TOWARDS MARKER ASSISTED SELECTION FOR NEMATODE RESISTANCE IN SOYBEAN

BY

CHARLOTTE MARIA SUSANNA MIENIE

Submitted in fulfilment of the
requirements for the degree of
Doctor of Philosophy

in the
Department of Botany
University of Natal
Pietermaritzburg

2000

ACKNOWLEDGEMENTS

I would like to thank the following:

My supervisor, Prof J van Staden, and co-supervisor, Prof FC Botha, for their guidance and support.

Dr MA Smit for continual help and encouragement.

The Nematology Division, ARC-Grain Crops Institute, and in particular Mss Driekie Fourie and Erna Venter for outstanding technical assistance.

Ms Rita Terblanche for technical assistance and invaluable emotional support.

Dr Mike Butterfield and Mr Pierre van Rooyen for invaluable assistance with statistical analyses.

Mss Alta Enslin for design of figures and Sophie Swanepoel for technical editing of the manuscript.

The Protein Research Trust and the Agricultural Research Council for financial assistance.

Dr KW Pakendorf for his support.

My husband, Tienus, and my family for their continual belief in me and their never failing support.

DECLARATION

I hereby declare that this thesis, unless otherwise acknowledged to the contrary in the text, is the result of my own investigation, under the supervision of Professor J van Staden, Department of Botany, University of Natal, Pietermaritzburg.

Charlotte MS Mienie

We certify that the above statement is correct.

Supervisor : Prof J van Staden

Co-supervisor : Prof FC Botha

ABSTRACT

Meloidogyne javanica is the most widely spread nematode pest on soybean in South Africa. Only a few registered cultivars have some resistance to this nematode and there is an urgent need for an efficient breeding programme for resistant cultivars of all maturity groups. However, breeding is hampered by laborious screening procedures for selection of resistant lines. The objective of this study was to develop an economically viable molecular marker system for application in selection procedures. Three techniques of marker identification were investigated, i.e. RAPD, RFLP and AFLP analysis. The RAPD technique proved to be applicable in fingerprinting soybean varieties, but was too sensitive for interplant variation to be used as an efficient system for identification of molecular markers linked to nematode resistance. Both RFLP and AFLP screening identified markers linked to gall index variation in a segregating population of 60 F_2 progeny from a cross between a resistant cultivar, Gazelle, and a highly susceptible variety, Prima. A codominant RFLP marker(B212) was linked significantly to resistance and explained 62% of the variation in gall index. Seven AFLP markers were linked significantly to the resistance trait, of which four were linked in repulsion phase and three in coupling phase. All seven AFLP markers mapped to LG-F on the public soybean molecular map. The QTL for resistance mapped between markers E-ACC/M-CTC2 linked in coupling phase, B212 and E-AAC/M-CAT1, linked in repulsion phase. These two AFLP markers bracketing the major resistance QTL were successfully converted to SCARs. Marker E-ACC/M-CTC2 was converted to a codominant SCAR marker SOJA6, which acounted for 41% of variation in gall index in the mapping population. Marker E-AAC/M-CAT1 was converted to a dominant SCAR marker (SOJA7) and explained 42% of gall index variation in the mapping population. These two markers mapped approximately 3.8 cM and 2.4 cM respectively from the resistance QTL. This study represents the first report of the development of PCR-based sequence specific markers linked to resistance to *M. javanica* in soybean.

PUBLICATIONS FROM THIS THESIS TO DATE

MIENIE, C.M.S., SMIT, M.A. & PRETORIUS, P.J. 1995. The use of random amplified polymorphic DNA for identification of South African soybean cultivars. *Field Crops Research* 43:43-49.

CONFERENCES FROM THIS THESIS TO DATE

MIENIE, C.M.S. 1997. The development of molecular markers for nematode resistance in soybean. ARC Biotechnology Symposium, Roodeplaat. September.

MIENIE, C.M.S., PAKENDORF, K.W. & FOURIE, H.D. 1997. Development of a molecular marker for resistance to javanese root knot nematode in soybean: genetic basis of resistance. Second Conference, Biotechnology and Development in South Africa, Grahamstown. January.

MIENIE, C.M.S. 2000. Molecular markers linked to nematode resistance in soybean. Southern African Plant Breeders' Association, Harare, Zimbabwe. March.

LIST OF ABBREVIATIONS

A	Adenine
AFLP	Amplified Fragment Length Polymorphism
AP-PCR	Arbitrarily primed polymerase chain reaction
ASAP	Allele-specific associated primers
bp	Basepairs
BSA	Bulked segregant analysis
C	Cytosine
$CaCl_2$	Calciumchloride
CAPS	Cleaved amplified polymorphic sequences
cDNA	Complementary deoxyribonucleic acid
cM	Centimorgan
CTAB	Cetyltrimethylammonium bromide
DAF	Digital amplification fingerprinting
dATP	Deoxy-adenosine triphosphate
dCTP	Deoxy-cytosine triphosphate
DGGE	Denaturing gradient gel electrophoresis
dGTP	Deoxy-guanosine triphosphate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DP	Donor parent
dTTP	Deoxy-thymine triphosphate
dUTP	Deoxy-uridine triphosphate
EDTA	Ethylenediaminetetra-acetic acid disodium salt
G	Guanine
GDE	Genetic diversity estimates
HCl	Hydrochloric acid
IPTG	Isopropyl-ß-D-thiogalactopyranoside
IRA	Inter-repeat amplification
ISSR	Inter-simple sequence repeat

K	Potassium
kb	Kilobases
KCl	Potassiumchloride
LB-medium	Luria-Bertani medium
LG	Linkage group
LOD score	$\log_{10}$ of the odds ratio
Ma	*Meloidogyne arenaria* ((Neal) Chitwood)
MAAP	Multiple arbitrary amplicon profiling
MAS	Marker assisted selection
Mg^{2+}	Magnesium ions
$MgCl_2$	Magnesiumchloride
Mi	*Meloidogyne incognita* ((Kofoid and White) Chitwood)
min	minute(s)
Mj	*Meloidogyne javanica*
MP-PCR	Microsatellite-primed polymerase chain reaction
mRNA	Messenger ribonucleic acid
N	Nitrogen
NaCl	Sodiumchloride
NaOCl	Sodium hypochlorite
NaOH	Sodiumhydroxide
NIL	Near-isogenic lines
P	Phosphorus
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
pg	picogram
QTL	Quantitative Trait Loci
RAMP	Random amplified microsatellite polymorphism
RAPD	Random Amplified Polymorphic DNA
Rf-value	Reproduction factor
RFLP	Restriction Fragment Length Polymorphism
RIL	Recombinant inbred lines

RP	Recurrent parent
rRNA	Ribosomal ribonucleic acid
SAP	Specific amplicon polymorphism
SCAR	Sequence Characterized Region
SD	Standard deviation
SDS	Sodiumdodecylsulphate
sec	Seconds
SRFA	Selected restriction fragment amplification
SSR	Simple sequence repeat
STS	Sequence tagged site
T	Thymine
tec-MAAP	Template endonuclease-cleaved multiple arbitrary amplicon profiling
Tm	Melting point
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol
UPGMA	Unweighted pair group method with arithmetic mean
UPOV	International Union for the Protection of New Varieties of Plants
X-Gal	5-bromo-4-chloro-3-indolyl-ß-D-galactoside

TABLE OF CONTENTS

CHAPTER 3

CHAPTER 1

LITERATURE REVIEW

1.1 INTRODUCTION

Soybeans can be cultivated on a variety of soil types, but in South Africa it is limited to heavier soil types due to the high risk of nematode infection in lighter soil types. Only a few of the registered South African cultivars have some resistance to root-knot nematodes, but these cultivars are not representative of the wide variety of maturity groups needed. Root-knot nematodes can be controlled with nematicides, but these are expensive and harmful to the environment. It would be economically more viable to control nematodes with biological methods and there is thus an urgent need for more cultivars from other maturity groups with resistance to nematodes. This literature review will briefly touch the extent of the pathology and biology of nematode resistance in soybean.

Since the beginning of agriculture man has selected individual plants based on higher yield, better taste or appearance. The constant quest for something better and the ever increasing demand for better yield led to manipulation of the genetic material of wild types and domestication of all currently known crops - and the birth of plant breeding. Plant breeders have traditionally improved crop varieties based on selecting a desirable phenotype from the progeny of crosses made, accomplishing admirable success in a virtual vacuum of basic knowledge of the link between plant genetics and biology. The existence of this link was in fact exploited early in plant breeding - conventional breeding makes use of indirect selection on the basis of phenotypic markers linked to certain characteristics to follow these in a breeding program. Knowledge of the chromosomal location of a gene(s) affecting simple or complex traits can facilitate breeding efforts. Genetic markers have been used in plant breeding since the early 1900s and classical genetic maps were constructed through observation of phenological characters. The discovery of genetic linkage by Morgan (1911), i.e. that Mendelian genetic factors which lie close together on a chromosome are usually cotransmitted from parent to progeny, set the stage for the application in crop improvement (PATERSON, *et al.*, 1991). During the early 1970's the close linkage between genes for genetic male sterility and anthocyanin pigments in seedling leaves of sunflower was used extensively to produce hybrid seed in France and Romania (FICK, 1978). Fertile male plants carrying the anthocyanin pigments could be identified and removed easily from the population, to reduce the risk of cross pollination. The applications of these morphological characteristics is, however, very limited and over the years newer and more sophisticated techniques were developed for easier selection of desired traits. Mapped DNA markers

provide plant breeders with a tool to manage the genetic control of complex traits much more efficiently in breeding programs. It serves as a valuable link between plant breeding, plant genetics and biology.

DNA marker technology is developing at a dazzling pace and it is difficult to keep up with the newest technology and vast volume of literature. This review is an attempt to give a short overview of the technological developments of the past few decades. The methodology of compilation of genetic maps will be briefly discussed. It will also look into refining of these technologies for use as an economic alternative to the ever increasing cost and time aspects of plant breeding. The current applications in marker assisted selection (MAS), with the emphasis on inherent problems of soybean breeding and the potential of MAS in soybean improvement will be highlighted.

1.2 ISSUES CONCERNING THE NEMATODE PROBLEM

KINLOCH (1974) indicated that a yield loss of 53-90% was experienced in cultivars susceptible to *Meloidogyne incognita (Mi)* and a 32-40% loss in resistant genotypes. These figures were obtained in *Mi*-infested fields in Florida. About 29% of losses in soybean yield in the southern United States in 1984 was attributed to nematode damage (MULROONEY, 1986). Yield losses in soybean can be limited by using various methods, eg. utilization of nematicides, or crop rotation. Nematicides currently available in South Africa are however, extremely poisonous and harmful to the environment, and crop rotation is not very effective in the control of nematode populations in the field.

Meloidogyne javanica is widespread throughout South Africa and was found in 106 of 136 districts that were surveyed (KLEYNHANS, 1991). *M. incognita* (consisting of a variety of host races) showed a distribution in 67 out of the 136 agricultural districts investigated. *M. hapla* (45 out of 136) and *M. arenaria* (32 out of 136) are less abundant and are mostly found on horticultural plants. The major root-knot nematodes causing severe loss of soybean yield are *M. incognita* race 2 and 4 and *M. javanica*.

1.3 PATHOLOGY OF NEMATODE RESISTANCE

1.3.1 Biology and distribution

Nematodes comprise a large phylum of animals that includes plant and animal parasites as well as many free-living species. Plant parasitic nematodes are obligate parasites and live from nutrients obtained from

the cytoplasm of living plant cells (WILLIAMSON and HUSSEY, 1996). These nematodes can be either ectoparasites, living outside the host causing severe root damage, or endoparasites - either migratory or sedentary in the roots. Migratory nematodes move through the root, causing massive cellular necrosis (WILLIAMSON and HUSSEY, 1996). The most important economic damage in crop plants is caused by sedentary endoparasites of the family Heteroderidae. The Heteroderidae can be divided into two groups: the cyst nematodes, which include the soybean cyst nematode, *Heterodera glycines*, and the root-knot nematodes (genus *Meloidogyne*). The soybean cyst nematode is causing the most damage to soybean crops worldwide, especially in the USA, but fortunately has not yet been found in South Africa. The root-knot nematodes use thousands of plant species as hosts and cause severe losses in yield of many crops. Taxonomy of the genus was revised in 1949 by CHITWOOD. The root-knot nematodes of the genus *Meloidogyne* are more widely distributed throughout the world than any other group of plant-parasitic nematodes (SASSER, 1977). Three species of the genus, namely, *Meloidogyne incognita* (KOFOID and WHITE, 1919; CHITWOOD, 1949), *Meloidogyne javanica* (TREUB, 1885; CHITWOOD, 1949) and *Meloidogyne arenaria* (NEAL, 1889; CHITWOOD, 1949) are the most common and widespread in warm temperate, subtropical and tropical regions and are found between 35°South and 35°North latitudes. *Meloidogyne hapla* (CHITWOOD, 1949) is more prevalent in cooler climates. *Meloidogyne incognita* has been divided into four races and *Meloidogyne arenaria* into two races based on the North Carolina host differentials. No evidence could be found for host races among *M. javanica* and *M. hapla* populations studied (TAYLOR and SASSER, 1978). *M. javanica* appeared to be stable with reference to its reaction on the differentials, always infecting resistant tobacco, watermelon and tomato, and failed to attack cotton, pepper and peanut. Identity of a species can only be verified by microscopic examination of the perineal patterns of the nematodes.

The life cycle of root-knot nematodes involves a series of four juvenile stages. The second-stage juveniles (J2) are infective and penetrate the roots just above the root cap. They migrate intercellularly and establish a permanent feeding site. Salivary secretions from the dorsal oesophageal gland of the nematode trigger cellular responses such as synchronous nuclear division without cytokinesis, cell wall invagination and hypertrophy. The procambial cells adjacent to the head of the nematode develop into giant cells (syncitia), formed by enlargement (hypertrophy) of cells. Giant cells are multinucleate and have dense cytoplasm, with minute feeding tubes from which the nematode draws nutrients. These sedentary nematodes develop into adult females, producing thousands of eggs on the root surface in a protective, gelatinous matrix. They reproduce by mitotic parthenogenesis, generating clonal progeny by mitotic divisions. Swelling and division of cortical cells around the nematode lead to the formation of the characteristic galls on the plant

roots. The life cycle of root-knot nematodes is longer in soybean than in tomato and may be 39 days or longer (RIGGS and SCHMITT, 1987), depending on temperature. For *M. javanica* the life cycle time at 26.1°C was 21 days (TAYLOR and SASSER, 1978). Depending on the maturity group and planting date of soybean, as many as four generations of root-knot nematodes may be produced in one growing season.

1.3.2 Plant-nematode interaction

The rootknot nematode affects the function of the root system of plants and causes morphological and physiological changes in the roots. The upper parts of the plant show nonspecific symptoms of a defective root system. A series of complex changes in the host follows after infection of roots (WILLIAMSON and HUSSEY, 1996). Phytohormone levels are abnormally high in infected roots, as well as induction of several known plant defence genes (including peroxidase, chitinase, lipoxygenase and proteinase inhibitors)(WILLIAMSON and HUSSEY, 1996). Levels of extensin, a family of glycoproteins that form a major part of plant cell walls and are induced during defence responses, are significantly higher after infection with *Meloidogyne javanica* (NIEBEL, *et al.*, 1993). Extensin may play a role in the alterations in the feeding site or gall development. Strategies to resolve this complex response include identification of early changes in gene expression after infection (WILLIAMSON and HUSSEY, 1996) and the analysis of promoter elements and putative transcription factors using transgenic plants. Characterization of nematode secretions is also a major field of research, but is hampered by the minute amounts of exudate available.

Resistance to root-knot nematode species may be defined as some characteristic of a plant inhibiting reproduction of the nematodes. Tolerant plants have characteristics which reduce damage to growth or yield of a plant infected with these species (TAYLOR and SASSER, 1978). To be of real practical value in controlling root-knot nematodes, a resistant cultivar must be at least 90% effective against reproduction, compared to susceptible cultivars of the same species. Infective J2 of *Meloidogyne* species initially penetrated roots of resistant soybean cultivars in greater numbers than roots of susceptible plants (HERMAN, *et al.*, 1991), which precludes a physical barrier to penetration as a common form of resistance. However, 27% fewer J2 are present in roots of resistant plants 14 days after infection. This has been correlated with emigration of the nematodes from the roots (HERMAN, *et al.*, 1991). The stimulus for the migration of the nematodes from the resistant plants remains unknown. The resistant reaction in tomato cultivars produced little or no gall formation after *M. incognita* and *M. javanica* inoculation. However, some egg production was detected even in the absence of galling (ROBERTS and THOMASON, 1986).

Variability of reproduction of *M. javanica* isolates were also found on resistant tomatoes. Histological studies of 19 soybean cultivars infected with *M. incognita* revealed four types of response: (1) Formation of multinucleate giant cells with dense cytoplasm which were optimal for nematode reproduction; (2) Giant cells with thinner cell walls and less dense cytoplasm, which were less than optimal for reproduction; (3) Giant cells that are small with many inclusions, and were associated with poor nematode reproduction; (4) Marked necrosis of cells around the head of the larva, without larval development (DROPKIN and NELSON, 1960). There seems to be a difference in resistance to nematode reproductive ability and gall formation, and different genes were found associated with the different physiological responses in groundnut (GARCIA, *et al.*, 1996). Resistance to reproduction of nematodes can be monitored by counting the eggs harvested from a plant after at least two to three nematode life cycles (approximately 60 days) and the level of gall formation can be used as measure of host response (GARCIA, *et al.*, 1996). Larvae infecting roots of resistant plants may :(1) develop to mature females, but with no viable egg production; (2) develop to mature males; (3) have arrested development; (4) be killed by an immune reaction; or (5) leave the root and infect another root (TAYLOR and SASSER, 1978).

Root galling after infection reduces the cross-sectional area of xylem vessels and this restricts water flow (SIJMONS, 1993). Heavily infected plant roots are shorter than uninfected roots, with less branch roots and root hairs. Root deformity and inefficiency cause stunted growth, wilting in dry weather and other symptoms typical of shortage of water and nutrients. Wilting of infected plants are often seen in fields during hot, dry weather. It seems that alternate high and low soil moisture reduces the efficiency of galled roots, but even heavily infected plants can grow fairly well if irrigated frequently (TAYLOR and SASSER, 1978). Loss of root efficiency leads to a reduction in growth of the plant with a reduction in yield. Infection of plants with root-knot nematodes also lead to secondary infections by bacteria, fungi and viruses. The physiological changes in host plants due to nematode infection may be responsible for alteration of plant susceptibility to other pathogens.

1.4 HISTORICAL DEVELOPMENT OF MARKER TECHNOLOGY

One of the first attempts to associate biochemical quality with phenotype or performance were with protein profiling and isoenzyme analysis, which were employed successfully for linking desirable traits to specific protein bands, and are still widely employed in varietal identification and breeding programs. Although these proteins are transcripts of the DNA, environmental factors can profoundly influence both quantitative and qualitative levels of expression. These factors can detract from the reproducibility of the technique

and can mask the genotype of the plant. The number of genetic markers provided by isoenzyme analysis is also insufficient. The discovery of restriction enzymes in 1973 and subsequent development of recombinant DNA technology has had an enormous impact on the science of plant breeding (TANKSLEY, *et al.*, 1989). The potential impact of restriction fragment length polymorphism (RFLP) mapping on eukaryotic genetics was first described in 1980 by researchers in human genetics (BOTSTEIN, *et al.*, 1980).

In plant breeding, crop improvement is achieved by years of direct and indirect selection of desirable traits. One of the problems facing the plant breeder is the uncertainty of whole plant assessment as an indicator of genetic potential. Segregating progeny obtained from a cross between two parents are mosaics of both parents and carry desirable as well as undesirable traits from both parents. The technique of backcross breeding is used to recover the more desirable phenotype of the recurrent parent after introduction of a specific trait from the donor parent. The desired phenotype has to be selected in several cycles of backcrossing and is often difficult to identify under uncontrolled conditions. The process of selection would be much improved if whole plant assessment could be backed by direct analysis of the genetic composition of the plants through the use of molecular markers. Molecular markers exploits the discovery that Mendelian genetic factors which lie close together on a chromosome are usually cotransmitted from parent to progeny. If the desired gene(s) are tightly linked to a DNA marker, the segregating population of plants can be screened in the seed or seedling stage - before the trait is expressed - for the presence of the gene(s) of interest (TANKSLEY, *et al.*, 1989). A major application can be found in breeding for disease resistance. Simultaneous screening of plants for resistance to several different pathogens can normally be difficult and impractical. Breeders are also unable to screen for resistance to new pathogens because of quarantine restrictions. It is now possible to detect resistance genes by means of their linkage to DNA markers and minimize the need to inoculate the plants with the pathogen (TANKSLEY, *et al.*, 1989). Knowledge of the RFLP genotypes of many loci throughout the genome yields an estimate of the composition of an individual's chromosomes in terms of its parents. This not only shows which portions of the genome are derived from each of the parents, but also the regions where crossovers took place. Estimations based on computer simulations indicate that a recurrent parent genotype can be reconstructed in only three generations of 30 individuals each based on whole genome selection (TANKSLEY, *et al.*, 1989).

Our ability to perform linkage analysis is affected by the level of variation in a species, which differs markedly between species (BURR, 1994). HELENTJARIS, *et al.* (1985) compared the level of

polymorphism on the basis of RFLP among a number of plant species and noted that self-pollinating species showed much less variation than out-crossing species. This is consistent with the hypothesis of NEI and LI (1979) that heterozygosity is a function of the effective species population size and the mutation rate, as it could be expected that self-pollination would reduce the effective population size (BURR, 1994).

1.4.1 Mapping populations

The choice and development of a suitable mapping population for the construction of molecular maps is critical and will depend on the breeding system of the particular plant (KOCHERT, 1994). The goal of the mapping will determine which parents should be chosen for crossing, the size of the population, how the cross is advanced and which generations will be used for DNA and phenotypic analysis (YOUNG, 1994). In order to identify markers for a specific trait the two parents must differ significantly for the trait of interest, with the ideal situation where no recombination occurs between the marker locus and the quantitative locus (SOLLER, *et al.*, 1976). If one or both of the lines to be crossed are not at fixation for alternative alleles of the quantitative trait locus, the differences between marker genotypes will be less than in the situation of complete fixation (SOLLER, *et al.*, 1976). The choice of population will also be influenced by the number of genes involved, i.e. a single dominant gene or several genes underlying quantitative traits. It is also critical that sufficient DNA sequence polymorphism exists between the two parents. Most RFLP maps have been constructed using F_2 populations or backcross populations derived from crosses between inbred parent lines. It is possible to reconstitute F_2 mapping populations by selfing and growing F_3 plants for DNA extraction, combining material from several of the F_3 plants. Recombinant inbred (RI) lines can be developed by selfing individual F_2 or backcross plants for six or more generations, using single seed descent. The resultant RI lines will be largely homozygous and can be propagated by seeds for use by other laboratories as a mapping population. When a map containing 100-200 well-dispersed markers has been constructed, virtually any new marker will be linked to one previously mapped (KOCHERT, 1994).

The population size is of great importance and a rather controversial issue. The traditional approach for mapping quantitative traits involved relatively large numbers for progeny testing. According to HANSON (1959) the minimum size of a backcross or testcross population for the determination of linkage should be 25-35 if the recombination value is 0.2 with a degree of independence of 0.05 or 0.025 respectively. SOLLER, *et al.* (1976) studied the detection of linkage between a quantitative locus and a marker locus which were both segregating in a backcross or F_2 population. They found that the backcross design was

more sensitive when the dominant quantitative locus was linked to the dominant marker locus, but the F_2 design was preferable in the absence of dominance. In most cases the F_2 design required fewer offspring than the backcross design and a total of about 2000 offspring would have been sufficient. In contrast, later experimental designs used much smaller populations for mapping, averaging between 60 and 100. MICHELMORE, *et al.* (1991) used two F_2 populations of lettuce comprising 66 and 80 individuals respectively. WANG and PATERSON (1994) suggested that QTLs with phenotypic effects of 0.75-1.0 SD (standard deviation) or larger should be detectable in backcross-, F_2- and recombinant-inbred populations of 100 to 200 plants. WANG, *et al.* (1991) used 60 F_2 plants for mapping rye, and SONG, *et al.* (1991) used 95 F_2 individual cabbage plants for construction of a detailed linkage map of *Brassica rapa* (syn. *Campestris*).

The resolution of the map and the ability to determine marker order is largely dependent on population size. Clearly, the larger the population the better, but it may be limited by technical problems eg. the number of seeds available, the number of DNA samples that can be analysed reasonably with a specific technique, or the phenotypic screening of individual plants. The number of progeny scored with DNA markers can be substantially reduced by bulked segregant analysis (BSA) (MICHELMORE, *et al.*, 1991) or selective genotyping (LANDER and BOTSTEIN, 1989), used alone or in combination. Selective genotyping involves growing a larger population, but genotyping only the individuals with phenotypes that deviate the most from the mean of the population. In a population with a continuous variation in phenotype, about 5% of the total population will have phenotypes more than 2SD from the mean, contributing about 28% of the total linkage information (LANDER and BOTSTEIN, 1989), and 33% will have phenotypes 1SD from the mean, contributing 81% of linkage information. By growing a larger population and genotyping only the extremes, the same total linkage information could be obtained.

It is not always appreciated that the success of marker development is totally dependent on the ability for accurate evaluation of the phenotype of the individual plants in the mapping population. As the main reason for the development of a marker for a specific trait is often the fact that the trait is difficult to evaluate, it can influence the success of obtaining reliable linkage between the trait and marker.

1.4.2 Bulked segregant analysis (BSA)

The classical method to identify or map specific genes made use of near isogenic lines (NILs), which differ theoretically only for the gene under investigation. NILs are produced by repeated backcrossing of the

progeny to one of the parents for several generations to produce a line that is theoretically identical to one of the parents except for the specific allele introduced. This approach can take several years and is very time-consuming and expensive. A valuable and much used alternative procedure simulating NILs known as bulked segregant analysis was proposed by MICHELMORE, *et al.* (1991). As for selective genotyping, this method also makes use of plants from the extremes of the phenotypic spectrum of a segregating population, but differs in that it pools the DNA from these plants in two bulks. F_2 lettuce plants segregating for resistance to downy mildew were divided into two groups which were homozygous for resistance and susceptibility. Heterozygotes were identified by progeny tests and excluded from the bulks. DNA from several plants from each group was pooled for analysis. Each bulk contained individuals that were identical for a particular trait or genomic region, but seemingly heterozygous at all other regions. The principle of DNA pooling is the grouping together of informative individuals in order to study a selectable marker linked to a particular gene of interest in a randomised genetic background of unlinked loci (WANG and PATERSON, 1994). BSA does not reveal novel types of variation, but allows the rapid screening of many loci and therefore the identification of informative polymorphisms (MICHELMORE, *et al.*, 1991). The technique can also be applied to other types of populations. HALEY, *et al.* (1993) used this technique in a backcross population that segregated for a gene for rust resistance in dry beans. This combination of backcrossing and bulked segregant analysis strengthened the identification of markers tightly linked to the gene under investigation (MICHELMORE, *et al.*, 1991). Markers could be reliably identified in a 25cM window either side of the targeted locus. With this technique they identified three RAPD markers in lettuce linked to a gene encoding resistance against downy mildew.

