University of Tennessee, Knoxville

TRACE: Tennessee Research and Creative Exchange

Doctoral Dissertations

Graduate School

12-1998

Applications of molecular markers in genetic analysis

Farshid Ghassemi

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

Recommended Citation

Ghassemi, Farshid, "Applications of molecular markers in genetic analysis. " PhD diss., University of Tennessee, 1998.
https://trace.tennessee.edu/utk_graddiss/7493

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a dissertation written by Farshid Ghassemi entitled "Applications of molecular markers in genetic analysis." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Plant, Soil and Environmental Sciences.

Peter M. Gresshoff, Major Professor

We have read this dissertation and recommend its acceptance:

Fred L. Allen, Daniel M. Roberts, Gary Stacey, Dennis R. West

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

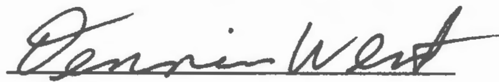
(Original signatures are on file with official student records.)

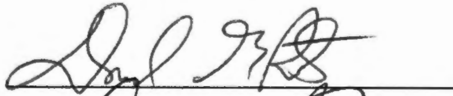
To the Graduate Council:

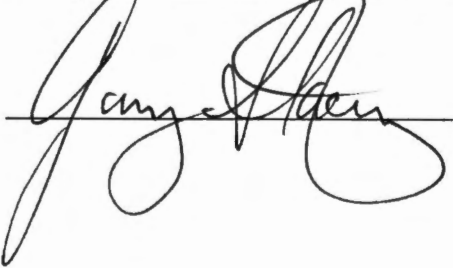
I am submitting herewith a dissertation written by Farshid Ghassemi entitled "Applications of Molecular Markers in Genetic Analysis." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Plant and Soil Science.

Peter M. Gresshoff, Major Professor

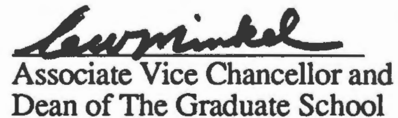
We have read this dissertation and recommend its acceptance:







Accepted for the Council:



Associate Vice Chancellor and Dean of The Graduate School

Applications of Molecular Markers in Genetic Analysis

A Dissertation
Presented for the
Doctor of Philosophy
Degree

The University of Tennessee, Knoxville

Farshid Ghassemi

December 1998

Ag-VetMed
Thesis
98b
.G48

Dedication

To the "past", "present", and "future".

In the memory of my mother, Jalile Ahmadi, and my father, Ahmad-Ali Ghassemi, who in the "past" continuously encouraged and supported me to come here to pursue my Ph. D. degree but both passed away during my stay here and could not witness my graduation.

To my wife, Dr. Nazila Hosseini Tehrani, who is supporting and bringing me peace of mind at "present" by taking immense responsibilities for raising our children. I am proud of my wife for being awarded her recent Ph. D. degree. She is truly a role model for a wonderful mother and wife.

To my brothers, Jamshid and Farhad Ghassemi, and their families who have never stopped their enthusiasm and support for me.

To my lovely children, Pedram, Parham, and Pegah Ghassemi, who even contributed in writing and editing of my dissertation! I encourage them to follow a career in their "future" to serve effectively human being, as does a scientist.

بنام خدا

این پایان نامه دکتری را به پدر گرامی ام، مرحوم احمدعلی قاسمی، و مادر بزرگوارم، مرحومه جلیله احمدی، اهدا می نمایم. روحشان شاد باد.

همچنین این رساله را به همسر عزیزم، دکتر نازیلا حسینی تهرانی، و فرزندان برومندم، پدرام، پرهام، و پگاه قاسمی تقدیم می نمایم.

Acknowledgments

I would like to thank many people who contributed greatly in this work by providing me financial and most importantly emotional supports. I would like to thank Dr. Peter M. Gresshoff, Professor and Chair, not only for financial support but also for his active mind and inspiring scientific and social discussions throughout my stay under his supervision in the PMG lab, Racheff Chair of Excellence. I am also honored and grateful of my committee members Drs. Fred L. Allen, Daniel M. Roberts, Gary Stacey, and Dennis R. West for their time, attention, and patience as well as valuable comments on my dissertation. I have my special thank for Dr. Lloyd M. Callahan for my association with him in DNA fingerprinting of grasses. I was really inspired by his knowledge and character integrity.

I would like to acknowledge the Ministry of Culture and Higher Education of Islamic Republic of Iran for a four-year scholarship. I am thankful of the Racheff Chair of Excellence and Tennessee Soybean Promotion Board for additional financial supports. I also appreciate Dr. Fred L. Allen for his timely financial support. I acknowledge the University of Tennessee for awarding me the Scholarly Activities Research Incentive Fund (SARIF) fellowship. Finally, I would like to thank Drs. W. Franklin Harris, Patricia Cox, and particularly Otto J. Schwarz who arranged a teaching assistantship at the end of my way to graduation.

I am very grateful of Ms. Deborah Landau-Ellis, The University of Tennessee at Knoxville for providing F_2 seeds of Bragg and *G. soja* cross as well as for her wonderful personality and valuable help which never came short and set an extraordinary example for me.

I would like to thank Dr. Gordon K. Lark and Kevin Chase, University of Utah, for kindly providing me the RIL seeds and analysis of segregation data in RILs and to thank Dr. James Orf, University of Minnesota, for providing additional RIL seeds. I appreciate Dr. Norman Weeden, Cornell University, for providing me with update on pea genetic linkage map.

My special acknowledgment goes to Mr. Mostafa Rahbar a fellow Iranian at the ELI, UTK, whose help was a turning point for me to come to the USA. I am also grateful of many Iranians, particularly Drs. Bahman Yazdi-Samadi and Abbas-Ali Zali, who contributed significantly to my success in pursuing my Ph. D. degree.

There are so many people who provided a friendly environment that made working in the PMG lab enjoyable. I thank all of them for their help and friendship, especially, I enjoyed very much my association with Dr. Raymond McDonnell whose friendship and help was not limited only at work. I appreciate Drs. Jiri Stiller, Ru-Ju Chian, and Qunyi Jiang for their great friendship and personality. I will never forget the memorable social events in the PMG lab which organized by Dr. Peter M. Gresshoff.

Finally, I would like to thank Ms. Janice Crocket, Principal Secretary at PMG, for her friendly and charming personality and for her continuous help toward me and other PMG staff which set an example as a perfect friend and secretary.

ABSTRACT

Restriction fragment length polymorphism (RFLP) and microsatellite molecular markers were used to map two soybean nodulation genes, *enod2* and leghemoglobin (*lbc3*). In addition, a high annealing temperature DNA amplification fingerprinting (DAF) method was developed for DNA fingerprinting of soybean cyst nematode (SCN), *Mychorrizae*, aphid, centipedegrass, and bermudagrass samples.

Recombinant inbred lines (RILs) as well as an F_2 segregating population of soybean *Glycine max* (L. Merr) facilitated the mapping of two expressed sequence tags (EST) involved in early nodulation and subsequent nitrogen fixation in soybean. For the early nodulin gene *enod2*, the parents of RILs, Minsoy and Noir1, showed a polymorphism (5.5 vs. 5.9 kb) after *Eco*RV digestion. RFLP patterns of 42 RILs were analyzed using the MAPMAKER program linking *enod2* to the seed coat color gene, *I*, with a distance of 11.1 cM on linkage group U3 of RIL map. *Enod2* and *I* are located close to *Rhg*4 , a soybean cyst nematode (SCN) resistance gene, and a locus for seed coat hardness. The molecular marker pA110 and seed coat color were used to integrate *enod2* on an F_2 segregating population (72 plants) generated from a cross between cultivar Bragg and *G. soja* (Sieb and Zucc), PI468.397. *Enod2* was mapped in the same order as on the RIL map but 18.5 cM from the I locus. A microsatellite from the 5' region of *enod2B* was mapped in the same position, demonstrating that *enod2B* and not *enod2A* was mapped. An RFLP for *lbc3* (leghemoglobin) segregated independently from *enod2* and the *nts*-1 supernodulating locus suggesting that in soybean, symbiotically significant loci (including *rj1*, *Rj2*, and *rj6*) are not clustered.

To overcome potential problems caused by mismatch priming and secondary DNA structure and taking advantage of high primer-template ratios used in DAF reactions, annealing temperature of 55°C were used with single short arbitrary oligonucleotide as well as mini-hairpin primers to provide high resolution DNA profiles of soybean. Initially, high annealing temperatures for three arbitrary octamer primers in polymerase chain reaction (PCR) were tested for DNA fingerprinting of two soybean cultivars, Minsoy and Noir1. Fifteen PCR programs differing in levels of annealing temperature (47, 55, and 60°C), denaturation, annealing, and extension time (30, 60, and 120 second), and presence/ absence of extension step (+/- 72°C) were tested. The number of bands (amplification products) ranged from 7 (Program 10) to 51 (Program 3). The average ramping

temperature for heating and cooling were calculated 1.42 and 1.27 sec/°C, respectively. Intensity of the silver-stained bands in a 10% polyacrylamide gel was high for the most PCR programs. Program 15, DAF-15, (95°C/30 sec, 55°C/120, and 72°C/30 sec) generated a complex DNA fingerprinting profiles for tested primers in Minsoy and Noir1. These profiles contained an average of 42 sharp and highly intense bands using both octamer primers 8-4 and 8-8 for DNA amplification. Using high annealing temperature increased stringency of primer-template annealing, avoided potential mismatching and hybrid molecule formation, and consequently improved reproducibility of DNA fingerprinting.

Newly-developed high annealing temperature DAF was used successfully and detected markers linked to the *enod2* gene and analyzed DNA fingerprinting of soybean cyst nematode (SCN), *Mycorrhizae*, aphid, centipedegrass, and bermudagrass samples. RFLP patterns of 41 homozygous F_2 individuals for *enod2* gene were set into two bulks of 26 and 15 with RFLP patterns identical to their parental patterns Bragg and *G. soja*, respectively. Screening of the bulks B and S with 31 primers resulted in detection of four polymorphic bands using primers HpC29 and HpC30 and DAF-15 program. Due to low number of polymorphic bands in the B and S bulks, sub-pools were generated and screened. B1 and S1 sub-pools were tested with total 196 primers of which 32 were used for screening of sub-pools B3 versus S2. Primers Hp30, HpC22 and HpC30 generated 1, 1 and 4 polymorphic markers, respectively, in the B3 vs. S2. The major screening was focused on the B1 versus S1 sub-pools which resulted in screening of 196 mini-hairpin and unstructured primers of which a set of 9 primers detected 20 polymorphic bands. Primer HpD25 generated polymorphic bands with 920B1, 320B1, 220S1, and 185B1 base pairs which were reliable and reproducible. These bands are promising bands for further analysis such as cloning and generating SCAR markers in the region of genome containing the *enod2* gene.

Key Words: nitrogen fixation, RFLP, recombinant inbred lines, integration mapping, annealing temperature, PCR, DNA fingerprinting, arbitrary primers, soybean.

Abbreviations: RFLP, restriction fragment length polymorphism; SCN, soybean cyst nematode; RILs, recombinant inbred lines; *CHS*, chalcone synthase; QTL, quantitative trait locus; *lbc3*, leghemoglobin gene; DAF, DNA amplification fingerprinting; PCR, polymerase chain reaction

Table of Contents

Part	Page
1. Introduction to Legume-Bacterium Symbiosis	1
The Preinfection Stage	2
The infection and Nodule Formation Stage	6
Nodule Function and N_2 Fixation	7
Nitrate Inhibition of Nodulation	9
Autoregulation of Nodulation	10
Legume Nodulation Mutants	10
The nodulin Concept	12
The Genetics of Nodulins	18
Molecular Mapping	21
Molecular Genetic Linkage Map in Soybean	23
Molecular Genetic Linkage Maps in Other Legumes	28
Outline of Dissertation	34
References	37
2. Molecular Markers	63
Types of Molecular Markers	65
Class 1. Hybridization-Based Molecular Markers	65
RFLP	65
Mini and Microsatellite Probes	66
Class 2. PCR-Based Molecular Markers, Arbitrary Primers	67
RAPD Marker	67
Randomly Amplified Microsatellite Polymorphism	68
Advantages and Disadvantages	69
AP-PCR Markers	70
DAF Markers	72
DAF- Mini-Hairpin Primers	74
DAF- ASAP	76
DAF- TechMAAP	78

Part	Page

Amplified Fragment Length Polymorphism (AFLP) 79
Arbitrary Sequence Oligonucleotide Fingerprinting (ASOF) 80
Class 3. PCR-Based Molecular Markers, Known Sequence 81
Cleaved Amplified Polymorphic Sequence (CAPS) 81
Genetic Bit Analysis (GBA) 82
Microsatellite or Simple Sequence repeats (SSR) Markers 83
References 83

3. The Early *enod2* and the Leghemoglobin (lbc3) Genes Segregate Independently from Other Known Soybean Symbiotic Genes. 95

Abstract 96
Introduction 96
Materials and Methods 99
Results 100
Discussion 110
References 114

4. High Annealing Temperature Using Short Arbitrary Primers Provides Robust and High Resolution DNA Amplification Fingerprinting 121

Abstract 122
Introduction 122
DNA Factor 123
Thermocycler Factor 123
Primer Factor 124
Magnesium Factor 124
Other Factors 124
Annealing Temperature 125
Materials and Methods 126
DNA Quantification 127
DNA Amplification 127
PCR Programs 127

Part	**Page**

Gel Separation 128
Results and Discussion 130
New Programs with Other Primers 138
References 143

5. Applications of DNA Amplification Fingerprinting 148

Bulked Segregant Analysis (BSA) 149
Materials and Methods 151
Results and Discussion 152
DNA Fingerprinting 175
Race Distinction of SCN 175
DNA Fingerprinting of Mycorrhizal Fungi 180
DNA Amplification Fingerprinting of Aphid 182
DNA Amplification Fingerprinting of Bermudagrass 182
Genetic Stability in Centipedegrass 187
Concluding Remarks 190
References 190

6. Conclusions 195

Appendices 199

Appendix A 200
Figure A-1. Complete sequence of *enod2A* genomic clone 200
Figure A-2. Complete sequence of *enod2B* genomic clone 202

Appendix B 204
Table B-1. Segregation data for seed coat color and RFLPs for *enod2* and pA110 in F_2 population 204

Appendix C 206
Table C-1. List of 10-, 9-, 8-, 7-, 6-, 5, 4, and 3-mer primers. 206

Part **Page**

Table C-2. List of series A, B, and C for octamer (8-mer) and mini-hairpin primers. 209

Table C-3. List of mini-hairpin primers. 213

Table C-4. List of octamer primers obtained from Oak Ridge National Laboratory; OR primers. 214

Table C-5. List of series A to N decamer (10-mer) primers obtained from Operon Technology Inc. 215

Vita 217

List of Figures

Figure **Page**

1-1. Structure of Nod factors of different *Rhizobia* 5

1-2. Major differences between indeterminate and determinate nodules 8

1-3. Schematic representation of developmental and metabolic events associated with nodulin expression 13

1-4. Promoter elements of the soybean *enod2B* gene 19

1-5. Schematic representation of the *cis*-regulatory elements in the soybean *lbc3* promoter 20

1-6. Soybean genetic linkage map constructed in the Uniuversity of Utah 24

1-7. Soybean genetic linkage map constructed in the Northern Arizona University 25

1-8. A genetic linkage map of pea (*Pisum sativum*) 29

1-9. A genetic linkage map of common bean 32

1-10. A genetic linkage map of *Lotus japonicus* 33

1-11. Integration mapping approach in soybean 36

2-1. General DAF reaction 73

2-2. A DAF profile for human and grass samples 75

2-3. Structure of linear and mini-hairpin primers 77

3-1. Autoradiograph of Southern hybridization for *enod2*, pA110, and *lbc3* probes 101

Figure	Page

3-2. Segregation of restriction fragment length polymorphism (RFLP) for *enod2* 103

3-3. Linkage maps of *enod2* and *lbc3* genes 107

3-4. Mapping of the *enod2* by an AT microsatellite 109

4-1. Polyacrylamide gels for 15 PCR programs using 8-4 primer for soybean DNA template cultivar Minsoy 132

4-2. Polyacrylamide gels for 15 PCR programs using 8-4 primer for soybean DNA template cultivar Noir1 133

4-3. Polyacrylamide gels for original DAF, P# 3, and J programs 139

4-4. Polyacrylamide gels for Programs 8 and original 3 140

4-5. Polyacrylamide gels for Programs 11 and 14 141

4-6. Polyacrylamide gels for Programs 9 and 15 142

5-1. Silver-stained 10% polyacrylamide DAF gels for two bulks B and S using 55°C annealing temperature DAF 166

5-2. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1, S1, B3 and S2 using 55°C annealing temperature DAF and Hp10 and Hp30 primers 169

5-3. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1, S1, B3 and S2 using 55°C annealing temperature DAF and HpC22 and HpC30 primers 170

5-4. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1 and S1 using 55°C annealing temperature DAF and octamer primers series A and B 172

Figure	Page

5-5. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1 and S1 using 55°C annealing temperature DAF and octamer primers seri B 173

5-6. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1 and S1 using 55°C annealing temperature DAF and mini-hairpin primers 174

5-7. Silver-stained 5% polyacrylamide DAF gels for 3, 14, and 5 SCN races using original DAF program and 8-mer primers 177

5-8. Silver-stained 5% polyacrylamide DAF gels for 3, 14, and 5 SCN races using original DAF program and 8- and 10-mer primers 178

5-9. Silver-stained 5% polyacrylamide DAF gels for 3, 14, and 5 SCN races using original DAF program and octamer primers 179

5-10. Silver-stained 10% polyacrylamide gels of DNA amplification fingerprinting using original DAF program and 8-4 primer for mycorrhizal fungi 181

5-11. Silver-stained 10% polyacrylamide gels of DNA amplification fingerprinting using original DAF program and OR30 primer for 1-19 aphid samples 183

5-12. Silver-stained 10% polyacrylamide gels of DAF for unknown bermudagrass and standard cultivars using 55°C annealing temperature DAF 185

5-13. Silver-stained 10% polyacrylamide gels of DNA amplification fingerprinting using 55°C annealing temperature DAF for centipedegrass samples 188

A-1. Complete sequence of *enod2A* genomic clone 200

A-2. Complete sequence of *enod2B* genomic clone 202

List of Tables

Table **Page**

3-1. Five independent *Chi* Square tests 106

3-2. Linkage and genetic data for soybean symbiosis and symbiosis-related genes 113

4-1. Time course (seconds) of cycles for new DAF programs in the MJR thermocycler 131

4-2. Effect of annealing temperature on generation of total bands in 15 programs 135

4-3. Ramping of cooling and heating for the high annealing temperature DAF programs used in the MJR machine 136

5-1. RFLP patterns for *enod2* in a F_2 population 153

5-2. Results of primary screening of primers for the *enod2* gene in bulked segregant analysis (BSA) using original DAF program 155

5-3. Plant # of F_2 individuals in sub-pools of B(26) and S(15) bulks 167

5-4. Bulks and sub-bulks of F_2 individuals for the soybean *enod2* gene 168

5-5. Size of polymorphic bands (base pairs) resulted from screening of BSA for the *enod2* gene 171

5-6. Results of polymorphic bands (base pairs) detected in soybean cyst nematode (SCN) races 3, 14, and 5 176

B-1. Segregation patterns for *enod2*, seed coat color, and molecular marker pA110 in a soybean F_2 population 204

C-1. List of 10- to 3-mer primers 206

Table	Page
C-2. List of series A, B, and C for octamer and mini-hairpin primers	209
C-3. List of mini-hairpin primers	213
C-4. List of Oak Ridge (OR) primers	214
C-5. List of series A to N decamer (Operon) primers	215

Part 1

Introduction to Legume-Bacterium Symbiosis

Nitrogen (N) is one of the limiting factor in plant growth and development affecting crop production, food, and ultimately the fate of the human population and societies. Traditionally, agronomically significant N come from soil fertility (mineralization, deposits, etc.), organic manure, rain water (especially after thunderstorms), fertilizer application and symbioses. Atmospheric nitrogen can be converted to ammonia through the energy-demanding Haber-Bosch process for nitrogen fertilizer production. Nitrogen can also be fixed naturally through either non-biological (like photochemical) reactions and lightening or biological (like microbial nitrogenase reduction) mechanisms. Biological nitrogen fixation is carried out by free living, associated, or symbiotic diazotrophic eubacteria. The association of plants and a variety of nitrogen-fixing organisms is responsible for reducing 120 million tons of atmospheric N_2 to ammonia each year. Symbiotic N_2 fixation is the consequence of plant-microbe interactions leading to formation of specialized plant organ, the nodule, either in roots or stems.

Nodule formation occurs in plants of the family *Leguminoseae* interacting symbiotically with nodulating bacteria of the genera *Rhizobium*, *Bradyrhizobium*, *Azorhizobium*, and *Sinorhizobium*. Other than legume plants, there is a symbiotic relationship between *Rhizobium* and the plant genus *Parasponia* (Trinick, 1979), as well as between filamentous *Frankia* (Baker and Mullin, 1992; Simonet *et al.*, 1990) and members of the genera *Alnus*, *Casuarina*, and *Elaeagnus* (Mullin *et al.*, 1990).

Legume-bacterium symbiosis provides a model to investigate the mechanisms of interaction between the plant and the bacteria. In this symbiosis, atmospheric nitrogen is reduced by the bacterial nitrogenase enzyme to ammonia, a nitrogen source for the plant, while the plant provides carbon compound in the form of sugars and carboxylic acids such as malate and succinate for the bacteria (Mellor and Werner, 1990). Success of symbiosis depends on compatibility of the bacterium and the plant and development of several steps involving preinfection, infection and nodule initiation, and nodule function (Rolfe and Gresshoff, 1988).

The Preinfection Stage

The onset of plant-bacteria interactions begins with exudation of plant flavonoid-type compounds (e.g., luteolin, genestein, naringenin, and daidzein) in the rhizosphere as chemoattractants (Dixon and Lamb, 1990; Sutherland *et al.*, 1989; Halverston and

Stacey, 1986; Lynn and Chang, 1990; Peters and Verma, 1990) inducing bacterial nodulation (*nod*) genes (Freiberg *et al.*, 1997; Orgambide *et al.*, 1996; Philip-Hollingsworth *et al.*, 1995; Caetano-Anollés *et al.*, 1988; Zaat *et al.*, 1988, Kosslak *et al.*, 1987). After chemotaxis, rhizobia attach to root hairs through two steps (Dazzo *et al.* 1984; Smit *et al.*, 1987). It was believed that plant lectins bind to bacterial surface polysaccharides facilitating the attachment (Smit *et al.*, 1987; Hirsch, 1992). However, recent studies indicated that this was not the case and flavonoids such as daidzein, coumestrol, and naringenin were dominant compounds in exudates of infected common bean (Bolanos-Vasquez and Werner, 1997). the concentration of these flavonoids were markedly enhanced after inoculation by bacteria. Attachment happens either via cellulose fibrils or fimbriae (Vesper and Bauer, 1986). Entry of bacteria occurs at the growing root hair tip because the cell wall is thinner and expanding. Root hair deformation and curling occur within 6-18 hour in response to bacterial *nod* gene products (Kondorosi, 1992; Hollingsworth *et al.*, 1990; Lerouge, *et al.*, 1990; Schmidt *et al.*, 1988).

Bacterial *nod* genes in fast-growing (*Rhizobium* spp.) and in slow-growing (*Bradyrhizobium* spp) rhizobia are located on a large plasmid, *sym* plasmid, and chromosome, respectively. Flavonoids, isoflavonoids, and chalcones are the inducers of *nod* genes in fast-growing species, while in slow-growing species, *nod* genes are induced by isoflavones (Sanjuan *et al.*, 1994: Stacey *et al.*, 1993; Kondorosi, 1992; Györgypal *et al.*, 1991; Sadowsky *et al.*, 1991; Peters and Verma, 1990).

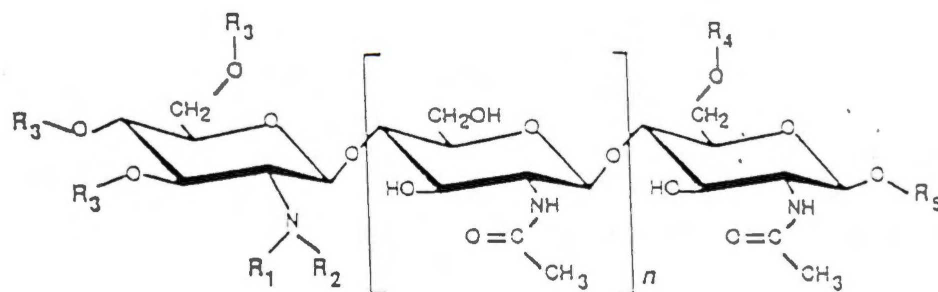
Previously, bacterial nodulation genes (*nod*, *nol*, and *noe*) were characterized into two groups: (a) the "common" *nodABCI J* genes which were thought to be essential for nodulation to occur, and (b) host specific genes (hsn) which were found only in specific strains determining the host range of bacterium. Mutations in the *nodABC* completely arrested nodulation including root hair curling, cortical cell division, and infection threads formation (Long, 1989). In contrast, mutation in *hsn* genes, such as the *nodI* and *nodJ* show only a slight delay or lessening in nodulation (Kondorosi *et al.*, 1985).

The "common" *nod* genes have been detected in all rhizobia and are functionally comparable from one *Rhizobium* species to another (see Kondorosi, 1991; Fisher and Long, 1992). The regulatory *nodD* gene product found in all bacteria acts together with the flavonoids inducing other nodulation genes (Franssen *et al.*, 1992; Long 1989). Host specificity is controlled by plant flavonoids, bacterial lipo-chitin oligosaccharides,

LCOs (the product of common and *hsn/nod* genes), and possibly plant lectins (Spaink *et al.*, 1995; Bloemberg *et al.*, 1995; Diaz *et al.*, 1995). The eventual signal for nodule initiation is the result of an interplay of "common" and host-specific nodulation genes (Stacey, 1995) producing glycolipid molecules which initiate cell division and infection cascades in the host plant by yet unexplained mechanisms. These signal molecules are commonly called Nod factors (Figure 1-1) involved in root hair deformation (*Had*), hair curling (*Hac*), and onset of cortical cell division (*Ccd*) which establish nodule initiation (*Noi*).

Recently, it has been shown that NodC is responsible for biosynthesis of rhizobial lipochitin oligosaccharides, LCOs, (Kamst *et al.*, 1997). Different *Rhizobium* species produce different length of the LCO which can influence their activities on host plant. This indicates that NodC contributes to the host specificity. Furthermore, NodA proteins of *R. meliloti* and *R. tropici* determine the N-acylation of nod factors by different fatty acids (Debelle *et al.*, 1996). This allelic variation of common nodA gene is an indication of host range specificity. In addition, NodB is a chitooligosaccharide deacetylase which is required for infection and nodulation of alfalfa (Roche *et al.*, 1996). The recent studies suggest that variation in *nodABC* is a genetic mechanism in signaling variation which controls the host range.

The Nod factors was first characterized in *Rhizobium meliloti* (Lerouge et al., 1990). The structure of this molecule consists of a chitin-like glucosamine backbone which is either acylated or sulfated. Host specificity of the Nod factors in different rhizobia is controlled by the number of glucosamine residues differing in the length and extent of the saturation of fatty acid, acylation and sulfation of side chain. For example, Nod factor for *Bradyrhizobium japonicum*, a soybean symbiont, contains a pentaglucosamine backbone, but not sulfated, and a 2-O-methyl fucose on the reducing end (Stacey *et al.*, 1995). The Nod factor molecule is controlled by *nodABC* and other specific nod genes (Spaink, 1992; Hirsch, 1992). It was postulated that the action of nod factor is involved in association with a lectin receptor (Lugtenberg *et al.*, 1991). The specificity of the action depends on the structure of the individual Nod factor including length and saturation of the acyl side chain and the structure of the side groups. In addition, other rhizobial molecules which are important in the nodulation process are lipooligosaccharides (LPS), exopolysaccharides (EPS), and capsular polysaccharides



Species	R_1	R_2	R_3	R_4	R_5	n	References
Rhizobium meliloti	-H	-C16:2 (2,9)[a] or -C16:3 (2,4,9)	$-COCH_3$ (C-6)[b] or -H	$-SO_3H$	-H	1,2,3	Lerouge et al. (1990) Schultze et al. (1992)
Rhizobium leguminosarum bv *viciae*	-H	-C18:4 (2,4,6,11) or -C18:1 (11)	$-COCH_3$ (C-6)	-H or $-COCH_3$[c]	-H	2,3	Spaink et al. (1991) Firmin et al. (1993)
Bradyrhizobium japonicum	-H	-C18:1 (9), -C18:1 (9,Me), -C16:1 (9), or -C16:0	$-COCH_3$ (O-6) or -H	2-O-Methylfucosyl group	-H	3	Sanjuan et al. (1992) Carlson et al. (1993)
Bradyrhizobium elkanii	-H or Me	-C18:1	$-COCH_3$ (C-6), -H, or Cb[d,e]	2-O-Methylfucosyl or fucosyl group	-H or Gro[f]	2,3	Carlson et al. (1993)
Rhizobium sp strain NGR234	Me	-C18:1 or -C16:0	Cb (O-3 and/or O-4)[b] or -H	Sulfated or acetylated 2-O-methylfucosyl group	-H	3	Price et al. (1992)
Azorhizobium caulinodans strain ORS571	Me	-C18:1 or -C18:0	Cb (O-6) or -H	D-Arabinosyl or -H	-H	2,3	Mergaert et al. (1993)
Rhizobium fredii	-H	-C18:1	-H	Fucosyl or 2-*O*-methylfucosyl group	-H	1,2,3	Bec-Ferte et al. (1994)
Rhizobium tropici	Me	-C18:1	-H	$-SO_3H$	-H	3	Poupot et al. (1993)

Structure of Nod Factors of Different Rhizobia.

The number of the *N*-acetylglucosamine residues can vary between three and five. The substitutions at positions R_1, R_2, R_3, R_4, and R_5 among the different rhizobia are indicated.

[a] The numbers in parentheses indicate the positions of the double bonds in the fatty acids.

[b] O-n indicates the position of the substitution on the *N*-acylglucosamine residue.

[c] This substitution is present only in Nod factors of *R. leguminosarum* bv *viciae* strain TOM.

[d] Cb indicates carbamyl group.

[e] The position of the carbamyl group could be 0-3, 0-4, or 0-6.

[f] Gro indicates glyceryl group.

Figure 1-1. Structure of Nod factors of different *Rhizobia*. (This figure is reproduced from Mylona *et al.*, 1995).

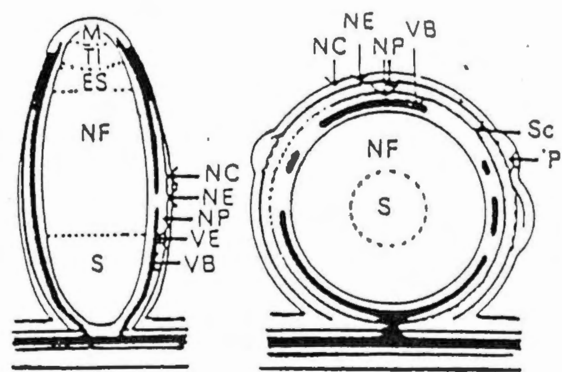
(CPS). These cell surface molecules are involved in the bacteria-legume signaling and in the infection process.

The Infection and Nodule Formation Stage

After root hair curling, rhizobia enter the root hair through an infection thread. Successful formation of infection thread appears to require an involvement of *Rhizobium nod*, *ndv*, *lps*, and *exo* genes (Dénarié and Cullimore, 1993). The *exo* and *lps* gene products are involved in biosynthesis of exopolysaccharides (EPS) and lipopolysaccharides (LPS), respectively (Finan *et al.*, 1985; Dylan *et al.*, 1986). Bacteria *ndv* mutants exhibit decreased mobility and increased phage sensitivity (Dylan *et al.*, 1986). Rhizobial acidic EPS is essential for N_2 fixation in indeterminate nodules, e.g., alfalfa or pea, but is not important in determinate nodules, e.g., soybean, *Phaseolus*, or *Lotus*. (Borthakur *et al.*, 1986). LPS synthesis also affects nodule development except in *R. meliloti* deficient in LPS which induce N_2-fixing nodules on alfalfa and white clover (Blauenfeldt *et al.*, 1994; Caetano-Anollés and Gresshoff 1992; Clover *et al.*, 1989). Furthermore, *lps* mutants of *R. leguminosarum bv. viciae* induce nodules on pea or vetch although they are Fix^- (non-nitrogen fixing) because the bacteria are not released from the infection threads (Brewin *et al.*, 1990; Priefer, 1989). Different responses to *exo* and *lps* in determinate and indeterminate nodules have not been explained yet.

The host plant but not the rhizobial strain is responsible for the type and morphology of the nodule. In general, the nodules contain infected, uninfected cells , parenchyma, endodermis, and vascular bundles (Gresshoff and Delves, 1986). Indeterminate and young determinate nodules maintain a zone of cell division.

After formation of the nodule primordium, a persistent nodule meristem is initiated in indeterminate, e.g., pea, clover, alfalfa, or vetch, unlike determinate, e.g., soybean, mungbean, or common bean, nodules (Figure 1-2). A primordium is formed from cells of the root outer cortex and inner cortex in determinate and indeterminate nodule species, respectively. Infection threads are broader and penetrate a longer distance in indeterminate (elongate) *vs.* determinate (spherical) nodules. Infection threads in both types of nodules grow towards the nodule primordium and bacteria are released into the plant cells, enclosed by a plant membrane called peribacteroid or symbiosome membrane



	Indeterminate	Determinate
Site of initial cell divisions	Inner cortex	Outer cortex
Nodule growth	Cell division; persistent meristem	Cell expansion
Effect of *exo* mutant	Fix$^-$, 'empty' nodules	Fix$^+$, normal nodules
Effect of *lps* mutant	Fix$^-$ or $^+$, normal in overall appearance	Fix$^-$, abnormal nodules
Infection thread	Broad	Narrow
Origin	Temperate regions	Subtropical and tropical regions
Transport	Amides	Ureides
nod gene inducers	Flavones, flavanones	Isoflavones, mainly
Examples	Alfalfa, pea, vetch	Soybean, bean, *Lotus*

M, meristem; TI, thread invasion zone; ES; early symbiotic zone; NF, nitrogen-fixing zone; S, senescent zone; NC, nodule cortex; NE, nodule endodermis; NP, nodule parenchyma; VE, vascular endodermis; VB, vascular bundle; Sc, sclerenchyma; P, periderm.

Figure 1-2. Major differences between indeterminate and determinate nodules, modified from Sutton, 1983.

(Roth *et al.*, 1988). The rhizobia Nod factor, NodRm-1 and NodR1v, can induce cortical cell division leading to formation of nodule on alfalfa seedlings (Truchet *et al.*, 1991). It seems unlikely that a large and complex molecule like the sulphated glycolipid NodRm-1 can diffuse across plant membrane. In addition, formation of spontaneous nodules in alfalfa (Truchet *et al.*, 1989; Joshi *et al.*, 1991) suggests that an independent signal other than Nod factor is responsible for cortical cell division and nodule formation. Nodulation in the absence of *Rhizobium* (Nar) phenotype is dominant, heritable, and caused by an alfalfa gene (Caetano-Anollés *et al.*, 1992). The gene product could be involved in a signal transduction started with perception of the Nod factor in nodule development. The morphogenic signal in the infected roots is transduced from the root hair to the cortical cells leading to cell division. This signal may involve a change in the membrane potential and Ca^{2+} levels. After addition of Nod factor, root hair membrane depolarize rapidly (Ehrhardt *et al.*, 1992), vacuoles change their shape, and cytoskeletal rearrangements occur within root hair cells (Allen *et al.*, 1991). Cortical cell division and root hair curling in soybean roots is also induced by purified common nod factor at nM concentration or lower level (Stacey *et al.*, 1995).

Plant hormones as an endogenous growth regulators can also function as signals for morphogenesis. Application of plant hormones such as cytokinin induced cortical cell division in soybean, cowpea, and alfalfa (Bauer *et al.*, 1985) whether these are specially involved in nodule initiation is not clear. Cytokinin is also responsible for formation of bacteria-free nodules on alfalfa roots (Long and Cooper, 1988). Bacteria can also produce auxins, gibberellins, and cytokinins. Addition of flavonoids stimulates IAA production in *R. meliloti* (Prinsen *et al.*, 1991). Treatment with *nod*-gene inducing flavonoids produce different cytokinin in *R. meliloti* suggesting that cytokinin production may be NodD regulated while *nodA*, *B*, *C*, or *D* mutants in *B. japonicum* still produce cytokinins (Taller and Sturtevant, 1991). Application of auxin transport inhibitors on the alfalfa roots led to endogenous hormone imbalance causing cell divisions, formation of psuedonodules, and expression of early nodulins (Hirsch *et al.*, 1989).

Nodule Function and N_2 Fixation

In indeterminate nodules, the nodule meristem results in central nodule tissues consisting of infected and uninfected cells, peripheral vascular bundles, and nodule

parenchyma (Hirsch, 1992).In determinate nodules, the first cell division occurs in the hypodermal region of the root and later in the pericycle and inner cortex. Eventually, these dividing tissues come together forming the central nodule tissues (Rolfe and Gresshoff, 1988). *Rhizobium* and *Bradyrhizobium* will differentiate into bacteroids after release into infected host cells. Fixation of atmospheric nitrogen begins in the bacteroid state by nitrogenase which may represent up to ten percent of the soluble bacteroid protein (Werner, 1992).

Nitrogenase is a two component-metalloenzyme encoded by rhizobial *nif* genes (*nifH,D,K*). Both components, an iron-containing (Fe) protein and a molybdenum-iron (MoFe) protein, are required for the nitrogen reduction. Substrate binding sites and reduction resides in the MoFe domain while an ATP binding site is in the Fe domain (Dean and Jacobsen, 1992). After reduction of nitrogen, ammonium (NH_4) is exported through bacteroid and peribacteroid membranes into the host cell cytoplasm. Then, the ammonium is biosynthesized by enzymes such as, glutamine synthetase, glutamate synthase, aspartate aminotransferase, and asparagine synthetase (Werner, 1992) to provide the "building blocks" for common metabolites such as amino acids, nucleotides, vitamins, and secondary products.

Nitrate Inhibition of Nodulation

Although the number of nodules per plant and nodule morphology is genetically controlled (Fujita et al, 1991), the root nodule symbiosis is inhibited by high concentration of nitrate in the soil (see Carroll and Mathews, 1990). This suppresses accumulation of rhizobia on root hairs, root hair curling, infection thread formation, and nodule development (Streeter, 1988). In soybean, similar effects on root hair deformation, bacterial attachment, and infection thread formation have been demonstrated (Carroll and Mathews, 1990). Application of 4 mM nitrate in the soil caused reduction in the nodule mass per plant and in bacterial nitrogenase activity as well (Carroll *et al.*, 1985a). Nitrate inhibition can be explained with nitrite toxicity, carbohydrate, and oxygen deprivation in the nodule. Soybean mutants capable of nodulation in the presence of high nitrate concentration suggests a link between the deregulation of control over nodule number and the nitrate inhibition of nodulation (Carroll *et al.*, 1985a and b: Carroll and Mathews, 1990).

Autoregulation of Nodulation

In general, plants control and optimize the formation of nodules through a single and responsive mechanism. This suppresses nodule formation in younger parts of the root where once a number of nodules has formed (see Gresshoff, 1993; Gresshoff et al., 1989; Caetano-Anollés and Bauer, 1988; Carroll *et al.*, 1985; Olsson *et al.*, 1989). Removal of nodule from both soybean (Caetano-Anollés and Gresshoff, 1991c) and alfalfa (Caetano-Anollés and Gresshoff, 1991 b, c) showed different responses in both time and space. In soybean, new nodules emerged in the region of original nodulation on the primary root where the original bacterial inoculum was applied. This indicates that soybeans arrest some nodule primordia during early nodule ontogeny. In alfalfa, in contrast to soybean, after nodule excision, nodules emerged in young root tip regions where it was expected to be the site of new nodule formation. These results were confirmed by histological studies which showed that soybean roots harbored abundant early nodulation stages (Mathews *et al*, 1989; Gerahty *et al.*, 1992) compared to alfalfa with only a small number of immature nodule primordia (Caetano-Anollés and Gresshoff, 1991 b, c).

Autoregulation blocks nodule formation in a soybean cultivar, Bragg, but not in its supernodulating mutant, nts (Carroll and Mathews, 1990). It appears that the number of infection events in the parental and the mutant is the same but the nodule formation is blocked in the parental type while deregulated in the mutant (Carroll and Mathews, 1990). Supernodulation was suppressed when shoot extracts from wild type injected into mutant nts (Gresshoff *et al.*, 1988). This regulatory suppressor in wild type was shown to be systemic by using split root system study (Olsson, *et al.*, 1989) suggesting that the autoregulation inhibitor in the wild type regulates nodule formation and growth.

Legume Nodulation Mutants

To study the nodule symbiosis required genetic analysis of the bacteria and the plant. Using transposon mutagenesis and gene isolation a set of nodulation and nitrogen fixation genes in bacteria was characterized. The function of some of these genes involved in the synthesis of the nod factors in *Rhizobium meliloti* has been defined (Verma, 1992; Truchet *et al.*, 1991; Lerouge *et al.*, 1990).

Several symbiotic legume mutants have been developed that are either supernodulating or non-nodulating (see Caetano-Anollés and Gresshoff, 1991). In soybean, nodulation mutants were developed by using ethyl methane sulfonate (EMS) chemical mutagenesis (Carroll *et al.*, 1985a). The non-nodulating and supernodulating mutants are controlled by single Mendelian recessive genes (Caetano-Anollés and Gresshoff, 1991). The non-nodulating mutants, *nod49*, *nod772*, (Carroll *et al.*, 1986) and *rj1* (Williams and Lynch, 1954) are in the same complementation group while the non-nodulating mutant *nod139* formed a new class (Carroll *et al.*, 1986). In all non-nodulating mutants root hair deformation did not occur and only few subepidermal cell divisions with no infection threads (pseudoinfection) were observed. In *nod139* even pseudoinfections did not occur, suggesting that the block in nodulation for *nod49*, *nod772*, and *rj1* is at a later stage than in *nod139*. The roots genotype of mutants *nod49* and *nod139* controlled the shoot genotype by using grafting experiment of parental and mutant root and shoot stocks exchange (Mathews *et al.*, 1992; Delves *et al.*, 1986).

The supernodulating mutant can form three to forty times more nodules than the wild type and is nitrate tolerant (Carroll *et al.*, 1985a, b). It also forms more lateral roots than the wild type due to possibly hormone imbalance in the mutant. Both the mutants and the wild type grow at similar rate when there is no inoculation and nitrate is used as the nitrogen source (Carroll and Mathews, 1990). The *nts* mutant showed supernodulating phenotype when its shoots were grafted on the parental, Bragg, root stocks while parental shoots on the mutant rootstocks gave normal nodulation phenotype (Delves *et al.*, 1987a, b). In contrast to soybean, pea autoregulation of nodule number is controlled by the root in supernodulating *nod3* (Carroll and Mathews, 1990),.but also the shoot (see Sagan and Gresshoff, 1996).

Other supernodulating mutants have been developed using the same procedure as in soybean by Carroll *et al* in 1985. Supernodulating mutant in bean (*Phaseolus vulgaris*) shared most properties with the soybean mutant (Buttery and Park, 1989; Park and Buttery, 1997). Other soybean supernodulating mutants were developed by chemical mutagenesis in Elgin cultivar (Buzzell *et al.*, 1990) and Enrei cultivar (Akao and Kouchi, 1992) as well as Williams (Gremaud *et al.*, 1989). EMS mutagenesis on faba bean seed induced a supernodulating phenotype having 3 to 5 times more nodules than the wild type (Duc, 1995). A hypernodulation phenotype was recovered from mutagenized M2 population of *Lotus japonicus*, a legume model, and designated as

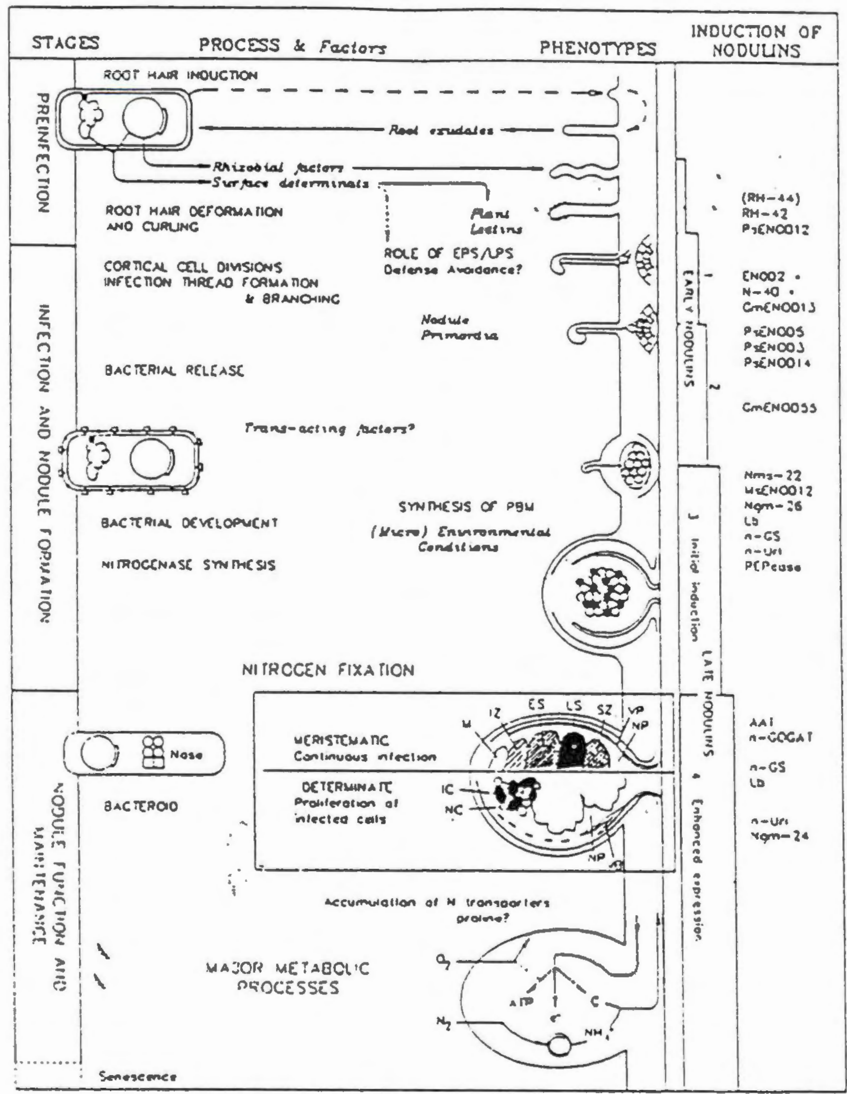
Nod++ (Szczyglowski *et. al.*, 1998). In pea, nodulation phenotype of plants supplied with NH_4^+ was similar to that reported for supernodulating mutants (Waterer *et al.*, 1992). All these mutants shows Mendelian single locus inheritance for nodulation.

The Nodulin Concept

'Nodulin genes' was first defined as nodule-specific genes expressed during nodule development (Legocki and Verma, 1979, 1980) but molecular evidence has shown the expression of nodulins in other plant organs such as the flower (Nap and Bisseling, 1990). Thus the nodulin genes define as plant genes which are differentially and temporally expressed or enhanced during nodule development and nitrogen fixation process (Figure 1-3). Almost 200 nodulin genes have been identified from different legumes so far.

Nodulin genes that are expressed in the early stage of nodule development, preinfection, infection, and cortical cell division, are named early nodulin (*ENOD*) genes. The majority of nodulin genes expressed around the onset of nitrogen fixation are late nodulin (*NOD*) genes which are involved in specific biochemical pathways (Verma, 1992).

A set of legume nodulin cDNA clones has been isolated from soybean, *Glycine max*, (Franssen *et al.*, 1987; Kouchi *et al.*, 1990; Delauney and Verma, 1988; Nirunsuksiri and Sengupta-Gopalan, 1990; Sandal *et al.*, 1987; Thummler and Verma, 1987; Fortin *et al.*, 1987; Fuller *et al.*, 1983; Katinakis and Verma, 1985; Lee *et al.* 1983; Legocki and Verma, 1979 and 1980), pea, *Pisum sativum*, (Scheres *et al.*, 1990; Nap, 1988; Tingey *et al.*, 1987; Govers *et al.*, 1986) clover, vetch, *Vicia sativa* (Moreman *et al.*, 1987) faba bean, *Vicia faba*, (Perlick *et al.*, 1996 and 1997; Küster *et al.* 1994), alfalfa, *Medicago sativa*, (Dickstein *et al.*, 1988), lupin, *Lupinus luteus*, (Konieczny *et al.*, 1988), *Medicago truncatula* (Wilson *et al.*, 1994), *Sesbania rostrata* (Goormachtig *et al.*, 1995; Strittmatter *et al.*, 1989) winged bean (Manen *et al.*, 1991), French bean, *Phaseolus vulgaris*, (Lara *et al.*, 1983) and *Lotus japonicus*. In general, expression of early nodulin genes has been used to study the mode of action of Nod factors and the early signal exchange between the plant and bacteria. Soybean early nodulin GsENOD2, one of the most studied nodulin, has been characterized in detail (Franssen *et al.*, 1987, Gloudemans *et al.*, 1987). This early nodulin function is not related to the infection



Schematic representation of developmental and metabolic events associated to nodulin expression. Distinct stages, processes, and factors *(left* and *center)* are shown in relation to nodule developmental phenotypes. The figure also depicts the sequence of expression of nodulins and nodule-enhanced activities *(right)* in an arbitrary time scale. At least four steps in nodulin induction are indicated (1–4), with representative examples of several species (* = ubiquitous nodulin type). Nodule types are represented in the middle (MERISTEMATIC: M = meristem; IZ = infection, ES = early symbiotic, LS = late symbiotic and SZ = senescing zones; VP = vascular bundle; NP = nodule parenchyma. DETERMINATE: IC = infected and NC = uninfected cell; VB = vascular bundle; NP = nodule parenchyma/inner cortex.) Major metabolic processes *(bottom)* are described as linked pathways of carbon, nitrogen, and oxygen metabolism, in which several nodulins participate.

Figure 1-3. Schematic representation of developmental and metabolic events associated with nodulin expression. This figure is reproduced from Sanchez *et al.*, 1991.

process because its expression occurs in soybean pseudonodules which lack intracellular bacteria and infection threads (Franssen *et al.*, 1987). Different lipo-chitin oligosaccharide (LCO) molecules from soybean symbionts were tested for soybean ENOD2 mRNA expression. A cooperative action of at least two LCO molecules, an active and a non-specific LCO, were found to be sufficient for induction (Minami *et al.*, 1996). Furthermore, alfalfa early nodulin MsENOD2 was expressed in the *R. meliloti exo* mutant induced nodules which devoid of infection threads (Dickstein *et al.*, 1988). Therefore, early nodulin ENOD2 function is not associated with infection process but it is involved in nodule development and organogenesis. *ENOD2* transcripts of soybean and pea accumulate in the nodule parenchyma as well as in cells surrounding the vascular bundle connecting the nodule with the root central cylinder (van de Wiel *et al.*, 1990). In the Part 2 of this dissertation, ENOD2 will be discussed in detail.

In contrast to the *enod2* gene, other early nodulin gene expression studies have shown that they might play a role in infection process. The pea early nodulin PsENOD5 gene encoding a proline-rich protein (van de Wiel *et al.*, 1990) and the PsENOD7 gene are expressed in infected cells during nodule maturation (Kozik *et al.*, 1996; Franssen *et al.*, 1992). Another pea early nodulin PsENOD5 gene is expressed in the infected cells, and is also detectable at low level in the invasion zone (van de Wiel *et al.*, 1990). The early nodulin VfENOD-GRP3 transcript was detected predominantly in the interzone II-III region of faba bean root nodule (Kuster *et al.*, 1995).

The early ENOD40 mRNA accumulates in the nodule pericycle of the vascular bundle at 40 hours after inoculation with either purified or chemically synthesized lipochitooligosaccharide (LCO) Nod factors (Minami *et al.*, 1996). Expression of French bean *ENOD40* is comparable with soybean *ENOD40* and occurs in the root pericycle, nodule perimordia, pericycle of vascular bundles, and uninfected cells of mature nodules (Papadopoulou *et al.*, 1996). First expression of the early nodulin *GmENOD55* occurs after release of bacteria in plant cells and is restricted to the infected cell type (Blank *et al.*, 1993).

Whether nodulins are essential for nodulation and function is still an open question. Many may be lateral responses, rather than causes. For example, in *M. sativa*, *MsENOD12* transcript was localized in the epidermis of infected roots (Journet *et al.*, 1994). Similarly, Transcription of *M. truncatula ENOD12* occurs 3 to 6 hours after

inoculation in a zone of differentiating root epidermal cells which is located close to the growing root tip ahead of the infection zone (Pichon *et al.*, 1992). Genetic analysis of this gene in F_1 and F_2 revealed a null allele in several offsprings which were similar to wild-type parents in viability, nodule development, nodule structure, and nitrogen fixation efficiency suggesting that *MsENOD12* is not required for symbiotic nitrogen fixation (Csanadi *et al.*, 1994).

Those nodulins with known functions are mostly late nodulins which participate in nodule metabolism. An exception regarding known function is the early nodulin peroxidase in *Medicago truncatula, Mtrip1*, which is expressed in the pericycle of uninoculated roots (Cook *et al.*, 1995). In soybean, many late nodulin genes such as *Ngm-20, Ngm-21, Ngm-23, Ngm-24, Ngm-26, Ngm-44, Ngm-56, Ngm-93*, sucrose synthase (*Ngm-100*), uricase (*Ngm-35*), and leghemoglobin have been characterized.

Function of the nitrogen-fixing nodule depends on the expression of the late nodulins involved in nitrogen and carbon nodule metabolisms. The expression of leghemoglobin (Lb) facilitate oxygen diffusion in the host cell and is essential for effective symbiosis (Appleby, 1984; Wittenberg and Wittenberg, 1990). leghemoglobin is suggested to be a symbiotic molecule in the way that the globin moiety is encoded be the host plant and the heme moiety is made by the bacteria (Lee and Verma, 1984). Supporting this idea, a *Rhizobium* mutant defective in heme biosynthesis resulted in non-functional leghemoglobin (Nadler, 1981). In contrast, a similar mutant in *Bradyrhizobium japonicum* induced effective nodules in soybean (Guerinot and Chelm, 1985) suggesting that some plants can synthesize leghemoglobin at low level to maintain nodule function. Recent studies have shown that the heme precursor was synthesized in the plant cells while the subsequent heme biosynthesis steps are completed by the bacteria (Sangwan and O'Brian, 1991). This suggests that heme biosynthesis is spatially separated between the two partners.

Several studies have shown that leghemoglobin is located in the cytoplasm of the nodule's infected cells, but not inside peribacteroid membrane compartment (Robertson *et al.*, 1984; Nguyen *et al.*, 1985). Only low level of leghemoglobin may be expressed in the uninfected cells of soybean (VandenBosch and Newcomb, 1988) although other data tend not to support this finding (Kouchi *et al.*, 1990). Expression of leghemoglobin in infected cells maintains the oxygen concentration at a level that is not toxic to the

bacterial nitrogenase inside the bacteroid. Thus, leghemoglobin is considered as a defense molecule against nitrogenase toxicity (Verma *et al.*, 1990). At least four leghemoglobin genes are expressed at slightly different times of nodule formation indicating developmental control (Jensen *et al.*, 1988).

Uricase, a key enzyme in the ureide biosynthetic pathway for assimilation of ammonia, is the second most abundant late nodulin in the cytoplasm of root nodules. Soybean nodule uricase with 35 kDa molecular weight, Ngm-35, is encoded by a gene totally different from those for root and leaf uricase. Using a cDNA clone as probes in Southern hybridizations, several fragments homologous to uricase were identified in soybean suggesting the existence of a small number of genes (Nguyen *et al.*, 1985). The soybean n-uricase is localized in the peroxisomes of the uninfected cells (Bergmann *et al.*, 1983).

Another late nodulin with known function is glutamine synthetase (GS) which is involved in nitrogen assimilation. Multiple glutamine synthetase isoenzymes exist in legume plants suggesting the presence of multigene family expression (Dunn *et al.*, 1988; Hirel *et al.*, 1987; Tingey *et al.*, 1987). In French Bean, *Phaseolus vulgaris*, different subunits of glutamine synthetase ranging from 41 to 45 kDa has been identified (Bennett *et al.*, 1989). Two types of glutamine synthetase exist in the root nodules of which one is expressed as a nodule-specific (GS-n) and the other whose expression is enhanced significantly during symbiosis (Forde and Cullimore, 1989). Both types of the glutamine synthetase isoforms expression are increased before bacterial nitrogenase activity (Dunn *et al.*, 1988; Padilla *et al.*, 1987).

Sucrose synthase (SS), a late nodulin, is involved in the flow of carbon to nodule and bacteroid metabolism. A soybean nodule-specific cDNA clone encoding the sucrose synthase subunit with 100 kD molecular weight has been isolated and characterized (Morell and Copeland, 1985). Based on partial sequence analysis, there is 73% homology at amino acid level between soybean nodule and maize sucrose synthase (Thummler and Verma, 1987). Soybean nodule sucrose synthase is inactivated after incubation with free heme suggesting that free heme regulates the activity of the sucrose synthase and the metabolism of the carbon (Thummler and Verma, 1987).

It was shown that the expression of some late nodulins was regulated in different ways to maintain a balance of metabolites. For example, nodule sucrose synthase activity increased rapidly during nodule development and decreased as the leghemoglobin degradation occurred during senescence (Thummler and Verma, 1987). The uricase and sucrose synthase mRNA levels are increased 5-6 and 4- fold, respectively, when soybean callus was exposed to 4% oxygen concentration. In contrast, the leghemoglobin c3, nodulin-22, and nodulin-44 mRNAs were not expressed in response to neither atmospheric nor low oxygen levels (Xue *et al.*, 1991). Expression of early N-75 and N-38 nodulins occurred within 7 days while late leghemoglobin and N-24 were expressed by 10 days after inoculation (Kouchi *et al.*, 1990). The late nodulin *GmN-56* gene, expressed at the onset of nitrogen fixation together with leghemoglobin and other late nodulin, encodes a protein homologous to isopropylmalate synthase and homocitrate synthase (Kouchi and Hata, 1995).

Some of the nodulin genes are associated with peribacteroid (a symbiosome) membrane (PBM) where fixed nitrogen (ammonia) is transported into host cell cytoplasm. The soybean late nodulins *Ngm-24, Ngm-25, Ngm-26,* highly expressed and members of a gene family, are not immunoprecipitated with antibody to the soluble fraction of nodules suggesting that these proteins are associated with membrane (Richter *et al.* 1991). Nodulin 26 is an integral symbiosome and ion channel protein which transports both cations and anions (Weaver *et al.*, 1994). This nodulin shares high sequence homology with several proteins characterized in other plants and species suggesting a similar role and common ancestor for this nodulin. Recently, it was found that water crossed the soybean NOD 26 of root nodule symbiosome membrane with a single channel pathway, while larger solutes such as formamide and glycerol appeared to cross the membrane by a pathway different from the one for water (Rivers *et al.*, 1997).

