rice)	gibberellin 20-oxidase 2
PsGA20ox	AAB67838	*Pisum sativum* (pea)	gibberellin 20-oxidase
PvGA20ox	AAC49756	*Phaseolus vulgaris* (common bean)	gibberellin 20-oxidase
AtGA3ox	AAC83647	*Arabidopsis thaliana*	gibberellin 3β-hydroxylase
HvGA3ox	AAT49061	*Hordeum vulgaris* (barley)	gibberellin 3β-hydroxylase
NtGA3ox	BAA89316	*Nicotiana tabacum* (tobacco)	gibberellin 3β-hydroxylase
PsGA3ox	AAC86820	*Pisum sativum* (pea)	gibberellin 3β-hydroxylase
SoGA3ox	AAN87570	*Spinacia oleracea* (spinach)	gibberellin 3β-hydroxylase
AtGA2Box	CAB41007	*Arabidopsis thaliana*	gibberellin 2β-hydroxylase
HvGA2Box	AAT49062	*Hordeum vulgare* (barley)	gibberellin 2β-hydroxylase
NtGA2Box	BAD17856	*Nicotiana tabacum* (tobacco)	gibberellin 2β-hydroxylase
PsGA2Box	AAF08609	*Pisum sativum* (pea)	gibberellin 2β- hydroxylase
VaGA2Box	BAD99511	*Vigna angularis* (adzuki bean)	gibberellin 2β-hydroxylase
AtF3H	Q95818	*Arabidopsis thaliana*	flavanone-3-hydroxylase
DianF3H	Q05964	*Dianthus caryophyllus* (clove pink)	flavanone-3-hydroxylase
GmF3H	AAT94365	*Glycine max* (soybean)	flavanone-3-hydroxylase
HvF3H	P28038	*Hordeum vulgare* (barley)	flavanone-3-hydroxylase
ZmF3H	AAA91227	*Zea mays* (corn)	flavanone-3-hydroxylase
AtFLS	BAB10452	*Arabidopsis thaliana*	flavonol synthase

Sequence Name	GB Accession	Species	Description
NsFLS	BAC10995	*Nierembergia* sp. NB17 (cupflower)	flavonol synthase
OsFLS	XP_467968	*Oryza sativa* (rice)	flavonol synthase
RhFLS	BAC66468	*Rosa* hybrid cultivar "Kardinal"	flavonol synthase
StFLS	CAA60392	*Solanum tuberosum* (potato)	flavonol synthase
AgFNS	AAX21537	*Apium graveolens* (celery)	flavone synthase I
CmFNS	AAX21538	*Conium maculatum* (hemlock)	flavone synthase I
DaucFNS	AAX21536	*Daucus carota* (carrot)	flavone synthase I
PcFNS	AAP57393	*Petroselinum crispum* (parsley)	flavone synthase I
AtLDOX	Q96323	*Arabidopsis thaliana*	leucoanthocyanidin dioxygenase
CsLDOX	AAT02642	*Citrus sinensis* (orange)	leucoanthocyanidin dioxygenase
DaucLDOX	AAD56580	*Daucus carota* (carrot)	leucoanthocyanidin dioxygenase
OsLDOX	CAA69252	*Oryza sativa* (rice)	leucoanthocyanidin dioxygenase
ZmLDOX	P41213	*Zea mays* (corn)	leucoanthocyanidin dioxygenase
AbH6H	BAA78340	*Atropa belladonna* (belladonna)	hyoscyamine 6β-hydroxylase
AnistH6H	AAQ75700	*Anisodus tanguticus*	hyoscyamine 6β-hydroxylase
DmH6H	AAQ04302	*Datura metel* (Hindu datura)	hyoscyamine 6β-hydroxylase
HnH6H	P24397	*Hyoscyamus niger* (henbane)	hyoscyamine 6β-hydroxylase
SpH6H	AAY53932	*Scopolia parviflora*	hyoscyamine 6β-hydroxylase
AtACCox	Q06588	*Arabidopsis thaliana*	ACC oxidase 1
NtACCox	BAA83466	*Nicotiana tabacum* (tobacco)	ACC oxidase
OsACCox1	Q40634	*Oryza sativa* (rice)	ACC oxidase 1
PsACCox	P31239	*Pisum sativum* (pea)	ACC oxidase
SoACCox	AAS09956	*Saccharum officinarum* (sugarcane)	ACC oxidase
StrepIPNS	Q53932	*Streptomyces cattleya*	isopenicillin N synthase

Figure 1: Alignment of Protein Sequences Used in the Construction of the Plant Dioxygenase Gene Tree

The alignment of the protein sequences used in tree construction is displayed in CLUSTAL W format. Asterisks indicate identical amino acids and colons and periods indicate similar amino acids. Prior to alignment, the plant dioxygenase sequences were truncated at their N- and C-termini to ensure that the most highly conserved portions of the sequences were used. Default parameters of CLUSTALW Version 1.8 (Thompson *et al.*, 1994) were used for the alignment, except that the BLOSUM series of matrices were selected for the pairwise and multiple alignment parameters. In addition, the gap open penalty was reduced to five. Manual editing was carried out with BioEdit Version 7.0.5.2 (Hall, 1999).