CHAPARRO, *et al.* (1994) used a combination of RAPD markers and bulked segregant analysis to compile a genetic map of peach. This method was very effective in the mapping of specific loci. The DNA of 8-12 F_2 seedlings was bulked for analysis. Polymorphisms were confirmed through comparison of the phenotypes of the parents and segregation analysis of 96 F_2 plants that were not included in the bulks. Polymorphic fragments differing significantly from the expected 3:1 relationship in the 96 F_2 plants were not used for mapping purposes. The results indicated that less RAPD polymorphisms could be identified if heterozygote plants were included in the bulks. If heterozygotes for a dominant phenotype were included in the bulks, the RAPD markers linked to the recessive allele could not be observed. The F_2 progeny of peach comprises a mixture of dominant markers in linkage or repulsion and result in homologue specific maps (two maps per chromosome). For the meaningful use of RAPD markers in the selection of specific characteristics, these maps should be lined up. This can be done with testcrosses, the use of single band

RAPD markers as RFLP probe (codominant markers) or genotype analysis of F_2 and F_3 progeny (CHAPARRO, *et al.*, 1994).

MIKLAS, *et al.* (1996) used a combination of the methods of BSA and selective genotyping in a stepwise fashion for the identification of QTLs conditioning disease resistance in common bean. Their mapping strategy followed a five-step process: (1) RAPDs polymorphic between the two parents were identified, (2) DNA bulks from the extremes were tested with these polymorphic primers, (3) individual lines within the bulks were characterized with RAPDs amplified between the bulks, (4) RAPDs that cosegregated with disease reaction in at least five of six (83%) of the lines were mapped in the entire population, and (5) selectively mapped markers and mean disease scores for each line were regressed to ascertain RAPD-QTL associations.

The size of a backcross population and the amount of individuals included in the bulks for analysis differs between various authors. HALEY, *et al.* (1994) evaluated 70 individual F_2 plants for rust resistance and made separate bulks of four resistant and four susceptible plants. They did not do any progeny testing *a priori* to distinguish homozygous and heterozygous plants. WANG and PATERSON (1994) recommended separate bulking of ten plants of the phenotypic extremes to avoid detecting false positive markers. It is also very important to use equal amounts of DNA from each individual. However, MIKLAS, *et al.* (1993) used as little as three plants per bulk with excellent results, with only one out of 931 fragments amplified by 167 primers being polymorphic.

A similar but alternative approach for DNA pooling was followed by GIOVANNONI, *et al.* (1991). They described a method where DNA pools from a segregating population could be constructed on the basis of mapped molecular markers. A target segment of the genome was selected which contains the gene of interest or the segment of the genome where more markers are required. Two DNA pools are constructed consisting of F_2 plants homozygous for the markers flanking the segment of interest from the two parents. Homozygous pools can also be compared to heterozygous pools. A sufficiently large number of individuals in each pool assures that the pools are essentially homogeneous for all genomic loci except those adjacent to the target interval. Proof of the localisation of the marker is obtained through segregation analysis in the individual plants (GIOVANNONI, *et al.*, 1991). GIOVANNONI, *et al.* (1991) used tomato as a test species and combined equal amounts of isolated DNA of seven plants for the homogeneous pools and 14 plants for the heterogeneous pools. The basic difference between this method and the bulking method of MICHELMORE, *et al.* (1991) lies in the possible applications. Basing the pooling

strategy on phenotype selects a single genetic point in a population segregating for the target trait. Selection and use of DNA pools based on existing marker data can target a defined genomic interval for filling in gaps in the map with more markers, or to isolate markers in intervals likely to contain genes of interest. Different combinations of individuals from the same F_2 population can be used to target any interval in the genome. The number of individuals used in the pools is important. Larger numbers of individuals increases the probability that the two pools will only differ in the target region, but also the probability that individuals will occur in the pool with a double crossover within the target interval. GIOVANNONI, *et al.* (1991) recommended a pool size of more than five. REITER, *et al.* (1992) used this approach to construct a high density map in a selected genome region of *Arabidopsis thaliana* with recombinant inbred lines. They constructed two DNA pools from six different RI lines each and identified 23 RAPD polymorphisms mapping to chromosome 1, four mapped to other regions and an additional five polymorphisms did not segregate in Mendelian fashion.

1.4.3 Quantitative Trait Loci (QTL)

Quantitative traits are characterized by complex inheritance patterns with continuous distribution of phenotypes in segregating populations (CONCIBIDO, *et al.*, 1996b). This variation can be explained by independent actions of many discrete genetic factors, each contributing a relatively small effect to the overall phenotype. This makes breeding and analysis of the genes involved extremely difficult. Several powerful DNA marker analyses allow for the resolution of multigenic traits into individual Mendelian components (PATERSON, *et al.*, 1988), and can be utilized to obtain a high-resolution map around a quantitative trait locus, such as bulked segregant analysis (MICHELMORE, *et al.*, 1991), comparative genome mapping (TANKSLEY, *et al.*, 1988) and integrated mapping (STAM, 1993). Comparative genome mapping is a strategy that utilizes mapping information from one taxon to predict linkage relationships in related taxa. This affords many benefits to crop genome analysis, including greater utility of existing DNA probes, effectively increasing the density of genetic markers in many species simultaneously. It also offers new opportunities for studying plant evolution. Recent results suggest that it may have even greater utility than previously envisioned, reaching directly into the molecular dissection of complex traits that are the basis of agricultural productivity (PATERSON, 1996). Integrated or "Join" mapping integrates linkage maps that were developed in independent populations. Genes and markers that do not segregate in different mapping populations can thus be placed onto a joint map by combining information from multiple mapping populations (CONCIBIDO, *et al.*, 1996b). Comparative genome mapping with RFLPs from mungbean and common bean was used to increase marker density on linkage group G of soybean, leading to one RFLP

marker every 2.6 cM (centimorgan - 1 cM is defined as the distance along the chromosome which gives a recombination frequency of one percent).

High density molecular genetic maps make it possible to resolve complex traits into their individual genetic components (TANKSLEY, *et al.*, 1989). Linkage of a DNA marker to a QTL is done by making a cross between two plants differing for one or more characters. Segregating progeny (F_2, backcross or recombinant inbred lines (RIL)) are obtained and evaluated for the character of interest as well as for their genotypes at DNA marker loci throughout the genome.

Associations between the trait and the segregating markers are done with statistical methods, which were developed and refined by several authors since the early 1900s. The statistical procedures used in identification of linked QTL and the compilation of genetic maps received considerable attention, and evolved through several types of analyses. Genetic mapping of QTLs is based on the simple idea that genetic markers that tend to be transmitted together with values of the trait are likely to be close to a gene affecting that trait. Thus an association is sought between marker variants and trait values, with higher levels of association suggesting closer genetic map distance (DOERGE, *et al.*, 1994). The ability to detect a QTL with an RFLP marker is a function of the magnitude of the effect of the QTL on the desired character, the size of the population studied, and the recombination frequency between the marker and the QTL (TANKSLEY, *et al.*, 1989). SOLLER, *et al.* (1976) have shown that it should be possible to detect a codominant QTL responsible for 1% of the F_2 phenotypic variance in a population of 1000 individuals. The authors used analysis of variance in their calculations. WELLER (1986) presented a statistical analysis using maximum likelihood techniques for mapping of QTLs and the estimation of parameters under conditions of partial linkage. He studied the segregation of plant height and an esterase marker in a tomato F_2 population of 1596 progeny. He postulated that a codominant QTL of 1SD will be responsible for about 10% of the phenotypic variation in the population. The method of WELLER (1986) was not effective in distinguishing between complete and partial linkage in samples of only 500 individuals or for quantitative loci with effects less than one phenotypic standard deviation. The method was also more effective for codominant than for dominant loci. LUO and KEARSEY (1989) modified the maximum likelihood method to be used with a smaller sample size of 500 plants when the heritability of the quantitative trait is not less than 0.1.

The work with the greatest practical impact on statistical issues of marker association has been that of LANDER and BOTSTEIN (1989). They described a set of analytical methods that modified and extended

the classical theory for mapping QTLs. The traditional approach to mapping QTLs involved studying single genetic markers one at a time, and made use of analysis of variance for detection of linkage between loci. This method lead to an underestimation of the phenotypic effects of QTLs, the genetic locations of the QTLs were not well resolved and the number of progeny was larger than necessary. It could not distinguish between tight linkage to a QTL with small effect and loose linkage to a QTL with large effect. These problems could be overcome by adapting the method of likelihood of the odds (LOD) scores applied in human genetics to interval mapping of QTLs (LANDER and BOTSTEIN, 1989). The traditional approach with linear regression analysis of phenotype on genotype, is a special case of the method of maximum likelihood. The evidence for a QTL are given by the LOD score which indicates the probability of the data to have arisen assuming the presence of a QTL than assuming its absence (LANDER and BOTSTEIN, 1989). If genetic markers have been scored throughout the genome, the method of maximum likelihood can estimate the phenotypic effect and the LOD score for a putative QTL at any given genetic location. Thus, at each position in the genome, one computes the 'most likely' phenotypic effect of a putative QTL affecting a trait (the effect which maximizes the likelihood of the observed data arising) and the odds ratio (the chance that the data would arise from a QTL with this effect, divided by the chance that it would arise given no linked QTL (PATERSON, *et al.*, 1988). The LOD score, defined as the $\log_{10}$ of the odds ratio, summarizes the strength of evidence in favour of the existence of a QTL with this effect at this position. If the LOD score exceeds a pre-determined threshold, the presence of a QTL is inferred. The threshold for the LOD score depends on the size of the genome and the density of markers genotyped. Interval mapping allows inference about points throughout the genome and avoids confounding phenotypic effects with recombination, by using information from flanking genetic markers (PATERSON, *et al.*, 1988). This method (interval mapping) combined with selective genotyping of individuals from the extremes can decrease the number of progeny seven-fold (LANDER and BOTSTEIN, 1989). When the selective genotyping approach is followed, standard linear regression procedures cannot be followed as the biased selection of progeny would lead to a gross overestimation of phenotypic effects. Specially developed programs like MAPMAKER/QTL (LINCOLN, *et al.*, 1992), which is the most widely used program today can analyse these data with the genotypes of non-extreme individuals entered as missing data.

1.4.4 DNA based molecular techniques

Molecular techniques currently available include Southern Blot based RFLP (restriction fragment length polymorphism) and PCR-based analysis methods, which include RAPD (random amplified polymorphic DNA), AFLP (amplified fragment length polymorphism) and microsatellite analysis, as well as quite a few

derivations of these techniques. These three techniques are most frequently used, with other techniques derived from these techniques, also found in some applications. Both RAPD and AFLP methods result in mainly dominant markers, i.e. bands present or absent, whereas RFLP results in a codominant marker with bands differing in size.

In 1996 LU, *et al.* compared the use of RFLP with various PCR-based techniques (RAPD, AFLP and microsatellite-AFLP) regarding their informativeness and applicability for genetic diversity analysis. Among the ten genotypes studied, the PCR-based methods proved to be much more informative than cDNA-RFLP analysis. Approximately two-thirds of randomly chosen cDNA probes detected at least one difference between the ten pea lines studied. The PCR based methods gave similar high levels of polymorphism (47-68%). In genetic diversity comparison, the trees from all the molecular marker techniques were significantly correlated, and agreed with the tree formed from RFLP data. It was concluded that the PCR-based methods could be used as alternatives to replace RFLP in genetic diversity assessment. The number of markers used influenced the assessment of genetic diversity. Precision improved as more marker loci were detected (LU, *et al.*, 1996). The number of bands required to estimate genetic distance among genotypes, will be a function of the genetic relationship among the genotypes in a database, as discrimination among more closely related individuals will likely require more bands than discrimination among distantly related individuals (TIVANG, *et al.*, 1992).

In 1997 a network of laboratories in Europe tested the reproducibility of RAPD, AFLP and SSR (JONES, *et al.*, 1997). They found RAPD difficult to reproduce, with small differences obtained in the sizing of SSR markers. AFLP was found to be as reproducible as RFLP patterns. The various marker systems were compared for two metrics: (1) Expected heterozygosity, which is a function of the marker system's ability to distinguish between genotypes. (2) The multiplex ratio of a marker system, which defines the number of loci (or bands) simultaneously analysed per experiment (POWELL, *et al.*, 1996). SR markers had the highest expected heterozygosity, while AFLP markers had the highest effective multiplex ratio. Genetic similarity estimates of soybean lines were highly correlated based on RFLP, AFLP and SSR marker data. RAPD data produced higher estimates of interspecific similarities (POWELL, *et al.*, 1996).

1.4.4.1 Restriction Fragment Length Polymorphism (RFLP)

RFLPs have their origin in (1) base sequence changes which add or eliminate restriction sites, or (2) DNA rearrangements by insertion or deletion of pieces of DNA, and are naturally occurring, simply inherited,

Mendelian characters (KOCHERT, 1994). The most RFLP variability in plants is apparently caused by genome rearrangements (KOCHERT, 1994) for which evidence is derived from (1) observations that 6-cutter restriction enzymes reveal more polymorphisms than do 4-cutter enzymes, (2) the amount of RFLP variation detected correlates with the average length of fragments produced by that enzyme, and (3) RFLPs detected by one enzyme tend to be detected by multiple enzymes. The variability found with RFLPs in a given species could be correlated with variability found previously with isozyme marker analysis (CHASE, *et al.*, 1991; HELENTJARIS, *et al.*, 1985). Less variability was observed between species and inbred varieties of tomato than inbred lines of maize. This could be due to the fact that one species is self-pollinating while the other is usually cross-pollinated (HELENTJARIS, *et al.*, 1985). Both cDNA and random genomic libraries have been used as sources for RFLP probes. Single or low copy number clones are most useful for RFLP map construction. Repeated sequences are present in high copy number and often appear as smears on autoradiograms. Low copy number probes will result in fewer bands on the autoradiogram, but if the map is transferred to a different segregating population, a different subset of bands could be polymorphic and the chromosomal location of these would be unknown (KOCHERT, 1994). *PstI* digestions of genomic DNA in library construction can partially eliminate repeated sequences as these sequences are often methylated at cytosine sequences. *PstI* is sensitive to methylated sequences and will not cleave at these sites, which would leave the repeated sequences as large fragments ligating poorly into plasmids and transforming with low efficiency. The advantage of a cDNA library is that you are mapping actual genes as well as intervening sequences (KOCHERT, 1994). RFLP markers normally behave in a codominant manner and are apparently free of epistatic effects. They are phenotype neutral and a virtually limitless number of markers can be monitored in a single population (TANKSLEY, *et al.*, 1989). In maize it was found that a number of different alleles could be detected at a single locus when a clone was hybridized to genomic DNA of several different maize lines (HELENTJARIS, *et al.*, 1985). The heterozygous hybrids possessed all of the fragments. One clone could also detect more than one independently segregating locus by cross hybridization to related sequences at other loci.

Mapping by RFLP analysis consists of DNA isolation from a suitable set of plants, digestion of the DNA with various restriction enzymes, separation of the restriction fragments by agarose gel electrophoresis, transfer of the separated fragments to a filter by Southern Blotting (SOUTHERN, 1975), detection of individual restriction fragments by nucleic acid hybridization with a probe (radioactively labelled or detected with chemiluminescent methods), and scoring of RFLPs by direct observation of autoradiograms.

RFLPs are being used extensively in the production of high density molecular maps of a wide diversity of crops, determination of phylogenetic relationships between varieties within a species, interspecific relationships, genomic evolution studies, marker assisted selection in breeding programs for selecting progeny for specific traits, and many more (BONIERBALE, *et al.*, 1988).

The construction of a RFLP linkage map is based on the estimation of recombination frequencies between genetic loci and on the determination of the linear order of loci in linkage groups (RITTER, *et al.*, 1990). The distance on a linkage map between two markers is determined by measuring the recombination frequency. Linked markers are then aggregated in linkage groups. The number of linkage groups is equivalent to the chromosome number for that species (RITTER, *et al.*, 1990). Most RFLP maps have been obtained from segregating populations, F_2 and/or backcross populations derived from homozygous inbred lines, but also from crosses between heterozygous parents.

Over the past few years RFLP maps have been developed for all the most important crop species. These include maps for crops with long histories of genetic studies such as maize (HELENTJARIS, 1987; COE and GARDINER, 1994), soybean (SHOEMAKER, 1994), tomato (TANKSLEY, 1994a), and rice (TANKSLEY, 1994b). Maps are also being developed for less studied crops such as lettuce (LANDRY, *et al.*, 1987), potato (BONIERBALE, *et al*, 1988), *Brassica rapa* (SONG, *et al.*, 1991), rye (WANG, *et al.*, 1991), peanut (HALWARD, *et al.*, 1994) and common bean (VALLEJOS, 1994) to name but a few.

Numerous examples of the use of RFLP markers to map important characteristics of plants can be found in the literature, eg. disease resistance - markers for wheat powdery mildew resistance genes *Pm1* and *Pm2*, which were tagged using near-isogenic lines (HARTL, *et al.*, 1995). Three genetic loci were identified for control of resistance to wheat streak mosaic virus in maize (McMULLEN, *et al.*, 1994), or morphological characteristics, eg. hard seededness of soybean (KEIM, *et al.*, 1990b) and seed protein and oil content in soybean (DIERS, *et al.*, 1992a).

Conventional RFLP analysis is limited by several factors. Firstly, it requires a relatively large amount of DNA for restriction digestion. Secondly, the analysis is relatively slow and expensive. By comparison to polymerase chain reaction-based techniques, RFLP markers would be too labour- and time-intensive to be practical in screening large numbers of individuals required by marker-assisted selection (MUDGE, *et al.*, 1997) in a breeding population. Thirdly, the maintenance and distribution of probes has proven to be

time consuming and often error prone. Fourthly, the level of variability in a species must be readily detectable by this method (HELENTJARIS, *et al.*, 1985).

1.4.4.2 Polymerase-Chain-Reaction-based mapping methods (PCR)

The discovery of a temperature tolerant bacteria *Thermus aquaticus* and the subsequent isolation of the temperature resistant DNA polymerase, caused a dramatic explosion of techniques and applications based on the polymerase chain reaction (SAIKI, *et al.*, 1988). These techniques include RAPD (random amplified polymorphic DNA), AFLP (amplified fragment length polymorphism), DAF (DNA amplification fingerprinting) - (CAETANO-ANOLLES, *et al.*, 1991), MAAP (multiple arbitrary amplicon profiling) (CAETANO-ANOLLES, *et al.*, 1993b) and IRA or SSR (inter-repeat amplification or simple sequence repeats) (ZIETKIEWICZ, *et al.*, 1994).

PCR amplification can be used to generate DNA fragments suitable for use as genetic markers. The principle of the technique involves the use of two primers, recognising sites some distance from one another and the amplification of the sequence flanked by them. Any variation in the size of the fragments generated will have to result from changes occurring between the two primers in a space of a few kilobases. PCR primers can be synthesized with specific sequences as used in minisatellites or microsatellites, or it can be of random sequence as used in RAPDs and AFLPs.

Advantages of application of PCR-based techniques in marker detection include the ease of separation and detection of amplified products, usually without the need for radioisotopes. Dissemination of sequence-based methods only involves publication of the DNA sequence along with the mapping results. Individual laboratories can then synthesise their own oligonucleotides (KOCHERT, 1994). Other advantages are the minute amounts of template DNA necessary for analysis, the speed with which results can be obtained and the large number of loci detected (LU, *et al.*, 1996). The techniques are also more sensitive, which makes it possible to detect low-frequency polymorphism. The major advantage of the PCR-based methods is that they can easily be applied to a large number of samples and can be automated (LU, *et al.*, 1996).

1.4.4.2(i) Microsatellites (SSR, Simple Sequence Repeats)

The use of minisatellites in genomic fingerprinting was discovered in by JEFFREYS, *et al.* (1985). Since then numerous variations of the technique evolved. JEFFREYS, *et al.* (1985) originally developed a probe

based on a tandem-repeat of a core sequence which could detect many highly variable loci simultaneously and could provide an individual-specific DNA fingerprint for use in human genetic analysis. Oligonucleotide probes comprised variations of the GAT(C)A simple repeats, with an optimal length of 20 bases (ALI, *et al.*, 1986). Simple sequences were apparently non-coding. They consisted of tandemly organized short repetitive DNA which tended to be hypervariable in copy number (SCHÄFER, *et al.*, 1988). Additional synthetic probes were constructed consisting of $(CT)_8$, $(CAC)_5$ and $(TCC)_5$. Genomic DNA was digested with various restriction endonucleases, followed by separation of the fragments with agarose gelelectrophoresis and Southern hybridisation with ^{32}P-labelled probes. The technique was adapted to use non-radioactive detection methods and the use of these probes were also exploited for fingerprinting of plant and fungal genomes (BIERWERTH, *et al.*, 1992). The technique was soon extended to other repetitive DNAs, and three groups of markers could be distinguished: (1) M13 repeat probes, (2) simple repetitive sequences and (3) minisatellite DNAs (LU, *et al.*, 1996). Due to their repetitive nature, all of these markers generated complex banding patterns after Southern hybridisation.

Microsatellites are tandem repeats of short sequences (2-6 bp) such as $(GT)_n$ or $(CAC)_n$. The allelic variability is a result of different copy numbers of the tandem repeat at different alleles of the same locus. The use of microsatellites was well established for application in human and mammalian genetics, but their practicality in plants was also demonstrated (AKKAYA, *et al.*, 1992). This study demonstrated the segregation of microsatellites as codominant markers and the prevalence of polymorphism in soybean. The use of microsatellites as probes in hybridisation to genomic DNA led to complex patterns which had great application in DNA fingerprinting.

The preferred method used in mapping studies was to design PCR primers complementary to single copy DNA flanking the repeated element, according to available sequence data. The DNA sequences flanking microsatellites are generally conserved within individuals of the same species, allowing selection of PCR primers that will amplify the intervening SSR in all genotypes (AKKAYA, *et al.*, 1995). This means that sequencing data of the genome had to be known for the design of the primers. The amplification products were analysed for length differences by electrophoresis, usually on a sequencing gel. Microsatellite markers were developed for diverse crop species including soybean, maize, wheat, brassica, barley, grape, sunflower, avocado and *Arabidopsis thaliana* (AKKAYA, *et al.*, 1995). SSR loci could be linked to several traits in soybean and were found to be relatively randomly distributed throughout the genome, although a limited amount of clusters were detected. In *Phaseolus* it was observed that all of the di-nucleotide repeats and most of the tetra-nucleotide repeats were primarily found in non-coding regions of

the genome. Most of the tri-nucleotide repeats were found in coding regions (YU, *et al.*, 1998). Mendelian segregation was confirmed in a F_6 recombinant inbred population.

The less costly and more widely available agarose gel system for detection of SSR polymorphism was successfully used in a study of 94 elite maize inbred lines (SENIOR, *et al.*, 1998). A special agarose (Metaphor, FMC Bioproducts) were used at a concentration of 4% (m/v). A cluster analysis placed the inbred lines in nine clusters that corresponded to major heterotic groups of market classes for North American maize. A unique fingerprint for each inbred line could be obtained from as few as five SSR loci.

There were several advantages using microsatellite markers: (1) They were codominant and (2) PCR based which meant that automation of analysis was possible. (3) They were multi-allelic and hypervariable and (4) appeared to be randomly and uniformly distributed throughout eukariotic genomes. (5) They were accessible to other research laboratories via published primer sequences (YU, *et al.*, 1998).

The major disadvantage of microsatellites is the cost of establishing polymorphic primer sites (BURR, 1994), as sequence data of the genome was essential. The method is also tedious, time consuming and required the use of radioisotopes.

An automated SSR system using fluorescent labelling of alleles was developed for cultivar identification of soybean (DIWAN and CREGAN, 1997). Twenty loci successfully distinguished modern soybean cultivars that were identical for morphological and pigmentation traits, as well as seven genotypes that were indistinguishable with RFLP probes. Pedigrees of seven cultivars were studied to estimate stability of SSRs across generations. Six of these had one locus in the progeny with an allele(s) that was not present in either parent. These new alleles were most likely the result of mutation (DIWAN and CREGAN, 1997). In order to avoid difficulty with mutation, it was recommended that DNA fingerprint data should be determined from the bulk of 30-50 plants of a cultivar.

The SSR derived primers could also be used in a PCR reaction for the amplification of genomic DNA between the repeat sequences, called microsatellite-primed PCR (SSR-PCR) (ZINK and NAGL, 1996). This technique was applied to different *Phaseolus* species. The plant genomic DNA was restricted with the enzyme *HaeIII* before amplification with single synthetic oligonucleotide primers complementary to repeat sequences. The amplified products were electrophoresed in agarose gels, stained with ethidium

bromide. This technique resulted in RAPD-like banding patterns and seemed to be less sensitive than the conventional DNA fingerprinting (ZINK and NAGL, 1996).

Anchored microsatellites (or ISSR-PCR - inter-simple sequence repeat-PCR) made use of primers complementary to simple sequence repeats (SSRs) with variable three-base 'anchors' at their 5' end (CHARTERS, *et al.*, 1996). No prior sequence knowledge was required. The use of this technique was evaluated in various oilseed rape cultivars. Amplification products were separated on polyacrylamide gels and detected with silver nitrate staining. The use of only two primers could discriminate 16 of the 20 cultivars studied (CHARTERS, *et al.*, 1996). It was concluded that the anchored SSR-PCR was highly informative and reproducible in fingerprinting oilseed rape populations, although intra-cultivar variation should be investigated before using banding profiles from pooled samples. The technique was also found to be very effective for analysis of potato cultivars (PREVOST and WILKINSON, 1999), used together with a horizontal electrophoresis of PCR products in pre-cast polyacrylamide gels and stained with silver nitrate. The profiles generated were highly reproducible. As few as two primers were able to distinguish all potato cultivars.

Cloning and molecular characterization of products obtained by RAMP (random amplified microsatellite polymorphism) and MP-PCR (microsatellite primed-PCR, SSR-PCR) elucidated the sequences responsible for the polymorphisms generated (DÁVILA, *et al.*, 1999). RAMPs were produced by using decamer 5'-anchored oligonucleotides containing a dinucleotide repeat (e.g. $GC(CA)_4$ OR $CCGG(AC)_8$) in combination with arbitrary decamer primers (the same as used in RAPD). When 5'-anchored oligonucleotides were used as primers, the polymorphism was produced in the variation of the number of repeats of the core sequence at each locus. When 3'-anchored oligonucleotides were used, polymorphism was attributed not to variation at the priming site, but to the variation of the inter-repeat sequence.

1.4.4.2(ii) Random Amplified Polymorphic DNA (RAPD)

This method which is based on random priming of genomic DNA is used almost exclusively in higher plants (alternatively named AP-PCR (arbitrarily primed polymerase chain reaction). The method was developed in two laboratories at the same time (WILLIAMS, *et al.*, 1990; WELSH and McCLELLAND, 1990). It depends on the observation that single short oligonucleotide primers can frequently recognize similar sequences that are opposed to each other at distances close enough for the intervening sequence to be amplified in the PCR (BURR, 1994). It has been shown that single primers of 8-10 nucleotides in length

will produce from one to several amplified fragments (WILLIAMS, *et al.*, 1990; WELSH and McCLELLAND, 1990). The primer sequences are totally random and no sequence information is needed. The primers are designed to contain at least 50% guanosines (G) and cytosines (C) and to lack internal inverted repeats. As only one primer is used in the reaction, a sequence will be amplified only if the random primer matches the genomic template at two sites, one on each complementary strand, that bracket a short sequence of template DNA. The amplification products are separated on an agarose gel and stained with ethidium bromide. Polymorphisms are the result of insertions, deletions, or simple base changes in either or both priming sites, or insertions between primer sites that make them too far apart for the segment to be amplified. This type of polymorphism is detected as the presence or absence of a band. Less frequently codominant markers are generated with insertions or deletions between primer sites that result in bands differing in length. The polymorphisms are simply inherited in Mendelian fashion. The presence and absence of a specific band are interpreted as corresponding to two alleles at a locus on a chromosome (SKROCH, *et al.*, 1991).