Expression of the nodulin is induced under stimuli other than bacterial Nod factors. The early nodulin gene *SrENOD2* from *Sesbania rostrata* was inducable after cytokinin application (Dehio and de Bruijn, 1992). Upon application of auxin transport inhibitor, the early nodulin ENOD2 was expressed in alfalfa nodules (Hirsch *et al.*, 1989). Nodule metabolites can also affect expression of some nodulins involved in nodule metabolism. Expression of soybean glutamine synthetase localized in the cytoplasm of the infected cells of nodule is increased after external application of ammonia (Hirel *et al.*, 1987). Activity of glutamine synthetase was elevated after ammonia application on the *Trifolium*

roots (Reynolds *et al.*, 1990). In contrast, *Phaseolus vulgaris* glutamine synthetase gene expression was not responsive to externally supplied ammonia (Cock *et al.*, 1990). It appears that glutamine synthetase expression is regulated in many ways depending on different isoforms, legume species, and physiological condition in legume plants.

The Genetics of Nodulins

Study of nodulin gene promoters in transgenic legumes has identified a number of cis-regulatory elements essential for nodule specific expression. Significant DNA elements were identified in soybean *enod2B* gene promoter (Lauridsen *et al.*, 1993). Activity of GUS under *enod2B* promoter was found in both determinate *Lotus* and indeterminate *Trifolium* nodules. A positive and a qualitative element were identified by deletion and hybrid promoter analysis in *L. corniculus* (Figure 1-4).

The soybean leghemoglobin, lbc3, promoter studied in transgenic *Lotus corniculatus* plants contains a strong positive element (SPE), a weak positive element (WPE), an organ specific element (OSE), and a negative element (Figure 1-5). Two conserved motifs, CTCTT and AAAGAT, within the organ-specific-element are important for its function (Ramlov *et al.*, 1993). The function of cis-elements was further studied by their interaction with transcription factors. A nodule-specific nuclear protein, NAT2, was identified that interacted with two AT-rich DNA sequences in the weak positive element (Jensen *et al.*, 1988). In addition, the NAT2 binding site fused to a -139 lbc3 promoter could activate the construct (Laursen *et al.*, 1994). The A-T rich DNA elements with the ability to bind to trans-acting factors have been identified in the soybean *GmN-23* and French bean glutamine synthetase, *gln-g*, (Forde *et al.*, 1990; Jacobsen *et al.*, 1990). These trans-acting factors were found in nodule extracts, but also in roots and leaves. Biochemical and DNA analysis of some of these factors showed functional relationships with human nuclear proteins, HMG1, (Jacobsen *et al.*, 1990) suggesting that chromatin structure might be an important controlling factor for organ-specific expression. Furthermore, a rhizobial trans-acting factor has been identified that interacts with DNA sequences in the promoter regions of *Sesbania rostrata* leghemoglobin gene (Welters *et al.*, 1990). It is yet unclear, how this bacterial protein traverse the multiple bacterial and plant membranes. Based on deduced amino acid sequences and homology studies in the data base, a possible function has been defined for some nodulin genes. For example, using differential hybridization, a cDNA clone of *GsENOD2* was isolated from a

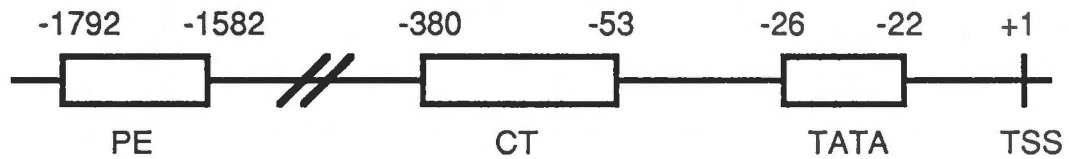


Figure 1-4. Promoter elements of soybean *enod2B* gene. PE, positive elements; possible tissue specific element. CT, cell type containing conserved sequences of CTCTT and AAAGAT. TSS, transcript start site.

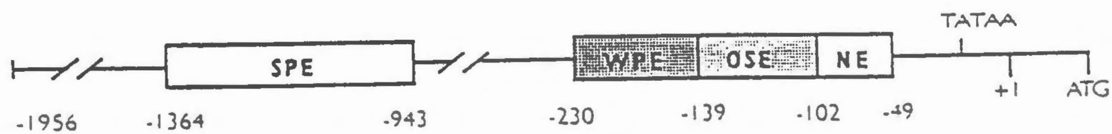


Figure 1-5. Schematic representation of the cis-regulatory elements in the soybean lbc3 promoter. SPE: strong positive element; WPE: weak positive element; OSE: organ specific element; NE: negative element. This figure reproduced from Stougaard *et al.*, 1990.

soybean nodule library. Amino acid sequence analysis showed that this gene is a hydroxyproline-rich cell wall protein (Franssen *et al.*, 1987). The other early nodulin genes such as *PsENOD3* and *PsENOD14* are assumed to be proteins involved in metal transport based on their amino acid sequences (Scheres *et al.*, 1990).

The alfalfa late nodulin gene *Nms-25* was expressed 9 days after inoculation similar to the leghemoglobin time of expression suggesting a similar regulation. The protein deduced from cDNA sequence contained a signal sequence which might direct the protein into the symbiosome membrane (Kiss *et al.*, 1990). This gene consists 13 exons and 12 introns of which most of exons are similar to each other suggesting exon-shuffling. The promoter region contains the common promoter elements of plant genes (Vegh *et al.*, 1990).

The *GmN-23* nodulin promoter contains two positive distal (PE-A, -320 to -298) and proximal (PE-B, -257 to -165) enhancer elements. In PE-A and PE-B, two 12-bp sequence motifs are found to be core of the enhancer elements, InvA and InvB, respectively. The nodule-specific trans-acting factor binding site NAT2 is present in PE-A (Jorgensen *et al.*, 1991).

Leghemoglobin (Lb) proteins are encoded by a gene family in the legumes. In soybean, at least four major Leghemoglobin genes have been identified, *Lba*, *Lbc1*, *Lbc2*, and *Lbc3* (Lee *et al.*, 1983). Furthermore, Leghemoglobin psuedogenes and truncated genes exist in the soybean (Brisson and Verma, 1983). However, it is unknown whether these different Leghemoglobins have different roles in the nodule. Indeed, their function in nitrogen fixation is only verified by biochemical studies with isolated bacteroids. As yet there is no leghemoglobin-defective mutant, altered in Lb itself. Perhaps the multi-gene nature prevents such phenotype.

Molecular Mapping

The purpose of mapping is to isolate and determine the function of a gene of interest. In addition, co-segregation study of a locus and a gene with known-function and already mapped on the genetic linkage map may result in function determination of the locus under study. Molecular mapping can also be used for saturation of a region with

molecular markers which in turn provides useful information for anchoring BAC and YAC clones containing DNA sequences for other genes in the same region of genome.

In plants, there are numerous genes and it is a tedious task to correlate genes with functions. One of the approach for gene isolation is map-based cloning strategy in which, first, a mutant impaired in the biological process in question is induced. Then, the gene suffered the mutation is mapped using molecular markers which should be close to the target gene (Gresshoff and Landau-Ellis, 1994; Gresshoff, 1993). The next step is to analyze the complex genome and clone a large DNA fragment bearing the gene of interest. This can be done by cloning genomic fragments into yeast artificial chromosome (YAC) or bacterial artificial chromosome (BAC) and screen them with the help of tightly linked molecular markers (Funke and Kolchinsky, 1994). Finally, complementation of mutant can be conducted by tranformation of the mutant with the isolated gene to regain original wild-type phenotype. The advantage of this strategy over ambiguous biochemical approaches is that one can be sure that the isolated gene is really involved and essential in the biological activity. Therefore, the initial requirement for map-based cloning is to construct a genetic linkage map conferring the distance and order of genetic markers on the chromosome.

Furthermore, molecular mapping provides information regarding DNA markers that are linked to a trait which is important for plant and animal breeders. This linked marker would be a useful tool for selection of the trait in the early stage of growth saving time and money for the breeders. This approach is called marker-assisted breeding.

Other benefits of molecular mapping are to determine genetic linkage of quantitative trait loci, OTLs, (Keim *et al.*, 1990), pedigree mapping (Shoemaker *et al.*, 1994), and integration mapping (Shoemaker and Specht, 1995; Ghassemi and Gresshoff, 1998). In QTL mapping, number and frequency of alleles and loci involved in the trait can be estimated as well as contribution of each locus in total variation for the trait. Pedigree mapping assists the breeders to not only determine the genotype of an individual but also to evaluate the genetic composition of entire population.

Integration mapping refers to mapping a locus on a genetic linkage map using molecular and conventional markers from another genetic linkage map. Some of the new genetic linkage maps have been saturated through a concept called "synteny" mapping. In this

approach, one can use the information of markers from one map and integrate to the other. This concept can be applied not only within a given species but also between species or even between genera and organisms.

Molecular Genetic Linkage Map in Soybean

Most of the genetic linkage maps in eukaryotes have been constructed based on the restriction fragment length polymorphism (RFLP) approach. In recent years, many molecular markers generated by polymerase chain reaction (PCR) have been added to the present genetic linkage maps. These PCR-based molecular markers consist of randomly amplified polymorphic DNA, RAPD, (Williams *et al.*, 1990; Welsh and McClelland, 1990), DNA amplification fingerprinting, DAF, (Caetano-Anollés *et al.*, 1991d), microsatellite or simple sequence repeat, SSR, (Jeffreys *et al.*, 1985; Cregan *et al.*, 1994), and amplified fragment length polymorphism, AFLP, (Vos *et al.*, 1995). In addition, there are morphological and isoenzyme markers on the present genetic linkage maps.

In soybean, several genetic linkage maps have been constructed by several groups in the University of Utah (Lark *et al.*, 1993; Mansur *et al.*, 1996), the Iowa State University (Shoemaker and Specht, 1993), the Pioneer Hi-Bred International Inc. (Webb *et al.*, 1995), and the Northern Arizona University (Keim *et al.*, 1997). These maps were generated based on an initial cross between parents which might be either intra- or inter-species, *G. max vs. G. max* or *G. max vs. G. soja*, respectively. The disadvantage of inter-species cross is that recombination rate is lower than intra-species cross due to lack of complete chromosome pairing in parts of genome during meiosis division.

The soybean genetic linkage map constructed in the University of Utah (Figure 1-6) consists of 35 linkage groups and 377 markers covering about 2,000 centiMorgans (cM) and is expected to define another 1,000 cM of the genome (Mansur *et al.*, 1996). This map was developed from a recombinant inbred line population consisted of 284 F_7-derived lines generated from single seed descent of a cross between Minsoy, PI27890, and Noir1, PI290136 (Mansur *et al.*, 1993).

A genetic linkage map developed in the Northern Arizona University consists of 840 markers and 28 linkage groups covering 3441 cM of the soybean genome, Figure 1-7,

U1a

Marker	cM
1)Sct46	5.0
3*)A60	33.2
4)LE28H8	1.6
4)L50ND10	0.0
1)Scaa3	2.0
1)Sct1	28.4
3*)K375	1.0
4)R189	8.3
4*)G815	22.7
3)A132	
	102.1cM

U1b

Marker	cM
4)G214NH5	1.3
4)G214NT12	26.3
4)G214MT8	0.5
4)G214MH15	0.0
2)BLT49b	4.9
3)K11MB20	
	33.0cM

U1c

Marker	cM
1)Sct34	23.7
4)G214MH1	2.1
1)Satt70	1.3
2)BLT57	1.4
4)L201	0.5
4)G214MH19	18.5
4)G214MT10	23.4
4)G214MB9	4.5
4)C9b	0.5
4)G214MH18	0.3
1)SOYPRP1	3.6
1)Satt1	2.0
Fr1	20.2
4)R51	0.8
3)A315	29.1
Ep	26.8
3)A975	
	158.7cM

U2a

Marker	cM
4)C9c	2.9
3*)A711	15.3
4)L163	11.6
3)K274	0.5
4)G214NT19	1.8
2)BLT49a	0.0
4)R28	0.0
1*)Satt45	1.6
3)A510b	5.7
3)B124b	22.1
3*)A427	
	61.8cM

U2b

Marker	cM
4)G214NH2	35.6
4)L50NT19	2.9
4)L50MT13	8.4
4)G214MT25	1.9
4)G214NB14	3.1
3)A459	0.0
3)B124a	9.2
2)BLT39	13.6
2)BLT10	0.2
2)BLT7	
	75.0cM

U2c

Marker	cM
4)L50ME29	13.3
4)CHI	17.5
1)GMGLPSI2	
	30.8cM

U3

Marker	cM
4)L144	14.2
2)BLT36a	16.8
3)A65	1.1
3*)A505	0.1
2)BLT53NT18	14.6
1)Sat40	0.1
3)K443	5.3
3*)A690dr	10.5
4)G214NT1	0.7
4)G214NH12	22.3
1)Satt89	34.7
3*)A111	10.3
3)K401b	4.2
1)Gmenod2b	8.4
3)BLT24	0.9
*I	2.1
3)A262b	2.6
4)L199b	17.6
3*)A110b	
	166.3cM

U4

Marker	cM
4*)T28	12.4
3)K11a	6.9
3)K11MB5	24.3
3*)A109c	32.3
3*)A262a	3.5
4)T92	15.4
2)BLT43	12.8
4)G214MB3	3.2
4)G214MT2	1.2
4)G214MH3	1.8
1)Sct26	11.5
4)L204NT4	
	125.3cM

U5

Marker	cM
4*)L156	8.9
2)BLT36b	8.6
3)A510c	0.5
4*)R17	2.4
4)L2b	1.1
4)L50MD28	19.3
4)G214MT6	1.4
4)G214MB2	4.8
1*)Satt12	20.3
4)T5	23.3
4*)L2a	0.0
3*)A235a	2.7
4)L154	14.1
3)A690E	4.5
3)A586	1.4
3)A378	
	113.4cM

U6

Marker	cM
3)A363a	7.1
3)A455	11.2
1)Satt22	18.2
4)G214MH11	0.0
4)G214MT13	8.2
2*)BLT15	28.8
Rpg4	13.7
4)L50ME23	15.1
3*)B162	0.0
4*)L103a	4.2
1)Satt80	11.0
3)A280	6.6
2)BLT4	4.2
1*)Satt9	
	128.3cM

U7

Marker	cM
4)G214NB6	10.6
1)Satt42	2.2
3)A329	2.9
1)Satt73	0.8
3)A53b	0.5
3)A110a	0.0
4)L194b	0.4
4)G214NH8	1.0
4)G214NT11	8.1
4)R183	0.0
3)A64b	2.7
1)Satt50	12.7
4)T153b	4.8
3)A262c	22.9
2)BLT53NT19	4.2
3)K401a	5.4
3)K636	13.6
4)T155	2.6
3)B170	
	95.5cM

U8

Marker	cM
4)C53	2.7
4)L58	14.4
1)Satt71	15.7
3)A295	18.1
1*)Sat36	11.0
4)NP8	3.4
3*)A235b	1.8
1)Satt32	7.1
3)K227	10.1
4*)R13	29.8
4*)M373	
	114.0cM

U9

Marker	cM
3)A676	38.3
4*)C56	12.7
3*)K365	1.3
1)Satt79	0.0
1*)Sct28	1.3
2)BLT29	1.4
3*)A397	1.3
1)Satt100	0.6
3)A109a	0.0
4*)L50ND15	4.0
2)BLT53MT11	5.3
2)BLT32	6.2
4*)L148	0.0
4)R92	10.2
4)L59	11.2
1)SoyAC7L	2.7
3*)A426	
	96.5cM

U10a

Marker	cM
3)A748	16.0
2)BLT19a	8.8
3)A131	0.6
1)Satt52	0.0
3*)A404	12.2
2)BLT46	1.9
2)BLT53NT3	12.3
3*)A89	
	51.8cM

U10b

Marker	cM
4)L192	6.0
1)Sat42	7.7
3)A63	
	13.2cM

U11

Marker	cM
3)A60b	1.9
4)R79	18.3
4)L204NT2	1.8
1)Sat3	5.5
3)A584	35.7
2)BLT25	1.4
4)M121	
	64.6cM

U12

Marker	cM
Fr2	6.7
1)Sat1	0.9
4)G214NT14	0.2
4)G214NH6	1.7
4)L26g	3.4
3)K11NB9	11.5
4)L204NT14	23.0
3)A141	
	47.5cM

U13

Marker	cM
3*)K265	17.7
3*)A186	0.0
4)L63	2.3
1*)HSP176	0.8
3*)K644a	14.9
3*)A708	22.9
4)L195	
	58.6cM

U14

Marker	cM
4)L50MT3	3.7
4)L50ME31	21.8
*G173Tb	6.3
1*)Satt6	3.1
3*)A489	6.0
4)L103b	0.4
3*)K385	9.3
3)A802	0.3
3)B174a	1.2
3*)A363b	
	52.1cM

U15

Marker	cM
1)Sat39	7.6
W1	17.5
2)BLT30	3.1
4)G214MT9	0.6
1)Satt40	0.1
1)Satt30	1.4
4)G214MH12	4.9
1)Gmrubp	
	35.1cM

U16

Marker	cM
1)Satt2	25.6
3)A124	3.3
4)G214MH6	2.4
3)A401	6.8
3)A64a	5.1
4)L72	
	43.2cM

U17

Marker	cM
3)A109b	5.6
2)BLT2	2.6
4)L204NT11	15.4
3*)A955a	9.9
1)Satt49	5.5
3*)K644b	
	39.1cM

U18

Marker	cM
4)G214MH2	3.8
4)G214MB15	0.0
4)G214MT	2.7
4)L194a	2.6
3)A53a	22.3
3)A135b	0.0
4)T183	
	31.5cM

U19

Marker	cM
4)L50ND21	1.6
3)K11MB11	7.1
1)Satt41	29.2
1)Sat69	22.3
3*)A343	11.5
3*)A135a	0.3
2)BLT13	
	72.0cM

U20

Marker	cM
3)A262d	1.5
4)L199a	2.9
3)A59	29.2
3)A121	
	33.7cM

U21

Marker	cM
2)BLT27	1.1
4)T10	11.1
3)A955b	15.2
1)Scaa1	
	27.4cM

U22

Marker	cM
1)SOYGPATR	3.0
3)A351	7.0
3)A463	13.7
3)K1	
	23.7cM

U23

Marker	cM
3)A381	2.8
4)R249	13.5
4)L185	
	16.2cM

U24

Marker	cM
3*)A661	0.6
1*)Sat20	16.2
*)R	
	16.8cM

U25

Marker	cM
3)A725	15.8
1)Satt95	
	15.8cM

U26

Marker	cM
3)A234	8.7
4)G214MH14	1.6
4)G214MB13	0.8
4)G214MT5	
	11.0cM

U27

Marker	cM
4)R56	11.4
4)L50ME32	
	11.4cM

U28

Marker	cM
4)L14	6.7
3)A74	1.8
4)L175	
	8.5cM

U29

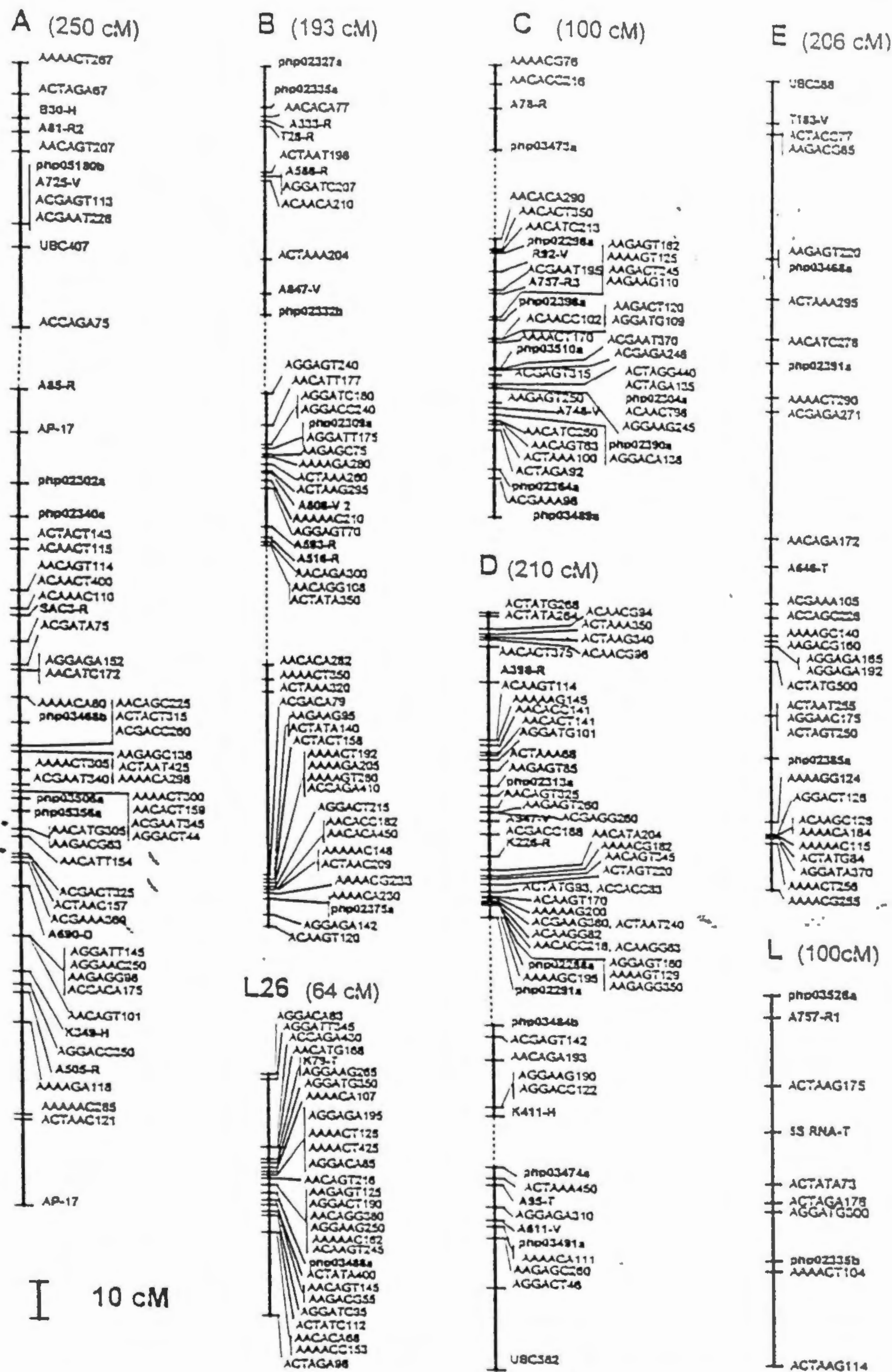
Marker	cM
4)T98	1.8
2)BLT53MT18	1.9
2)BLT53NT17	
	3.7cM

U30

Marker	cM
4)L191	2.7
4)G173Ta	0.4
4)T270	

Linkage groups and map distances for the RI population, Minsoy - Noir 1. Linkages and map distances were determined as described in Materials and Methods. The total linkage in this map is 1981 cM. SSR markers (1) were identified using the primer sequences in Table 1; RFLP probes are available from B. Matthews (2), R. Shoemaker (3), or K.G. Lark (4) (see Materials and Methods). (*) denotes markers (RFLP & SSR) that were used in comparisons between linkage groups in this map and the ISU, interspecific, map (see text). BLT53, G214, K11, L50, L26, and L204 are probes which hybridize with multiple loci and have a dominant genotype. We have changed the nomenclature for these loci from the one used previously (Lark et al., 1993). In each case, the loci now are designated by two alphabetical indices and one numerical index. The first alphabetical index, in italics, represents the parent in which the dominant fragment is found (*M* = Minsoy; *N* = Noir). The second is the restriction enzyme used to digest the plant DNA (B = *Bcl*I; D = *Dra*I; E = *Eco*RI; H = *Hind*III; T = *Taq*I). The numerical index corresponds to the rank of the molecular weight of the restriction fragment, in the parent which displays the dominant marker (polymorphic fragment), relative to all of the fragments found for that parent (i.e., 1 if it is the largest restriction fragment, 2 if it is the second largest etc.). Thus the marker G214*M*H3, in linkage group U4, is represented by the third largest radioactive restriction fragment which is found when Minsoy DNA is digested with *Hind*III and the digest hybridized to the radioactive probe G214. Previously, we described several QTL which interacted to affect height (Lark et al., 1995). These included L50t3, now designated L50*M*T3; BLT531, now BLT*N*T17; 214H10 now G214*N*H6; 214H16 now G214*M*H12; and L50Dr9 now L50*N*D13).

Figure 1-6. Soybean genetic linkage map constructed in the University of Utah. This figure is reproduced from Mansur *et al.*, 1996.



A genetic map of soybean. Twenty-eight genetic linkage groups are shown with either their homologous name from the public soybean map (e.g., A) or an arbitrarily assigned number (e.g., L1). Some marker groups had been previously assigned to the same linkage group (Shoemaker and Specht, 1995) but were not significantly linked in this study (A, B, C, D, and I). These have been retained as a single group but attached by a dashed line to indicate the lack of linkage in this study. Markers indicated in bold text were evaluated in the full 300 RIL population and used in the construction of the scaffold map. This map and updated versions can be found in SOYBASE (http://probe.nalusda.gov:8300/cgi-bin/dbrun/soybase?find + Map_collection).

Figure 1-7. Soybean genetic linkage map constructed in the Northern Arizona University. This figure is reproduced from Keim *et. al.*, 1997.

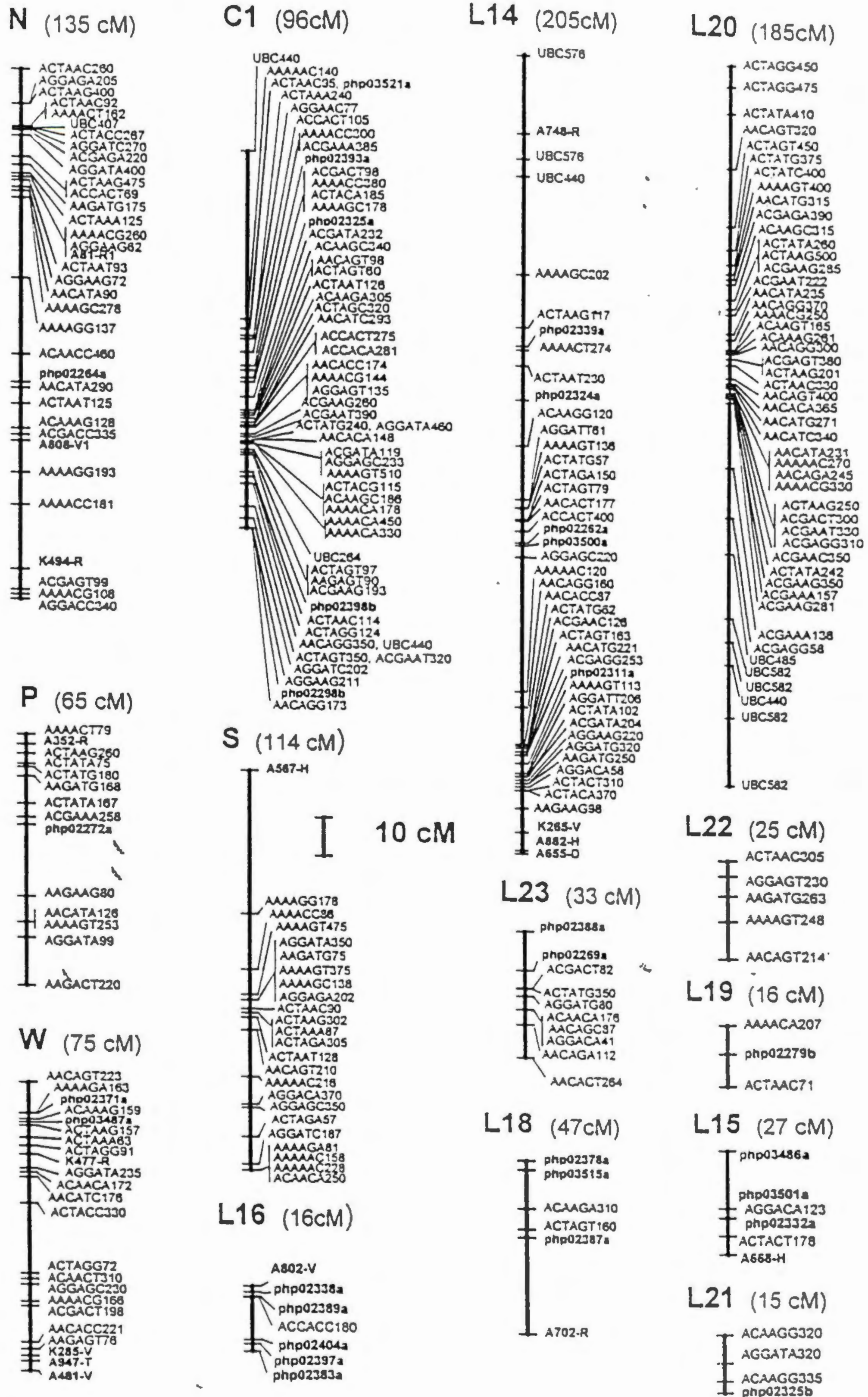


Figure 1-7 (continued)

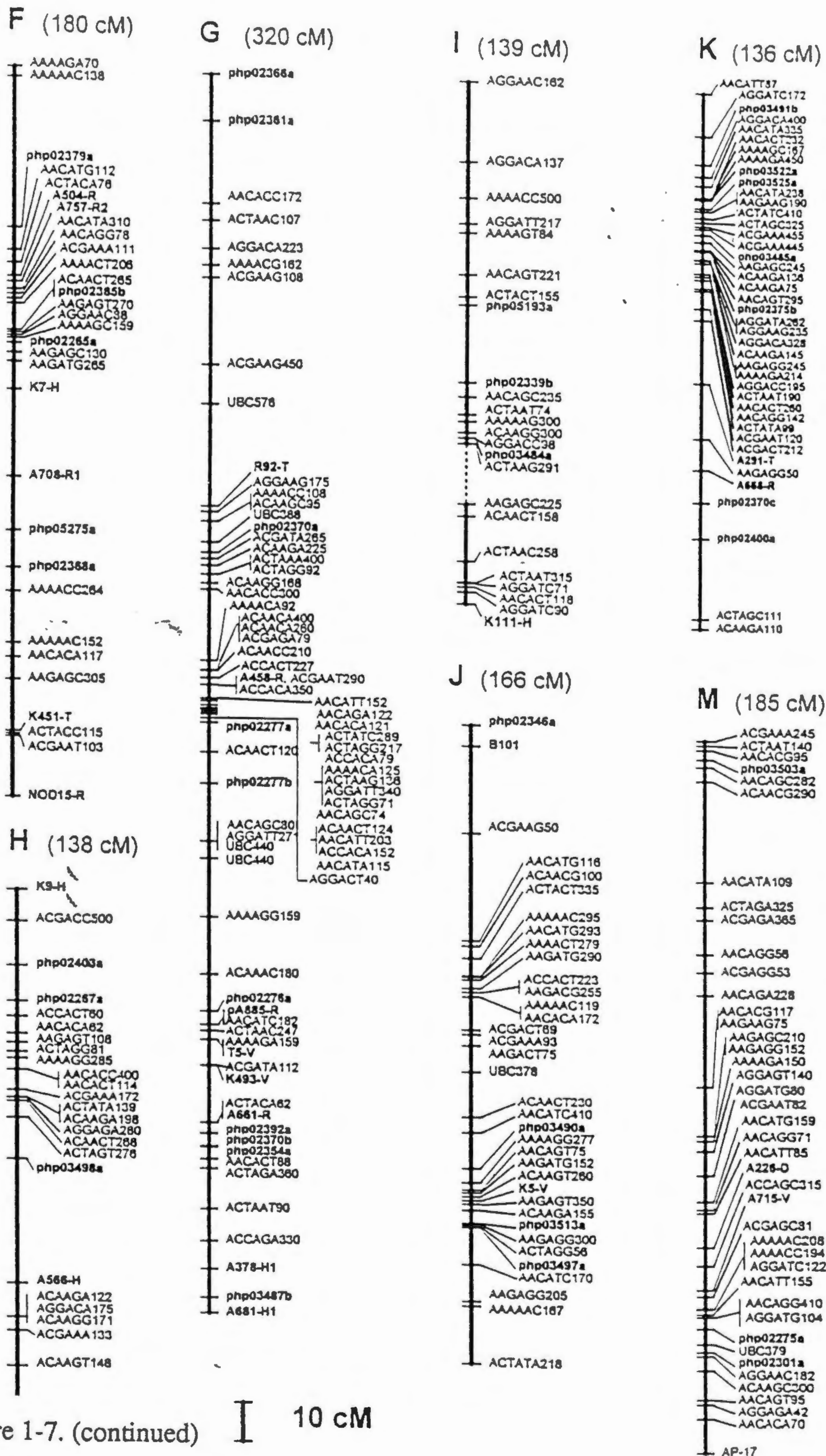


Figure 1-7. (continued)

(Keim *et al.*, 1997). The markers include 165 RFLP, 25 RAPD, and 650 AFLP markers of which AFLP markers were very useful markers to saturate the genetic linkage map. This map was developed initially based on the RFLP markers from a 300 RIL population generated from a cross between BSR110 X PI437.654. Then, the map was saturated further with AFLP markers based on a 42 RIL subset.

Molecular Genetic Linkage Maps in Other Legumes

In this part, examples of genetic linkage maps constructed on other legumes such as alfalfa (Kiss *et al.*, 1993), common or French bean (Nodari *et al.*, 1993), pea (Weeden *et al.*, 1996), lentil (Weeden *et al.*, 1992), chickpea (Simon and Muehlbauer, 1997), azuki bean (Kaga *et al.*, 1996), and *Lotus japonicus* (Jiang and Gresshoff, 1997) are reviewed.

A genetic map of alfalfa, *Medicago sativa*, has been constructed using morphological, isozymes, RFLP, and RAPD markers (Kiss *et al.*, 1993). Mapping was conducted on 138 F_2 segregating individuals obtained from one self-pollinated F_1 plant. More than 1,000 genetic markers were mapped to eight linkage groups corresponding to the eight chromosomes of alfalfa. The alfalfa genome size is about 550 cM, and the physical equivalent of 1 cM is 1,500 kb. More than 20 nodulin genes have been mapped including *ENOD2*, *ENOD12*, *Nod-22*, *Nod-25*, leghemoglobin, glutamine synthase, and ineffective nodulation, *in6*. The later, a recessive Mendelian locus, was mapped on linkage group 7 and linked to two RFLP markers (Kiss *et al.*, 1993).

In pea, a genetic linkage map (Figure 1-8) consisting of eight linkage groups with more than 500 classical, isozyme, RFLP, AFLP, and SCAR markers has been constructed (Weeden *et al.*, 1996). About 30 nodulation genes, *sym*, have been identified by mutagenesis (Duc and Messager, 1989; Kneen *et al.*, 1994; Sagan *et al.*, 1994). These genes are randomly distributed on the eight pea linkage groups except for *sym2*, *sym5*, *sym19*, *nod3*, and a major leghemoglobin locus which are clustered on the linkage group I (Temnykh *et al.*, 1995; Weeden *et al.*, 1990). The late nodulin genes such as glutamine synthetase, leghemoglobin, and *PsNOD6* have been also described (Kardailsky *et al.*, 1993; Nap, 1988; Tingey *et al.*, 1987).

In common or French bean (*Phaseolus vulgaris*, 2n=2x=22) an RFLP-based genetic

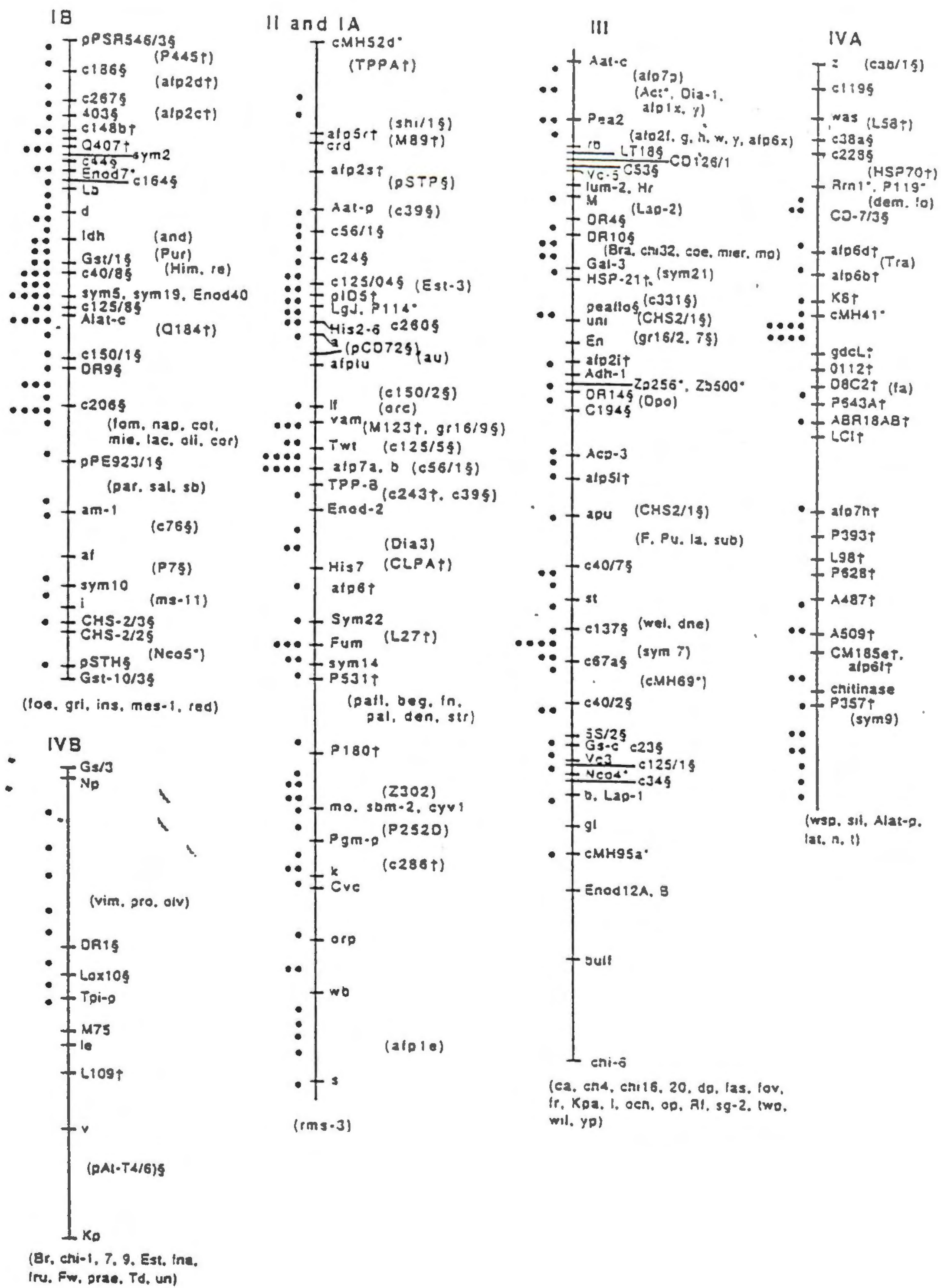


Figure 1-8. A genetic linkage map of pea (*Pisum sativum*). Eight linkage groups are designated as IB, II and IA, II, IVA, IVB, V, VI, VII. Dashed lines represent lack of markers in the specific region. This figure reproduced from Weeden *et al.*, 1996.

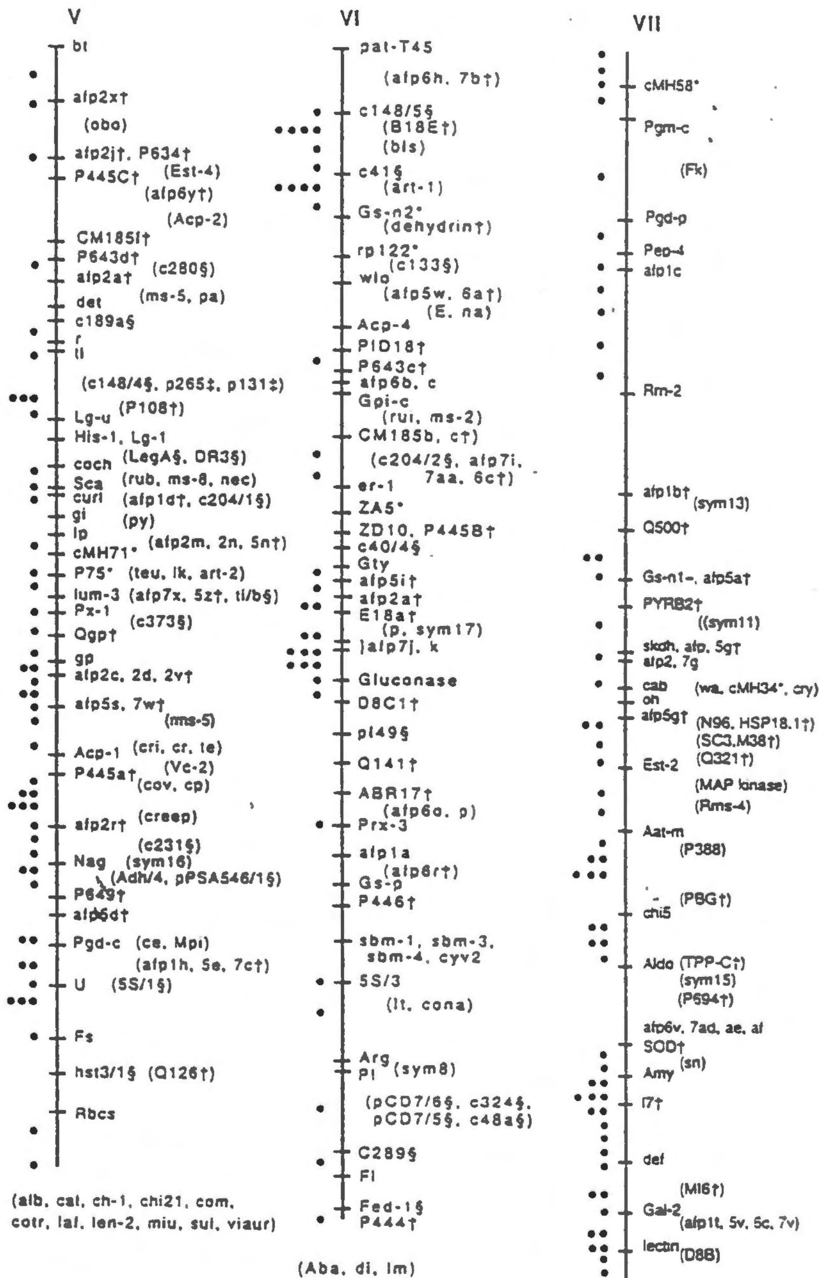


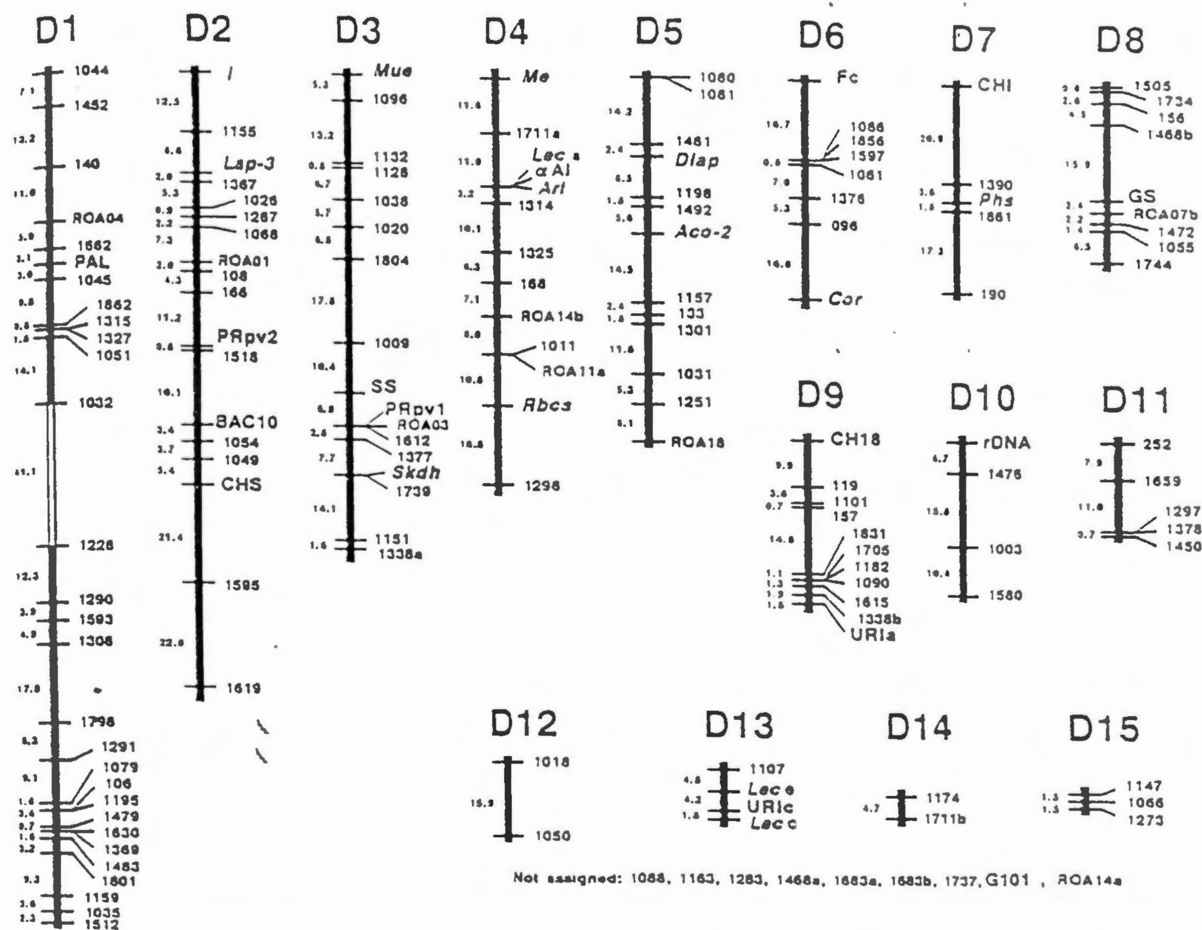
Figure 1-8. (continued)

linkage map (Figure 1-9) has been developed consisting of 15 linkage groups and 152 markers covering more than 800 cM (Nodari *et al.*, 1993). The markers include 115 RFLP, 7 isozyme, 8 RAPD, 2 morphological loci, and 19 known genes as well as one virus resistance gene. The map was constructed from an F_2 segregating population.

In chickpea (*Cicer* sp.) a genetic map consisting 9 morphological, 27 isoenzyme, 10 RFLP, and 45 RAPD markers covering 550 cM has been developed (Simon and Muehlbauer, 1997). This genetic linkage map containing 10 linkage groups representing chickpea eight chromosomes generated from a segregating population of an interspecific cross between *C. arietinum* and *C. reticulatum*. Comparison of this map with those from pea and lentil revealed that the chickpea map has 5 regions with the similar gene orders found in pea genetic linkage map. There is less similarity between the pea and the lentil genetic linkage maps which is consistent with the evolutionary distances between genomes of these three genera.

In azuki bean, *Vigna angularis*, a genetic linkage map has been developed with 132 markers consisting 108 RAPD, 19 RFLP, and 5 morphological markers (Kaga *et al.*, 1996). This map bears 14 linkage groups covering 1250 cM and was constructed based on an F_2 segregating population of an interspecific cross between azuki bean and its wild type, *V. nakashimae*. A comparison of the genetic linkage maps of azuki bean, mungbean, and cowpea using 20 common RFLP markers showed that some of the markers belong to the same linkage groups of the respective maps suggesting that some of the genomic regions are conserved among the three *Vigna* species (Kaga *et al.*, 1996).

Recently, *Lotus japonicus* has been considered as the model legume plant because of its advantages such as small genome size (about 400 Mb/haploid genome), self-fertile diploid (2n=12), short generation period, easy emasculation and cross hybridization (Jiang and Gresshoff, 1997), and high frequency of regeneration and transformation (Stiller and Gresshoff, 1997). Primary classical and genetic linkage map of this model legume (Figure 1-10) consists of 9 linkage groups containing more than 50 molecular and phenotypical markers covering more than 350 cM (Jiang and Gresshoff, in press). This map has been constructed based on mostly DNA amplification fingerprinting (DAF) markers screened in a F_2 population of 100 individuals resulted from a cross between Gifu and Funakura ecotypes.



RFLP-based linkage map of common bean. Linkage group numbers are indicated on *top*. Kosambi map distance are indicated at the *left side* of each interval between two markers. Genomic markers are numbered from *001* to *999* (from *Eco*RI-*Bam*HI library) and from *1001* to *1862* (from *Pst*I library) (the GUC prefix was removed for the sake of simplicity). *Aco-2*, *Diap*, *Lap-3*, *Me*, *Mue*, *Rbcs*, and *Skdh* are isozyme loci. RAPD marker loci are designated by *ROA* followed by a *number* (see text for explanations). Morpho-agronomic traits are *C*, *Cor*, and *I* (see text for explanations). The following markers represent sequences coding for products of known function: *Arl*, *αAI*, BAC10, CHI, CHS, CH18, GS, *Lec*, PAL, *Phs*, PRPv1, PRPv2, rDNA, SS, and URI (see text for explanations). The *open bar* in linkage group D1 represents an interval with a LOD score above 3.0 but with a recombination frequency above 30%

Figure 1-9. A linkage map of common bean. This figure reproduced from Nodari *et al.*, 1993.

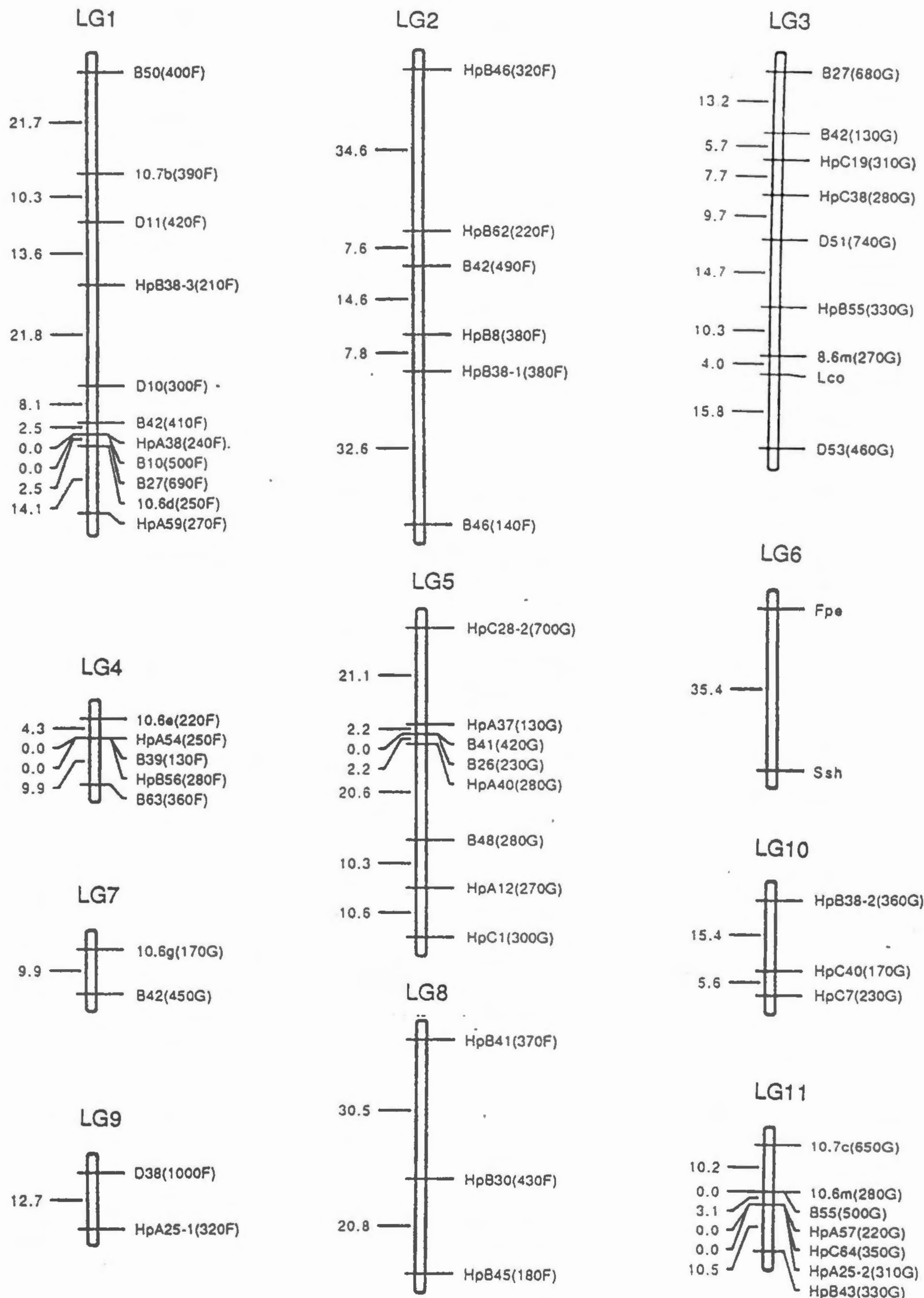


Figure 1-10. A genetic linkage map of *Lotus japonicus*. This figure reproduced from Jiang and Gresshoff, 1998.

Outline of Dissertation

The objective of this study was to map two soybean nodule-specific genes, an early *enod2* and a late leghemoglobin, on the soybean genetic linkage map using RFLP technology, integration mapping, microsatellite marker, bulked segregant analysis (BSA), and DNA amplification fingerprinting (DAF) approaches. First, RIL population was considered for mapping the *enod2* and leghemoglobin because this map contains several hundreds of molecular markers and gives the better linkage estimation. Second, using flanking markers on the genetic linkage map for *enod2*, I mapped this gene on a F2 segregating population generated in the Plant Molecular Genetic lab, University of Tennessee, Knoxville. Since this gene mapped in an interesting soybean genome containing soybean cyst nematode (SCN) resistance gene, I used DAF molecular marker and BSA to saturate the *enod2* region and detect molecular markers linked to *enod2* gene. However DAF marker generated a controversy of irreproducibility of the results. Thus, I developed a high annealing temperature DAF to increase reproducibility and reliability of DAF markers and I could apply this high stringency of DAF to a variety of organisms such as soybean, mungbean, bermudagrass, centipedegrass, garlic, and human.

During the end of last decade, RFLP technology had been extensively used to construct a genetic linkage map in plants. The term RFLP, restriction fragment length polymorphism, is defined as the differences in molecular weight of homologous fragments of restriction enzyme-digested genomic DNA of two genetically distinct individuals. Several reasons might be involved in generating the differences such as base pair changes, rearrangement of DNA sequence, and/or insertion/deletion events at the restriction site. These differences may be codominant and inherited in a simple Mendelian fashion. Once a difference for the gene of interest was detected between parental genotypes, the location of the gene on the genetic map can be determined by cosegregation study of the gene and other markers on the map.

In addition to RFLP, a microsatellite marker was used to determine which of the two copies of the *enod2* gene was mapped. Since only one of the genomic clone contains a microsatellite motif, (AT)17, cosegregation of microsatellite and RFLP markers in parental, RIL, and segregating individuals were studied.

In this study two soybean recombinant inbred line (RIL) and F_2 segregating population were used. The significance of RIL population is its immortality and one can reproduce and share the plant materials and DNA indefinitely because of near homozygosity of the lines. In addition, there is more chance of segregation and recombination in RIL population due to several meiosis events during the inbreeding process. It has been shown that the proportion of recombination among self-pollinated inbreds is about twice the rate observed in F2 segregating population (Haldane and Waddington, 1931) resulting in decrease of linkage errors for target gene and marker association in RIL population compared to F2 or backcross populations.

To assure the reliability of the mapping, the integration mapping concept was used to determine the location of the *enod2* gene on the genetic linkage map of an F2 segregating population generated at the University of Tennessee, Knoxville, from a cross between *Glycine max* cultivar Bragg and *G. soja*, PI468.397. In this approach, markers flanking the gene on the genetic linkage map were used to study the cosegregation patterns of the flanking markers and the gene in the F2 population and integrate the mapping information from one to another population (Figure 1-11)

Bulked segregant analysis (BSA; Michelmore *et al.*, 1991) was used to link additional molecular markers to the *enod2* gene. This methodology has the advantage of reducing identification of unlinked markers to the target region. BSA has been used successfully to link molecular markers to resistance genes in different crop species such as lettuce (Michelmore *et al.*, 1991), onion (de Vries *et al.*, 1992), common bean (Miklas *et al.*, 1993), and tomato (van der Beek *et al.*, 1994).

In BSA strategy two DNA bulks from individuals segregating in a population are pooled. Each pool consists of individuals which differ for a specific phenotype, genotype, or individuals at either extreme of a segregating population for a quantitative trait locus (QTL). There is a minor chance that polymorphism appears for a region unlinked to the target locus because many individuals are pooled to generate the bulks.

The two bulk DNA can be screened for detection of polymorphic markers using any molecular marker technology such as RFLP, RAPD, AFLP, microsatellite, and DAF. Upon detection of polymorphism between parents and the two bulks, the percentage of recombination can be determined in the segregating individuals to determine genetic

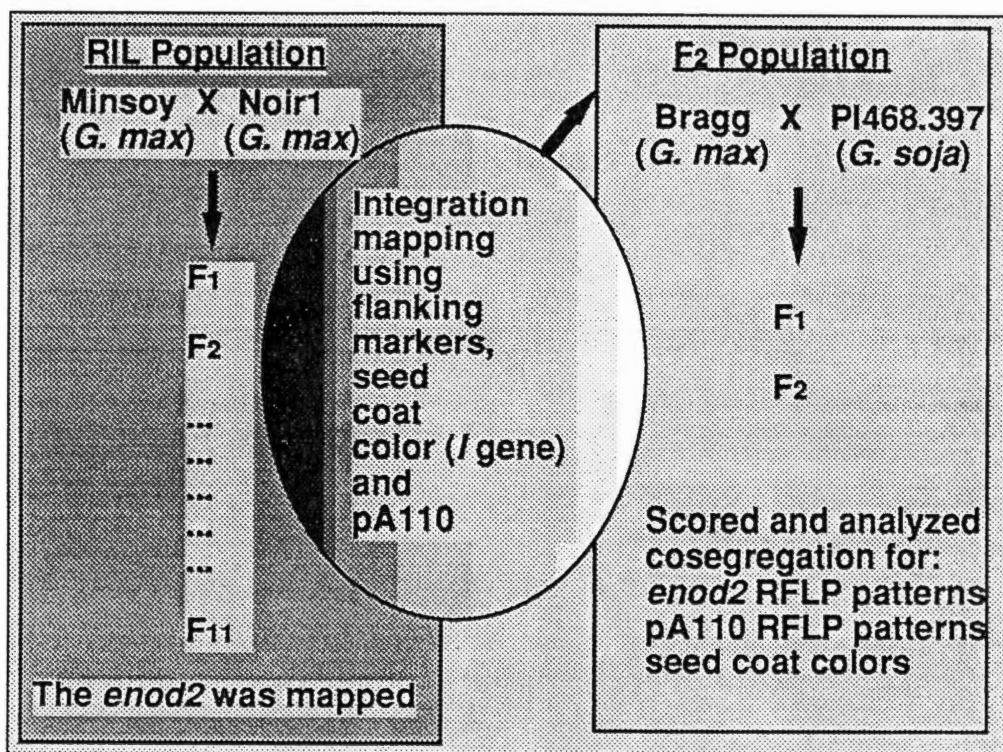


Figure 1-11. Integration mapping approach in soybean. Information obtained from *enod2* gene mapping on the RIL population (at the left) was used to determine flanking markers (at the middle). Using this information and cosegregation analysis of flanking markers for the *enod2*, this gene was mapped on the F_2 population.

distance between molecular marker and the target locus. To integrate this information onto a genetic linkage map, cosegregation of the molecular marker and other molecular markers on the map can be studied.