```
PsGA2Box        PLVDLSK----------PDAKT--LIVKACEDFGFFKVINHGIPLDAISQLESEAFKFFS
NtGA2Box        PLIDLSK----------PDSKN--LIVKACEEFGFFKVINHSVPTEFITKLS-EAIKFFS
AtGA2Box        PVIDMSD----------PESKH--ALVKACEDFGFFKVINHGVSAELVSVLEHETVDFFS
VaGA2Box        PTIDLSL----------ERSKLSETVVKACEEYGFFKVVNHSVAKEVISRLEEEGTAFFS
HvGA2Box        PAVDLSS----------PGAAA--DVVRACERFGFFSVVNHGVPAGVVDRLEAEAVRFFA
SAN1A           PIIDLSPITNHTVSDPSAIESLVKEIGRACQEWGFFQVTNHGVPLTLRQNIEKASKLFFA
SAN1B           PIIDLSPITNHTVSHPSPIEGLVKEIGNACKEWGFFQVINHGVSLTLRQNIEKASKLFFA
At3g19000       PTIDLSSLEDT----HHDKTAIAKEIAEACKRWGFFQVINHGLPSALRHRVEKTAAEFFN
At3g19010       PVIDLSRLDD-----PEDVQNVISEIGDACEKWGFFQVLNHGVPSDARQRVEKTVKMFFD
Osdiox1         PVIDLSPLLAAGDGDADGVDALAAEVGRASRDWGFFVVVRHGVPAEAVARAAEAQRTFFA
AtGA20ox1       PLIDLQNLLS----DPSSTLDASRLISEACKKHGFFLVVNHGISEELISDAHEYTSRFFD
PvGA20ox        PLIDLRGFLSG---DPVATMEAARMVGEACQKHGFFLVVNHGIDANLISHAHSYMDDFFE
PsGA20ox        PPIDLKAFLSD---DPKSISNACSKVNHACKKHGFFLVVNHGVDNKLIAQAHKLVDEFFC
HvGA20ox        PLIDIGGMLSG---DPRAAAEVTRLVGEACERHGFFQVVNHGIDAQLLADAHRCVDAFFT
OsGA20ox2       PVVDVGVLRDG---DAEGLRRAAAQVAAACATHGFFQVSEHGVDAALARAALDGASDFFR
AtF3H           PVISLAGIDDVD----GKRGEICRQIVEACENWGIFQVVDHGVDTNLVADMTRLARDFFA
DianF3H         PVISLAGIDGE------KRGEICRKIVEACEDWGIFQVVDHGVGDDLIADMTRLAREFFA
GmF3H           PVISLAGIDEVD----GRRREICEKIVEACENWGIFQVVDHGVDQQLVAEMTRLAKEFFA
HvF3H           PLISLHGIDG------ARRAQIRDRVAAACEDWGIFQVIDHGVDADLIADMTRLAREFFA
ZmF3H           PVVSLEGIDGP-----RRRAEIRGRVAAACEDWGIFQVVDHGVDAALVADMARLARDFFA
PcFNS           PIISLAGLDDDSD---GRRPEICRKIVKACEDWGIFQVVDHGIDSGLISEMTRLSREFFA
AgFNS           PIISLAGLDDDSN---GRRAEICRKIVEAFEEWGIFQVVDHGIDSGLISEMSRLSREFFA
CmFNS           PIISLAGLENDSD---GRRPEICRKIVEAFENWGIFQVADHGIDSALISEMSRLSREFFA
DaucFNS         PIISLAGIDDDSN---GRRPEVCRKIVEAFEDWGIFQVVDHGIDSGLIAEMSRLSREFFA
AnistH6H        PIIDLQ----------QDHHFVVQQITKACKDFGLFQVINHGFPEKLMVETMEVCKEFFA
SpH6H           PMIDLQ----------QEHHLVVQQITKACQDFGLFQVINHGFPEKLMAETMKVCKEFFA
AbH6H           PIIDLQ----------QDHLVVVQQITKACQDFGLFQVINHGLPEKLMAETMDVCKEFFA
HnH6H           PIIDLQ----------QHHHLLVQQITKACQDFGLFQVINHGFPEELMLETMEVCKEFFA
DmH6H           PIIDLQ----------QDHLLIVQQITKACQDFGLFQVINHGVPEKLMVEAMEVYKEFFA
PsACCox         PIVDMGKLNT------EDRKSTMELIKDACENWGFFECVNHGISIEMMDTVEKLTKEHYK
NtACCox         PIINLEKLNG------SERADTMEMIKDACENWGFFELVNHGIPHEVMDTVEKMTKGHYK
AtACCox         PIINLEKLNG------EERAITMEKIKDACENWGFFECVNHGISLELLDKVEKMTKEHYK
OsACCox1        PVINMELLAG------EERPAAMEQLDDACENWGFFEILNHGISTELMDEVEKMTKDHYK
SoACCox         PIIDMGLLGG------EERPAAMELLRDACESWGFFEILNHGISTELMDEVEKLTKDHYK
AtLDOX          PTIDLKNIESDDEK---IRENCIEELKKASLDWGVMHLINHGIPADLMERVKKAGEEFFS
CsLDOX          PTIDLKEIDSEDRV---EREKCREELKKAAMDWGVMHLVNHGISDDLTERVKRAGQAFFD
DaucLDOX        PTVDIADILSEDKA---VREKCYERIKDAAVEWGVMHLVNHGISNELMDRVRVAGQAFFA
ZmLDOX          PVVDISPFLDSSSQQ-QQRDECVEAVRAAAADWGVMHIAGHGIPAELMDRLRAAGTAFFA
OsLDOX          PVVDISAFDNDGD----GRHACVEAVRAAAEEWGVMHIAGHGLPGDVLGRLRAAGEAFFA
StFLS           PVIDISN-------VDDDEEKLVKEIVEASKEWGIFQVINHGIPDEVIENLQKVGKEFFE
NsFLS           PVIDLAPRVVGDEQHDHDDVEVVKQIADASKEWGIFQVINHGIPNDVIADLQKVGKEFFE
RhFLS           PIIDFSD---------PDEEKLLKQIFEASTDWGMYQIVNHDISNEAIAKLQAVGKEFFE
OsFLS           PVIDMAAP------------ESAARVAEAAAEWGLFQVVNHGVPAAAVAELQRVGREFFA
AtFLS           PVIDLSNP---------DEELVASAVVKASQEWGIFQVVNHGIPTELILRLLQVGMEFFE
NtGA3ox         PVIDLDNYINNN------NINVLEHIGQACKKWGAFQIINHNISERLLQDIELAGKSLFS
PsGA3ox         PVIDLN------------DPNASKLIGLACKTWGVYQVMNHGIPLSLLEDIQWLGQTLFS
SoGA3ox         PVINLQ------------DPQAQQLVGLACRSWGVFQVTNHGIQKSLLDDIEAAGKSLFA
AtGA3ox         PLIDLS------------DIHVATLVGHACTTWGAFQITNHGVPSRLLDDIEFLTGSLFR
HvGA3ox         PVVDMR------------DPCAAEAVALAAQDWGAFLLQGHGVPLELLARVEAAIAGMFA
StrepIPNS       PTIDISPQLFG--TDPTPRRTSRGRSTRPARGSGFFYASHHGIDVRRLQTWSNESTTMTD
                * :..                          .    *      *..