RAPD analysis is well suited for use in the large-throughput systems needed for plant selection in breeding programs, population genetics and studies of biodiversity, and lends itself readily to automation of the process. RAPD analysis is therefore used in several studies for fingerprinting varieties as well as construction of genomic maps and marker detection. Combining the use of NILs or bulked segregant analysis (MICHELMORE, *et al.*, 1991) with RAPD analysis can shorten the process of finding a marker linked to a specific trait from several years to a couple of weeks. Only 300 PCR reactions were required by MICHELMORE, *et al.* (1991) to obtain three markers linked to the target locus in lettuce. MARTIN, *et al.* (1991) surveyed 144 primers for linkage with resistance to *Pseudomonas syringae* pv. tomato in two tomato NILs. They identified and confirmed three markers in about a month. The authors estimated that if each primer generated four products from independent genomic sites, testing 100 primers would yield a marker within an expected distance of 1.9cM from any target gene in tomato, with the upper 95% confidence limit 5.6cM. If higher numbers of primers are surveyed, closer markers can be found with an inverse relationship between number of primers and distance from the target locus (MARTIN, *et al.*, 1991). The probability of obtaining a marker in NILs also depends on the genome size and the degree of DNA sequence divergence between the NILs in the region surrounding the target locus.

Various authors differ greatly in their conclusions on the applicability of RAPDs, although the majority found it to be a valuable tool in plant breeding. REITER, *et al.* (1992) used *Arabidopsis thaliana* to demonstrate the use of RAPD markers for constructing genetic maps. The method is of great value especially for the

speed with which RAPD markers can be generated. DEVOS and GALE (1992) tested the applicability of RAPD markers in genetic analysis of wheat. Due to the non-homoeologous, non-dose and dominant behaviour of RAPDs, they concluded that it was not worthwhile for production of genetic markers and construction of linkage maps in wheat. In contrast, HE, *et al.* (1992) found the application of RAPDs in wheat feasible to utilize in marker-based selection in a breeding program. They conducted a similar study, but used a denaturing gradient gel electrophoresis system (DGGE) for detection of fragments. The system revealed that a number of different DNA species were contained in single bands as resolved by agarose. Over 38% readily detectable and reproducible polymorphisms between two wheat lines, and a high level of polymorphism between commercial varieties and breeding lines were recorded. The survey was broadened to application in other cereal crop species, namely barley and oat (DWEIKAT, *et al.*, 1993), and believed that the combination of techniques were superior to RFLP and RAPD combined with other electrophoresis techniques. DGGE allows the resolution of sequence differences among fragments of similar size, and takes advantage of the fact that even single base-pair differences will alter fragment melting properties (Tm), and altered migration rate.

RAPD markers were also linked to 11 resistance genes to Hessian fly in wheat using NILs (DWEIKAT, *et al.*, 1997), the *PM1* gene for resistance to powdery mildew in wheat (HU, *et al.*, 1997), resistance to *Phytophthora fragariae* in strawberry (HAYMES, *et al.*, 1997), sunflower rust resistance genes (LAWSON, *et al.*, 1996), to list but a few.

RAPD was used extensively in common bean (*Phaseolus vulgaris* L.) for mapping of specific traits. Near-isogenic lines in combination with RFLP and RAPD were used to identify markers linked to resistance to anthracnose (caused by *Colletotrichum lindemuthianum*), and confirmed in a backcross population (ADAM-BLONDON, *et al.*, 1994). The RAPD marker was converted into a SCAR, resulting in a codominant marker. Selection against a repulsion-phase RAPD marker linked to resistance to common bean mosaic virus (BCMV) proved to be more effective in selecting homozygous resistant individuals than selection for a coupling-phase marker (HALEY, *et al.*, 1994). This was found even where the repulsion-phase marker had greater linkage distances than the coupling-phase marker. JOHNSON, *et al.* (1995) found however, that their coupling-phase marker linked to resistance to rust in common bean was most useful for selection of resistant BC_nF_1 individuals during traditional backcross breeding. The repulsion-phase marker was more effective for selecting homozygous-resistant individuals in F_2 or later segregating generations. RAPD markers were also found which were linked to various rust resistance genes in bean (MIKLAS, *et al.*, 1993;

HALEY, *et al.*, 1993), and QTLs for resistance to common bacterial blight and bean golden mosaic virus (MIKLAS, *et al.*, 1996).

BEAUMONT, *et al.* (1996) compared the ability of RFLPs and RAPDs to create a genetic linkage map of maize. Most of the RFLPs (80%) could be placed on a linkage map with a high level of certainty (LOD>4). Due to their dominant nature, only between 37% and 59% of RAPD markers could be placed on the linkage map with the same LOD score. Combined data from RAPD and RFLP increased the level of information provided by RAPDs.

The use of RAPD in fingerprinting varieties from several plant species has been extensively studied. RAPDs are useful for classification of japonica rice cultivars, but it was concluded that many primers would be needed to resolve closely related cultivars (MACKILL, 1995). Analysis of 100 accessions of *Stylosanthes scabra* led to the conclusion that seven of these were not *S. scabra* after all (LIU, 1997). Comparison of clustering results based on RAPD with results from morphological-agronomical characters did not always match. This could be attributed to the influence of environmental factors on the morphological-agronomical characters. RAPD were also used for fingerprinting varieties in soybean (MIENIE, *et al.*, 1995; THOMPSON, *et al.*, 1998), avocado (FIEDLER, *et al.*, 1998), walnut (NICESE, *et al.*, 1998). A computer generated key was developed for identification of Canadian registered oat cultivars with 13 selected amplified fragments in 53 cultivars (GUILLIN, *et al.*, 1998).

RAPD (AP-PCR) was also used for fingerprinting strains of bacteria (CANCILLA, *et al.*, 1992). The method was modified to incorporate ^{32}P or a fluorescent label, followed by electrophoresis in a polyacrylamide-urea gel. The flourescent labelled fragments could be analysed by an automated DNA sequencer. Closely related strains of *Lactococcus lactis* produced almost identical fingerprints, but could be differentiated from each other. The technique was able to detect strain relationships and to differentiate unambiguously between strains that were not closely related. The automated DNA sequencer allowed computer storage of data, providing a basis for the compilation of a reference library of fingerprints.

Inheritance of RAPD markers was studied in an interspecific cross in the genus *Stylosanthes* (KAZAN, *et al.*, 1993a). Ninety 90 primers were tested which all amplified polymorphisms in the two parents, but when these were screened against some F_2 progeny, only 35 primers displayed easily resolved and scorable bands. Segregation of 55 loci amplified fit a 3:1 ratio. Eight loci deviated significantly from the expected 3:1 ratio. In seven out of ten additional loci tested, the loci did not segregate at all, but displayed maternal

inheritance. The distorted segregation ratios were attributed to the existence of four genomes in the progeny, possible linkages between markers and genes operating in prezygotic and postzygotic phases of reproduction, preferential chromosome elimination, preferential fertilization or selective elimination of particular zygotes. Segregation distortion was also observed for loci correlated with pollen-viability variation in sunflower (QUILLET, *et al.*, 1995). This phenomena is population dependent and is not due to the marker technique used, but to a segregation distortion of the gametes or zygotes leading to the F_2 progenies (EUJAYL, *et al.*, 1997), especially in wide crosses. In two wide crosses of lentil, 83% of RAPD markers showed segregation distortion, which was also observed for isozyme and morphological loci, in one of the crosses. In contrast, there was only 10% distortion in the second cross (EUJAYL, *et al.*, 1997).

Before using RAPD analysis the advantages and disadvantages of the technique for the specific application must be carefully weighed. RAPD analysis is fast and easy to perform, does not need any radioactivity and large numbers of samples can be analysed. Analysis can also be done on very limited amounts of DNA. A universal set of primers can be used for genomic analysis in a wide variety of species (WILLIAMS, *et al.*, 1990). Application of the technique is however, limited in that the priming polymorphism appears to be based on mismatches with target sequences so that alleles are either present or absent (BURR, 1994), leading to dominant markers. Segregating progeny can thus only be scored for the presence or absence of a marker and heterozygotes can not be distinguished from homozygotes. Since there is no guarantee that the dominant allele will be present in a second population, it is not always possible to use a RAPD locus in a second population. Possibly the greatest disadvantage of the technique is the apparent utmost sensitivity for changes in reaction conditions, which makes it difficult to duplicate results in different laboratories. Several reports emphasized this problem. DEVOS and GALE (1992) reported that the amplification reaction was sensitive to template DNA concentration, Mg^{2+}, *Taq* polymerase and denaturing temperature. CHEN, *et al.* (1997) studied the reproducibility of differential amplification of root and leaf DNA in soybean. They found that bands over 2 kb and less than 400 bp, were generally less stable and confirmed findings of inconsistent amplification with different thermal cyclers, different batches of the same enzyme, different enzymes and even primers with the same sequence from different sources. The period of cycling times also affected banding patterns. Differences were observed in DNA from the two different organs, with two types of variation - presence of absence of a band, or differences in intensity of bands. The second type of variation was more sensitive for all factors tested. Southern hybridisation indicated that these bands were related to repeated sequences. Polymorphic loci between root and leaf DNA samples were inconsistent among lines, and they suggested that methylation patterns could be involved in the differences observed, but that it required further study.

RAPD band reproducibility was measured as 76%, with the data scoring error at 2% in a study of replicate reactions with bean DNA (SKROCH and NIENHUIS, 1995). The reproducibility could be correlated with band quality, i.e. bold, medium or faint bands, with faint bands the least consistent. The results indicated that there is variation in reproducibility among primers and this should be taken into account when selecting primers for generation of DNA markers.

RAPD amplified products often contain repetitive sequences which render it unsuitable as hybridization probes. This can be overcome by converting the RAPD fragment to an RFLP by cloning the amplified product. Alternatively the ends of the fragment can be sequenced and stable primers synthesized to amplify this fragment preferentially in several genotypes (PARAN and MICHELMORE, 1993).

1.4.4.2(iii) Amplified Fragment Length Polymorphism (AFLP)

The details of the AFLP technique were first published by VOS, *et al.* (1995) after the initial patent application by ZABEAU and VOS (1993) from Keygene N.V., Wageningen, The Netherlands. It was initially called selected restriction fragment amplification (SRFA). The AFLP technique combines the use of restriction enzymes, generating restriction fragments as in the RFLP technique, with the ease of the PCR reaction with selective oligonucleotides. The technique involves three steps: (1) Digestion of the genomic DNA with two restriction enzymes. The one enzyme was chosen for frequent cutting in the genome, *MseI*, and the other cuts less frequently and was chosen for its reliability, *EcoRI*. The restricted fragments are ligated to oligonucleotide adapters containing a core sequence and an enzyme specific sequence. (2) Selective amplification is done in two steps. A preamplification reaction uses two primers each having a single selective nucleotide (primer+1). This product is used as template in the second amplification reaction with primers having longer selective extensions at the 3' ends. The two oligonucleotides used in the reactions corresponds to the *MseI*-ends and *EcoRI*-ends generated in the restriction-ligation reaction. Only DNA fragments with nucleotides flanking the restriction sites that match the selective nucleotides of the primers are amplified during PCR. (3) The amplified fragments are resolved with denaturing polyacrylamide gel electrophoresis (PAGE). Predominant amplification of restriction fragments, which have a rare cutter sequence at one end and a frequent cutter sequence at the other end, results in a complex banding pattern. Typically 50-100 restriction fragments are amplified and detected on denaturing PAGE (VOS, *et al.*, 1995). The number of fragments amplified is determined by the two enzymes used in the restriction digestion of the genomic DNA. The number of selective nucleotides added to the PCR primers also determines the number of fragments amplified. This

phenomenon is exploited in adaptation for the efficient amplification of DNA from organisms with different genome sizes. Adding selective nucleotides to the primers reduced the number of bands 4-fold with each additional selective base (VOS, *et al.*, 1995). DNA from organisms with smaller genomes are more efficiently amplified with primers with less selective bases, e.g. for bacterial or fungal DNA one or two selective bases with each primer were used. Plant species with intermediate genome sizes amplify an optimal number of bands with primers+3 selective nucleotides. The system was also optimized for plant species with very large genomes, like *Astroemeria* (Inca lily) (HAN, *et al.*, 1999). The nuclear content of this species ranged from 37 to 79 pg with a haploid genome size of 25 pg relative to the genome of *Arabidopsis* of 0.04 pg. Reproducible fingerprints were found with preamplification with four selective nucleotides (each primer+2) followed by selective amplification with primers+4 selective nucleotides. Originally the method was used with radioactively labelled oligonucleotides, but the method was found to be more effective without the use of radioisotopes, and combined with silver staining of the gels (CHO, *et al.*, 1996). As with RAPD, no prior knowledge of DNA sequence is required.

The basic difference between RFLP and AFLP polymorphisms is that for RFLP, an area of DNA is scanned that is defined by the number of nucleotides in the restriction sites, whereas for the AFLP technique, an additional number of nucleotides defined by the 3' selective nucleotides is scanned (BECKER, VOS, *et al.*, 1995). Therefore, it is expected that AFLP markers will detect more point mutations per 100 nucleotides than RFLPs. Both procedures should detect the same frequency of insertions and deletions. Mapping of AFLP markers onto an existing RFLP map of barley, filled in gaps between the RFLP markers, but seldom interrupted RFLP clusters, grouping next to them.

AFLP has since been applied in fingerprinting varieties of various species, genome mapping and marker-assisted breeding. AFLP revealed a large number of polymorphisms in rice (MAHESWARAN, *et al.*, 1997). Using only 20 pairs or selective primer combinations, 945 bands were amplified, of which 208 (22%) were polymorphic. The majority of markers showed Mendelian segregation in a doubled haploid population. A much lower order of polymorphism was observed (11.3%) in a barley doubled haploid population used for generating a genomic map with RFLP and AFLP markers (BECKER, *et al.*, 1995). All of the 114 bands could be scored as present or absent, with none of them cosegregating. The segregation patterns of the AFLP loci revealed that there was slightly more alleles (52%) from the one parent in the progeny than from the other parent (48%). This was in accordance with RFLP data. Of the 118 AFLP loci, 94% displayed the expected 1:1 segregation pattern, while seven markers showed distorted segregation. Some of the markers showing distorted segregation were linked to RFLP markers that also showed abnormal patterns.

Other researchers were able to identify a higher incidence of cosegregating loci with AFLP markers. MAHESWARAN, *et al.* (1997) identified 22 codominant loci from a total of 945 loci from 20 primer combinations in a doubled haploid rice population. A very high degree of segregation distortion (65%) was observed in doubled haploid lines of *Brassica oleracea*, much higher than that observed in F_2 populations of the same species (VOORRIPS, *et al.*, 1997). It seems that, as for all of the other techniques discussed, great variation can be found in the results obtained from different species, and the applicability of the technique should be evaluated for each new species under investigation.

The construction of genetic maps for various crops were done with analyses of segregating populations with a number of AFLP primer combinations. RFLP markers with known chromosomal locations could be combined with AFLP data and mapped together on one genetic map (NANDI, *et al.*, 1997). These RFLP markers were used as anchor markers for the AFLP map. Based on the RFLP anchors, AFLP linkage groups were then associated with specific chromosomes. This approach was also followed by BECKER, *et al.* (1995) and MAHESWARAN, *et al.* (1997).

Genetic maps from different potato genotypes were aligned by verifying the identity of comigrating markers in the different populations (ROUPPE VAN DER VOORT, *et al.*, 1997a). Because the AFLP technique produced a large number of bands per reaction, it was difficult to recognise different allelic products as amplified from a single locus. AFLP markers were thus mapped as alleles rather than loci. The ability to align maps would depend on the number of alleles shared among the markers segregating in different mapping populations. Two factors contributed to correct identification of identical markers: (a) AFLP markers were amplified under highly stringent conditions and it was unlikely that two amplification products of identical size could arise from mismatches in primer-template annealing during PCR. (b) The mobility of a PCR product could be estimated very accurately in a sequencing gel. The probability of coincident comigration arising by chance was estimated at 0.03. The identity of comigrating markers were confirmed through cloning and sequencing of twenty putatively homologous markers. Nineteen of these were shown to be identical (ROUPPE VAN DER VOORT, *et al.*, 1997a).

Although AFLP is regarded as mainly a dominant marker system, with very few codominant markers identified, ROUPPE VAN DER VOORT, *et al.* (1997a) recognised three types of segregation patterns of AFLP products in segregating rice populations: (1) Where an amplification product was found in one parental clone and segregated as presence/absence polymorphism in the offspring, the underlying genetic model was assumed to be *Aa x aa* or *aa x Aa*. (2) Where amplification products were found in both

parental clones at identical positions in the gel and one band was double the density of the other, the genetic model was assumed to be *Aa x AA* or *AA x Aa*. (3) Where both types of polymorphisms (presence/absence polymorphisms and band intensity polymorphisms) were observed in the offspring descending from parents with the weaker band intensity phenotype, the underlying genetic model was supposed to be *Aa x Aa*.

The AFLP technique was also modified to allow display of mRNA fingerprints and could be used to isolate sequences mapping to deleted chromosome segments in hexaploid wheat (MONEY, *et al.*, 1996). As was the case with RAPD, differences in banding patterns were observed in AFLP of seed and leaf DNA of wheat (DONINI, *et al.*, 1997). It was most likely that the differences were the result of DNA methylation differences between organs, as the methylation sensitive enzyme *Sse*83871 was used in the analysis.

The applicability and extent of AFLP variation in soybean was studied by MAUGHAN, *et al.* (1996). They amplified 759 AFLP fragments with just 15 primer pairs and found 17% of the fragments to be polymorphic in *Glycine max.* The average number of fragments detected ranged from 19 to 86. The levels of polymorphism ranged from 23% to as high as 64%. AFLP fragment sizes ranged from 35 to 400 bp. Inheritance of polymorphic fragments in a segregating population displayed simple Mendelian patterns.

The large number of fragments generated with a small number of primer pairs makes the AFLP technique superior in variety analysis. In lentil a much higher level of polymorphism was detected with AFLP than with RAPD (SHARMA, *et al.*, 1996). The use of 148 AFLP fragments generated with four primer combinations was able to discriminate between lentil genotypes which could not be distinguished using 88 RAPDs. TOHME, *et al.* (1996) found AFLP a very reliable technique for studies of genetic diversity, permitting greater insights into the genetic structure of wild beans than had been possible with other methods of analysis. Other methods such as RFLP divided the bean germplasm into two major gene pools, the Mesoamerican and southern Andean gene pool. Analysis with AFLP confirmed these findings, and revealed additional gene pools from Colombia, and the northern Andes of Ecuador and northern Peru. In the southern Andean gene pool, more discrete groups were formed which were associated with certain regions (TOHME, *et al.*, 1996).

AFLP markers are very useful in assessing genetic diversity in barley (SCHUT, *et al.*, 1997). Each of the eight primer combinations tested was able to identify all 31 lines uniquely. Another study found the level of polymorphism in wild barley to be 76% using AFLP analysis (PAKNIYAT, *et al.*, 1997). Twelve markers

were also found to be significantly associated with salt tolerance in the 30 barley lines. The genotypes analysed were grouped together according to area of origin in a dendrogram of AFLP markers. Genetic relationships between cultivated cassava and six wild taxa from the same genus (*Manihot*) were estimated using AFLP markers (ROA, *et al.*, 1997). Species-specific markers, which might be useful in germplasm classification, were suggested by the unique presence of AFLP products in samples of each of three wild species. AFLP were also used successfully in fingerprinting of 32 genotypes of Indian and Kenyan tea with five primer combinations (PAUL, *et al.*, 1997). A dendrogram constructed on the basis of band sharing clearly separated the three populations of tea into China type, Assam type and Cambod type. Genetic markers and maps were lacking in sunflower before 1994 (HONGTRAKUL, *et al.*, 1997). Fingerprints were produced for 24 public inbred lines of sunflower (*Helianthus annuus* L.) using six AFLP primer combinations. Principal-coordinate and cluster analysis separated the lines into two groups, one for B-lines and another for R-lines. Similar findings were reported in local studies of South African breeding lines (unpublished results). Although heterotic groups undoubtedly exist in sunflower, none of these studies could as yet define the groups unequivocally.

Determination of predictive estimates of heterosis or genetic variance among progeny from specific parental combinations was also attempted in wheat breeding lines (BARRETT, *et al.*, 1998). The traditional pedigree method was compared with AFLP analysis. Comparison of the genetic diversity estimates (GDE) of both methods suggested that the AFLP technique might have more utility than GDE_{PED} for identifying parental combinations with maximum allelic variation. Similar hierarchical patterns of genetic diversity among the 43 cultivars were observed with both methods. The influence of the use of different enzymes in the restriction reaction was studied. AFLP fragments from hypomethylated portions of the genome (generated with *PstI:MseI*) were more highly associated with GDE_{PED} than were fragments generated with the methylation insensitive combination *EcoRI:MseI* (BARRETT, *et al.*, 1998), and the mean diversity level detected with *PstI* significantly lower than the mean *EcoRI*-based estimate. Low levels of methylation are associated with high levels of gene expression, which suggests that hypomethylated regions may exert more influence on phenotype. The use of a methylation sensitive enzyme (*PstI*) targeted the monitoring of allelic diversity of expressed genes.

This phenomenon also affected the use of AFLP for diversity analysis for plant variety registration purposes. The criticism against the use of AFLP in cereals was that they tend to cluster in areas of low recombination, such as the pericentromeric regions, which have a high content of repetitive DNA, and so do not provide genome coverage (LAW, *et al.*, 1998). The use of the methylation insensitive enzyme

EcoRI biased the population of fragments to the repetitive fraction as up to 80% of the cereal genome consists of highly repetitive DNA. LAW, *et al.* (1998) tried to avoid this problem by using the methylation sensitive enzyme *SseI* in variety identification. According to guidelines compiled by UPOV (International Union for the Protection of New Varieties of Plants) every new plant variety must undergo statutory testing to show distinctness (D) from others, that they demonstrate uniformity (U) and stability (S) in the characteristics that distinguish them. According to the guidelines, the variety must differ by at least one character (usually phenotypic characters, mostly morphological traits). Various criteria for the distinctness of varieties in terms of AFLP band differences were considered, and it was concluded that the optimal number of polymorphic bands would be between *v* and 2*v*, where *v* was the number of varieties tested. The results indicated that AFLP could be used in determination of distinctness between wheat cultivars, but uniformity over generations and inter-plant variation were not examined.

Other applications included the estimation of outcrossing rate in breeding populations of *Eucalyptus urophylla*, which is an open-pollinated plant (GAIOTTO, *et al.*, 1997). Empirical analysis suggested that a minimum number of 18 dominant markers were necessary to achieve estimates of outcrossing rate. The genomic contribution of parents to populations advanced through recurrent selection was estimated with AFLP markers (VANTOAI, *et al.*, 1997) in soybean. The AFLP markers provided a relatively inexpensive technique for precise estimates of the parental contribution with a small number of primer pairs.

Another form of the AFLP technique combined it with simple sequence repeats (microsatellite-AFLP), where one of the two amplification primers was replaced by a compound simple sequence repeat in the PCR reaction (LU, *et al.*, 1996).

AFLP analysis is quick, robust, requires minimal preliminary work and is capable of detecting >50 discrete genetic loci in a single PCR reaction (MAUGHAN, *et al.*, 1996). These markers appear to be inherited in a stable Mendelian manner. However, due to their dominant nature, AFLP markers provide less information per locus than codominant markers such as RFLP. Despite this fact, AFLP analysis is extremely efficient in detecting markers for map-based applications because they allow the simultaneous analysis of a large number of bands in a single reaction (BECKER, *et al.*, 1995).

The AFLP technique combined with bulked segregant analysis was used to enrich a part of the potato genome with markers more closely linked to the resistance gene to *Phytophthora infestans* (MEKSEM, *et al.*, 1995). Twenty-nine of approximately 3200 informative AFLP loci displayed linkage to the *R1* locus.

A high-resolution map could be constructed for the segment of the chromosome bordered by two RFLP loci, and which included the *R1* locus. PCR based markers were developed for the RFLP loci, but attempts to make SCARs from two AFLP loci were unsuccessful, as polymorphism was lost in the subsequent amplification reaction and the fragments were either too short for Southern hybridisation or revealed repetitive fragment patterns. AFLP markers were successfully isolated and cloned from silver stained-gels of amplification products of rice genomic DNA (CHO, *et al.*, 1996). Specific bands were excised directly from the polyacrylamide gel and used in a PCR reaction. The amplified bands were cloned into a TA vector and sequenced. Southern analysis with the amplified bands indicated that they were single copy sequences and demonstrated Mendelian segregation. Occasionally, different AFLP fragments were cloned from a sample which should contain only one AFLP marker (ROUPPE VAN DER VOORT, *et al.*, 1997a). This could be the result of minor products resulting from less specific annealing temperatures during the final stages of the PCR profile.

Other examples of AFLP markers linked to disease resistance traits, are resistance against leaf rust (*Melampsora larici-populina*) in *Populus* (CERVERA, *et al.*, 1996), cyst nematode in potato (ROUPPE VAN DER VOORT, *et al.*, 1997b), two resistance genes to clubroot (*Plasmodiophora brassicae*) in *Brassica* (VOORRIPS, *et al.*, 1997), and other traits such as QTLs linked to submergence tolerance in rice (NANDI, *et al.*, 1997). AFLP was found to be more efficient in mapping the melon genome than RAPD or microsatellite markers (WANG, *et al.*, 1997), as well as the rice genome (NANDI, *et al.*, 1997).

1.4.4.3 DNA Chip technology

A new technology currently being developed is at the forefront of the functional genomics revolution and promises to have an even greater global impact on Biotechnology than the discovery of recombinant DNA. DNA chips exploits the principle of complementary hybridization of nucleic acid strands through specific base pairing. It is a dramatic breaktrough in miniaturisation where thousands of individual gene sequences are printed in a high-density array on a glass microscope slide, providing a practical economical tool for studying gene expression on a large scale (DERISI, *et al.*, 1997). Two complementary types of DNA chips have been developed in parallel in the USA; 'synthesised' DNA chips (FODOR, 1997) and DNA microarray or 'spotted' DNA chips (SHALON, *et al.*, 1996). The former is commercially available and very expensive. Synthetic chips have their greatest application where information on the target DNA is known, for example HIV array resistance screening or measuring the relative expression level of specific target sequences.

In contrast, the DNA microarray chips developed at Stanford University uses a simpler technology. Small droplets of genomic DNA, cDNA clones or PCR samples are spotted onto a microscope slide. Differential expression of genes can be monitored on duplicate slides, using mRNA extracted from two plants under different conditions, with different expression of genes. Thus in one experiment, resistance genes expressed from different parts of chromosomes can be identified and cloned. DNA chips thus have great potential in the discovery of genes for specific traits, simplification of varietal fingerprinting and developing markers for marker assisted selection.