To accomplish saturation of the genomic region, DNA amplification fingerprinting, DAF, (Caetano-Anollés *et al.*, 1991d) methodology was approached. This PCR-based DNA amplification, a powerful technique to generate molecular markers linked to a gene of interest, applies very short arbitrary primers (ranging from 5 to 10 bases), low template DNA concentration, polyacrylamide gel electrophoresis (PAGE), and silver staining visualization. First, a robust and high annealing temperature DAF was developed by manipulating the annealing temperature and time period of denaturation, extension, and annealing steps. Then the two pools of template DNA were screened with arbitrary octamer and hairpin primers.

DNA amplification fingerprinting was tested in addition to soybean for a variety of organisms, such as soybean cyst nematode (SCN), mycorrhizal fungi, aphid, bermudagrass, and centipedegrass to show the universal applications of the DAF.

References

Akao, S. and Kouchi, H. 1992. A supernodulating mutant isolated from soybean cultivar Enrei. Soil Sci. Plant Nutr. 38: 183-187.

Allen, N. S., Ehrhardt, D. W., Long, S. R. 1991. Early cellular responses to perfusion of soluble nod gene dependent factors in the *Rhizobium meliloti*-alfalfa symbiosis. Plant Physiol. (Supplement) 96: 42.

Appleby, C. A. 1984. Leghemoglobin and *Rhizobium* respiration. Ann. Rev. Plant Physiol. 35: 443-478.

Baker, D. D. and Mullin, B. C. 1992. Actinorhizal symbioses. In: Biological Nitrogen Fixation (Stacey, G., Burris, R. H., Evans, H. J., eds.) Chapman and Hall, New York, pp. 259-292.

Bauer, W. D., Bhuvaneswari, T. V., Calvert, H. E., Law, I. J., Malik, N. S. A., and Vesper, S. J. 1985. Recognition and infection by slow-growing *rhizobia*. In: Nitrogen Fixation Research Progress (eds. H. J. Evans, P. J. Bottomley, and W. E. Newton). Martinus Nijhoff Publishers, Dordrecht, Netherlands, pp. 247-253

Bennett, M. J., Lightfoot, D. A., and Cullimore, J. V. 1989. cDNA sequence and differential expression of the gene encoding the glutamine synthetase g-polypeptide of *Phaseolus vulgaris* L. Plant Mol. Biol. 12: 553.

Bergmann, H., Preddie, E., and Verma, D. P. S. 1983. Nodulin-35: a subunit of specific uricase (uricaseII) induced and localized in the uninfected cells of soybean nodules. EMBO J. 2: 2333.

Blank, C., de., Mylona, P., Yang, W. C., Katinakis, P., Bisseling, T., and Franssen, H. 1993. Characterization of the soybean early nodulin cDNA clone *GmENOD55*. Plant Mol. Biol. 22: 1167-1171.

Blauenfeldt, J., Joshi, P. A., Gresshoff, P. M., and Caetano-Anollés, G. 1994. Nodulation of white clover (*Trifolium repens*) in the absence of *Rhizobium*. Protoplsma 179: 106-110.

Bloemberg, G. V., Kamst, E., Harteveld, M., van der Drift, K. M. G. M., Haverkamp, J., Thomas-Oates, J. E., Lugtenberg, B. J. J., and Spaink, H. P. 1995. A central domain of *Rhizobium* NodE protein mediates host specificity by determining the hydrophobicity of fatty acyl moieties of nodulation factors. Mol. Microbiol. 16: 1123-1136.

Bolanos-Vassquez, M. C. and Werner, D. 1997. Effects of *Rhizobium tropici*, *R. etli*, and *R. leguminosarum bv. phaseoli* on *nod* gene-inducing flavonoids in root exudates of *Phaseolus vulgaris*. Mol. Plant-Microbe Interac. 10: 339-346.

Borthakur, D., Barber, C. E., Lamb, J. E., Daniels, M. J., Downie, J. A., and Johnston, A. W. B. 1986. A mutation that blocks exopolysaccharide synthesis prevents nodulation of peas by *Rhizobium leguminosarum* but not of beans by *R.*

phaseoli and is corrected by cloned DNA from *Rhizobium* or the phytopathogen *Xanthmonas*. Mole. Gen. Genet. 203: 320-323.

Brewin, N. Rae, A., L., Perotto, S., Knox, J. P., Roberts, K., Legal, M. F., Sindhu, S. S., Wood, E. A., and Kannenberg, E. L. 1990. Immunological dissection of the plant-microbe interface in pea nodules. In: Nitrogen Fixation: Achievements and Objectives (ed. P. M. Gresshoff, L. E. Roth, G. Stacey, and W. E. Newton), pp. 227-234. Chapman and Hall, New York.

Brisson, N. and Verma, D. P. S. 1982. Soybean leghemoglobin gene family: Normal, pseudo, and truncated genes. Proc. Natl. Acad. Sci. USA 79: 4055-4059.

Buttery, B. R. and Park, S. J. 1989. Effect of nitrogen, inoculation and grafting on the expression of supernodulating in a mutant of *Phaseolus vulgaris* L. Can. J. Plant Sci. 70: 375-381.

Buzzell, R. I., Buttery, B. R., and Ablett, G. 1990. Supernodulating mutants in Elgin 87 soybean. In: Nitrogen Fixation: Achievements and Objectives (eds. P. M. Gresshoff, L. E. Roth, G. Stacey, and W. E. Newton), p 726. Chapman and Hall, New York and London.

Caetano-Anollés, G. 1997. Molecular dissection and improvement of the nodule symbiosis in legumes. Field Crop Res. 53: 47-68.

Caetano-Anollés, G. and Gresshoff P. M. 1992. Plant genetic suppression of the non-nodulating phenotype of *Rhizobium meliloti* host-range nodH mutants: gene-for-gene interaction in the alfalfa-*Rhizobium* symbiosis? Theor. Appl. Genet. 84: 624-632.

Caetano-Anollés, G., Joshi, P. A., and Gresshoff P. M. 1992. Nodulation in the absence of *Rhizobium*. In: Current Topics in Plant Molecular Biology, vol. 1, Plant Biotechnology and Development (ed., P. M. Gresshoff) CRC Press, Boca Raton, pp. 61-70.

Caetano-Anollés, G. and Gresshoff P. M. 1991a. Efficiency of nodule initiation and autoregulatory responses in a supernodulating soybean mutant. Appl. Environ. Microbiol. 57: 2205-2210.

Caetano-Anollés, G. and Gresshoff P. M. 1991b. Excision of nodules induced by *Rhizobium meliloti* exopolysaccharide mutants releases autoregulation in alfalfa. J. Plant Physiol. 138: 765-767.

Caetano-Anollés, G. and Gresshoff P. M. 1991c. Alfalfa controls nodulation during the onset of *Rhizobium*-induced cortical cell division. Plant Physol. 95: 366-373.

Caetano-Anollés, G., Bassam, B., and Gresshoff P. M. 1991d. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Biotechnology 9: 553-557.

Caetano-Anollés, G., Crist-Estes, D. K., and Bauer, W. D. 1988. Chemotaxis of *Rhizobium meliloti* to the plant flavone luteolin requires functional nodulation genes. J. Bacteriol. 170: 3164-3169.

Caetano-Anollés, G. and Bauer, W. D. 1988. Feedback regulation of nodule formation in alfalfa. Planta 175: 546-547.

Carroll, B. J. and Mathews, A. 1990. Nitrate inhibition of nodulation in legumes. In: Molecular Biology of Nitrogen Fixation: Achievements and Objectives (eds. P. M. Gresshoff, L. E. Roth, G. Stacey, and W. E. Newton), pp 158-180. Chapman and Hall, New York and London.

Carroll, B. J., McNeil, D. L., and Gresshoff, P. G. 1985a. A supernodulation and nitrate tolerant symbiotic (*nts*) soybean mutant. Plant Physiol. 78: 34-40.

Carroll, B. J., McNeil, D. L., and Gresshoff, P. G. 1985b. Isolation and properties of soybean (*Glycine max* L. Merr.) mutants that nodulate in the presence of high nitrate concentrations. Proc. Natl. Acad. Sci. USA 82: 4162-4166.

Carroll, B. J., McNeil, D. L., and Gresshoff, P. G. 1986. Mutagenesis of soybean (*Glycine max* L. Merr.) and the isolation of non-nodulating mutants. Plant Sci. 40: 109-114.

Clover, R. H., Kieber, J., and Signer, E. R. 1989. Lipopolysaccharide mutants of *Rhizobium meliloti* are not defective in symbiosis. J. Bacteriol. 171: 3961-3967.

Cock, J. M., Mould, R. M., Bennett, M. J., and Cullimore, J. V. 1990. Expression of glutamine synthetase genes in roots and nodules of *Phaseolus vulgaris* following changes in the ammonium supply and infection with various *Rhizobium* mutants. Plant Mol. Biol. 14: 549-560.

Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, N., and VandenBosch, K. 1995. Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti* . Plant Cell 7: 43-55.

Cregan, P. B., Bhagwat, A. A., Akkaya, M. S., and Rongwen, J. 1994. Microsatellite fingerprinting and mapping of soybean. Methods Mol. Cell. Biol. 5: 49-61.

Csanadi, G., Szecsi, J., Kalo, P., Kiss, P., Endre, G., Kondorosi, A., and Kiss, G, Y. 1994. *ENOD12*, an early nodulin gene, is not required for nodule formation and efficient nitrogen fixation in alfalfa. Plant Cell 6: 201-213.

Dazzo, F. B., Truchet, G. L., Sherwood, J. E., Hrabak, E. M., Abe, M., and Pankratz, S. H. 1984. Specific phases of root hair attachment in the *Rhizobium trifolii*-clover symbiosis. Appl. Environ. Microbiol. 48: 1140-1150.

Dean, D. R. and Jacobsen, M. R. 1992. Biochemical genetics of nitrogenase. In: Biological Nitrogen Fixation (eds. G. Stacey and R. H. Burris), pp. 763-834. Chapman and Hall, New York.

Debelle, F., Plazanet, C., Roche, P., Pujol, C., Savagnac, A., Rosenberg, C., Prome, J. C., and Denarie, J. 1996. The NodA proteins of *Rhizobium meliloti* and *Rhizobium tropici* specify the N-acylation of Nod factors by different fatty acids. Mol. Microbiol. 22: 303-314.

Dehio, C. and de Bruijn, F. J. 1992. The early nodulin gene *SrEnod2* from *Sesbania rostrata* is inducible by cytokinin. The Plant Journal 2: 117-128.

Delauney, A. J. and Verma, D. P. S. 1988. Cloned nodulin genes for symbiotic nitrogen fixation. Plant Mol. Biol. Rep. 6: 279-285.

Delves, A. C., Higgins, A., and Gresshoff, P. M. 1987a. Shoot control of supernodulation in a number of mutant soybeans (*Glycine max* L. Merr.) Austral. J. Plant Physiol. 14: 689-694.

Delves, A. C., Higgins, A., and Gresshoff, P. M. 1987b. Supernodulation in interspecific grafts between *Glycine max* (soybean) and *Glycine soja*. J. Plant Physiol. 128: 473-438.

Delves, A. C., Mathews, A., Day, D. A., Carter, A. S., and Gresshoff, P. M. 1986. Regulation of the soybean-*Rhizobium* nodule symbiosis by shoot and root factors. Plant Physiol. 82: 588-590.

Dénarié, J. and Cullimore, J. 1993. Lipo-oligosaccharide nodulation factors: a new class of signaling molecules mediating recognition and morphogenesis. Cell 74: 951-954.

de Vries, J., Jongerius, R., Sandbrink, H., and Lindhout, P. 1992. RAPD markers assist in resistance breeding. Prophyta 2: 50-51.

Diaz, C. L., Logman, G. J. J., Stam, J. C., and Kijne, J. W. 1995. Sugar-binding activity of pea lectin expressed in white clover hairy roots. Plant Physiol. 109: 1167-1177.

Dickstein, R., Bisseling, T., Reinhold, V. N., and Ausubel, F. M. 1987. Expression of nodule-specific genes in alfalfa root nodules blocked at an early stage of development. Genes and Development 2: 677-687.

Dixon, R. A. and Lamb, C. J. 1990. Molecular communication in interactions between plants and microbial pathogens. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41:339-367.

Duc, G. and Messager, A. 1989. Mutagenesis of pea (*Pisum sativum* L.) and the isolation of mutants for nodulation and nitrogen fixation. Plant Sci. 60: 207-213.

Duc, G. 1995. Mutagenesis of faba bean (*Vicia faba* L.) and the identification of five different genes controlling no nodulation, ineffective nodulation or supernodulation. Euphytica 83: 147-152.

Dunn, K., Dickstein, R., Feinbaum, R., Burnett, B. K., Peterman, T. K., Thoidis, G., Goodman, H. M., and Ausubel, F. M. 1988. Developmental regulation of nodule-specific genes in alfalfa root nodules. Mol. Plant-Microbe. Interac. 1: 66-74.

Dylan, T., Ielpi, L., Stanfield, S., Kashyap, L., Douglas, C., Yanofsky, M., Nester, E., Helinski, D. R., and Ditta, G. 1986. *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. Proc. Natl. Acad. Sci. USA 83: 4403-4407.

Ehrhardt, D. W., Atkinson, E. M., and Long, S. R. 1992. Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. Science 256: 998-1000.

Finan, T., M., Hirsch, A. M., Leigh, J. A., Johansen, E., Kuldau, G. A., Deegan, S., Walker, G. C., and Signer, E. R. 1985. Symbiotic mutant of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell 40: 869-877.

Fisher, R. F. and Long, S. R. 1992. *Rhizobium*-plant signal exchange. Nature 357: 655-660.

Forde, B. G., Freeman, J., Oliver, J. E, and Pineda, M. 1990. Nuclear factors interact with conserved A/T rich elements upstream of a nodule-enhanced glutamine synthetase gene from french bean. Plant Cell 2: 925-939.

Forde, B. G. and Cullimore, J. V. 1989. The molecular biology of Glutamine synthetase in higher plants. In: Oxford Surveys of Plant Molecular and Cell Biology (ed. B. J. Miflin), Oxford, Clarendon, 6: 247-296.

Fortin, M. G., Morrison, N. A., and Verma, D. P. S. 1987. Nodulin-26, a peribacteroid membrane nodulin is expressed independently of the development of the peribacteroid compartment. Nucl. Acid Res. 15: 813-824.

Franssen, H. J., Nap, J. P., and Bisseling, T. 1992. Nodulin in root nodule development. In: Biological Nitrogen Fixation (eds, Stacey, G. Burris, R. H.). Chapman and Hall, New York, pp. 598-624.

Franssen, H. J., Nap, J. P., Gloudemans, T., Stiekema, W., van Dam, H., Govers, F., Louwerse, J., van Kammen, A., and Bisseling, T. 1987. Characterization of cDNA for nodulin-75 of soybean: a gene product involved in early stages of root nodule development. Proc. Natl. Acad. Sci. USA 84: 4495-4499.

Freiberg, C., Fellay, R., Bairoch, A. Broughton, W. J., Rosenthal, A., and Perret, X. 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387: 394-401.

Fujita, K, Masuda, T., and Ogata, S. 1991. Analysis of factors controlling dinitrogen fixation in wild and cultivated soybean (*Glycine max*) plants by reciprocal grafting. Soil Sci. Plant Nutr. 37: 233-240.

Fuller, F., Kunstner, P. W., Nguyen, T., and Verma, D. P. S. 1983. Soybean nodulin genes: Analysis of cDNA clones reveals several major tissue-specific sequences in nitrogen-fixing root nodules. Proc. Natl. Acad. Sci. USA 80: 2594-2598.

Funke, R. P. and Kolchinsky, A. 1994. Plant yeast artificial chromosome libraries and their use: status and some strategic considerations. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, USA, pp. 125-134.

Gerathy, N., Caetano-Anollés, G., Joshi, P. A., and Gresshoff, P. M. 1992. Anatomical analysis of nodule development in soybean reveals an additional autoregulatory control point. Plant Sci. 85: 1-7.

Ghassemi F. and Gresshoff, P. M. The early *enod2* and the leghemoglobin (*lbc3*) genes segregate independently from other known soybean symbiotic genes. Mol. Plant-Microbe Interact. 11: 6-13.

Gloudemans, T. and Bisseling, T. 1989. Plant gene expression in early stages *Rhizobium*-legume symbiosis. Plant Sci. 65: 1-14.

Goormachtig, S., Valerio-Lepiniec, M., Szcyglowski, K., Van Montagu, M., Holsters, M., and de Bruijn, F. J. 1995. Use of differential display to identify novel *Sesbania rostrata* genes enhanced by *Azorhizobium caulinodans* infection. Mol. Plant-Microbe Interact. 8: 816-824.

Govers, F., Moerman, M., Downie, J. A., Hooykaas, P., Franssen, H. J., and Louwerse, J. 1986. *Rhizobium nod* genes are involved in inducing an early nodulin gene. Nature 323: 564- 566.

Gremaud, M. F. and Harper, J. E. 1989. Selection and initial characterization of partially nitrate tolerant nodulation mutants of soybean. Plant Physiol. 89: 169-173.

Gresshoff, P. M. and Landau-Ellis, D. 1994. Molecular mapping of soybean nodulation genes. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, USA, pp. 97-112.

Gresshoff, P. M. 1993. Molecular genetic analysis of nodulation genes in soybean. Plant Breed. Rev. 11: 275-318.

Gresshoff, P. M., Olsson, J. E., Li, Z. Z., and Caetano-Anollés, G. 1989. Root-shoot interaction in the control of nodulation in soybean. Curr. Top. Plant Biochem. Physiol. 8: 125-139.

Gresshoff, P. M., Krotsky, A., Mathews, A., Day, D. A., Schuller, K. A., Olsson, J. E., Delves, A. C., and Carroll, B. J. 1988. Suppression of the symbiotic supernodulation symptoms of soybeans. J. Plant Physiol. 132: 417-423.

Gresshoff, P. M. and Delves, A. C. 1986. Plant genetic approaches to symbiotic nodulation and nitrogen fixation in legumes. In Plant Gene Research III (eds. A. D. Blonstein and P. J. King). Springer-Verlag, Vienna, pp. 159-206.

Györgypal Z., Kiss, G. B., and Kondorosi, A. 1991. Transduction of plant signal molecules by the *Rhizobium* Nod D proteins. BioEssays 13: 1-7.

Guerinot, M. L. and Chelm, B. 1985. Bacterial cAMP and hemes in the *Rhizobium*-legume symbiosis. In: Nitrogen Fixation Research Progress (eds. H. J. Evans, P. J. Bottomley, and W. E. Newton), p. 220. Martinus Nijhoff, Dordrecht, Netherlands.

Haldane, J. B. and Waddington, C. H. 1931. Inbreeding and linkage. Genetics 16: 357-374.

Hirel, B., Bouet, C., King, B., Layzell, D., Jacobs, F., and Verma, D. P. S. 1987. Glutamine synthetase genes are regulated by ammonia provided externally or by symbiotic nitrogen fixation. EMBO J. 5: 1167-1171.

Hirsch, A. M. 1992. Development biology of legume nodulation. New Phytol. 122: 211-237.

Hirsch, A. M., Bhuvaneswari, T. V., Torrey, J. G., and Bisseling, T. 1989. Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. Proc. Natl. Acad. Sci. USA 86: 1244-1248.

Halverston, L. J. and Stacey G. 1986. Signal exchange in plant-microbe interactions. Microbiol. Rev. 50: 193-225.

Hollingsworth, R. I., Squartini, A., Philip-Hollingsworth, S, and Dazzo, F., B. 1990. Had and noi signals from *Rhizobium trifolii*. In Nitrogen Fixation: Achievements and

Objectives (eds. P. M. Gresshoff, L. E. Roth, W. E. Newton, G. Stacey), p. 260. Chapman and Hall, New York.

Jacobsen, K., Laursen, N. B., Ostergaard, J. E., Marcker, A., and Poulsen, C. 1990. HMG1-like proteins from leaf and nodule nuclei interact with different A-T motifs in soybean nodulin promoters. Plant Cell 2: 85-94.

Jeffreys, A. J., Wilson, V., and Thein, S. L. 1985. Hypervariable 'minisatellite' regions in human DNA. Nature 314: 67-73.

Jensen, E., Ø, Stougaard, J., Jorgensen, J. E., Sandal, N., de Bruijn, F., J., 1988. Regulation of nodule-specific plant genes. In: Nitrogen Fixation: Hundred Years After (eds. H. Bothe, F. J. de Bruijn, W. E. Newton). Gustav Fischer, Stuttgart:, pp. 605-609.

Jiang, Q. and Gresshoff, P. M. 1997. Classical and molecular genetics of the model legume *Lotus japonicus*. Mol. Plant-Microbe. Interact. 10: 559-568.

Jiang, Q. and Gresshoff, P. M. 1998. A genetic map of *Lotus japonicus*. Mol. Plant-Microbe. Interact. (in press).

Jorgensen, J. E., Stougaard, J., and Marker, K. A. 1991. A two-component nodule-specific enhancer in the soybean *N23* gene promoter. Plant Cell 3: 819-827.

Joshi, P., A., Caetano-Anollés, G., Graham, E. T., and Gresshoff P. M. 1992. Ontogeny and ultrastructure of spontaneous nodules in alfalfa (*Medicago sativa*). Protoplasma 162: 1-11.

Journet, E. P., Pichon, M., Dediev, A., de Billy, F., Truchet, G., and Barker, D. G. 1994. *Rhizobium meliloti* Nod factors elicit cell-specific transcription of the *ENOD12* gene in transgenic alfalfa. Plant J. 6: 241-249.

Kaga, A., Ohnishi, M., Ishii, T., and Kamijima, O. 1996. A genetic linkage map of azuki bean constructed with molecular and morphological markers using an

interspecific population (*Vigna angularis XV. nakashimae.*). Theor. Appl. Genet. 93: 658-663.

Kamst, E., Pilling, J., Raamsdonk, L. M., Lugtenberg, B. J., and Spaink, H. P. 1997. *Rhizobium* nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in Nod factor biosynthesis. J. Bacteriol. 179: 2103-2108.

Kardailsky, I., Yang, W. C., Zalensky, A., van Kammen, A., Bisseling, T. 1993. The pea late nodulin gene *PsNOD6* is homologous to the early nodulin genes *PsENOD314* and is expressed after the leghemoglobin genes. Plant Mol. Biol. 23: 1029-1037.

Katinakis, P. and Verma, D. P. S. 1985. Nodulin-24 gene of soybean codes for a peptide of the peribacteroid membrane and was generated by tandem duplication of an insertion element. Proc. Natl. Acad. Sci. USA 82: 4157-4161.

Keim, P., Schupp, J. M., Travis, S. E., Clayton, K., and Webb, D. M. 1997. A high-density soybean genetic map based upon AFLP markers. Crop Sci. 37: 537-543.

Kiss, G. B., Csanádi, G., Kálmán, K., Kaló, P., and Ökrész, L. 1993. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozymes and morphological markers. Mol. Gen. Genet. 238: 129-137.

Kiss, G. B., Vincze, E., Vegh, Z., Toth, G., and Soos, J. 1990. Identification and cDNA cloning of a new nodule-specific gene, *Nms-25* (nodulin-25) of *Medicago sativa* Plant Mol. Biol. 14: 467-475).

Kneen, B. E., Weeden, N. F., and LaRue, T. A. 1994. Non-nodulating mutants of *Pisum sativum* (L.) cv. Sparkle. J. Hered. 85: 129-133

Kondorosi, A. 1992. Regulation of nodulation genes in rhizobia. In Molecular Signals in Plant-Microbe Communications (ed. D. P. S. Verma). CRC Press, Boca Raton, FL, pp. 325-340.

Kondorosi, A., Kondorosi, E., John, M., Schmidt, J., and Schell, J. 1991. The role of nodulation genes in bacterium-plant communication. In: Genetic Engineering, vol. 13 (ed. by J. K. Setlow). Plenum Press, New York, pp. 115-136.

Kondorosi, A., Horváth, B., Göttfert, M., Putnoky, P., Rostas, K., Györgypal, Z., Kondorosi, E., Török, I., Bachem, C., John, M., Schmidt, J., and Schell, J. 1985. Identification and organization of *Rhizobium meliloti*i genes relevant to the intitation and development of nodules. In: Nitrogen Fixation Research progress (eds. H. J. Evans, P. J. Bottomley, and W. e. Newton). Martinus Nijhoff, Dordrecht/ Boston/ Lancaster., pp. 73-79.

Konieczny, A., Szcyglowski, K., Boron, L., Przybylska, M., and Legocki, A. B. 1988. Expression of lupin nodulin genes during root nodule development. Plant Sci. 55: 145-149.

Kosslak, R. M., Brookland, R., Barkei, J., Parren, H., and Applebaum, E. R. 1987. Induction of *Bradyrhizobium japonicum* common nod genes by isoflavones from *Glycine max*. Proc. Natl. Acad. Sci. USA 84: 7428-7432.

Kouchi, H. and Hata, S. 1995. *GmN56*, a novel nodule-specific cDNA from soybean root nodule encodes a protein homologous to isopropylmalate synthase and homocitrate synthase. Mol. Plant Microbe Interact. 8: 172-176.

Kouchi, H., Tsukamoto, M., and Tajima, S. 1990. Differential expression of nodule-specific (nodulin) genes in the infected, uninfected and cortical cells if soybean (*Glycine max*) root nodules. J. Plant Physiol. 135: 608-617.

Kozik, A., Matvienko, M., Scheres, B., Paruvangada, V. G., Bisseling, T., van Kammen, A., Ellis, T. H. N., LaRue, T., and Weeden, N. 1996. The pea early nodulin gene *PsENOD7* maps in the region of linkage group I containing *sym2* and leghemoglobin. Plant Mol. Biol. 31: 149-156.

Kuster, H., Schroder, G., Fruhling, M., Pich, U., Rieping, M., Schubert, I., Perlick, A. M., and Puhler, A. 1995. The nodule-specific *VfENOD-GRP3* gene encoding a glycine-rich early nodulin is located on chromosome I of *Vicia faba* L. and is

predominantly expressed in the interzone II-III of root nodules. Plant Mol. Biol. 28:405-421.

Lara, M., Cullimore, J. V., Lea, P. J., Miflin, B. J., Johnston, A. W. B., and Lamb, J. W. 1983. Appearance of a novel form of plant glutamine synthetase during nodule development in *Phaseolus vulgaris* L. Planta 157: 254-258.

Lark, K. G., Weisemann, J. M., Mathews, B. F., Palmer, R., Chase, K., Macalma, T. 1993. A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir1'. Theor. Appl. Genet. 86: 901-906.

Lauridsen, P., Franssen, H., Stougaard, J., Bisseling, T., and Marcker, K. A. 1993. Conserved regulation of the soybean early nodulin *ENOD2* gene promoter in determinate and indeterminate transgenic root nodules. Plant J. 3: 483-492.

Laursen, N. B., Larsen, K., Knudsen, J. Y., Hoffmann, H. J., Poulsen, C., Marcker, K. A., and Jensen, E. O. 1994. A protein binding AT-rich sequence in the soybean leghemoglobin c3 promoter is a general *cis* element that requires proximal DNA elements to stimulate transcription. Plant Cell 6: 659-668.

Lee, J. S., Brown, G. G., and Verma, D. P. S. 1983. Chromosomal arrangement of leghemoglobin genes in soybean. Nucl. Acids Res. 11: 5541-5553.

Lee, J. S. and Verma, D. P. S. 1984. Structure and chromosomal arrangement of leghemoglobin genes in kidney bean suggest divergence in soybean leghemoglobin gene loci following tetraploidization. EMBO J. 3: 2745-2752.

Legocki, R. P. and Verma D. P. S. 1980. Identification of 'nodule-specific' host proteins (nodulins) involved in the development of *Rhizobium*-legume symbiosis. Cell 20: 153-163.

Legocki, R. P. and Verma D. P. S. 1979. A nodule-specific plant protein (nodulin-35) from soybean. Science 205: 190-192.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., and Truchet, G. 1990. Symbiotic host-specificity of *Rhizobium melilotii* is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature 344: 781-84.

Long, S. R. 1989. *Rhizobium*-legume nodulation: life together underground. Cell 56:203-214

Long, S. R. and Cooper, J. 1988. Overview of symbiosis. In: Molecular Genetics of Plant-Microbe Interactions (ed. by R. Palacios and D. P. S. Verma). APS Press, Minnesota, pp. 163-178.

Lugtenberg, B. J. J., Diaz, C., Smit, G., de Pater, S., and Kijne, J. W. 1991. Role of lectin in the *Rhizobium*-legume symbiosis. In: Advances in Molecular Genetics of Plant Microbe Interactions, Vol. 1. (eds. H. Henneke and D. P. S. Verma). Kluwer Academic Publishers, Dordrecht, pp. 174-181

Lynn, D. G. and Chang, M. 1990. Phenolic signals in cohabitation: implications for plant development. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41: 497-526.

Manen, J. F., Simon, P., Van Slooten, J. C., Osteras, M., Frutiger, S., and Hughes, G. J. 1991. A nodulin specifically expressed in senescent nodule of winged bean is a protease inhibitor. Plant Cell 3: 259-270.

Mansur, L. M., Orf, J. H., Chase, K., Jarvik, T., Cregan, P. B., and Lark, K. G. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean [*Glycine max* (L.) Merr.]. Crop Sci. 36: 1327-1336.

Mansur, L. M., Lark, K. G., Kross, H., and Oliveira, A. 1993. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor Appl Genet 86: 907-913.

Mathews, A., Carroll, B. J., and Gresshoff, P. M. 1992. Studies on the root control of non-nodulation and plant growth of non-nodulating mutants and a supernodulating mutant of soybean (*Glycine max* L. Merr.) Plant Sci. 83: 35-43.

Mathews, A., Carroll, B. J., and Gresshoff, P. M. 1989. Development of *Bradyrhizobium* infections in a supernodulating and non-nodulating mutant of soybean (*Glycine max* L. Merr.) Protoplasma 150: 40-47.

Mellor, R. B. and Werner, D. 1990. Legume nodule biochemistry and function. In: Molecular Biology of Symbiotic Nitrogen Fixation (ed. P. M. Gresshoff). CRC Press Inc., Boca Raton, FL, USA, pp. 111-129.

Michelmore, R. W., Paran, I., Kesseli, R V. 1991. Identification of markers linked to disease resistant genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating population. Proc. Natl. Acad. Sci. USA 88: 9828-9832.

Miklas, P. N., Stavely, J. R., and Kelly, J. D. (1993). Identification and potential use of molecular marker for rust resistance in common bean. Theor. Appl. Genet. 85: 745-749.

Minami, E., Kouchi, H., Carlson, R. W., Kolli, V. K., Day, R. B., Ogawa, T., and Stacey, G. 1996. Cooperative action of lipo-chitin nodulation signals on the induction of the early nodulin, ENOD2, in soybean roots. Mol. Plant Microbe Interact. 9: 574-583.

Minami, E., Kouchi, H., Cohn, J. R., Ogawa, T., and Stacey, G. 1996. Expression of the early nodulin, ENOD40, in soybean roots in response to various lipo-chitin signal molecules. Plant J. 10: 23-32.

Moerman, M., Nap, J. P., Govers, F., Schilperoort, R., van Kammen, A., and Bisseling, T. 1987. *Rhizobium nod* genes are involved in the induction of two early nodulin genes in *Vicia sativa* root nodules. Plant Mol. Biol. 9: 171-179.

Morell, M. and Copeland, L. 1985. Sucrose synthase of soybean nodules. Plant Physiol. 78: 149-154.

Mullin, B. C., Swensen, S. M., and Goeting-Minesky, P. 1990. Hypothesis for the evolution of actinorhizal symbioses. In: Nitrogen Fixation: Achievements and

Objectives, (eds. Gresshoff, P. M., Roth, E., Stacey, G., and Newton, W. E., eds.), Chapman and Hall, New York and London, pp 781-787.

Mylona, P., Pawlowski, K., and Bisseling, T. 1995. Symbiotic nitrogen fixation. Plant Cell 7: 869-885.

Nadler, K. D. 1981. A mutant strain of *Rhizobium leguminosarum* with an abnormality in heme synthesis. In: Current Perspective in Nitrogen Fixation (eds. A. H. Gibson and W. E. Newton). Australian Academy of Science, Canberra, Australia, p. 414.

Nap, J. P. and Bisseling, T. 1990a. Nodulin function and nodulin gene regulation in root nodule development. In: Molecular Biology of Symbiotic Nitrogen Fixation (ed. P. M. Gresshoff). CRC Press, Boca Raton, FL, pp. 181-229.

Nap, J. P. and Bisseling, T. 1990b. Developmental biology of a plant-prokaryote symbiosis. Science 250: 948-954.

Nap, J. P. 1988. Isolation and analysis of a leghemoglobin gene from pea (*Pisum sativum*). In: Nodulins in Root Nodule Development, Ph. D. thesis, pp. 19-28. Wageningen Agricultural University, Netherlands.

Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H., and Verma, D. P. S. 1985. Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc. Natl. Acad. Sci. USA 82: 5040-5044.

Nirunsuksiri, W. and Sengupta-Gopalan, C. 1990. Charaterization of a novel nodulin gene in soybean that shares sequence similarity to the gene for nodulin-24. Plant Mol. Biol. 15: 835-849.

Nodari, R. O., Tsai, S. M., Gilbertson, R. L., and Gepts, P. Toward an integrated linkage map of common bean 2. Development of an RFLP-based linkage map. Theor. Appl. Genet. 85: 513-520.

Olsson, J. E., Nakao, P., Bohlool, B. B., and Gresshoff, P. M. 1989. Lack of systemic suppression of nodulation in split root systems of supernodulating soybean (*Glycine max* L. Merr.). Plant Physiol. 90: 1347-1352.

Orgambide, G. G., Philip-Hollingsworth, S., Mateos, P. F., Hollingsworth, R. I., and Dazzo, F. B. 1995. Subnanomolar concentration of membrane chitolipooligosaccharides from *Rhizobium leguminosarium bv. trifolii.* are fully capable of eliciting symbiosis-related responses on white clover. Plant and Soil 186: 93-98.

Padilla, J. E., Campos, F., Conde, V., Lara, M., and Sánchez, F. 1987. Nodule-specific glutamine synthetase is expressed before the onset of nitrogen fixation in *Phaseolus vulgaris* L. Plant Mol. Biol. 9: 65-74.

Papadopoulou, K., Roussis, A., and Katinakis, P. 1996. *Phaseolus ENOD40* is involved in symbiotic and non-symbiotic organogenetic processes: expression during nodule and lateral root development. Plant Mol. Biol. 30: 403-417.

Park, S. j. and Buttery, B. R. 1997. Complementation of nodulation genes of various mutants in common bean (*Phaseolus vulgaris* L.). J. Hered. 88: 543-545.

Perlick, A. M., Frühling, M., Schröder, G., Albus, U., Frosch, S. C., Becker, J., Böhner, S., Quandt, H. J., Broer, I., Küster, H., and Pühler, A. 1997. Analysis of genes expressed in root nodules of broad bean (*Vicia faba* L.). In: Biological Fixation of Nitrogen for Ecology and Sustainable Agriculture (eds. A. Legocki, H. Bothe, and A. Pühler). Springer-Verlag, Berlin, Heidelberg, Germany, pp. 91-94.

Perlick, A. M., Frühling, M., Schröder, G., Frosch, S. C., and Pühler, A. 1996. The Broad bean gene *VfNod32* encodes a nodulin with sequence similarities to chitinase that is homologous to $(a/b)_8$-barrel-type seed protein. Plant Physiol. 110: 147-154.

Peters, N. K. and Verma, D. P. S. 1990. Phenolic compounds as regulators of gene expression in plant-microbe interactions. Mol. Plant-Microbe Interact. 3: 4-8.

Philip-Hollingsworth, S., Orgambide, G. G., Bradford, J. J., Smith, D. K., Hollingsworth, R. I., and Dazzo, F. B. 1995. Mutation or increased copy number of *nodE* has no effect on the spectrum of chitolipooligosaccharide nod factors made by *Rhizobium leguminosarium bv. trifolii.* J. Biol. Chem. 36: 20968-20977.

Pichon, M., Journet, E. P., Dedieu, A., deBilly, F., Truchet, G. and Barker, D. G. 1992. *Rhizobium melilotii* elicits transient expression of the early nodulin gene *ENOD12* in the differentiating root epidermis of transgenic alfalfa. Plant Cell 4: 1199-1211.

Priefer, U. B. 1989. Genes involved in lipooligosaccharide production and symbiosis are clustered on the chromosome of *Rhizobium leguminosarium bv. viciae* VF39. J. Bacteriol. 171: 6161-6168.

Prinsen, E., Ghauvaux, N., Schmidt, J., John, M., Wieneke, U., De Greef, J., Schell, J., and Van Onckelen, H. 1991. Stimulation of indole-3-acetic acid production in *Rhizobium* by flavonoids. FEBS Letters 282: 53-55.

Ramlov, K. B., Laursen, N. B., Stougaard, J., and Marcker, K. A. 1993. Plant J. 4: 577-580.

Reynolds, P. H. S., Boland, M. J., Naughton, G. S., More, R. D., and Jones, W. T. 1990. Induction of ammonium assimilation: leguminous roots compared with nodules using split root system. Physiol Plant. 79: 359-367.

Richter, H. E., Sandal, N. N., Marcker, K. A., and Sengupta-Gopalan, C. 1991. Characterization and genomic organization of a highly expressed late nodulin gene subfamily in soybeans. Mol. Gen. Genet. 229: 445-452.

Rivers, R. L., Dean, R. M., Chandy, G., Hall, J. E., Roberts, D. M., and Zeidel, M. L. 1997. Functional analysis of nodulin 26, an aquaporin in soybean root nodule symbiosomes. J. Biol. Chem. 272: 16256-16261.

Robertson, J. G., Wells, B., Bisseling, T., Farnden, K. J. F., and Johnson, A. W. B. 1984. Immuno-gold localization of leghemoglobin in cytoplasm in nitrogen-fixing root nodules of pea (*Pisum sativum*). Nature 311: 254-256.

Roche, P., Maillet, F., Plazanet, C., Debelle, F., Ferro, M., Truchet, G., Prome, J. C., and Denarie, J. 1996. The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants. Proc. natl. Acad. Sci. USA 93: 15305-15310.

Rolfe, B. G. and Gresshoff, P. M. 1988. Genetic analysis of legume nodule initiation. Ann. Rev. Plant Physiol. Plant Mol. Biol. 39: 297-319.

Roth, E., Jean, K., and Stacey, G. 1988. Homology in the endosymbiotic system: the term 'symbiosome'. In; Molecular Genetics of Plant-Microbe Interactions, (eds., R. Palacios and D. P. S. Verma). APS Press, St. Paul, MN, pp. 220-225.

Sadowsky, M. J., Cregan, P. B., Göttfert, M., Sharma, A., Gerhold, D., Rodriguez,-Quinones, F., Keyser, H. H., Hennecke, H., and Stacey, G. 1991. The *Bradyrhizobium japonicum nol A* gene and its involvement in the genotype-specific nodulation of soybeans. Proc. Natl. Acad. Sci. USA 88: 637-641.

Sagan, M., Huguet, T., and Duc, G. 1994. Phenotypic characterization and classification of nodulation mutants of pea (*Pisum sativum* L.). Plant Sci. 100: 59-70.

Sagan, M. and Gresshoff, P. M. 1996. Developmental mapping of nodulation events in pea (*Pisum sativum* L.) using supernodulating plant genotypes and bacterial variability reveals both plant and *Rhizobium* control of nodulation regulation. Plant Sci. 117: 167-179.

Sanchez, F., Padilla, J. E., Prez, H., and M. Lara. 1991. Control of nodulin genes in root-nodule development and metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42: 507-528.

Sandal, N. N., Bojsen, K., and Marcker, K. A. 1987. A small family of nodule specific genes from soybean. Nucl. Acids Res. 15: 1507-1519.

Sangwan, I. and O'Brian, M. R. 1991. Evidence for interorganismic heme biosynthetic pathway in symbiotic soybean root nodules. Science 251: 1220-1222.

Sanjuan, J., Grob, P., Göttfert, M., Henneke, H., and Stacey, G. 1994. NodW is essential for full expression of the common nodulation genes in *Bradyrhizobium japonicum*.. Mol. Plant-Microbe Interact. 7:364-369.

Schmidt, J., Wingender, R., John, M., Wieneke, U., and Schell, J. 1988. *Rhizobium meliloti nodA* and *nodB* genes involved in generating compounds that stimulate mitosis of plant cells. Proc. Natl. Acad. SCI. USA 85: 8578-82.

Scheres, B., van Engelen, F., van der Knaap, E., van de Wiel, C., van Kammen, A., and Bisseling, T. 1990. Sequential induction of nodulin gene expression in the developing pea nodule. Plant Cell 2: 687-700.

Shoemaker, R. C., Lorenzen, L. L., Diers, B. W., and Olson, T. C. Genome mapping and agriculture. 1994. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, pp. 1-10.

Shoemaker, R. C. and Specht, J. E. 1995. Integration of the soybean molecular and classical genetic linkage groups. Crop Sci. 35: 436-446.

Simon, C. J. and Muehlbauer, F. J. 1997. Construction of a chickpea linkage map and its comparison with maps of pea and lentil. J. Hered. 88: 115-119.

Simonet, P., Normand, P., Hirsch, A. M., and Akkermans, A. D. L. 1990. the genetics of the *Frankia*-actinorhizal symbiosis. In Nitrogen Fixation: Achievements and Objectives (ed. P. M. Gresshoff). CRC Press, Boca Raton, pp. 77-109.

Smit, G., Kijne, J. W., and Lugtenberg, B. J. J. 1987. Involvement of both cellulose fibrilis and a Ca^{2+}-dependent adhesion in the attachment of *Rhizobium leguminosarum* to pea root hair tips. J. Bacteriol. 169: 4294-4301.

Spaink, H. P., Bloemberg, G. V., van Brussel, A. A. N., Lugtenberg, B. J. J., , Van der Drift, K. M. G. M., Haverkamp, J., and Thomas-Oates, J. E. 1995. Host specificity of *Rhizobium leguminosarum* is determined by the hydrophobicity of highly unsaturated fatty acyl moieties of the nodulation factors. Mol. Plant-Microbe Interac. 8: 155-164.

Spaink, H. P. 1992. Rhizobial lipo-oligosaccharides: answers and questions. Plant Mol. Biol. 20: 977-986.

Stacey, G., Luka, S., Sanjuan, J., Banfalvi, Z., Nieuwkoop, A. J., and Carlson, R. 1994. *nodZ*, a unique host-specific nodulation gene, is involved in the fucosylation of the lipo-oligosaccharide nodulation signal of *Bradyrhizobium japonicum*. J. Bacteriol. 176: 620-633.

Stacey, G. 1995. *Bradyrhizobium japonicum* nodulation genetics. FEMS-Microbiol. Lett. 127: 1-9.

Stacey, G., Sanjuan, J., Luka, S., Dockendorff, T., and Carlson, R. 1995. Signal exchange in the *Bradyrhizobium*-soybean symbiosis. Soil Biol. Biochem. 27: 473-483

Stiller, J., Martirani, L., Tuppale, S., Chian, Ru-Ju, Chiurazzi, M., and Gresshoff, P. M. 1997. High frequency transformation and regeneration of transgenic plants in the model legume *Lotus japonicus*. J. Experiment. Botany 48: 1357-1365.

Streeter, J. 1988. Inhibition of legume nodule formation and N_2 fixation by nitrate. CRC Crit. Rev. Plant Sci. 7: 1-23.

Stritmatter, G., Chia, T. F., Trinh, T. H., Katagiri, F., Kuhlemeier, C. and Chua, N. H. 1989. Characterization of nodule-specific cDNA clones from *Sesbania rostrata* expression of the corresponding genes during the initial stages of stem nodules and root nodule formation. Mol. Plant-Microbe Interac. 2: 122- 127.

Sutherland, T. D., Bassam, B. J., Schuller, L. J., and Gresshoff, P. M. 1990. Early nodulation signals of the wild type and symbiotic mutants of soybean (*Glycine max*). Mol. Plant-Microbe Interact. 3: 122-128.

Szczyglowski, K., Shaw, R. S., Wopereis, J., Copeland, Sue, Hamburger, D., Kasiborski, B., Dazzo, F. B., and de Bruijn, F. J. 1998. Nodule organogenesis and symbiotic mutants of model legume *Lotus japonicus*. Mol. Plant-Microbe Interac. 11: 684-697.

Taller, B. J. and Sturtevant, D. B. 1991. Cytokinin production by rhizobia. In: Advances in Molecular Genetics of Plant-Microbe Interactions, vol. 1 (ed. by H. Hennecke and D. P. S. Verma). Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 215-221.

Temnykh, S. V., Kneen, B. E., Weeden, N. F., and LaRue, T. A. 1995. Localization of *nod-3*, a gene conditioning hypernodulation, and identification of a novel translocation in *Pisum sativum* L. cv. 'Rondo'. J. Hered. 86: 303-305.

Thummler, F. and Verma, D. P. S. 1987. Nodulin 100 of soybean is the subunit of sucrose synthase regulated by availability of free heme in nodules. J. Biol. Chem. 262: 14730-14736.

Tingey, S. V., Walker, E. L., and Coruzzi, G. M. 1987. Glutamine synthetase genes of pea encode distinct polypeptides which are differentially expressed in leaves, roots, and nodules. EMBO J. 6: 1-9.

Trinick, M. J. 1979. Structure of nitrogen-fixing nodules formed by *Rhizobium* on roots of *Parasponia andersonii*. Can. J. Microbiol. 25: 565-578.

Truchet, G., Roche, P., Lerouge, P., Vasse, J., Camut, S., De Billy, F., Promé, J-C, and Dénarié, J. 1991. Sulphated lipooligosaccharide signals of *R. meliloti* elicit root nodule organogenesis in alfalfa. Nature 351: 670-673.

Truchet, G., Barker, D. G., Camut, S., De Billy, F., Vasse, J., and Huguet, T. 1989. Alfalfa nodulation in the absence of *Rhizobium* . Mol. Gen. Genet. 219-65-68.

VandenBosch, K. A. and Newcomb, E. H. 1988. The occurrence of leghemoglobin protein in the uninfected interstitial cells of soybean root nodules. Planta 175: 442-451.

van der Beek, J. G., Pet, F., and Lindhout, P. 1994. Resistance to powdery mildew (*Oidium lycopersicum*) in *Lycopersicon hirsutum* is controlled by an incompletely dominant gene *Ol-1* on chromosome 6. Theor. Appl. Genet. 89: 467-473.

van de Wiel, C. Scheres, B. Franssen, H., van Lierop, M. J., van Lammeren, A., van Kammen, A., and Bisseling, T. 1990. The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. EMBO J. 9: 1-7.

Vegh, Z., Vincze, E., Kadirov, R., Toth, G., and Kiss, G. B. 1990. The nucleotide sequence of a nodule-specific gene, *Nms-25* of *Medicago sativa*: its primary evolution via exon-shuffling and retrotransposon-mediated DNA rearrangements. Plant Mol. Biol. 15: 295-306.

Verma, D. P. S. 1992. Signals in root nodule organogenesis and endocytosis of *Rhizobium*. Plant Cell 4: 373-382.

Verma, D. P. S., Miao, G. H., Cheon, C. I., and Suzuki, H. 1990. Genesis of root nodules and function of nodulins. In: Advances in Molecular Genetics of Plant-Microbe Interactions (eds. H. Henneke and D. P. S. Verma), Vol. 1. Kluwer Academic Press, Dordrecht, Netherlands, pp. 291-299.

Vesper, J. S. and Bauer, W. D. 1986. Role of pili (fimbrae) in attachment of *Bradyrhizobium japonicum* on soybean roots. Appl. Environ. Microbiol. 52: 134-141.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zebeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. 23: 4407-4414.

Waterer, J. G., Vessey, J. K., and Raper, C. D. Jr. 1992. Stimulation of nodulation in field peas (*Pisum sativum*) by low concentrations of ammonium in hydroponic culture. Physiolgia Plantarum 86: 215-220.

Weaver, C. D., Shomer, N. H., Louis, C. F., and Roberts, D. M. 1994. Nodulin 26, a nodule-specific symbiosome membrane protein from soybean, is an ion channel. J. Biol. Chem. 269: 17858-17862.

Webb, D. M., Baltazar, B. M., Rao-Arelli, A. P., Schupp, J., Clayton, K., Keim, P., and Beavis, W. D. 1995. Genetic mapping of soybean cyst nematode race-3 resistance loci in the soybean PI437.654. Theor. Appl. Genet. 91:574-581.

Weeden, N. F., Swiecicki, W. K., Timmerman-Vaughan, G. M., Ellis, T. H. N., and Ambrose, M. 1996. The current pea linkage map. *Pisum* Genet. 28: 1-4.

Weeden, N. F., Muehlbauer, F. J., and Ladizinsky, G. 1992. Extensive conservation of linkage relationships between pea and lentil genetic maps. J. Hered. 83: 123-129.

Weeden, N. F., Kneen, B. E., and LaRue, T. A. 1990. Genetic analysis of *sym* genes and other nodule-related genes in *Pisum sativum*. In: Nitrogen Fixation: Achievements and Objectives (eds. P. M. Gresshoff, L. E. Roth, G. Stacey, and W. E. Newton). Chapman and Hall, New York and London, pp 323-330.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucl. Acids Res. 18: 7213-7218.

Welters, P., Felix, G., Schell, J., and de Bruijn, F. J. 1990. Specific interactions of rhizobial proteins (trans-acting factors) with distinct DNA sequences in the promoter regions of *Sesbania rostrata* leghemoglobin (*lb*) genes.In Nitrogen Fixation: Achievements and Objectives (eds. P. M. Gresshoff, L. E. Roth, W. E. Newton, G. Stacey). Chapman and Hall, New York, p. 775.

Werner, D. 1992. Physiology of the nitrogen-fixing legume nodules: compartments and functions. In: Biological Nitrogen Fixation (eds. G. Stacey and R. H. Burris). Chapman and Hall, New York, pp. 399-431.

Williams, L. E. and Lynch, D. L. 1954. Inheritance of a non-nodulating character in the soybean. Agron. J. 46: 28-29.

Wilson, R. C., Long, F., Maruoka, E. M., and Cooper, J. B. 1994. A new proline-rich early nodulin from *Medicago trancatula* is highly expressed in nodule meristamic cells. Plant Cell 6: 1266-1275.

Wittenberg, J. B. and Wittenberg, B. A. 1990. Mechanisms of cytoplasmic hemoglobin and myoglobin function. Ann. Rev. Biophys. Chem. 19: 217-241.

Xue, Z. T., Larsen, K., and Jochimsen, B. U. 1991. Oxygen regulation of uricase and sucrose synthase synthesis in soybean callus tissue is exerted at the mRNA level. Plant Mol. Biol. 16:899-906.

Zaat, S. A. J., Wijffelman, C. A., Mulders, I. H. M., Van Brussel, A. A. N., and Lugtenberg, B. J. J. 1988. root exudates of various host plants of *Rhizobium leguminosarum* contain different sets of inducers of *Rhizobium* nodulation genes. Plant Physiol. 86: 1298-1303.

Part 2

Molecular Markers

Molecular markers are defined as markers at the DNA level or protein level (allozymes) which can be generated through different techniques. These molecular markers are used to map a gene of interest, (see Gresshoff and Landau-Ellis, 1994), to facilitate gene isolation, differentiate genotypes, tag a phenotype of interest for breeding purposes, provide evolutionary relatedness and diversity of a given organism at different systematic classifications, and solve forensic cases.

A genetic map can be constructed by the study of cosegregation of phenotypic and/or molecular genetic marker in segregating F_2 or recombinant inbred line populations. Independent segregation of markers indicates that these markers are on two chromosomes or on one chromosome with 50 percent recombination rate in that chromosome region. In contrast, linkage of markers concludes that these markers are on the same chromosome. The distance between two markers is determined by the rate of recombination events in region. A complete genetic map requires several hundred markers to cover the entire genome. Order of genetic markers is determined in a large population because many recombination events are needed to obtain statistical support for a distance of about 5 cM apart. Construction of physical map is required for determination of genetic order of loci less than 1 cM apart. [One cM (centimorgan) is equal to one percent of recombination].

Genetic analysis of quantitative trait loci (QTLs) is possible by genomic mapping approach where phenotypic effects can be correlated with segregating molecular markers to map the QTL on a genetic map. Using synteny mapping, in cereal crops such as rice, maize, and sorghum, some important agronomic traits (large seeds and day length insensitivity) have been associated independently with QTLs (Paterson *et al.*, 1995). Hence, important genes for a QTL can be mapped in a less characterized species by reference to the markers on the map of a more characterized species.

Molecular marker technologies have been used widely in genotype differentiation of a vast varieties of organisms. Crop cultivar identification allows farmers to be assured that cultivars introduced for sale are of correct genotypes. Traditionally, morphological markers provided identification of varietal genotype and purity. However, molecular markers reveal genetic differences more quickly and accurately, eliminating effects of the environment. This provides significant advantages in reliability and discrimination reducing time and cost. Molecular markers are also used for plant variety protection to

maintain inbred parents of hybrids as trade secrets, to screen varieties sold by competitor breeders, seed producers, and farmers who sell their excess of seeds for replanting. This protection provide assurance for private investment to continue their efforts into breeding programs.

Application of molecular markers to diversity question may concern artificial or cultivated populations including accessions, collections, germplasm and breeding lines. DNA markers can resolve how many different genetic classes are present and the genetic similarities among them, how much diversity is present in those classes and their evolutionary relationships with wild types. In plant breeding programs, study of genetic relationship is useful because it provides more efficient parental selection (Anderson *et al.*, 1993).

Marker-assisted selection enhances the efficiency of selection for a trait of interest. Exotic germplasm is one of the most important source of genes with highly qualitative effects on traits such as disease and stress resistance. These kinds of genes can be transferred with crossing and backcrossing to the desired cultivar. Molecular markers availability in marker-breeding approach provides an increase in the use of exotic germplasm and consequently widening the gene pool for breeding programs (see Allen, 1994; Shoemaker *et al.*, 1994)).

Types of Molecular Markers

There are three classes of molecular markers: (1) hybridization-based markers, (2) arbitrary-primed PCR and other PCR-based markers, and (3) sequence targeted and single locus PCR.

Class 1. Hybridization-Based Molecular Markers

RFLP

This class consists of restriction fragment length polymorphism, RFLP, (Botstein *et al.*, 1980) technology in which genomic DNA is digested with restriction enzyme which cleave the DNA at specific sequences. Subsequently, different length of DNA is generated in different genotypes if a change in DNA sequence is generated by point

mutation, insertion, deletion, or/and DNA rearrangement at restriction site of one genotype. The digested DNA can be separated by agarose gel electrophoresis and transferred to Nylon or nitrocellulose membranes. The polymorphisms in genotypes can be detected upon hybridization where a radioactive-labeled DNA probe forms a double-stranded DNA with homologous DNA on the membrane. Autoradiography is used to visualize the RFLP on an X-ray film. The membrane is reusable for other hybridizations after washing off the DNA probe. One of the advantages of a RFLP markers is their codominant nature. The technology is easy and reproducible and there is no need for DNA sequence information. However, there are disadvantages such as low number of data points, high cost, need for large amount of DNA, and lack of automation (also phospho-imager usage can help). There are so many references of RFLP available in data bases that it is impossible to five a comprehensive review here (see Shoemaker and Olson, 1993; Landau-Ellis *et al.*, 1991).

Mini and Microsatellite Probes

Minisatellite or microsatellite probes can also be used to generate variable number of tandem repeats (VNTR). This technique is a derivative of RFLP analysis but differs in the type of hybridization probe applied to detect DNA polymorphisms. The probe creates complex banding patterns recognizing multiple genomic loci simultaneously. Each of these loci is characterized by an array of tandemly DNA repeats which occur in different numbers at different loci. (Weising and Kahl, 1998)

Two types of multilocus probes are mainly used. The first type is a cloned DNA fragment or synthetic oligonucleotide which is complementary to tandem repeats of a sequence about 10 to 60 base pair long called minisatellite. This approach was first used for DNA fingerprinting thirteen years ago (Jeffreys *et al.*, 1985a, b). Minisatellites have been cloned from many organisms such as rice (Winberg *et al.*, 1993), *Arabidopsis* (Tourmente *et al.*, 1994), tomato (Broun and Tanksley 1993 and 1996), and fungus (Meyer *et al.*, 1991). The second type of probe is a short oligonucleotide which is complementary to tandem repeats of about one to five base pairs called microsatellite, simple sequence repeats (SSRs), or short tandem repeats (STRs). Both minisatellite and microsatellite have been used for fingerprinting of numerous animal, plant, and fungal species (see Rosewich and McDonald 1994; Weising *et al.*, 1995).

Class 2. PCR-based Molecular Markers, Arbitrary Primers

In this class of molecular markers, PCR-based techniques use arbitrary or semi-arbitrary primers for amplification of DNA. One of the great advantages of class 2 is the lack of need for sequence information from the genome under study. There are different approaches in this class varying in stringency of the PCR conditions, sequence and length of the primers, and the way of fragment separation and visualization.

Arbitrarily-primed PCR (AP-PCR) is a powerful technique that generates fingerprints of genome under conditions where the primers will anneal the template DNA even when the matching is imperfect. Amplification of genomic DNA results in multiple amplification products from loci distributed throughout the genome. Based on the specific amplification conditions, product separation, and detection technique, the arbitrary primer amplification methods were termed randomly amplified polymorphic DNA (RAPD; Williams *et al.*, 1990), arbitrarily primed PCR (AP-PCR; Welsh and McClelland, 1990), or DNA amplification fingerprinting (DAF; Caetano-Anollés *et al.*, 1991). These techniques are different in the length of primers used, the primer-to-template ratios, and the way that amplification products are detected and resolved. These techniques have been used for gene mapping, taxonomy, phylogenetics, clinical epidemiology, and detection of mutations in cancer.

RAPD Marker

RAPD markers have been extensively used for more than 3,000 cases in mapping traits in segregating populations and near isogenic lines, mapping traits using bulk segregant analysis, generating genetic linkage maps, saturating regions of a genome with markers, and fingerprinting of genetic materials. In addition, RAPD analysis has been applied for measurement of genetic distances between individuals, germplasm analysis, and evaluating parental contributions in backcrosses.

RAPD marker application requires an optimization particularly for DNA, magnesium, and primer concentration, cycling conditions, and the type and amount of thermostable DNA polymerase. Factors that affect reproducibility of RAPD profiles within and between laboratories are DNA quality and concentration, type of thermocycler, primer

quality and concentration, magnesium concentration, type of DNA polymerase, and pipetting accuracy.