```

```
PsGA2Box        -LPQTEKEKAGPANP-----------FGYGNKRIGLNGD-IGWIEYLLLTTNQDHN---
NtGA2Box        -SPLSEKEKAGPPDP-----------FGYGNKQIGCNGD-SGWVEHILVSTNSEFNYY-
AtGA2Box        -LPKSEKTQVAG-YP-----------FGYGNSKIGRNGD-VGWVEYLLMNANHDSGSG-
VaGA2Box        -KTTSEKRQAGPANP-----------FGYGCRNIGPNGD-MGHLEYLLLHTN-PLSVS-
HvGA2Box        -STQAEKDASGPADP-----------FGYGSKRIGRNGD-MGWVEYLLLAIDRDTLS--
SAN1A           -QTLEEKRKVSRNES----------SPMGYYDTEHTKNVR--DWKEVFDFLAKDPTF-IP
SAN1B           -QSLEEKRKVSRDES----------SPMGYYDTEHTKNIR--DWKEVFDFLAKDPTF-VP
At3g19000       -LTTEEKRKVKRDEV----------NPMGYHDEEHTKNVR--DWKEIFDFFLQDSTI-VP
At3g19010       -LPMEEKIKVKRDDV----------NPVGYHDGEHTKNVK--DWKEVFDIYFKDPMV-IP
Osdiox1         -LPPERRAAVARSEA----------APMGYYASEHTKNVR--DWKEVFDLVPRQTPP-PP
AtGA20ox1       -MPLSEKQRVLRKSG----------ESVGYASSFTGRFSTKLPWKETLSFRFCDDMSRSK
PvGA20ox        -VPLTQKQRAQRKTG----------EHCGYASSFTGRFSSKLPWKETLSFQFSAEEKSST
PsGA20ox        -MQLSEKQRAQRKIG----------EHCGYANSFIGRFSSKLPWKETLSFRYSADES-CR
HvGA20ox        -MPLPEKQRALRRPG----------ESCGYASSFTGRFASKLPWKETLSFRSCPSDP--A
OsGA20ox2       -LPLAEKRRARRVPG----------TVSGYTSAHADRFASKLPWKETLSFGFH-DRAAAP
AtF3H           -LPPEDKLRFDMSGG----------KKGGFIVSSHLQGEAVQDWREIVTYFSYPVRN---
DianF3H         -LPAEEKLRFDMSGG----------KKGGFIVSSHLQGEVVQDWREIVTYFSYPTNS---
GmF3H           -LPPDEKLRFDMSGA----------KKGGFIVSSHLQGESVQDWREIVTYFSYPKRE---
HvF3H           -LPAEDKLRYDMSGG----------KKGGFIVSSHLQGEAVQDWREIVTYFSYPVKA---
ZmF3H           -LPPEDKLRFDMSGG----------KKGGFIVSSHLQGEAVQDWREIVTYFSYPVKA---
PcFNS           -LPAEEKLEYDTTGG----------KRGGFTISTVLQGDDAMDWREFVTYFSYPINA---
AgFNS           -LPAEEKLVYDTTGE----------KKGGFTISTHLQGDDVRDWREFVTYFSYPISA---
CmFNS           -LPAEEKLRYDTTGG----------KRGGFTISTHLQGDDVRDWREFVTYFSYPIDA---
DaucFNS         -LPAEEKLRYDTTGG----------KRGGFTISTHLQGDDVKDWREFVVYFSYPVDA---
AnistH6H        -LPAEENEKLQPKGKPAKFELPLEQKAKLYIEGEQLSNEEFFYWKDTLAHGCHPLDE---
SpH6H           -LPAEEKEKLQPKGEPAKFELPLEQKAKLYIEGEQLSNEEFFYWKDTLAHGCHPLDE---
AbH6H           -LPAEEKEKLQPKGEPAKFELPLEQKAKLYVEGEQLSDEAFLYWKDTLAHGCHPLDE---
HnH6H           -LPAEEKEKFKPKGEAAKFELPLEQKAKLYVEGEQLSNEEFLYWKDTLAHGCHPLDQ---
DmH6H           -LPAEEKEKFQPKGEPAKFELPLEQKAKLYVEGERRCNEEFLYWKDTLAHGCYPLHE---
PsACCox         -KCMEQRFKEMVAT-----------KGLECVQSEIDD---LDWESTFFLRHLPVSS---
NtACCox         -KCMEQRFKELVAS-----------KGLEAVQAEVTD---LDWESTFFLRHLPVSN---
AtACCox         -KCMEERFKESIKN-----------RGLDSLRSEVND---VDWESTFYLKHLPVSN---
OsACCox1        -RVREQRFLEFASKT----------LKEGCDDVNKAEK---LDWESTFFVRHLPESN---
SoACCox         -RVREQRFLEFASKT----------LKD-AQDVKAAEN---LDWESTFFVRHLPESN---
AtLDOX          -LSVEEKEKYANDQATG--------KIQGYGSKLANNASGQLEWEDYFFHLAYPEEK---
CsLDOX          -QPVEEKEKYANEQASG--------KIQGYGSKLANNASGQLEWEDYFFHLIYPEDK---
DaucLDOX        -EPIGEKEKYANDPGTG--------MIQGYGSKLANNASGQLEWEDYFFHLVYPEEK---
ZmLDOX          -LPVQDKEAYANDPAAG--------RLQGYGSRLATNTCGQREWEDYLFHLVHPDGL---
OsLDOX          -LPIAEKEAYANDPAAG--------RLQGYGSKLAANASGKREWEDYLFHLVHPDHL---
StFLS           EVPQEEKELIAKKPGAQ--------SLEGYGTSLQKEIEGKKGWVDHLFHKIWPPSA---
NsFLS           NVPQEEKELIAKTPGSN--------EIEGYGTSLQKEVEGKKGWVDHLFHKIWPPSS---
RhFLS           -LPHEEKEVYAKDPNSK--------SVEGYGTFLQKELEGKKGWVDHLFHKIWPPSA---
OsFLS           -LPQEEKARYAMDASSG--------KMEGYGSKLQKDLEGKKAWADFFFHNVAPPAM---
AtFLS           -LPETEKEAVAKPEDSL--------DIEGYRTKYQKDLEGRNAWVDHLFHRIWPPSR---
NtGA3ox         -LPMQQKLKAARSPE----------GVTGYGVARISSFFSKLMWSEGFTIVGSPLE----