1.4.5 Sequence Characterized Amplified Region (SCAR) development

A SCAR (sequence characterized amplified region or allele-specific associated primers (ASAPs) or sequence-tagged-site (STS)) is a genomic DNA fragment at a single defined locus which can be identified by PCR amplification with a specific pair of oligonucleotides (PARAN and MICHELMORE, 1993), usually derived from markers identified with other techniques, such as RAPD, AFLP or RFLP. The advantages of SCARs are: (1) it amplifies a single locus; (2) it is less sensitive to changes in reaction conditions and thus more reproducible; and (3) the polymorphism can be observed as a dominant marker (presence or absence of a band) or a polymorphism differing in length of fragments (codominant marker).

Mapped RFLP or RAPD markers can be converted to a SCAR by synthesising primers that uniquely amplify portions of the sequence of a known gene or mapped marker (WILLIAMS, *et al.*, 1991). The SCAR or ASAP uses longer (17-25mer) primers in pairs to specifically amplify the DNA fragment linked to the gene of interest, with the main objective of creating a stable, easy to use and reliable marker. SCARs are obtained by cloning the fragment and sequencing the ends of a RAPD, RFLP probe or AFLP fragment and developing longer primers from this data. Three scenarios are possible: The SCAR could be detected as a dominant marker in the parents, i.e. present or absent. In a most favourable case, fragments of different lengths will be amplified between the two parents of the mapping population, creating a codominant marker. If the two parents produce identically sized products, these fragments can be digested with a series of restriction enzymes and the size of the restriction fragments compared by gel electrophoresis (also termed CAPS - cleaved amplified polymorphic sequences, TSUMURA and TOMARU, 1999). Since most PCR products are relatively small, 4-cutter enzyme sites would be more likely to be present. This approach has been used in studies of fungal genetics (BUCHKO and KLASSEN, 1990; CUBETA, *et al.*, 1991) as well as in plants. CUBETA *et al.*(1991) were able to identify anastomosis groups of binucleate *Rhizoctonia* species by restriction analysis of identical amplified fragments of a region of DNA coding for

a portion of the 25S rRNA. Their findings were consistent with prior groupings based on hyphal anastomosis.

Many SCARs have since been developed for various traits in diverse crops. A SCAR was developed for resistance to anthracnose in common bean (ADAM-BLONDON, *et al.*, 1994). The SCAR amplified by the designed primer pair was not polymorphic between the two parents. An informative polymorphism was observed after *DdeI* digest of the PCR product, suggesting that the RAPD polymorphism was caused by a mismatch in one of the two primer-targeted sequences. A RAPD fragment linked to the *Lr9* leaf rust resistance gene of wheat was cloned, sequenced and specific primers synthesized (SCHACHERMAYR, *et al.*, 1994). Only resistant lines showed an amplified product with these primers at stringent reaction conditions and the SCAR could be successfully applied in marker assisted selection. Several examples of SCAR development have been reported in common bean, including a SCAR marker linked to angular leaf spot resistance (SARTORATO, *et al.*, 1999) from a RAPD band; a SCAR linked to gene *Ur-11*, conferring resistance to the bean rust fungus (BOONE, *et al.*, 1999); and three SCAR markers linked to resistance to bean common mosaic virus (BCMV), *Colletotrichum lindemuthianum* (Sacc. & Magnus) Lams.-Scrib, and *Uromyces appendiculatus* (Pers.) Unger var. *Appendiculatus*, respectively (MELOTTO and KELLY, 1998).

SCARs were also developed successfully from RFLP probes (TALBERT, *et al.*, 1994; HITTALMANI, *et al.*, 1995; WILLIAMS, *et al.*, 1991). A total of 37 primer sets were designed from mapped RFLP clones for wheat. Of these, 29 directed successful amplification of wheat genomic DNA and 23 amplified products that mapped to the expected chromosome group. Nine of the primer sets generated products that showed polymorphic banding patterns upon digestion with either *HinfI* or *HhaI* restriction enzymes (TALBERT, *et al.*, 1994). Similarly, the products amplified with specific primers designed from the RFLP probe, RG64, linked to rice blast resistance, were not polymorphic between the varieties examined (HITTALMANI, *et al.*, 1995). Cleavage of the amplified products with *HaeIII* generated a polymorphic fragment, called a specific amplicon polymorphism (SAP), between the resistant and susceptible genotypes. The segregation pattern of the SAP marker was the same as that of the RFLP marker in an F_2 population. SCAR primers were designed corresponding to 30 Indica rice genomic clones (WILLIAMS, *et al.*, 1991). Size polymorphisms were observed between PCR products from Indica and Japonica varieties, and among wild *Oryza* species. Identical products were amplified between closely related Indica lines, and were digested with 4-cutter restriction endonucleases. In a random sequence, 4-cutter sites should occur every 256 bp (4^4). Sites for certain enzymes such as *AluI* and *RsaI* were more common than others and more useful in preliminary

surveys. Digestion of the PCR products with *AluI, MaeI* and *MaeIII* yielded size differences between the varieties (WILLIAMS, *et al.*, 1991). An important feature of these RFLPs detected in PCR products was that they were not subject to artifacts caused by differences in methylation of the DNA. When RFLPs are detected by Southern hybridization, only a minority of DNA sequence alterations responsible for these RFLPs lie within regions hybridizing with the probes. The majority lie at unknown sites outside the hybridizing regions, with the consequence that these types of SCARs were not always successful in amplifying polymorphisms. Mutations responsible for RFLPs detected by PCR lie within the amplified segments themselves and may be characterized fully by sequencing the PCR products of fragments thereof (WILLIAMS, *et al.*, 1991).

AFLP fragments can also be isolated from the polyacrylamide gels and SCAR primers designed. The fragments can be cut directly from the gel after localisation with autoradiography (MEKSEM, *et al.*, 1995), or after silver-staining (CHO, *et al.*, 1996). In both methods the fragment was eluted or used in a PCR reaction directly, amplified with the same primers used in the AFLP reaction, and cloned in an appropriate vector. Some authors recommended the further purification of the fragment with a few rounds of PCR and agarose gel purification (ROUPPE VAN DER VOORT, *et al.*, 1997a; PADILLA, *et al.*, 1994). The fragments were cloned and sequenced and specific primers were designed. As for other types of markers, the subsequent PCR products with the specific markers could exhibit no polymorphism (MEKSEM, *et al.*, 1995) and need to be digested with restriction enzymes to identify polymorphisms. However, due to the small size of AFLP fragments, they might not contain restriction sites, which can complicate the analysis. SCARs should be verified for copy number and segregation in Southern Blot analysis with genomic DNA from parents and progeny.

SCAR markers were developed for use in soybean (PADILLA, *et al.*, 1994; ZHANG, *et al.*, 1998). SCAR primers were developed from fragments isolated from silver-stained polyacrylamide gels of polymorphic bands linked to a root nodulation locus (*nts*). Fragments were generated by template endonuclease-cleaved multiple arbitrary amplicon profiling (tec-MAAP) - template soybean DNA was restricted with three endonucleases prior to amplification with short (7-8 bp) arbitrary oligonucleotides (PADILLA, *et al.*, 1994). A codominant SCAR marker linked to *Rsa*, a single dominant gene for resistance to soybean mosaic virus was developed (ZHANG, *et al.*, 1998) from RAPD products. The RAPD fragment was cloned and used in Southern analysis for verification. SCAR primers designed from sequence data from this fragment, amplified codominant fragments in the two parents used in developing the mapping population. Identical bands were amplified with SCAR primers developed from a RAPD marker linked to ToMV resistance in

tomato (DAX, *et al.*, 1998). Restriction digestion with *HindIII* lead to formation of polymorphic bands cosegregating with susceptibility or resistance in a F_2 population. Homozygous and heterozygous plants could be distinguished.

1.5 MARKER ASSISTED SELECTION

Plant breeding strategies for introduction of specific traits into existing varieties of annual crop species typically include a backcrossing regime, with simultaneous selection for several traits over a number of generations. The development of a new cultivar requires between five and ten years for most annual crop species. In some situations, genetic advance has been limited due to the complex and ambiguous nature of the trait(s) and its response to environmental cues (LEE, 1995). The probability of selecting superior genotypes is low for low to moderate heritability traits. Plant breeders cope with this problem by producing and testing progeny from numerous crosses, using low selection intensities, using replicated testing, testing advanced generations, and using recurrent selection (KNAPP, 1998). To improve a quantitative trait such as yield performance under specific conditions, phenotypic selection cannot begin until later generations (e.g. F_6 in common bean) where sufficient homozygosity and seed is available for replicated trials (SCHNEIDER, *et al.*, 1996). Incorporation of disease resistance traits poses its own problems. Many pathogens consist of several races, each controlled by different resistance genes in the plant. Gene pyramiding has been suggested as a strategy for stable disease resistance against variable plant pathogens, but incorporating more than one gene into a single genotype is time-consuming and difficult to select (YOUNG and KELLY, 1996). Epistatic interactions between resistance genes require extensive test-crossing with different races of the pathogen, while ensuring that the genotype meets other agronomical requirements. The problem is even greater when the pathogen is not indigenous and restricted by quarantine conditions. Even when the pathogen is present in a specific area, disease development is often dependent on environmental factors.

Marker-assisted selection (MAS) has emerged as a strategy for increasing selection gains. General applications include parent selection, recovery of recurrent parent genome in backcrossing programs, early generation trait selection and multiple trait selection (KNAPP, 1998). It may provide new solutions for selecting and maintaining durable genotypes (HITTALMANI, *et al.*, 1995). Breeding disease-resistant genotypes using marker-assisted selection requires that: (1) the resistance gene(s) be tagged by closely linked molecular markers; (2) the linkage be stable across generations and populations; and (3) an efficient way of screening large populations for molecular markers be available (HITTALMANI, *et al.*, 1995).

Marker-assisted selection would be more competitive with traditional phenotypic selection when improving a trait with low heritability (XIE and XU, 1998). Marker-aided recurrent selection is especially valuable in speeding the breeding process, in selection of immature individuals, and in characteristics that are difficult or expensive to measure such as drought tolerance or pest resistance. The strategy followed strongly depends on the breeding purpose, the available resources of the breeder, the nature of the genome of the species, and the nature of the trait to be improved (XIE and XU, 1998). MAS should be most effective in early generations of selection from progeny of crosses between inbred lines. Heritability is usually lowest and linkage disequilibrium greatest on these generations. The paradox is that the power for mapping QTL decreases as heritability decreases and is lowest for traits where MAS has the greatest theoretical impact (KNAPP, 1998). KNAPP (1998) concluded that MAS substantially decreases the resources needed to accomplish a selection goal for a low to moderate heritability trait when the selection goal and the selection intensity are high. A breeder using phenotypic selection must test 1.0 to 16.7 times more progeny than a breeder using MAS to be assured of selecting one or more superior genotypes.

LANDER and BOTSTEIN (1989) found that the effectiveness of MAS on a particular trait is inversely proportional to the heritability of that trait. This was supported by results from a field study of drought resistance in common bean (SCHNEIDER, *et al.*, 1996). The effectiveness of MAS vs. conventional phenotypic selection was tested with data from two locations. Results indicated that MAS was a better indicator of improved performance while phenotypic selection was a better indicator of below average performance. Although MAS proved effective in one population, it was not effective in improving yield performance in another population. The heritability estimates for yield in the latter population were three times greater than in the first population, which supported the conclusions of LANDER and BOTSTEIN (1989).

The efficiency of application of a SCAR marker is illustrated beautifully by a study on resistance to bacterial blight in common bean (YU, *et al.*, 1999). One hundred and thirty eight lines were tested for presence of the SCAR band and also tested for resistance in the greenhouse. An accuracy of 82% was obtained with only five plants misclassified as resistant. In this particular case the cost of MAS was about one third the cost of the greenhouse test. MAS was used successfully to pyramid four bacterial blight resistance genes in rice (HUANG, *et al.*, 1997). The pyramid lines showed a wider spectrum and a higher level of resistance than lines with only a single gene. Markers were also developed for recessive genes. MAS was applied very effectively in this case where it would have been very difficult or impossible to pyramid multiple resistance genes using conventional breeding methods.

1.5.1 *Towards marker assisted selection in soybean*

Soybean is regarded as a stable tetraploid with diploidized genomes. Because it behaves as a diploid, the chomosome number is regarded as 2n=40 (SHOEMAKER, 1994). The soybean genome contains an estimated 1.29 x 10^9 bp to 1.81 x 10^9 bp for 1n content. The genome contains about 40-60% repetitive sequences (GURLEY, *et al.*, 1979). The soybean genome is highly conserved with a narrow genetic base. Only ten plant introductions contributed more than 80% of the northern genetic pool, while only seven contributed the same share to the southern gene pool (DELANNAY, *et al.*, 1983). Progress with the construction of a genetic linkage map in soybean was slow before the introduction of RFLP technology, with only 40 classical markers in 17 linkage groups covering 420 cM on the linkage map in 1987 (PALMER and KILEN, 1987). The first biochemical markers to be mapped in soybean were isozymes, with the first report of genetic linkage between a biochemical locus (isozyme) and a morphological locus in 1985 (KIANG and CHIANG, 1985). They found linkage between the pubescence colour locus (*t*) and the ß-amylase locus (*Am3*) with 31,88% recombination frequency. RFLP makers for soybean were introduced in the late 1980s. A genomic library as a source of probes for RFLP analysis was constructed by KEIM and SHOEMAKER (1988), using the restriction endonuclease *PstI*, a methylation-sensitive enzyme. Approximately 40% of random genomic probes detected polymorphisms with RFLP analysis. Several genetic maps were constructed with variation in the amount of linkage groups identified over the next decade. KEIM, *et al.* (1990a) mapped 130 RFLP markers in an interspecific cross between *G. max* and *G. soja* to 26 linkage groups, covering 1200 cM. The map was expanded to 252 markers in 31 linkage groups covering 2147 cM (DIERS, *et al.*, 1992a) in 1992. LARK, *et al.* (1993) constructed the first linkage map from an intraspecific cross comprising 31 linkage groups consisting of 132 RFLP, isozyme, morphological and biochemical markers spanning 1550 cM. In 1994 the RFLP linkage map of soybean included 20 linkage groups of three or more markers each and four linkage groups containing only two-point linkages. The linkage map encompassed approximately 2900cM (SHOEMAKER, 1994). The current map is accessible through the internet at 'http://probe.nalusda.gov:8300/cgi-bin/browse/soybase' or complete data in SoyBase at 'http://probe.nalusda.gov:8000/plant/ aboutsoybase.html'. In 1997 a high-density AFLP map was published for soybean (KEIM, *et al.*, 1997) based on RFLP, RAPD and AFLP markers. The authors constructed the map on a RIL population by first using 300 plants for developing a "scaffold" map with RFLP. This anchored map was further populated with AFLP markers on a smaller population of 42 plants. The result was a map with 840 markers, consisting of 165 RFLP, 25 RAPD and 650 AFLP markers, spread over 28 linkage groups representing 3441 cM distance. Although clustering of AFLP markers did occur, the markers could be mapped to every linkage group and were well distributed relative to other

marker systems. One possible explanation for non-uniform distribution of the AFLP markers could be the occurrence of reduced recombination in these chromosomal regions. The RFLP probes used in the study were derived from hypomethylated regions (*PstI* restriction fragments), and might have discriminated against centric and favouring euchromatic regions. The AFLP markers were based on the restriction enzymes *EcoRI* and *MseI*, which are insensitive to methylation patterns. This might have led to a greater sampling of centromere regions of chromosomes than with RFLP (KEIM, *et al.*, 1997). The discovery and mapping of more than 700 SSR markers on the current public soybean map will make the application of MAS in plant breeding even more efficient and cost-effective (BOERMA and MIAN, 1999).

The feasibility of application of the map developed with the interspecific cross to physiologically distant soybean genotypes was studied in 1993 (SKORUPSKA, *et al.*, 1993). They surveyed 108 genotypes comprising ancestral genotypes, breeding lines and elite cultivars with 83 RFLP probes. Forty-six percent of the probes were informative. Thirty-five percent had a probability of detecting polymorphism between any two random genotypes with a frequency above 0.3. TAMULONIS, *et al.* (1997a) found the frequency of polymorphism in soybean to be 42% as analysed with the RFLP technique.

Integrating classical qualitative markers into the molecular map was a potentially efficient process for soybeans, since the germplasm collection contains an extensive number of NILs. Gene mapping with NILs is based on the premise that when a conventional marker is introgressed from a donor parent (DP) into a recurrent parent (RP) through backcrossing, the resultant NIL retains a small number of DP-specific molecular markers in its genome. Most of these markers will be linked to the introgressed conventional marker (MUEHLBAUER, *et al.*, 1991). The DP/NIL/RP sets are genotyped for their allelic status at molecular marker loci of interest. If the NIL possessed the DP allele at one or more molecular marker loci, then one could presume linkage between the molecular marker and the introgressed trait. Verification of the linkage would necessitate cosegregation analysis in an F_2 or F_3 population. The same strategy was also used to map phytophtora resistance (*Rps*) and the locus for ineffective nodulation (*Rj2*) (DIERS, *et al.*, 1992b). The authors found linkage between RFLP markers and loci *Rps1, Rps2, Rps3, Rps4, Rps5* and *Rj2*. Linkage was also found between *Rps2* and *Rj2*. A RAPD marker linked to *Rps4* was found by BYRUM, *et al.* (1993) in the same NIL and confirmed in an F_2 population (44 plants). In another study a combination of RAPD analysis with restriction enzyme digests increased the number of informative bands and led to identification of markers tightly linked to the supernodulation locus in an F_2 population (CAETANO-ANOLLÉS, *et al.*, 1993b).

KEIM, *et al.* (1990b) reported five independent genomic regions containing putative QTLs for seed hardness. They analysed 60 F_2 plants from an interspecific cross segregating for the trait with 72 RFLP probes. This data was correlated with the phenotype of the F_4 families. The five markers and their epistatic interactions explained 71% of the variation in the hard-seeded trait. MANSUR, *et al.* (1993) mapped QTL for reproductive, morphological and seed traits in a segregating soybean population of a cross between two soybean cultivars 'Minsoy' and 'Noir 1'. QTLs for developmental and morphological traits (development stages R1, R5 and R8, plant height, canopy height and leaf area) tended to be clustered in three intervals, two of which were also associated with seed yield. Seed oil content was not correlated to any of the other traits, while seed protein content had a negative genetic correlation with seed yield. Transgressive segregation was observed for all the traits, indicating that the two parental genotypes had achieved similar phenotypes through different gene combinations (MANSUR, *et al.*, 1993). In a further study seed oil and protein were found to be inversely correlated (LARK, *et al.*, 1994). By analysing DNA from extreme phenotypes of recombinant inbred lines and plotting cumulative distributions of the trait values, the authors found that QTLs for maturity acted independent and additively. Cumulative distributions of values for oil and protein content linked to RFLP marker R183 were not compatible with an additive model, but fit the hypothesis for epistatic action. For this type of analysis two parameters are essential: (1) The genotypic variation must be large relative to the environmental effects, with high heritability. (2) The population should be large (LARK, *et al.*, 1994).

In a population derived from a *Glycine max* x *Glycine soja* cross, QTLs have been identified which affected iron efficiency, hard seededness, protein, oil, maturity, height, lodging, days to flowering, seed-filling period, stem diameter, stem length, canopy height, leaf width and leaf length (BURTON, 1997). The level of consistency of QTL across environments was trait specific and population specific (LEE, *et al.*, 1996). The authors evaluated 120 F_4 lines for segregation at 155 RFLP loci over four different locations. With single-factor analysis of variance, 11 markers could be associated with plant height and eight with lodging. Only two of these markers for plant height and one for lodging were detected at all locations. However, QTL for maturity were more consistent, with four out of five markers detected at all locations. The QTL linked to oil and protein content were the subject of a similar study across environments (BRUMMER, *et al.*, 1997). The study evaluated eight different populations for genetic markers linked to seed protein and oil content. The identified QTL were sensitive to both environment and genetic background although some common QTL were identified in multiple populations across several years.

Markers for other QTL in soybean include RFLP linked to water use efficiency and leaf ash (MIAN, *et al.*, 1996) and chlorimuron ethyl sensitivity (MIAN, *et al.*, 1997). Frequently a large number of seemingly unrelated traits map to the same Linkage Group (LG) (IMSANDE, *et al.*, 1998). A total of 18 traits were associated with LG-L, including seed weight, first flower, R5, pod maturity, reproduction period, hard seed, pod dehiscence, oil, protein, linolenate, leaf ash, chlorimuron ethyl sensitivity, stem diameter, plant height, leaf length, lodging, yield/height and height/lodging. QTL for disease resistance seemed to be clustered on certain LGs and were located within limited distances on those LGs. LG-F contains *Rps3* and markers associated with resistance to soybean cyst nematode, soybean mosaic virus, peanut mottle virus, *Pseudomonas syringae*, *Meloidogyne javanica*, *Meloidogyne arenaria*, corn ear worm and *Phytophtora megasperma*. LG-G contains *Rps4* and markers for resistance to soybean cyst nematode, sudden death syndrome, iron deficiency, manganese toxicity and *Meloidogyne incognita*. LG-J is associated with *Rj2*, *Rps2*, *Rmd*, seven resistance gene analogs and a marker for resistance to soybean cyst nematode (IMSANDE, *et al.*, 1998).

Efficiency of application of molecular markers in soybean breeding programs has had variable success. An early QTL identified conditioning iron deficiency chlorosis were not effective in selection for the trait among additional lines from the same population (BOERMA and MIAN, 1999). MUDGE, *et al,* (1997) reported 98% accuracy in prediction of resistance or susceptible phenotype in selection for resistance to soybean cyst nematode. The use of these SSR markers flanking the QTL for resistance to this nematode is one of the first successful applications of MAS in both commercial and public soybean breeding (BOERMA and MIAN, 1999). The narrow genetic base of soybean breeding lines makes it critical to be able to identify parents which would produce transgressive progeny with superior agronomic traits. The use of DNA markers for predicting the best parental combinations that would produce such superior genetic hybrids for yield and other agronomic traits, has generally been unsuccessful (BOERMA and MIAN, 1999). However, there was a trend for DNA markers to identify groups of crosses that produced progeny with superior genetic variance in yield. This strategy is currently under investigation (BOERMA and MIAN, 1999).

DNA fingerprinting for identification of varieties in soybean was attempted with isozyme analysis, RFLP analysis and hybridization with oligonucleotide probes for simple repetitive sequences (YANAGISAWA, *et al.*, 1994). The highest polymorphic frequency were obtained with $(AAT)_6$ as a probe, with which all of the 47 soybean cultivars tested could be distinguished. PRABHU, *et al.* (1997) compared DNA amplification fingerprinting (DAF) with RFLP analysis of 10 soybean genotypes and concluded that both

methods could distinguish the varieties with similar efficiency. DIWAN and CREGAN (1997) used automated sizing and fluorescent labelled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Large numbers of SSR markers were developed by designing primers for sequences flanking SSR regions in the soybean genome. Primers were labelled with fluorescent tags, genomic DNA was amplified and the products analysed on an automatic DNA sequencer. All the soybean cultivars used in the study could be readily distinguished. Subsequent studies were aimed at establishing a standard set of SSR loci for use in soybean plant variety identification. Such loci were selected based upon (1) a high level of informativeness (gene diversity), (2) position in separate linkage groups, (3) the production of discrete products with minimal "stutter" bands, and (4) the ability to permit multiplex PCR amplification.

1.5.2 Nature of nematode resistance genetics in soybean

The soybean cyst nematode is one of the most destructive pests of soybean in the USA (CONCIBIDO, *et al.*, 1994) and was the focus of attention of many studies. Inheritance of resistance to soybean cyst nematode (SCN, *Heterodera glycines Ichinohe*) is multigenic (at least ten genes (RIGGS and SCHMITT, 1987)) and complex (CONCIBIDO, *et al.*, 1994). Classical genetic studies found that the resistance was carried by at least five major genes: one dominant and four recessive. BOUTIN, *et al.* (1992) analysed four pairs of near isogenic lines, differing only for SCN resistance with 60 uniformly spaced RFLP markers. Of these, 52 RFLP probes were informative and ten of these clones showed differences between the resistant and susceptible lines. The ten markers mapped to four different linkage groups. These results were used as an indication and needed confirmation with F_2 populations from crosses between resistant and susceptible lines. CONCIBIDO, *et al.* (1994) identified two markers associated with resistance to race 3 of *H. glycines*. A total of 56 F_2 lines of a cross between inbred lines for resistance and sensitivity respectively were used for linkage analysis with DNA polymorphisms. Both F_2 and F_3 lines were tested for resistance. The F_2 genotypes for each polymorphic RFLP marker were contrasted with SCN disease response with statistical methods (regression analysis, analysis of variance and MAPMAKER-QTL) to identify loci associated with the disease. Due to the small size of the population the authors suggested that only loci with moderately high effects could be uncovered. Two unlinked RFLP markers pA85 and pB32 accounted for 51.7% of disease response (CONCIBIDO, *et al.*, 1994) and linked to linkage group G in soybean (CONCIBIDO, *et al.*, 1996a). Two microsatellite markers, BARC-Satt038 and BARC-Satt130, flanking the major SCN resistance locus were identified (MUDGE, *et al.*, 1997). These polymerase chain reaction-based markers were less expensive and required less labour than the RFLP markers. In a more recent study (QIU, *et al.*, 1999) found additional QTL associated with resistance to

three races of SCN. Five RFLP markers, A593 and T005 on LG-B, A018 on LG-E, and K014 and B072 on LG-H were significantly linked to loci for resistance to Race 1, which jointly explained 57.7% of the phenotypic variation. Three markers (B072 and K014 on LG-H and T005 on LG-B were associated with resistance to Race 5 and together explained 21.4% of phenotypic variation. Two markers (K011 on LG-I and A963 on LG-E) were associated with resistance to Race 5 and explained 14.0% of the total phenotypic variation. Three RFLP markers located on LG-B and H were linked to loci conferring resistance to both Race 1 and Race 3 isolates. The genomic region on LG-H may contain a cluster of unique, but closely adjacent QTLs, but also suggested the possible existence of pleiotropism among loci for SCN resistance in the population studied (QIU, *et al.*, 1999).

A root-knot nematode resistance gene (*Mi*) was identified in tomato in the 1940's and transferred to inbred tomato lines from the wild tomato, *Lycopersicon peruvianum*, using conventional breeding techniques. This gene confers resistance to most races of *M. incognita*, as well as *M. javanica* and *M. arenaria* but not to *M. hapla*. There is however, a breakdown of the resistance when the soil temperature is above 28°C (DROPKIN, 1969). At this temperature only 2% of the larvae within the roots developed further, with 87% at 33°C. The host-parasite relationship developed either toward resistance or susceptibility during the first 24-48 hours after penetration, and then was not reversible by shifts in temperature. There are also reports of selected virulent isolates of *Meloidogyne* spp. normally controlled by the gene that are able to reproduce on tomatoes bearing the *Mi* gene (ROBERTS and THOMASON, 1986). *Mi* is a single dominant gene and maps in the tomato genome on chromosome 6, tightly linked to the isoenzyme marker for acid phosphatase (*Aps-1*). MESSEGUER, *et al.* (1991) built a high resolution RFLP map around the *Mi*-gene in order to isolate the gene with chromosome walking and to clone the gene. A single dominant heat-stable resistance gene which is expressed at 30°C was later found by CAP, *et al.* (1993) in *Lycopersicon peruvianum*. The heat-stable resistance locus LA2157 was mapped on chromosome 6 and localized in the resistance genes' cluster close to *Mi-1* (VEREMIS, *et al.*, 1999). The gene was expressed at 32°C.