RAPD analysis offers advantages such as lack of need for prior DNA sequence information, radioactive material, and expensive equipment other than a thermocycler and a transilluminator. Arbitrary primers are universal and work for any genome resulting multiplex detection of polymorphisms. In addition, a RAPD test is simple and requires only small amount of genomic DNA (30 ng).

Despite many advantages of RAPD markers, there are some limitations. Reproducibility is one of disadvantages of RAPD markers and has been a subject of considerable discussion (Skroch and Nienhuis, 1995). In case of genome mapping, a proper statistical method should be used to determine and confirm segregation ratios of 3:1 in an F_2 or 1:1 in a backcross population.

RAPD is not an appropriate technique when two genomes with extremely small difference in their genomes are being compared. In general, RAPD is not able to detect a single mutation or a very small deletion. However, this technique can identify efficiently dispersed differences constituting a significant portion of the genomes. In near-isogenic lines generating from several backcross containing 1-10% of the donor genome, RAPD can be efficiently used to detect marker linked to segments of donor genome. Furthermore, RAPD can be applied for bulked segregant analysis to identify markers closely linked to a trait or a gene of interest (Michelmore *et al.*, 1991).

RAPD tends to underestimate genetic distances between more distantly related individuals particularly in inter-specific comparisons (Powell *et al.*, 1996). Therefore, in taxonomic studies above the species level, RAPD should be used with caution. Similar mobility of an incorrect band is another problem which can be overcome by using Southern hybridization for further verification. This is seldomly done.

Randomly amplified microsatellite polymorphism (RAMPO)

RAMPO is relatively a new technique which is based on the combination of a PCR-based, RAPD, and a hybridization-based, RFLP, techniques (Richardson *et al.*, 1995; Cifarelli *et al.*, 1995; Ender *et al.*, 1996). In the first step, a genomic DNA is amplified

using an arbitrary decamer primer and RAPD fingerprinting. In the second step, a labeled simple sequence repeat (SSR), called also microsatellite, oligonucleotide serves as a probe to detect complementary sequences upon Southern hybridization on the RAPD fragments. In the last step, positive hybridized RAPD fragments are cloned and sequenced to analyze the nature of the simple sequence repeats.

Advantages and Disadvantages

Advantages of RAMPO are high discriminatory potential, versatility of application to DNA of various complexity and origin, rapid screening of a genome, and excellent reproducibility. In contrast to RFLP, there is no need for neither sequence information nor laborious cloning.

RAMPO application in chickpea cultivars (*Cicer arietinum*) produced 35 detectable hybridization signals using 38 microsatellite probes and 14 different restriction enzymes (Sharma *et al.*, 1995). Oligonucleotide probes complementary to all possible microsatellites of the mono, di, tri, and tetranucleotide repeats, e. g. $[A]_{18}$, $[ATG]_6$, $[AAG]_6$, $[GTG]_6$, $[GGAT]_4$, and $[AAAC]_4$, were used to screen *Eco*RI and *Hin*dIII-restricted *Arabidopsis* and yeast DNA by southern analysis (Depeiges *et al.*, 1995). Only 9 out of 49 probes generated clear fingerprints indicating that different results obtained with chickpea and *Arabidopsis* are most likely due to differences in the size and complexity of the two genomes.

Level of polymorphism and pattern complexity detected by DNA fingerprinting with both mini- and microsatellite probes depends on population or species under study and its reproductive biology such as selfing, outcrossing, apomixis, and vegetative propagation. The sequence of probe used for hybridization is also important for levels of polymorphism.

The RAMPO discriminatory potential has suggested its use in fingerprinting of human and other species. Unrelated individuals as well as first order relatives could be differentiated with an individual-specific fingerprint (Jeffreys *et al.*, 1985a, b; 1991), consequently, paternity and forensic testing became the first application areas for DNA fingerprinting. The stability of fingerprints made this technique highly useful for identification and differentiation of banana (Kaemmer *et al.*, 1992), *Achillea*, a

micropropagated plant, (Wallner et al., 1996), ***Rubus***, an apomictic plant, (Kraft *et al.*, 1996), and asexually propagating phytopathogenic fungi (Morjane *et al.*, 1994; Sastry *et al.*, 1995). In vegetatively propagated plants, this fingerprinting technique fulfills maximum variation between cultivars, minimum variation within cultivars, stability over time, and independence from environmental factors. In contrast, in sexually outcrossing propagated plants, fingerprints are usually variable within cultivars.

The RAMPO fingerprinting have been also used to study genetic relatedness between organisms (Kaemmer *et al.*, 1995) and genetic diversity among and within populations (Piquot *et al.*, 1996) using the extent of band sharing between individuals. Furthermore, high sensitivity of the RAMPO enabled its use for linkage analysis and genome mapping with rapidly screening of the genome with few probes in human (Wells *et al.*, 1989) and fungi (Romao and Hamer 1992).

Disadvantages of RAMPO are relatively complex experimental protocol, need of higher amount of DNA compared to PCR-based techniques. In RAMPO, there is insufficient allelic information provided by multilocus banding patterns. Hence, banding pattern is considered as dominant marker. Mutation rate is an additional problem in genetic mapping of human. Mutation in minisatellites occurs at rates from 0.5-1% per gamete and generation (Jeffreys *et al.*, 1991), but can be as high as 5% (Jeffreys *et al.*, 1989). Another disadvantage is considerable proportions of nonparental bands observed in plants (Rogstad 1994) and fungi (DeScenzo and Harrington 1994). More serious problem for mapping is the tendency of mini- and microsatellite-derived fingerprint bands to occur in clusters contrast to the much shorter PCR-detected microsatellites (Bell and Ecker 1994). Clustering of fingerprints in pea occurred for four $(GAA)_5$ polymorphic bands and mapped to the same linkage group within (Dirlewanger *et al.*, 1994), while in tomato, (GATA)4 bands formed several clusters in certain linkage groups (Arens *et al.*, 1995). Moreover, clustering has happened in other organisms such as dog (Jeffreys and Morton 1987) and swan (Meng *et al.*, 1990).

AP-PCR Markers

Four types of changes can be distinguished in comparison of particular DNA fingerprint: (1) new bands in a particular sample or loss of bands; (2) molecular change of amplified fragments reflected by change in the mobility of bands; (3) increase in intensity of a

band; (4) decrease in intensity of a band. These changes can indicate polymorphic markers for mapping or somatic mutations in comparison of normal and tumor tissue. AP-PCR can detect changes such as allelic losses or gains and deletion or insertion mutations which occur in neoplasms (Perucho *et al.*, 1995). This molecular marker has potential to study tumor development and progression and possibly provides clues for the understanding of carcinogenesis. Tumor-specific genetic changes can be easily detected by comparing DNA fingerprints of normal and tumor tissue from the same individual. Such polymorphic bands resulted from somatic mutations have been cloned and characterized.

AP-PCR has been also used for RNA fingerprinting which brings an exciting area to study differential gene expression (Welsh *et al*, 1995; Suzuki *et al.*, 1995; Ralph *et al.*, 1993; Wada-Kiyama *et al.*, 1992). A 5' anchor primer such as oligo (dT)CA is used for reverse transcription and an arbitrary primer for priming the second strand cDNA (Liang and Pardee, 1992). RNA arbitrarily primed PCR (RAP-PCR) can provide a complex molecular phenotype reflecting changes in the abundance of hundreds of RNA species under different conditions. Differential gene expression occurs at different biological situations such as different types of tissues and cells, and cells responding to growth factors, hormones, stress, heterologous expression of particular genes. RNA fingerprints can be compared from different treatments to conclude regulation of the gene under study. Differentially-expressed DNA fragments can be isolated, cloned, and characterized.

In RAP-PCR, fingerprinting of total cellular RNA is achieved by synthesis of first-strand cDNA using an arbitrary primer. Using a thermostable DNA polymerase, synthesis of second-strand is initiated at the sites where the arbitrary primer finds the best matches. Weaker matches at one end of the amplified sequence can be compensated for by very good matches at the other end. A collection of molecules flanked at their 3' and 5' ends by the exact sequence and complement of the arbitrary primer is resulted after these two steps. These strands serve as template for high stringency PCR amplification resulting in RNA fingerprinting similar to genomic DNA fingerprinting. Open reading frames are found in about 30% of products because the primers are internal to the transcripts. Difference in intensity of the RAP-PCR products is due to different ratio of RNA abundance. Intensity of different bands within the same

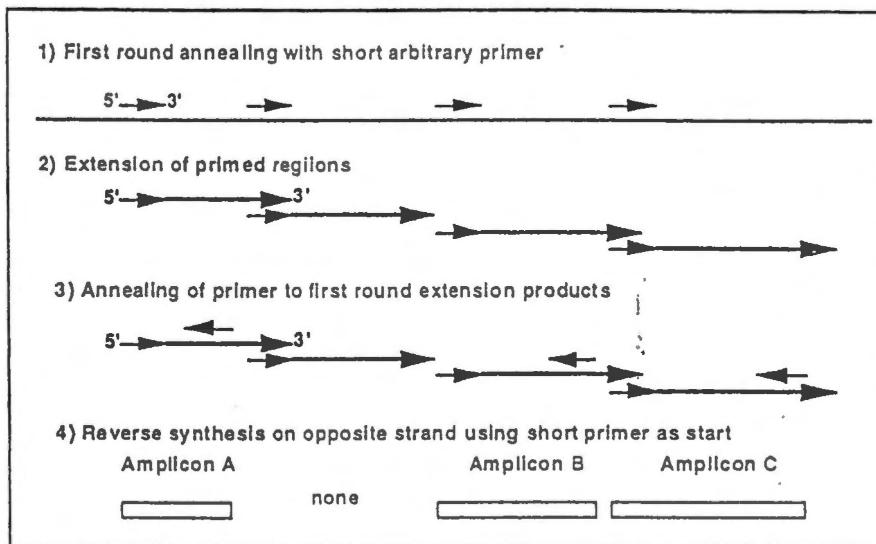
fingerprint varies independently while the intensity of a band between fingerprints might be proportional to the concentration of its corresponding template sequence.

Differential gene expression can be studied with other methods of which there are two major categories, subtractive hybridization (Aasheim *et al.*, 1994) and differential screening (Maser and Calvet, 1995). Each one has its strengths and weaknesses. One of the problems with subtractive hybridization is that abundant genes hybridize faster and to a greater rate of completion than low abundant genes. Thus, many regulatory genes that are in low abundance can not be detected due to low completion rate of hybridization. RAP-PCR, in contrast, leads to partial abundance normalization and consequently rare transcripts can be sampled. Another problem of subtractive hybridization is that transcripts which do not show significant differences can be easily missed. However, using an improved version called represential difference analysis, RDA, might overcome this problem (Lisitsyn *et al.*, 1993).

In differential screening (Maser and Calvet, 1995), radioactive probes made from cDNA of two cell types are used to screen a cDNA library prepared from one of the two cell types. Usually, clones from the library hybridize to one or the other but not to both probes. Similar to subtractive hybridization, low abundance messages do not provide enough probe signals to allow favorable hybridization. An alternative approach for AP-PCR for detecting differentially expressed genes is based on differential hybridization of complex cDNA probes to dot blot clones (Bernard *et al.*, 1996).

DAF Markers

DAF is initiated by a template screening phase including primer-template-enzyme interactions (Figure 2-1). In the first round of reaction, annealing of a single arbitrary primer to complementary sequence of DNA template generates single-stranded amplification products which contain palindromic termini causing template-template interactions and formation of hairpin loops and duplexes. In the second round, the primer has to recognize and displace these structures and allow enzyme anchoring and primer extension. DNA strands containing complementary inverted sequence for the primer can be successfully amplified. Different species of amplification products tend to establish and equilibrium in subsequent rounds of amplification while the rare duplexes of primer-template are transformed into amplification products.



General DAF reaction for production of multiple arbitrary amplicon profiling. (Inverted primer sequence exits at each end of the amplicon. Amplicons differ in size and sequence composition).

Figure 2-1. General DAF reaction.

The mass ratio between primer and the template is one of the most important variable in the amplification reaction providing adjustment for the overall stringency of the amplification reaction. One of the major differences in DAF and RAPD is the ratio of primer-to-template which is 5-50,000 and 1, respectively. Length of the arbitrary primer is also important in influencing the rate of mismatch priming in the amplification reaction.

Three types of DAF can be used to generate fingerprints: (1) short or mini-hairpin primed DAF, (2) arbitrary signature from amplification products (ASAP), and (3) template endonuclease cleaved multiple arbitrary amplicon profiling (tecMAAP). Arbitrary primers containing hairpin loop structure at their 5' termini can be used for fingerprinting of any type and complexity of genome (Caetano-Anollés and Gresshoff, 1994) lowering some of the potential limitations of the fingerprinting reaction. In ASAP, original amplification products generated by PCR are again subjected to PCR amplification using new primers (Caetano-Anollés and Gresshoff, 1996). The primers for ASAP are either mini-hairpin or primers that their sequence differ significantly from those used to generate initial amplification products. In tecMAAP, DNA template is first subjected to endonuclease digestion and then amplified one or more arbitrary primer (Caetano-Anollés *et al.*, 1993). These three DNA fingerprinting techniques are used for a variety of templates including whole genome, plasmids, cloned DNA, and PCR products resulting in increased levels of polymorphisms. Amplification products are separated on a polyacrylamide gel and visualized by a silver staining procedure (Figure 2-2).

DAF- Mini-hairpin Primers

Use of very short primers 5-6 base in DAF produces relatively simple banding patterns that resemble those generated in RAPD analysis. This is due to existence of palindromic termini in the amplification products causing formation of hairpin loops and subsequently lowering primer annealing efficiency. Using stable mini-hairpin primers with a loop of 3-4 base at 5' end and a 2 base stem, complication of primer-template interaction can be minimized (Caetano-Anollés and Gresshoff, 1994). Stability of structure is based on the existence of a hairpin-turn region determined by the helical motif of the stem region, the loop-closing sequence, and stacking of a loop B form structure (Hirao *et al.*, 1994). Mini-hairpin primers, HP-NNN, can be synthesized by

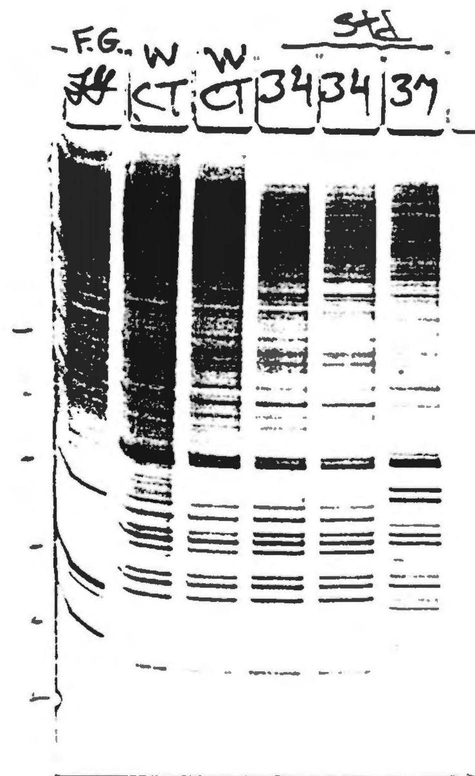


Figure 2-2. A DAF profile for human and grass samples. PCR was run for human DNA (F. G., Farshid Ghassemi) and grass samples (WCT, 34, and 37) using an octamer primer (OcB49) under high annealing temperature (55°C) condition which is discussed in Part 4 of this dissertation. DNA fragments were separated on a 10% polyacrylamide gel and stained with silver.

adding an arbitrary 3 bases core containing one of 64 possible sequence at 3' end of a constant mini-hairpin sequence. In amplification process, the core region and either the 3' terminal palindrome or the loop of the mini-hairpin structure are important but not the 5' palindromic sequence. This suggests that the hairpin structure remains in tight conformation and does not form a duplex with the template (Figure 2-3).

The advantage of these primers is to interfere with formation of hairpin loop structure in the amplification products increasing efficiency of primer-template annealing (Caetano-Anollés and Gresshoff, 1994). Thus, these primer can generate reliable DNA fingerprinting profiles from small template molecules such as plasmids, cloned DNA, and PCR products. In addition, annealing of mini-hairpin primers are influenced by secondary structure of DNA and interaction between amplicon termini.

DAF-ASAP

This kind of DNA amplification fingerprinting is based on the reamplification of PCR fingerprints using mini-hairpin or standard arbitrary primers (Caetano-Anollés and Gresshoff, 1996). ASAP is a two-step process which additional amplification of initial preselected amplicon is scanned with a substantially different primer than the one used for the first step of amplification.

Specific primers can be designed to amplify particular target sequences or interspersed repetitive sequences in the genome. Primers complementary to simple sequence repeats present in microsatellite loci can generate a simple profile by reamplification of DNA fingerprints. The advantage of this approach is that microsatellite loci are codominant and highly polymorphic regions with multiple allelic forms.

Another approach is to use a primer complimentary to sequence repeats anchored with arbitrary 5' or 3' sequences (Zietkiewicz *et al.*, 1994; Wu *et al.*, 1994; Meyer *et al.*, 1993). One limitation of this approach is that the primer anneals both the sequence repeats and unrelated arbitrary sequences resulting in DNA profiles with relatively high complexity and it is difficult to interpret the co-dominant loci (Weising *et al*, 1995). It is possible to obtain a simple DNA amplification profile representing only microsatellite loci and codominant markers by using primers which are anchored at their 5' termini

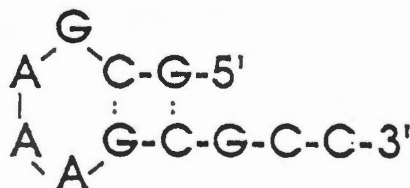
Structure of linear and mini-hairpin DAF primers

A. linear primer:

5'-GTAACGCC-3'

[try to prevent internal homology or high AT content]

B. mini-hairpin primer:



Structure of linear and minihairpin primer 8-4 and HP10, respectively. (The mini-hairpin will close early during ramping and will not be a major determinant in annealing. However, the nature of the hairpin affects the amplification pattern. We believe that the hairpin requires the presence of at least one "A" residue to permit bending and stacking)

Figure 2-3. Structure of linear and mini-hairpin primers.

with degenerate nucleotide and high stringent amplification conditions to avoid mismatch priming.

The ASAP analysis was used to generate markers linked to soybean supernodulating locus, *nts-1*, a trait that segregates as a single recessive Mendelian locus (Caetano-Anollés and Gresshoff, 1996; Kolchinsky *et al.*, 1997). This approach was coupled with bulk segregant analysis to find markers associated with *nts-1* to screen soybean yeast artificial chromosome (YAC) and bacterial artificial chromosome (BAC) libraries and anchor initial contigs for positional cloning strategy. After excision of the bands from gels, cloning, and sequencing (Men and Gresshoff, 1998; Weaver *et al.*, 1994), two polymorphic DNA converted to a sequence-characterized amplified region, SCAR, (Men and Gresshoff, 1998; Jiang and Sink, 1997; Grattapaglia *et al.*, 1996) to use for screening a large number of F_2 segregating population. One molecular marker was linked to supernodulating gene with 26 cM away from *nts-1* locus. ASAP analysis with only a few number of primers identified several markers associated with target gene despite the extremely conserved sequence of the *nts-1* region.

DAF- tecMAAP

DNA amplification fingerprinting is generated with enzymatic digestion of DNA template or amplification products. This increases the information contents such as number of bands and polymorphisms produced by a particular primer. These MAAP markers are dominant and can be used for construction of genetic linkage maps, sequence-tagged markers, and positional cloning approaches.

Detection of polymorphic DNA is enhanced in comparison of closely related organisms of various origin with pretreatment of endonuclease digestion of template DNA. Mutants and their wild types are easily distinguished by using this technique. Several MAAP markers linked to EMS-induced soybean supernodulating mutant, *nts-1*, isolated from Bragg cultivar (Caetano-Anollés *et al.*, 1993). After screening only 19 primers, 42 polymorphisms were found between nts-1 mutant and Bragg cultivar indicating that either restricted DNA template is an extremely sensitive technique or EMS mutagenesis is able to induce more extensive DNA alteration than previously expected.

One of the limitation of MAAP is that digestion of DNA template sometimes causes both disappearance and appearance of bands. Disappearance of bands can be easily explained by cleavage at restriction sites destructs the primer annealing target sites but appearance of new bands is unexpected. These bands might be amplified from those amplicons which are normally hidden by the secondary structure of DNA template. Alternatively, digestion of DNA might cause the extending primer to jump to another template during amplification to produce a hybrid product (Pääbo *et al.*, 1990). Furthermore, new bands could arise from a change in the amplification reaction because not all possible amplicons are actually amplified (Arnheim and Erlich, 1992).

The tecMAAP can be used as a tool in exploration of specific genome regions. In isolation of a given gene, this technique allows us to directly land on the locus or at least close enough to target gene. In subsequent step, the target gene can be identified using a cloned DNA fragment. This approach of gene isolation is independent of molecular maps and genome size. After cloning and sequencing, a MAAP marker can also provide an anchor for chromosome walking, marker-assisted breeding, bulk segregant analysis, and distinction of closely related organisms, cell lines, and individuals.

Amplified Fragment Length Polymorphism (AFLP)

AFLP is based on the combination of DNA restriction with endonucleases and subsequent PCR amplification (Vos *et al.*, 1995; Zabeau and Vos, 1993). In the first step, genomic DNA is restricted with two endonucleases, a frequent and a rare cutter. The frequent cutter generates small fragments which are in the optimal size for amplification and separation on the gels. The rare cutter limits the number of fragments to be amplified. In the second step, the ends of the restricted fragments are ligated to double-stranded adopters. Restriction site and adopter sequences serve as primer sites for AFLP amplification. There are also additional selective bases at the 3' end of primers which allow amplification of only a subset of restriction fragments. In the last step, those fragments with complementary sequences for the restriction site and selective nucleotides at the 3' end are amplified and generate DNA fingerprints which can be separated on denaturing polyacrylamide gels.

AFLP for small-size genome organisms, such as bacteria and fungi, uses only 2 selective bases for each primer. For organisms with complex genomes, it requires more

than 2 selective bases as well as an additional preamplification step which uses only common primers lacking the selective bases to generate optimal primer selectivity and increase the amount of template.

AFLP fragments can be detected by labeling one of the two primers by phosphorylating the 5'end with polynucleotide kinase. In addition, other methods of labeling cab be applied. The reason for labeling one fragment is that if two strands are labeled, doublets will be detected because of slightly different mobility of the two fragments on the sequencing gels.

AFLP is a technique can be used for different kind of organisms and complexity. The great advantage of AFLP over other random fingerprinting techniques, such as RAPD, AP-PCR, and DAF, is high rate of reproducibility and reliability of fingerprints. A limited number of AFLP primers can generate a large number of primer combinations which in turn amplify a unique set of fragments.

Arbitrary Sequence Oligonucleotide Fingerprinting (ASOF)

ASOF is a combination of PCR- and hybridization-based techniques using miniature arrays of oligonucleotide bound to a solid surface as probes (Beattie *et al.*, 1995). Genomic DNA is amplified by PCR to generate a set of random amplification products using one of the arbitrary-primed amplification techniques such as RAPD, AP-PCR, and DAF. Alternative choice for genomic DNA is to prepare a set of sequence tagged site (STS) fragments using targeted multiplex PCR.

The genomic DNA fragments are labeled and then hybridized to a genosensor array containing several hundred to few thousand arbitrary oligonucleotide probes. Polymorphic DNA can be determined by differences in the hybridization fingerprints of different individuals. These differences are caused by either sequence variations within the priming sites or variations within amplified sequences that hybridized to arrayed probes.

ASOF is different from RAPD, AP-PCR, and DAF in analysis of DNA fragment by hybridization fingerprinting not gel electrophoresis. Based on the statistical estimation, number of polymorphisms detected by ASOF is much greater than other techniques. A

great advantage of ASOF is that after determination of which arbitrary sequence probes provide reproducible results for a given genome, a genosensor can then be fabricated with those informative probes to analyze a large number of samples.

Class 3. PCR-based Molecular Markers, Known Sequence

In this class of molecular markers, PCR is used to amplify a specific single locus target which requires prior knowledge of the target sequence. Three sources of target sequences in animal and plants are chloroplast, mitochondrial, and nuclear genomes. DNA sequences contain phylogenetic and frequency information and are very important for ecological and evolutionary investigations. Although, sequencing is laborious and insufficient polymorphisms are detected, with new sequencing techniques, single pass sequencing of EST may yield sufficient polymorphism to allow mapping. A number of PCR-based techniques, such as CAPS, GBA, and SSR, can be used to screen markers easily.

Cleaved Amplified Polymorphic Sequence (CAPS)

In CAPS analysis or sometimes referred as PCR-RFLP, amplification products are digested with endonucleases to produce restriction site polymorphisms (Konieczny and Ausubel 1993; Williams *et al.*, 1991). There are several advantages for CAP markers. First, size of restricted and unrestricted of amplification products can be adjusted arbitrarily by proper placement of the PCR primers. Second, CAPS markers are co-dominant, i.e., a heterozygous genotype contains both cleaved and uncleaved fragments inherited from its parental genotypes. Third, CAPS technique requires only small amount of DNA which can be isolated in plant from a portion of a single leaf using a rapid DNA isolation protocol. Fourth, the technique is simple and robust because an application product is always obtained. Fifth, CAPS markers can be assayed relatively fast and have the potential for automation.

This approach has been used for gene mapping in *Arabidopsis* and DAN fingerprinting in bacteria. A set of 18 primer pairs were designed to generate CAPS markers in both Columbia and Landsburg ecotypes of *Arabidopsis thaliana*. Amplification products were restricted with a set of endonucleases to detect specific patterns for the ecotypes. After detection of polymorphisms, only 28 F_2 individuals were required to map

confidently a gene to one of the ten *Arabidopsis* chromosome. A set of 74 CAPS marker have been added to the genetic linkage map using recombinant inbred lines recently developed from Columbia x Landsburg cross.

Recently, CAPS markers were used to do fingerprinting in strains of the human bacterial pathogen *Staphylococcus epidermidis* (Calderwood *et al.*, 1996). Seven pairs of primers corresponding to 5 sequenced bacterial genes were designed to amplify DNA from 33 strains. Amplification products were digested with endonucleases resulted in detection of seven polymorphisms. Classification of strains was consistent with epidemiologic information as well as with classification based on the pulsed field gel electrophoresis.

DNA fingerprinting of bacterial strains using CAPS approach is technically simpler and more reproducible and informative than previously used techniques such as biotyping, antimicrobial sensitivity profiles, serotyping, multilocus electrophoresis, plasmid profiles, ribotyping, and pulse field gel electrophoresis. The CAPS approach can be readily used for DNA fingerprinting of individuals, strains, or any organism of which DAN sequence information is available. Alternative methods for bacterial DNA fingerprinting is PCR-based DNA amplification using arbitrary primers which was previously discussed.

Genetic bit analysis (GBA)

GBA is a non-radioactive single-base sequencing method which can be used in detection of a single nucleotide polymorphism of known sequences. The GBA relies on the robust ability of DNA polymerase to differentiate single nucleotide differences in a way that is automatable and simple to apply (Nikiforov *et al.*, 1994; Nikiforov and Rogers, 1995). First, Target DNA sequence is amplified by PCR using one exonuclease-resistant primer per set of primer pairs. Then, a single-stranded DNA template produced by exonuclease digestion is hybridized to a pre-synthesized GBA primer immobilized in the well of a microtiter plate. Using biotinylated or fluoroseinated ddNTPs and the Klenow fragment of DNA polymerase, the GBA primer is extended and then detected using enzyme-linked immunoassay (ELISA) colorimetry. Finally, sample genotypes are determined after analysis of colorimetric data.

Microsatellite or Simple Sequence Repeats (SSR) Markers

SSR is based on the amplification of variable number of tandem repeats (VNTR), also called microsatellite or short tandem repeat (STR), loci using primers that are flanking and complementary to a conserved regions of the SSR loci (Jeffreys *et al.*, 1988). This allows to amplify the entire SSR locus resulting in PCR products which are different in size based on the number of repeated DNA motifs present in the SSR locus (see Cregan *et al.*, 1994). The repetitive DNA units are only 2 to 5 base pairs in length, such as $(AT)_{17}$, $(TAA)_{22}$, etc. The high rate of polymorphism is as a result of differences in the number of such short repeat motifs. The occurrence frequency of a short repeat such as $(CA)_{n>10}$ is about 50,000 times in human genome. After PCR using ^{32}P-labeled nucleotide or end-labeled primer, amplification products are separated on a sequencing gel and visualized by autoradiography (Beckman and Soller, 1990). Alternatively, a non-radioactive method can be used for visualization of amplified microsatellite loci by silver staining of polyacrylamide gels (Ghassemi and Gresshoff, 1998). Amplification products differed even in one repeat can be differentiated on a sequencing gel. For example, two genotypes with different number of repeats, $(AT)_{17}$ vs. $(AT)_{18}$, generates amplification products with 2 bases different in length.

SSR loci are quite abundant in a variety of plant species, such as soybean, rice, *Brassica*, mango, avocado, and coca. SSR markers are co-dominant and used for genetic mapping and DNA fingerprinting. SSR markers might be a useful tool in genetic analysis of important crops such as tomato and wheat in which genetic variations are low.

References

Aasheim, H. C., Deggerdal, A., Smeland, E. B., and Hornes, E. 1994. A simple subtraction method for the isolation of cell-specific genes using magnetic monodisperse polymer particles. Biotechniques 16: 716-721.

Aldrich, J. and Cullis, C. A. 1993. RAPD analysis in flax: Optimization of yield and reproducibility using KlenTaq1 DNA polymerase, chelex 100, and gel purification of genomic DNA. Plant Mol. Biol. Rep. 11: 128--141.

Allen, F. L. 1994. Usefulness of plant genome mapping to plant breeding. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, pp. 11-18.

Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D., and Sorrells, M. E. 1993. Optimizing parental selection for genetic linkage maps. Genome 36: 181-186.

Arens, P., Odinot, P., Van Heusden, A. W., Lindhout, P., and Vosman, B. 1995. GATA- and GACA-repeats are not evenly distributed throughout the tomato genome. Genome 38: 84-90.

Arnheim, N. and Erlich, H. 1992. Polymerase chain reaction strategy. Annu. Rev. Biochem. 61: 131-156.

Arnold, M. L., Buckner, C. M., and Robinson, J. J. 1991. Pollen mediated introgression and hybrid speciation in Louisiana irises. Proc. Natl. Acad. Sci. USA 88: 1398-1402.

Bassam, B. J., Caetano-Anollés,G., and Gresshoff, P. M. 1992. DAN amplification fingerprinting of bacteria. Appl. Microbiol. Biotechnol. 38: 70-76.

Beattie, K. L., Beattie, W., Meng, L., Turner, S., Bishop, C., Dao, D. Coral, R., Smith, D., and McIntyre, P. 1995. Clin. Chem. 41: 700-706.

Beckman, J. S. and Soller, M. 1990. Towards an unified approach to genetic mapping of eucaryotes based on sequence tagged microsatellite sites. Bio/Technol. 8: 197-250.

Bell, C. J. and Ecker, J. R. 1994. Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19: 137-144.

Bernard, K., Auphan, N., Granjeaud, S., Victerero, G., Schmitt-Verhulst, A-M, Jordan, B., and Nguyen, C. 1996. Nucleic Acids Res. 24: 1435-1442.

Blanchard, M. M., Taillon-Miller, P., Nowotny, P., and Nowotny, V. 1993. PCR buffer optimization with uniform temperature regimen to facilitate automation. PCR Meth. Appl. 2: 234-240.

Botstein, D., White, R. L., Skolnick, M. H., and Davis, R. W. 1980. Construction of genetic map in man using restriction fragment length polymorphisms. Amer. J. Hum. Genet.

Broun, P. and Tanksley, S. D. 1996. Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol. Gen. Genet. 250: 39-49.

Broun, P. and Tanksley, S. D. 1993. Characterization of tomato DNA clones with sequence similarity to human minisatellite 33.6 and 33.15. Plant Mol. Biol. 23: 231-242.

Caetano-Anollés, G. and Gresshoff, P. M. 1996. Generation of sequence signatures from DNA amplification fingerprints with mini-hairpin and microsatellite primers. Biotechniques 20: 1044-1056.

Caetano-Anollés, G., Callahan, L. M., Williams, P. E., Weaver, K. R., and Gresshoff, P. M. 1995. DNA amplification fingerprinting analysis of bermudagrass (*Cynodon*): genetic relationships between species and interspecific crosses. Theor. Appl. Genet. 91: 228-235.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1994. Buffer components tailor DNA amplification with arbitrary primers. PCR Meth. Appl. 4: 59-61.

Caetano-Anollés, G. and Gresshoff, P. M. 1994. DAN amplification fingerprinting using mini-hairpin oligonucleotide primers Bio/Technology 12: 619-623.

Caetano-Anollés, G. 1993. Amplifying DNA with arbitrary oligonucleotide primers. PCR Meth. Appl. 3: 85-94.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1993. Enhanced detection of polymorphic DNA by multiple arbitrary amplicon profiling of endonuclease-digested DNA: identification of markers tightly linked to the supernodulation locus in soybean. Mol. Gen. Genet. 241: 57-64.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1992. Primer template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. Mol. Gen. Genet. 235: 157-165.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Bio/Tech. 9: 553-557.

Calderwood, S. B., Baker, M. A., Caroll, P. A., Michel, J. L., Arbiet, R. D., and Ausubel, F. M. 1996. Use of cleaved amplified polymorphic sequences to distinguish strains of *Staphylococcus epidemidis*. J. Clin. Microbiol. 34: 2860-2865.

Cifarelli, R. A., Gallitelli, M., and Cellini, F. 1995. Random amplified hybridization microsatellite (RAHM): isolation of a new class of microsatellite-containing DNA clones. Nucleic Acids Res. 23: 3802-3803.

Cregan, P. B., Akkaya, M. S., Bhagwat, A. A., Lavi, U., and Rongwen, J. 1994. Length polymorphisms of simple sequence repeat (SSR) DNA as molecular markers in plants. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, USA, pp. 47-56.

Dellaporta, S. J., Wood, J., and Hicks, J. B. 1983. A plant DNA mini-preparation. Version II., Plant Mol. Biol. Reporter 1: 19-21.

Depeiges, A., Goubely, C., Lenor, A., Cocherel, S., Picard, G., Raynal, M., Grellet, F., and Delseny, M. 1995. Theor. Appl. Genet. 91: 160-168.

DeScenzo, R. A. and Harrington, T. C. 1994. Use of (CAT)-5 as a DNA fingerprinting probe for fungi. Phytopath. 84: 534-540.

Dirlewanger, E., Isaac, P. G., Ranade, S., Belajouza, M., Cousin, R., and De Vienne, D. 1994. Restriction fragment length polymorphism analysis of loci associated with disease resistance genes and developmental traits in *Pisum sativum* L. Theor. Appl. Genet. 88: 17-27.

Ender, A., Schwenk, K., Stadler, T., Streit, B., and Schierwater, B. 1996. RAPD identification of microsatellites in Daphnia. Mol. Ecol. 5: 437-441.

Frohman, M. A., Dush, M. K., and Martin, G. R. 1988. rapid production of full length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Sci. Acad. USA 85: 8998-9002.

Ghassemi, F. and Gresshoff, P. M. The early *enod2* and the leghemoglobin (*lbc3*) genes segregate independently from other known soybean symbiotic genes. Mol. Plant-Microbe Interac. 11: 6-13.

Grattapaglia, D. Bertolucci, F. L. G., Penchel, R., and Sederoff, R. R. 1996. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. Genetics 144: 1205-1214.

Gresshoff, P. M. and Caetano-Anollés, G. 1997. Commercial applications of DNA profiling by amplification with arbitrary oligonucleotide primers. In "Technology Transfer of Plant Biotechnology" (Ed. P. M. Gresshoff). CRC Press, Boca Raton, FL, USA, pp. 131-148.

Gresshoff, P. M. and Landau-Ellis, D. 1994. Molecular mapping of soybean nodulation genes. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, USA, pp. 97-112.

Gyllensten, U. B. and Erlich, H. A. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to the direct sequencing of the HLA-DQA locus. Proc. Natl. Acad. Sci. USA. 85: 7652-7656.

Hardeys, H., Balick, M., and Schierwater, B. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol. Eco. 1: 55-63.

Hirao, I., Kawai, G., Yoshizazwa, S., Nishimura, Y., Ishido, Y., Watanabe, K., and Miura, K. 1994. Most compact hairpin-turn structure exerted by a short DNA fragment, d(GCGAAGC). Nucleic Acids Res. 22: 572-582.

Hu, J. and Quiros, C. F. 1991. identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell Rep. 10: 505-511.

Jeffreys, A. J., Wilson, V., and Thein, S. L. 1985a. Hypervariable 'minisatellite' regions in human DNA. Nature 314: 67-73.

Jeffreys, A. J., Wilson, V., and Thein, S. L. 1985b. Individual-specific 'fingerprints' of human DNA. Nature 316: 76-79.

Jeffreys, A. J. and Morton, D. B. 1987. DNA fingerprints of dogs and cats. Animal Genet. 18: 1-15.

Jeffreys, A. J., Royle, N. J., Wilson, V., and Wong, Z. 1988. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. Nature 332: 278-281.

Jeffreys, A. J., Turner, M., and Debenham, P. 1991. The efficiency of multilocus DNA fingerprint probes for individualization and establishment of family relationships, determined from excessive casework. Am. J. Hum. Genet. 48: 824-840.

Jiang, C. X. and Sink, K. C. 1997. RAPD and SCAR markers linked to the sex expression locus M in asparagus. Euphytica 94: 329-333.

Kaemmer, D. Weising, k., Beyemann, B., Börner, T., Epplen, J. T., and Kahl, G. 1995. Oligonucleotide fingerprinting of tomato DNA. Plant Breeding 114: 12-17.

Kolchinsky A., Landau-Ellis, D., and Gresshoff, P. M. 1997. Map order and linkage distance of molecular markers close to the supernodulation (*nts-1*) locus of soybean. Mol. Gen. Genet. 245: 29-36.

Konieczny, A. and Ausubel, F. M. 1993. A procedure for mapping *Arabidopsis* mutations using recombinant ecotype-specific PCR-based markers. Plant J. 4: 403-410.

Kraft, T., Nybom, H., and Werlemark, G. 1996. DNA fingerprint variation in some blackberry species (*Rubus* subg. *Rubus*, *Rosaceae*). Plant Sys. Evol. 199: 93-108.

Landau-Ellis, D., Angermüller, S. A., Shoemaker, R., and Gresshoff, P. M. 1991. The genetic locus controlling supernodulation in soybean (*Glycine max* L.) co-segregates tightly with a cloned molecular marker. Mol. Gen. Genet. 228: 221-226.

Levi, A., Rowland, I. J., and Hartung, J. S. 1993. Production of reliable randomly amplified polymorphic DNA (RAPD) markers from DNA of woody plants. HortScience 28: 1188-1190.

Liang, P and Pardee, A. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971.

Lisitsyn, N., Lisitsyn, N., and Wigler, M. 1993. Cloning the differences between two complex genomes. Science 259: 946-951.

Maser, R. L. and Calvet, J. P. 1995. Analysis of differential gene expression in the kidney by differential cDNA screening, subtractive cloning, and mRNA differential display. Semin. Nephrol. 15: 29-42.

Men, A. E. and Gresshoff, P. M. 1998. Efficient cloning of DAF polymorphic markers from silver-stained polyacrylamide gels. BioTechniques 24: 593-595.

Meng, A., Carter, R. E., and Parkin, D. T. 1990. The variability of DNA fingerprints in three species of swan. Hered. 64: 73-80.

Meyer, W., Mitchel, T. G., Freedman, E. Z., and Vilgalys, R. 1993. Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. J. Clin. Microbiol. 31: 2274-2280.

Meyer, W., Koch, A., Niemann, C., Beyermann, B., Epplen, J. T., and Börner, T. 1991. Differentiation Of species and strains among filamentous fungi by DNA fingerprinting. Curr. Genet. 19: 239-242.

Michelmore, R. W., Paran, I., and Kesseli, R. V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating population. Proc. Natl. Acad. Sci. USA 88: 9828-9832.

Morjane, H., Geistlinger, J., Harrabi, M., Weising, K., and Kahl, G. 1994. Oligonucleotide fingerprinting detects genetic diversity among *Ascochyta rabiei* isolates from a single chickpea field in Tunisia. Curr. Genet. 26: 191-197.

Nikiforov, T. T. and Rogers, Y-H. 1995. The use of 96-well polystyrene plates for DNA hybridization-based assays: an evaluation of different approaches to oligonucleotide immobilization. Anal. Biochem. 227: 201-209.

Nikiforov, T. T., Rendle, R. B., Goelet, P., Rogers, Y-H., Kotewicz, Ml. L., Anderson, S., Trainor, G. L., Knapp, M. R. 1994. Genetic bit analysis: a solid phase method for typing single nucleotide polymorphisms. Nucl. Acids Res. 22: 4167-4175.

Pääbo, S., Irwin, D. M., and Wilson, A. C. 1990. DNA damage promotes jumping between templates during enzymatic amplification. J. Biol. Chem. 265: 4718-4721.

Paterson, A. H. Molecular dissection of quantitative traits: progress and prospects. Genome Res. 5: 321-333.

Perucho, M., Welsh, J., Peinado, M. A., Ionov, Y., and McClelland, M. 1995. Fingerprinting Of DNA and RNA by arbitrary primed polymerase chain reaction: applications in cancer research. Meth. Enzymol. 254: 275-290.

Piquot, Y., Saumitou-Laprade, P., Petit, D., Vernet, P., and Epplen, J. T. 1996. Genotypic diversity revealed by allozymes and oligonucleotide DNA fingerprinting in French populations of the aquatic macrophyte *Sparganium erectum*. Mol. Ecol. 5: 251-258.

Powell, W., Machray, G. c., Provan, J. 1996. Polymorphisms revealed by simple sequence repeats. Trends in Plant Sci. 1: 215-222.

Ralph, D. McClelland, M., Welsh, J. 1993. RNA fingerprinting using arbitrary primed PCR identifies differentially regulated RNAs in mink lung (Mv1Lu) cells growth arrested by transforming growth factor beta 1. Proc. Natl. Acad. Sci. USA 90: 10710-10714.

Richardson, T., Cato, S., Ramser, J., Kahl, G., and Weising, K. 1995. Hybridization of microsatellites to RAPD: a new source of polymorphic markers. Nucl. Acids Res. 23: 3798-3799.

Rogstad, S. H. 1994. Inheritance in turnip of variable-number tandem-repeat genetic markers revealed with synthetic repetitive DNA probes. Theor. Appl. Genet. 89: 824-830.

Romao, J. and Hamer, J. E. 1992. Genetic organization of a repeated DNA sequence family in the rice blast fungus. Proc. Natl. Acad. Sci. USA 89: 5316-5320.

Rosewich, U. L. and McDonald, B. A. 1994. DNA fingerprinting in fungi. Meth. Mol. Cell Biol. 5: 41-48.

Sastry, J. G., Ramakrishna, W., Sivaramakrishnan, S., Thakur, R. P., Gupta, V. S., and Rajekar, P. K. 1995. DNA fingerprinting detects genetic variability in the pearl

millet downy mildew pathogen (*Sclerospora graminicola*). Theor. Appl. Genet. 91: 856-891.

Sharma, P. C., Winter, P., Bünger, T., Hüttel, B., Weigand, F., Weising, K., and Kahl, G. 1995. Abundance and polymorphism of di-, and tetra-nucleotide tandem repeats in chickpea (*Cicer arietinum* L.). Theor. Appl. Genet. 90: 90-96.

Shoemaker, R. C., Lorenzen, L. L., Diers, B. W., and Olson, T. C. Genome mapping and agriculture. 1994. In: Plant Genome Analysis- A CRC series of current topics in plant molecular biology, (ed. P. M. Gresshoff). CRC Press, Inc., Boca Raton, FL, USA, pp. 1-10.

Shoemaker, R. C. and Olson, T. C. 1993. Molecular linkage map of soybean (*Glycine max* L. Merr). In "Genetic Maps: Locus maps of complex genomes", (ed. O'Brien, S. J.). Cold Spring Harbor Laboratory Press, New York, pp. 6.131-6.138.

Skroch, P. and Nienhuis, J. 1995. Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. Theor. Appl. Genet. 91: 1086-1091.

Suzuki, Y., Wanaka, A., Tohyama, M., and Takagi, T. 1995. Identification of differentially expressed mRNAs during neuronal differentiation of P19 embryonal carcinoma cells. Neurosci. Res. 23: 65-71.

Tourmente, S., Deragon, J. M., Lafleuriel, J., Tutois, S., Cuvillier, C., Espagnol, M. C., and Picard, G. 1994. Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human minisatellite core sequence. Nucleic Acids Res. 22: 3317-3321.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zebeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. 23: 4407-4414.

Wada-Kiyama, Y., Peters, B., and Naguchi, C. T. 1992. The epsilon-globin gene silencer. Characterization by in vitro transcription. J. Biol. Chem. 267: 11532-11538.

Wallner, E., Weising, K., Rompf, R., Kahl, G., and Kopp, B. 1996. Oligonucleotide fingerprinting and RAPD analysis of *Achillea* species: characterization and long-term monitoring of micropropagated clones. Plant Cell Reports 15: 647-652.

Watt, W. B. 1994. Allozymes in evolutionary genetics: self-imposed burden or extraordinary tool? Genetics 136: 11-16.

Weaver, K., and Callahan, L. M. 1994. Isolation and cloning of DNA amplification products from silver-stained polyacrylamide gels. BioTechniques 16: 226-227.

Weising, K., Atkinson, R. G., and Gardner, R. C. 1995. PCR Methods Applic. 4: 249-255.

Weising, K. and Kahl, G. 1998. Hybridization-based microsatellite fingerprinting of plants and fungi. In: "DNA Markers: Protocols, Applications and Overviews" (eds. G. Caetano-Anollés and P. M. Gresshoff). Wiley-Liss, Inc., New York, pp. 27-54.

Weising, K., Atkinson, R. G., Gardner, R. C. 1995. Genomic fingerprinting by microsatellite-primed PCR: a critical evaluation. PCR Methods Appl. 4: 249-255.

Weising, K., Nybom, H., Wolff, K., and Meyer, W. 1995. DNA Fingerprinting in Plants and Fungi. CRC Press, Boca Raton, USA.

Wells, R. A., Green, P., and Reeders, S. T. 1989. Simultaneous genetic mapping of multiple human minisatellite sequences using DNA fingerprinting. Genomics 5: 761-772.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218.

Welsh, J., Chada, K., Dalal, S. S., Cheng, R., Ralph, D., and McClelland, M. 1992. Nucleic Acids Res. 20: 4965-4970.

Welsh, J., Rampino, N., McClelland, M., and Perucho, M. 1995. Nucleic acid fingerprinting by PCR-based methods: applications to problems in aging and mutagenesis. Mutat. Res. 338: 215-229.

Winberg, B. C., Zhou, Z., Dallas, J. F., McIntyre, C. L., and Gustafson, J. P. 1993. Characterization of minisatellite sequences from *Oryza sativa*. Genome 36: 978-983.

Williams, M. N. V., Pande, N., Nair, M., Mohan, M., and Benneet, J. 1991. Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa* L.) genomic DNA. Theor. Appl. Genet. 82: 489-498.

Wu, K., Jones, R., Danneberger, L., and Scolnik, P. A. 1994. Detection of microsatellite polymorphisms without cloning. Nucleic Acids Res. 22: 3257-3258.

Zabeau, M. and Vos, P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application No. 92402629.7. Publication No. EP 0 534 858 A1.

Zietkiewicz, E., Rafalski, A., Labuda, D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction. Genomics 20: 176-183.

Part 3

The Early *enod2* and the Leghemoglobin (lbc3) Genes Segregate Independently from Other Known Soybean Symbiotic Genes

Abstract

Recombinant inbred lines (RILs) as well as an F_2 segregating population of soybean *Glycine max* (L. Merr) facilitated the mapping of two expressed sequence tags (EST) involved in early nodulation and subsequent nitrogen fixation in soybean. For the early nodulin gene *enod2*, the parents of RILs, Minsoy and Noir1, showed a polymorphism (5.5 vs. 5.9 kb) after *Eco*RV digestion. RFLP patterns of 42 RILs were analyzed using the MAPMAKER program linking *enod2* to the seed coat color gene, *I*, with a distance of 11.1 cM on linkage group U3 of RIL map. *Enod2* and *I* are located close to *Rhg*4 , a soybean cyst nematode (SCN) resistance gene, and a locus for seed coat hardness. The molecular marker pA110 and seed coat color were used to integrate *enod2* on an F_2 segregating population (72 plants) generated from a cross between cultivar Bragg and *G. soja* (Sieb and Zucc), PI468.397. *Enod2* was mapped in the same order as on the RIL map but 18.5 cM from the *I* locus on the TN map. A microsatellite from the 5' region of *enod2B* was mapped in the same position, demonstrating that *enod2B* and not *enod2A* was mapped. An RFLP for *lbc3* (leghemoglobin) segregated independently from *enod2* and the *nts*-1 supernodulating locus suggesting that in soybean, symbiotically significant loci (including rj_1, Rj_2, and rj_6) are not clustered in soybean.

Introduction

Genetic properties of both bacteria (e. g., *Rhizobium*, *Bradyrhizobium*) and host plants are essential for the development of nitrogen-fixing root nodules. Plant control of nodulation was confirmed by the isolation of symbiotically altered mutants (see Gresshoff, 1993; Caetano-Anollés and Gresshoff, 1991) and by the demonstration of nodule-specific or nodule-enhanced proteins, namely nodulins, such as leghemoglobin, uricase or glutamine synthase (Verma and Delauney, 1988). Such nodulins are classified into early and late nodulins based on their temporal sequence of induction. Early nodulin genes are expressed during infection and nodule morphogenesis (Gloudemans and Bisseling, 1989) and may shed light on events related to cell division, hormonal responses, and plant defense mechanisms as well as the signal transduction chain connecting nod-factor recognition and nodule initiation. Late nodulin gene expression is correlated with the onset of nitrogen fixation (Govers *et al.*, 1985, 1987) and reveals information about carbon/nitrogen metabolism and transport, symbiosome membrane

transport, functions related to the maintenance of the intracellular symbiosis, and oxygen regulation and transport.

ENOD2, one of the first discovered early nodulins, was initially characterized in soybean (Franssen *et al.*, 1987). The gene product is a hydroxy-proline rich cell wall protein with molecular weight of 75,000 daltons, and has the alternative label of GmN-75 nodulin. It was postulated to be involved in the oxygen barrier inside of legume nodules (Pawlowski and Bisseling, 1996, Long, 1996), although there is no evidence for ENOD2 function other than the localization of the protein. Indeed no difference was observed in the level of the *enod2* protein and mRNA of alfalfa nodules under different concentrations of oxygen (Hunt *et al.*, 1995). Two soybean genomic clones, *enod2*A and *enod2*B, of 16 kb and 25 kb were isolated from a soybean genomic library in lambda charon 35 (Franssen *et al.*, 1990). These two clones contained no introns and were identical in their coding domains but differed in the 5' regions. The deduced amino acid sequences revealed two pentapeptide proline-rich repeats (PRO-PRO-GLU-TYR-GLN and PRO-PRO-HIS-GLU-LYS) which were similar to pentapeptide repeats of a soybean proline-rich protein designated SbPRP1 (Hong *et al.*, 1987). The strong conservation of coding sequence suggests recent gene duplication, genome tetraploidization, or functional restriction.

Using a soybean *enod2* cDNA as a probe, homologous *enod2* sequences were isolated from pea (Govers *et al.*, 1987), white clover, bird's foot trefoil, vetch (Moreman *et al.*, 1987), alfalfa (Dickstein *et al.*, 1988), common bean (Sanchez *et al.*, 1988) and *Sesbania rostrata* (Strittmatter *et al.*, 1989). The major difference among mRNAs from different plants was predominantly in the length of the gene.

It is possible that ENOD2 represents a pericycle-specific protein whose expression is increased during early nodulation. The protein has similarity to other cell wall proteins, some of which expressed during lateral root induction or plant defense responses (Lamb *et al.*, 1989). Expression of the *enod2* is initiated in cells at the base of the nodule and ultimately in the nodule parenchyma, uninfected cells that surround the infection zone (Allen *et at.*, 1991; Kouchi *et al.*, 1989, van de Wiel *et al.*, 1990). *Enod2* gene expression occurred in empty nodules on alfalfa roots induced by exopolysaccharide-deficient (*Exo*$^-$) mutant of *Rhizobium meliloti* which lacked infection threads and intracellular bacteroids (Dickstein *et al.*, 1988). Application of auxin transport inhibitors

such as NPA and TIBA, resulted in empty nodule-like structures which expressed *enod2* (Hirsch *et al.*, 1989). Alfalfa variants which formed nodules in the absence of *Rhizobium* (NAR) also expressed *enod2* (Truchet *et al.*, 1989). The protein therefore is expressed independently of bacterial infection and indeed may be reflective of pericycle cell proliferation. *Sesbania rostrata enod2* (*SrEnod2*) was expressed after exogenous application of cytokinins to roots and stems (Dehio and de Bruijn, 1992).

Molecular mapping of nodulation-related genes may permit the determination of function, if complete co-segregation with mutant loci is observed. With an increasing database containing both QTL and single gene loci associated with agronomic parameters, it is likely that nodulin genes with secondary function are revealed (Mansur *et al.*, 1996). Furthermore, knowledge obtained from mapping may allow the discovery of related mechanisms of a multigene family (possibly with different, but related function) and an elucidation of evolutionary relatedness in syntenic linkage groups.

Most mapping is done in segregating F_2 populations, but these have the disadvantage of being finite (see Landau-Ellis *et al.*, 1991; Keim *et al.*, 1990; Weeden *et al.*, 1990). While ample supplies of DNA can be collected, eventually they run out. Recombinant inbred lines (RILs), through single seed descent, fix homozygosity in genomic regions, resulting in (more or less) stable seed lines, to be shared indefinitely by the scientific community. Such RILs derived from F_7 seed are available for soybean (Lark *et al.*, 1993; Mansur *et al*, 1993a/b and 1996) and *Lotus japonicus* (Jiang and Gresshoff, 1997). RILs also have the advantage that living tissue can be harvested, permitting physical mapping using high molecular weight DNA derived from protoplasts (Funke *et al.*, 1993) or the isolation of telomeres, satellites, or related chromosomal regions (Kolchinsky and Gresshoff, 1995).

Of course, most parents used to make RILs do not differ in all desired phenotypic properties. Hence a strategy of map-integration is used, in which molecular markers associated with a mutant locus (i. e., *nts-1*, Landau-Ellis *et al.*, 1991) were mapped in a specific cross close to an RFLP marker (pA381), which then in turn was found on the RIL map (Filatov and Gresshoff, 1997, submitted). Thereby, the mutant locus was transferred by association through map integration. We tested this approach here, attempting to confirm conserved marker order and mapping distances for the region around the *enod2* gene. Furthermore, we wanted to evaluate the hypothesis developed

from mapping in pea that symbiotically important genes are clustered (as on chromosome 1 of pea, Weeden *et al.*, 1990).

Materials and Methods

Parent plants of the Utah soybean (*Glycine max* L. Merr) RIL population, Minsoy and Noir1 cultivars, and RIL plants (at F_{12}) were grown in an environmentally-controlled greenhouse. Unifoliolates and young trifoliolates [0.5 g (FW)] were harvested and immediately placed in liquid nitrogen. Average DNA yield was 100 µg/g FW measured by DNA fluorometer (Model TKO 100, Hoefer Scientific Instruments, San Francisco, USA). After DNA isolation (Dellaporta *et. al.*, 1983), samples from parents were digested with seven restriction enzymes, *Bcl*I, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, and *Taq*I using buffers recommended by the manufacturer. Restricted genomic DNA (5 µg/lane) was electrophoresed on 0.9% agarose gels and then blotted onto Zeta Probe Nylon membranes by vacuum blotting (Sambrook *et al.*, 1989).

Enod2 cDNA cloned in pUC18 (confirmed by partial DNA sequencing), provided by Ton Bisseling (Wageningen, Netherlands), was isolated (Sambrook *et al.*, 1989) and restricted with *Pst*I. Following the isolation of the *enod2* insert, Southern blot hybridization was carried out with ^{32}P-labeled (Boehringer Mannheim Random Primer Labeling Kit) *enod2* probe and parental DNA blots. After detection of polymorphisms in the parents, Southern hybridization was carried out for 42 RILs.

Parents of the F_2 population for this map integration were *G. max* cultivar Bragg and the ancestral soybean *G. soja*, PI468.397. These were previously used to map the supernodulation gene, *nts-1*, (Landau-Ellis *et al.*, 1991). DNA samples from these parents were digested with six restriction nucleases, namely, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Taq*I, and *Xba*I. Parental crosses were made and verified as described by Landau-Ellis *et al.* (1991) using morphological and RFLP markers. After detection of polymorphisms in parents, Southern hybridization was carried out for 88 F_2 plants generated from a cross between Bragg and PI468.397. Inheritance of RFLP banding patterns of *enod2* in the RILs and the F_2 population was determined with two independent Chi Square tests.

The MAPMAKER program (Lander *et . al.*, 1987) was used to map *enod2* and *lbc3* in the RIL and the F_2 populations. In the case of the F_2 population, cosegregation of *enod2*, pA110 molecular marker, and seed coat color (*I*) of 72 plants was studied.

The *lbc3* gene was cloned in pAR10 as a 3.5-kb *Bam*HI fragment carrying a 2-kb 5' region of the *lbc3* gene (provided by Dr. Jens Stougaard, Aarhus, Denmark). After detection of RFLP with *Taq*I restriction of Minsoy and Noir1, 226 RILs were screened for RFLP patterns.

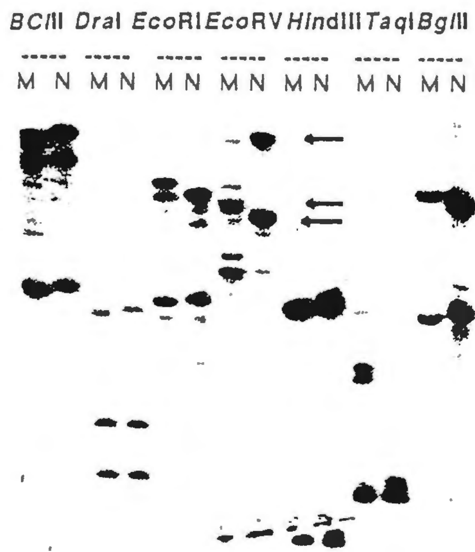
Cosegregation of RFLP marker and microsatellite marker in parental materials, RIL, and F_2 populations was studied to determine which *enod2* copy we mapped. The microsatellite represents 17 repeats of AT from nucleotide 866 through nucleotide 900 in the 5' upstream region of the genomic *enod2B* clone of cultivar Wayne. PCR conditions for amplification of the microsatellite marker were: 6 min/ 95°C and 35 times of 30 sec/ 95°C, 1 min/ 47°C, and 2 min/ 72°C in an Ericomp Twin Block thermocycler (San Diego, CA, USA). The PCR reaction mix contained 50 ng DNA template, 1 mM Mg^{2+}, 0.15 µM of each 3' and 5' primers, 100 µM of each dNTP, 0.1 µL of 3,000 Ci/mmol a-^{32}P dATP, 1x PCR buffer (50 mM KCl, 10 mM Tris-HCl pH 8.3, 1% Triton X-100), and 1 unit Taq DNA polymerase (Perkin Elmer, Roche Molecular System, Inc., New Jersey, USA). Each sample (3.5 µL) was added to 3.5 µL stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF) before loading in a 6% acrylamide sequencing gel containing 8M urea. The polymorphism was detected by autoradiography as well as silver staining. All photos of autoradiographs were taken by Alpha Innotech IS-1000 digital imaging system, San Leandro, CA, USA.