PsGA3ox         -LPSHQKHKATRSPD----------GVSGYGIARISSFFPKLMWYEGFTIVGSPLD----
SoGA3ox         -LPVNQKLKAARSSC----------GVTGYGPAGISSFFPKRMWSEGFTILGSPLD----
AtGA3ox         -LPVQRKLKAARSEN----------GVSGYGVARIASFFNKKMWSEGFTVIGSPLH----
HvGA3ox         -LPASEKMRAVRRPG----------DSCGYGSPPISSFFSKCMWSEGYTFSPANLRS---
StrepIPNS       QRSTTWRSTRYNENNS-------HVRNGYYMARPGRETVESWCYLNPSFGEDHPMMKAG-
```

```
PsGA2Box          ------------FSLYGE---DIHKFRGLLKDYKCAMRN--MACEILDLMAEGLK-IQP
NtGA2Box          -----------KFASILGV---NPENIRAAVNDYVSAVKK--MACEILEMLAEGLK-IHP
AtGA2Box          -----------PLFPSLLK---SPGTFRNALEEYTTSVRK--MTFDVLEKITDGLG-IKP
VaGA2Box          -----------ERSQTIAS---DPTKFSSAVNDYIEAAKG--LTCEILDLVAEGLW-VQD
HvGA2Box          --------------KASPA---PSSALREAINAYVSAMRG--LARTVLEMVAEGLG-VSP
SAN1A             LTSDEHDDRVNQWTNQSP---QYPPLFRVVTQEYIQEMEK--LSFKLLELIALSLG--LE
SAN1B             LTSDEHDNRLTQWTNPSP---QYPPHFRDIIKEYVEEMEK--LSFKLMELIALNLG--LE
At3g19000         ASPEPEDTELRKLTNQWP---QNPSHFREVCQEYAREVEK--LAFRLLELVSISLG--LP
At3g19010         STTDPEDEGLRLVYNKWP---QSPSDFREACEVYARHAEK--LAFKLLELISLSLG--LP
Osdiox1           TTAVADGD--LVFDNKWP---DDLPGFREAMEEYGEAVEE--LAFKLLELIARSLG--LR
AtGA20ox1         SVQD-------YFCDALG---HGFQPFGKVYQEYCEAMSS--LSLKIMELLGLSLG--VK
PvGA20ox          IVKD-------YLCNTLG---QEFEQFGRVYQDYCDAMSN--LSLGIMELLGMSLG--VG
PsGA20ox          TVED-------YFVNIMG---EDFRQFGIVYQKYCEAMSN--LSLGIMELLGMSLG--VG
HvGA20ox          LVVD-------YIVATLG---EDHRRLGEVYARYCSEMSR--LSLEIMEVLGESLG--VG
OsGA20ox2         VVAD-------YFSSTLG---PDFAPMGRVYQKYCEEMKE--LSLTIMELLELSLG--VE
AtF3H             -----------RDYSRWP---DKPEGWVKVTEEYSERLMS--LACKLLEVLSEAMG--LE
DianF3H           -----------RDYTRWP---DKPEGWIKVTEEYSNKLMT--LACTLLGVLSEAMG--LE
GmF3H             -----------RDYSRWP---DTPEGWRSVTEEYSDKVMG--LACKLMEVLSEAMG--LE
HvF3H             -----------RDYGRWP---EKPAGWCAVVERYSERLMG--LSCNLMGVLSEAMG--LE
ZmF3H             -----------RDYSRWP---DKPAAWRAVVERYSEQLMA--LACRLLGVLSEAMG--LD
PcFNS             -----------RDYSRWP---KKPEGWRSTTEVYSEKLMV--LGAKLLEVLSEAMG--LE
AgFNS             -----------RDYSRWP---KKPEGWRSTTEVYSEKLMV--LGAKLLEVLSEAMG--LE
CmFNS             -----------RDCSRWP---DKPEGWRSITEVYSERLMV--LGAKLLEVLSEAMG--LE
DaucFNS           -----------RDYSRCP---DKPEGWRSVTEVYSEKLMA--LGAKLLEVLSEAMG--LE
AnistH6H          -----------ELINSWP---EKPATYREVAAKYSVEARK--LTMRMLDYICEGLG--LK
SpH6H             -----------KLVNSWP---EKPATYREVVAKYSVEVRK--LTMRMLDYICEGLE--LK
AbH6H             -----------ELVNSWP---EKPATYREVVAKYSVEVRK--LTMRILDYICEGLG--LK
HnH6H             -----------DLVNSWP---EKPAKYREVVAKYSVEVRK--LTMRMLDYICEGLG--LK
DmH6H             -----------ELLNSWP---EKPPTYRDVIAKYSVEVRK--LTMRILDYICEGLG--LK
PsACCox           -------------ISEIP---DLDDDYRKVMKEFALKLEE--LAEELLDLLCENLG--LE
NtACCox           -------------ICEVP---DLDDQYREVMRDFAQRLEK--LAEELLDLLCENLG--LE
AtACCox           -------------ISDVP---DLDDDYRTLMKDFAGKIEK--LSEELLDLLCENLG--LE
OsACCox1          -------------IADIP---DLDDDYRRLMKRFAAELET--LAERLLDLLCENLG--LE
SoACCox           -------------IAEIP---DLDDDYRRAMRQFAGELEA--LAERLLDLLCENLG--LD
AtLDOX            -----------RDLSIWP---KTPSDYIEATSEYAKCLRL--LATKVFKALSVGLG--LE
CsLDOX            -----------RDMSIWP---KTPSDYTEATSEYARQLRS--LATKILAVLSLGLG--LE
DaucLDOX          -----------ADLSIWP---KRPQDYIPATREYAKELRG--LATKLLSALSFGLG--LE
ZmLDOX            -----------ADHALWP---AYPPDYIAATRDFGRRTRD--LASTLLAILSMGLLGTDR
OsLDOX            -----------ADHSLWP---ANPPEYVPVSRDFGGRVRT--LASKLLAILSLGLG--LP