LUZZI, *et al.* (1987) tested 2370 soybean genotypes for resistance to *M. incognita* (Mi), *M. arenaria* (Ma) and *M. javanica* (Mj) based on root galling and nematode reproduction. Sixty one genotypes were found to be resistant to Mi, 56 to Ma and 61 to Mj. In the next screening levels different inoculum densities were used with Forrest and Gordon as resistant checks. 'Amredo', PI96354, PI408088 and PI417444 were identified as having better resistance against Mi than Forrest, which was previously believed to be resistant to certain nematodes. PI230977 displayed a lower gall index in the Mj screening than 'Gordon', the resistant check. The fact that different cultivars had different resistance patterns against the three species

of *Meloidogyne*, led to the assumption that more than one gene were involved. In this study, gall indices were positively correlated with reproduction of all three species. There were, however, exceptions in some soybean lines, which had low gall indices and high reproductive ability of nematodes, or *vice versa*. There was an apparent genetic independence between galling and nematode reproduction in soybean (HUSSEY and BOERMA, 1981). This suggests that separate genes and mechanisms could be involved in the resistance to root galling and nematode reproduction, as was observed for groundnut (GARCIA, *et al.*, 1996) in the case of resistance to *Meloidogyne arenaria* .

Data from LUZZI, *et al.* (1994a) indicated quantitative inheritance of resistance to *Meloidogyne incognita* ((Kofoid and White) Chitwood). Their data suggested that the two plant introductions used in the analysis had resistance genes at the same loci and differed from Forrest by at least one gene. Neither resistance nor susceptibility was dominant. Forrest was previously found to possess one additive gene (LUZZI, *et al.*, 1994b). Resistance to peanut root-knot nematode (*Meloidogyne arenaria* (Neal) Chitwood) was also inherited quantitatively with different genes in the lines tested (LUZZI, *et al.*, 1995a). TAMULONIS, *et al.* (1997a) identified two QTL which mapped to LG-O and LG-G of the USDA/ARS-Iowa State University soybean map. The major resistance QTL was linked to marker G248A-1 and the second in the interval from K493H-1 to Cs008D-1 and was dominant with respect to resistance. Together the two QTL accounted for 39% of the variation in Mi galling. The authors speculated that the major resistance QTL located on LG-O could be the *Rmi1* gene found in Forrest (LUZZI, *et al.*, 1994b). A two QTL model was also proposed for resistance to *M. arenaria* (TAMULONIS, *et al.*, 1997c). One QTL was mapped at 0 cM recombination with marker B212V-1 on linkage group F and was additive to partially dominant. The gene action for the QTL located on LG-E was dominant. Together the QTLs accounted for 51% of variation in gall number in $F_{2:3}$ families from a cross between PI200538 and 'CNS'.

Inheritance of resistance to *Meloidogyne javanica* was determined in crosses of a susceptible genotype 'CNS', and three resistant genotypes, 'Gordon', PI80466 and PI230977 (LUZZI, *et al.*, 1995b). Both F_2 and F_3 populations were evaluated for gall formation in a greenhouse screening procedure. The correlation of gall number on F_2 and the mean of their F_3 progeny was 0.26-0.29. Inheritance was probably also quantitative, with moderate to high heritability (estimates ranging from 0.48 to 0.76). The authors proposed that the different parents might possess resistance genes at different loci or different alleles at the same loci. Resistance was not maternally inherited. TAMULONIS, *et al.* (1997b) identified RFLP markers linked to genes affecting resistance to *M. javanica*. Eighty-four F_2 progeny from a cross between 'CNS' (susceptible) and PI230977 (resistant) were used to map 86 RFLP markers. Two QTL conditioning

resistance to root galling of Mj were identified. The marker B212-1 accounted for 46% of the variation in gall number and mapped to linkage group F, and marker A725-2 accounted for 13% and mapped to LG-D1. No studies were done to determine the linkage to resistance to reproductive ability of Mj on soybean. Marker B212-1 is within a cluster of other disease resistance loci on LG-F (TAMULONIS, *et al.*, 1997b). The major QTL mapped to the same location as the resistance to *M. arenaria* at marker B212V-1 on linkage group F (TAMULONIS, *et al.*, 1997c), and the QTL for Ma and Mj could be the same gene. TAMULONIS, *et al.* (1997a) concluded through comparative mapping with common bean and mungbean that root-knot nematode resistance QTL were located on duplicated segments found on linkage groups F,E, D1, O and G.

1.6 DEVELOPMENT OF A MARKER FOR SELECTION OF RESISTANCE TO *M. JAVANICA* IN SOYBEAN

In a study conducted on widely cultivated South African varieties (FOURIE and MCDONALD, 1999), only one cultivar was identified with significant resistance to *M. javanica*. This variety, Gazelle (Maturity Group VIII), originated in Zimbabwe. Soybean varieties from maturity groups IV to VII are most widely cultivated in South Africa, and it will therefore be advantageous to get higher levels of nematode resistance in these locally adapted cultivars. Traditional breeding methods had to rely on field selection in which complex nematode populations can be found. This leads to non-specific breeding for resistance. Greenhouse selection can only be done with the use of cultivated nematode populations, which require careful maintenance and extra greenhouse space. Selection based on phenotype is influenced by numerous environmental factors, such as population density and aggressiveness, humidity, soil and nutritional factors, etc. The use of a molecular marker would enable breeders to select resistant plants without the need for labour intensive nematode evaluations. A genetic marker linked to the resistance gene(s) can identify the resistance on genotype instead of phenotype and can eliminate nematode screening of breeding material in early generations.

An RFLP marker linked to the resistance trait was identified in a soybean population (TAMULONIS, *et al.*, 1997b), but the applicability in foreign soybean breeding populations was not established. This needs to be evaluated in the local breeding material. RFLP markers are very reliable, but tedious and impractical to use in large populations. It would thus be preferable to develop a marker detected by PCR amplification methods, which is much easier to use by laboratory technicians, relatively inexpensive to perform and time saving. The options are the development of a SCAR of the RFLP marker, provided the marker is

polymorphic between the breeding parents, or the identification of additional linked markers with PCR amplification techniques such as RAPD or AFLP.

In order to reach this objective it is necessary to explore and optimize the use of RAPD in local soybean cultivars. A segregating population needs to be developed for linkage analysis of the phenotype with molecular markers. Both RAPD and AFLP analysis techniques will be evaluated for the identification of linked markers to nematode resistance in the segregating population and used for creation of a high density map around the resistance trait. SCAR markers will be designed from sequencing data of polymorphic fragments for use in marker assisted selection in breeding programs. Markers linked to the resistance QTL will be placed on the public soybean map using known RFLP probes.

CHAPTER 2

OPTIMIZATION OF THE RANDOM AMPLIFIED POLYMORPHIC DNA TECHNIQUE FOR SOYBEAN: USE OF RANDOM AMPLIFIED POLYMORPHIC DNA FOR IDENTIFICATION OF SOUTH AFRICAN SOYBEAN CULTIVARS

2.1. INTRODUCTION

Traditional means for identifying soybean cultivars involves extensive observations of morphological traits of mature plants and lacks the ability to prove uniqueness of closely related cultivars. Multigenic characteristics such as yield potential and photoperiodic sensitivity, traits important to soybean producers, are difficult to assess and consequently cannot be used in the registration and identification of unique cultivars. The development of more detailed genetic analysis to establish distinctness and uniformity is becoming increasingly important, especially for private breeding companies who need a method of documenting how their cultivars differ from those of competitors.

Several biochemical approaches have been used to distinguish soybean cultivars. These include isoenzyme analysis (LARSEN and BENSON, 1970; KIANG and GORMAN, 1983; DOONG and KIANG, 1987), RFLPs of nuclear DNA (APUYA, *et al.*, 1988; KEIM, *et al.*, 1989), chloroplast DNA (CLOSE, *et al.*, 1989), mitochondrial DNA (GRABAU, *et al.*, 1992) and ribosomal gene variation (DOYLE and BEACHY, 1985; DOYLE, 1988). Only low levels of diversity among soybean cultivars were detected by these means. GRABAU, *et al.* (1992) identified four cytoplasmic groups among 138 cultivars, including modern materials as well as older traditional lines. CLOSE, *et al.* (1989) divided 53 cultivars and plant introductions in six cytoplasmic groups based on RFLP patterns of chloroplast DNA. Ribosomal DNA did not vary among cultivated soybean lines; however, some variation was observed in wild relatives (DOYLE and BEACHY, 1985; DOYLE, 1988). Using randomly chosen DNA clones as probes, APUYA, *et al.* (1988) found that one in five probes revealed a polymorphism in genomic DNA of widely distant cultivars of soybean (*Glycine max* (L.) Merrill). Complex probes corresponding to repeated DNA revealed different polymorphisms in different cultivars and a single probe could be used to distinguish the five cultivars. Low molecular diversity was found among soybean cultivars (KEIM, *et al.*, 1989).

The major limitations of the afore-mentioned techniques are the low frequency of polymorphism and the laborious procedures involved. In addition, isoenzymes and other proteins may be influenced by environment, tissue source and developmental stage. The randomly amplified polymorphic DNA (RAPD)

method described by WILLIAMS, *et al.* (1990) provides a faster and less expensive alternative to RFLP analysis. This technique has been used in varietal identification of many crops including *Stylosanthes* (KAZAN, *et al.*, 1993b), papaya (STILES, *et al.*, 1993), celery (YANG and QUIROS, 1993) and apple (KOLLER, *et al.*, 1993). There are a few reports on the use of RAPDs in characterization of soybean (LARK, *et al.*, 1992; PRABHU and GRESSHOFF, 1994; PAIVA, *et al.*, 1994; CAETANO-ANOLLÉS, *et al.*, 1993b). WILLIAMS, *et al.* (1990) differentiated between two soybean species with RAPD analysis. LARK, *et al.* (1992) compared four soybean species with RAPD analysis consisting of 11 domesticated cultivars, nine wild soybeans and five perennials with RAPD analysis. The four species were well separated on a phylogram. The *G. max* cultivars sharing a common genetic background were clustered together. From adding or removing some plants with particular genetic components, other genetic relationships could be revealed.

The main objective of this study was to optimize the RAPD system for soybean and determine the level of polymorphism detected by the technique. Detection of an adequate number of polymorphism between breeding lines is an essential prerequisite for applicability in marker development. The applicability of the RAPD technique in the identification of local soybean cultivars was tested and the phylogenetic relationship among cultivars determined. The criteria for an efficient identification system are reliability, repeatability and economic viability.

2.2. MATERIALS AND METHODS

2.2.1 Materials

The morphological characteristics of 37 locally available soybean cultivars used in this study are listed in Table 2.1. Seed of these cultivars were obtained from plants grown in National cultivar trials. Primers for amplification were acquired from OPERON Technologies, Alameda, California, USA. Taq DNA polymerase was obtained from Promega Corporation, Madison, WI, USA. SeaKem LE agarose was used

Table 2.1 Morphological characteristics of soybean cultivars

CULTIVAR	GROWTH HABIT[1]	HILUM COLOUR[2]	FLOWER COLOUR[3]	PUBESCENCE[4]	MATURITY GROUP[5]	SEED COMPANY OR COUNTRY OF ORIGIN
A5308	D	F	P	B	V	Asgrow, Argentina
A5409	D	G	P	G	V	Asgrow, USA
A5678	D	IB	W	B	V	Asgrow, USA
A7119	D	IB	P	B	VII	Asgrow, USA
Bakgat	D	BI	P	B	III	Sensako
Bamboes	I	B	W	B	V	Saffola
Braxton	D	B	P	B	VII	USA
Columbus	I+D	IB	P	B	IV	USA
Crawford	D	IB	P	B	IV	USA
Duiker	I	Y	P	G	V	Zimbabwe
Dumela	D	B	P	B	VI	National Seeds
Edgar	D	LB+Y	P	F	VII	USA
Forrest	D	BI	W	B	V	USA
Geduld	I	B	W	B	VII	SA-ARC
HighveldTop	I	IB	P	B	V	Bührman Seeds
Hennops	I	B	P	B	VII	Saffola
Hutton	D	B	P	B	VIII	USA
Ibis	D	G	P	F	VIII	USA
Impala	I	G	P	G	VIII	Zimbabwe
Knap	I	IB	P	B	V	Bührman Seeds
Komati	D	IB	W	B	V	Zimbabwe
N80-2317	D	IB	W	B	VIII	SA-ARC
PAN494	I	LB	P	G	IV	Pannar
PAN577G	D	F	P	G	V	Pannar

CULTIVAR	GROWTH HABIT[1]	HILUM COLOUR[2]	FLOWER COLOUR[3]	PUBESCENCE[4]	MATURITY GROUP[5]	SEED COMPANY OR COUNTRY OF ORIGIN
PAN581	D	Bl	P	B	V	Pannar
PAN790	D	Bl	W	B	VII	Pannar
PAN812	D	Bl	P	B	VIII	Pannar
PAN855	I	Y	P	G	VIII	Pannar
Prima	I	B	P	B	V	Bührman Seeds
Ransom	D	B	P	B	V	Pannar
Roan	D	LB+Y	W	G	VIII	Zimbabwe
SNK60	D	Bl	P	B	VII	Sensako
SSS3	I	LB-DB	P	G	VIII	SA-ARC
Sabie	D	B	P	G	V	Zimbabwe
Success	I+D	IB	P	B	IV	Bührman Seeds
Usutu	I	IB	W	B	III	SA-ARC
Wilger	I	B	P	B	V	Saffola

1: D: Determined growth habit I: Indetermined growth habit
2: B: Brown Bl: Black
IB: Imperfect black LB: Light brown
DB: Dark brown F: Fawn
G: Grey Y: Yellow
3: P: Purple W: White
4: B:Brown G: Grey
F: Fawn
5: Grouped according to USA reference standards.

in all experiments, unless otherwise specified and was obtained from FMC Bioproducts, Rockland, Maine, USA.

2.2.2 *Extraction of DNA*

Mature seeds were surface sterilized with 70% (v/v) ethanol for 1 min, soaked in 1.75% (m/v) NaOCl for 10 min, followed by three rinses of sterile distilled water. The seeds were soaked overnight in sterile water. DNA was extracted from 5 to 10 embryos of each cultivar to account for interplant variation using a modification of the method of ROGERS and BENDICH (1988). Tissue was cut up and ground to a fine suspension in 2 x CTAB buffer (2% (m/v) cetyltrimethylammonium bromide (CTAB), 100 mM Tris-HCl (pH8.0), 20 mM EDTA, 1.4 M $CaCl_2$, 1% (m/v) polyvinylpirolidone) at 65°C. The suspension was kept at 65°C for 5 min and an equal volume of chloroform : isoamylalcohol (24:1 (v/v)) was added. The suspension was mixed thoroughly and the resulting emulsion was centrifuged for 30 sec (at 12 000 x g). One-fifth volume of a 5% (m/v) CTAB solution (5% (m/v) CTAB, 0.35 M NaCl) was added to the supernatant. The chloroform: isoamylalcohol extraction was repeated and an equal volume of CTAB precipitation buffer (1% (m/v) CTAB, 50 mM Tris-HCl (pH 8.0), 10 mM EDTA) was added. The suspension was left at 4°C for about 15 min and centrifuged at 12 000 x g for 1 min. The DNA pellet was rehydrated in high-salt TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 M NaCl). The DNA was precipitated overnight in two and a half volumes of absolute ethanol at -20°C. The precipitate was collected by centrifugation at 12 000 x g for 15 min and washed with 70% (v/v) ethanol for 5 min. The DNA pellet was lyophilized and rehydrated in 0.1 x TE buffer (1 mM Tris-HCl (pH8.0), 0.1 mM EDTA) and treated with 0.1 µg μL^{-1} DNase-free RNase, prepared by heating a stock solution of 10 mg mL^{-1} RNase (Roche Boehringer Mannheim, Randburg, South Africa) for 10 min at 94°C. The concentration of the DNA was determined by UV spectrophotometry at 260 nm.

2.2.3 *Amplification of DNA*

The polymerase chain reaction was carried out according to the method of WILLIAMS, *et al.* (1990) in 25 µL reactions containing 25 ng soybean DNA, 5 pmol primer, 2 mM $MgCl_2$, 100 µM each of dATP, dCTP, dGTP and dTTP, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1% (v/v) Triton X-100, 0.5 U Taq DNA polymerase. Amplification was done in a Hybaid Thermal Cycler (Hybaid Limited, United Kingdom) for 45 cycles, each cycle consisting of 94°C for 1 min, 36°C for 1.5 min and 72°C for 2.5 min. Products were analysed by

electrophoresis in 1.5% (m/v) agarose gels at 80 V for 2.5 h or 3% (m/v) Nusieve agarose gels at 60 V for 4 h, containing 1 µg mL^{-1} ethidium bromide and photographed under UV light with Polaroid 667 film.

Optimization of the following parameters was done: $MgCl_2$ concentration (Fig. 2.1a), DNA concentration (Fig. 2.1b), primer concentration (Fig. 2.1c) and Taq polymerase concentration (Fig. 2.1d). Amplification with primer OPD-02 was performed with a slightly different program to get more reproducible results; 94 °C for 1 min, 36 °C for 1.5 min and 72 °C for 2 min for 6 cycles, followed by 40 cycles consisting of 94 °C for 1 min, 46 °C for 1.5 min and 72 °C for 2 min. The reaction mix contained 3 mM $MgCl_2$. Products were analysed on an agarose gel as above.

For optimization, 120 primers (Kits A through F) and 60 combinations of 2 primers each were used in separate amplification reactions with DNA from 10 different cultivars to identify polymorphic bands. Each reaction was repeated at least twice to test for reproducible polymorphisms. Repeatability was tested in two ways. Two concentrations of DNA differing tenfold (2.5 ng and 25 ng) were used in amplification reactions (Fig. 2.2a). Only fragments that amplified similarly at both concentrations of DNA were considered to be useful polymorphisms. Polymorphisms identified with this procedure were confirmed by repeating the reactions at least three times with 25 ng DNA (results not shown). To verify that the polymorphisms were not artifacts from accidental seed mixture, new DNA extractions were used in a separate set of amplification reactions. DNA from different sources of seeds were amplified (Fig. 2.2b) to determine whether the production environment had any influence on the banding patterns.

2.2.4 Data analysis

Polymorphisms among cultivars were scored as present or absent and calculated using the index of genetic distance (1-F). The number of shared fragments (F) between two cultivars was estimated using the method of WANG and TANKSLEY (1989) based on the theory of NEI and LI (1979), using the formula $F=2m_{xy}/(m_x+m_y)$, where m_{xy} is the number of RAPD fragments shared by the two cultivars x and y and m_x and m_y are the total number of fragments scored in each cultivar. A cluster analysis was performed using the Proc Tree procedure of the SAS computer program (SAS Institute, 1992), based on the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and a phenogram was constructed.

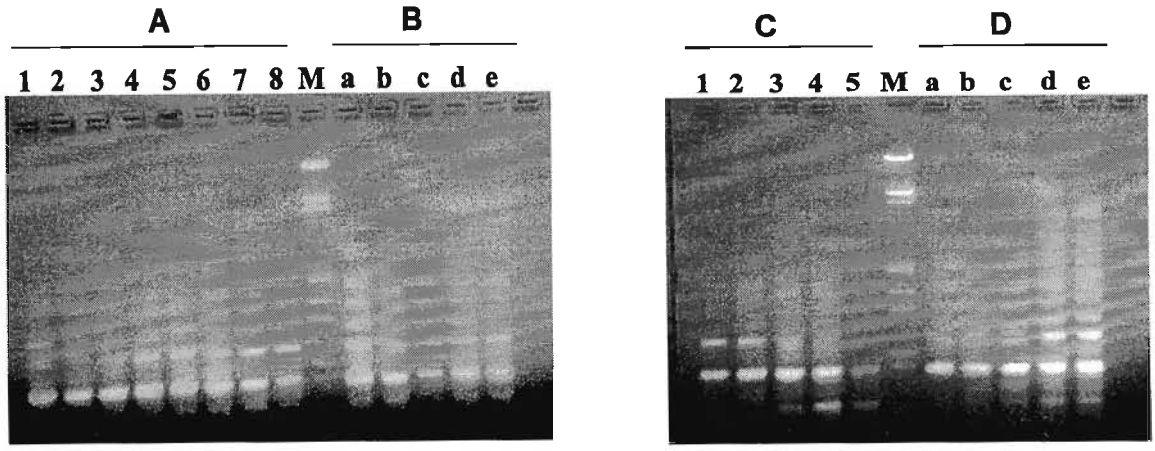


Figure 2.1: Optimization of conditions for PCR reactions with RAPD primers.

(A) $MgCl_2$ concentration 1) 1.5 mM, 2) 1.75 mM, 3) 1.9 mM, 4) 2.0 mM, 5) 2.1 mM, 6) 2.25 mM, 7) 2.5 mM and 8) 3.0 mM. M=Molecular weight marker (λDNA - *EcoRI/ HindIII*).

(B) DNA concentration 1) 2.5 ng, b) 10 ng, c) 25 ng, d) 50 ng and e) 100 ng.

(C) Primer concentration 1) 2.5 pmol, 2) 5 pmol, 3) 10 pmol, 4) 20 pmol and 5) 50 pmol.

(D) Taq DNA polymerase concentration a) 1.25 U, b) 0.5 U, c) 0.75 U, d) 1 U and e) 2U.

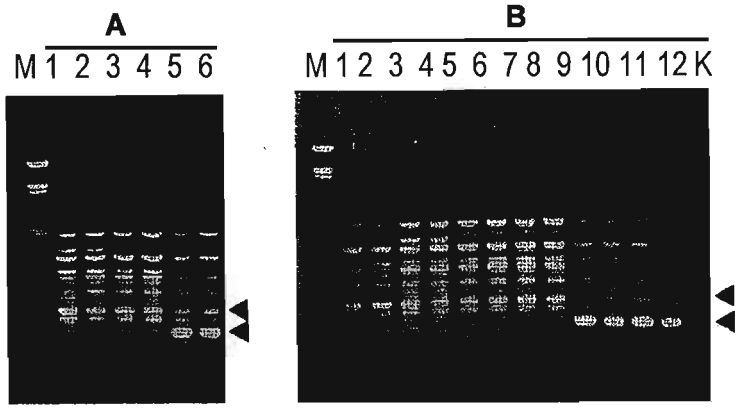


Figure 2.2: RAPD analysis of different cultivars with the same primer and difference in DNA concentration (A) and comparison of banding patterns from different sources of DNA (B).

(A) Fragments generated by primer OPB-07 with Prima (lanes 1 and 2), Knap (lanes 3 and 4) and Ibis (lanes 5 and 6). Lanes 1, 3 and 5 were done with 2.5 ng DNA and lanes 2, 4 and 6 with 25 ng DNA. M: Molecular weight marker (λ DNA digested with *EcoRI*, *HindIII*).

(B) DNA (25 ng) of seed from different sources amplified with OPB-07: Prima (lanes 1 to 4), Knap (lanes 5 to 8) and Ibis (lanes 9 to 12). M: Molecular marker, K: Control without DNA.

2.3. RESULTS

2.3.1 RAPD banding patterns

Different parameters were optimized for the RAPD reaction (Fig. 2.1). Mg^{2+} ion concentration did not have a dramatic effect on banding patterns in the range tested. The recommended concentration of 2 mM was therefore used. The same banding patterns were observed over a range from 2.5 ng to 100 ng DNA per reaction. A primer concentration of 10 to 50 pmol gave more bands, possibly due to non-specific binding. The recommended concentration of 5 pmol per reaction was used in all experiments, as this was the lowest concentration giving a clear banding pattern, taking into account economic considerations. Increase of the Taq DNA polymerase concentration lead to amplification of more fragments and the formation of a non-specific smear. The optimum concentration was determined to be 0.5 U to 0.75 U per reaction.

In the RAPD analysis of DNA of the different cultivars, 42 (35%) of the primers tested yielded differences in banding patterns among cultivars. The number of bands amplified by individual primers varied from single bands to complex patterns of more than 10 bands. Minor bands tended to be inconsistent on replication. Fragments generated by the primers ranged in size from 300 bp to approximately 2 kb. Combinations of primers (OPF-10 with primers from Kit A through C) gave a different banding pattern as the individual primers. Adjustment of Mg^{2+} ion concentration had no positive effect on the number of consistent polymorphisms generated, except in the reaction of OPD-02, where a higher Mg^{2+} concentration as well as an adjustment of the temperature program gave a more reproducible pattern (results not shown). Only primers answering to both criteria for reproducibility were used in the statistical analysis. Figure 2.1a shows consistent banding patterns of three cultivars at a tenfold difference in DNA concentration with primer OPB-07. Amplification reactions with DNA of seed from different sources in the country gave consistent banding patterns for the cultivars tested (Fig. 2.2b). Seed impurities can, however, pose some problems as the technique is very sensitive.

A total of 830 bands were amplified, giving an average of 4.6 bands per primer. Differences in banding patterns were scored as polymorphisms only if they were amplified consistently. Fourteen primers revealed a total of 22 polymorphisms that could be scored reliably (Table 2.2). The size of the polymorphic fragments ranged from approximately 365 bp to 1100 bp. These polymorphisms could be used to differentiate

Table 2.2 Amplified polymorphisms used for identification of cultivars

PRIMER	SEQUENCE (5'-3')	POLYMORPHIC BANDS (bp)
OPA-07	GAAACGGGTG	810
OPA-19	CAAACGTCGG	660, 620
OPB-03	CATCCCCCTG	1100
OPB-07	GGTGACGCAG	680, 365
OPC-02	GTGAGGCGTC	710, 450, 355
OPC-07	GTCCCGACGA	560
OPC-15	GACGGATCAG	1020, 830, 560
OPC-20	ACTTCGCCAC	710
OPD-02	GGACCCAACC	800
OPD-16	AGGGCGTAAG	810, 690
OPE-01	CCCAAGGTCC	780
OPE-14	TGCGGCTGAG	600
OPF-10	GGAAGCTTGG	920, 810
OPA-12+OPF-14	TCGGCGATAG	820
	TGCTGCAGGT	

among the 37 cultivars. The remaining amplified fragments were either monomorphic for all the cultivars or were polymorphic but were not reliable. The sequences of the selected primers are listed in Table 2.2.

2.3.2 Statistical analysis

The DNA fingerprints generated were used to obtain 1-F values of the South African soybean cultivars. These data were used in a cluster analysis (Fig. 2.3) to generate a phenogram. The analysis revealed a similarity of >80% between the cultivars tested. Four main groups were clustered together (pairwise clusters do not mean single band differences).

2.4 DISCUSSION

The results indicate that RAPD markers can be used for cultivar identification in soybean, although relatively few polymorphisms could be amplified reliably with the primers tested. The 14 primers selected amplified 22 polymorphic bands that differentiated the 37 cultivars. The frequency of polymorphisms

detected by the 180 primers was 35% or 2.6% of the total of 830 bands generated. As could be expected, the frequency of polymorphisms detected is relatively low because the soybean genome is highly conserved among cultivars (APUYA, *et al.*, 1988), with a narrow genetic base.

In comparison to other techniques, RAPD analysis is relatively simple and inexpensive. Preparation of samples is easy and requires minimum time. One disadvantage is its sensitivity to changes in reaction conditions, which might lead to differences in results from different laboratories. This can be overcome, however, by thorough standardisation of conditions. RAPD markers can be converted to SCARs (sequence characterized amplified regions), which give a more stable reaction that is less sensitive to variation in reaction conditions. The primers selected proved to give reproducible banding patterns at a tenfold difference in DNA concentration. The primers also gave consistent results with DNA from seed produced in different environments.