Results

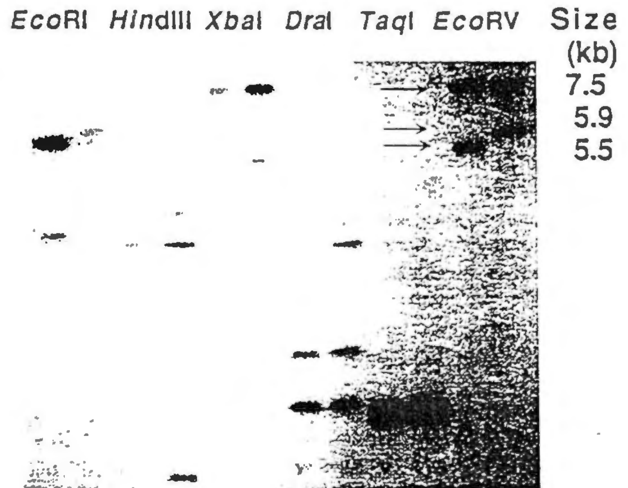
Restriction fragment length polymorphisms (RFLPs) derived from Southern hybridizations probed with a 600-bp *Pst*I fragment of the *enod2* clone (pENOD2) with digested genomic DNA from Minsoy (M), Noir1 (N) (Figure 3-1A), and their RILs (Figure 3-2), were detected. In most cases, the two or more copies of *enod2* were consistent with the ancient tetraploidy nature of soybean (Polzin *et al.*, 1994; Singh and Hymowitz, 1988) and the fact that two related but distinct genomic clones were available. However, a few faint bands were detectable stemming presumably from

Figure 3-1. Autoradiograph of Southern hybridization for *enod2*, pA110, and *lbc3* probes. A, Minsoy (M) and Noir1 (N) probed with ENOD2 cDNA clone (pENOD2). Two copies of *enod2* loci were detected in most cases. Restriction fragment length polymorphism (RFLP) from *Eco*RV restriction (5.9 vs. 5.5 kb) indicated by arrows, was used for screening recombinant inbred lines (RILs). B, Bragg (B) and *G. soja* (S) probed with pENOD2. RFLP from *Eco*RV restriction (5.9 vs. 5.5 kb) and monomorphic band (7.5 kb), indicated by arrows, was used for screening F_2 individuals. C, Bragg (B) and *G. soja* probed with flanking marker pA110. RFLP from *Xba*I restriction (12.0 vs. 6.0 kb) and monomorphic band (8.5) was used for screening F_2 population. D, autoradiograph of *Taq*I restriction (3.0 vs. 2.6 kb) was used for screening RILs.

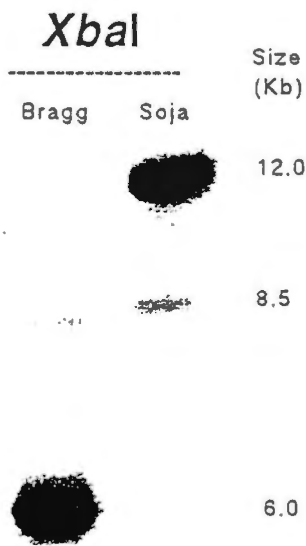
A



B



C



D

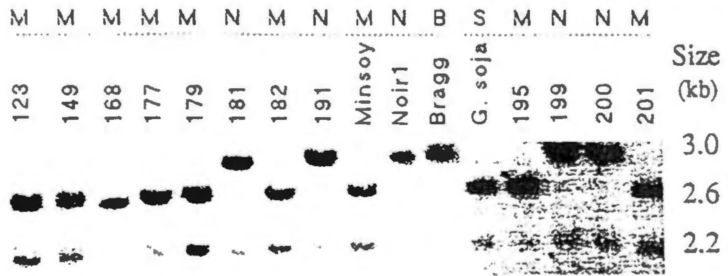
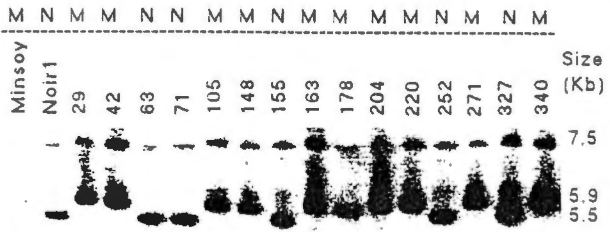
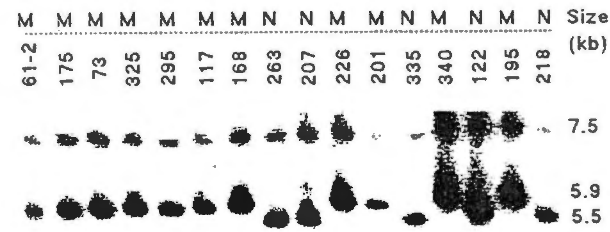
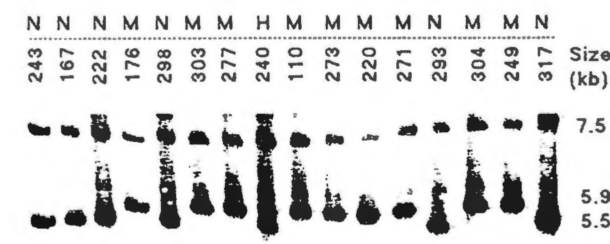


Figure 3-2. Segregation of restriction fragment length polymorphism (RFLP) for *enod2*. A, *Eco*RV restricted DNA of Minsoy, Noir1, and recombinant inbred lines (RILs) was probed with pENOD2 clone. Top row labels the RFLP patterns (M or N type); numbers are the same as original ones obtained from the University of Utah. B, *Eco*RV restriction of the parents of two populations (RIL and F_2) and some of the RILs and F_2 individuals. Those F_2 individuals with either 5.9- or 5.5-kb bands were scored as B and S type patterns, respectively. Heterozygotes possessed all three bands and were scored as H.

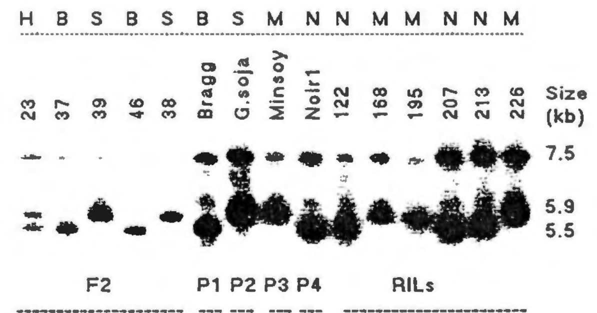
A







B



similarity of *enod2* to other cell-wall-protein genes. Four RFLPs were detected from restriction with *Eco*RI, *Eco*RV, *Taq*I, and *Bgl*II. In the case of *Eco*RV, the sizes of polymorphic bands were 5.9 and 5.5 kb for Minsoy and Noir1, respectively. A monomorphic band was observed at 7.5 kb. Identical RFLP patterns for *enod2* (5.5 vs. 5.9 kb) and *lbc3* (3.0 vs. In most cases, the two or more copies of *enod2* were consistent with the ancient tetraploidy nature of soybean (Polzin *et al.*, 1994; Singh and Hymowitz, 1988) and the fact that two related but distinct genomic clones were available. However, a few faint bands were detectable stemming presumably from similarity of *enod2* to other cell-wall-protein genes. Four RFLPs were detected from restriction with *Eco*RI, *Eco*RV, *Taq*I, and *Bgl*II. In the case of *Eco*RV, the sizes of polymorphic bands were 5.9 and 5.5 kb for Minsoy and Noir1, respectively. A monomorphic band was observed at 7.5 kb. Identical RFLP patterns for *enod2* (5.5 vs. 5.9 kb) and *lbc3* (3.0 vs. 2.6 kb) in Minsoy and *G. soja* may indicate common origin in this region; likewise, Noir1 and Bragg also share common patterns.

Although several RFLPs were detected with other enzymes, none of them (but *Eco*RV) were clearly codominant which is valuable to distinguish heterozygous segregants. Inheritance of the *enod2* RFLP was studied using a *Chi* Square test to examine whether the observed RFLP banding patterns in RILs were consistent with an expected 1:1 Mendelian ratio (Table 3-1). No significant differences between expected and observed values indicated that this RFLP was inherited as a single locus.

We used the MAPMAKER program to analyze the data for *enod2* segregation. *enod2* was mapped to linkage group U3 (Mansur *et. al.*, 1996) (Figure 3-3A). Flanking are two molecular markers of K401b and BLT24 with 7.4 and 7.3 cM distance, respectively. In addition, *enod2* is linked to the "*I*" gene encoding seed coat color at 11.1 cM distance (LOD score of 6.34). The "*I*" gene was postulated to encode chalcone synthase genes (*CHS*) or modulators of their activity (Wang *et. al.*, 1994). This central gene of isoflavone/flavone biosynthesis (Estabrook and Sengupta-Gopalan, 1991) is important for plant pigmentation, plant disease response (phytoalexins), and nod-gene inducer biosynthesis (Kosslak *et al.*, 1987; Sutherland *et al.*, 1990). Further analysis revealed strong interactions between the *enod2* region and a region on linkage group U16 close to the molecular marker pA401 to explain quantitative variation for agronomically important traits such as yield and pod filling period (Dr. Gordon Lark,

Table 3-1. Five independent *Chi* Square tests[a].

Character	RFLP patterns	Observed values	Expected values	Mendelian ratios	Calculated chi squares	Table chi squares
enod2 RIL population	M	25	20.5	1		
	N	16	20.5	1	1.560	3.840 $\alpha = 5\%$ NS
enod2 F_2 population	B	26	22	1		
	H	47	44	2		
	S	15	22	1	3.217	5.991 $\alpha = 5\%$ NS
F_2 seed coat color	B (yellow)	16	22.75	1		
	V (variegated)	57	45.5	2		
	S (black)	18	22.75	1	5.902	5.991 $\alpha = 5\%$ NS
F_2 pA-110	B	15	18.25	1		
	H	41	36.5	2		
	S	17	18.25	1	1.220	5.991 $\alpha = 5\%$ NS
lbc3 RIL population	M	121	113	1		
	N	105	113	1	0.994	3.840 $\alpha = 5\%$ NS

[a] Tests were for the inheritance of restriction fragment length polymorphism (RFLP) banding patterns of *enod2* gene in the recombinant inbred line (RIL) and the F_2 populations, for seed coat color and pA-110 RFLP in the F_2 population, and the *lbc3* gene in the RIL population of soybean. M, Minsoy; N, Noir1; B, Bragg; H, heterozygote; S, *G. soja;* and V, variegated color (gray-brown). NS, non-significant differences.

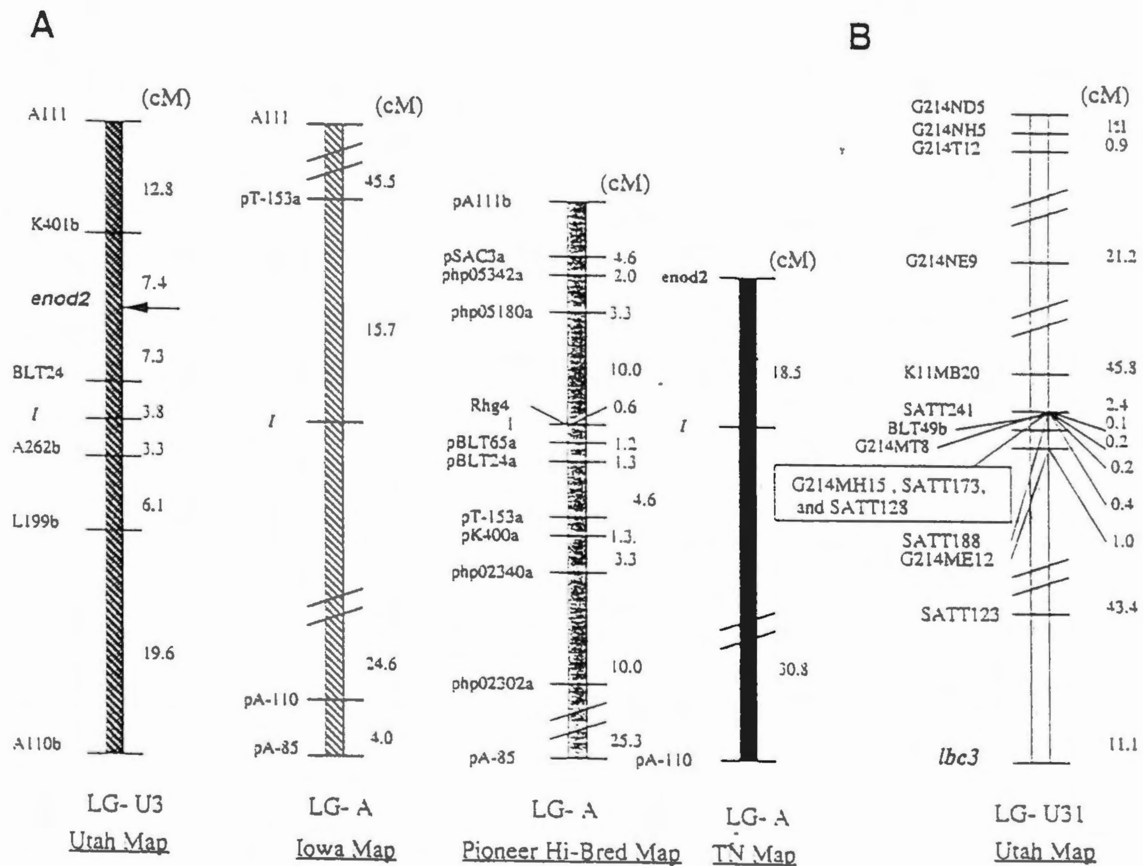


Fig. 3. Linkage maps of *enod2* and *lbc3* genes. A, integration mapping of the soybean *enod2* gene. Utah map is generated from a recombinant inbred line population. Iowa, Pioneer Hi-Bred, and TN (The University of Tennessee, Knoxville) maps all are generated from F_2 segregating populations. A region on linkage group (LG) U16 containing five molecular markers has strong interaction with the *enod2* region. Segments are aligned through a common positioning of the "*I*" locus on each LG. There is inconsistency in order of BLT24 and pT-153a markers in Pioneer Hi-Bred map compared with Utah and Iowa maps, caused by either possible inversion or the MAPMAKER program. B, Linkage of the leghemoglobin (*lbc3*) gene on LG U31 of the Utah map. Single sequence repeat and restriction fragment length polymorphism markers are indicated.

Figure 3-3. Linkage maps of *enod2* and *lbc3* genes.

personal communication). The *enod2* gene was located in a region close to the *Rhg4* gene, a soybean cyst nematode (SCN) resistance gene (Weisemann *et al.*, 1992; Yazdi-Samadi *et al.*, 1996, Webb *et al.*, 1995; Concibido *et al.*, 1994; Matson and Williams, 1965) and a QTL for seed coat hardness (Keim *et al.*, 1990).

The flanking RFLP marker pA110 and seed coat color were used to integrate *enod2* onto an F_2 segregating population map (Shoemaker and Olson, 1993). The results of Southern blots probed with the *enod2* clone for parental materials, Bragg and *G. soja*, PI468.397 (Figure 3-1C) , and F_2 individuals showed an inheritance pattern consistent with expected 1:2:1 Mendelian ratio (Table 3-1). Since F_2 seed coat color phenotypes were distinguishable only after seed formation, F_3 seed coat color was scored as yellow (Bragg pattern), gray-brown (heterozygote pattern), and black (*soja* pattern) with a ratio of 16: 57: 18, respectively. F_2 segregants probed with the pA110 clone revealed a polymorphism of 6 vs. 12 kb (Bragg *vs.* PI468.397) after *Xba*I restriction and segregation of 15:41:17 (Bragg, heterozygote, and soja pattern, respectively). Significant differences for expected *vs.* observed ratio were found neither seed coat color nor for pA110 (Table 3-1). Cosegregation study of *enod2*, pA110, and seed coat color of 72 F_2 individuals resulted in integration of *enod2* on the Iowa and TN *G. max* x *G. soja* partial map in the same order as on the RIL map, but 18.5 cM from the *I* locus (Figure 3-3A). The numerical differences are within the expected range of error caused by differences parents and sample sizes.

To check the hypothesis of clustering of nodulation genes, one soybean leghemoglobin gene, *lbc3*, was hybridized to *Taq*I restricted genomic blot of Minsoy, Noir1, and 226 RILs (Figure 3-1D). An RFLP for *lbc3* was mapped at the end of linkage group U31 (Fig. 3-3B) and segregated independently from the *enod2* locus.

The *enod2* RFLP marker and the microsatellite marker cosegregated in the RIL population, and the F_2 population (Figure 3-4) indicating that the *enod2* RFLP which we mapped was the *enod2B* copy. The size of single sequenced repeat (SSR) polymorphic bands were 184 *vs.* 189 bp in Minsoy vs. Noir1 and 165 *vs.* 264 bp in Bragg *vs. G. soja*, respectively. We note with interest the 99-bp size difference in Bragg and *G. soja*, suggesting perhaps that *enod2B* is transcriptionally inactive and that the monomorphic fragment (7.5kb) contained the functional *enod2A* copy.

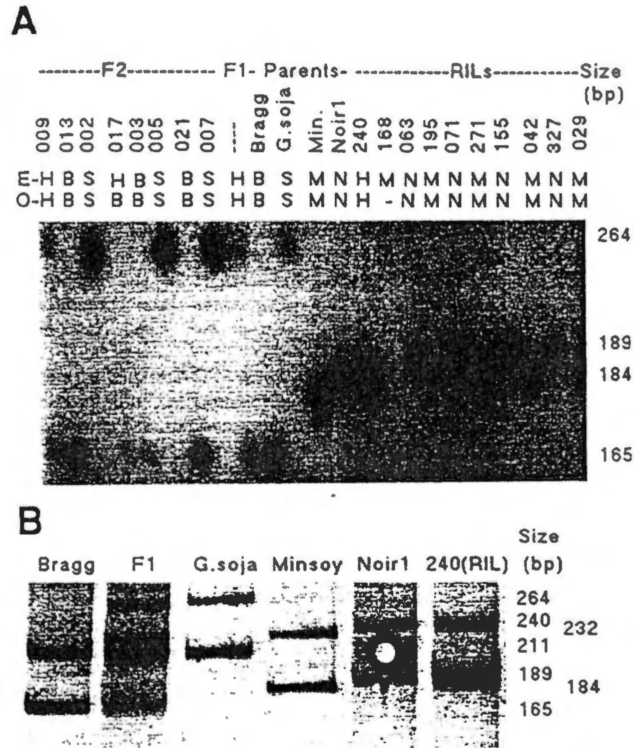


Fig. 4. Mapping of the *enod2* by an AT microsatellite. A, Autoradiogram of polyacrylamide gel for the *enod2b* microsatellite in F_2, F_1, parents, and recombinant inbred lines (RILs). B, Bragg; S, *Glycine soja*; M, Minsoy; N, Noirl; H, heterozygote; E, expected; and O, observed banding patterns. No amplification product was observed for RIL # 168. B, A silver-stained 10% polyacrylamide gel run in a Mini-proteanII gel electrophoresis apparatus (BioRad, Hercules, CA). Additional band at 232 bp may be a heteroduplex or mismatched product.

Figure 3-4. Mapping of the *enod2* by an AT microsatellite.

Discussion

We mapped the *enod2B* gene on the linkage group U3 of Utah map in a region of soybean genome that seems to contain a number of genes encoding monogenic or polygenic traits (QTL) which may be associated with cell wall proteins, e.g., seed hardness (Keim *et al.*, 1990), pod filling period, and soybean cyst nematode resistance genes (Weisemann *et al.*, 1992; Webb *et al.*, 1995). Furthermore, *enod2* was linked to the *I* gene which encodes seed coat color and affects chalcone synthase (CHS) gene expression (Wang *et al.*, 1994).

Seed hardness in soybean is a quantitative trait locus (QTL) which is encoded by three genes affecting germination rate, viability, and quality of stored seed (Kilen *et al.*, 1978; Rolston, 1978; Potts *et al.*, 1978). This trait has been genetically analyzed with 72 RFLP markers (Keim *et al.*, 1990). Five independent RFLPs were associated with seed hardness trait. However, a major portion of variation (32%) was explained by the genomic region containing the *I* locus.

In our study, the *enod2* gene was mapped near the *I* locus with a distance of 11.5 cM on the RIL map (Utah) and 18.5 cM on the *G. max* x *G. soja* map (TN). It is possible that this genomic region is specialized for genes controlling cell wall components. The ENOD2 protein is similar to another class of structural cell wall proteins in soybean, namely hydroxyproline-rich glycoproteins (HRGPs). An example of this class of protein is the extensins which are localized in palisade, epidermal and hourglass cells of soybean seed coat (Cassab and Varner, 1987). The *enod2* gene is expressed in cortical cells of nodule which has been postulated to contain the variable oxygen barrier. Shape and size of these cells may be factors in regulating O_2 permeability (Layzell *et al.*, 1993). Thus, the *enod2* gene product, being a hydroxyproline-rich cell wall protein, might be a specific structural protein which plays a role in generating small intercellular spaces in cortical tissue. However, no differences were observed in levels of the ENOD2 protein and ENOD2 mRNA under different concentrations of oxygen subjected to alfalfa nodules, weakening the hypothesis of ENOD2 involvement in oxygen barrier (Hunt *et. al.*, 1995). Exogenously applied cytokinins induced the *Sesbania rostrata enod2* gene (*SrEnod2*) expression in root and stem nodules (Dehio and de Bruijn, 1992). The *SrEnod2* was also induced in tumors generated by wild-type *Agrobacterium tumefaciens*, but not by *A. tumefaciens* mutant of cytokinin biosynthesis.

The *I* locus was linked to two RFLP markers BLT24 and BLT65 with recombination rates of 4.4% and 4.0%, respectively. Gene order was determined as BLT24- *I*- BLT65 (Weisemann *et al.*, 1992). The two molecular markers must also be linked to *Rhg4*, soybean cyst nematode (SCN) race 3 resistance gene, since this gene had been tightly linked to the *I* locus with 0.35% recombination rate (Matson and Williams 1965; Webb *et al.*, 1995). In another study, two independent RFLP markers, pA32 and pA85, were found to be associated with SCN race3 resistance (Concibido *et al.*, 1994). These two markers contributed 51.7% of the total variation in race 3 SCN disease response, with pA85 accounting for 21.4%. It was mapped with a distance of 10.9 cM from the *I* locus on linkage group A of the USDA/Iowa soybean map (Figure 3-3).

The *I* gene seems to be involved in the distribution of anthocyanin pigments (reviewed in Palmer and Kilen, 1978). When a dominant *I* allele is present, production of anthocyanin is inhibited in the epidermal layer of the developing seed coat resulting in yellow seed color. In yellow-coated seed (*I*) activity of chalcone synthase (CHS) was 7- to 10-fold less than in the pigmented (i/i) seed coats. Chalcone synthase mRNA was barely detectable in seed coats carrying the *I* allele. Association of *CHS* and *I* locus was further confirmed with detection of multiple restriction site polymorphisms in genomic DNA blots of the *CHS* gene family in near-isogenic lines of *I* locus (Wang *et al.*, 1994).

Although the order of the *enod2* map is the same in all the maps, the differences in distance might be due to different parents, environmental effects during recombination, or experimental error. Sequence analysis of the *enod2B* genomic clone of cultivar Wayne revealed a microsatellite (AT) of 17 repeats between nucleotide 866 and nucleotide 900 (about 2 kb upstream from the ATG start site for *enod2*). This single sequence repeat (SSR) was detected by PCR primers and detected a polymorphism between Minsoy and Noir1. Mansur *et al.* (1996) mapped this *enod2* associated SSR to 9.3 cM from *I*, confirming our RFLP based data. Of special interest is the larger size difference of the *enod2B* SSR found between Bragg and *G. soja*, 165 and 264 bp, respectively. We do not know whether this region is involved in gene control.

The second copy of the *enod2* gene (represented by the 7.5 kb monomorphic *Eco*RV fragment) seems to be highly conserved in *G. max* cultivars (Minsoy, Noir1, and Bragg) and wild-type *G. soja*. We postulate that it maps in a homeologous region of

soybean. It would be of value to test whether both loci are actively involved in expression of ENOD2.

In pea, *sym* cluster 1 on chromosome 1 consists of glutamine synthetase, leghemoglobin genes, *sym-2* (strain specificity), *sym-5* (low nodule number), *sym-18* (strain specificity and late flowering), and *sym-19* (non-nodulating mutant). In soybean, *Rj2*, controlling ineffective nodulation, was integrated from the classical genetic map onto the USDA/Iowa map (Polzin *et al.*, 1994) in a cluster of disease resistance QTLs (Kanazin *et al.*, 1996). However, an RFLP for one of the leghemoglobin genes (*lbc3*) segregated independently from the *enod2B* locus, and the *nts-1* supernodulation locus mapped on linkage group H of the USDA/Iowa map; Landau-Ellis *et al.*, 1991, (Table 3-2). Two loci for P^{34} protein kinase (*cdc2*) were mapped onto separate linkage groups (U20 and U25), as was the region of the supernodulation locus *nts-1* (U23- close to pA381) (Filatov and Gresshoff, 1997, submitted). Moreover, *nod139* (*rj6*) and *nod49* (*rj1*) (non-nodulating mutants) as well as *nod49* and *nts-1* (supernodulating mutant) segregated independently (Table 3-2) suggesting that in soybean, symbiotically significant loci are not clustered.

Table 3-2. Linkage and genetic data for soybean symbiosis and symbiosis-related genes.

Gene Symbol	Function	Linkage Group
cdc2-M	P34 protein kinase	U20
cdc2-N	P34 protein kinase	U25
pA381- pPV (*nts-1* region)	Supernodulation mutant	U23
enod2	cell wall protein	U3
lbc3	O2 carrier protein	U31
Rj2	ineffective nodulation	Iowa Linkage Group J
rj1	non-nodulation	Chromosome 3
rj6	non-nodulation	Segregates independently from *rj1* and *nts-1*

References

Allen, T., Raja, S., and Dunn, K. 1991. Cells expressing ENOD2 show differential spatial organization during the development of alfalfa root nodules. Mol. Plant-Microbe Interact. 4: 139-146.

Caetano-Anollés, G. and Gresshoff P. M. 1991. Plant genetic control of nodulation. Annual Review of Microbiology 45: 345-382.

Cassab, G. I. and Varner, J. E. 1987. Immunocytolocalization of extensin in developing soybean seed coats by immunogold silver staining and by tissue printing on nitrocellulose paper. J. Cell. Biol. 105: 2581- 2588.

Concibido, V. C., Denny, R. L., Boutin, S. R., Hautea, R., Orf, J. H., and Young, N. D. 1994. DNA marker analysis of loci underlying resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe). Crop Sci. 34: 240- 246.

Dehio, C. and de Bruijn, F. J. 1992. The early nodulin gene *SrEnod2* from *Sesbania rostrata* is inducible by cytokinin. The Plant Journal 2: 117-128.

Dellaporta, S. J., Wood, J., and Hicks, J. B. 1983. A plant DNA mini-preparation. Version II., Plant Mol. Biol. Reporter 1: 19-21.

Dickstein, R., Bisseling, T., Reinhold, V. N., and Ausubel, F. M. 1987. Expression of nodule-specific genes in alfalfa root nodules blocked at an early stage of development. Genes and Development 2: 677-687.

Diers, B. W., Mansur, L., Imsande, J., and Shoemaker, R. C. 1992. Mapping *Phytophthora* resistance loci in soybean with restriction fragment length polymorphism markers. Crop Sci. 32: 377-383.

Estabrook, E. M. and Sengupta-Gopalan, C. 1991. Differential expression of phenylalanine ammonia-lyase and chalcone synthase during soybean nodule development. Plant Cell 3: 299-308.

Franssen, H. J., Nap, J. P., Gloudemans, T., Stiekema, W., van Dam, H., Govers, F., Louwerse, J., van Kammen, A., and Bisseling, T. 1987 Characterization of cDNA for nodulin-75 of soybean: a gene product involved in early stages of root nodule development. Proc. Natl. Acad. Sci. USA 84: 4495-4499.

Franssen, H. J., Thompson, D. V., Idler, K., Kormelink, R., van Kammen, A., and Bisseling, T. 1990. Nucleotide sequence of two soybean *ENOD2* early nodulin genes encoding Ngm-75. Plant Mol. Biol. 14: 103-106.

Fuchsman, W. H. and Appleby, C. A. 1979. Separation and determination of the relative concentrations of the homogenous components of soybean leghemoglobin by isoelecteric focusing. Biochem. Biophys. Acta 579: 314-324.

Funke, R., Kolchinsky, A., and Gresshoff, P. M. 1993. Physical mapping of a region in the soybean (*Glycine max*) genome containing duplicated sequences. Plant Mol. Biol. 22: 437-446.

Gloudemans, T. and Bisseling, T. 1989. Plant gene expression in early stages *Rhizobium*-legume symbiosis. Plant Sci. 65: 1-14.

Govers, F., Nap, J. P., Moerman, M., Franssen, H. J., Van Kammen, A., and Bisseling, T. 1987. cDNA cloning and developmental expression of pea nodulin genes. Plant Mol. Biol. 8: 425-436.

Govers, F., Moreman, M., Downie, J. A., Hooykaas, P., Franssen, H. J., and Louwerse, J. 1986. *Rhizobium nod* genes are involved in inducing an early nodulin gene. Nature 323: 564- 566.

Govers, F., Gloudemans, T., Moreman, M., van Kammen, A., and Bisseling, T. 1985. Expression of plant genes during the developmental of pea root nodules. EMBO Journal 4: 861-867.

Gresshoff, P. M. 1993. Molecular genetic analysis of nodulation genes in soybean. Plant Breeding Rev. 11:275-318.

Hirsch, A. M., Bhuvaneswari, T. V., Torrey, J. G., and Bisseling, T. 1989. Early nodulin genes are induced in alfalfa root out-growths elicited by auxin transport inhibitors. Proc. Natl. Acad. Sci. USA 86: 1244-1248.

Hong, J. C., Nagao, R. T. and Key, J. L. 1987. Characterization and sequence analysis of a developmentally regulated putative cell wall protein gene isolated from soybean. J. Biol. Chem. 262: 8367- 8376.

Hunt, S., Wycoff, K. L., Layzell, D. B., and Hirsch, A. M. 1995. Is the early nodulin ENOD2 associated with regulation of the nodule permeability to oxygen? In "Nitrogen Fixation: Fundamentals and Applications". Xth International Congress on Nitrogen Fixation St. Petersburg, edited by Tikhonovich, I. A., Provorov, N. A., Romanov, V. I., and Newton, W. E. Kluwer Academic Publishers. p. 504.

Jiang, Q. and Gresshoff, P. M. 1997. Classical and molecular genetics of the model legume *Lotus japonicus*. Molecular Plant-Microbe Interac. 10: 59-68.

Kanazin, V., Marek, L. F., and Shoemaker, R. C. 1996. Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA 93: 11746-11750.

Keim, P., Diers, B. W., and Shoemaker, R. C. 1990. Genetic analysis of soybean hard seededness with molecular markers. Theor. Appl. Genet., 79: 465- 469.

Kilen, T. C. and Hartwig, E. E. 1978. An inheritance study of impermeable seed in soybeans. Field Crops Res. 1: 65- 70.

Kolchinsky, A. and Gresshoff, P. M. 1995. A major satellite DNA of soybean is a 92-base pairs tandem repeat. Theor. Appl. Genet. 90: 621-626.

Kosslak, R. M., Bookland, R., Barkei, J., Paaren, H. E., and Appelbaum, E. R. 1987. Induction of *Bradyrhizobium japonicum* common *nod* genes by isoflavones isolated from *Glycine max*. Proc. Natl. Acad. Sci. USA 84: 7428-7432.

Kouchi, H., Tsukamoto, M., and Tajima, S. 1989. Differential expression of nodule-specific (nodulin) genes in the infected, uninfected and cortical cells if soybean (*Glycine max*) root nodules. J. Plant Physiol. 135: 608-617.

Lamb, C. J., Lawton, M. A., Dron, M., and Dixon, R. A. 1989. Signals and transduction mechanisms for activation of plant defenses against microbial attack. Cell 56: 215-224.

Landau-Ellis, D., Angermüller, S. A., Shoemaker, R., and Gresshoff, P. M. 1991. The genetic locus controlling supernodulation in soybean (*Glycine max* L.) co-segregates tightly with a cloned molecular marker. Mol. Gen. Genet. 228: 221-226.

Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S. E., and Newburg, L. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174-181.

Lark, K. G., Weisemann, J. M., Mathews, B. F., Palmer, R., Chase, K., Macalma, T. 1993. A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir1'. Theor. Appl. Genet. 86: 901-906.

Layzell, D. B., Diaz Del Castillo, L., Hunt, S., Kuzma, M., van Cauwenberghe, O., and Oresnik, I. 1993. The regulation of oxygen and its role in regulating nodule metabolism. In "New Horizons in Nitrogen Fixation". Palacios, R., Mora, J., and Newton, W. E. Kluwer Academic Publishers pp. 393-398.

Long, S. R. 1996. *Rhizobium* symbiosis: nod factors in perspective. Plant Cell 8: 1885-1898.

Mansur, L. M., Orf, J. H., Chase, K., Jarvik, T., Cregan, P. B., and Lark, K. G. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean [*Glycine max* (L.) Merr.]. Crop Sci. 36: 1327-1336.

Mansur, L. M., Orf, J., and Lark, K. G. 1993. Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbreds of soybean (*Glycine max* L. Merr.). Theor. Appl. Genet. 86: 914-918.

Mansur, L. M., Lark, K. G., Kross, H., and Oliveira, A. 1993. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor Appl Genet 86: 907-913.

Matson, A. L. and Williams, L. F. 1965. Evidence of a fourth gene for resistance to the soybean cyst nematode. Crop Sci. 5: 477.

Moreman, M., Nap, J. P., Govers, F., Schilperoort, R., van Kammen, A., and Bisseling, T. 1987. *Rhizobium nod* genes are involved in the induction of two early nodulin genes in *Vicia sativa* root nodules. Plant Mol. Biol. 9: 171-179.

Palmer, R. G. and Kilen, T. C. 1978. Quantitative genetics and cytogenetics. In J. R. Wilcox, ed., Soybean: Improvement, Production, and Uses, Ed 2. American Society of Agronomy, Madison, WI, pp 135-209.

Pawlowski, K. and Bisseling, T. 1996. Rhizobial and actinorhizobial symbioses: what are the shared features? Plant Cell 8: 1899-1913.

Potts, H. C., Duangpatra, J., Hairston, W. G., and Delouche, J. C. 1978. Some influences of hard seededness on soybean quality. Crop Sci. 18: 221- 224.

Polzin, K. M., Lohnes, C. D., and Shoemaker, R. C. 1994. Integration of *Rps2*, *Rmd*, and *Rj2* into linkage group J of the soybean molecular map. Journal of Heredity 85: 300-303.

Rolston, M. P. 1978. Water impermeable seed dormancy. Bot. Rev. 44: 365- 396.

Sambrook, J. Fritsch, E. F., and Maniatis. 1989. Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Sanchez, F., Quinto, C., Vazquez, H., Spaink, H., Wijffelman, C. A., Cevallos, M. A. 1988. The symbiotic association of *Phaseolus vulgaris* and *Rhizobium legominosarum bv. phaseoli*. In: "Molecular Genetics of Plant- Microbe Interactions" (eds. R. Palacios and D. P. S. Verma). APS Press, St. Paul, Minnesota, pp. 370-375.

Shoemaker, R. C. and Olson, T. C. 1993. Molecular linkage map of soybean (*Glycine max* L. Merr). In: "Genetic Maps: Locus maps of complex genomes" (ed. O'Brien, S. J.). Cold Spring Harbor Laboratory Press, New York, pp. 6.131-6.138.

Singh, R. J. and Hymowitz, T. 1988. The genomic relationship between *Glycine max* (L.) Merr. and *G. soja* Sieb and Zucc. as revealed by pachytene chromosome analysis. Theor. Appl. Genet. 76: 705-711.

Stritmatter, G., Chia, T. F., Trinh, T. H., Katagiri, F., Kuhlemeier, C. and Chua, N. H. 1989. Characterization of nodule-specific cDNA clones from *Sesbania rostrata* and expression of the corresponding genes during the initial stages of stem nodules and root nodule formation. Mol. Plant-Microbe Interac. 2: 122- 127.

Sutherland, T., Bassam, B. J., Schuller, L. J., and Gresshoff, P. M. 1990. Early nodulation signals of wild type and symbiotic mutants of soybean (*Glycine max*) . Mol. Plant-Microbe Interact. 3: 122-128.

Truchet, G., Barker, D. G., Camut, S., de Billy, F., Vasse, J., and Huguet, T. 1989. Alfalfa nodulation in the absence of *Rhizobium*. Mol. Gen. Genet. 219: 65-68.

van de Wiel, C. Scheres, B. Franssen, H., van Lierop, M. J., van Lammeren, A., van Kammen, A., and Bisseling, T. 1990. The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. EMBO J. 9: 1-7.

Verma, D. P. S. and Delauney, A. J. 1988. Root nodule symbiosis: nodulins and nodulin genes. In "Temporal and Spatial Regulation of Plant Genes" (eds. Verma, D. P. S. and Goldberg, R. B.). New York : Springer-Verlag, pp. 69-199.

Wang, C. S., Todd, J. J., and Vodkin, L. O. 1994. Chalcone synthase mRNA and activity are reduced in yellow soybean seed coats with dominant *I* alleles. Plant Physiol. 105: 739- 748.

Webb, D. M., Baltazar, B. M., Rao-Arelli, A. P., Schupp, J., Clayton, K., Keim, P., and Beavis, W. D. 1995. Genetic mapping of soybean cyst nematode race-3 resistance loci in the soybean PI437.654. Theor. Appl. Genet. 91:574-581.

Weeden, N. F., Kneen, B. E., and LaRue, T. A. 1990. In: "Nitrogen Fixation: Achievements and Objectives" (eds. P. M. Gresshoff, L. E. Roth, G. Stacey, and W. E. Newton). Chapman and Hall, New York and London, pp 323-330.

Weisemann, J. M., Mathews, B. F., and Devine, T. E. 1992. Molecular markers located proximal to the soybean cyst nematode resistance gene, *Rhg*4. Theor. Appl. Genet. 85: 136-138.

Yazdi-Samadi, B., Qian, D., Ghassemi, F., Allen, F. L., and Gresshoff, P. M. 1996. Genetics and molecular biology of resistance to soybean cyst nematode. Physiol. Mol. Biol. Plants 2:25-48.

Part 4

High Annealing Temperature Using Short Arbitrary Primers Provides Robust and High Resolution DNA Amplification Fingerprinting

Abstract

To overcome potential problems caused by mismatch priming and secondary DNA structure, and taking advantage of high primer-template ratios used in DNA amplification fingerprinting (DAF) reactions, annealing temperature of 55°C were used with single short arbitrary oligonucleotide as well as mini-hairpin primers to provide high resolution DNA profiles of soybean (*Glycine max* L. Merr.). Initially, high annealing temperatures for three arbitrary octamer primers in polymerase chain reaction (PCR) were tested for DNA fingerprinting of two soybean cultivars, Minsoy and Noir1. Fifteen PCR programs differing in annealing temperature (47, 55, and 60°C), denaturation, annealing, and extension time (30, 60, and 120 second), and presence/absence of an extension step (+/- 72°C) were tested. The number of scorable bands (amplification products) after 10% PAGE and DNA silver staining ranged from 7 to 51. The average ramping temperature for heating and cooling were calculated 1.42 and 1.27 sec/°C, respectively. Intensity of the silver-stained bands in a 10% polyacrylamide gel was high for the most PCR programs. Program 15 (95°C/30 sec, 55°C/120, and 72°C/30 sec) generated a complex DNA fingerprinting profiles for tested primers in Minsoy and Noir1. These profiles contained an average of 42 sharp and highly intense bands using both octamer primers 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3') for DNA amplification. Using high annealing temperature increased stringency of primer-template annealing, avoided potential mismatching and hybrid molecule formation, and consequently improved reproducibility of DNA fingerprinting.

Introduction

Since the development of polymerase chain reaction, PCR, (Saiki *et al.*, 1985; Mullis *et al.*, 1986; Mullis and Faloona, 1987) DNA amplification has been widely used for variety of purposes such as, direct cloning from genomic or cDNA (Rashtchian *et al.*, 1992a), *in vitro* mutagenesis of DNA (Rashtchian *et al.*, 1992b), prenatal diagnosis of genetic diseases, analysis of RNA transcript structure (Frohman *et al.*, 1988), differential display of RNA (Liang and Pardee, 1992), direct sequencing of genomic DNA (Gyllensten and Erlich, 1988) or cDNA (Ohara *et al.*, 1989), molecular maker mapping, and DNA fingerprinting using arbitrary primers (Williams *et al.*, 1990; Welsh and McClelland, 1990; Caetano-Anollés *et al.*, 1991, Vos *et al.*, 1995).

DNA fingerprinting has been used for forensic cases, cultivar identification (Hu and Quiros, 1991), development of genetic linkage map (Reiter *et al*, 1992; Jiang and Gresshoff, 1997), gene tagging (Paran *et al.*, 1991; Martin *et al.*, 1991), pedigree analysis (Welsh *et al.*, 1991; Caetano-Anollés *et al*, 1995), analysis of genetic diversity and relationships (Stiles *et al.*, 1993; Wilde *et al.*, 1992; Prabhu *et al.*, 1997), and evolution and ecological genetic studies (Arnold *et al.*, 1991; Hadrys *et al.*, 1992).

DNA fingerprinting using arbitrary primers depends on several factors including DNA concentration, PCR buffer components (Caetano-Anollés *et al*, 1994; Blanchard *et al.*, 1993), length and GC content of the primer (Williams *et al.*, 1990), magnesium concentration (Caetano-Anollés *et al*, 1993), type of bacterial thermostable DNA polymerases (Bassam *et al.*, 1992; Schierwater and Ender, 1993; Aldrich and Cullis, 1993), and PCR program.

DNA Factor

DNA concentration should be in the range of certain critical concentration of genomic DNA to generate reproducible results. Below the required range, DNA profiles consist of incomplete amplification products and are not reproducible (Williams *et al.*, 1993). Using excessive DNA concentration usually produces a low quality DNA fingerprinting profile with poor resolution and smears. It is best to perform a serial dilution of each genomic DNA to identify the useful rang of DNA concentrations. The range of DNA concentration recommended for DAF and RAPD is 0.01-2 and 0.2-1 ng/µL, respectively (Bassam and Bently, 1994). The increased range of DAF is presumed to stem from the higher primer concentration. Indeed the primer/template ratio of DAF seems to permit a widened "window" of activity such as use of shorter primers (even 5 mers occasionally work) and higher annealing temperatures (this study), cf. Williams *et al.*, 1990, who stated above 45°C, RAPD products can not be obtained.

Thermocycler Factor

DAF and other PCR-based DNA fingerprinting techniques are sensitive to thermal conditions (Ramping) of the thermocycler. DNA amplification profiles sometimes vary between different thermocyclers and units of the same model even set to identical cycling parameters. In some machines the temperature is different in the cells located in the

interior part than those in the vicinity of the thermal block; such instruments should be avoided. This problem can be overcome by measuring the actual ramping temperature, the time required for heating or cooling in a given range of temperature, in the reference machine using a data logger or strip chart recorder. Then, this profile can be used to set the other thermocycler to such cycling parameters.

Primer Factor

Another important factor in DNA fingerprinting is quality and concentration of the primer. Commercial primers usually provide consistent results but storage for a long period of the time results in decrease in the quality of the primer. This is especially after frequent thaw and freeze cycles. It is recommended to thaw the primer completely before use and mix it with the reaction mixture very well to provide homogenous PCR condition. Primer purity may be assessed by running 20% PAGE and silver staining (Caetano-Anollés *et al.*, 1994)

Magnesium Factor

Magnesium is a co-factor for enzymatic activity of all DNA polymerases, such as AmpliTaq and Stoffel (PE Inc.). These two types of enzymes have different requirements for magnesium (Mg^{+2}) concentration. Low concentration of Mg^{+2} generates low number of amplification products while excess of Mg^{+2} ions produces non-specific amplification products. Hence, the Mg^{+2} concentration should be optimal in the amplification reaction. Magnesium concentration is also optimal for DNA templates of different complexity. Thus, 1 ng of *E. coli* DNA behaves differently than 1 ng of soybean DNA amplified with the same primer and Mg^{+2} concentration.

Other Factors

Other factors such as choice for DNA polymerase and pipetting accuracy affect DNA fingerprinting profiles in different techniques. It is well known that the activity of thermostable DNA polymerases varies considerably depending on their origin (Bassam *et al.*, 1992; Schierwater and Ender, 1993). Two types of DNA polymerases, *Taq* polymerase and Stoffel fragment, are used for DNA amplification fingerprinting. Stoffel fragment is a truncated derivative of native form which lacks the 5'-3' exonuclease

activity. This form shows better thermostability and more magnesium tolerance than the native one (Bassam *et al.*, 1992). In addition, the Stoffel fragment generates more reproducible DNA fingerprinting profiles compare to *Taq* polymerase. However, the size distribution of amplification products tends to be lower molecular weight when Stoffel fragment is used. For more satisfactory results, it is better to compare the two enzymes side-by-side.

PCR-based DNA fingerprinting techniques are very sensitive to pipetting error and changes in PCR conditions and components leading to irreproducible banding patterns (Levi *et al.*, 1993; Caetano-Anollés *et al.*, 1992). Pipetting of a small amount of volumes, especially under 3 µL, can be a source of errors resulting in inconsistent amplification products. It is recommended to dilute various stock solutions to a concentration that would allow pipetting of 5 µL or more. In addition, one problem might be inaccuracy of the pipette itself which can be solved by regular service and adjustments.

Annealing Temperature

Polymerase chain reaction is a complex and dynamic process of kinetic and thermodynamic reactions which are changed by relative concentrations and activities of all components during each cycle. In an arbitrary-primed DNA amplification, the success of reaction depends on the reliable and simultaneous primer annealing at many locations of the DNA template. Using short primers between 8 and 12 nucleotides accomplishes the success of reaction. However, use of a short arbitrary primer has a disadvantage of imperfect events of annealing, called mismatching, between primer and template. It is well-known that primers in PCR can initiate amplification after annealing to template sequences which are not perfectly complementary. The mismatches are most common at the 5' end although they can occur at any nucleotide position in the primer-template complex (Somer and Tautz, 1989; Kwok *et al.*, 1990; Huang *et al.*, 1992; Caetano-Anollés *et al.*, 1992).

It is shown that as the cycle number in a DAF reaction increased, secondary DAF of primary products is possible through the "opening" of otherwise conciled primer sites by increased template leads. Hybrid molecules, especially stemming from repeated

eukaryotic DNA regions may produce novel templates possibly not found in parental profiles.

It is possible to minimize mismatching by using stringent PCR conditions favoring perfect annealing. This can be achieved by coupling high annealing temperature and low primer, enzyme, and buffer ion concentrations. Nevertheless, rare mismatching events still occur under high stringent conditions, only less frequently and less predictably resulting in amplification products which are present in one experiment but are absent in other identical reactions. These rare mismatches and consequent amplification are due to low efficiency interaction of primer and template in early thermocycling reaction and thereafter amplification will continue and complete with high efficiency.

One of factors which determines stringency of the PCR condition is annealing temperature of primer and DNA template which differs in DAF (30°C) and RAPD (35-45°C) techniques. In general, RAPD thermocycling programs use standard PCR conditions but with lower annealing temperature. A typical RAPD program is as follows: an initial high denaturing temperature at 94°C for 5 min followed by 35 cycles of 94°C for 1 min, 45°C annealing step for 1 min, and 72°C extension step for 2-3 min. At the end of cycling, a single final extension step of 72°C for 5-10 min is usually used.

In DAF, the extension step is omitted and a typical program is as follows: an initial denaturing temperature of 95°C for 5 min followed by 35 cycles of 96°C melting step for 1 sec and 30°C annealing step for 1 sec. At the end of cycling, a single final extension step of 72°C for 5 min is usually used.

We report here a high annealing temperature (55°C) DAF for soybean DNA samples to obtain robust amplification products and improved banding pattern resolution.

Materials and Methods

Plant materials consisted of two soybean, *Glycine max* L. (Merr), cultivars Minsoy and Noir1 which were grown in greenhouse under controlled conditions. Genomic DNA from young leaves was isolated according to Dellaporta *et al.* (1983). Concentration of DNA was measured in a TKO100 Fluoremeter (Hoeffer Scientific Instruments, San Francisco, CA).

Two template soybean DNA samples, Minsoy and Noir1, and three arbitrary octamer primers, 8-4 (5'GTAACGCC3'), 8-8 (5'GAAACGCC3'), and 8-9 (5'GTTACGCC3'), were tested. Several of the programs were compared in parallel to each other for the 8-4 primer. Original DAF (Caetano-Anollés *et al.*, 1991) program (P#3) in an Ericomp thermocycler was used as a control.

DNA quantification

After DNA isolation, the working DNA concentration was prepared as follows:

1. Measure the DNA stock dissolved in water using a Hoefer Fluorimeter Model TK100.
2. Dilute the DNA to a level of 50 ng/µL and, then, measure it the same way as in step 1.
3. Repeat the step 2 but to a level of 25 ng/µL
4. Dilute the material from step 3 twenty-five fold to a working concentration of the DNA template of 1 ng/µL.

DNA Amplification

Polymerase chain reaction (PCR) was carried out in MJ Research and Ericomp thermocyclers (San Diego, CA). Reaction was run in a 20 µL total volume consisting 0.15 ng/µL template DNA, 3 µM primer, 0.2 units/µL of AmpliTaq Stoffel fragment DNA polymerase (PE, Norwalk, CT), 200 µM of each deoxynucleotide triphosphate, 1.25 mM $MgCl_2$ and Stoffel buffer (10 mM KCl, 10 mM Tris-HCl; pH 8.3) when 8-mer primers were used. In case of mini-hairpin primers we used 4 mM $MgSO_4$ and TTNK10 buffer (10 mM KCl, 4 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 20 mM Tris-HCl; pH 8.3). Amplification was carried out in 500 µL plastic tubes, suited for the Ericomp wells. Amplification volume was overlaid with mineral oil (Mallinckrodt, U.S.P.).

PCR Programs

Fifteen PCR programs used in this study differed in levels of annealing temperature (30, 47, 55, and 60°C), denaturation, annealing, and extension time (30, 60, 120 seconds), and presence/absence of extension step (+/- 72°C) to increase stringency of DAF. P#3 and N-DAF, the original DAF programs with annealing temperature of 30°C for 1 sec and melting temperature of 95°C for 1 sec in Twin-Block (Ericomp, San Diego, CA) and MJR thermocyclers, respectively, were used as controls. Four programs, 7, J, 8, and 3, had 30°C annealing temperature but differed in their ramping times of annealing and melting temperature change, and the presence/absence of a 72°C extension step. Six programs designed with 47°C annealing temperature considered the effect of different ramping times for annealing, melting, and the absence/presence of the 72°C extension step. In addition, four programs differing in melting time and absence/presence of a 72°C extension step were used with 55°C annealing temperature for 120 sec. Annealing temperature of 60°C for 120 sec was used as the extreme temperature limit for DNA amplification.

Total amplification time, number of detectable bands, and intensity of the bands were recorded for each program. Cooling and heating (ramping) for each range of temperature in the experiments were also recorded.

Gel Separation

Denaturing polyacrylamide gels were prepared in MiniProtean II gel rigs from BioRad Inc. To do so, the following recipe of 10% acrylamide gel was used.

Acrylamide	39.2 g
PDA	0.8 g
Urea	40 g
TBE (10x)	40 mL
Glycerol	20 mL
distilled water	as needed
Total:	400 mL

Gels were loaded with a two-fold dilution from the amplification mix. At times amplification mixes needed to be as is, at other times dilution was needed. We suggest that extensive staining in upper gel regions can be reduced by dilution. Also make sure at all times that all chemicals are properly dissolved. Urea crystals, for example, become nearly invisible before they are properly in solution. Undissolved urea crystals will ruin a gel.

The following steps were used to run a gel:

1. 150 μL of fresh 10% ammonium persulfate and 15 μL of TEMED were added to 10 mL gel solution (4°C) while stirring.
2. Gel solution was loaded in gel rigs using 10-mL syringe and syringe filter (Millipore type), and then, a 13-well comb was inserted in each gel.
3. After >30 minutes, the gel rigs were prerun in 1x TBE buffer in cold room at 300 V (PowerPac 300, BioRad Inc.) for 15 minutes.
4. Before loading the samples, wells were cleaned thoroughly twice using a syringe. [this step is important to get nice bands].
5. 3 μL of sample (amplification products) was added to 3 μL of loading buffer in a microtiter plate and then, each two wells were cleaned again before loading the samples. In addition, 3 μL of twenty-fold dilution of 90 ng/μL molecular weight marker (Bio Ventures, Inc.) was loaded.
6. The gels were run at room temperature at 300 V (PowerPac 300, BioRad Inc.) for 35 minutes.
7. Then, the gels were fixed in 7.5 % acetic acid for 10 minutes while shaking.
8. The gels were washed three times with deionized water while shaking.
9. Silver staining was carried out by adding 300 μL of 37 % formaldehyde to 200 mL of 0.1% silver nitrate solution just before pouring on the gels. Staining took 20 minutes while shaking.
10. Meanwhile, developer, 3 % sodium carbonate solution (200 mL), was prepared and chilled to 8-10°C while stirring.
11. The silver solution was discarded in a waste bottle and then the gels were rinsed in deionized water for 5 seconds.
12. 150 μL of 0.4% sodium thiosulfate and 600 μL of 37% formaldehyde were added to chilled developer while stirring.

13. The developer was poured while the gels were shaking vigorously for first 30 seconds and then they were shaken gently until majority of the bands had appeared.
14. The developing process was stopped with cold 7.5 % acetic acid for 3 minutes after pouring off the developer. The gels were washed for 10 minutes while shaking.
15 To avoid either possible bending or cracking of the gels, we suggest that keep the gels in anti-cracking solution (10 % acetic acid, 35 % ethanol, and 1 % glycerol) for >60 minutes.
16. The gels were hung off the metal bench top by magnets to dry out for several hours.
17. Dried gels are kept in photo albums for future reference. We have kept gels for 6 years without fading. Failure to rinse gels properly may produce urea crystals in the form of beautiful webs across the gel. While artistically attractive, such problems should be avoided. Dried gels can be used to isolate indicative DNA bands as described by Weaver *et al* (1994). Dried gels also can be used for presentation directly on an overhead projector, usually under a glass plate to prevent warping. Recently, an improved technique has been used for cloning of a band excised from wet gels of DAF, AFLP, and differential display RT-PCR (Men and Gresshoff, 1998)

Results and Discussion

Total time for each of the fifteen PCR programs and number of the bands on the polyacrylamide gels were recorded (Table 4-1). Intensity and sharpness of the bands for DNA fingerprinting profile were evaluated (Table 4-1; Figures 4-1 and -2). Effect of time course of each step, addition of an extension step, ramping, and annealing temperature are discussed in the following sections.

Total time for programs ranged from 85 (original N-DAF program) to 313 (Program 2) minutes. It appeared that total time had an effect on the low intensity of the bands in programs N, 7, 8, and 6, indicating that time was not sufficient to complete amplification products. Although intensity of the bands was low in Program 10 as well, this was probably due to the too high annealing temperature, 60°C, which did not allowed the template and primer to interact. When extension steps of 72°C for 30 sec and 72°C for 60 sec were added to the Programs 7 and 8, respectively, designated as Programs J and 3, the intensity of the bands was increased showing that the extension step was useful to obtain further amplification products in programs with 30°C annealing temperature (Table 4-1). In contrast, Programs 4, 5, 11, 12, and 9, lacking the

Table 4-1. Time course (seconds) of cycles for new DAF programs in the MJR Thermocycler. P# 3, the original DAF program, was run in a twin-block thermocycler (Ericomp) as a control.

Programs	Denaturation Temperature	---------- Annealing	Temperature	----------		Extension Temperature	Total Time	Result Summary	
	95°C	30°C	47°C	55°C	60°C	72°C	Min	No. of Total Bands	Intensity of Bands
P# 3	1	1	-	-	-	-	160	46	high
N-DAF	1	1	-	-	-	-	085	32	low
FG-DAF7	30	30	-	-	-	-	123	42	low
FG-DAF(J)	30	30	-	-	-	30	159	52	high
FG-DAF8	60	60	-	-	-	-	160	38	low
FG-DAF3	60	60	-	-	-	60	214	40	low
FG-DAF6	30	-	30	-	-	-	114	37	low
FG-DAF4	60	-	60	-	-	-	145	46	high
FG-DAF5	120	-	120	-	-	-	221	41	high
FG-DAF2	120	-	120	-	-	120	313	42	high
FG-DAF11	30	-	120	-	-	-	166	41	high
FG-DAF14	30	-	120	-	-	30	205	50	high
FG-DAF12	30	-	-	120	-	-	160	43	high
FG-DAF15	30	-	-	120	-	30	200	42	high
FG-DAF9	120	-	-	120	-	-	213	48	high
FG-DAF13	30	-	-	120	-	120	256	48	high
FG-DAF10	120	-	-	-	120	-	211	07	low

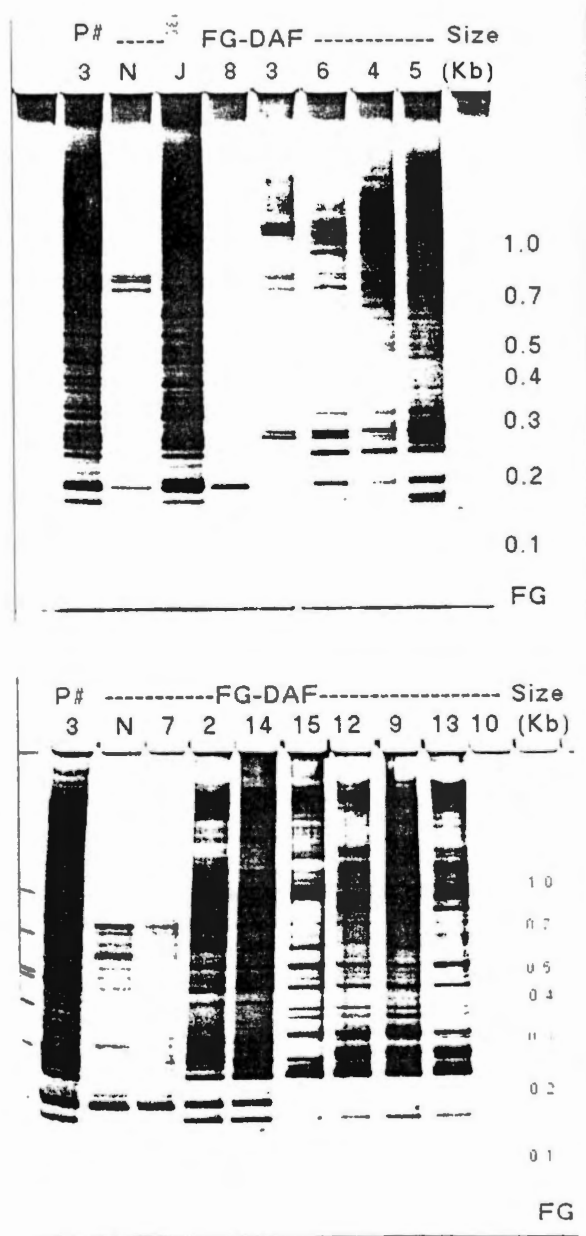


Figure 4-1. Polyacrylamide gels for 15 PCR programs using 8-4 (5'GTAACGCC3') primer for soybean DNA template cultivar Minsoy. P# 3 (shown in both panels) and N-DAF are the original DAF program run in twin-block Ericomp and MJR thermocyclers, respectively.