StFLS             -----------INYRYWP---KNPPSYREANEEYAKWLRK--VADGIFRSLSLGLG--LE
NsFLS             -----------INYRYWP---KNPPSYREANEVYGKKLRE--VVDKIFKSLSLGLG--LE
RhFLS             -----------INYCFWP---KNPASYREANEEYAKNLHK--VVEKLFKLLSLGLG--LE
OsFLS             -----------VNHDIWP---SHPAGYREANEEYCKHMQR--LARKLFEHLSTALG--LD
AtFLS             -----------VNHKFWP---KNPPEYIEVNEEYASHIKK--LSEKIMEWLSEGLG--LR
NtGA3ox           -----------HARQIWP---HDYQKFCDVIEEYEREMEK--LAGRLMWLMLGSLG--IT
PsGA3ox           -----------HFRELWP---QDYTRFCDIVVQYDETMKK--LAGTLMCLMLDSLG--IT
SoGA3ox           -----------HARQLWP---NNYNKFCDIIEKYQKEMNQ--LAKKLMQLVVGSLG--IS
AtGA3ox           -----------DFRKLWP---SHHLKYCEIIEEYEEHMQK--LAAKLMWFALGSLG--VE
HvGA3ox           -----------DLRKLWPKAGHDYRHFCAVMEEFHREMRV--LADKLLELFLVALG--LT
StrepIPNS         --------TPMHEVNVWPDE-ERHPDFGSFGEQYHREVSASRRCCCGASRWRRQAGESSS
                                                          :
```

```
PsGA2Box        KNVFSKLVMDK--------QSDCLFRVNHYPACPELA-----------------INGEN
NtGA2Box        RNVFSKLLMDE--------KSDSVFRLNHYPPCTEIQ-----------------QFNDN
AtGA2Box        RNTLSKLVSDQ--------NTDSILRLNHYPPCPLSNKK----------------TNGGK
VaGA2Box        NFSLSNLIKDV--------HSDSLLRINQYPPVSLTGNKNSDTSKLEPHQLHNQNNNNNN
HvGA2Box        RGALADMVTGE--------ASDQVFRVNHYPPCPLLQG-----------------LPPNC
SAN1A           AKRFEEFFIK---------DQTSFIRLNHYPPCPYPDL----------------------
SAN1B           AKRFEEFFIK---------DQTSFLRLNYYPPCPYPHL----------------------
At3g19000       GDRLTGFFN----------EQTSFLRFNHYPPCPNPEL----------------------
At3g19010       KERFHDYFK----------EQMSFFRINRYPPCPRPDL----------------------
Osdiox1         PDRLHGFFKD---------DQTTFIRLNHYPPCPSPDL----------------------
AtGA20ox1       RDYFREFFE----------ENDSIMRLNYYPPCIKPDL----------------------
PvGA20ox        KACFREFFE----------ENDSIMRLNYYPPCQKPDL----------------------
PsGA20ox        KEYFREFFE----------GNESVMRLNYYPPCKNPDL----------------------
HvGA20ox        RAHYRRFFE----------GNESIMRLNYYPPCQRPLE---------------------
OsGA20ox2       RGYYREFFA----------DSSSIMRCNYYPPCPEPER---------------------
AtF3H           KESLTNACVDM----------DQKIVVNYYPKCPQPDL----------------------
DianF3H         LEALTKACVDM----------DQKIVVNYYPKCPQPDL----------------------
GmF3H           KEGLSKACVDM----------DQKVVVNYYPKCPQPDL----------------------
HvF3H           TEALAKACVDM----------DQKVVVNFYPRCPQPDL----------------------
ZmF3H           TEALARACVDM----------DQKVVVNFYPRCPQPDL----------------------
PcFNS           KGDLTKACVDM----------EQKVLINYYPTCPQPDL----------------------
AgFNS           KEALTKACVEM----------EQKVLINYYPTCPEPDL----------------------
CmFNS           KEALTKACVNM----------EQKVLINYYPTCPEPDL----------------------
DaucFNS         KEALTEACVNM----------EQKVLINYYPTCPQPDL----------------------
AnistH6H        LG-YFDNELSQ----------IQMMLTNYYPPCPDPSS----------------------
SpH6H           LG-YFDNELSQ----------IQMMLTNYYPPCPDPSS----------------------
AbH6H           LG-YFDNELSQ----------IQMMLTNYYPPCPDPSS----------------------
HnH6H           LG-YFDNELSQ----------IQMMLTNYYPPCPDPSS----------------------
DmH6H           LG-YFDNELTQ----------IQMLLANYYPSCPDPST----------------------
PsACCox         KGYLKKAFYGS-------KGPNFGTKVSNYPPCPKPEL----------------------
NtACCox         KGYLKKIFYGT-------QGPNFGAKVSNYPPCPQPDL----------------------
AtACCox         KGYLKKVFYGS-------KRPTFGTKVSNYPPCPNPDL----------------------
OsACCox1        KGYLTKAFRGP------AGAPTFGTKVSSYPPCPRPDL----------------------