Relatedness of genotypes were tested by means of a cluster analysis for similarity of polymorphisms. Despite the diversity in geographic background such as the USA, Argentina, Zimbabwe and South Africa, results showed all cultivars to be within 80% similarity. Four main clusters were formed. The Zimbabwe cultivars namely Roan, Duiker, Impala, and Komati (all from the same breeding program) were grouped within 90% similarity. Four of the South African cultivars namely Knap, Highveld Top, Prima, and Success had common parentage. Highveld Top and Prima grouped within 95% similarity, with Success at 87% in the same main group. Knap however, showed a similarity distance equal to the maximum reported. In order to demonstrate the relatedness of Knap therefore, a greater portion of the genome will have to be amplified. Because most cultivars are privately owned, the ancestral background could not be revealed.

Determination of genetic relationships among cultivars based on polymerase chain reaction must be handled with caution. RAPD-amplified fragments do not necessarily represent the entire genome and addition or elimination of primer data can cause apparent reshuffling of relationships between cultivars.

2.5 CONCLUSIONS

It was demonstrated that the RAPD technique could be efficiently applied for fingerprinting of soybean cultivars. The usefulness of RAPD in establishing genetic relatedness in soybeans, however, was not illustrated convincingly in this study, and need to be investigated further. Despite the fact that the chosen cultivars represented a wide diversity in geographic background, a relatively high degree of genetic

similarity (80%) was reported for the chosen primers. Variation in the banding patterns was not enough to clearly demonstrate genetic relatedness. A greater portion of the genome may be needed to do this. This imply a search for additional primers over and above the 180 tested in this study.

The RAPD technology was thoroughly optimized and proved efficient and reliable in distinguishing 100% of the soybean (*Glycine max* L. Merr.) genotypes used in this study. The polymorphisms repeated consistently for the selected primers, and was not affected when using seeds from varied cropping environments. The level of polymorphism detected was about 35%, which compared well to the level obtained in other RFLP studies (42%) (TAMULONIS, *et al.*, 1997b). Therefore, it was decided to use this methodology to develop a marker linked to nematode resistance. This is outlined in Chapter 3.

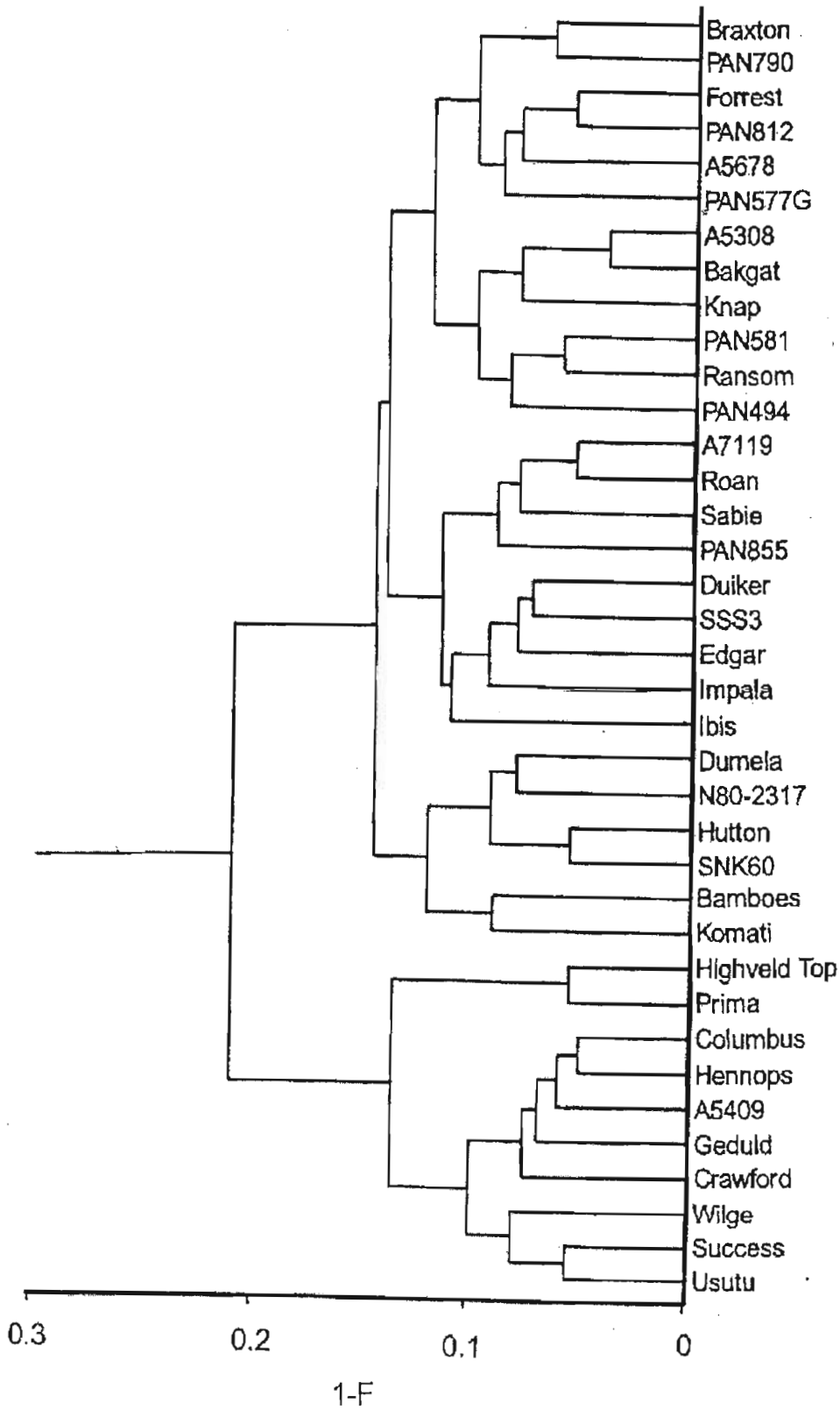


Figure 2.3: Phenogram of soybean cultivars generated with a cluster analysis of 1-F values from pairwise comparisons of RAPD fragments between cultivars.

CHAPTER 3

DIFFICULTY IN DEVELOPMENT OF MAPPING POPULATIONS

3.1 INTRODUCTION

The crucial part of any marker development programme is the phenotypic characterization of the mapping population. The precision with which each genotype can be identified determines the efficiency with which the location of the markers linked to the genes can be achieved. This can be problematic as the best application for marker assisted selection is usually when the phenotype is difficult to evaluate, or where a high genotype by environment interaction exists (DE O. ZIMMERMANN, 1993). Marker assisted selection improves the efficiency of selection for the target traits, and it can speed up the procurement of lines with the desired combination of genes. It can also assist in the introduction of recessive genes in backcrossing programmes and the pyramiding of desirable genes in a superior variety (DE O. ZIMMERMANN, 1993). Both of these would be difficult and tedious with conventional breeding methods.

Breeding for nematode resistance is certainly no exception. Additional factors to be considered are problems inherent to the inoculation technique such as uniformity of inoculum, aggressiveness of the nematode population on the soybean genotype, egg hatching and nematode feeding habits; environmental effects such as temperature, soil temperature and humidity, soil composition and light intensity (HUSSEY and BOERMA, 1981). The evaluation method of the reaction of the plant to nematode infestation is an additional, rather controversial, issue. These factors, together with the difficulties in obtaining true crosses in soybean, complicates the development of a reliable segregating population.

The first critical step in marker development for a specific trait would be the choice of parents displaying contrasting phenotypes. These varieties are crossed according to traditional methods to obtain a mapping population which can be F_2s, backcrosses, recombinant inbred lines, etc., which is then used for genotyping the phenotypic trait. The DNA of each plant is extracted and polymorphisms identified with a suitable method, which are then linked to the phenotype with statistical methods.

3.2 DEVELOPMENT OF A MAPPING POPULATION : PG3-1, PG3-2

3.2.1 Materials and methods

3.2.1.1 Plant material

Seed for the cultivars Gazelle and Prima were kindly supplied by the soybean division of the Oil and Protein Seed Centre, Grain Crops Institute, Agricultural Research Council, Potchefstroom, South Africa. Preliminary screening of various cultivars for resistance to *Meloidogyne javanica* was conducted by the nematology division of the Summer Grain Centre from the same Institute. Gazelle was identified as resistant and Prima as extremely susceptible (FOURIE, *et al.*, 1999). The heritage of Gazelle is illustrated in Figure 3.1.

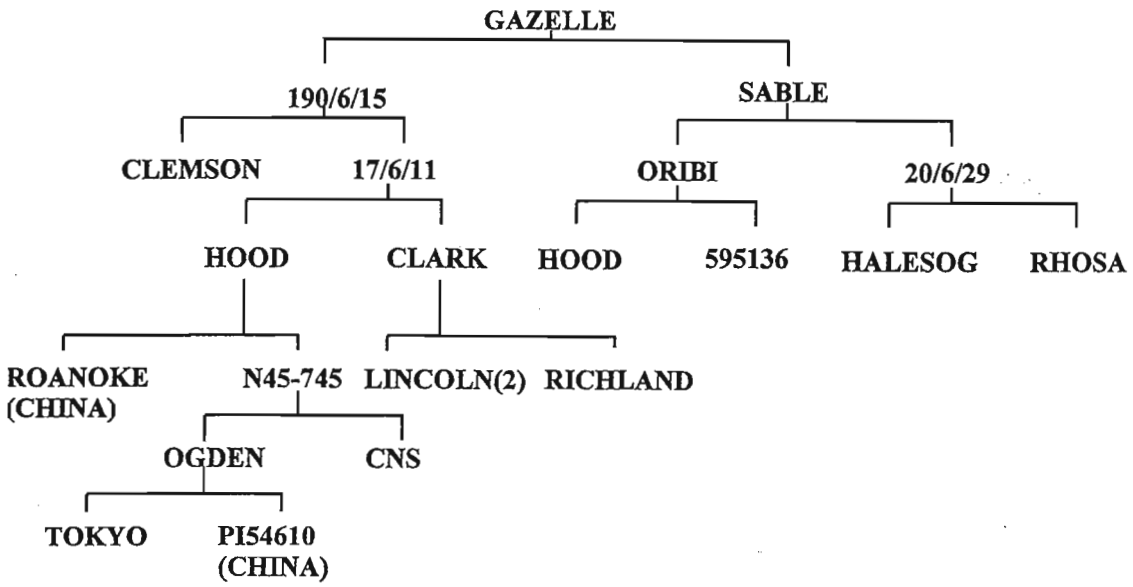


Figure 3.1: Heritage of the soybean cultivar Gazelle (R. TATTERSFIELD, personal communication, SeedCoop, Rattray Arnold Research Station, Harare, Zimbabwe).

3.2.1.2 *Greenhouse screening procedure*

Crosses were made between Prima (susceptible) and Gazelle (resistant) plants. Sandy soil containing less than 10% (m/m) clay was fumigated with methyl bromide before steam sterilization. Ten putative F_1 seeds were obtained. Seeds were coated with Thiulin (thiram - organic compound, Bayer, Isando, South Africa) and planted in 4 L pots in a greenhouse. Gazelle and Prima plants (parental genotypes) were randomly distributed between the test plants as positive and negative controls. Greenhouse conditions were maintained at 25°C day and 20°C night respectively. When trials were conducted during the winter season, the daylength was extended to 16 h with artificial light (±340 µmol m^{-2} s^{-1}) for counteracting photoperiodisity in soybean (RAPER and CRAMER, 1987).

Meloidogyne javanica eggs were collected from tomato plants (cv. Moneymaker) with 1% (m/v) NaOCl (RIEKERT, 1995). Twelve days after planting each seedling was inoculated with 5 000 nematode eggs (HUSSEY and BOERMA, 1981) at a depth of approximately 5 cm next to the roots. Plants were grown in a greenhouse and nutrient solution (2N:1P:2K with micro-elements, supplemented with urea) added weekly. Water was provided daily through a drip irrigation system in the saucers at the bottom of the pots.

Seeds were harvested from all putative F_1 plants. One hundred and twenty eight F_2 plants (progeny originated from 2 F_1 plants from a single pod) were planted in a complete randomized design with positive and negative control plants randomly distributed between them. The plants were inoculated with 10 000 *Meloidogyne javanica* eggs each. The number of eggs were increased (compared to that used with F_1 plants) to achieve a higher disease pressure on these plants.

3.2.1.3 *Resistance evaluation*

Plants were grown to seed maturity (approximately 115 days post inoculation) and the shoots excised just before the plants ripened. At this stage viable seed could be obtained, while the root systems were still intact. The root systems had to be evaluated before degeneration occured and sustainability of the nematode population would be declining. The seeds were left to dry on the upper parts of the plants in brown paper bags. The soil was washed from each root system and gall formation evaluated. Due to the long growth period and heavy infestation, it was impossible to count individual galls on the roots. An alternative classification system was used to analyse the nematode infestation. Root systems were classified relative to the extremes, Gazelle and Prima (Figure 3.2), in 5 main classes, with 1=very little or

no gall formation (Gazelle) 3, 5 and 7 intermediate and 9=severe gall formation (Prima), giving an estimated gall index for each plant. All root systems were weighed and cut into pieces, and the nematode eggs extracted with 1% (m/v) NaOCl. Eggs and larvae were extracted from the suspension through a set of sieves including from top to bottom: 710, 250, 75, 63, 45 and 10 m mesh sizes respectively. The total number of eggs produced was determined for each root system by counting an aliquot of the egg suspension and multiplying it by the appropriate dilution factor. These were used to calculate a reproduction factor (Rf-value) by dividing the total egg count by the initial number of eggs used for inoculation.

A B

Figure 3.2: Root systems of (A) Gazelle and (B) Prima soybean plants 115 days after inoculation with 10 000 *M. javanica* eggs.

Possible statistical relationships between the F_2 populations and controls were determined with t-tests and analysis of variance.

3.2.1.4 DNA extraction with CTAB

DNA was isolated from lyophilized leaf material by a modified CTAB extraction procedure (DELLAPORTA, *et al.*, 1983). Leaf material was ground to a fine powder with silica gel and suspended in CTAB extraction buffer (100 mM Tris-HCl (pH 8.0), 20 mM EDTA, 1.4 M NaCl, 2% (m/v) CTAB, 0.2% (v/v) ß-Mercaptoethanol) and incubated for 1 h at 65 C. The suspension was extracted with chloroform : isoamylalcohol (24:1 (v/v)) and the phases separated by centrifugation at 10 000 x g for 10 min. The DNA was precipitated from the aqueous phase with 0.66 volume isopropanol at room temperature, and centrifuged for 15 min at 12 000 x g. The precipitate was washed with 70% (v/v) ethanol and airdried. The

pellet was resuspended in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) and precipitated with 0.75 M ammoniumacetate and 2 volumes absolute ethanol after a chloroform : isoamylalcohol (24:1 (v/v)) extraction. After an overnight incubation at -20°C, the DNA was recovered by centrifugation at 12 000 x g for 15 min and washed twice with 70% (v/v) ethanol for 5 min. The ethanol was removed and the pellet airdried. The DNA was resuspended in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) and treated with 0.1 μg μL^{-1} DNase-free RNase, prepared by heating a stock solution of 10 mg mL^{-1} RNase (Roche Boehringer Mannheim, Randburg, South Africa) for 10 min at 94°C. The concentration of the DNA was determined spectrophotometrically at 260 nm.

3.2.1.5 DNA extraction with SDS

An alternative extraction method was followed for small samples of lyophilized leaf material (EDWARDS, *et al.*, 1991). Leaf material was ground as before and suspended in buffer (200 mM Tris-HCl (pH 8.0), 250 mM NaCl, 25 mM EDTA, 0.5% (m/v) SDS (sodium dodecyl sulphate)). The suspension was incubated at 60°C for 30 min and extracted with an equal volume of chloroform : isoamylalcohol (24:1 (v/v)). The phases were separated by centrifugation at 10 000 x g and the DNA precipitated from the aqueous phase with 2 volumes ice-cold absolute ethanol. The DNA was recovered by centrifugation at 12 000 x g for 15 min and washed with 70% (v/v) ethanol for 5 min. The pellet was resuspended in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) and treated with RNase A as described above. The chloroform : isoamyl-alcohol (24:1 (v/v)) extraction was repeated and the DNA precipitated with 0.75 M ammoniumacetate and 2 volumes absolute ethanol overnight. The tubes were centrifuged at 12 000 x g and the pellets were airdried before resuspension in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA). The concentration of the DNA was determined spectrophotometrically at 260 nm. The quality of the DNA was verified by electrophoresis of 1 μL of the DNA sample on an 0.8% (m/v) agarose gel at 60 V for approximately 2 h.

3.2.1.6 RAPD analysis

DNA from leaf material of the parent and hybrid plants was subjected to RAPD analysis (Chapter 2.2.3) for determination of hybrid authenticity.

The RAPD technique was used for screening DNA from leaf material of F_2 plants for markers linked to the resistance trait. The concept of bulked segregant analysis was followed (MICHELMORE, *et al.*, 1991). DNA from six plants from each of the extremes (Table 3.1) of the phenotypic values, taking both gall index

and Rf-values into account, were constructed by mixing approximately equal amounts (m/m) of each individual DNA sample. Samples were diluted to approximately 25 ng μL^{-1}. Primers amplifying polymorphisms between the bulks were tested on individuals of the bulks, followed by testing of more randomly selected individual plants.

Table 3.1 Composition of resistant and susceptible DNA bulks.

RESISTANT BULK			SUSCEPTIBLE BULK		
PLANT	RF	GI	PLANT	RF	GI
PG3-2-68	1.59	2	PG3-2-39	44.94	9
PG3-1-81	0.83	2	PG3-1-20	46.27	9
PG3-1-90	3.22	2	PG3-1-7	58.03	9
PG3-2-56	3.78	2	PG3-1-19	78.66	9
PG3-2-67	3.78	2	PG3-2-32	77.88	9
PG3-2-64	0.72	3	PG3-2-49	44.59	7

3.2.1.7 DNA marker analysis

Linkage to the resistance trait was determined with a general linear model of the STATGRAPHICS Plus computer program (Manugistics, Rockville, Maryland, USA, 1998), using genetic marker data as the independent and gall index or reproduction factor values as the dependent variable. The association between the DNA marker and the trait was considered significant if the probability was <0.05. The coefficient of determination (R^2) was used as a measure of the magnitude of association. Interaction between markers was determined with a two-way analysis of variance from the same program (STATGRAPHICS Plus). The analysis was applied to log transformed reproduction factor data, initially with results from the 12 individual plants of the bulks. Putative markers were tested on the larger population using the same statistical methods. A significant association between a marker and the response to reproductive support was declared if the probability was <0.05.

3.2.1.8 RFLP analysis

The restriction enzyme *(TaqI)* was purchased from Roche Boehringer Mannheim, Randburg, South Africa. Positively charged (magnacharge) MSI Nylon membranes were obtained from Micron Separations, Westborough, MA, USA. Soybean genomic DNA probe B212, as developed and mapped by Dr R.C. Shoemaker of Iowa State University, Ames, Iowa, were supplied by Biogenetic Services, Inc., Brookings,

SD, USA. The probe were cloned in the *PstI* site of the pBS+ vector. Labelling and detection kits, as well as ready to use DIG-labelled molecular weight marker III and X-ray film were obtained from Roche Boehringer Mannheim. T3 (5'-ATT AAC CCT CAC TAA AGG GA-3') and T7 (5'-TAA TAC GAC TCA CTA TAG GG-3') promoter primers were supplied by Promega Corporation, Madison, WI, USA. Hybridization and detection were carried out using DIG Easy Hyb, Anti-digoxigenin-AP Fab fragments, DIG wash and block buffer set and CDP-Star™ from Roche Boehringer Mannheim.

3.2.1.8(i) *Restriction digestion*

Approximately 10 µg plant genomic DNA was digested for 6 h with 5 U of *TaqI at 65°C*. The restricted fragments were separated by electrophoresis in a 0.8% (m/v) agarose gel at 20 V for 16 h. DIG-labelled molecular weight marker was co-electrophoresed for easy determination of fragment sizes after detection.

3.2.1.8(ii) Southern transfer

DNA fragments were blotted onto Nylon membranes according to the method developed by SOUTHERN (1975). The DNA was depurinated with 250 mM HCl for 10 min and denatured in the gel with 0.5 M NaOH, 1.5 M NaCl. The gel was neutralised in 0.5 M Tris-HCl (pH 7.5), 3 M NaCl and the DNA transferred to the membrane overnight in 10 x SSC (1.5 M NaCl, 150 mM tri-sodiumcitrate, pH 7). DNA was immobilised on the membrane by baking at 80°C for 1 h. Membranes were stored in sealed plastic bags at 4°C.

3.2.1.8(iii) Probe labelling

The plasmid DNA containing the probe fragment was isolated from overnight bacterial cultures with a mini-prep procedure. Overnight cultures were centrifuged at 3 500 x g and resuspended in STET (50 mM Tris-HCl (pH 8.0), 50 mM EDTA, 8% (m/v) sucrose, 5% (v/v) Triton-X100). The bacterial cell walls were lysed with addition of 10 µg lysozyme and heating at 94°C for 1 min. The samples were cooled immediately on ice and the nuclear DNA was removed by centrifugation at 12 000 x g for 20 min. The plasmid DNA was precipitated with 0.66 volumes of isopropanol and resuspended in TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA). Presence of the probe fragment (correct size according to information from the supplier) was verified with restriction digestion of 10 µL of plasmid with 0.5 U *PstI* at 37°C for 2 h, followed by agarose gel (1.5% (m/v)) electrophoresis. Correct amplification of the probe insert of 2 000 bp with T7 and T3 promoters was verified before labelling the probe. The probe insert was amplified in a total reaction

volume of 25 µL containing 1 µL plasmid DNA (approximately 10-100 pg), Taq DNA buffer (10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1% (v/v) Triton X-100) 0.03 µM of each primer T3 and T7, 3 mM $MgCl_2$, 100 µM each of dATP, dCTP, dGTP and dTTP, 1 U Taq DNA polymerase. Amplification was done in a Hybaid Thermal Cycler (Hybaid Limited, United Kingdom) for 30 cycles, each cycle consisting of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 1 min. The insert was labelled with DIG-dUTP using the same temperature program, with the PCR DIG probe synthesis kit from Roche Boehringer Mannheim. The DIG-labelled dUTP mix was diluted with unlabelled dNTP to a ratio of 2 (labelled):3 (unlabelled) for optimum labelling efficiency. Labelling was carried out in a total reaction volume of 50 µL or 25 µL containing plasmid DNA (approximately 10-100 pg), 1 x reaction buffer with $MgCl_2$ (Expand™ High Fidelity buffer), 200 µM dNTP (2 µL DIG-dUTP; 3 µL dNTP in the 50 µL reaction volume), 0.3 µM each T3 and T7 primers, 2.6 U Taq DNA polymerase (or 1.3 U in 25 µL reaction). Efficiency of labelling was verified against a non-labelled control reaction on a 1.5% (m/v) agarose gel.

3.2.1.8(iv) Hybridization

Pre-hybridization of genomic DNA blots was conducted in DIG Easy Hyb solution (Roche Boehringer Mannheim) in a volume of 20 mL cm^{-2} at 50°C for 30 min. The hybridization temperature was determined empirically, as the exact GC content of the probe was not known, and the actual temperature for hybridization in DIG Easy Hyb should be approximately 20-25°C lower than the calculated T_m value, according to the manufacturer's instructions. An approximate T_m value was calculated using the formula

$$T_m = 49.82+0.41(\%G+C) - (600/l)$$ [l=length of hybrid in base pairs]

(Roche Boehringer Mannheim)

with estimated values of %G+C=50% and l=2000 bp.

The probe DNA was denatured at 94°C for 5 min and immediately cooled on ice before dilution in DIG Easy Hyb solution at a concentration of 0.5-0.75 µL mL^{-1}. The membranes were hybridized overnight in a volume of 3.5 mL 100cm^{-2} of this solution at 50°C.

The membranes were washed twice, 5 min per wash, in 2 x SSC-SDS (0.3 M NaCl, 30 mM tri-sodiumcitrate, 0.1% (m/v) SDS) at room temperature, followed by two washes of 15 min each in 0.5 x SSC-SDS (0.075 M NaCl, 7.5 mM tri-sodiumcitrate, 0.1% (m/v) SDS) to remove unbound probe.

3.2.1.8(v) Detection with DIG

Detection of hybridized DNA fragment was carried out with the DIG Wash and Block buffer set (Roche Boehringer Mannheim) according to the manufacturer's instructions. The membranes were equilibrated in washing buffer (0.1 M maleic acid, 0.15 M NaCl, pH 7.5, 0.3% (v/v) Tween 20) and blocked by gently agitating the membranes in blocking solution (1% (v/v) blocking solution in maleic acid buffer) for 30 min. The bound digoxigenin-labelled probe was detected immunologically with anti-DIG-AP antibodies conjugated to alkaline phosphatase at a concentration of 1:20 000 in blocking solution. The membranes were washed twice for 15 min each and equilibrated to detection buffer (0.1 M Tris-HCl (pH 9.5), 0.1 M NaCl). The alkaline phosphatase conjugates was detected with a chemiluminescent substrate, CDP-Star™ (0.25 mM) and exposed to X-ray film overnight at room temperature.

3.2.2 Results

3.2.2.1 F_1 characterization : PG3-1, PG3-2

Due to severe fungal disease (*Fusarium* spp.) of the F_1 plants, no reliable phenotypic evaluation of these plants could be obtained. DNA of the parents was screened for polymorphisms with RAPD analysis with a subset of 200 10-mer primers. Thirty of these primers amplified polymorphisms between the two parents and were subsequently used in further analyses. A number of putative F_1 plants were subjected to RAPD analysis, and two of the hybrid plants (PG3-1 and PG3-2), originating from crosses with Prima as seed parent and Gazelle as pollen parent, were used in further studies (Table 3.2). Four of the fragments (OPA4-2, OPA5-1, OPC16-1 and OPD16-2) were amplified in both PG3-1 and PG3-2 and could have been inherited only from the pollen parent, indicating authenticity of the hybrids. Progeny from these plants were used for DNA analysis and detection of linkage to the nematode resistance trait.

Table 3.2 RAPD analysis of F_1 soybean plants

BAND	GAZELLE	PRIMA	PG3-1	PG3-2	BAND	GAZELLE	PRIMA	PG3-1	PG3-2
A4-1	0*	1*	1	1	D16-2	1	0	1	1
A4-2	1	0	1	1	D16-3	0	1	1	1
A4-3	0	1	1	1	F14-1	1	0	0	1
A5-1	1	0	1	1	F14-2	0	1	1	1
A5-2	0	1	1	1	G7-1	1	0	0	0
A8-1	0	1	1	1	H12-1	1	0	0	0
A8-2	0	1	1	1	E12-1	0	1	1	1
A8-3	0	1	1	1	E12-2	0	1	1	1
A9-1	1	0	0	1	E12-3	0	1	1	1
A20-1	0	1	1	1	E20-1	0	1	1	1
A20-2	0	1	1	1	E20-2	1	0	0	0
A20-3	1	0	0	1	F7-1	0	1	1	1
A20-4	0	1	1	1	F7-2	1	0	0	0
C16-1	1	0	1	1	I14-1	0	1	1	1
D16-1	0	1	1	0	I14-2	1	0	0	0

*0 : Fragment absent; 1 : Fragment present

3.2.2.2 F_2 characterization : PG3-1, PG3-2

3.2.2.2(i) Phenotypic results of F_2 : PG3-1, PG3-2

Complete results for the phenotypic screening of individual plants (PG3-1 and PG3-2 progeny) are listed in Appendix A. The Rf-values for both the Prima control plants and the F_2 population exhibited a large variation (Figure 3.3). The Rf-value calculated for Gazelle was low and varied between 0.25 and 2.31 with an average value of 0.95±0.63, which indicated good resistance to the nematode. Large variation was found among Prima plants with values ranging from 7.00 to 78.40 with an average of 36.40±26.76, which indicated severe susceptibility to the nematode. The large variation in values of replicate plants could in part be attributed to environmental factors (eg. placement of pots in the greenhouse, watering, light intensity, etc.). The F_2 population displayed values intermediate to the two parents (Figure 3.3), with the average of the population (15.31±14.50) within the lower limit of standard deviation of the Prima controls.