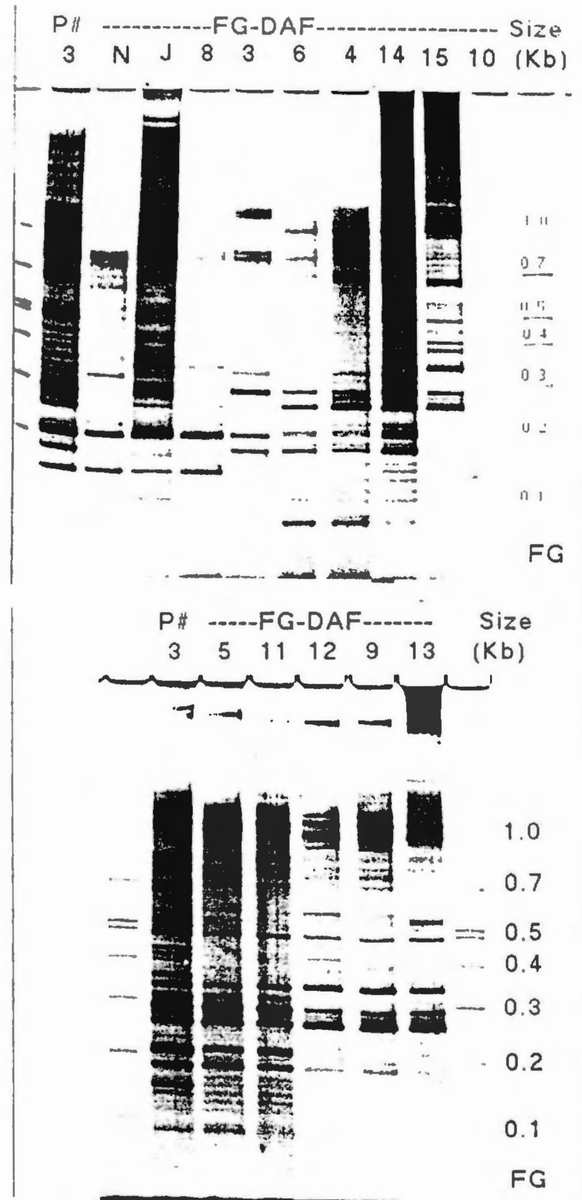


Figure 4-2. Polyacrylamide gels for 15 PCR programs using 8-4 (5'GTAACGCC3') primer for soybean DNA template cultivar Noir1. P# 3 (shown in both panels) and N-DAF are the original DAF program run in twin-block Ericomp and MJR thermocyclers, respectively.

extension step but having high annealing temperature and longer cycling time, generated bands with high intensity. This is probably due to an increase in the time of annealing and/or melting steps (Table 4-1) providing enough time for template-primer interaction and consequent generation of sufficient amplification products. Total number of the bands was high for all programs, ranging from 32 to 52 except for Program 10 with annealing temperature of 60°C which showed only 7 bands on average. Number of total bands changed from 42 to 52 when a 72°C/ 30 sec extension step was added to the Program 7. A similar conclusion, an increase in number of bands from 38 to 51, came from addition of extension step to the Program 8 with 30°C annealing temperature (Table 4-1) indicating a positive effect of extension step on increase of number of bands. However, addition of extension step in Programs 5 and 11 with 47°C annealing temperature generating Programs 2 and 14, respectively, did not apparently increased the number of the bands (Table 4-1). A similar conclusion came from the programs with 55°C annealing temperature, i.e. Programs 12 vs. 15 and 9 vs. 13 (Table 4-1). However the general quality such as sharpness and low level of background was apparently improved in the programs with extension step, i.e., 2, 14, 13, and 15 programs (Figures 4-1 and -2).

To examine the effect of annealing temperature on the total number of bands, the average number of bands was calculated 46, 49, and 45, and 7 over the 30°C, 47°C, 55°C, and 60°C annealing temperatures, respectively (Table 4-2). The original program in twin-block thermocycler (P# 3) and N-DAF in MJR machine containing identical cycling steps but differed in ramping showed 46 and 32 bands, respectively. Based on theoretical expectation, low annealing temperature, i.e. 30°C, should generate more bands than the higher annealing temperature. However, this is not the case here where average of the number of bands for 30°C, 47°C, and 55°C generated 46, 49, and 45, respectively, indicating that annealing temperature apparently did not affect the average number of bands (Table 4-2). Obviously, these conclusions are based on the subjective evaluation rather than objective and empirical analysis.

Ramping for cooling and heating rates were determined for the MJR thermocycle (Table 4-3) as well as in a twin-block thermocycler where the original DAF program, P#3, was run. The average of cooling and heating rates for MJR machine was 1.27 and 1.42 sec/°C, respectively. Cooling and heating ramps in the control machine, twin-block thermocycler, 2.15 and 1.65 sec/°C, respectively, were much longer than the MJR

Table 4-2. Effect of annealing temperature on generation of total bands in 15 programs. P# 3 and N-DAF, the original DAF programs with annealing temperature of 30°C for 1 sec and melting temperature of 95°C for 1 sec in Twin-Block (Ericomp, San Diego, CA) and MJR thermocyclers PT200, respectively, were used as controls.

Annealing Temperature (°C)	Program	Average Number of Bands
30	P#3 as control	46
30	N-DAF as control	32
30	7, J, 8, and 3	46
47	6, 4, 5, 2, 11, and 14	44
55	12, 15, 9, and 13	45
60	10	07

Table 4-3. Ramping of cooling and heating for the high annealing temperature DAF programs used in the MJR machine. Ramping of cooling and heating for the original DAF program, P#3, in a twin-block thermocycler (Ericomp) were 2.15 and 1.65 seconds/1°C, respectively.

Range (°C)	Δ° C	Time (sec)	seconds / 1°C
Cooling			
95-60	35	51	1.31
95-55	40	55	1.37
95-47	48	60	1.25
95-30	65	75	1.15
Average			1.27
Heating			
72-95	23	41	1.78
60-95	35	45	1.29
55-95	40	47	1.17
47-95	48	50	1.04
30-95	65	53	0.81
55-72	17	40	2.35
47-72	25	43	1.72
30-72	42	49	1.17
Average			1.42

thermocycler. This might explain why the P#3 generated high intensity and more numerous bands than the similar program, N-DAF, in the MJR machine (Figures 4-1 and -2; Table 4-2). This indicates that thermocycler ramping is a major factor in successful DNA amplification, and that each program in a new machine should be adapted to the original program to generate reproducible DNA amplification profiles, not only in terms of maximum and minimum temperatures, but also ramping rates.

When DNA amplification profiles of the original DAF, P#3, and other programs with annealing temperature of 30°C, i.e., N-DAF, 7, J, 8, and 3, were compared, only program J showed comparable profiles with high intensity and number of bands (Fig. 4-1 and -2). This might be due to presence of 30 sec extension step in J compared with N-DAF, 7, and 8 which lack an extension step. However, program 3 did not generate a comparable profile although it had an extension step, but with twice longer melting, annealing, and extension times (60 sec each). This might be the reason for the failure of program 3. One might speculate that longer time periods especially for melting temperature, lowered the efficiency of the Stoffel fragment DNA polymerase enzyme. Contrasting this reasoning, Program 2, with 120 sec for each step, showed an interesting DNA amplification profile (Figure 4-1), indicating that longer melting temperature alone may not the cause for problem with the Program 3.

Comparing the programs with 47°C annealing temperature, there was a consistent trend between the quality of DNA profiles and the annealing temperature time courses. The intensity (from low to high) and the number of bands (from 37 to 46) increased when the annealing period increased from 30 (Program 6) to 60 seconds (Program 4). Even the subjective quality of the DAF profile from the Program 4 was improved further when the 60 sec annealing period was modified to 120 sec for Program 5 (Figures 4-1 and -2). Modification of melting time appeared not to affect the quality of the DAF profiles as shown for Program 5 vs. 11 (Figure 4-2). Addition of extension step in Program 11 generated a DAF profile with an average of 50 bands with high intensity in Program 14. Thus, the addition of extension step increased the quality of DAF profile in programs with 47°C as well as those with 30°C annealing temperature.

To increase the stringency of the DAF conditions, four programs (12, 15, 9, and 13) were tested with 55°C annealing temperature which all showed high quality and high number of bands, ranged from 42 to 48, DAF profiles (Figures 4-1 and -2). However,

Program 15 generated more consistent results in terms of quality, sharpness, intensity, and resolution. This consistency might be due to an extra extension step (compared to Programs 12 and 9) and shorter melting temperature (compared to Program 9).

New Programs with other Primers

Several of the new programs were tested using arbitrary octamer primers other than 8-4 (5'GTAACGCC3'), i. e., either primer 8-8 (5'GAAACGCC3') or 8-9 (5'GTTACGCC3'). Primer 8-4 was also run in parallel to the other primer as control.

The original DAF program, P#3, in a twin-block was run as a control using Minsoy and Noir1 templates and two 8-4 and 8-9 primers (Figure 4-3A). Each of the two templates and primers were equally mixed prior to PCR to test whether either primers or templates are working by comparing the mixture and the normal DNA fingerprinting profiles. For example, if the Minsoy DNA is contaminated with an amplification inhibitor, the mix reaction will amplify the Noir1 template because the inhibitor is diluted in the mix and the DAF profile of the mix tends toward profile of Noir1 alone. This is also true for testing whether the primers either are contaminated or degraded. Since the primer is used for both template, failure or incomplete amplification can be judged by comparison of mix and individual DAF profiles.

Primer 8-9 showed the same quality of banding pattern as primer 8-4 except for the Noir1 template which generated a lower number of bands and low intensity of the bands. This might be due to a lower template concentration of Noir1 compared to Minsoy because the quality in mixture profile was between Minsoy and Noir1 profiles. This suggests that Minsoy template concentration could "compensate" the Noir1 concentration.

Program J was run using the 8-8 primer and showed almost similar quality of DAF profile as primer 8-4 (Figure 4-3B). The same conclusion can be reached, when Programs 8, 3, 14, and 15 were tested using primers 8-4 and 8-8 (Figures 4-4, -5B, and -6B) indicating that these programs work with different primer sequences. For further confirmations, Programs 11 and 9 were tested against primer 8-9 (Figures 4-5A and 4-6A). Although the quality and number of the bands were poor in primer 8-9

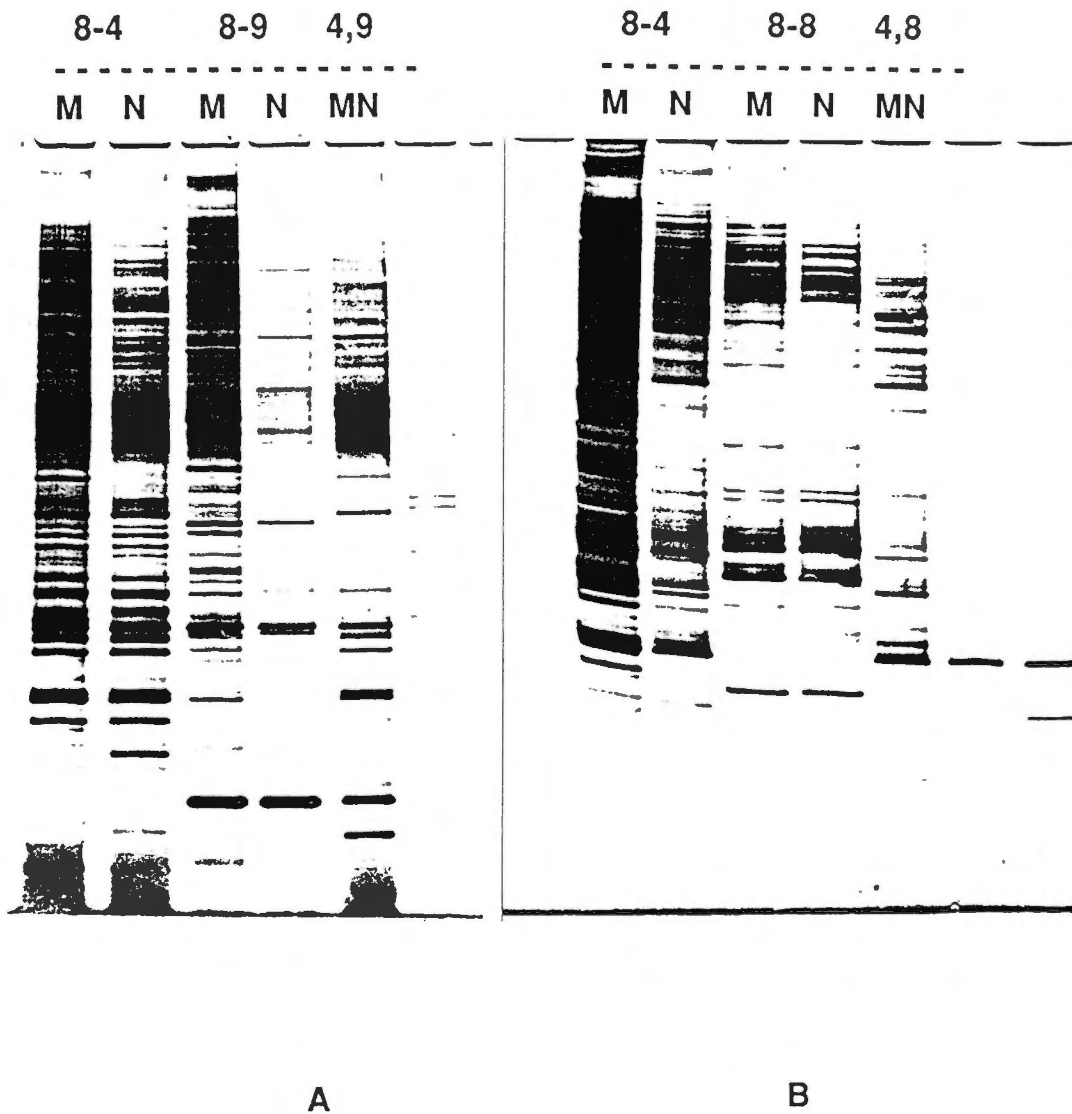


Figure 4-3. Polyacrylamide gels for original DAF, P# 3, (Panel A) and J (Panel B) programs using 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3'), 8-9 (5'GTTACGCC3') primers for soybean DNA template cultivars of Minsoy (M) and Noir1 (N). MN is a mixture of Minsoy and Noir1 and 4,8 or 4,9 are a mixture of primers 8-4 and 8-8 or 8-4 and 8-9, respectively, which were combined with equal volumes in a single reaction.

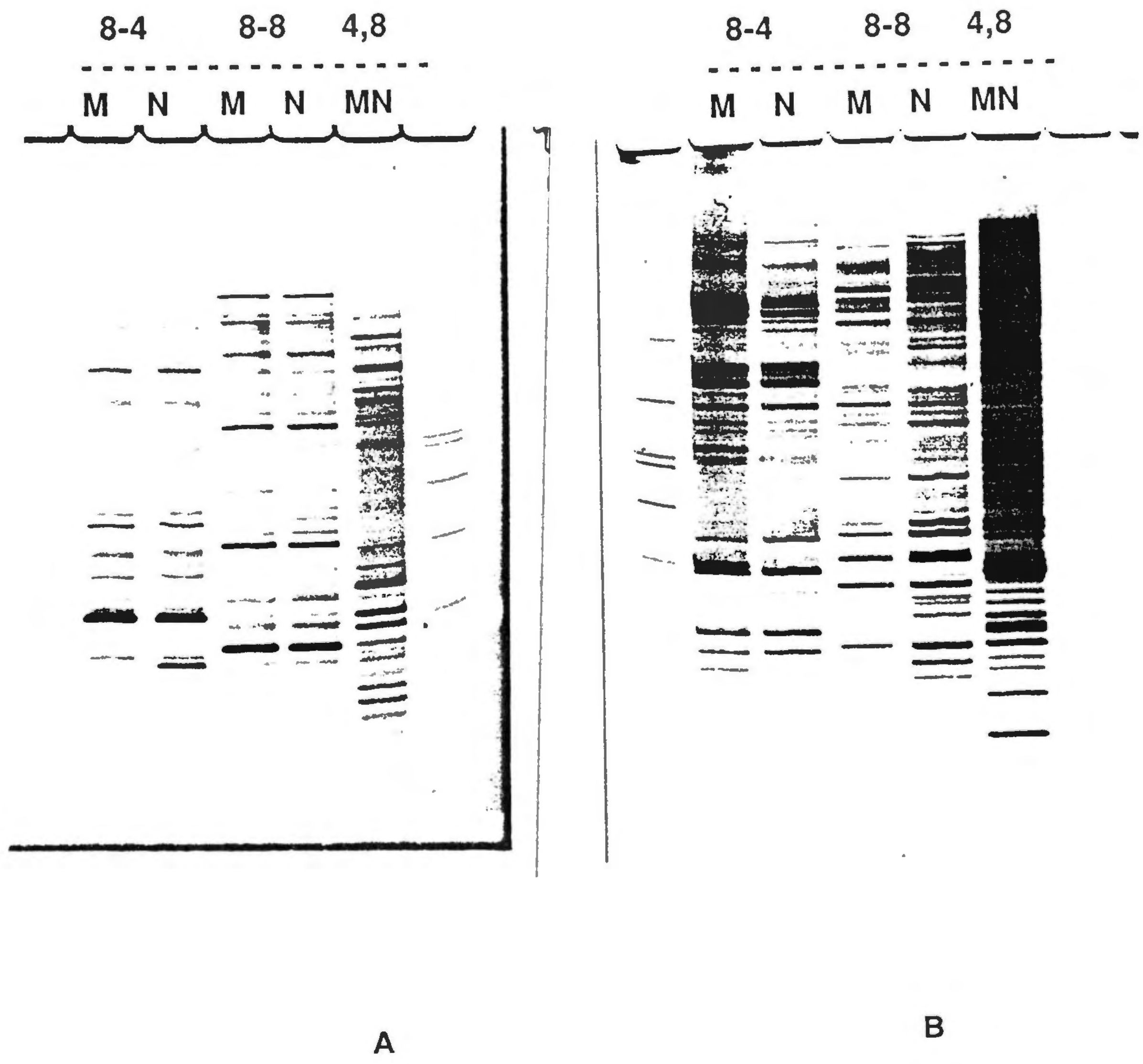


Figure 4-4. Polyacrylamide gels for programs 8 (Panel A) and original 3 (Panel B) using 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3') primers for soybean DNA template cultivars of Minsoy (M) and Noir1 (N). MN and primer mixture as in Figure 4-3.

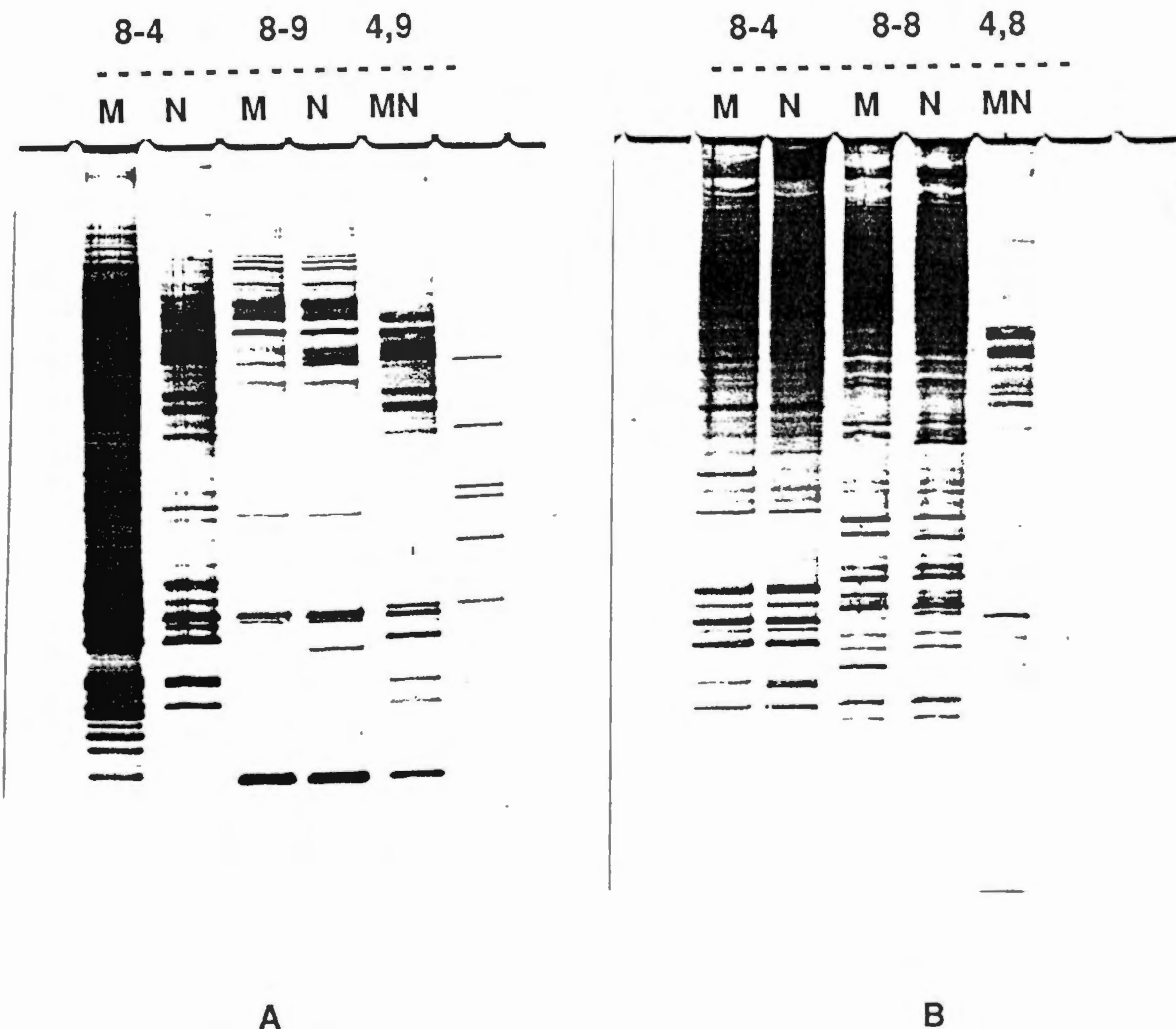


Figure 4-5. Polyacrylamide gels for Programs 11 (Panel A) and 14 (Panel B) using 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3'), 8-9 (5'GTTACGCC3') primers for soybean DNA template cultivars of Minsoy (M) and Noirl (N). MN and primer mixture as in Figure 4-3.

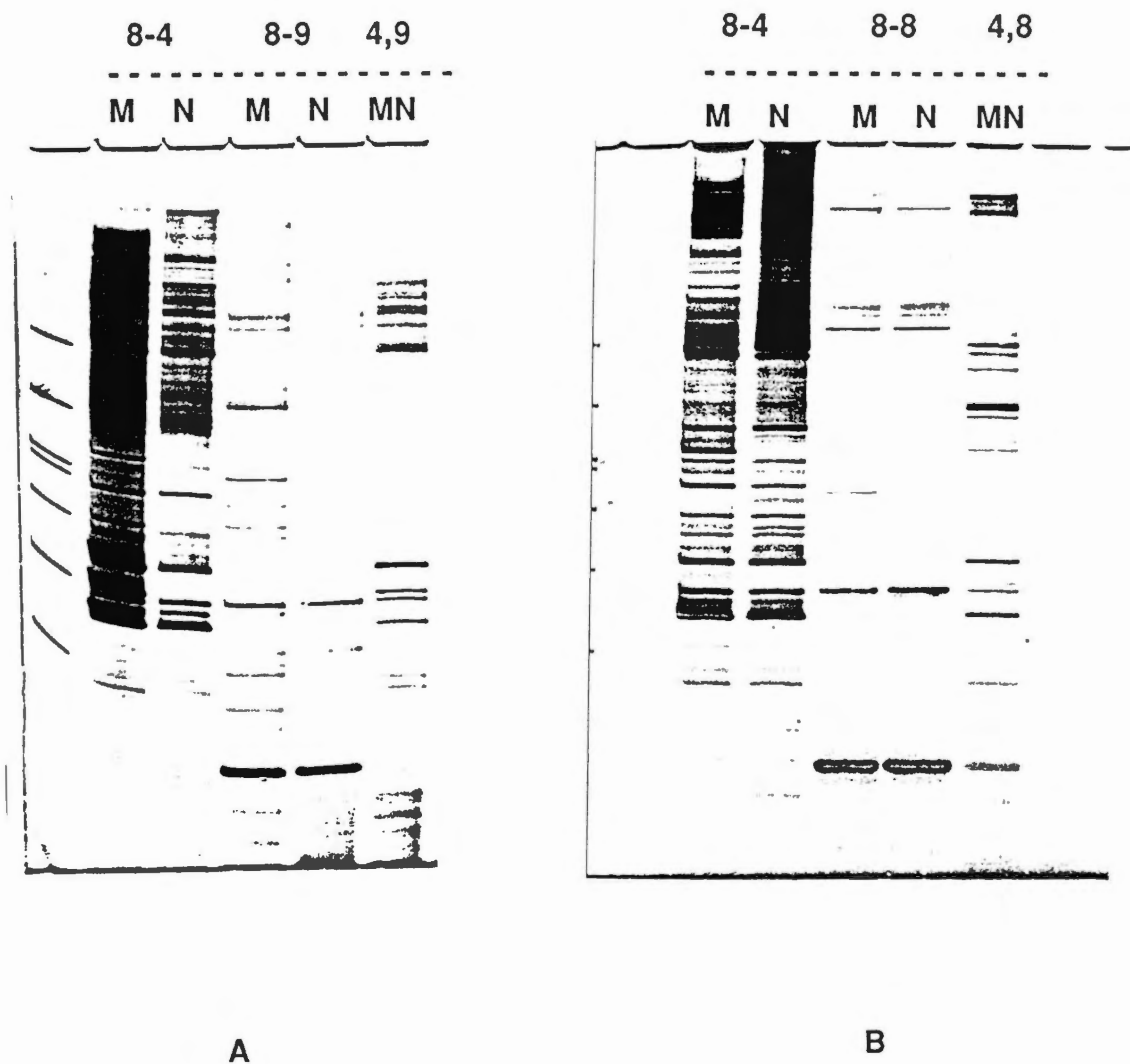


Figure 4-6. Polyacrylamide gels for Programs 9 (Panel A) and 15 (Panel B) using 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3'), 8-9 (5'GTTACGCC3') primers for soybean DNA template cultivars of Minsoy (M) and Noir1 (N). MN and primer mixture as in Figure 4-3.

compared to 8-4, the amplification was successful and generated an acceptable DAF profile.

In summary, Program # 15, (95°C/ 30 sec, 55°C/ 120 sec, 72°C/ 30 sec) resulted in sharp, high intensity, and large number of amplification products while running at high stringency annealing temperature. However, alternative programs are suggested (Programs 13, 9, and 12). Because the PCR is a very complex phenomenon to which many factors make a contribution, it is still possible to improve more the high annealing temperature DAF by manipulating other components of the reaction.

References

Aldrich, J. and Cullis, C. A. 1993. RAPD analysis in flax: Optimization of yield and reproducibility using KlenTaq1 DNA polymerase, chelex 100, and gel purification of genomic DNA. Plant Mol. Biol. Rep. 11: 128--141.

Arnold, M. L., Buckner, C. M., and Robinson, J. J. 1991. Pollen mediated introgression and hybrid speciation in Louisiana irises. Proc. Natl. Acad. Sci. USA 88: 1398-1402.

Bassam, B. J. and Bently, S. 1994. DNA fingerprinting using arbitrary primer technology (APT): A tool or a torment. Australian Biotechnol. 4: 232-236.

Bassam, B. J., Caetano-Anollés,G., and Gresshoff, P. M. 1992. DNA amplification fingerprinting of bacteria. Appl. Microbiol. Biotechnol. 38: 70-76.

Blanchard, M. M., Taillon-Miller, P., Nowotny, P., and Nowotny, V. 1993. PCR buffer optimization with uniform temperature regimen to facilitate automation. PCR Meth. Appl. 2: 234-240.

Caetano-Anollés, G., Callahan, L. M., Williams, P. E., Weaver, K. R., and Gresshoff, P. M. 1995. DNA amplification fingerprinting analysis of bermudagrass

(*Cynodon*): genetic relationships between species and interspecific crosses. Theor. Appl. Genet. 91: 228-235.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1994. Buffer components tailor DNA amplification with arbitrary primers. PCR Meth. Appl. 4: 59-61.

Caetano-Anollés, G. 1993. Amplifying DNA with arbitrary oligonucleotide primers. PCR Meth. Appl. 3: 85-94.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1992. Primer template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. Mol. Gen. Genet. 235: 157-165.

Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Bio/Tech. 9: 553-557.

Dellaporta, S. J., Wood, J., and Hicks, J. B. 1983. A plant DNA mini-preparation. Version II., Plant Mol. Biol. Reporter 1: 19-21.

Frohman, M. A., Dush, M. K., and Martin, G. R. 1988. rapid production of full length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Sci. Acad. USA 85: 8998-9002.

Gyllensten, U. B. and Erlich, H. A. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to the direct sequencing of the HLA-DQA locus. Proc. Natl. Acad. Sci. USA. 85: 7652-7656.

Hardeys, H., Balick, M., and Schierwater, B. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol. Eco. 1: 55-63.

Hu, J. and Quiros, C. F. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell Rep. 10: 505-511.

Huang, M. N., Arnheim, N., and Goodman, M. F. 1992. Extension of base mispairs by *Taq* DNA polymerase: Implications for single nucleotide discrimination in PCR. Nucl. Acids Res. 20: 4567-4573.

Jiang, Q. and Gresshoff, P. M. 1997. Classical and molecular genetics of the model legume *Lotus japonicus*. Mol. Plant-Microbe Interact. 10: 59-68.

Kwok, S., Kellog, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J. J. 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nucl. Acids Res. 18: 999-1005.

Levi, A., Rowland,I. J., and Hartung, J. S. 1993. Production of reliable randomly amplified polymorphic DNA (RAPD) markers from DNA of woody plants. HortScience 28: 1188-1190.

Liang, P and Pardee, A. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971.

Martin, G. B., Williams, J. G. K., Tanksley, S. D. 1991. Rapid identification of markers linked to a *Psuedomonas* resistance gene in tomato by using random primers and near-isogenic lines. Proc. Natl. Acad. Sci. USA 88: 2336-2340.

Men, A. E. and Gresshoff, P. M. 1998. Efficient cloning of DAF polymorphic markers from silver-stained polyacrylamide gels. BioTechniques 24: 593-595.

Mullis, K. B. and Faloona, F. A. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed reaction. Methods Enzymol. 255: 335-350.

Mullis, K. B. and Faloona, F. A., Scharf, S., Saiki, R., Horn, G., and Erlich, H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symp. Quant. Biol. 51: 263-273.

Ohara, O., Dorit, R. L., and Gilbert, W. 1989. One-sided polymerase chain reaction: the amplification of cDNA. Proc. Natl. Acad. Sci. USA. 85: 5673-5677.

Paran, I., Kesseli, R., Michelmore, R. 1991. Identification of restriction fragment length polymorphism and random amplified polymorphic DNA markers linked to downy mildew resistance genes in lettuce, using near-isogenic lines. Genome 30: 1021-1027.

Prabhu, R. R., Webb, D., Jessen, H., Luk, S., Smith, S., and P. M. Gresshoff. 1997. Genetic relatedness among soybean genotypes using DNA amplification fingerprinting (DAF), RFLP, and pedigree. Crop Sci. 37: 1590-1595.

Rashtchian, A., Buchman, G. W. Schuster, D. M., Berninger, M. S. 1992a. Uracil DNA glycosylase-mediated cloning of polymerase chain reaction-amplified DNA: application to genomic and cDNA cloning. Anal. Biochem. 206: 91-97.

Rashtchian, A., Thornton, C. G., and Heidecker, G. 1992b. A novel method for site-directed mutagenesis using PCR and uracil DNA glycosylase. PCR Meth. Appl. 2: 124-130.

Reiter, R. S., Williams, J. G. K., Feldmann, K. A., Rafalski, J. A., Tingey, S. V., and Scolnik, P. A. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Natl. Acad. USA 89: 1477-1481.

Saiki, P. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-1354.

Schierwater, B. and Ender, A. 1993. Different thermostable DNA polymerases may amplify different RAPD products. Nucl. Acids Res. 21: 4647-4648.

Somer, R. and Tautz, D. 1989. Minimal homology requirements for PCR primers. Nucl. Acids Res. 17: 6749.

Stiles, J. I., Lemme, C., Sondur, S., Morshidi, M. B., and Manshardt, R. 1993. Using randomly amplified polymorphic DNA for evaluating genetic relationship among papaya cultivars. Theor. Appl. Genet. 85: 697-701.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zebeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. 23: 4407-4414.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucl. Acids Res. 18: 7213-7218.

Welsh, J., Peterson, C., and McClelland, M. 1991. Polymorphisms generated by arbitrary primed PCR in the mouse: application to strain identification and genetic mapping. Nucl. Acids. Res. 19: 303-306.

Wilde, J., Waugh, R., and Powell, W. 1992. Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. Theor. Appl. Genet. 83: 871-877.

Williams J. G. K., Rafalski, J. A., and Tingey, S. V. 1993. In: "Methods in Enzymology" (ed. R. Wu). Academic Press, Orlando, FL, Vol. 218, pp. 704-740.

Williams J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531-6535.

Part 5

Applications of DNA Amplification Fingerprinting

In this part I will demonstrate applications of DNA amplification fingerprinting (DAF) in identification of linked markers to *enod2* gene using bulked segregant analysis (BSA) and DNA fingerprinting of soybean cyst nematode (SCN), mycorrhizal fungus, aphids, centipedegrass, and bermudagrass.

Bulked segregant analysis (BSA)

Genetic linkage map of soybean *enod2* gene discussed in Chapter 3 revealed that this gene is located in an interesting region of genome in approxity of *I* gene and most importantly of soybean cyst nematode (SCN) resistance gene. BSA, a method of choice for tagging a gene and saturating a genomic region of interest, is explained here to identify DAF markers linked to the *enod2* gene and to provide additional molecular markers in the genome containing this gene.

BSA is been widely used since its first introduction (Michelmore *et al.*, 1991) and is based on the comparison of two bulks (pools) of DNA generated from a segregating population. Each bulk is originated from individuals which have identical genotypes for gene or trait of interest but different for other genes. The two bulks are then screened with one or more molecular marker techniques such as RFLP, RAPD, DAF, and AFLP to detect polymorphic marker genetically linked to the gene of interest. Linkage between polymorphic marker and the target gene is confirmed and measured by using the individuals in segregating population of which the bulks were originated.

Study of cosegregation of the linked marker with other markers already mapped to a genetic linkage map results in mapping of the gene. In addition, molecular markers linked to the gene provides a starting point for map-based cloning of the gene. Furthermore, a linked marker would be a useful tool for the plant and animal breeders in marker-assisted breeding programs.

The BSA approach has been used in construction of genetic linkage map and saturation of a region of interest sparsely populated with markers not only in plant kingdom but also in animal kingdom including human. Any segregating population such as F_2 and advanced backcross population can be used in BSA. Using a F_2 population provides the greatest genetic window for the target locus while a greater focus around the locus is allowed when backcross population is considered for BSA. In a F_2 population,

heterozygous individuals are excluded from the analysis allowing a rapid identification of molecular markers from both parents. If the target locus is dominant, heterozygous and homozygous dominant genotypes can not be distinguished in the F_2 generation and have to be pooled. Therefore, only molecular markers associated in *cis* with the dominant target locus can be detected. In addition, analysis of F_3 individuals can be used to identify and exclude heterozygous genotypes for the target locus.

The BSA approach has been used in variety of plants for identification of markers linked to resistance genes to golden mosaic virus in common bean (Miklas *et al.*, 1996), leaf rust in wheat (Williams *et al.*, 1997), *Rhynchosporium* in barley (Barua *et al.*, 1993), nematode in peanut (Garcia *et al.*, 1996), virus x (Jong et al., 1997) and virus Y (Hamalainen *et al.*, 1997) in potato, downy mildew in lettuce (Michelmore *et al.*, 1991), *Leveillula* in tomato (Chunwongse *et al.*, 1997), *Phytophtora* in strawberry (Haymes *et al.*, 1997), scab in apple (Koller *et al.*, 1994), *tristeza* virus in citrus (Gmitter, *et al.*, 1996), *Rhizomania* in sugar beet (Giorio *et al.*, 1997), and *Melaspora* in *Populus* (Cervera *et al.*, 1996). Many of these resistance genes in peanut, sugar beet, tomato, potato, wheat, strawberry, and *Populus* have been successfully mapped.

The BSA facilitated saturation of a region of the genome conferring supernodulation (*nts-1*) locus of soybean using DAF markers (Kolchinsky *et al.*, 1997). In addition to nts-1, two ineffective nodulation phenotypes, *nod49* and *nod139*, in soybean were analyzed using BSA and DAF markers (Caetano-Anollés *et al.*, 1995). Then, the linked markers were confirmed by analysis of individuals in the segregating population and some of the markers were converted to sequence-characterized amplified regions (SCARs). These markers would be useful in high density mapping and in linking to cloned soybean DNA from bacterial and yeast artificial chromosome (BAC and YAC) libraries.

In pea, ineffective nodulation, *sym31*, mutation was linked to DAF markers using BSA approach (Men *et al.*, 1988, submitted). Three DAF markers were tightly linked to *sym31* gene and two of them flanking the gene were converted to sequence characterized amplified region (SCAR) which provides a tool for positional cloning of the *sym31* gene.

Quantitative trait loci (QTLs) controlling seed color in pea was mapped on four different genomic regions using BSA, molecular markers, and selection of distributional extremes (McCallum *et al.*, 1997). The BSA was more efficient than near isogenic lines (NILs) in finding eight RAPD markers linked to the dwarf gene in *Brassica napus* (Foisset *et al.*, 1995). Three dwarfing genes in oat were mapped on different chromosomes using BSA and RFLP markers (Milach *et al.*, 1997). Sex types in asparagus were studied by BSA, RAPD and SCAR (Jiang and Sink, 1997). Two RAPD markers linked to male type gene converted to SCAR markers for scoring female and male types in the mapping population. OTLs for growth and wood quality in *Eucalyptus grandis* (Grattapaglia *et al.*, 1996) and growth habit in apple (Hermat *et al.*, 1997) were mapped using BSA and RAPD markers.

In addition to plants, the BSA approach has been applied for genetic mapping and linkage analysis of a target gene in other organisms. In fungus, metalaxyl insensitivity in *Phytophthora infestans* was analyzed and found to be linked to RAPD and RFLP markers (Fabritius *et al.*, 1997). In chicken, the dominant white locus was mapped using microsatellite markers and BSA (Ruyter *et al.*, 1997). White phenotype in the axolotl of salamanders was linked to RAPD markers using BSA approach (Voss and Shaffer, 1996).

As reviewed briefly here, the BSA approach has been used for genetic analysis of a wide range of traits such plant resistance genes, supernodulation, ineffective nodulation, seed color, sex type, dwarf phenotype, and QTLs for wood quality and growth habit. All these examples can be identified by phenotype contrast to *enod2* gene which is the subject of this section.

Here, a novel case is presented for identification of a DAF marker to the *enod2* gene which does not condition any phenotype except that based on the deduced amino acid sequences it is suggested that the gene encodes a cell wall protein in the inner cortex of soybean nodule.

Materials and Methods

Parents of the segregating F_2 population were *G. max* cultivar Bragg, and the ancestral soybean *G. soja*, PI468.397. DNA samples from these parents were digested with 6

restriction nucleases, namely, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Taq*I, and *Xba*I. Parental crosses were made and verified as described by Landau-Ellis *et al.* (1991) using morphological and RFLP markers. After detection of polymorphisms in parents, Southern hybridization was carried out for 88 F_2 plants generated from a cross between Bragg and *G. soja*. Mendelian inheritance (1: 2: 1 ratios) of RFLP banding patterns of *enod2* in the F_2 population was confirmed with Chi Square test.

Since the *enod2* gene does not have a phenotype, homozygous individuals including recessive and dominant genotypes could only be identified by their RFLP patterns in the F_2 population. Hence, those F_2 individuals with RFLP patterns identical to homozygous parental patterns, Bragg and *G. soja*, were considered as bulks B and S, respectively.

DNA from each individual of two bulks was extracted and quantified separately as discussed in Chapter 3. Then, each DNA diluted at least twice followed by subsequent quantification to reach to a working concentration of 2 ng/µL. At the end, equal volume of DNA from each individual was mixed thoroughly to make working DNA concentration of bulks B and S for PCR.

Two bulks of DNA were initially screened for polymorphic bands using mini-hairpin and decamer primers and original DAF program, P# 3, as discussed in Chapter 4. After development of high annealing (55°C) temperature DAF, two bulks were screened further for identification of polymorphic bands using mini-hairpin and unstructured primers.

In addition, three sub-pools from B bulk (B1, B2, and B3) and two sub-pools from S bulk (S1 and S2) were randomly generated from either 7 or 8 F_2 individuals each. Then, B1 versus S1 and B3 versus S2 were screened for detection of polymorphic bands using mini-hairpin and unstructured primers and high annealing temperature DAF (DAF-15) program.

Results and Discussion

RFLP patterns of 41 homozygous F_2 individuals for *enod2* gene were set into two bulks of 26 and 15 with RFLP patterns identical to their parental patterns Bragg and *G. soja*, respectively (Table 5-1). Bulks B and S were primarily screened for identification of

Table 5-1. RFLP patterns for *enod2* in a F_2 population.

RFLP patterns of homozygous individuals for *enod2* gene in a F_2 segregating population generated from a cross between cultivar Bragg and the ancestral soybean *G. soja*, PI468.397. B (26 individuals) and S (15 individuals) are RFLP patterns identical to Bragg and *G. soja*, respectively. There were 47 heterozygous individuals which are not shown here.

Order	Plant #	RFLP Pattern	Order	Plant #	RFLP Pattern	Order	Plant #	RFLP Pattern
01	2	S	15	36	B	29	62	S
02	3	B	16	38	B	30	63	B
03	5	S	17	40	B	31	68	B
04	7	S	18	41	B	32	70	B
05	11	S	19	43	S	33	71	B
06	13	B	20	44	B	34	74	B
07	15	B	21	46	S	35	76	S
08	19	S	22	47	B	36	79	S
09	21	B	23	50	S	37	80	B
10	26	S	24	52	B	38	81	B
11	29	B	25	55	B	39	83	S
12	31	B	26	57	B	40	86	S
13	33	B	27	59	B	41	88	B
14	34	B	28	61	S			

polymorphic markers using 76 mini-hairpin and unstructured primers and original DAF program (P#3) run in a twin-block thermocycler (Table 5-2). A total of 59 polymorphic bands were detected but due to inconsistency and irreproducibility of DNA amplification stemmed from low annealing temperature (30°C) and mismatch of primer and template, the bulks were screened further using high annealing temperature 55°C (DAF-15) program run in the MJR thermocycler.

Screening of the bulks B and S with 31 primers resulted in detection of four polymorphic bands using primers HpC29 and HpC30 and DAF-15 program (Figure 5-1). Due to low number of polymorphic bands in the B and S bulks, sub-pools were generated and screened.

Results of randomly-selected sub-pools B1, B2, B3, S1, S2 are shown in Table 5-3. B1 and S1 were tested with total 196 primers of which 32 were used for screening of sub-pools B3 versus S2 (Table 5-4). Primers Hp30, HpC22 and HpC30 generated 1, 1 and 4 polymorphic markers, respectively, using DAF-15 program (Table 5-4; Figures 5-2 and -3).

The major screening was focused on the B1 versus S1 sub-pools which resulted in screening of 196 mini-hairpin and unstructured primers (Table 5-4) of which a set of 9 primers detected 20 polymorphic bands (Table 5-5; Figures 5-3, through -6). Primer HpD25 generated polymorphic bands with 920B1, 320B1, 220S1, and 185B1 base pairs which were reliable and reproducible (Figure 5-6). Although a large number of primers (196), different programs (P#3 and DAF-15), and different bulks and sizes (Table 5-3) were used to detect polymorphic bands linked to the *enod2* gene, only a portion of primers could detect 30 polymorphic bands. Some of these bands failed to reproduce in subsequent amplification due to either the complex nature of amplification or the size of bulks. However, four polymorphic bands generated from HpD25 were reamplified in second amplification indicating that these bands are promising bands for further analysis such as cloning and generating SCAR markers in the region of genome containing the *enod2* gene.

Table 5-2. Results of primary screening of primers for the *enod2* gene in bulked segregant analysis (BSA) using original DAF. Bulks B and S generated from 26 and 15 F_2 individuals with RFLP patterns identical to homozygous parental patterns, Bragg and *G. soja*, respectively. Primer sequences are listed in the appendix. Class I is presence *versus* absence of polymorphic band while Class II and Class III are presence of bands with different intensities.

Primer	HPs 53 (rerun)		HPA 46		HPA 47		HPA 48		HPA 49		HPA 42		HPA 43	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands			51	51	58	58	52	52	43	43	65	65	41	41
Class I														
Class II														
Class III					+0.7 +0.6 +0.2 +0.1	- - - -	-	+0.71	+0.7 +0.68	- -	+>1.0 -	- +0.8	+0.8 +0.05	- -
Gel	5 - 30 - 95		5 - 30 - 95		5 - 30 - 95		5 - 30 - 95		5 - 30 - 95		5 - 30 - 95		5 - 30 - 95	

Table 5-2 (continued)

Primer	HPA 54 (rerun)		HPA 55 (rerun)		HPA 56		HPA 57		HPA 58		HPA 59		HPA 60	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands					64	64	53	53	68	68	53	53	41	41
Class I													i	
Class II														
Class III					+ 0.7	-	-	+0.18	-	+0.09	-	+0.12	+0.6 +0.25 +0.12	- - -
Gel	5 - 30 - 95		5 - 30 - 95		5 - 30 - 95		5 - 30 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95	

Table 5-2 (continued)

Primer	HPA 44		HPA 45		HPA 60		HPA 65		HPA 67		HPA 70	
	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	44	44	44	44	45	45	34	34	37	37	47	47
Class I												
Class II									- -	+>1.0 +0.52		
Class III	+0.7	-	+0.7 +0.25	- -	+0.35 +0.3 +0.2	- - -	-	+0.15	-	+0.1	-	+0.28
Gel	5 - 30 - 95		5 - 30 - 95		6 - 5 - 95		6 - 5 - 95		6 - 5 - 95		6 - 5 - 95	

Table 5-2 (continued)

Primer	HPA 61		HPA 62 (rerun)		HPA 63		HPA 64 (rerun)		HPA 65		HPA 66		HPA 67	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	51	51			51	51			Incomplete	50	55	55	Incomplete	38
Class I									-	+0.15 ?				
Class II					+0.4	-								
Class III	+0.7	-			-	+0.12								
Gel	5 - 31 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95	

Table 5-2 (continued)

Primer	HPA 68		HPA 69		HPA 70		HPA 71		HPA 72		HPB 1		HPB 2	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	25	25	48	48	57	57	40	40	49	49	42	42	36	36
Class I														
Class II	+0.8	-						- -	+1.2 +0.18						
Class III														
Gel	5 - 31 - 95		5 - 31 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95	

Table 5-2 (continued)

Primer	HPB 3		HPB 4		HPB 5		HPB 6		HPB 7		HPB 8		HPB 9	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	53	53	52	52	46	46	58	58	60	60	40	40	48	48
Class I														
Class II													-	+0.6
Class III							+1.0	-			-	+0.3		
Gel	6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95	

Table 5-2 (continued)

Primer	HPB 10		HPB 11 (rerun)		HPB 12		HPB 13		HPB 18		HPB 19		HPB 20	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	40	40			45	45	47	47	40	40	48	Incomplete	60	60
Class I														
Class II														
Class III					+0.7	-			+0.85	-				
Gel	6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 2 - 95		6 - 5 - 95		6 - 5 - 95		6 - 5 - 95	

Table 5-2 (continued)

Primer	HPB 21		HPB 22		HPB 23		HPB 24		HPB 25		HPB 26 (rerun)		HPB 27	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	41	41	44	44	43	Incomplete	50	50	54	54			49	49
Class I														
Class II							-	+0.2						
Class III							-	+0.3	+>0.1	-				
Gel	6 - 5 - 95		6 - 5 - 95		6 - 5 - 95		6 - 5 - 95		6 - 5 - 95		6 - 7 - 95		6 - 7 - 95	

Table 5-2 (continued)

Primer	HPB 28		HPB 29		HPB 34 (rerun)		HPB 35 (rerun)		HPB 36		HPB 37		HPA 36	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	53	53	55	55					51	51	57	57	Incomplete	37
Class I														
Class II														
Class III											- +0.42	+0.17 _		
Gel	6 - 7 - 95		6 - 7 - 95		6 - 7 - 95		6 - 7 - 95		6 - 7 - 95		6 - 7 - 95		5 - 25 - 95	

Table 5-2 (continued)

Primer	HPA 37 (rerun)		HPA 38		HPA 39		10.6 d		10.6f		10.7h		10.6e	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands			Incom plete	43	41	41	31	31	34	34	31	31	23	23
Class I														
Class II														
Class III					-	+0.52	- - - -	+0.6 +0.33 +0.32 +0.05					- -	0.3 0.28
Gel	5 - 25 - 95		5 - 25 - 95		5 - 25 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95		5 - 31 - 95	

Table 5-2 (continued)

Primer	HPC 22		HPC 23		HPC 25		HPC 29		HPs 50		HPs 51		HPs 52	
	B	S	B	S	B	S	B	S	B	S	B	S	B	S
# of monomorphic bands	41	41	51	51	52	52	49	49	54	54	33	33	51	51
Class I														
Class II														
Class III									-	+0.22	+0.7 +0.3 +0.2	- - -	- +0.8	>1.0 -
Gel	8 - 9 - 95		8 - 9 - 95		8 - 9 - 95		8 - 9 - 95		5 - 30 - 95		5 - 30 - 95		5 - 30 - 95	

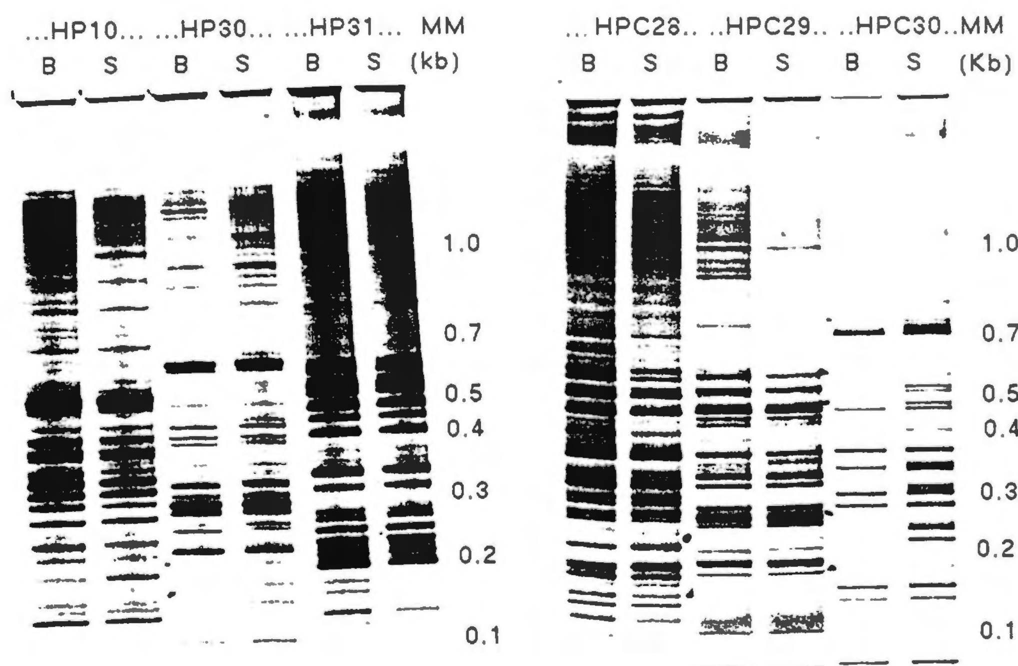


Figure 5-1. Silver-stained 10% polyacrylamide DAF gels for two bulks B and S using 55°C annealing temperature DAF. B and S were generated from 26 and 15 F_2 individuals with RFLP patterns identical to homozygous parental patterns, Bragg and *G. soja*, respectively. Mini-hairpin primers were Hp10 (5'GCGAAGCCTG3'), Hp30 (5'GCGAAGCCTT3'), Hp31 (5'GCGAAGCCAG3'), HpC28 (5'GCGAGAGCTGA3'), HpC29 (5'GCGAGAGCTAC3'), HpC30 (5'30GCGAGAGCTAT3'). MM is moleculr marker ranged from 0.1 to 1.0 kb.

Table 5-3. Plant # of F_2 individuals in sub-pools of B(26) and S(15) bulks.

Sub-pools	1	2	3	4	5	6	7	8
B1	33	40	41	59	71	80	88	-
B2	13	21	44	47	55	68	70	
B3	29	34	36	38	52	57	63	74
S1	05	26	46	61	62	76	79	-
S2	02	07	11	19	43	50	83	86

Table 5-4. Bulks and sub-bulks of F_2 individuals for the soybean *enod2* gene. These are based on the banding patterns identical to Bragg (B) and *G. soja* (S) screened with mini hairpin and unstructured (octamer and decamer) primers. A subset of the bulks were randomly selected and designated as B1, S1, B3, and S2.

Bulks	Number of Individuals	Number of Mini Hairpin Primers	Number of Unstructured Primers	Number of Total Primers
B S	26 15	25	6	31
B1 S1	7 7	106	90	196
B3 S2	8 8	32	-	32

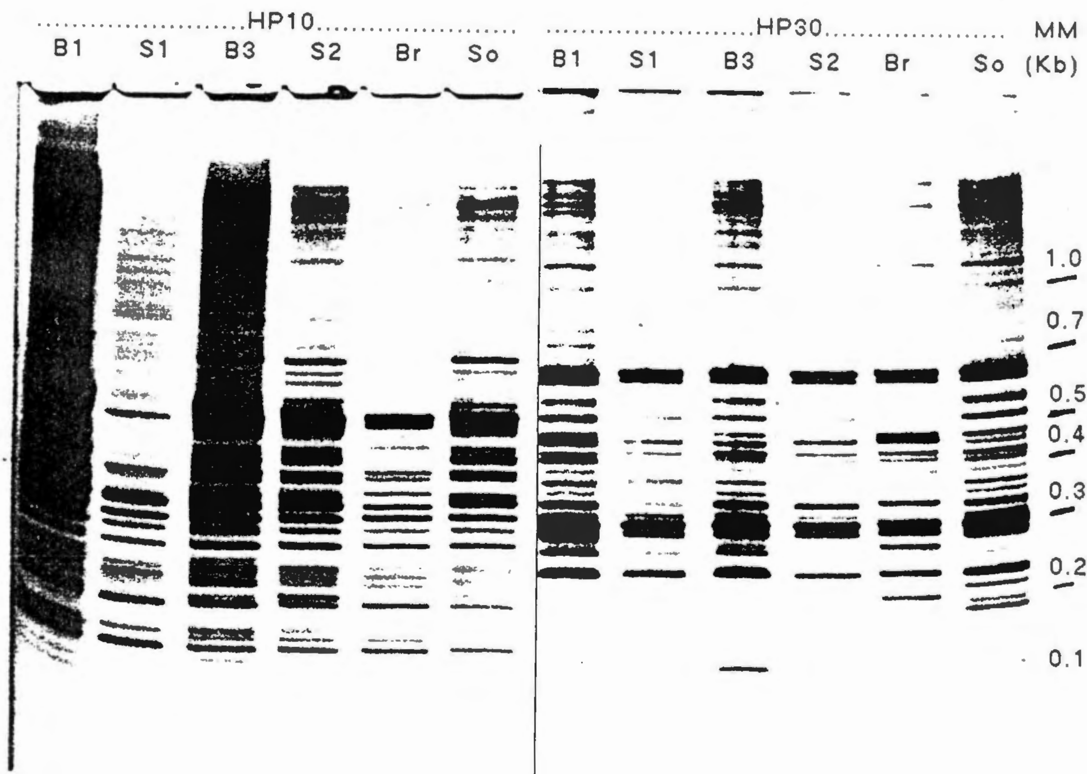


Figure 5-2. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1, S1, B3, and S2 along with parents Bragg (Br) and G. *soja* (S0) using 55°C annealing temperature DAF. Primers Hp10 and Hp30 as in Figure 5-1. MM is moleculr marker ranged from 0.1 to 1.0 kb.

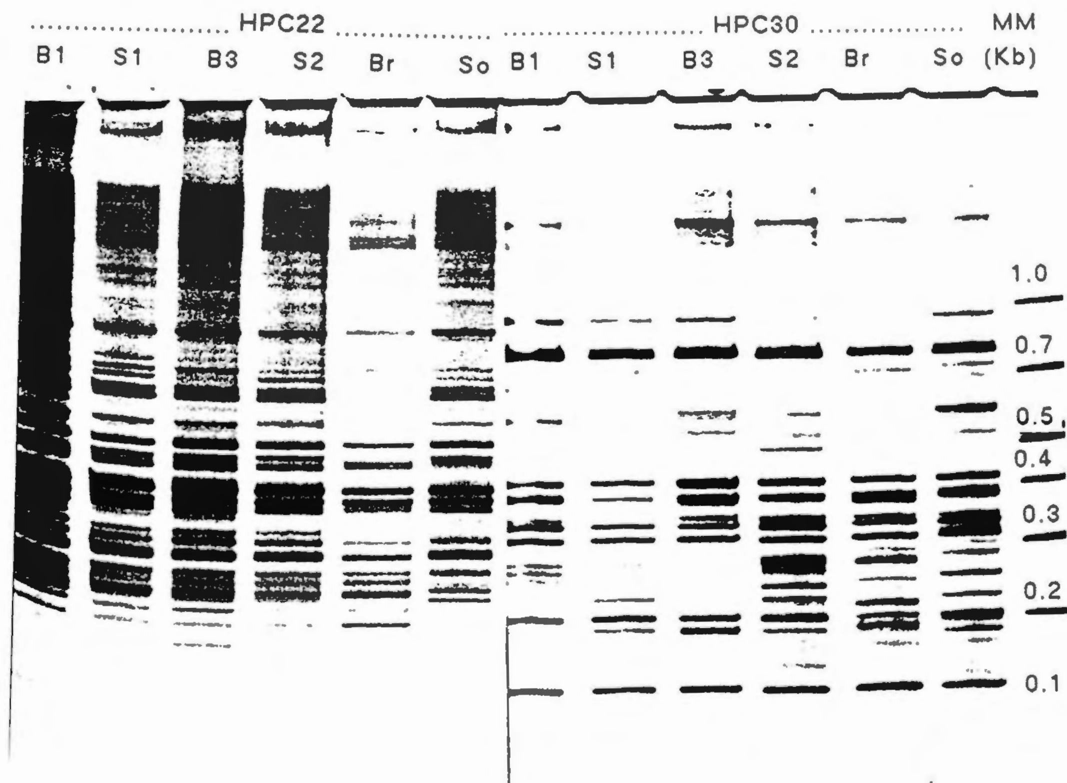


Figure 5-3. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1, S1, B3, and S2 along with parents Bragg (Br) and *G. soja* (SO) using 55°C annealing temperature DAF. Primers used for amplification were HpC22 (5'GCGAGAGCTTT3') and HpC30 (5'GCGAGAGCTAT3'). MM is moleculr marker ranged from 0.1 to 1.0 kb.