SoACCox         RGYLARAFRGPR-----TGAPTFGTKVSSYPPCPRPDL----------------------
AtLDOX          PDRLEKEVGG-------LEELLLQMKINYYPKCPQPEL----------------------
CsLDOX          EGRLEKEVGG-------LEELLLQMKINYYPKCPQPEL----------------------
DaucLDOX        EGRLEKEVGG-------MEELILQMKINYYPKCPQPEL----------------------
ZmLDOX          GDALEKALTTTTTRTAADDDLLLQLKINYYPRCPQPEL----------------------
OsLDOX          EETLERRLRG-HELAGVDDDLLLQLKINYYPRCPRPDL----------------------
StFLS           GHEMMEAAGS--------EDIVYMLKINYYPPCPRPDL----------------------
NsFLS           AHEMKEAAGG--------DDIVYLLKINYYPPCPRPDL----------------------
RhFLS           AQELKKAVGG--------DDLVYLLKINYYPPCPRPDL----------------------
OsFLS           GGAMWEAFGG--------DELVFLHKINFYPPCPEPEL----------------------
AtFLS           HEALKEGLGG--------ETIEYLMKINYYPPCPDPEL----------------------
NtGA3ox         KEEVK-WAVCP---KGESKGGSAALQLNSYPACPDPDR----------------------
PsGA3ox         KEDIK-WAGS----KAQFEKACAALQLNSYPSCPDPDH----------------------
SoGA3ox         NQDIMNWADLL---EGAN----GAMQLNSYPIRPDPNR----------------------
AtGA3ox         EKDIQ-WAGP----NSDFQGTQAAIQLNHYPKCPEPDR----------------------
HvGA3ox         GEQVA-AVESE---HKIAETMTATMHLNWYPKCPDPKR----------------------
StrepIPNS       NEVTEEDTLS----------AVSMIRYPYLDPYPEAAIKTG----------------PDG
```

```
PsGA2Box        -LIGFGEHTDPQIISILRSN-NTSGFQISLRD-GSWISVPPDHSSFFINVGDSLQVMTNG
NtGA2Box        NLIGFGEHTDPQIISVLRSN-NTSGLQILLKN-GHWISVPPDPNSFFVNVGDSLQVMTNG
AtGA2Box        NVIGFGEHTDPQIISVLRSN-NTSGLQINLND-GSWISVPPDHTSFFFNVGDSLQVMTNG
VaGA2Box        NNIGFGEHSDPQILTIMRSN-NVDGLQISTHD-GLWIPVPPDPNQFFVMVGDALQVLTNG
HvGA2Box        SVTGFGEHTDPQLVSILHSN-GTAGLQVALHD-GRWVSVPPNRDAFFVNVGDSLQVLTNG
SAN1A           -ALGVGRHKDPGALTILAQD-EVGGLEVRRKRDQEWIRVKPTPNAYIINIGDTVQVWSND
SAN1B           -ALGVGRHKDSGPLTILAQD-EVGGLEVRPKADQDWIRVKPIPNAYIINLGDMIQVWSND
At3g19000       -ALGVGRHKDGGALTVLAQD-SVGGLQVSRRSDGQWIPVKPISDALIINMGNCIQVWTND
At3g19010       -ALGVGHHKDADVISLLAQD-DVGGLQVSRRSDGVWFPIRPVPNALVINIGNCMEIWTND
Osdiox1         -ALGVGRHKDAGALTVLYQD-DVGGLDVRRRSDGEWVRVRPVPHSFIINVGDIIQVWSND
AtGA20ox1       -TLGTGPHCDPTSLTILHQD-HVNGLQVF--VENQWRSIRPNPKAFVVNIGDTFMALSND
PvGA20ox        -TLGTGPHCDPTSLTILHQD-QVGGLQVF--VDNEWHSINPNFNAFVVNIGDTFMALSNG
PsGA20ox        -AFGTGPHCDPTSLTILHQD-QVEGLQVL--VDGIWHSVVPKEDAFVVNIGDTFMALSNG
HvGA20ox        -TLGTGPHCDPTSLTILHQD-DVGGLQVH--TDGRWRSIRPRADAFVVNIGDTFMALSNG
OsGA20ox2       -TLGTGPHCDPTALTILLQD-DVGGLEVL--VDGEWRPVSPVPGAMVINIGDTFMALSNG
AtF3H           -TLGLKRHTDPGTITLLLQD-QVGGLQATRDNGKTWITVQPVEGAFVVNLGDHGHFLSNG
DianF3H         -TLGLKRHTDPGTITLLLQD-QVGGLQATRDGGKTWITVQPVPGAFVVNLGDHGHFLSNG
GmF3H           -TLGLKRHTDPGTITLLLQD-QVGGLQATRDNGKTWITVQPVEAAFVVNLGDHAHYLSNG
HvF3H           -TLGLKRHTDPGTITLLLQD-LVGGLQATRDGGKNWITVQPISGAFVVNLGDHGHFMSNG
ZmF3H           -TLGLKRHTDPGTITLLLQD-LVGGLQATRDGGRTWITVQPVEGAFVVNLGDHGHLLSNG
PcFNS           -TLGVRRHTDPGTITILLQD-MVGGLQATRDGGKTWITVQPVEGAFVVNLGDHGHYLSNG
AgFNS           -TLGVRRHTDPGTITILLQD-MVGGLQATRDGGKTWITVQPVEGAFVVNLGDHGHYLSNG
CmFNS           -TLGVRRHTDPGTITVLLQD-MVGGLQATRDGGKTWITVQPVEGAFVVNLGDHGHYLSNG
DaucFNS         -TLGVRRHTDPGTITILLQD-MVGGLQATRDGGKTWITVQPVEGAFVVNLGDHGHYLSNG
AnistH6H        -TLGSGGHYDGNLITLLQQD-LPG-LQQLIVKDDNWIAVEPIPTAFVINLGLTLKVITNE
SpH6H           -TLGSGGHYDGNLITLLQQD-LPG-LQQLIVKDGNWIAVEPNPTAFVINLGLTLKVITNE
AbH6H           -TLGSGGHYDGNLITLLQQN-LPG-LQQLIE-DAKWIAVEPIPTAFVVNLGLTLKVITNE
HnH6H           -TLGSGGHYDGNLITLLQQD-LPG-LQQLIVKDATWIAVQPIPTAFVVNLGLTLKVITNE
DmH6H           -TIGSGGHYDGNLITLLQQD-LVG-LQQLIVKDDKWIAVEPIPTAFVVNLGLTLKVMSNE