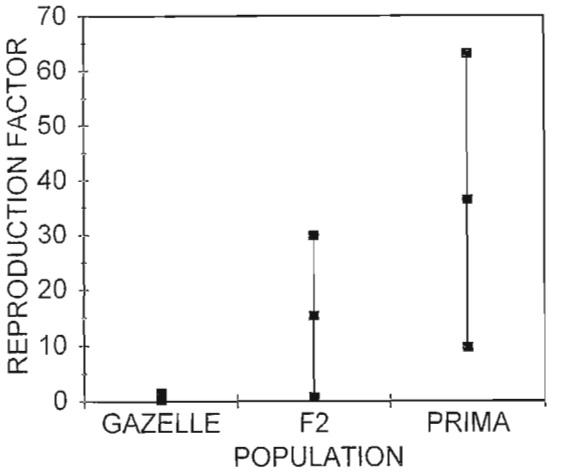


Figure 3.3: Distribution of reproduction values of nematodes for parent and F_2 populations, indicating standard deviation of the mean values.

Analysis of variance indicated a statistically significant difference (P=0.05) between the gall index values as well as the Rf-values of the two parent populations. Comparison of means of the Prima and F_2 progeny with a t-test indicated a statistically significant difference between the means of the two populations at the 95% confidence level (t=4.4). The frequency distribution of the plants for the gall index values (Figure 3.4) indicated a normal distribution over the five groups. With this evaluation method, no F_2 plants were found

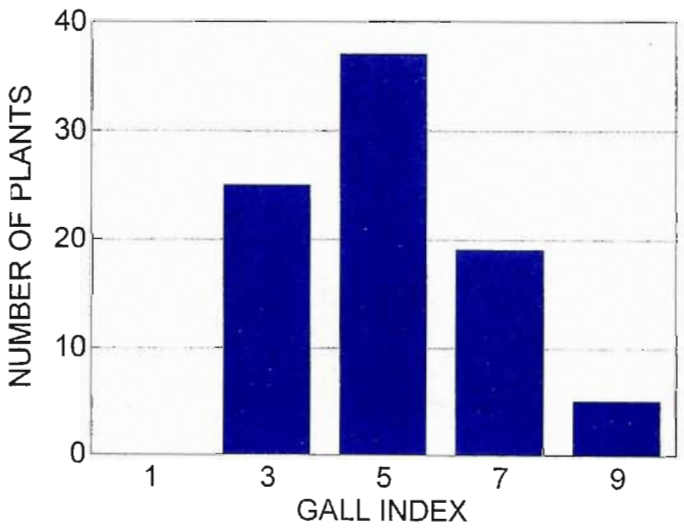


Figure 3.4: Frequency distribution of F_2 plants (PG3-1 and PG3-2) according to gall index.

with the same level of resistance as the resistant parent, Gazelle. Five F_2 plants showed severe gall formation resembling the susceptible parent, Prima. Analysis of the frequency distribution of the Rf-values for the F_2 population (Figure 3.5) displayed a continuous variation with a few plants in the upper range of the resistant parent, Gazelle. A few plants had extremely high Rf-values, indicating severe susceptibility to nematode reproduction.

3.2.2.2(ii) RAPD analysis of F_2 : PG3-1, PG3-2

DNA of the two parents, Gazelle and Prima, were screened with a further 320 primers, totalling 520 10-mer primers, for the presence of polymorphisms. These primers detected a total of 2983 loci, corresponding to an average of 5.7 loci/primer. Two hundred and sixteen of these primers (41.5%) gave rise to polymorphic bands between the two parents. Twenty of the polymorphisms detected could be classified as possible codominant loci. These (216) primers were used to screen the two bulk DNA samples comprising individual plants from the two phenotypic extremes. Fifty one of the primers amplified polymorphic fragments between the two bulks. Some problems were encountered with reproducibility of amplification of certain bands.

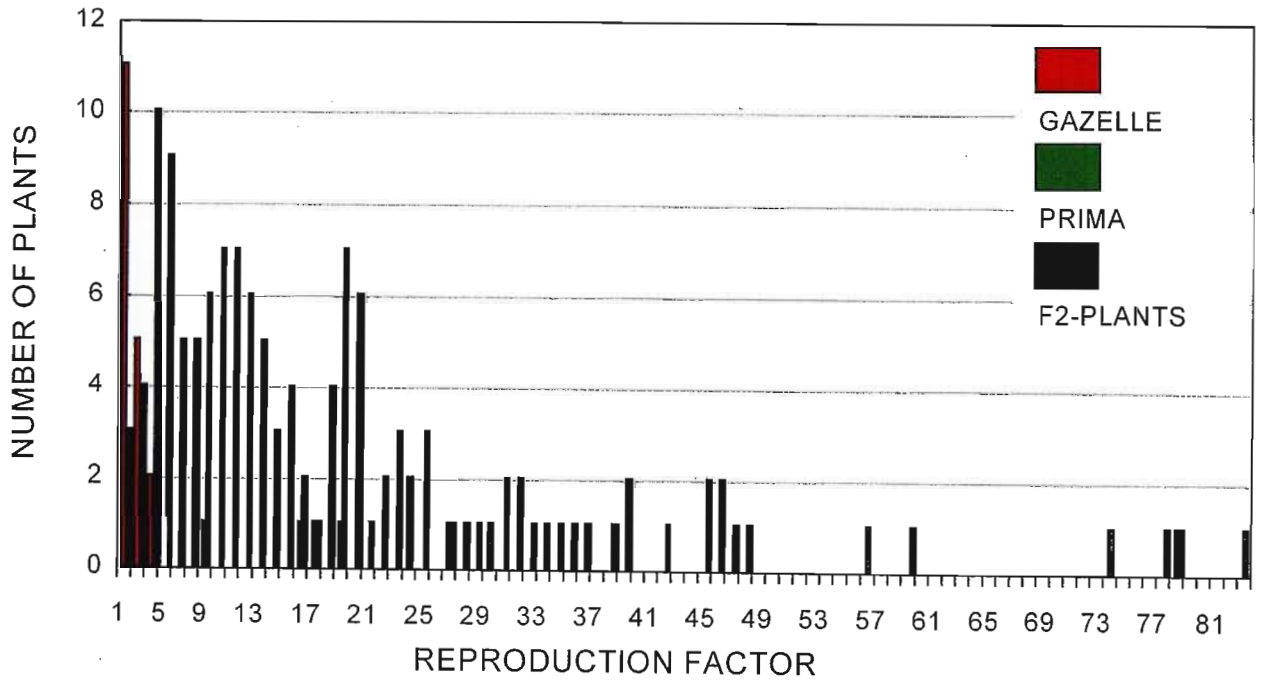


Figure 3.5: Frequency distribution of Rf-values of parent plants and F_2 progeny.

The incidence of polymorphism (41.5%) detected with the RAPD technique was considered sufficiently high to identify potential markers. The use of the two bulks grouped informative individuals in order to equalise the genetic background so that the groups differed theoretically mainly for the gene(s) in question. The effect could clearly be seen in the number of polymorphisms detected between the two bulks (from 51 primers), which was much less than between the two parents (216). RAPD analysis was repeated with these 51 primers with DNA of the 12 individual plants included in the 2 bulks (Table 3.1).

Table 3.3 Statistical analysis of potential markers

	F_2 BULKS[1]		F_2 POPULATION[2]	
FRAGMENT	R^2	P	R^2	P
OPC-16(1)	0.428	0.01*	0.01	0.21
OPR-04(2)	0.329	0.03*	0.01	0.21
OPT-08(1)	0.428	0.01*	0.00	0.95

[1] : 12 Individuals from the two DNA pools.

[2] : 52 individual F_2 plants.

Three putative markers were indicated (Table 3.3) with analysis of variance of the Rf-values of individual plants included in the two DNA pools. Fragments OPT-08(1) and OPC-16(1) each could account for 42.8% of variation and OPR-04(2) for 32.9%. As this analysis was done on a biased selection of plants from the two extremes, this would be expected to be a gross overestimation of the contribution of the markers to the phenotypic effect. OPT-08(1) was linked in coupling phase and the other two, OPC-16(1) and OPR-04(2), were linked in repulsion phase. OPT-08(1) differed quantitatively between the plants. This quantitative difference was repeatable, and could be due to more than one fragment with the same molecular weight migrating together. The other two polymorphisms were not always reproducible. RAPD analysis was repeated with these 3 primers with 40 additional randomly chosen F_2 plants, but no linkage with $P<0.05$ to the resistance trait could be confirmed (Table 3.3), based on reproduction factor values. Likewise, no significant linkage could be established between markers and gall index values of the F_2 population.

3.2.2.3 Re-evaluation of F_1

The DNA of F_1 (PG3-1, PG3-2) and parent plants were subjected to RFLP analysis for confirmation of authenticity (Figure 3.6). Both F_1 plants analysed displayed the same banding pattern as Prima, used as

seed parent. According to these results the plants tested were not true hybrids. Likewise, no segregation of bands could be confirmed in randomly selected F_2 plants (results not shown).

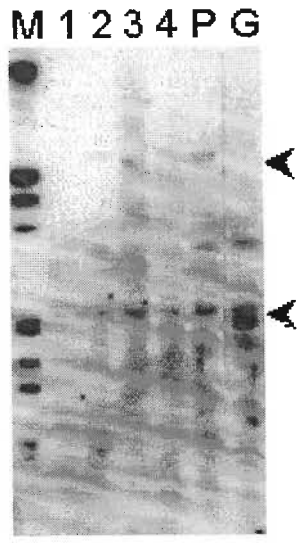


Figure 3.6 RFLP analysis of selected F_1 plants with probe B212. G:Gazelle; P:Prima; 1:PG1-1, 2:PG2-2, 3:PG3-1; 4:PG3-2, M:Molecular weight marker.

3.2.3 *Discussion*

RAPD analysis of the F_1 plants indicated hybrid authenticity in two plants and these were therefore used in the development of a segregating population. The F_2 progeny displayed large variation in their phenotypic response to the nematodes, with a normal distribution pattern. Comparison of means indicated a significant difference between the phenotypic patterns of the putative F_2 population and the parental controls (P=0.000). The DNA from these plants were subjected to RAPD analysis, which displayed polymorphic patterns, although problems occurred with reproducibility of the amplification reactions. None of the polymorphic fragments could however, be linked to the resistance trait.

The results obtained with the RAPD technique on the F_2 population raised questions as to the applicability of the technique for marker identification in this population. The problem could be three pronged: (i) it could be a problem with the reproducibility of the RAPD technique as applied in this soybean population and with the current equipment; (ii) it could be that only a small part of the genome was being sampled; or (iii) there could be inherent problems with the population used for screening. In the first case, both the

reproducibility and applicability of the technique were thoroughly tested in the previous Chapter. Thus, this possibility was ruled out. This led to a decision to re-evaluate the screening population and subsequent questioning of its authenticity. It was decided to re-evaluate the F_1 hybrids with an amplification independent technique.

The putative F_1 plants were subjected to RFLP analysis, using a probe which was polymorphic between the parents. The F_1 plants proved not to be true hybrids. It was therefore essential to develop a new mapping population.

3.3 DEVELOPMENT OF A MAPPING POPULATION : GP20-2

3.3.1 Materials and methods

A further series of crosses between Prima and Gazelle resulted in an additional number of F_1 seeds. These were planted in pots in the greenhouse and the first leaves were used for determination of the authenticity of the hybrids with RFLP analysis (Chapter 3.2.1.8). Two selected F_1 plants were inoculated with 5000 nematode eggs each and evaluated for the level of resistance to *M. javanica* 110 days after inoculation. Seed from one of the plants (GP20-2) were used in development of a mapping population.

Sixty seeds from GP20-2 were planted in a randomized pattern with Gazelle and Prima plants as positive and negative controls. All plants were inoculated with 7 500 nematode eggs approximately 14 days after planting. Ninety five days after inoculation, the plants were uprooted and evaluated for nematode resistance, according to both parameters (gall index and Rf-value of nematodes).

Five to 6 progeny from each of 29 F_2 plants were planted and inoculated with 5000 nematode eggs each and evaluated for resistance as before.

3.3.2 *Results*

3.3.2.1 *F_1 characterization* : GP20-2

3.3.2.1(i) *Phenotypic results of F_1 : GP20-2*

Two individual F_1 plants were evaluated for resistance to nematode reproduction. Egg counts per root system were 12 425 and 16 625 respectively, with an average of 14 525 eggs per root system. Reproduction factors of 2.49 and 3.33 respectively were calculated, giving an average of 2.91. Gall formation was not evaluated.

3.3.2.1(ii) *Genotypic characterization of F_1 : GP20-2*

Authenticity of the F_1 hybrids in the second round of crosses was established by comparing the banding patterns generated with RFLP analysis from leaf tissue with those of the parent plants (Figure 3.7). Four

of the hybrids tested, i.e. GP20-2, GP21-1, GP21-2 and GP23-1, displayed both polymorphic bands from Prima and Gazelle. This could only be explained by inheritance of bands from both parents, and these plants were therefore assumed to be true hybrids. Seeds from GP20-2 were used to establish a segregating F_2 population.

3.3.2.2 *Phenotypic results of F_2 progeny: GP20-2*

Complete phenotypical data for the GP20-2 F_2 population is tabled in Appendix B. Although the screening data between the first and second experiments could not be compared directly due to differences in the assay conditions (inoculate density, temperature, and other environmental factors), a strong correlation was observed between the control replicates of the resistant plants ($P<0.05$) for both measurements (gall index as well as Rf-values), with no statistically significant difference between the means of the two populations. The reproduction factor values for the resistant controls ranged between 0.60 and 1.26 (Appendix B) with an average of 0.91±0.34 (Figure 3.8b). The Prima controls again displayed a large variation in susceptible response according to reproduction values, ranging between 4.87 and 25.67

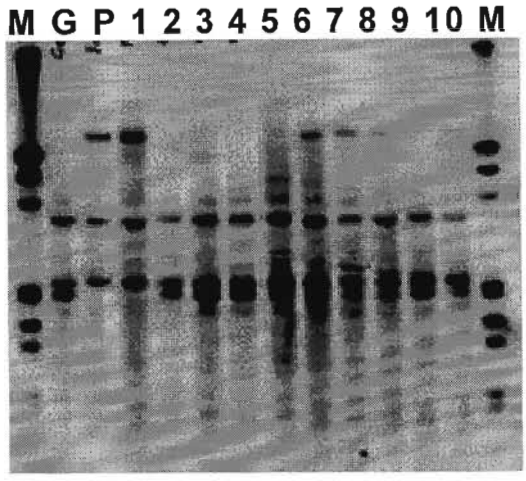
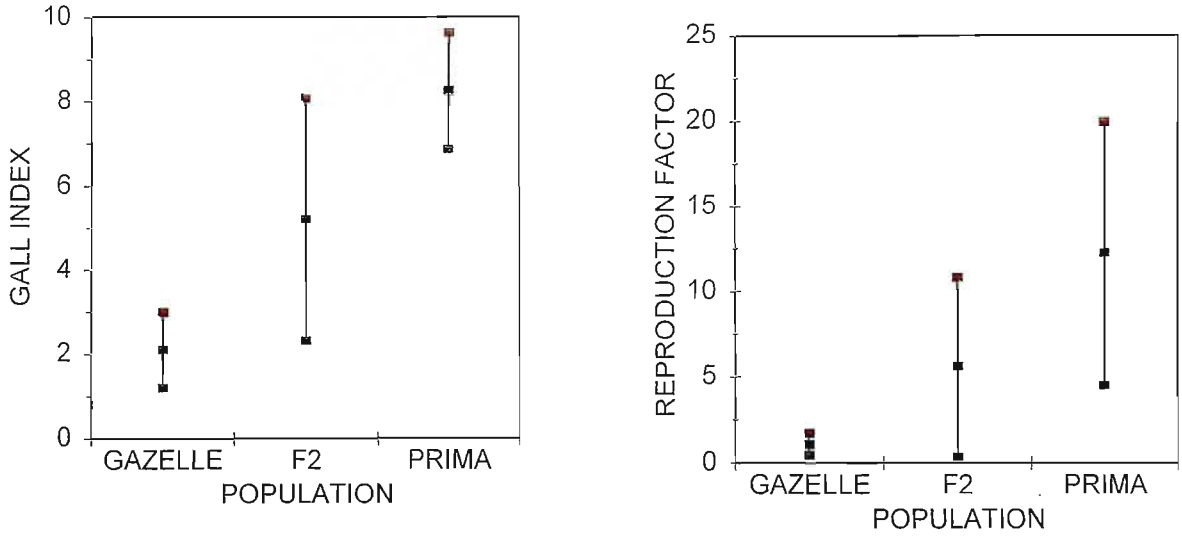


Figure 3.7 RFLP analysis of parental plants and 10 putative F_1 hybrid plants. M: Molecular weight marker; G:Gazelle; P:Prima; 1-10: F_1 hybrids (1)PG20-1, (2)GP5-1, (3)GP5-2, (4)GP6-1, (5)GP6-2, (6)GP20-2, (7)GP21-1, (8)GP21-2, (9)GP22-3, (10)GP23-1.

(Appendix B), with an average value of 13.64±7.19 (Figure 3.8b). The lower values compared to the first experiment can be ascribed to the shorter incubation period as well as the lower inoculate density. The gall index determinations displayed less variability in both control populations (Figure 3.8a).

The frequency distribution of the gall index values of the hybrid population was reclassified in 3 groups, i.e. resistant (1), intermediate (5) and highly susceptible (9) for illustrative purposes (Figure 3.9). The mean value of the F_2 population was more similar to the midparent mean than to either parental mean (Table 3.4) and had a frequency distribution significantly similar to a 1:2:1 relationship with χ^2=1.20 (P=0.55), indicating partially dominant inheritance. The Rf-values as determined for the individual F_2 plants displayed a continuous distribution (Figure 3.10). The mean value as calculated for the F_2 population matched that of the midparent mean (Table 3.4).



(a) (b)

Figure 3.8: Distribution of the (a) gall index and (b) reproduction factor values of nematodes for parent and F_2 populations, indicating standard deviation from the mean values.

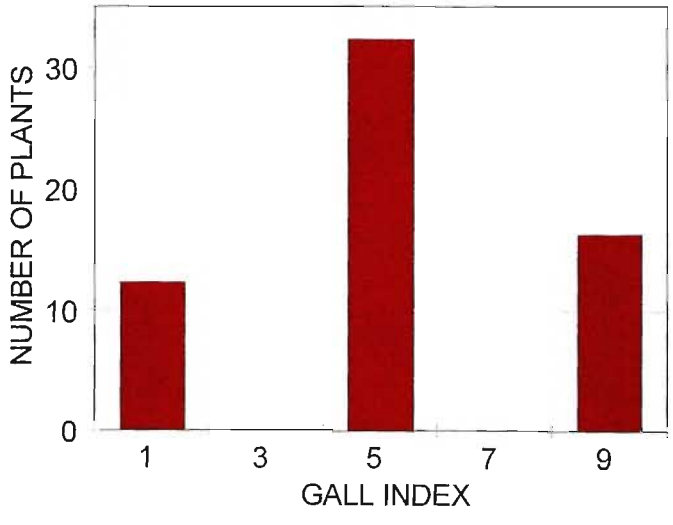


Figure 3.9: Frequency distribution of F_2 plants of the GP20-2 population according to gall index.

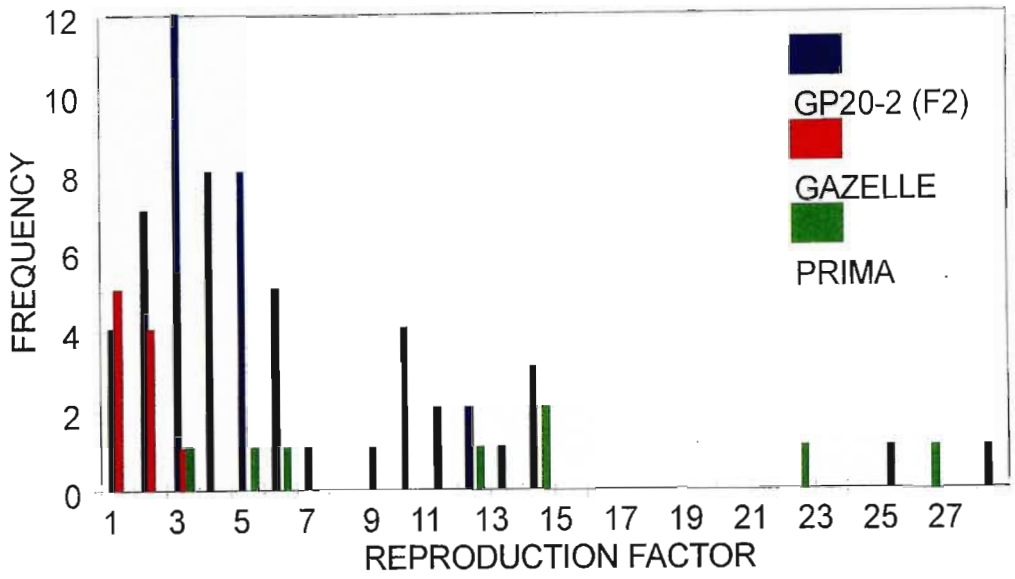


Figure 3.10: Frequency distribution of Prima, Gazelle and the GP20-2 F_2 population according to reproduction factor of the nematodes.

Table 3.4 Parent and progeny means for resistance evaluators

	P_1	MID	P_2	F_1	F_2
RF	0.91±0.28*	7.3	13.64±7.18	2.91	5.32±1.14
GI	1.4±0.90	5.3	8.5±1.18	ND	5.29±0.63

*95% confidence interval for the mean.

3.3.3 Discussion

True hybrids (according to RFLP analysis) with Gazelle as seed parent and Prima as pollen parent were obtained with the second batch of crosses made. During the performance of the crosses on these plants it was noticed that the pollen obtained from Gazelle was dry and much less abundant than that of Prima. This could be due to the low humidity and high temperatures experienced during the time of flowering. It would suggest that Prima was much better adapted to the greenhouse conditions than Gazelle.

Both the parameters evaluated, i.e. reproductive ability of nematodes and gall index of plant roots, displayed a continuous pattern of distribution. The gall index values could be classified in three groups with a 1:2:1 phenotypic ratio. This could indicate a partially dominant inheritance pattern with at least one gene involved in this particular cross. Closer examination of the Rf values of the F_1 and F_2 plants revealed that the F_1 plants had intermediate resistance, with 4 F_2 plants within the resistance level of Gazelle and 5 plants with a value above that of the mean value of Prima. Considering the formula (BURNS, 1976) $(¼)^n$ with n=2 genes, this data could fit a hypothesis of 2 genes contributing to the resistance trait. Thus, whether it is considered one gene with incomplete dominance or 2 genes, depends entirely on the level of analysis. The data was however, based solely on one cross. These findings are in accordance with results obtained by other groups. Resistance to *M. javanica* in three American genotypes, Gordon, PI80466 and PI230977 was found to be quantitatively inherited (LUZZI, *et al.*, 1995b). LUZZI, *et al.* (1995b) used more than one cross and therefore could make a more accurate suggestion on the inheritance of the trait in different plants, possibly different genes in the different breeding lines.

Phenotypic variation for the relationship between the two different parameters (gall index and Rf-value) for individual plants were observed. This could be explained by two possibilities : (1) The inheritance of resistance to gall formation could be independent of the resistance to support nematode reproduction. (2) Experimental variability in measurements between individual plants. The method of determination of reproductive ability of the nematodes inherently lends itself to experimental variation between duplicate samples. This is clear from the large variability observed within the control populations. The greatest variability was observed in susceptible populations, with more certainty of resistance in the resistant populations. Less variability was observed between replicates evaluated with the gall index measurements, where the evaluation was based on relative values, rather than exact numbers.

In view of these findings, it is clear that the use of determination of exact numbers for reproductive ability for individual plants holds a greater risk of experimental error in determination of plant phenotype than the relative gall index measurements. It is therefore recommended to use the gall index values in subsequent experiments, rather than the Rf-values.

3.4 CONCLUSIONS

In the first round of hybrid production, only RAPD analysis was used initially for establishment of authenticity of the hybrids. Subsequent RFLP analysis suggested that these plants were not true hybrids, although the RAPD analysis showed polymorphic bands in the progeny that were present only in the pollen parent. This discrepancy could be explained by a few possibilities. The RAPD analysis is very sensitive and could elucidate small variation in genetic material of individual plants. It is also possible that the Prima seed obtained were not pure, or that the cultivar was not true breeding due to genetic drift over a long period of time. This also placed a question over using exact values for individual plants obtained with the current method for phenotypic screening for nematode resistance. Less variation was observed with the relative gall index determinations than with the more specific calculation of the reproductive ability of the nematodes on individual plants. These issues should be taken into account when screening a second segregating population for DNA markers, following a different approach for construction of DNA pools for bulked segregant analysis. As problems did occur with reproducibility of RAPD fragments, and considering the apparent hypersensitivity for inter-plant variation, another method of screening should also be considered.

The re-evaluation of the purity of the germplasm seed of the cultivar Prima is strongly recommended.

Although over 1000 papers have been published on the applications of the RAPD technique since the first reports in 1990, and its numerous advantages, the disadvantages and pitfalls of the technique were highlighted in this study. In spite of all efforts to avoid variance in assay conditions, problems with reproducibility did occur. CHEN, *et al.* (1997) encountered problems with reproducibility with new batches of primers or enzymes, and detected differences between leaf and root DNA. The primers used in this study were from one batch and leaf DNA was used in all assays. The only factor that could account for variability in assay conditions, could be the use of different batches of enzyme and reaction buffer. This strengthens the postulation of hypersensitivity of the technique, and that it is essential to take extreme care in establishing reproducibility before application in marker technology. As a result the use of the RAPD technique was therefore not pursued further in this study.

The segregating population will be screened with RFLP as well as AFLP techniques for identification of markers linked to the nematode resistance trait. The gall index values obtained in the phenotypic screening will be used as basis for linkage analysis.

CHAPTER 4

IDENTIFICATION OF MARKERS LINKED TO NEMATODE RESISTANCE

4.1 INTRODUCTION

In the absence of near-isogenic lines for the nematode resistance trait, a combined selective mapping strategy of bulked segregant analysis and selective genotyping, reduces the time and cost involved in marker identification. By means of these approaches, polymorphic markers are first identified between the parents, followed by evaluation of the polymorphisms across two contrasting DNA pools (BSA). Markers are then tested for cosegregation in two groups of lines consisting of the most resistant and most susceptible lines within the mapping population. Only markers cosegregating within groups of lines are mapped across the entire population.