Table 5-5. Size of polymorphic bands (base pairs) resulted from screening of BSA for the *enod2* gene using octamer (Oc) and mini-hairpin (Hp) primers and high annealing temperature DAF (DAF-15) prgram run in a MJR thermocycler. B and S are bulks generated from pooling 26 and 15 F_2 individuals. B1, S1, B3, and S2 are sub-pools as in Table 5-3.

Primer	Polymorphic Band	Primer	Polymorphic Band
HpC29	230B 225S	HpC30	185S 180B
Hp30	520 B3	HpC22	190S2
HpC30	290B3 280S2 250B 240S2	OcA13	290B1
OcA43	330B1	OcB7	300B1 295S1 290S1 285B1
OcB19	120B1	OcB38	190S1 185B1 120S1
Hp10	390B1 330B1	Hp30	510B1 470S1
HpC22	180B1 170S1	HpD25	920B1 320B1 220S1 185B1

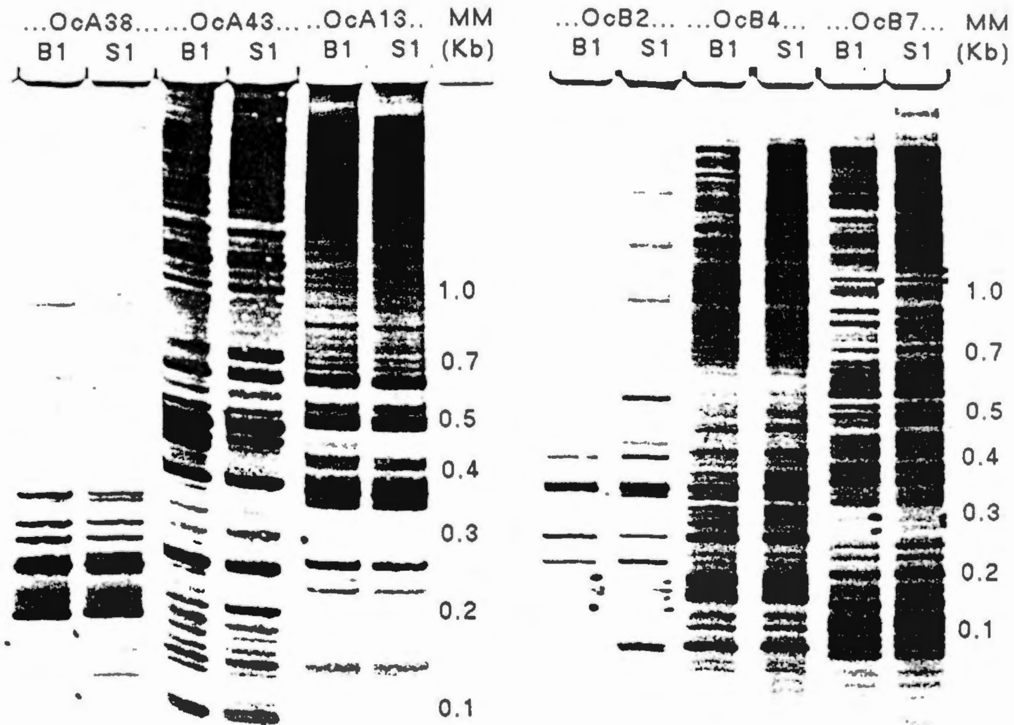


Figure 5-4. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1 and S1 using 55°C annealing temperature DAF. Octamer primers were OcA 38 (5'GCCCGGTT3'), OcA43 (5'GCCCGGGG3'), OcA13 (5'GCCCGCAC3'), OcB2 (5'GCAGGCCT3'), OcB4 (5'GCAGGCCA3'), and OcB7 (5'GCAGGCTG3'). MM is molecular marker ranged from 0.1 to 1.0 kb.

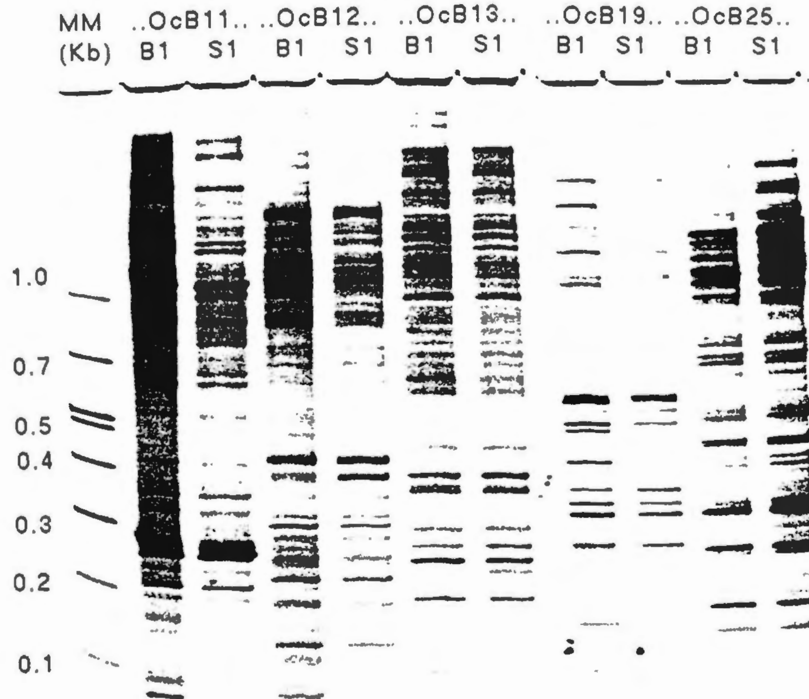


Figure 5-5. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1 and S1 using 55°C annealing temperature DAF. Octamer primers were OcB11 (5'GCAGGCGG3'), OcB12 (5'GCAGGCGA3'), OcB13 (5'GCAGGCAC3'), OcB19 (5'GCAGGTCG3'), AND OcB25 (5'GCAGGTGC3'). MM is moleculr marker ranged from 0.1 to 1.0 kb.

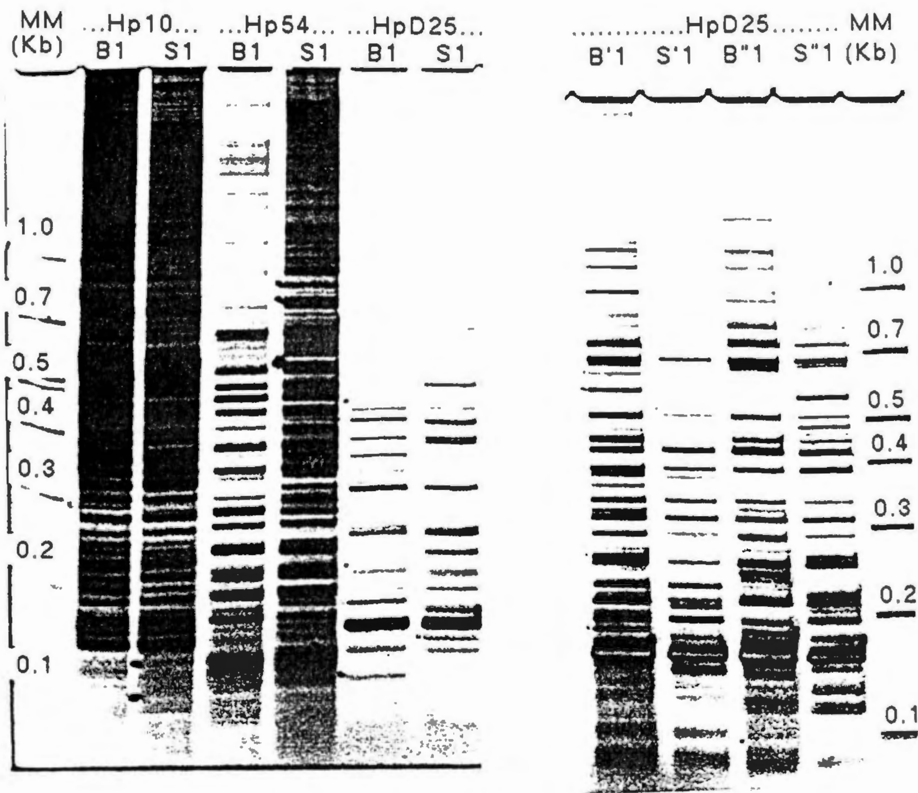


Figure 5-6. Silver-stained 10% polyacrylamide DAF gels for sub-pools B1 and S1 using 55°C annealing temperature DAF. B'1 and B"1 are second and third amplification reactions for B1 bulk, respectively. Mini-hairpin primers were Hp10 (5'GCGAAGCCTG3'), Hp54 (5'GCGAAGCTTT3'), and HpD25 (5'GCGATAGCTGC3'). Polymorphic bands for HpD25 at 920B1, 320B1, 220S1, and 185B1 are reproducible. MM is moleculr marker ranged from 0.1 to 1.0 kb.

DNA Fingerprinting

I used DAF technique for fingerprinting of not only plant DNA but also nematode, aphid and *Mycorrhizae* fungus DNA. The purposes of DNA fingerprinting were race distinction of soybean cyst nematode (SCN), type differentiation of aphids, isolate analysis of *Mycorrhizae* fungus, analysis of genetic stability of centipedegrasses, and commercial DNA fingerprinting of bermudagrasses.

Race Distinction of SCN

Question. Can DAF differentiate SCN races?

Soybean cyst nematode is one of the most widespread disease in the north central and southeastern regions of the USA which is responsible for the largest loss in the yield than any other soybean disease (see Yazdi-Samadi *et al.*, 1996). The purpose of this study was to differentiate three major SCN races 3, 14, and 5 using single arbitrary primers and DAF technology.

DNA extraction was conducted on the egg mass of each race using small scale DNA isolation procedure (Yoon *et al.*, 1991). Initial screening was performed for the SCN races 3 and 14 using 27 octamer and decamer primers. Total polymorphic bands were 75 of which 38 and 37 bands detected in the races 3 and 14, respectively. Average of polymorphic bands/primer was obtained 2.8 with average size 197.5 base pairs.

Primer screening was further conducted for the races 3, 14, and 5 using 13 primers of which 5 primers could detect 12 polymorphic bands (Table 5-6; Figures 5-7 to 5-9).

Conclusions

Primer 8-9 generated four polymorphic bands which could differentiate race 5 from the races 3 and 14. In all cases except for primer 8-27 which differentiate the race 3 from races 14 and 5, polymorphic bands were present in the races 3 and 14 but absent in the race 5 vice versa. This indicates that the races 3 and 14 are genetically related more closely than the race 5.

Table 5-6. Results of polymorphic bands (base pairs) detected in soybean cyst nematode (SCN) races 3, 14, and 5 using single arbitrary primers and original DAF method.

Primer	Sequence	Race 3	Race 14	Race 5
8-8	5'GAAACGCC3'	-	-	140 180
8-9	5'GTTACGCC3'	70 - 155 -	70 - 155 -	- 80 - 160
8-25	5'CGTGGTGG3'	- 205 180 130	- 205 180 130	210 - - -
8-27	5'CCTCGTGG3'	190	-	-
8-31	5'CCTGGTGC3'	230	230	-

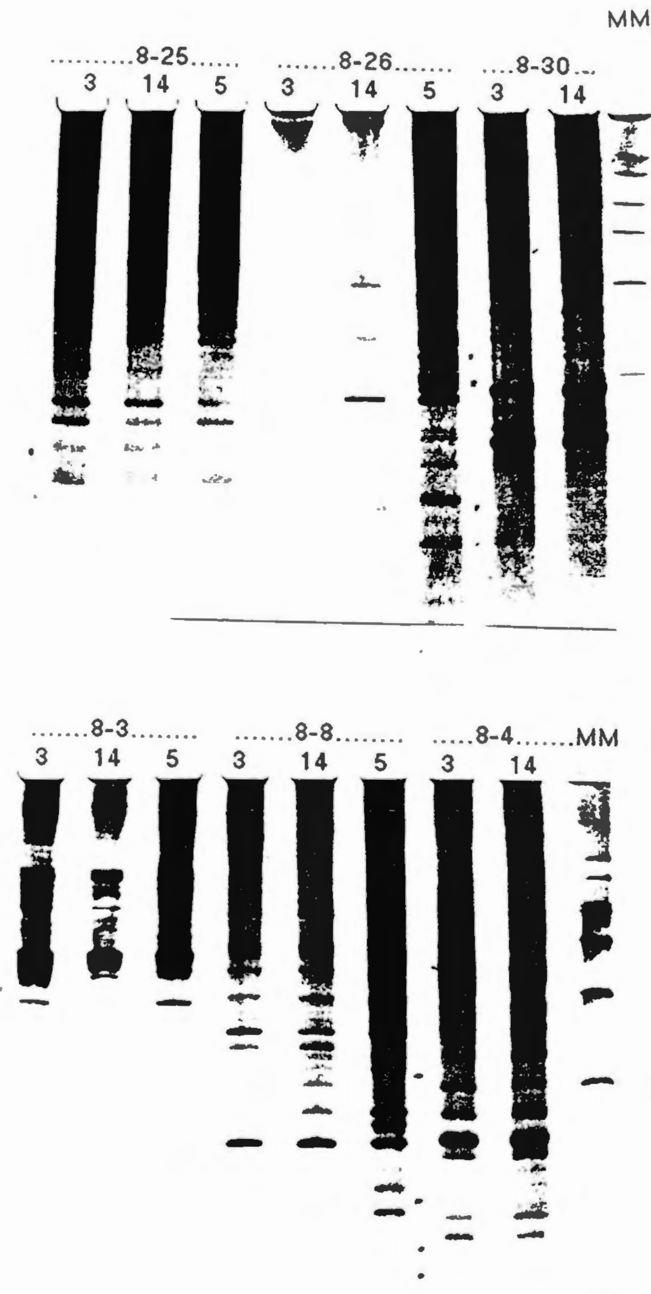


Figure 5-7. Silver-stained 5% polyacrylamide DAF gels for SCN races 3, 14, and 5 using original DAF program. Sequence of the 8-25, 8-26, 8-30, 8-3, 8-8, and 8-4 primers are listed in the appendix. MM, moleculr marker consists of seven bands 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, and 1.0 kb.

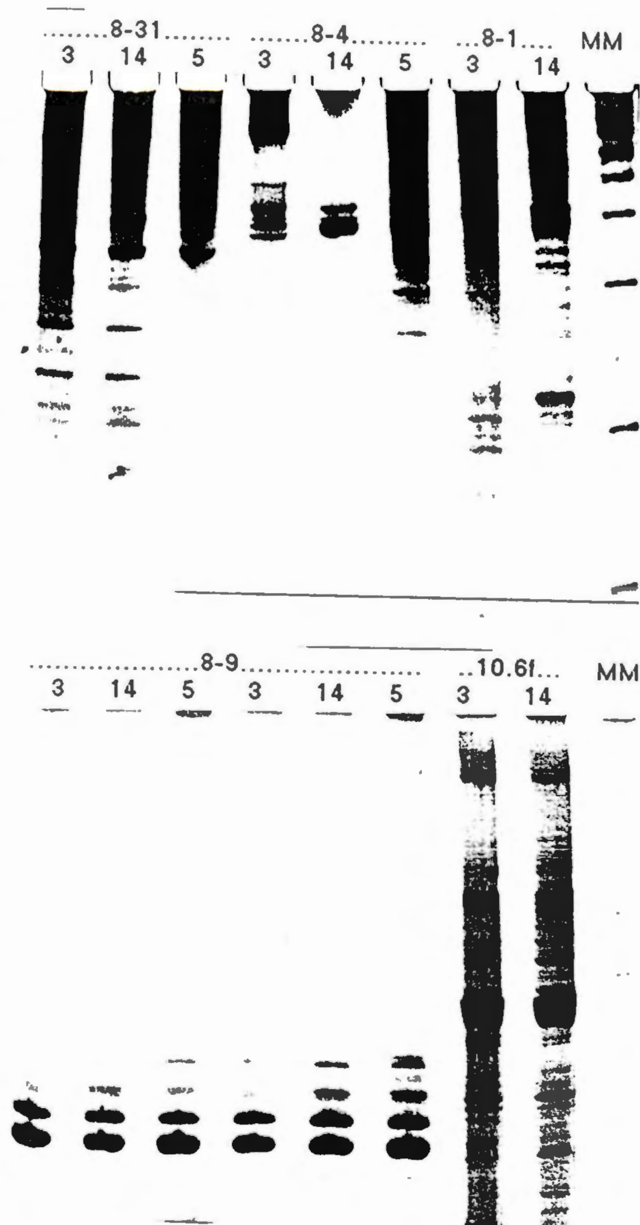


Figure 5-8. Silver-stained 5% polyacrylamide DAF gels for SCN races 3, 14, and 5 using original DAF program. Sequence of the 8-31, 8-4, 8-1, 8-9, and 10.6f primers are listed in the appendix. MM, moleculr marker consists of seven bands 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, and 1.0 kb.

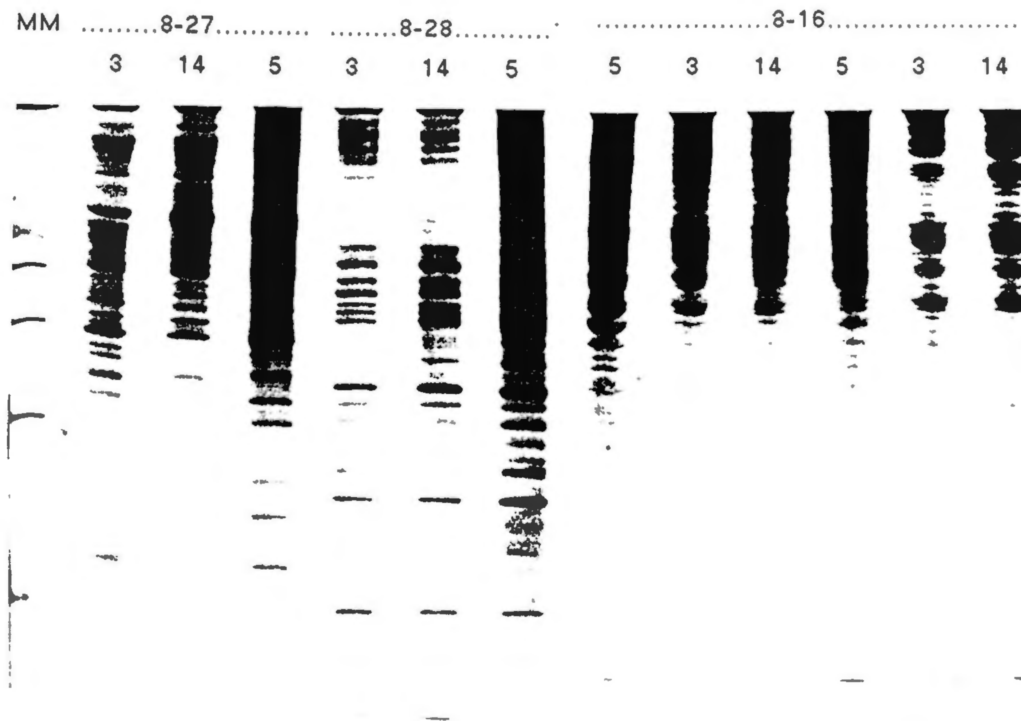


Figure 5-9. Silver-stained 5% polyacrylamide DAF gels for SCN races 3, 14, and 5 using original DAF program. Sequence of the 8-27, 8-28, and 8-16 primers are listed in the appendix. MM, moleculr marker consists of seven bands 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, and 1.0 kb.

DNA Fingerprinting of Mycorrhizal Fungi

Question. Can DAF be used as a distinctive tool for DNA fingerprinting of *Mycorrhizae*-associated loblolly pine roots?

In collaboration with Dr. E. G. O'Neill and R. A. Brewer at Oak Ridge National Laboratory (ORNL), loblolly pine (*Pinus taeda* L.) mycorrhizal roots infected by ectomycorrhizal fungi, and fruiting bodies of mycorrhizal fungus *Telophora terrestris* Pers. were collected and their DNAs were extracted. Several strains including some unknown fungi arising from stock cultures due to either contamination or mutation were compared (Gresshoff *et al.*, 1998). DNA from mycorrhizal roots, pine needles, fungal cultures and fruiting bodies were extracted and quantified. Primer 8-4 (GTAACGCC) effectively amplified the DNA samples using P#3 in a twin-block thermocycler (Ericomp). Additional octamer, decamer, and mini-hairpin primers which can be used for DAF are listed as 8-4, 8-5, 8-6, 10.6e, 10.6i, Hp10, and HpA33.

Comparison of mycorrhizal roots and needles showed different DNA amplification profiles (Figure 5-10). Some bands from fungal culture did not find in mycorrhizal roots suggesting either DNA competition or the true inoculum resulting in the mycorrhizal roots is not identical with the tested fungal culture. DNA amplification competition has been previously investigated in mixtures of DNA in symbiotic *Azolla-Anabaena* tissues (Eskew *et al.*, 1993)

Conclusions

The results of this study indicate that DAF can be used to differentiate mycorrhizal tissue and fungal isolate. The fungal-specific band at 210 base pairs can be excised and cloned (Men *et al.*, 1998; Weaver *et al.*, 1994) and generate a SCAR marker as an additional tool for analysis of mycorrhizal samples.

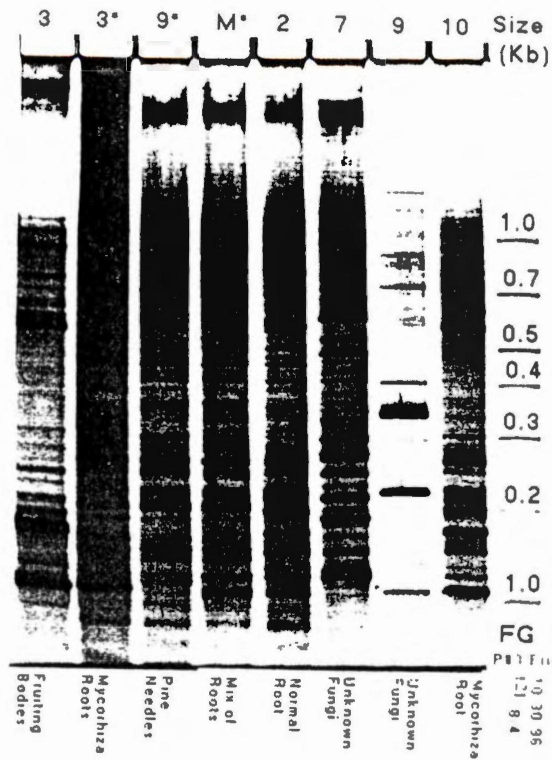


Figure 5-10. Silver-stained 10% polyacrylamide gels of DNA amplification fingerprinting using original DAF program and 8-4 primer (5'GTAACGCC3') for mycorrhizal tissue of loblolly pine and ectomycorrhizal fungal tissues. Lane 3, fruiting body of *Telophora terrestris*; lane 3*, mycorrhizal roots; lane 9*, loblolly pine needles; lane M, mixture of roots from different sources; lane 2, loblolly pine roots free of mycorrhizae; lane 7 and 9, unknown fungal isolates; lane 10, mycorrhizal root infected by a fungus. Molecular size is shown on the right.

DNA Amplification Fingerprinting of Aphid

Question. Can DAF be a choice of molecular marker for aphid DNA fingerprinting?

In collaboration with Dr. C. Niblett, University of Florida, 19 different types of aphids were studied to conduct DNA fingerprinting. DNA extraction was performed for each aphid type using small scale DNA isolation procedure (Yoon *et al.*, 1991). After quantitation, DNA of each type was diluted twice and quantified to obtain working concentration of 1 ng/μL. An octamer primer, OR30 (5'GTCCAAGA3'), was used to amplify the aphid's DNA in a twin-block thermocycler (Ericomp).

Conclusions

The aphid DNAs were effectively amplified for 15 samples but failed to amplify the samples 18 and 19 (Figure 5-11). The two remaining samples showed incomplete amplification with several bands. Failure and incomplete amplification might be arisen from either indigenous contaminant which is present in some of the aphid samples or general experiment errors for DAF. Indeed, additional step for DNA purification and consideration for sample replications might be useful to overcome these failures.

DNA Amplification of Bermudagrass

Question. Does DAF differentiate "off-type" and commercial cultivars of bermudagrasses?

Bermudagrass (*Cynadon* spp) is a warm-season grass that is widespread throughout the southern region of the USA. This type of grass consists of several species of which one is a sterile triploid (3n=27). Several cultivars have been developed from this triploid through clonal propagation and are of great economic importance in turf industry especially for golf courses. Recently, off-type patches have been observed in the golf courses causing sever problems and millions of dollars in loss. The purpose of using DAF was to provide a tool to distinguish off-types from standard turfgrass samples.

Samples from not only from different states in the USA but also from around the world,

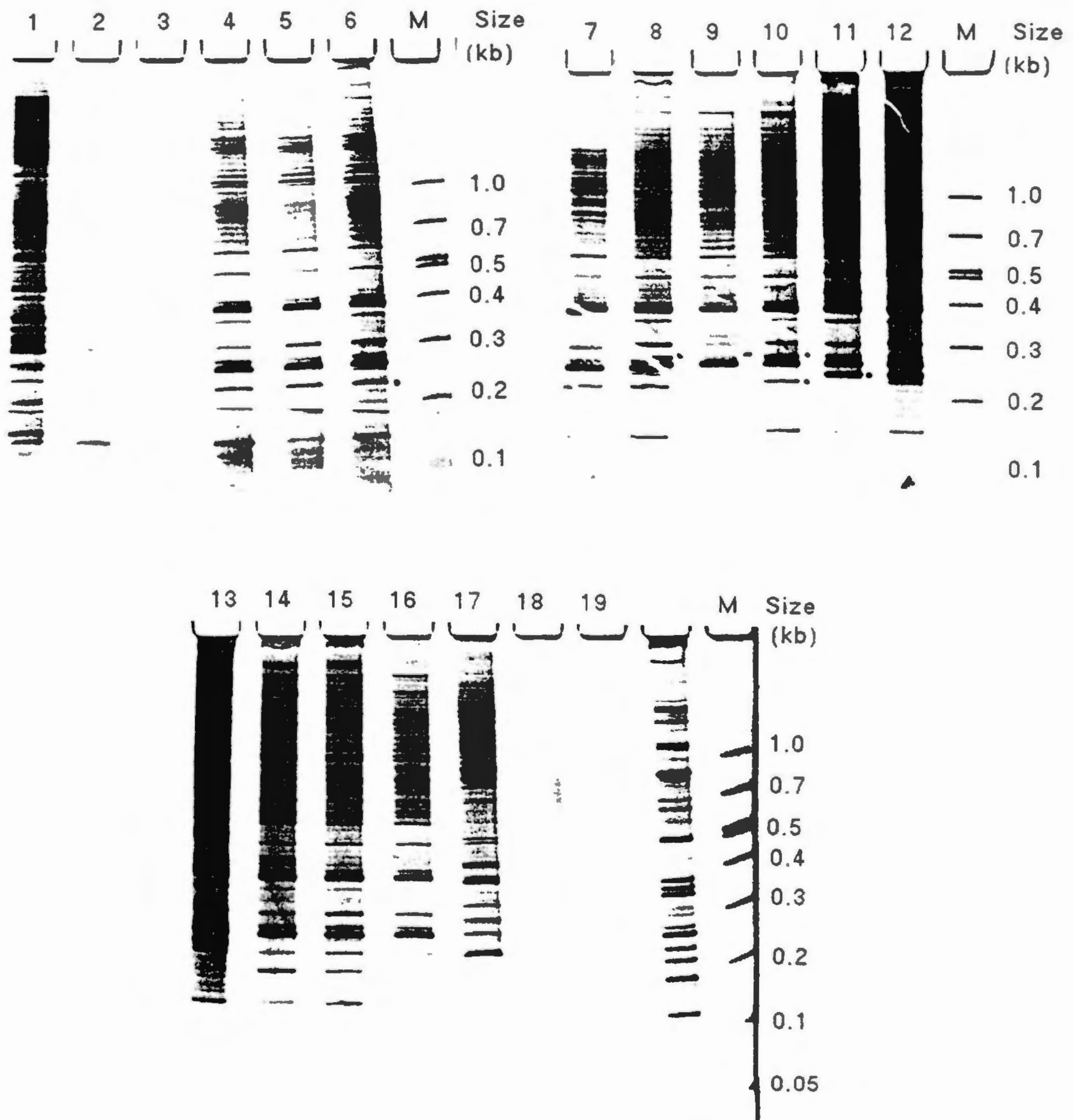


Figure 5-11. Silver-stained 10% polyacrylamide gels of DNA amplification fingerprinting using original DAF program aand OR30 (5'GTCCAAGA3') primer for 1-19 aphid samples. Molecuir marker ranged from 0.1 to 1.0 kb is shown on the right.

e.g., Australia, Philippines, and Singapore, were sent to our lab, Plant Molecular Genetic, to be analyzed. DNA was extracted from leaf samples using a modified protocol developed originally by Dellaporta *et al.* (1983). After twice dilution and subsequent quantification, DNA of off-types and standards were amplified using high annealing temperature program (DAF-15) in a MJ Research PT200 thermocycler. Standard samples, Tifway I(34), Tifway II (35), and Tifdwarf were obtained from Dr. E. Elsner, Athens, Georgia. Tifway II is a radiation-induced mutant of Tifway I and hence is closely related.

Here two examples are given to demonstrate applications of DAF in testing turfgrass samples. In first case, two samples A2 and B4 were contaminated with an off-type which had a course-leaf. DNA was extracted from course, fine, and a mix of course and fine leaves. DAF was conducted for these samples along with standard samples using an octamer primer 8-9 (5'GTTACGCC3'). Results of a 10% silver-stained polyacrylamide gel showed no differences between course and fine leaves (Figure 5-12A). Furthermore, DAF profiles were identical for A2, B4, and Tifway I indicating that either these two samples were the same as Tifway or the primer used in this study was not able to exclude these two samples. This primer could differentiate Tifway I and Tifway II by presence of a band (230 bp) which is absent in Tifway II (Figure 5-12A). Several polymorphic bands also were distinctive between Tifway I -Tifway II and Tifdwarf (Figure 5-12A).

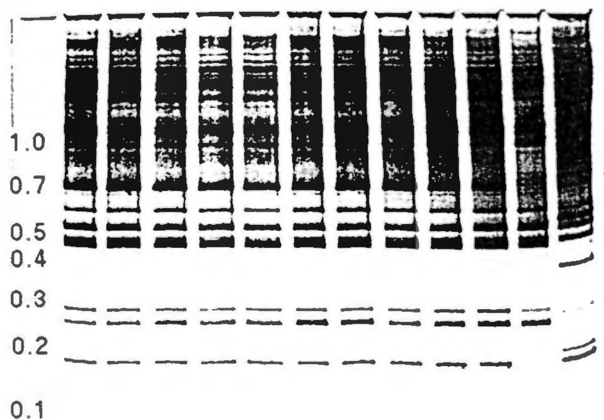
In second case, a sample (MB) was analyzed to determine whether it was a Tifway I. The sample and the standard DNAs were amplified in replications using a mini-hairpin primer Hp-10 (5'GCGAAGCCTG3'). Results of polyacrylamide gel revealed that MB DAF profiles were different from Tifway I in a polymorphic band at 430 bp. Another distinctive band which amplified with high intensity in MB was observed at 220 bp (5-12B).

Conclusions

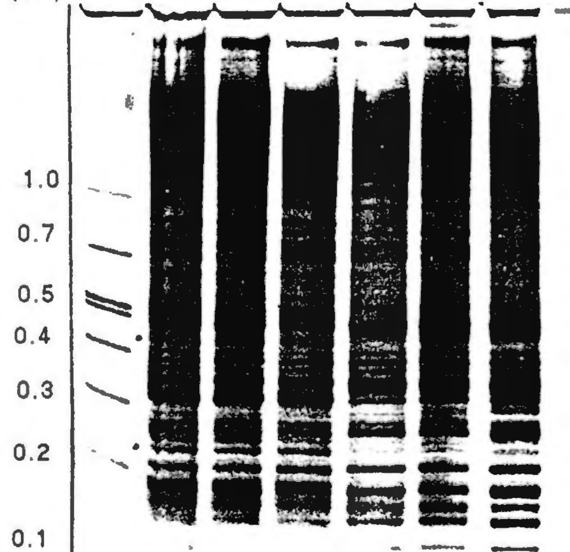
DAF was used to provide informative DNA fingerprinting for the off-type and commercial cultivars. Each sample was amplified and run in replication to avoid experimental errors. Newly-developed high annealing temperature at 55°C provided a robust and reliable DAF profiles in this study.

Figure 5-12. Silver-stained 10% polyacrylamide gels of DAF for unknown bermudagrass and standard cultivars using 55°C annealing temperature DAF and an octamer primer 8-9, 5'GTTACGCC3', (Panel A) and a mini-hairpin primer Hp10, 5'GCGAAGCCTG3', (Panel B). A2, B4, and MB are unknown samples. Standard samples are TI (Tifway I), TII (Tifway II), and Td (Tifdwarf). Mix is the mixture of course and fine leaves found in the both A2 and B4 samples. This mixture was used for DNA extraction and amplification. A'2, B'4, T'I, and MB' are duplication and MB" is the triplication of amplification. Moleculr marker ranged from 0.1 to 1.0 kb is shown on the right.

A



B



Genetic Stability in Centipedegrasses

Question. Can DAF detect mutation induced under different environmental conditions?

To determine genetic stability of different centipedegrasses including Tennessee Hardy (TH) which has tolerated extreme cold temperature of -17 to -31. Vegetative propagates of this cultivar and another cultivar, Oklawn (OK), were transplanted from 300 m to 600 m elevation and grown for 30 years. DNA fingerprints for these two and other cultivars have been previously studied (Weaver *et al.*, 1995).

In this study, DNA was extracted from Tennessee Hardy grown in Knoxville (THK), Tennessee Hardy grown in plateau (THP), Oklawn grown in Knoxville (OKK), Oklawn grown in plateau (OKP), Common, Centennial (Cent), Tennessee Tuff (TTf), A-320, A-321, A-322. DNA samples were amplified using two octamer primers 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3').in an oven thermocycler (Bios, New Haven, CT) for 35 cycles of 20 seconds at 96°C and 20 seconds at 30°C. Amplification products were separated in 10% polyacrylamide silver-stained gels

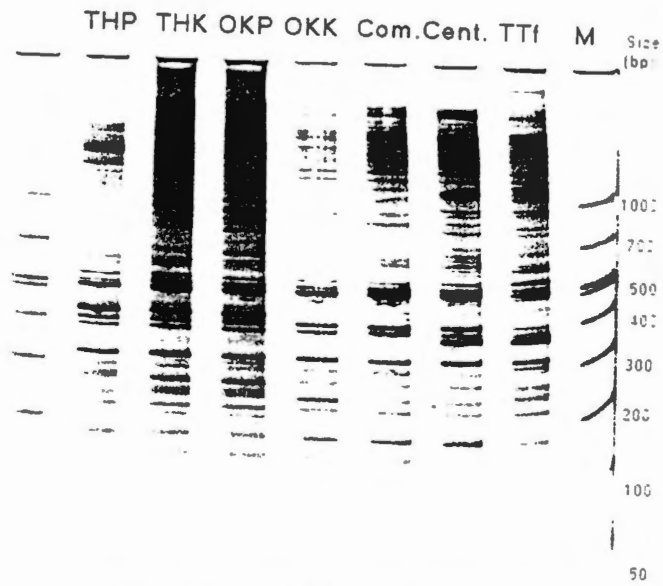
DAF profiles of TH from both elevations (THK and THP) showed no difference when primer 8-8 was used (Figure 5-13A). However, DAF profiles for OK amplified with the same primer exhibited a change in DNA fingerprint profiles (Figure 5-13A) indicating that natural selection was in process. Cultivars Com. Cent., and TTF were differentiated from each other using primer 8-8 (Figure 5-13A). Using primer 8-4 , all cultivar DNA fingerprints showed different banding patterns (Figure 5-13B).

Conclusions

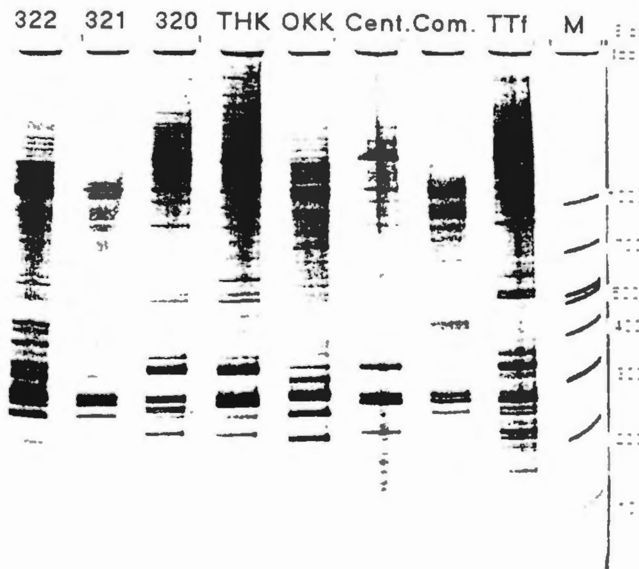
The two cultivars TH and OK grown at different elevation for 30 years exhibited different results indicating different genetic stability rates in these two cultivars. It is possible that THK and THP are also under natural selection but the primers used in this study were not able to differentiate THP and THK.

Figure 5-13. Silver-stained 10% polyacrylamide gels of DNA amplification fingerprinting using 55°C annealing temperature DAF and primer 8-8, 5'GAAACGCC3', (Panel A) and 8-4, 5'GTAACGCC3', (Panel B) for centipedegrass samples, THP (Tennessee Hardy grown in plateau) THK (Tennessee Hardy grown in Knoxville), OKP (Oklawn grown in plateau), OKK (Oklawn grown in Knoxville, Com. (Common), Cent. (Centennial), TTF (Tennessee Tuff), A-320, A-321, A-322.

A



B



Concluding Remarks

Bulked segregant analysis (BSA) was developed as a rapid approach for tagging a gene using molecular markers (RFLP, RAPD, DAF, and AFLP), and for saturating previously mapped regions with new molecular markers, and mapping a gene by targeting a specific region and anchoring it to a detailed genetic linkage map.

Prior to introduction of the BSA, near-isogenic lines (NIL) differing in the target gene was a method of choice to target molecular markers to a region (Young *et al.*, 1988). Generating NILs is time-consuming particularly in species with slow maturation and tedious backcrossing. In contrast, BSA allows efficient mapping without need for several crosses and screening of many individuals to find a linked marker to the target locus. The efficiency of BSA can be determined by the frequency at which linked markers are identified.

I used BSA to detect DAF markers linked to the *enod2* gene and provide additional markers in the genomic region. Although the ratio of polymorphic markers (0.1) detected in this study were far below the average (1.5) found in the other studies, four DAF markers were detected and found to be reproducible in different amplifications. These markers can be either mapped directly in the population or mapped after being cloned, sequenced and converted to SCAR markers.

I used DAF successfully and effectively for DNA fingerprinting of SCN, *Mycorrhizae*, aphid, centipedegrass, and bermudagrass samples. This success indicates that DAF is a ubiquitous technique which can be used for any kind of organism and any size of genome.

References

Barua, U. M., Chalmers, K. J., Hackett, C. A., Thomas, W. T., Powell, W. and Waugh, R. 1993. Identification of RAPD markers linked to a *Rhynchosporium secalis* resistance locus in barley using near isogenic lines and bulked segregant analysis. Heredity 71: 177-184.

Caetano-Anollés, G., Gresshoff, P. M., Elkan, G. H., and Upchurch, R. G. 1996. Advances in the positional cloning of nodulation genes in soybean. Plant and Soil 186: 1-7.

Cervera, M. T., Gusmao, J., Steenackers, M., Storme, V., Broeck, A., vanden, Montagu, M. van, Boerjan, W., Vanden-Broeck, A., Van Montagu, M., and Neale, D. B. 1996. The use of bulked segregant analysis to identify AFLP markers closely linked to *Melaspora larici-populina* resistance in *Populus*. In " Somatic Cell Genetics and Molecular Genetics of Trees" (eds. Ahuja, M. R. and Boerjan, W.). Kluwer Academic Publishers, Dordrecht, Netherlands pp. 205-210.

Chunwongse, J., Doganlar, S., Crossman, C., Jiang, J., and Tanksley, S. D. 1997. High-resolution genetic map of the *Lv* resistance locus in tomato. Theor. Appl. Genet. 95: 220-223.

Dellaporta, S. J., Wood, J., and Hicks, J. B. 1983. A plant DNA mini-preparation. Version II, Plant Mol. Biol. Reporter 1: 19-21.

Eskew, D. L., Caetano-Anollés, G., Bassam, B. J., and Gresshoff, P. M. 1993. DNA amplification fingerprinting of the *Azolla-Anabaena* symbiosis. Plant Mol. Biol. 21: 363-373.

Fabritius, A. L., Shottock, R. C., and Judelson, H. S. 1997. Genetic analysis of metalaxyl insensitivity loci in *Phytophthora infestans* using linked DNA markers. Phytopath. 87: 1034-1040.

Foisset, N., Delourme, R., Barret, P., and Renard, M. 1995. Molecular tagging of the dwarf BREIZH (*Bzh*) gene in *Brassica napus*. Theor Appl. Genet. 91: 756-761.

Garcia, G. M., Stalker, H. T., Shroeder, E., and Kochert, G. 1996. Identification of RAPD, SCAR, and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* into *Arachis hypogaea*. Genome 39: 836-845.

Giorio, G., Gallitelli, M., and Carriero, F. 1997. Molecular markers linked to *rhizomania* resistance in sugar beet, *Beta vulgaris*, from two different sources map to the same linkage group. Plant Breeding 116: 401-408.

Gmitter, F. G., Jr., Xiao, S. Y., Huang, S., Hu, X. L., Garnsey, S. M., and Deng, Z. 1996. A localized linkage map of the citrus tristeza virus resistance gene region. Theor. Appl. Genet. 92: 688-695.

Grattapaglia, D. Bertolucci, F. L. G., Penchel, R., and Sederoff, R. R. 1996. Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. Genetics 144: 1205-1214.

Gresshoff, P. M., Ghassemi, F., Brewer, R. A., and O'Neill, E. G. 1998. DNA amplification fingerprinting of mycorrhizal fungi and associated plant materials using arbitrary primers. In: "*Mycorrhizae* Manual: Structure, Function, and Molecular Biology", (eds. Varma and Hock), 2nd edition. Springer Verlag, Berlin, pp. 499-513.

Hermat, M., Weeden, N. F., Conner, P. J., and Brown, S. K. 1997. A DNA marker for columnar growth habit in apple contains a simple sequence repeat. J. Hort. Sci. 122: 347-349.

Haymes, K. M., Heneken, B., Davis, T. M., and van de Weg, W. E. 1997. Identification of RAPD markers linked to *Phytophtora fragariae* resistance gene (*Rpf1*) in the cultivated strawberry. Theor. Appl. Genet. 94: 1097- 1101.

Hamalainen, J. H., Watanebe, K. N., Valkonen, J. P. T., Arihara, A., Plaisted, R. L., Pehu, E., Miller, L., and Slack, S. A. 1997. Mapping and marker-assisted selection for a gene for extreme resistance to potato virus Y. Theor. Appl. Genet. 94: 192-197.

Jiang, C. X. and Sink, K. C. 1997. RAPD and SCAR markers linked to the sex expression locus *M* in asparagus. Euphytica 94: 329-333.

Jong, de, W., Forsyth, A., Leister, D., Gebhardt, C., Baulcombe, D. C., and De Jong, W. 1997. A potato hypersensitive resistance gene against potato virus X maps to a resistance gene cluster on chromosome 5. Theor. Appl. Genet. 95: 246-252.

Kolchinsky, A., Landau-Ellis, D., and Gresshoff, P. M. 1997. Map order and linkage distance of molecular markers close to the supernodulation (*nts-1*) locus of soybean. Mol. Gen. Genet. 245: 29-36.

Koller, B., Gianfranceschi, L., Segalias, N., McDermatt, J., and Gessler, C. 1994. DNA markers linked to *Malus floribunda* 821 scab resistance. Plant Mol. Biol. 26: 596-602.

Landau-Ellis, D., Angermüller, S. A., Shoemaker, R., and Gresshoff, P. M. 1991. The genetic locus controlling supernodulation in soybean (*Glycine max* L.) co-segregates tightly with a cloned molecular marker. Mol. Gen. Genet. 228: 221-226.

McCallum, J., Timmerman-Vaughan, G., Frew, T., and Russell, A. 1997. Biochemical and genetic linkage analysis of green seed color in field pea. J. Hort. Sci. 122: 218-225.

Men, A. E., Borisov, A. Y., Rozov, S. M., Ushakov, K. V., Tsyganov, V. E., Tikhanovich, I. A., and Gresshoff, P. M. 1998. Identification of DNA amplification fingerprinting (DAF) markers close to the symbiosis-ineffective *sym31* mutation of pea (*Pisum sativum*). Theor. Appl. Genet. (submitted).

Michelmore, R. W., Paran, I., and Kesseli, R. V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating population. Proc. Natl. Acad. Sci. USA 88: 9828-9832.

Miklas, P. N., Johnson, E., Stone, V., Beaver,. J. S., Montoya, C. Zapata, M. 1996. Selective mapping of QTL conditioning disease resistance in common bean. Crop Sci. 36: 1344-1351.

Milach, S. C. K., Rines, H. W., and Phillips, R. L. 1997. Molecular genetic mapping of dwarfing genes in oat. Theor. Appl. Genet. 95: 783-790.

Ruyter, S., Carolien, P., Gu, Z. L., and van der Poel, J. J. 1997. Bulked segregant analysis using microsatellites: mapping of the dominant white locus in the chicken. Poultry Sci. 76: 386-391.

Voss, S. R. and Shaffer, H. B. 1996. What insight into the developmental traits of urodeles does the study of interspecific hybrids provide? Int. J. Dev. Biol. 40: 885-893.

Weaver, K., and Callahan, L. M. 1994. Isolation and cloning of DNA amplification products from silver-stained polyacrylamide gels. BioTechniques 16: 226-227.

Weaver, K., Callahan, L. M., Caetano-Anollés, G., and Gresshoff, P. M.,1995. DNA amplification fingerprinting and hybridization analysis of centipedegrass. Crop Sci. 35: 881-885.

Williams, H. M., Hoisington, D., Singh, R. P., and Gonzalez-de Lion, D. 1997. Detection of quantitative trait loci associated with leaf rust resistance in bread wheat. Genome 40: 253-260.

Yazdi-Samadi, B., Qian, D., Ghassemi, F., Allen, F. L., and Gresshoff, P. M. 1996. Genetics and molecular biology of resistance to soybean cyst nematode. Physiol. Mol. Biol. Plants 2:25-48.

Yoon, C., Glawe, D. A., and Shaw, P. D. 1991. A method for rapid small-scale preparation of fungal DNA. Mycologia 83: 835-838.

Young, N. D., Zamir, D., Ganal, M. W., and Tanksley, S. D. 1988. Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics 120: 579-585.

Chapter 6

Conclusions

In this dissertation I have tried to show how molecular markers such as RFLP, microsatellite, and DAF can facilitate mapping a gene of interest and saturating a region of genome around the same gene with molecular markers as well as applications of DAF markers in different organisms. The development of molecular marker technology has greatly provided researchers working in plant genetics and evolution with an immense supply of genetic markers which enhances construction of genetic linkage maps and marker-assisted breeding programs. The plant and animal breeders have the ability to tag any trait or region of the genome. This has been accelerated by development of nucleic acid automation as well as rapid isolation procedures of nucleic acids at very early stage of growth or even from seeds in case of plant.

In this study, two expressed sequence tags (EST), the *enod2* and *lbc3* genes, involved in early nodulation and subsequent nitrogen fixation in soybean, respectively, were mapped in recombinant inbred lines (RILs) as well as an F_2 segregating population of soybean *Glycine max* (L. Merr). The *enod2* map location was near the seed coat color gene *I* locus with a distance of 11.5 cM on linkage group U3 of the RIL map (Utah) and 18.5 cM on the *G. max* x *G. soja* map (TN). *Enod2* and *I* are located close to *Rhg4* , a soybean cyst nematode (SCN) resistance gene, and a locus for seed coat hardness. The molecular marker pA110 and seed coat color were used to integrate *enod2* on an F_2 segregating population generated from a cross between cultivar Bragg and *G. soja* (Sieb and Zucc), PI468.397. *Enod2* was mapped in the same order as on the RIL map but 18.5 cM from the I locus on the TN map. A microsatellite from the 5' region of *enod2B* was mapped in the same position, demonstrating that *enod2B* and not *enod2A* was mapped. An RFLP for *lbc3* (leghemoglobin) segregated independently from *enod2* and the *nts*-1 supernodulating locus suggesting that in soybean, symbiotically significant loci (including rj_1, Rj_2, and rj_6) are not clustered in soybean.

To saturate the *enod2* region of the soybean genome, DAF marker was used to detect linked markers to the *enod2* gene. However, due to mismatch priming and secondary DNA structure in original low annealing temperature DAF, a high annealing temperature of 55°C was developed to overcome these potential problems. Fifteen PCR programs differing in annealing temperature (47, 55, and 60°C), denaturation, annealing, and extension time (30, 60, and 120 second), and presence/absence of extension step (+/- 72°C) were tested. These programs were tested for three arbitrary octamer and two soybean cultivars, Minsoy and Noir1. The number of scorable bands (amplification

products) after 10% PAGE and silver staining ranged from 7 to 51. The average ramping temperatures for heating and cooling were calculated 1.42 and 1.27 seconds/°C, respectively. Program 15 (95°C/30 sec, 55°C/120, and 72°C/30 sec) generated a complex DNA fingerprinting profiles for tested primers in Minsoy and Noir1. These profiles contained an average of 42 sharp and highly intense bands using both octamer primers 8-4 (5'GTAACGCC3') and 8-8 (5'GAAACGCC3') for DNA amplification. Using high annealing temperature increased stringency of primer-template annealing, avoided potential mismatching and hybrid molecule formation, and consequently improved reproducibility of DNA fingerprinting.

I used this high annealing temperature DAF and bulk segregant analysis (BSA) approach to detect DAF markers linked to the *enod2* gene and provide additional markers in the genomic region. Although the ratio of polymorphic markers (0.1) detected in this study were far below the average (1.5) found in the other studies, four DAF markers were detected and found to be reproducible in different replications of amplifications. These markers can be either mapped directly in the population or mapped after being cloned, sequenced and converted to SCAR markers. Furthermore, I used DAF successfully for DNA fingerprinting of soybean cyst nematode (SCN), *Mychorrizae*, aphid, centipedegrass, and bermudagrass samples. This success indicates that DAF is a ubiquitous technique which can be used for any kind of organism and any size of genome.

Future plan for the polymorphic markers detected for *enod2* using BSA approach would be reamplification of bands, cloning, sequencing, and development of SCAR markers for the *enod2* region. These potential SCARs might be a useful tool for anchoring BAC and YAC clones containing the *enod2* region and ultimately isolation of interesting genes, e. g. race 3 nematode resistance gene, which are located in this region of soybean genome

Molecular marker techniques such as AFLP, RAPD, AP-PCR, and DAF can detect a large number of polymorphic markers using a quick and relatively simple laboratory procedures which is directly accessible to the researchers in different disciplines. The speed of detection has been recently accelerated by development of automation and oligonucleotide arrays, also called DNA chips.

Fabricated arrays of oligonucleotides or even longer DNA probes can be used to screen rapidly a large number of samples including genomic DNA, sequence tag sites (STSs), expressed sequence tags (ESTs), and mRNA. A bulk of labeled cDNA or mRNA can be hybridized to an array of arbitrary probes to generate a hybridization fingerprint in which each mRNA may bind to few specific site within the array, called differential display (DD) on a chip. Fingerprinting and expression of mRNA can be studied directly by immobilizing probes corresponding to known gene sequence. It is possible to prepare a series of genosensors containing unique probes for each coding sequence. For example, in near future it is possible to represent all 100,000 human genes in a series of genosensors to analyze gene expression.

Hybridization fingerprints are digitally interpreted by computer which allows a rapid access to comprehensive sequence information. Recently, a new flowthrough sensor consisting of microchannel glass or porous silicon has been used. In this design, hybridization occurs in a three-dimensional volumes of silicon dioxide or channel array glass instead of 2-dimensional surface. Compared to flat surface, binding capacity per hybridization cell will increase about 100-fold causing improved detection sensitivity. Further improvement of flowthrough genosensor can still increase the speed of DNA fingerprinting and gene expression analysis.