PsACCox         -IKGLRAHTDAGGIILLFQDDKVSGLQLLKD--DQWIDVPPMRHSIVINLGDQLEVITNG
NtACCox         -IKGLRAHTDAGGIILLFQDDKVSGLQLLKD--GQWIDVPPMRHSIVVNLGDQLEVITNG
AtACCox         -VKGLRAHTDAGGIILLFQDDKVSGLQLLKD--GEWVDVPPVKHSIVVNLGDQLEVITNG
OsACCox1        -VKGLRAHTDRGGIILLFQDDSVGGLQLLKD--GEWVDVPPMRHSIVVNLGDQLEVITNG
SoACCox         -VSGLRAHTDAGGIILLFQDDRVGGLQLLKD--GAWVDVPPLRHSIVVNLGDQLEVITNG
AtLDOX          -ALGVEAHTDVSALTFILHN-MVPGLQLFYEG--KWVTAKCVPDSIVMHIGDTLEILSNG
CsLDOX          -ALGVEAHTDVSALTFILHN-MVPGLQLFYKD--KWVTAKCVPNSIILHIGDTIEILSNG
DaucLDOX        -ALGVEAHTDVSALSFILHN-TVPGLQLFYGG--KWVTAKCVPDSIIVHIGDTLEILSNG
ZmLDOX          -AVGVEAHTDVSALSFILHN-GVPGLQVLHGA--RWVTARHEPGTIIVHVGDALEILSNG
OsLDOX          -AVGVEAHTDVSALSFILHN-GVPGLQVHHAG--SWVTARPEPGTIVVHVGDALEILTNG
StFLS           -ALGVVAHTDMSYITLLVPN-EVQ---VFKDG--HWYDVNYIPNAIIVHIGDQVEILSNG
NsFLS           -ALGVVAHTDMSYITILVPN-EVQGLQVFKDG--HWYDVKYIPNALIVHIGDQVEILSNG
RhFLS           -ALGVVAHTDMSALTILVPN-DVQGLQACRDG--QWYDVKYIPNALVIHIGDQMEVMSNG
OsFLS           -TLGVAPHTDMSTFTVLVPN-DVQGLQVFKDG--HWYDVKYVPDALIIHIGDQIEILSNG
AtFLS           -VVGAPDHTDVNGITLLVAN-EALGLQAFKDN--QWIDAEYTTSGIIVIIGDQFLRMSNG
NtGA3ox         -AMGLAAHTDSTILTILHQN-NTSGLQVFKEGSG-WVTVPPFPGALVVNVGDLLHILSNG
PsGA3ox         -AMGLTPHTDSTFLTILSQN-DISGLQVNREGSG-WITVPPLQGGLVVNVGDLFHILSNG
SoGA3ox         -AMGLAAHTDSTLLTILHQS-NTTGLQVFRERSG-WVTVPPISGGLVINIGDLLHILSNG
AtGA3ox         -AMGLAAHTDSTLMTILYQN-NTAGLQVFRDDVG-WVTAPPVPGSLVVNVGDLLHILTNG
HvGA3ox         -ALGLIAHTDSGFFTFVLQS-LVPGLQLFRHGPDRWVTVPAVPGAMVVNVGDLFHILTNG
StrepIPNS       TRLSFEDHLDVSMITVLSKT-EVQNLQVETVDG--WQSLPTSGENFLINCGTYLGYLTND
                   .   * *   . .:                  *          ..  *      :*
```

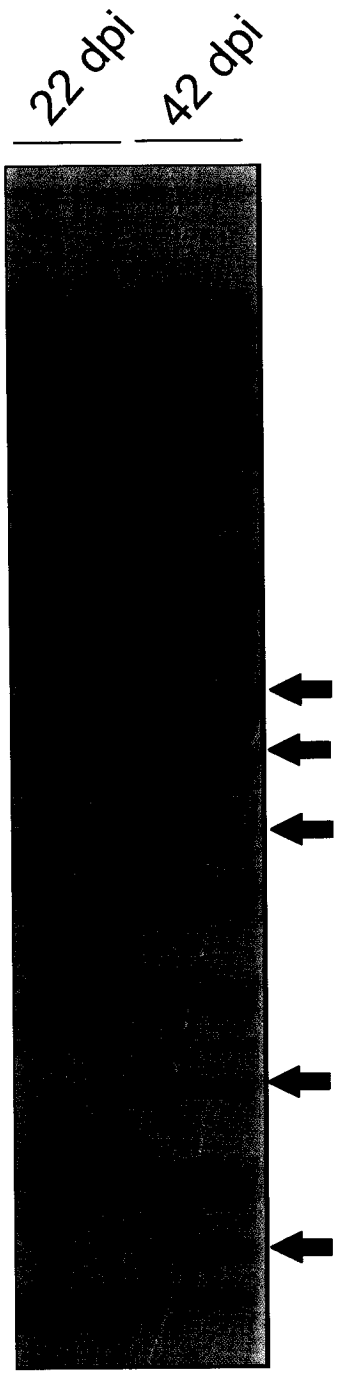
```
PsGA2Box        RFKSVRHRVLA-NGIDPRLSMIYF
NtGA2Box        KFKSVKHRVLA-NSMKSRLSMIYF
AtGA2Box        RFKSVRHRVLA-NCKKSRVSMIYF
VaGA2Box        RFRSVRHRVLT-NTTKARMSMMYF
HvGA2Box        RLRSVRHRVVAGNGLKSRVSMIYF
SAN1A           AYESVDHRVVVNSEK-ERLSIPFF
SAN1B           AYESVEHRVVVNSEK-ARFSIPFF
At3g19000       EYWSAEHRVVVNTSK-ERFSIPFF
At3g19010       KYWSAEHRVVVNTTR-ERYSIPFF
Osdiox1         RYESAEHRVAVNVEK-ERFSIPFF
AtGA20ox1       RYKSCLHRAVVNSES-ERKSLAFF
PvGA20ox        RYKSCLHRAVVNSKT-TRKSLAFF
PsGA20ox        MFKSCLHRAIVNDKI-VRKSLAFF
HvGA20ox        RYKSCLHRAVVNSRV-PRKSLAFF
OsGA20ox2       RYKSCLHRAVVNQRR-ERRSLAFF
AtF3H           RFKNADHQAVVNSNS-SRLSIATF
DianF3H         RFKNADHQAVVNSEC-SRLSIATF
GmF3H           RFKNADHQAVVNSNH-SRLSIATF
HvF3H           RFKNADHQAVVNGES-SRLSIATF
ZmF3H           RFKNADHQAVVNSEC-SRLSIATF
PcFNS           RFRNADHQAVVNSTS-SRLSIATF
AgFNS           RFRNADHQAVVNSTS-TRLSIATF
CmFNS           RFKNADHQAVVNSSS-SRLSIATF
DaucFNS         RFKNADHQAVVNSTS-SRLSIATF
AnistH6H        KFEGSIHRVVTDPTR-DRVSIATL
SpH6H           KFEGSIHRVVADPTR-DRVSIATL
AbH6H           KFEGSIHRVVTNPTR-DRVSIATL
HnH6H           KFEGSIHRVVTDPTR-DRVSIATL
DmH6H           KFEGSIHRVVTHPIR-NRISIGTL
PsACCox         KYKSVMHRVIAQTDG-ARMSIASF
NtACCox         KYKSVMHRVITQTDG-TRMSLASF
AtACCox         KYKSVEHRVLSQTDGEGRMSIASF
OsACCox1        RYKSVMHRVVAQTDG-NRMSIASF
SoACCox         RYKSVMHRVVAQPDG-NRMSIASF
AtLDOX          KYKSILHRGLVNKEK-VRISWAVF
CsLDOX          EYKSILHRGLVNKEK-VRISWAVF
DaucLDOX        KYKSILHRGLVNKEK-VRISWAVF
ZmLDOX          RYTSVLHRGLVNREA-VRISWVVF
OsLDOX          RYTSVLHRGLVSRDA-VRLSWVVF
StFLS           KYKSVYHRTTVNKYK-TRMSWPVF
NsFLS           KYKSVYHRTTVTKDK-TRMSWPVF
RhFLS           KFKAVLHRTTVSKDK-TRISWPVF
OsFLS           RYKAVLHRTTVDKDR-TRMSWPVF
AtFLS           KYKSVEHRAKMDKEK-TRISWPVF
NtGA3ox         LYPSVLHRAVVNRTR-HRLSVAYL
PsGA3ox         LYPSVLHRVLVNRTR-QRFSVAYL
SoGA3ox         RYPSVYHRAMVNRVQ-HRLSVAYL
AtGA3ox         IFPSVLHRARVNHVR-SRFSMAYL
HvGA3ox         RFHSVYHRAVVNRDS-DRISLGYF
StrepIPNS       YFPAPNHRVKYVNAE--RLSLPFF
                      *:          * *   :
```

To isolate more senescence-associated genes from soybean nodules, differential display RT-PCR was attempted as an alternative to cDNA library screening. A derivative of this method, RAP-PCR (Stratagene, La Jolla, CA), was carried out with primers A3 and B3 (Chapter 2, Table III) and several bands isolated for cloning and sequencing (Figure 1). BLAST searches were then carried out in an attempt to identify the genes that were upregulated. Of the 13 bands that were cloned and sequenced, six contained DNA species that did not correspond to rRNA. These BLAST matches are presented in Table I.

This study was only preliminary, and verification of whether the genes are really upregulated during senescence remains to be carried out by Northern analysis or an RT-PCR-based method. Genes of interest could then be further characterized in the context of root nodule senescence.

Figure 1: Comparison of 22 and 42 dpi mRNA Populations by RAP-PCR

RAP-PCR, a modified version of differential display RT-PCR, was used to compare the mRNA populations in 22 (pre-senescent) and 42 dpi (senescent) root nodules of soybean. The PCR products were loaded in duplicate on the gel. Since the goal was to isolate genes that were upregulated during nodule senescence, bands of interest were those that showed up in the products prepared from 42 dpi nodules, but not in those prepared from 22 dpi nodules. The autoradiograph displayed here is the result of an overnight exposure to BioMax MR film (Eastman Kodak Company, Rochester, NY). Black arrows highlight bands of interest that were cut out of the gel for cloning and sequencing.


22 dpi
42 dpi

APPENDIX 3:

RAP-PCR for Discovery of Genes Implicated in Nodule Senescence

Clone	BLAST Search Type	Accession Number(s) of Matches	Species	Description	Level of Similarity/ Identity	E-value
A3-1 (397)	BLASTX	NP_565435 AM13264 AAD15504 AAK96685	*Arabidopsis thaliana*	putative AAA-type ATPase	IDs: 118/125=94% +ves: 123/125=98%	1e-61
		NP_195376 CAB16835 CAB80324	*Arabidopsis thaliana*	ATPase-like protein	IDs: 116/125=92% +ves: 121/125=96%	6e-60
		AAL87170	*Oryza sativa*	putative AAA-type ATPase	IDs: 112/125=89% +ves: 122/125=97%	8e-60
A3-6 (247 bp)	BLASTX	CAA27618 S09552 P04672	*Glycine max*	nodulin 44 precursor	IDs: 106/106=100% +ves: 106/106=100% (Q: 13-330)	1e-59
					IDs: 67/71=94% +ves: 70/71=98% (Q: 13-225)	3e-35
					IDs: 19/20=95% +ves: 20/20=100% (Q: 271-330)	2e-04
B3-11/ B3-15 (376 bp/ 395 bp)	BLASTN	AFO83880	*Glycine max*	alternative oxidase precursor (*Aox1*) gene	IDs: 227/249=91%	8-9e-82

Clone	BLAST Search Type	Accession Number(s) of Matches	Species	Description	Level of Similarity/ Identity	E-value
B3-13 (402 bp)	BLASTX	NP_193133	*Arabidopsis thaliana*	unknown protein	IDs: 65/126=51% +ves: 91/126=72%	3e-34
		XP_477155 BAC84431	*Oryza sativa*	hypothetical protein	IDs: 60/127=47% +ves: 95/127=74%	1e-32
		BAD36475 BAD36226	*Oryza sativa*	SWIM zinc finger family protein-like	IDs: 46/116=39% +ves: 70/116=60%	3e-17
B3-17 (294 bp)	BLASTX	CAA70595	*Phaseolus vulgaris*	cinnamate 4-hydroxylase	IDs: 40/49=81% +ves: 47/49=95%	4e-16