Appendices

Appendix A

```
LOCUS       GMENOD2A     3072 bp    DNA             PLN       22-MAR-1995
DEFINITION  Soybean ENOD2A gene for Ngm-75.
ACCESSION   X16875
NID         g18575
KEYWORDS    ENOD2A gene; Ngm-75; nodulin.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Magnoliopsida; Rutanae; Sapindales; Fabaceae; Papilionoideae;
            Glycine.
REFERENCE   1  (bases 1 to 3072)
  AUTHORS   Franssen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-OCT-1989) Franssen H. J., Department of Molecular
            Biology, Agricultural University, Dreyenlaan 3, 6703 HA Wageningen,
            The Netherlands
REFERENCE   2  (bases 1 to 3072)
  AUTHORS   Franssen,H.J., Thompson,D.V., Idler,K., Kormelink,R., van Kammen,A.
            and Bisseling,T.
  TITLE     Nucleotide sequence of two soybean ENOD2 early nodulin genes
            encoding Ngm-75
  JOURNAL   Plant Mol. Biol. 14 (1), 103-106 (1990)
  MEDLINE   91322483
COMMENT     For early nodulin (N-75) mRNA,partial cds see .
FEATURES             Location/Qualifiers
     source          1..3072
                     /organism="Glycine max"
                     /strain="Wayne"
                     /db_xref="taxon:3847"
     misc_feature    1048..3049
                     /note="Homology region with GmENOD2B"
     misc_feature    1557..1571
                     /note="AGGA box"
     promoter        1604..1608
                     /note="TATA box"
     precursor_RNA   1630..2832
                     /note="transcript"
     sig_peptide     1653..1728
                     /note="put.signal peptide (AA -25 to -1)"
     CDS             1653..2582
                     /note="pre-pro polypeptide (AA -25 to 284)"
                     /codon_start=1
                     /db_xref="PID:g18576"
                     /db_xref="SWISS-PROT:P08297"
                     /translation="MTSVLHYSLLLLLLGVVILTTPVLANLKPRFFYEPPPIEKPPTY
                     EPPPFYKPPYYPPPVHHPPPEYQPPHEKTPPEYLPPPHEKPPPEYLPPHEKPPPEYQP
                     PHEKPPHENPPPEHQPPHEKPPEHQPPHEKPPPEYEPPHEKPPPEYQPPHEKPPPEYQ
                     PPHEKPPPEYQPPHEKPPPEHQPPHEKPPEHQPPHEKPPPEYQPPHEKPPPEYQPPQE
                     KPPHEKPPPEYQPPHEKPPPEHQPPHEKPPPVYPPPYEKPPPVYEPPYEKPPPVVYPP
                     PHEKPPIYEPPPLEKPPVYNPPPYGRYPPSKKN"
     mat_peptide     1729..2579
                     /note="mature peptide (AA 1-284)"
BASE COUNT     1020 a    731 c    405 g    916 t
ORIGIN
```

Figure A-1. Complete sequence of *enod2A* genomic clone.

```
   1 ggatccttac acaggccaga catccccaag ttctcaaata agacaaattt ggttgttctt
  61 ttcttaatat ttcacaggga gatgttctgt cttttgattt ggggatttca tttagcacat
 121 aacaaacagt taacaaaatt tcgccccacc aaaaagatgt tgcactagaa ctcaacatag
 181 tagctacaac taattctgta aaagttctgt tctttctttc agctttaccg ttcatttcag
 241 gtgaatatgg agcagttgtt tcatgtatga ttccatgcaa attataaaac tcattaaaca
 301 aactggaatc atactctgtg cctctatcac ttcgaagttt cttaattttc ttattgaatt
 361 gattttcaat ttctgttaca aataacttaa acatgtcaag cgcttcactt ttatttttca
 421 taagatatac atatatatat aatcagagca gtcatcaata aaagtgataa aatatcgttt
 481 tccatttctg gtcaacgttc catcaaattc acatatatca gaatgtatta aatccaatgg
 541 ctcagattct ttaactactg attttgtga tttttagtt attttagatt gactgcaaaa
 601 aacacacttt tcaaagtgat ttgaagatag ctttggaata aaacctaagt tactcatatt
 661 agatatgcaa cgactattta tatgacaaag tctagaatgc cagaattaaa atcacacagc
 721 atgtaagcag aaggagaaac tttattaata tcaagattca atttgaacat gccatcagtg
 781 gcgtaccctt tccctacaaa taccccattc ttggtcaaag taaataaatc tgcacctatg
 841 gtctgagtaa acccagcctt gtttaaaaga aaaccagaaa ccagattctt tctcatctct
 901 ggagtatgca tcacatcttt gagaatcaaa gtctttccag aggtaaactt cagttcaaca
 961 tctccagttc tagcaacagt agtggtgtgg gaatcaccca acaacacttt cttattttca
1021 acattcgtgt atgttttaaa catagcatga tctttatact tgtatttttt ttttgtttta

1081 gtttctatac ttaaaaattc tgttttatta tttttacgcc ttagttttct agcaatctaa
1141 aactgatata aaatagaagt ataacgacta aaaçataaaa aaaaaaaatt gtataaaaaa
1201 taaagcatat agctttcatt catatataag aactaaactg aaataccagt gtaagtataa
1261 gaactaatcg ataaattaag ccaaattaag ggtacatatt atttttaaga aaattaggcc
1321 gggtatatat ttttaaaaag gactatacac tatgtgacga tagaaataat aggtatgtag
1381 atgtatgtta agtattttct aatgtgtttt ttactttctc tatcacactt gttattttct
1441 cactattttt ttctcttgtt tctctgttat tttcactcta aaactggagt aatatgttta
1501 tgactacaac acattttgac atgacttagg attaacatat attatgataa aataactaaa
1561 gattgataac cttgatagaa aagcttctca tgtctcctct ccctataagt agtttcccat
1621 tgttatcact tttcatcagc acaagctaag acatgacttc tgtactacac tactcactcc
1681 tgctgctcct gcttggagtg gtgattctca ccactccagt gctagctaat ttgaagccac
1741 gcttcttcta tgagcctcct ccaattgaga aaccccccac ctatgaacct ccaccatttt
1801 ataagccccc atactaccca ccaccagtgc accaccctcc accagagtac caaccacccc
1861 atgaaaaaac accacctgag tatctacctc ctcctcatga gaaaccacca ccagaatacc
1921 tacctcctca tgagaaaccg ccaccagaat accaacctcc tcatgagaaa ccaccccatg
1981 agaatccacc accggagcac caaccacctc atgagaagcc accagagcac caaccacctc
2041 atgagaagcc accaccagag tatgaaccac ctcatgagaa accaccacca gaataccaac
2101 cacctcatga gaagccacca ccagaatacc aaccacctca tgagaaacca ccaccagaat
2161 accaaccacc tcatgagaag ccaccaccag agcaccaacc acctcatgag aagccaccag
2221 agcaccagcc acctcatgag aagccaccac cagagtatca accacctcat gagaaaccac
2281 caccagaata ccaacctcct caagaaaagc caccacatga aaaaccaccg ccagaatacc
2341 aacctcctca tgaaaagcca ccaccagaac accaacctcc ccatgaaaag ccaccaccag
2401 tgtacccacc cccttatgag aaaccaccac cagtgtatga accccccttat gagaagccac
2461 ccccagtagt gtatccacct cctcatgaga aaccacccat ttatgagcca ccgccattgg
2521 agaagccacc ggtctacaat cccccacctt atggccgcta tccaccatcc aagaaaaact
2581 aataaccact tgcctgcgtc acatgttttg gtctactcaa acttagacct gccctttgtc
2641 atataaagct tttgtttct gtttaagatc tcaagtacaa tatgtccctt ctgcatgcac
2701 tacttcttca aaataaaggc tttatgccta tgtataatac tctactttaa ttctcctttc
2761 accatcgata ttgtaatgtc aactactagt gtgggtttat ctatggctat aataagtttt
2821 tctttgtgtt tacttatgaa tctttgtttt taattgcatg ctaaaaattg gcaaaaacat
2881 atataattct gttcgtacat gttttatttt atgaacttca taagtaccgg taaagcaatg
2941 ataatgtgta aagttgcttg gtctatatat atgtttaaat acacatatct ctaaaccgtc
3001 aatgagaaat actctctgta cctgtttatt caacttggaa aactaaacca cataataaac
3061 caattattaa tc
```

Figure A-1 (continued)

```
LOCUS       GMENOD2B     4584 bp    DNA             PLN       17-FEB-1997
DEFINITION  Soybean ENOD2B gene for Ngm-75.
ACCESSION   X16876
NID         g18578
KEYWORDS    ENOD2B gene; Ngm-75; nodulin.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Magnoliopsida; Rutanae; Sapindales; Fabaceae; Papilionoideae;
            Glycine.
REFERENCE   1  (bases 1 to 2992)
  AUTHORS   Franssen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-OCT-1989) Franssen H. J., Department of Molecular
            Biology, Agricultural University, Dreyenlaan 3, 6703 HA Wageningen,
            The Netherlands
  REMARK    revised by [3]
REFERENCE   2  (bases 1 to 2992)
  AUTHORS   Franssen,H.J., Thompson,D.V., Idler,K., Kormelink,R., van Kammen,A.
            and Bisseling,T.
  TITLE     Nucleotide sequence of two soybean ENOD2 early nodulin genes
            encoding Ngm-75
  JOURNAL   Plant Mol. Biol. 14 (1), 103-106 (1990)
  MEDLINE   91322483
REFERENCE   3  (bases 1 to 4584)
  AUTHORS   Lauridsen,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-1992) Lauridsen, P., Institution Dept. of
            Molecular Biology, University of Aarhus, Gustav Wieds Vej 10,
            DK-8000 Aarhus C, Denmark
COMMENT     For early nodulin (N-75) mRNA,partial cds see .
FEATURES             Location/Qualifiers
     source          1..4584
                     /organism="Glycine max"
                     /strain="Wayne"
                     /db_xref="taxon:3847"
     old_sequence    1..1592
                     /citation=[1]
                     /replace=""
     misc_signal     3088..3101
                     /note="AGGA box"
     TATA_signal     3134..3138
     prim_transcript 3160..4362
                     /gene="ENOD2B"
     gene            3160..4362
                     /gene="ENOD2B"
     CDS             3183..4112
                     /gene="ENOD2B"
                     /codon_start=1
                     /product="Ngm-75, nodulin"
                     /db_xref="PID:g18579"
                     /db_xref="SWISS-PROT:P08297"
                     /translation="MTSVLHYSLLLLLLGVVILTTPVLANLKPRFFYEPPPIEKPPTY
                     EPPPFYKPPYYPPPVHHPPPEYQPPHEKTPPEYLPPPHEKPPPEYLPPHEKPPPEYQP
                     PHEKPPHENPPPEHQPPHEKPPEHQPPHEKPPPEYEPPHEKPPPEYQPPHEKPPPEYQ
                     PPHEKPPPEYQPPHEKPPPEHQPPHEKPPEHQPPHEKPPPEYQPPHEKPPPEYQPPQE
                     KPPHEKPPPEYQPPHEKPPPEHQPPHEKPPPVYPPPYEKPPPVYEPPYEKPPPVVYPP
                     PHEKPPIYEPPPLEKPPVYNPPPYGRYPPSKKN"
     sig_peptide     3183..3257
                     /gene="ENOD2B"
     mat_peptide     3258..4109
                     /gene="ENOD2B"
                     /product="Ngm-75, nodulin"
BASE COUNT     1623 a    835 c    573 g   1553 t
ORIGIN
        1 aagcttgaca aaagataaat gctttgtggg gtggcgtagc gtctttatgc agcaatggtt
       61 tatgtaattt atgtaatggg gtggtcactc ctagtgactg tcctctgtgt tatgattaat
      121 gaaatgtttt gctttttcga aaagaacaaa aaatccttaa gttcacccca tttgtaaata
      181 gtctcttaca ttgaattggg gttgaattat taaagaagaa atctcaacta cttatttatt
      241 ttaaatttca atcatttatt agtttaattt ttataaatca cttttctaaa tattaaaata
      301 taataaaact cttctaaaaa cataataaaa ttaataacta aaataaataa attatttttt
      361 attggtattt atttttgttt tttttttttc taaattcata ttcttttact tatgttttaa
```

Figure A-2. Complete sequence of *enod2B* genomic clone.

```
 421 tagacaaaaa ctgatttgta acaaaagatt aataaatgat tgacgtttaa gtttggtggt
 481 cagaatattt atttagtgtg tatatttgta aaaaggggta aagttcaaag tggttgagca
 541 ctttttttgtc ttaaaaagtc tcttatggga cataacaaac caaatgggat aaatttacat
 601 tttttttat ttatatagta ttttctttta gagacaaatt attcgaggca ctccatacat
 661 tgagtgcgaa tattttttc atccaaaatt ttaactttgg ttaatcacat tataaaatat
 721 tagccttttt tattaaacta aacttccgtt agtaggacaa tttttccttt gttaaaacaa
 781 caaaatttag gcaaaagact aaaagagtaa ttcaaccacc ttaaaataca gaataatcgt
 841 cattaaataa aattaatagt aatacatata tatatatata tatatatata tatatatata
 901 attattaagt tgtttaaatt tcaacttttc atttcaatca aaatgacatg cataatatat
 961 gttaatcaga aaattaataa tttaaaatat aataaattca ttcacatata taatattaat
1021 taaaatatta aaatacatct taacaactta ataatttatt atatgcgtaa ttgtgacaat
1081 ttcaagcaat taatgttatt atttttaatt aataattaaa gttttattgt catacattgt
1141 tagcctccat gtataagaaa actattcctt taacctataa aaaactatca ttaaaatatt
1201 ttttaagata attattataa aaatcaacaa acttattaat aatatatgat tcaataataa
1261 tatataaaat ctttgcatct aacataaatt ataataatat tacaattttt tcctttaaat
1321 caattttaca ttttaaaaaa tcaaattaaa ttcatatccg actattgctg cgcatgatag
1381 gctctaaaag accatcccat tcagatatta atatcttatt caacgttaat ctgtgttctg
1441 ttagattcca aagattccag tgaatagtga tggctaagaa cagtttcttg acctttcgct
1501 aacaagcaag cctacctata caagctccaa ttattttctt ttttgaggat tgctccattt
1561 attttcaaaa gattagttaa catatcactt ctccgacaaa acatacatgc atctaaatgt
1621 ggcagcatgc taaagttttg gtgaggctat agtaaaatat gaaataaaga tttgaagttt
1681 cagcccaata taaaaaaaaa attaattcct tctgaaatga aaagagtatc aaagaagata
1741 taatcagtaa aatctttttc ataagcattg atctggatac atcaactttg atgcgttgga
1801 aatactgtgc tcaagtttga cagcaattct tggaattttt tcgccacaac agaagctcca
1861 gacgattatg atttatgacc ttatatgatg ttagttacgt gaaagtaatt agaatcgcat
1921 ttgctaacta ttagcaattt tttttttaa gctaatgcaa gtgacagaat cttaggtctc
1981 tataatttga acctgtggcg gtggaactcg tacttcatgt gctgaaaaga acttgatatt
2041 tttttaaggg aaataatata tatcaatgct cctaagtcct aaactttatc ttctttggca
2101 gctaaattta ctttaaaaag aaataagatt aaataacttt ttcttacaag aaaatatatt
2161 taattattaa ttgttaagtt taacgtcttt ttatacattt atttgtttta aattccagtc
2221 atcttttttaa cataattcca atcatttatt agttttactt ttataaacaa taaaacataa
2281 ttaattttca gattaaaaaa tagatagaag tttttaatt gttttttatt atcaaatttc
2341 aattttaaca tattttataa tagataaaat gaattgtaac aaattaatga ttgaccttat
2401 agataagtaa tttagccaac aactttttta gtattaaatt gatagaaaaa ttaagctata
2461 tttggggggg gggggggtca agtttaatga agttaaagtt cattgaatat atttgtaaaa
2521 aaagataaag ggtttaaggt ctaatagaga taatatttaa ggacttaatt aattatttga
2581 tctttatact tgtatttttt ttttgtttta gtttctatac ttaaaaattc tgttttatta
2641 tttttacgcc ttagttttct agcaatctaa aactgatata aaatagaagt ataacgacta
2701 aaacataaaa aaaaaaaatt gtataaaaaa taaagcatat agctttcatt catatataag
2761 aactaaactg aaataccagt gtaagtataa gaactaatcg ataaattaag ccaaattaag
2821 ggtacatatt atttttaaga aaattaggcc gggtatatat ttttaaaaag gactatacac
2881 tatgtgacga tagaaataat aggtatgtag atgtatgtta agtattttct aatgtgtttt
2941 ttactttctc tatcacactt gttattttct cactattttt ttctcttgtt tctctgttat
3001 tttcactcta aaactggagt aatatgttta tgactacaac acattttgac atgacttagg
3061 attaacatat attatgataa aataactaaa gattgataac cttgataaag aagcttctca
3121 tgtctcctct ccctataagt agtttcccat tgttatcact tttcatcagc acaagctaag
3181 acatgacttc tgtactacac tactcactcc tgctgctcct gcttggagtg gtgattctca
3241 ccactccagt gctagctaat ttgaagccac gcttcttcta tgagcctcct ccaattgaga
3301 aacccccccac ctatgaacct ccaccatttt ataagccccc atactaccca ccaccagtgc
3361 accaccctcc accagagtac caaccacccc atgaaaaaac accacctgag tatctacctc
3421 ctcctcatga gaaaccacca ccagaatacc tacctcctca tgagaaaccg ccaccagaat
3481 accaacctcc tcatgagaaa ccaccccatg agaatccacc accggagcac caaccacctc
3541 atgagaagcc accagagcac caaccacctc atgagaagcc accaccagag tatgaaccac
3601 ctcatgagaa accaccacca gaataccaac cacctcatga gaagccacca ccagaatacc
3661 aaccacctca tgagaaacca ccaccagaat accaaccacc tcatgagaag ccaccaccag
3721 agcaccaacc acctcatgag aagccaccag agcaccagcc acctcatgag aagccaccac
3781 cagagtatca accacctcat gagaaaccac caccagaata ccaacctcct caagaaaagc
3841 caccacatga aaaaccaccg ccagaatacc aacctcctca tgaaaagcca ccaccagaac
3901 accaacctcc ccatgaaaag ccaccaccag tgtacccacc cccttatgag aaaccaccac
3961 cagtgtatga acccccttat gagaagccac ccccagtagt gtatccacct cctcatgaga
4021 aaccacccat ttatgagcca ccgccattgg agaagccacc ggtctacaat cccccacctt
4081 atggccgcta tccaccatcc aagaaaaact aataaccact tgcctgcgtc acatgttttg
4141 gtctactcaa acttagacct gccctttgtc atataaagct ttctgtttct gtttaagatc
4201 tcaagtacaa tatgtccctt ctgcatgcac tacttcttca aaataaaggc tttatgccta
4261 tgtataatac tctactttaa ttctcctttc accatcgata ttgtaatgtc aactactagt
4321 gtgggtttat ctatggctat aataagtttt tctttgtgtt tacttatgag tctttgtttt
4381 taattgcatg ctaaaaattg gcaaaaacat atataattct gttcgtacat gttttatttt
4441 atgaacttca taagtaccgg taaagcaatg ataatgtgta aagttgcttg gtctatatat
4501 atgtttaaat acacatatct ctaaacctgt caatgagaaa tactctcttg taccttgttt
4561 attcaacttg ggagactaaa ccta
```

Figure A-2 (continued)

Appendix B

Table B-1. Segregation patterns for *enod2*, seed coat color, and molecular marker pA110 in a soybean F_2 population generated from a cross between *Glycine max*, cultivar Bragg and the ancestral soybean *G. soja*, PI468.397. B, S, and H are patterns identical to Bragg, *G. soja*, and heterozygous patterns. Bl, Y, and V are black, yellow, and varigated seed coat colors, respectively.

Plant #	*enod2*	Seed Coat Color	pA110	Plant #	*enod2*	Seed Coat Color	pA110	Plant #	*enod2*	Seed Coat Color	pA110
01	H	V	H	02	S	Bl	-	03	B	V	H
04	H	V	H	05	S	Bl	H	06	H	V	-
07	S	Bl	-	08	H	Bl	S	09	H	V	H
10	-	V	B	11	S	Bl	H	12	H	V	S
13	B	V	H	14	H	V	H	15	B	V	-
16	H	V	S	17	H	V	B	18	H	V	H
19	S	Bl	H	20	H	V	S	21	B	V	-
22	H	V	-	23	H	V	S	24	H	V	H
25	H	V	-	26	S	Bl	S	27	H	V	-
28	H	V	-	29	B	V	-	30	H	V	-
31	B	V	B	32	H	V	H	33	B	Y	B
34	B	Y	B	35	H	V	H	36	B	Y	B
37	H	V	H	38	B	Y	H	39	H	Y	H
40	B	V	B	41	B	Y	B	42	H	V	-
43	S	Bl	H	44	B	Y	B	45	H	V	B
46	S	Bl	-	47	B	V	-	48	H	V	B
49	H	Bl	S	50	S	Bl	S	51	H	V	H
52	B	Y	B	53	H	V	B	54	H	V	-
55	B	Y	H	56	H	V	S	57	B	V	H

Table B-1 (continued)

Plant #	*enod2*	Seed Coat Color	pA110	Plant #	*enod2*	Seed Coat Color	pA110	Plant #	*enod2*	Seed Coat Color	pA110
58	H	V	S	59	B	V	-	60	-	Bl	H
61	S	Bl	S	62	S	Bl	S	63	B	V	H
64	H	Y	H	65	H	V	H	66	H	V	-
67	H	V	H	68	B	V	B	69	H	V	H
70	B	V	H	71	B	V	H	72	H	V	H
73	H	V	H	74	B	V	H	75	-	V	B
76	S	Bl	S	77	H	V	H	78	H	V	H
79	S	Bl	S	80	B	V	S	81	B	Y	B
82	H	V	H	83	S	Bl	H	84	H	V	H
85	H	Y	S	86	S	Bl	S	87	H	V	H
88	B	V	H	89	H	V	H	90	H	Y	H
91	H	V	-								

Appendix C

Table C-1. List of 10-, 9-, 8-, 7-, 6-, 5, 4, and 3-mer primers in the Plant Molecular Genetics, The University of Tennessee, Knoxville, TN.

Name	Sequence	Code	Other names
8-1	GAGCCTGT	RQJE	8.6a, BB5
8-2	CCTGTGAG	LQOI	8.6b, BB8
8-3	AACGGGTG	ONMB	8.6c, KS1591
8-4	GTAACGCC	IMBR	8.6d, KS3591
8-5	GACGTAGG	NFME	8.6e, KS2591
8-6	GATCGCAG	LJKE	8.6f, KS4591
8-7	CTAACGCC	IMBQ	8.6g
8-8	GAAACGCC	IMBE	8.6h
8-9	GTTACGCC	IMFR	8.6i
8-10	GTATCGCC	IMPR	8.6j
8-11	GTAAGGCC	INBR	8.6k
8-12	GTAACCCC	IIBR	8.6l
8-13	GTAACGGC	JMBR	8.6m
8-14	GTAACGCG	MMBR	8.6n
8-15	TTCGTGCC	IOMS	8.6o
8-16	ACCCAACC	IBIH	8.6p
8-17	TGGTGAGG	NERO	8.6q
8-18	AGCGGACA	DEML	8.6r
8-19	TGCGTCCA	DKMO	8.6s
8-20	AATGCAGC	JDOB	8.7a, A059
8-21	GCTGGTGG	NROJ	8.7b, BB3
8-22	GAGGGTGG	NRNE	8.7c, G1 ✓
8-23	CCGAGCTG	OJEI	8.7d, B3 11/6
8-24	CCTGGTGG	NROI	8.7e
8-25	CGTGGTGG	NROM	8.7f
8-26	CCAGGTGG	NRLI	8.7g
8-27	CCTCGTGG	NRKI	8.7h
8-28	CCTGCTGG	NQOI	8.7i
8-29	CCTGGAGG	NEOI	8.7j
8-30	CCTGGTCG	MROI	8.7k
8-31	CCTGGTGC	JROI	8.7l
8-32	CAGCTCGG	NKJD	8.7m
8-33	CGCACGTC	KMDM	8.7n
8-34	GCAGGCCA	DJLJ	8.7o
8-35	GCCGGCTA	FJMJ	8.7p
8-36	GCAGGTGC	JRLJ	8.7q
8-37	GGGAGCTG	OJEN	8.7r
8-38	CGCGAAGG	NBMM	8.7s
8-39	CGCCGTCA	DRIM	8.7t
8-40	AAACTCAG	LKHB	8.3a
8-41	CTGGACTA	FHNQ	8.5b
8-42	CGCGGCCA	DJMM	8.9a, L60, L60'
8-43	AGCTTGTC	KOQL	8.5a, L59
8-44	ATATCGCC	IMPP	8.5c
8-45	GGACCCGC	JIHN	8.9b, G4
8-46	GGGGGGGG	NNNN	8.10a, B1
8-47	GCCCGCCC	IJIJ	8.10b, B3
8-48	CGCGCCGG	NIMM	8.10c, G3

Table C-1 (continued)

Name	Sequence	Code	Other names
3.7a	GCT	Qg	
4.5a	ACGT	RH	
4.5b	CTGT	RQ	E785
4.7a	TTGT	SR	
4.10a	GCGC	MM	
5.0a	TATAT	PPt	
5.2a	GATAT	PPg	
5.4a	AATGC	JPa	A026, BB1$^{11/20}$
5.4b	CTTGT	RSc	A028
5.4c	GGTAT	PRg	
5.4d	TATGG	NPt	
5.4e	TTTGG	NSt	
5.4f	GGTTT	SRg	
5.6a	CCTGT	RQc	BB2$^{11/20}$, E784
5.6b	AGCTG	OJa	B6$^{11/6}$
5.6c	TGCTG	OJt	B7$^{11/6}$
5.6d	ACCTG	OIa	B8$^{11/6}$
5.6e	AGGTG	ONa	B9$^{11/6}$
5.6f	AGCAG	LJa	B10$^{11/6}$
5.6g	AGCTC	KJa	B11$^{11/6}$
5.6h	CAGCT	QLc	
5.8a	ACGCC	IMa	
6.5a	GAATGC	JPE	
6.6a	GCCTGT	RQJ	BB3$^{11/20}$
6.6b	CCTGTG	OOI	BB6$^{11/20}$
6.6c	AACGCC	IMB	
6.7a	GAGCTG	OJE	B5$^{11/6}$
7.5a	TAACGCC	IMBt	
7.6a	AGCCTGT	RQJa	BB4$^{11/20}$
7.6b	CCTGTGA	ERQc	BB7$^{11/20}$
7.7a	CGAGCTG	OJEc	B4$^{11/6}$
7.7b	GGAGCTG	OJEg	
7.7c	CCAGCTG	OJDc	
7.7d	CGTGCTG	OJRc	
9.8a	GCCGAGCTG	OJEIg	B2$^{11/6}$
9.10a	CGGCGGCGG	NJMNc	B2
9.10b	GCCGCCGCC	IMJIg	G2
10.5a	AATGCAGCTG	OJDOB	A646
10.6a	AGTCAGCCAC	HILKL	A03
10.6b	AATCGGGCTG	OJNKB	A04
10.6c	AGGGGTCCTG	OIRNL	A05
10.6d	GAAACGGGTC	KNMBE	A07
10.6e	GTGACGTAGG	NFMER	A08
10.6f	GTGATCGCAG	LJKER	A10
10.6g	CAATCGCCGT	RIMPD	A11
10.6h	TCGGCGATAG	LPMNK	A12
10.6i	TCTGTGCTGG	NQOOK	A14

Table C-1 (continued)

10.6j	TTCCGAACCC	IHEIS	A15
10.6k	ACGCAGCGAA	BMLJH	A16
10.6l	GACCGCTTGT	RSJIE	A17
10.6m	AGGTGACCGT	RIERL	A18
10.6n	CAAACGTCGG	NKMBD	A19
10.6o	GTTGCGATCC	IPMOR	A20
10.7a	CAGGCCCTTC	KQIND	A01
10.7b	TGCCGAGCTG	OJEIO	A02,B1[11/6]
10.7c	GGTCCCTGAC	HOIKN	A06
10.7d	GGGTAACGCC	IMBRN	A09
10.7e	CAGCACCCAC	HIHJD	A13
10.7f	ACCCGAGCTG	OJEIH	
10.7g	TCCCGAGCTG	OJEIK	
10.7h	AGCCGAGCTG	OJEIL	
12.6a	AATGCAGCTGGC	JOJDOB	A728
12.7a	CGTGCCGAGCTG	OJEIOM	
15.6a	GAGCTGGGTAACGCC		
15.8a	TAGCTGGCCGAGCTG		
17.4a	GAAATCACTCCCAATTA		R2
17.4b	AATACGACTCACTATAG		R3
21.5a	CTCATAAGGGGGTTCATACAC		E182
21.5b	ACACCACCTGAGTATCTACCT		A277
21.8a	GGTGTCCTGGGTGGGGGGG		

Table C-2. List of series A, B, and C for octamer (8-mer) and mini-hairpin primers in the Plant Molecular Genetics, The University of Tennessee, Knoxville, TN.

PRIMER SERIES	CODE				
Octamer. A		GCCCGGGA	A44	GCAGGTTT	B22
GCCCG-NNN		GCCCGGAC	A45	GCAGGTTG	B23
GCCCGCCC	A1	GCCCGGAT	A46	GCAGGTTA	B24
GCCCGCCT	A2	GCCCGGAG	A47	GCAGGTGC	B25
GCCCGCCG	A3	GCCCGGAA	A48	GCAGGTGT	B26
GCCCGCCA	A4	GCCCGACC	A49	GCAGGTGG	B27
GCCCGCTC	A5	GCCCGACT	A50	GCAGGTGA	B28
GCCCGCTT	A6	GCCCGACG	A51	GCAGGTAC	B29
GCCCGCTG	A7	GCCCGACA	A52	GCAGGTAT	B30
GCCCGCTA	A8	GCCCGATC	A53	GCAGGTAG	B31
GCCCGCGC	A9	GCCCGATT	A54	GCAGGTAA	B32
GCCCGCGT	A10	GCCCGATG	A55	GCAGGGCC	B33
GCCCGCGG	A11	GCCCGATA	A56	GCAGGGCT	B34
GCCCGCGA	A12	GCCCGAGC	A57	GCAGGGCG	B35
GCCCGCAC	A13	GCCCGAGT	A58	GCAGGGCA	B36
GCCCGCAT	A14	GCCCGAGG	A59	GCAGGGTC	B37
GCCCGCAG	A15	GCCCGAGA	A60	GCAGGGTT	B38
GCCCGCAA	A16	GCCCGAAC	A61	GCAGGGTG	B39
GCCCGTCC	A17	GCCCGAAT	A62	GCAGGGTA	B40
GCCCGTCT	A18	GCCCGAAG	A63	GCAGGGGC	B41
GCCCGTCG	A19	GCCCGAAA	A64	GCAGGGGT	B42
GCCCGTCA	A20			GCAGGGGG	B43
GCCCGTTC	A21	PRIMER SERIES	CODE	GCAGGGGA	B44
GCCCGTTT	A22	**Octamer. B**		GCAGGGAC	B45
GCCCGTTG	A23	**GCAGG-NNN**		GCAGGGAT	B46
GCCCGTTA	A24	GCAGGCCC	B1	GCAGGGAG	B47
GCCCGTGC	A25	GCAGGCCT	B2	GCAGGGAA	B48
GCCCGTGT	A26	GCAGGCCG	B3	GCAGGACC	B49
GCCCGTGG	A27	GCAGGCCA	B4	GCAGGACT	B50
GCCCGTGA	A28	GCAGGCTC	B5	GCAGGACG	B51
GCCCGTAC	A29	GCAGGCTT	B6	GCAGGACA	B52
GCCCGTAT	A30	GCAGGCTG	B7	GCAGGATC	B53
GCCCGTAG	A31	GCAGGCTA	B8	GCAGGATT	B54
GCCCGTAA	A32	GCAGGCGC	B9	GCAGGATG	B55
GCCCGGCC	A33	GCAGGCGT	B10	GCAGGATA	B56
GCCCGGCT	A34	GCAGGCGG	B11	GCAGGAGC	B57
GCCCGGCG	A35	GCAGGCGA	B12	GCAGGAGT	B58
GCCCGGCA	A36	GCAGGCAC	B13	GCAGGAGG	B59
GCCCGGTC	A37	GCAGGCAT	B14	GCAGGAGA	B60
GCCCGGTT	A38	GCAGGCAG	B15	GCAGGAAC	B61
GCCCGGTG	A39	GCAGGCAA	B16	GCAGGAAT	B62
GCCCGGTA	A40	GCAGGTCC	B17	GCAGGAAG	B63
GCCCGGGC	A41	GCAGGTCT	B18	GCAGGAAA	B64
GCCCGGGT	A42	GCAGGTCG	B19		
GCCCGGGG	A43	GCAGGTCA	B20		
		GCAGGTTC	B21		

Table C-2 (continued)

PRIMER SERIES	CODE
Octamer. C	
CCGAG-NNN	
CCGAGCCC	C1
CCGAGCCT	C2
CCGAGCCG	C3
CCGAGCCA	C4
CCGAGCTC	C5
CCGAGCTT	C6
CCGAGCTG	C7
CCGAGCTA	C8
CCGAGCGC	C9
CCGAGCGT	C10
CCGAGCGG	C11
CCGAGCGA	C12
CCGAGCAC	C13
CCGAGCAT	C14
CCGAGCAG	C15
CCGAGCAA	C16
CCGAGTCC	C17
CCGAGTCT	C18
CCGAGTCG	C19
CCGAGTCA	C20
CCGAGTTC	C21
CCGAGTTT	C22
CCGAGTTG	C23
CCGAGTTA	C24
CCGAGTGC	C25
CCGAGTGT	C26
CCGAGTGG	C27
CCGAGTGA	C28
CCGAGTAC	C29
CCGAGTAT	C30
CCGAGTAG	C31
CCGAGTAA	C32
CCGAGGCC	C33
CCGAGGCT	C34
CCGAGGCG	C35
CCGAGGCA	C36
CCGAGGTC	C37
CCGAGGTT	C38
CCGAGGTG	C39
CCGAGGTA	C40
CCGAGGGC	C41
CCGAGGGT	C42
CCGAGGGG	C43
CCGAGGGA	C44
CCGAGGAC	C45
CCGAGGAT	C46
CCGAGGAG	C47
CCGAGGAA	C48
CCGAGACC	C49
CCGAGACT	C50
CCGAGACG	C51
CCGAGACA	C52
CCGAGATC	C53
CCGAGATT	C54
CCGAGATG	C55
CCGAGATA	C56
CCGAGAGC	C57
CCGAGAGT	C58
CCGAGAGG	C59
CCGAGAGA	C60
CCGAGAAC	C61
CCGAGAAT	C62
CCGAGAAG	C63
CCGAGAAA	C64

PRIMER SERIES	CODE
Octamer. D	
GATCG-NNN	
GATCGCCC	D1
GATCGCCT	D2
GATCGCCG	D3
GATCGCCA	D4
GATCGCTC	D5
GATCGCTT	D6
GATCGCTG	D7
GATCGCTA	D8
GATCGCGC	D9
GATCGCGT	D10
GATCGCGG	D11
GATCGCGA	D12
GATCGCAC	D13
GATCGCAT	D14
GATCGCAG	D15
GATCGCAA	D16
GATCGTCC	D17
GATCGTCT	D18
GATCGTCG	D19
GATCGTCA	D20
GATCGTTC	D21
GATCGTTT	D22
GATCGTTG	D23
GATCGTTA	D24
GATCGTGC	D25
GATCGTGT	D26
GATCGTGG	D27
GATCGTGA	D28
GATCGTAC	D29
GATCGTAT	D30
GATCGTAG	D31
GATCGTAA	D32
GATCGGCC	D33
GATCGGCT	D34
GATCGGCG	D35
GATCGGCA	D36
GATCGGTC	D37
GATCGGTT	D38
GATCGGTG	D39
GATCGGTA	D40
GATCGGGC	D41
GATCGGGT	D42
GATCGGGG	D43
GATCGGGA	D44
GATCGGAC	D45
GATCGGAT	D46
GATCGGAG	D47
GATCGGAA	D48
GATCGACC	D49
GATCGACT	D50
GATCGACG	D51
GATCGACA	D52
GATCGATC	D53
GATCGATT	D54
GATCGATG	D55
GATCGATA	D56
GATCGAGC	D57
GATCGAGT	D58
GATCGAGG	D59
GATCGAGA	D60
GATCGAAC	D61
GATCGAAT	D62
GATCGAAG	D63
GATCGAAA	D64

Table C-2 (continued)

		GCGAAAGC-GGG	hpA43	GCGACAGC-TTC	hpB21
PRIMER SERIES	CODE	GCGAAAGC-GGA	hpA44	GCGACAGC-TTT	hpB22
Mini-hairpin. hpA		GCGAAAGC-GAC	hpA45	GCGACAGC-TTG	hpB23
GCGAAAGC-NNN		GCGAAAGC-GAT	hpA46	GCGACAGC-TTA	hpB24
GCGAAAGC-CCC	hpA1	GCGAAAGC-GAG	hpA47	GCGACAGC-TGC	hpB25
GCGAAAGC-CCT	hpA2	GCGAAAGC-GAA	hpA48	GCGACAGC-TGT	hpB26
GCGAAAGC-CCG	hpA3	GCGAAAGC-ACC	hpA49	GCGACAGC-TGG	hpB27
GCGAAAGC-CCA	hpA4	GCGAAAGC-ACT	hpA50	GCGACAGC-TGA	hpB28
GCGAAAGC-CTC	hpA5	GCGAAAGC-ACG	hpA51	GCGACAGC-TAC	hpB29
GCGAAAGC-CTT	hpA6	GCGAAAGC-ACA	hpA52	GCGACAGC-TAT	hpB30
GCGAAAGC-CTG	hpA7	GCGAAAGC-ATC	hpA53	GCGACAGC-TAG	hpB31
GCGAAAGC-CTA	hpA8	GCGAAAGC-ATT	hpA54	GCGACAGC-TAA	hpB32
GCGAAAGC-CGC	hpA9	GCGAAAGC-ATG	hpA55	GCGACAGC-GCC	hpB33
GCGAAAGC-CGT	hpA10	GCGAAAGC-ATA	hpA56	GCGACAGC-GCT	hpB34
GCGAAAGC-CGG	hpA11	GCGAAAGC-AGC	hpA57	GCGACAGC-GCG	hpB35
GCGAAAGC-CGA	hpA12	GCGAAAGC-AGT	hpA58	GCGACAGC-GCA	hpB36
GCGAAAGC-CAC	hpA13	GCGAAAGC-AGG	hpA59	GCGACAGC-GTC	hpB37
GCGAAAGC-CAT	hpA14	GCGAAAGC-AGA	hpA60	GCGACAGC-GTT	hpB38
GCGAAAGC-CAG	hpA15	GCGAAAGC-AAC	hpA61	GCGACAGC-GTG	hpB39
GCGAAAGC-CAA	hpA16	GCGAAAGC-AAT	hpA62	GCGACAGC-GTA	hpB40
GCGAAAGC-TCC	hpA17	GCGAAAGC-AAG	hpA63	GCGACAGC-GGC	hpB41
GCGAAAGC-TCT	hpA18	GCGAAAGC-AAA	hpA64	GCGACAGC-GGT	hpB42
GCGAAAGC-TCG	hpA19			GCGACAGC-GGG	hpB43
GCGAAAGC-TCA	hpA20	PRIMER SERIES	CODE	GCGACAGC-GGA	hpB44
GCGAAAGC-TTC	hpA21	**Mini-hairpin. hpB**		GCGACAGC-GAC	hpB45
GCGAAAGC-TTT	hpA22	**GCGACAGC-NNN**		GCGACAGC-GAT	hpB46
GCGAAAGC-TTG	hpA23	GCGACAGC-CCC	hpB1	GCGACAGC-GAG	hpB47
GCGAAAGC-TTA	hpA24	GCGACAGC-CCT	hpB2	GCGACAGC-GAA	hpB48
GCGAAAGC-TGC	hpA25	GCGACAGC-CCG	hpB3	GCGACAGC-ACC	hpB49
GCGAAAGC-TGT	hpA26	GCGACAGC-CCA	hpB4	GCGACAGC-ACT	hpB50
GCGAAAGC-TGG	hpA27	GCGACAGC-CTC	hpB5	GCGACAGC-ACG	hpB51
GCGAAAGC-TGA	hpA28	GCGACAGC-CTT	hpB6	GCGACAGC-ACA	hpB52
GCGAAAGC-TAC	hpA29	GCGACAGC-CTG	hpB7	GCGACAGC-ATC	hpB53
GCGAAAGC-TAT	hpA30	GCGACAGC-CTA	hpB8	GCGACAGC-ATT	hpB54
GCGAAAGC-TAG	hpA31	GCGACAGC-CGC	hpB9	GCGACAGC-ATG	hpB55
GCGAAAGC-TAA	hpA32	GCGACAGC-CGT	hpB10	GCGACAGC-ATA	hpB56
GCGAAAGC-GCC	hpA33	GCGACAGC-CGG	hpB11	GCGACAGC-AGC	hpB57
GCGAAAGC-GCT	hpA34	GCGACAGC-CGA	hpB12	GCGACAGC-AGT	hpB58
GCGAAAGC-GCG	hpA35	GCGACAGC-CAC	hpB13	GCGACAGC-AGG	hpB59
GCGAAAGC-GCA	hpA36	GCGACAGC-CAT	hpB14	GCGACAGC-AGA	hpB60
GCGAAAGC-GTC	hpA37	GCGACAGC-CAG	hpB15	GCGACAGC-AAC	hpB61
GCGAAAGC-GTT	hpA38	GCGACAGC-CAA	hpB16	GCGACAGC-AAT	hpB62
GCGAAAGC-GTG	hpA39	GCGACAGC-TCC	hpB17	GCGACAGC-AAG	hpB63
GCGAAAGC-GTA	hpA40	GCGACAGC-TCT	hpB18	GCGACAGC-AAA	hpB64
GCGAAAGC-GGC	hpA41	GCGACAGC-TCG	hpB19		
GCGAAAGC-GGT	hpA42	GCGACAGC-TCA	hpB20		

Table C-2 (continued)

PRIMER SERIES	CODE	GCGAGAGC-GGA	hpC44	GCGATAGC-TTT	hpD22
Mini-hairpin. hpC		GCGAGAGC-GAC	hpC45	GCGATAGC-TTG	hpD23
GCGAGAGC-NNN		GCGAGAGC-GAT	hpC46	GCGATAGC-TTA	hpD24
GCGAGAGC-CCC	hpC1	GCGAGAGC-GAG	hpC47	GCGATAGC-TGC	hpD25
GCGAGAGC-CCT	hpC2	GCGAGAGC-GAA	hpC48	GCGATAGC-TGT	hpD26
GCGAGAGC-CCG	hpC3	GCGAGAGC-ACC	hpC49	GCGATAGC-TGG	hpD27
GCGAGAGC-CCA	hpC4	GCGAGAGC-ACT	hpC50	GCGATAGC-TGA	hpD28
GCGAGAGC-CTC	hpC5	GCGAGAGC-ACG	hpC51	GCGATAGC-TAC	hpD29
GCGAGAGC-CTT	hpC6	GCGAGAGC-ACA	hpC52	GCGATAGC-TAT	hpD30
GCGAGAGC-CTG	hpC7	GCGAGAGC-ATC	hpC53	GCGATAGC-TAG	hpD31
GCGAGAGC-CTA	hpC8	GCGAGAGC-ATT	hpC54	GCGATAGC-TAA	hpD32
GCGAGAGC-CGC	hpC9	GCGAGAGC-ATG	hpC55	GCGATAGC-GCC	hpD33
GCGAGAGC-CGT	hpC10	GCGAGAGC-ATA	hpC56	GCGATAGC-GCT	hpD34
GCGAGAGC-CGG	hpC11	GCGAGAGC-AGC	hpC57	GCGATAGC-GCG	hpD35
GCGAGAGC-CGA	hpC12	GCGAGAGC-AGT	hpC58	GCGATAGC-GCA	hpD36
GCGAGAGC-CAC	hpC13	GCGAGAGC-AGG	hpC59	GCGATAGC-GTC	hpD37
GCGAGAGC-CAT	hpC14	GCGAGAGC-AGA	hpC60	GCGATAGC-GTT	hpD38
GCGAGAGC-CAG	hpC15	GCGAGAGC-AAC	hpC61	GCGATAGC-GTG	hpD39
GCGAGAGC-CAA	hpC16	GCGAGAGC-AAT	hpC62	GCGATAGC-GTA	hpD40
GCGAGAGC-TCC	hpC17	GCGAGAGC-AAG	hpC63	GCGATAGC-GGC	hpD41
GCGAGAGC-TCT	hpC18	GCGAGAGC-AAA	hpC64	GCGATAGC-GGT	hpD42
GCGAGAGC-TCG	hpC19			GCGATAGC-GGG	hpD43
GCGAGAGC-TCA	hpC20	PRIMER SERIES	CODE	GCGATAGC-GGA	hpD44
GCGAGAGC-TTC	hpC21	Mini-hairpin. hpC	D	GCGATAGC-GAC	hpD45
GCGAGAGC-TTT	hpC22	GCGATAGC-NNN		GCGATAGC-GAT	hpD46
GCGAGAGC-TTG	hpC23	GCGATAGC-CCC	hpD1	GCGATAGC-GAG	hpD47
GCGAGAGC-TTA	hpC24	GCGATAGC-CCT	hpD2	GCGATAGC-GAA	hpD48
GCGAGAGC-TGC	hpC25	GCGATAGC-CCG	hpD3	GCGATAGC-ACC	hpD49
GCGAGAGC-TGT	hpC26	GCGATAGC-CCA	hpD4	GCGATAGC-ACT	hpD50
GCGAGAGC-TGG	hpC27	GCGATAGC-CTC	hpD5	GCGATAGC-ACG	hpD51
GCGAGAGC-TGA	hpC28	GCGATAGC-CTT	hpD6	GCGATAGC-ACA	hpD52
GCGAGAGC-TAC	hpC29	GCGATAGC-CTG	hpD7	GCGATAGC-ATC	hpD53
GCGAGAGC-TAT	hpC30	GCGATAGC-CTA	hpD8	GCGATAGC-ATT	hpD54
GCGAGAGC-TAG	hpC31	GCGATAGC-CGC	hpD9	GCGATAGC-ATG	hpD55
GCGAGAGC-TAA	hpC32	GCGATAGC-CGT	hpD10	GCGATAGC-ATA	hpD56
GCGAGAGC-GCC	hpC33	GCGATAGC-CGG	hpD11	GCGATAGC-AGC	hpD57
GCGAGAGC-GCT	hpC34	GCGATAGC-CGA	hpD12	GCGATAGC-AGT	hpD58
GCGAGAGC-GCG	hpC35	GCGATAGC-CAC	hpD13	GCGATAGC-AGG	hpD59
GCGAGAGC-GCA	hpC36	GCGATAGC-CAT	hpD14	GCGATAGC-AGA	hpD60
GCGAGAGC-GTC	hpC37	GCGATAGC-CAG	hpD15	GCGATAGC-AAC	hpD61
GCGAGAGC-GTT	hpC38	GCGATAGC-CAA	hpD16	GCGATAGC-AAT	hpD62
GCGAGAGC-GTG	hpC39	GCGATAGC-TCC	hpD17	GCGATAGC-AAG	hpD63
GCGAGAGC-GTA	hpC40	GCGATAGC-TCT	hpD18	GCGATAGC-AAA	hpD64
GCGAGAGC-GGC	hpC41	GCGATAGC-TCG	hpD19		
GCGAGAGC-GGT	hpC42	GCGATAGC-TCA	hpD20		
GCGAGAGC-GGG	hpC43	GCGATAGC-TTC	hpD21		

Table C-3. List of mini-hairpin primers in the Plant Molecular Genetics, The University of Tennessee, Knoxville.

- all mini hairpin primers used coontained the stem-loop GCGAAGC---

Primer	Hairpin Tag	Primer	Hairpin Tag	Primer	Hairpin Tag
HP8	GCT	HP53	TTC	HP80	ATC
HP10	CTG	HP54	TTT	HP81	ATT
HP28	CTA	HP55	TTA	HP82	ATA
HP29	CTC	HP56	TGC	HP83	AGC
HP30	CTT	HP57	TGT	HP84	AGT
HP31	CAG	HP58	TGA	HP85	AGA
HP32	CCG	HP59	TAC	HP86	AAC
HP33	CGG	HP60	TAT		
HP34	CAC	HP61	TAG		
HP35	CGC	HP62	TAA		
HP36	CCC	HP63	GCC		
HP37	CGT	HP64	GCG		
HP38	CCT	HP65	GCA		
HP39	CAT	HP66	GTC		
HP40	CGA	HP67	GTT		
HP41	CCA	HP68	GTA		
HP42	CAA	HP69	GGC		
HP43	GTG	HP70	GGT		
HP44	ATG	HP71	GGA		
HP45	TTG	HP72	GAC		
HP46	GGG	HP73	GAT		
HP47	AGG	HP74	GAG		
HP48	TGG	HP75	GAA		
HP49	TCC	HP76	ACC		
HP50	TCT	HP77	ACT		
HP51	TCG	HP78	ACG		
HP52	TCA	HP79	ACA		

Table C-4. List of octamer primers obtained for Oak Ridge National Laboratory; OR primers.

Correct Sequence

No.	Pos 1,2,3	Pos 6,7,8	No.	Pos 1,2,3	Pos 6,7,8
1	AAATGGAC	GTCCATTT	33	GAATGGAC	GTCCATTC
2	AATTGGAC	GTCCAATT	34	GATTGGAC	GTCCAATC
3	AAGTGGAC	GTCCACTT	35	GAGTGGAC	GTCCACTC
4	AACTGGAC	GTCCAGTT	36	GACTGGAC	GTCCAGTC
5	ATATGGAC	GTCCATAT	37	GTATGGAC	GTCCATAC
6	ATTTGGAC	GTCCAAAT	38	GTTTGGAC	GTCCAAAC
7	ATGTGGAC	GTCCACAT	39	GTGTGGAC	GTCCACAC
8	ATCTGGAC	GTCCAGAT	40	GTCTGGAC	GTCCAGAC
9	AGATGGAC	GTCCATCT	41	GGATGGAC	GTCCATCC
10	AGTTGGAC	GTCCAACT	42	GGTTGGAC	GTCCAACC
11	AGGTGGAC	GTCCACCT	43	GGGTGGAC	GTCCACCC
12	AGCTGGAC	GTCCAGCT	44	GGCTGGAC	GTCCAGCC
13	ACATGGAC	GTCCATGT	45	GCATGGAC	GTCCATGC
14	ACTTGGAC	GTCCAAGT	46	GCTTGGAC	GTCCAAGC
15	ACGTGGAC	GTCCACGT	47	GCGTGGAC	GTCCACGC
16	ACCTGGAC	GTCCAGGT	48	GCCTGGAC	GTCCAGGC
17	TAATGGAC	GTCCATTA	49	CAATGGAC	GTCCATTG
18	TATTGGAC	GTCCAATA	50	CATTGGAC	GTCCAATG
19	TAGTGGAC	GTCCACTA	51	CAGTGGAC	GTCCACTG
20	TACTGGAC	GTCCAGTA	52	CACTGGAC	GTCCAGTG
21	TTATGGAC	GTCCATAA	53	CTATGGAC	GTCCATAG
22	TTTTGGAC	GTCCAAAA	54	CTTTGGAC	GTCCAAAG
23	TTGTGGAC	GTCCACAA	55	CTGTGGAC	GTCCACAG
24	TTCTGGAC	GTCCAGAA	56	CTCTGGAC	GTCCAGAG
25	TGATGGAC	GTCCATCA	57	CGATGGAC	GTCCATCG
26	TGTTGGAC	GTCCAACA	58	CGTTGGAC	GTCCAACG
27	TGGTGGAC	GTCCACCA	59	CGGTGGAC	GTCCACCG
28	TGCTGGAC	GTCCAGCA	60	CGCTGGAC	GTCCAGCG
29	TCATGGAC	GTCCATGA	61	CCATGGAC	GTCCATGG
30	TCTTGGAC	GTCCAAGA	62	CCTTGGAC	GTCCAAGG
31	TCGTGGAC	GTCCACGA	63	CCGTGGAC	GTCCACGG
32	TCCTGGAC	GTCCAGGA	64	CCCTGGAC	GTCCAGGG

Table C-5. List of series A to N decamer (10-mer) primers obtained from Operon Technology Inc.

Operon 10-mer Kits

Each kit contains twenty 10-mer primers, as listed below, and contains 0.5 O.D. (approx. 15 µgm) per tube. The list price is $150.00 per kit.

KIT A

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPA-01	CAGGCCCTTC	OPA-06	GGTCCCTGAC	OPA-11	CAATCGCCGT	OPA-16	AGCCAGCGAA
OPA-02	TGCCGAGCTG	OPA-07	GAAACGGGTG	OPA-12	TCGGCGATAG	OPA-17	GACCGCTTGT
OPA-03	AGTCAGCCAC	OPA-08	GTGACGTAGG	OPA-13	CAGCACCCAC	OPA-18	AGGTGACCGT
OPA-04	AATCGGGCTG	OPA-09	GGGTAACGCC	OPA-14	TCTGTGCTGG	OPA-19	CAAACGTCGG
OPA-05	AGGGGTCTTG	OPA-10	GTGATCGCAG	OPA-15	TTCCGAACCC	OPA-20	GTTGCGATCC

KIT B

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPB-01	GTTTCGCTCC	OPB-06	TGCTCTGCCC	OPB-11	GTAGACCCGT	OPB-16	TTTGCCCGGA
OPB-02	TGATCCCTGG	+ OPB-07	GGTGACGCAG	OPB-12	CCTTGACGCA	OPB-17	AGGGAACGAG
OPB-03	CATCCCCCTG	OPB-08	GTCCACACGG	OPB-13	TTCCCCCGCT	OPB-18	CCACAGCAGT
* OPB-04	GGACTGGAGT	OPB-09	TGGGGGACTC	OPB-14	TCCGCTCTGG	OPB-19	ACCCCCGAAG
OPB-05	TGCGCCCTTC	OPB-10	CTGCTGGGAC	OPB-15	GGAGGGTGTT	OPB-20	GGACCCTTAC

KIT C

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPC-01	TTCGAGCCAG	OPC-06	GAACGGACTC	OPC-11	AAAGCTGCGG	OPC-16	CACACTCCAG
OPC-02	GTGAGGCGTC	OPC-07	GTCCCGACGA	OPC-12	TGTCATCCCC	OPC-17	TTCCCCCCAG
OPC-03	GGGGGTCTTT	OPC-08	TGGACCGGTG	OPC-13	AAGCCTCGTC	OPC-18	TGAGTGGGTG
OPC-04	CCGCATCTAC	OPC-09	CTCACCGTCC	OPC-14	TGCGTGCTTG	OPC-19	GTTGCCAGCC
OPC-05	GATGACCGCC	OPC-10	TGTCTGGGTG	OPC-15	GACGGATCAG	OPC-20	ACTTCGCCAC

KIT D

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPD-01	ACCGCGAAGG	OPD-06	ACCTGAACGG	OPD-11	AGCGCCATTG	OPD-16	AGGGCGTAAG
OPD-02	GGACCCAACC	OPD-07	TTGGCACGGG	OPD-12	CACCGTATCC	OPD-17	TTTCCCACGG
OPD-03	GTCGCCGTCA	OPD-08	GTGTGCCCCA	OPD-13	GGGGTGACGA	OPD-18	GAGAGCCAAC
OPD-04	TCTGGTGAGG	OPD-09	CTCTGGAGAC	OPD-14	CTTCCCCAAG	OPD-19	CTGGGGACTT
OPD-05	TGAGCGGACA	OPD-10	GGTCTACACC	OPD-15	CATCCGTGCT	OPD-20	ACCCGGTCAC

KIT E

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPE-01	CCCAAGGTCC	OPE-06	AAGACCCCTC	OPE-11	GAGTCTCAGG	OPE-16	GGTGACTGTG
OPE-02	GGTGCGGGAA	OPE-07	AGATGCAGCC	OPE-12	TTATCGCCCC	OPE-17	CTACTGCCGT
OPE-03	CCAGATGCAC	OPE-08	TCACCACGGT	OPE-13	CCCGATTCGG	OPE-18	GGACTGCAGA
OPE-04	GTGACATGCC	OPE-09	CTTCACCCGA	OPE-14	TGCGGCTGAG	OPE-19	ACGGCGTATG
OPE-05	TCAGGGAGGT	OPE-10	CACCAGGTGA	OPE-15	ACGCACAACC	OPE-20	AACGGTGACC

KIT F

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPF-01	ACGGATCCTG	OPF-06	GGGAATTCGG	OPF-11	TTGGTACCCC	OPF-16	GGAGTACTGG
OPF-02	GAGGATCCCT	OPF-07	CCGATATCCC	OPF-12	ACGGTACCAG	OPF-17	AACCCGGGAA
OPF-03	CCTGATCACC	OPF-08	GGGATATCGG	OPF-13	GGCTGCAGAA	OPF-18	TTCCCGGGTT
OPF-04	GGTGATCAGG	OPF-09	CCAAGCTTCC	OPF-14	TGCTGCAGGT	OPF-19	CCTCTAGACC
OPF-05	CCGAATTCCC	OPF-10	GGAAGCTTGG	OPF-15	CCAGTACTCC	OPF-20	GGTCTAGAGG

Table C-5 (continued)

KIT G

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPG-01	CTACGGACGA	OPG-06	GTGCCTAACC	OPG-11	TGCCCGTCGT	OPG-16	AGCGTCCTCC
OPG-02	GGCACTGAGG	OPG-07	GAACCTGCGG	OPG-12	CAGCTCACGA	OPG-17	ACGACCGACA
OPG-03	GAGCCCTCCA	* OPG-08	TCACGTCCAC	OPG-13	CTCTCCGCCA	OPG-18	GGCTCATGTG
OPG-04	AGCGTGTCTG	OPG-09	CTCACGTCAC	OPG-14	GGATGAGACC	OPG-19	GTCAGGGCAA
OPG-05	CTGAGACGGA	OPG-10	AGGGCCGTCT	OPG-15	ACTGGGACTC	OPG-20	TCTCCCTCAG

KIT H

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPH-01	GGTCGGAGAA	OPH-06	ACGCATCGCA	OPH-11	CTTCCGCAGT	OPH-16	TCTCAGCTGG
OPH-02	TCGGACGTGA	OPH-07	CTGCATCGTG	OPH-12	ACGCGCATGT	OPH-17	CACTCTCCTC
OPH-03	AGACGTCCAC	OPH-08	GAAACACCCC	OPH-13	GACGCCACAC	OPH-18	GAATCGGCCA
OPH-04	GGAAGTCGCC	OPH-09	TGTAGCTGGG	OPH-14	ACCAGGTTGG	OPH-19	CTGACCAGCC
OPH-05	AGTCGTCCCC	OPH-10	CCTACGTCAG	OPH-15	AATGGCGCAG	OPH-20	GGGAGACATC

KIT I

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPI-01	ACCTGGACAC	OPI-06	AAGGCGGCAG	OPI-11	ACATGCCGTG	OPI-16	TCTCCGCCCT
OPI-02	GGAGGAGAGG	OPI-07	CAGCGACAAG	OPI-12	AGAGGGCACA	OPI-17	GGTGGTGATG
OPI-03	CAGAAGCCCA	OPI-08	TTTGCCCGGT	OPI-13	CTGGGGCTGA	OPI-18	TGCCCAGCCT
OPI-04	CCGCCTAGTC	OPI-09	TGGAGAGCAG	OPI-14	TGACGGCGGT	OPI-19	AATGCGGGAG
OPI-05	TGTTCCACGG	OPI-10	ACAACGCGAG	OPI-15	TCATCCGAGG	OPI-20	AAAGTGCGGG

KIT J

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPJ-01	CCCGGCATAA	OPJ-06	TCGTTCCGCA	OPJ-11	ACTCCTGCGA	OPJ-16	CTGCTTAGGG
OPJ-02	CCCGTTGGGA	OPJ-07	CCTCTCGACA	OPJ-12	GTCCCGTGGT	OPJ-17	ACGCCAGTTC
OPJ-03	TCTCCGCTTG	OPJ-08	CATACCGTGG	OPJ-13	CCACACTACC	OPJ-18	TGGTCGCAGA
OPJ-04	CCGAACACGG	OPJ-09	TGAGCCTCAC	OPJ-14	CACCCGGATG	OPJ-19	GGACACCACT
OPJ-05	CTCCATGGGG	OPJ-10	AAGCCCGAGG	OPJ-15	TGTAGCAGGG	OPJ-20	AAGCGGCCTC

KIT K

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPK-01	CATTCGAGCC	OPK-06	CACCTTTCCC	OPK-11	AATGCCCCAG	OPK-16	GAGCGTCGAA
OPK-02	GTCTCCGCAA	OPK-07	AGCGAGCAAG	OPK-12	TGGCCCTCAC	OPK-17	CCCAGCTGTG
OPK-03	CCAGCTTAGG	OPK-08	GAACACTGGG	OPK-13	GGTTGTACCC	OPK-18	CCTAGTCGAG
OPK-04	CCGCCCAAAC	OPK-09	CCCTACCGAC	OPK-14	CCCGCTACAC	OPK-19	CACAGGCGGA
OPK-05	TCTGTCGAGG	OPK-10	GTGCAACGTG	OPK-15	CTCCTGCCAA	OPK-20	GTGTCGCGAG

KIT L

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPL-01	GGCATGACCT	OPL-06	GAGGGAAGAG	OPL-11	ACGATGAGCC	OPL-16	AGGTTGCAGG
OPL-02	TGGGCGTCAA	OPL-07	AGGCGGGAAC	OPL-12	GGGCGGTACT	OPL-17	AGCCTGAGCC
OPL-03	CCAGCAGCTT	OPL-08	AGCAGGTGGA	OPL-13	ACCGCCTGCT	OPL-18	ACCACCCACC
OPL-04	GACTGCACAC	OPL-09	TGCGAGAGTC	OPL-14	GTGACAGGCT	OPL-19	GAGTGGTGAC
OPL-05	ACGCAGGCAC	OPL-10	TGGGAGATGG	OPL-15	AAGAGAGGGG	OPL-20	TGGTGGACCA

KIT M

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPM-01	GTTGGTGGCT	OPM-06	CTGGGCAACT	OPM-11	GTCCACTGTG	OPM-16	GTAACCAGCC
OPM-02	ACAACGCCTC	OPM-07	CCGTGACTCA	OPM-12	GGGACGTTGG	OPM-17	TCAGTCCGGG
OPM-03	GGGGGATGAG	OPM-08	TCTGTTCCCC	OPM-13	GGTGGTCAAG	OPM-18	CACCATCCGT
OPM-04	GGCGGTTGTC	OPM-09	GTCTTGCGGA	OPM-14	AGGGTCGTTC	OPM-19	CCTTCAGGCA
OPM-05	GGGAACGTGT	OPM-10	TCTGGCGCAC	OPM-15	GACCTACCAC	OPM-20	AGGTCTTGGG

KIT N

code	5' to 3'	code	5' to 3'	code	5' to 3'	code	5' to 3'
OPN-01	CTCACGTTGG	OPN-06	CACACCCACA	* OPN-11	TCGCCGCAAA	OPN-16	AAGCGACCTC
OPN-02	ACCAGGGGCA	OPN-07	CAGCCCAGAG	* OPN-12	CACAGACACC	OPN-17	CATTGGGGAG
OPN-03	GGTACTCCCC	OPN-08	ACCTCAGCTC	OPN-13	AGCGTCACTC	OPN-18	GGTGAGGTCA
OPN-04	GACCGACCCA	* OPN-09	TGCCGGCTTG	OPN-14	TCGTGCGGGT	OPN-19	GTCCGTACTG
OPN-05	ACTGAACGCC	* OPN-10	ACAACTGGGG	OPN-15	CAGCGACTGT	OPN-20	GGTGCTCCGT

Vita

Farshid Ghassemi was born in Shahroud, Iran, on August 7, 1954. He attended elementary, middle, and high schools as a top honor student in Tehran, Iran. Then, he was accepted as a second rank student in the Department of Agronomy, College of Agriculture, University of Tehran after passing successfully the national Iranian university entrance test administered by the Ministry of Culture and Higher Education in 1972. As a top honor student, he graduated in 1976 and entered the master of science program majoring in plant breeding in the same department.

He became more and more interested in soybean genetics as he was working on his master project titled "Diallele Cross Analysis in Soybean". He successfully determined the gene action effects and heterosis phenomenon in eight quantitative traits of soybean.

After graduation as a first honor master student in 1980, he became a faculty member of Mazandaran University (1981-1984) and then continued his career in the University of Tehran (1984-1989). After passing a very competitive test, he was awarded a scholarship to pursue his Ph. D. degree abroad.

Recently, he has been offered several post-doctoral position of which he has accepted a position in the Department of Pharmacology, School of Medicine, Yale University, to work on site-directed mutagenesis of serotonin transporter protein.

5338 1508 40
08-04-99 MRB