The most common method for detecting and placing molecular markers on a linkage map is by RFLP analysis. A relatively dense linkage map is available for soybean, based mainly on RFLP markers which can be used as anchor markers to locate newly developed markers on the existing linkage map. It is thus essential to do at least a limited number of RFLP analyses on a new mapping population to act as a scaffold map in linking new markers. The technique is however, laborious and time consuming, with relatively low frequency of polymorphism due to the low genetic diversity present in cultivated soybean. AFLP is a PCR-based technique capable of detecting more than 50 loci in a single reaction, making it very valuable for detecting markers in a relatively short time. The use of this technique in combination with BSA and selective genotyping, can be a powerful tool in generating molecular markers for specific traits in segregating populations with minimal cost and time requirements.

4.2 MATERIALS AND METHODS

4.2.1 RFLP analysis

DNA from Gazelle and Prima was screened for polymorphisms with 53 soybean cDNA RFLP probes. DNA was digested with 1 of 4 enzymes (Table 4.1) and analysed with the Southern hybridization technique (Chapter 3.2.1.8). DNA from 60 F_2 individuals from the GP20-2 population was analysed for segregation with probes identified as polymorphic between the parents.

Table 4.1 Soybean probes screened in this study

PROBE	LINKAGE GROUP	ENZYME	INSERT SIZE	PROBE	LINKAGE GROUP	ENZYME	INSERT SIZE
A006	B	EcoRV	2100	A597	E	EcoRV	1500
A036	H	EcoRI	1100	A664	I,V	HindIII	1700
A063	C	EcoRI	700	A685	P	EcoRI	2000
A081	O	EcoRI	1800	A725	W	TaqI	2100
A096	A	HindIII	1400	A747	D	EcoRI	2100
A102	I	TaqI	1600	A808	N	EcoRV	800
A112	G	TaqI	1600	A810	H	HindIII	1700
A122	U	TaqI	3000	A878	O	TaqI	900
A199	K,J	HindIII	1300	A946	C,M	EcoRI	2100
A204	J	TaqI	1300	B122	J	EcoRI	1000
A235	G,D	TaqI	1000	B142	P,D	EcoRV	2500
A242	P,E	TaqI	1200	B153	B	TaqI	700
A257	D	EcoRV	1200	B202	F	EcoRI	1600
A264	L	EcoRI	1800	B212	F	TaqI	2000
A315	K	EcoRI	1400	B216	G	EcoRV	1400
A351	M	TaqI	1400	K007	H	HindIII	1600
A374	E	HindIII	1300	K019	D	EcoRI/EcoRV	2500
A426	C,C,G	TaqI	800	K227	M	EcoRI	600
A461	L	TaqI	700	K250	X	HindIII	2100
A486	A	EcoRI	1900	K258	R	EcoRI/HindIII	700
A505	A	EcoRI	700	K365	C	HindIII	900
A515	I	EcoRI	1800	K387	K	HindIII	1200
A519	B,D,C	TaqI	1600	K411	D	EcoRV	1600
A566	F	TaqI	1100	K418	N,K	TaqI	900
A567	S	HindIII	1100	K494	N,K	EcoRI	1700
A588	B	EcoRI	800	K636	A	EcoRI	1400
				K647	Q	EcoRV	1400

4.2.2 *AFLP analysis*

4.2.2.1 *Materials*

Screening of AFLP selective primers for polymorphisms with putative linkage to nematode resistance was conducted with DNA samples from the two parents, Prima and Gazelle, and two bulk samples (Table 4.2). The bulk samples were constructed by mixing equal amounts (m/m) of DNA of 5 F_2 plants each corresponding on two factors: (a) Plants from each of the two homozygous genotypes (AA and BB), as determined with RFLP analysis with probe B212. (b) The gall index values of these plants, with the resistant plants having a value of 1 and the susceptible plants a value of 9.

Table 4.2 Composition of resistant and susceptible bulks of F_2 plants

RESISTANT BULK			SUSCEPTIBLE BULK		
PLANT	GI	GENOTYPE	PLANT	GI	GENOTYPE
GP20-2-12	1	AA	GP20-2-2	9	BB
GP20-2-15	1	AA	GP20-2-3	9	BB
GP20-2-19	1	AA	GP20-2-8	9	BB
GP20-2-22	1	AA	GP20-2-9	9	BB
GP20-2-25	1	AA	GP20-2-13	9	BB

AFLP analysis was conducted with the AFLP™ Analysis System I and AFLP Starter Primer Kit supplied by GibcoBRL - LifeTechnologies, Glasgow, United Kingdom. AFLP™ is a trademark of Keygene Inc. Wageningen, The Netherlands. Acrylamide was obtained from Bio-Rad, Hercules, CA, USA, bis-acrylamide from Sigma, St. Louis, Missouri, USA, and Urea from GibcoBRL. DNA fragments were visualised with a silver staining kit from Promega, Madison, WI. Taq polymerase was supplied by Promega. Primer sequences were as developed by Keygene, Inc., and are listed in Table 4.3.

Table 4.3 Primer sequences used for AFLP analysis.

NAME	TYPE	SEQUENCE (5'-3')
E-A	Primer+1	AGACTGCGTACCAATTCA
M-C	Primer+1	GACGATGAGTCCTGAGTAAC
E-AAC	Primer+3	GACTGCGTACCAATTCAAC
E-AAG	Primer+3	GACTGCGTACCAATTCAAG
E-ACA	Primer+3	GACTGCGTACCAATTCACA
E-ACT	Primer+3	GACTGCGTACCAATTCACT
E-ACC	Primer+3	GACTGCGTACCAATTCACC
E-ACG	Primer+3	GACTGCGTACCAATTCACG
E-AGC	Primer+3	GACTGCGTACCAATTCAGC
E-AGG	Primer+3	GACTGCGTACCAATTCAGG
M-CAA	Primer+3	GATGAGTCCTGAGTAACAA
M-CAC	Primer+3	GATGAGTCCTGAGTAACAC
M-CAG	Primer+3	GATGAGTCCTGAGTAACAG
M-CAT	Primer+3	GATGAGTCCTGAGTAACAT
M-CTA	Primer+3	GATGAGTCCTGAGTAACTA
M-CTC	Primer+3	GATGAGTCCTGAGTAACTC
M-CTG	Primer+3	GATGAGTCCTGAGTAACTG
M-CTT	Primer+3	GATGAGTCCTGAGTAACTT

4.2.2.2 Methods

A slightly modified protocol was followed according to the manufacturers instructions and as developed by ZABEAU and VOS (1993). Approximately 2 µg of soybean genomic DNA was double-digested with the two restriction enzymes, *EcoRI* and *MseI*, for approximately 4-5 h and ligated to *EcoRI* and *MseI* adapters overnight at 37°C.

Pre-selective PCR was carried out with primers+1 (Table 4.3) in a 50 µL volume containing 5 µL of the ligated DNA, pre-amp primer mix (GibcoBRL kit), Taq polymerase buffer, 2 mM $MgCl_2$ and Taq DNA polymerase (Promega). Samples were overlaid with mineral oil and amplified in a Hybaid Thermal Cycler

for 30 cycles of 30 sec at 94°C, 1 min at 56°C and 1 min at 72°C. Quality and quantity of pre-amplification products were determined with electrophoresis in a 1.5% (m/v) agarose gel.

For selective PCR the pre-amplification products were diluted 1:10. PCR was conducted in a 20 µL reaction mixture containing 5 µL pre-amplification product, 0.25 ng μL^{-1} *EcoRI*+3 primer (Table 4.3), 1.5 ng μL^{-1} *MseI* +3 primer, 2 mM $MgCl_2$, Taq polymerase buffer and 0.02 U μL^{-1} Taq DNA polymerase (Promega). After overlaying the samples with mineral oil, the samples were amplified for one cycle at 94°C for 30 sec, 65°C for 30 sec and 72°C for 1 min, after which the annealing temperature was lowered 1°C for each of 9 cycles, followed by 24 cycles of 94°C for 30 sec, 56°C for 30 sec and 72°C for 1 min. After amplification, reactions were stopped with an equal volume of loading buffer (95% (v/v) formamide, 20 mM EDTA, 0.025% (m/v) bromophenol blue, 0.025% (m/v) xylene cyanol) and denatured at 94°C for 3 min, followed immediately by chilling on ice. A 5% (m/v) denaturing polyacrylamide (19 acrylamide : 1 N,N'-methylene-bis-acrylamide ratio) gel was prepared with 7 M urea and 1 x TBE buffer (89 mM Tris-borate, 2.5 mM EDTA, pH 8.3). Two glass plates were prepared before casting the gel. One plate was treated with approximately 2 mL Wynn's C-Thru (WYNN Oil, Bramley, South Africa) and the other with bind silane (950 µL absolute ethanol, 5 µL acetic acid, 3 µL bind silane (Promega)). The gel was pre-run at constant 80 W for 30 min. PCR products (5 µL) were separated on the prepared gel at 80 W constant power for approximately 2 h using a standard DNA sequencing unit (C.B.S. Scientific Company, California, USA).

The separated amplified DNA fragments were visualized with a silver staining kit from Promega according to the manufacturers instructions. The gel was left upright overnight to air dry and photographed by exposing photographic paper (Kodak Polymax II RC) directly under the gel to about 20 sec of dim light. This produced a negative image of exactly the same size as the gel.

4.2.3 Statistical data analysis

4.2.3.1 DNA marker analysis

Two methods of analysis were used to identify markers associated with resistance to *M. javanica*. Data were analysed with a general linear model of the STATGRAPHICS Plus computer program (Manugistics, Rockville, Maryland, USA, 1998) using genetic marker data as the independent and gall index or $\log_{10}$ transformed reproduction factor values as the dependent variable. The association between the DNA

marker and the trait was considered significant if the probability was <0.05. The coefficient of determination (R^2) was used as a measure of the magnitude of association.

4.2.3.2 Mapping of markers

Interval mapping with MAPMAKER-EXP (LINCOLN, *et al.*, 1992) was used to link the markers to known RFLP markers on the existing soybean genomic map. Linkage data were used to assign markers to linkage groups if the LOD was ≥3.0 and the distance was ≤37 cM with Kosambi mapping function. The scan command of MAPMAKER-QTL (LINCOLN, *et al.*, 1992) was used to identify the position of putative QTL on linkage group F.

4.3 RESULTS

4.3.1 RFLP analysis

4.3.1.1 Screening with RFLP

The two parent plants, Gazelle and Prima, were screened for polymorphisms with 53 soybean cDNA probes (Table 4.1), using one of 4 restriction enzymes as recommended by the supplier for each probe. Sixteen of these probes (30%) produced polymorphic fragments between the two parents, with one of the probes, B142 recognising 2 loci, named B142-1 and B142-2. These 16 probes were screened for segregation in the GP20-2 population of 60 plants. Marker A808 was scored as a dominant marker. The patterns produced were compared with images downloaded from the soybean database (SOYBASE, 1995). Eight of these matched the hybridization images and were used in subsequent analyses as anchor markers on the soybean genomic map (Table 4.4). Three of the markers (K007, A685, K494) differed significantly from a 1:2:1 segregation pattern by χ^2 analysis ($P \leq 0.05$).

4.3.1.2 Mapping of RFLP markers

The complete RFLP analysis data for 60 F_2 plants with probe B212 is listed in Appendix C. Linkage analysis of the F_2 plants with probe B212 revealed that this marker was linked to gall index resistance with a magnitude of 62% (R^2= 0.615, P=0.000). This was in accordance with results obtained by TAMULONIS

Table 4.4 RFLP analysis of the segregating population (GP20-2)

PROBE	LINKAGE GROUP	PROBE	LINKAGE GROUP
B212	F	A878	O*
A063	C	A685	P*
K007	H*	K411	D
A006	B*	K258	R
A808	N*	K636	A*
A946	C/M	K647	Q
B142-1	P	K494	D
B142-2	D	K387	K
B122	J*		

*Linkage group not confirmed.

et al. (1997b). None of the other markers were linked to the resistance trait ($P \leq 0.05$). The segregation data was used for linkage analysis with MAPMAKER-EXP to construct a scaffold map for placement of AFLP markers. Nine of the markers could be assigned to known linkage groups (Table 4.4), with 8 of the RFLP markers unlinked to the anchor map (Figure 4.1).

4.3.2 AFLP analysis

The resistant and susceptible DNA bulks were constructed based on data obtained from both the gall index determinations as well as the RFLP patterns obtained with pB212 hybridised with *Taq I* digested genomic DNA from the individual F_2 plants (population GP20-2) (TAMULONIS *et al.*, 1997b). Theoretically, these pools would be targeting the chromosome interval around the resistance trait, and could be used to identify more closely linked AFLP markers to create a dense molecular map in this region.

4.3.2.1 Linkage of markers to resistance to M. javanica

Analysis of the two parental plants, Gazelle and Prima, with 64 primer pair combinations identified a total of 3814 loci, of which 377 (9.9%) were polymorphic (Appendix D). The average number of fragments

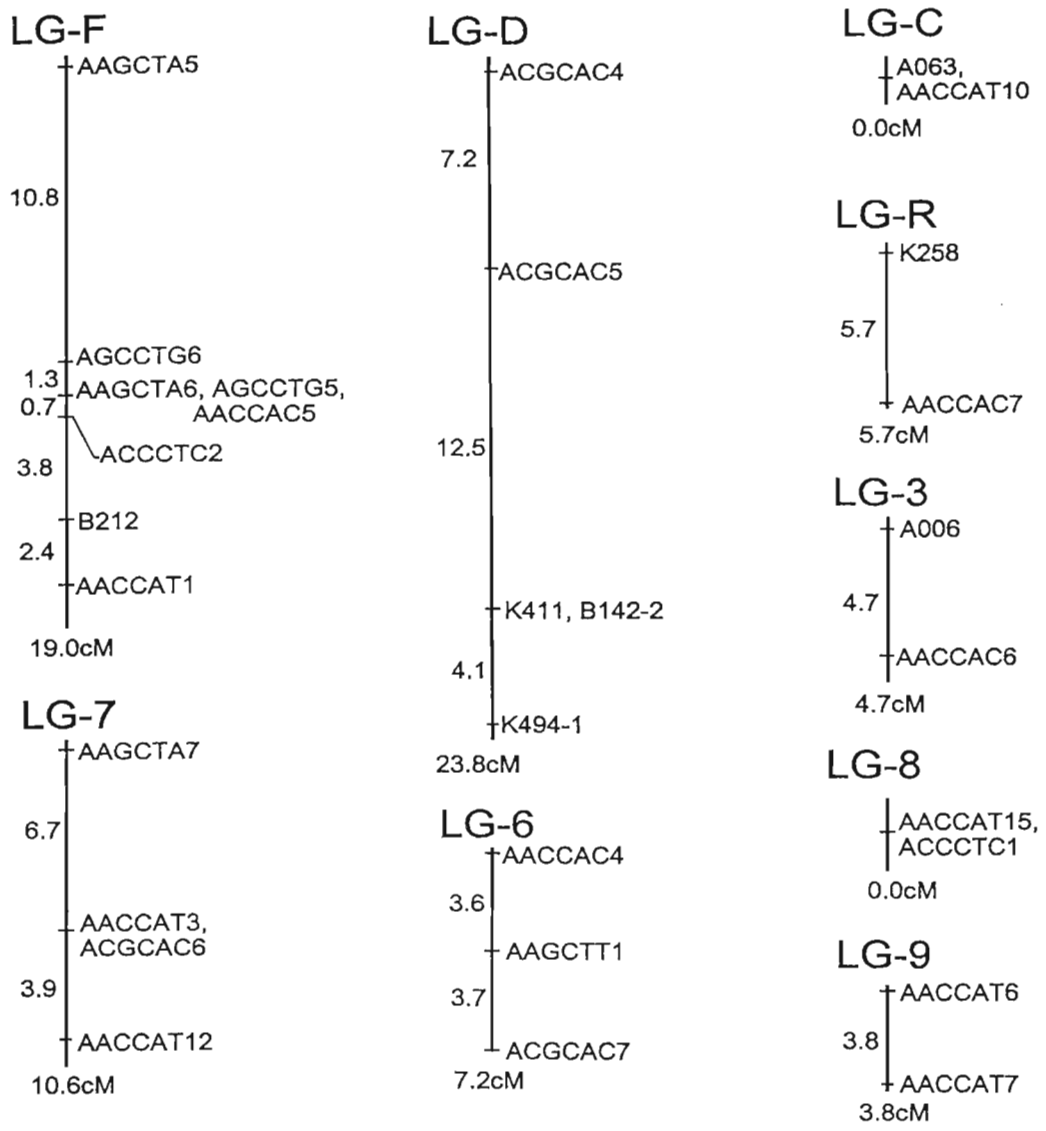


Figure 4.1 : A genetic map of soybean. Nine linkage groups are shown with either their homologous name from the public soybean map (C,D,F,R) or an arbitrarily assigned number. Distances are given in cM.

detected by single primer combinations ranged between 20 and 98 with an average of 59.6±18.2SD fragments. All primer combinations (100%) generated polymorphisms between the two plants with an average of 5.8±3.0SD polymorphisms per reaction. Sizes of fragments were compared to a 100 bp ladder and ranged between approximately 100 and 800 bp, as determined with denaturing PAGE (Chapter 4.2.2.2). A standardised notation of naming the *EcoRI*+3 selective nucleotides first and the *MseI*+3 selective nucleotides secondly will be used throughout further discussion.

Table 4.5 AFLP analysis data

	PARENTS	RESISTANT AND SUSCEPTIBLE BULKS	10 INDIVIDUALS OF BULKS	40 INDIVIDUALS
Primers tested	64	64	22	9
Total fragments	3814	3814	1604	594
Polymorphisms	377	377	176	63
Informative primers	64	22	9	5
Informative bands	-	47	16	7

Complete data of the AFLP analysis of the two bulks and 10 individual plants is presented in Appendix E. Twenty-two primer combinations amplified informative polymorphisms between the two bulk samples and the parent DNA (Table 4.5). Of the 377 polymorphisms amplified between the two parents, 47 were informative in the bulk samples. Fifteen markers were linked in coupling phase (originating from Gazelle) and 6 in repulsion phase (originating from Prima) (Appendix F). An additional 26 markers displayed quantitative differences in amplification in the 2 bulks, and were included in the next phase of screening. Selective genotyping of the 10 individuals from the most extreme resistant and susceptible plants revealed 16 polymorphic fragments cosegregating with the resistance trait (Table 4.6) with $P<0.05$, amplified with 9 primer combinations. These primer combinations were mapped across a larger segregating population of 40 additional plants. Clear segregation patterns were produced, of which the majority could be classified as dominant markers. Two sets of markers (E-AAG/M-CTA5 and E-AAG/M-CTA6; E-ACG/M-CAC5 and E-ACG/M-CAC6) comigrated and were also scored as co-dominant. Seven of the markers scored displayed a distorted segregation pattern (Appendix F) and differed significantly from a 3:1 segregation pattern by χ^2 analysis ($P \geq 0.05$).

Table 4.6 AFLP markers linked to nematode resistance

FRAGMENT		LINKAGE[1]	10 PLANTS		50 PLANTS	
	SIZE (bp)		R^2(%)	P	R^2(%)	P
E-AAC/M-CAC5	177	R	100	0.000**	30	0.000**
E-AAC/M-CAT1	246	R	100	0.000**	42	0.000**
E-AAG/M-CTA5	371	C	67	0.004**	23	0.000**
E-AAG/M-CTA6	369	R	100	0.000**	30	0.000**
E-ACC/M-CTC2	244	C	67	0.004**	25	0.000**
E-AGC/M-CTG5	84	R	100	0.000**	30	0.000**
E-AGC/M-CTG6	82	C	43	0.040*	21	0.000**

[1]: R-Repulsion phase; C-Coupling phase

Seven of the AFLP fragments were associated closely (P=0.000) with the gall index of the 50 individual F_2 plants. Three of the markers (E-AAG/M-CTA5, E-ACC/M-CTC2 and E-AGC/M-CTG6) were linked in coupling phase with the resistance gene, and 4 markers were linked in repulsion phase. Marker E-ACC/M-CTC2 accounted for 25% of gall index variation, with the 3 markers linked in coupling together accounting for 69%. The R^2 values for the markers linked in repulsion phase were larger, with E-AAC/M-CAT1 accounting for 42% of the variation. Together the 4 markers in repulsion phase accounted for 132% of variation in gall index.

4.3.2.2 Linkage map construction

Data from a total of 64 markers for 50 individual F_2 progeny (Appendix G) were used to construct a genetic map with MAPMAKER-EXP. The markers were grouped in 9 groups with 34 markers unlinked to any group (Figure 4.1). The 7 AFLP markers, identified by analysis of variance as closely linked to resistance to gall formation were grouped together with marker B212, which could be anchored on LG-F. The total distance covered by the markers on LG-F was 19.0 cM. Marker B212 was flanked by AFLP markers on both sides, with marker E-AAC/M-CAT1 (246 bp) mapping the closest to B212 at 2.4 cM with a LOD score of 8.8. The other 6 markers mapped to the opposite side of B212 with marker E-ACC/M-CTC2 (244 bp) at a distance of 3.8 cM. Three of the markers were completely linked to each other, i.e. distance 0 cM, because their marker types were identical. Marker E-AAG/M-CTA5 (371 bp) mapped the furthest away

from B212 at 16.6 cM. As only one known RFLP marker was polymorphic which could be linked to LG-F, the orientation of this map on the classical map could not be established.

Two RFLP markers were linked closely on LG-D, K411 and B142-2 (0 cM), which is in accordance with the published soybean genomic map (SOYBASE, 1995). Marker K494a (LG-T) and K494b (LG-K) did not produce the same banding pattern in this population as in the published samples, and was named K494-1. This marker was assigned to LG-D at a distance of 4.1 cM from K411 and B142-2. Two additional AFLP markers, E-ACG/M-CAC4 and E-ACG/M-CAC5 were mapped to the same linkage group. Two more AFLP markers could be assigned to known linkage groups (Figure 4.1), E-AAC/M-CAT10 to LG-C and E-AAC/M-CAC7 to LG-R. Five linkage groups did not include known anchor markers and were assigned arbitrarily chosen numbers. One RFLP marker, A006, and 12 AFLP markers were assigned to these linkage groups (Figure 4.1). The map as constructed with this limited amount of data spanned a total of 74.8 cM, including 30 markers on the 9 linkage groups.

As could be expected from the BSA methodology followed, the highest density of markers was found on LG-F, where the QTL with main effect on the resistance trait was localised (Figure 4.2). Analysis of LG-F for the position of the QTL for gall index resistance with MAPMAKER-QTL detected only one peak (LOD>10) between markers B212 and E-AAC/M-CAT1. No additional QTL could be identified on other linkage groups.

4.4 DISCUSSION

The RFLP marker B212 was closely linked to the resistance trait and accounted for 62% of variation in gall index. This association was even closer as that found by TAMULONIS, *et al.* (1997b) in their mapping population. This suggests that it could probably be the same gene or at least a gene located in the same region on LG-F. The other marker found by TAMULONIS, *et al.* (1997b), A725-2, situated on LG-D1, and which accounted for only 13% of gall variation, was not polymorphic in Gazelle and Prima. None of the other markers which mapped to LG-D were linked to the resistance trait. According to the various maps for LG-D/LG-D1 published on the internet (http://probe.nalusda.gov:8000/plant/ aboutsoybase.html), the distance between these markers (K411, B142-2 as anchor markers) could be between 150 cM and 450 cM, which would explain the lack of close linkage, especially if the QTL is situated on the far side of A725-2. As the main QTL with large effect (62%) was identified on LG-F, it was decided to concentrate on this

region for the development of a marker system which would be easily and economically applicable in marker assisted selection in a breeding program. As it was found that this marker was situated in a cluster of several other disease resistance loci (TAMULONIS, *et al.*, 1997b), a marker in this region could possibly also find application in detection of these loci.

A three step design was followed for identification of linkage of AFLP markers to the resistance trait, in which a process of elimination was used to minimize the samples to be analysed. The pooling of samples from the 2 extremes provided a crude simulation of NIL differing theoretically only at the region of the chromosome where the resistance trait is located. Two criteria were used in preparation of the DNA pools,

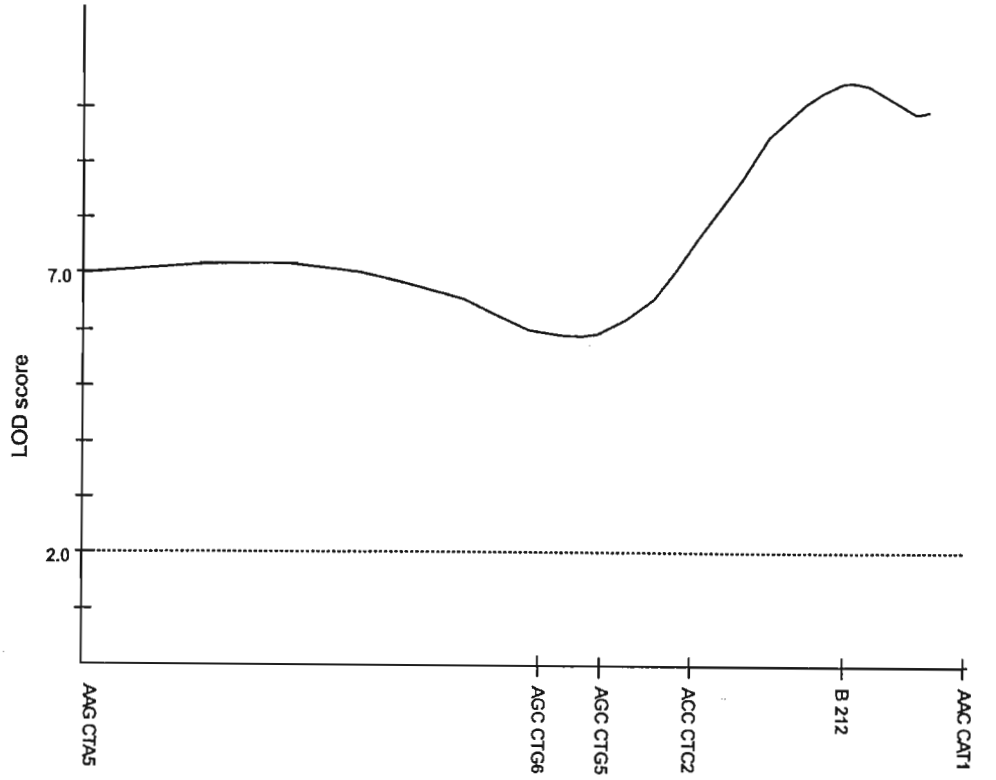


Figure 4.2 QTL likelihood plot indicating a peak between markers B212 and AAC CAT1. The horizontal dotted line at LOD=2 represents the minimum LOD required for significance. The horizontal scale of the plot is not exact.

with both phenotype and genotype (B212) taken into account. A high level of $P<0.05$ was chosen as significance level for linkage of markers in the ten individual plants from the 2 bulks. This high level was chosen to exclude the possibility of missing important loci. It was also acknowledged that this was a highly biased population, and there was a risk of missing possible distantly linked markers. The markers linked in the analysis of the 10 plants from the extremes were confirmed in the third stage with a larger set of progeny from the mapping population. The disadvantage of this strategy is the high possibility of missing QTLs located on other chromosomes. As the main QTL with large effect (62%) was identified on LG-F, it was decided to concentrate on this region for the development of a marker system which would be easily and economically applicable in marker assisted selection in a breeding program. All markers associated with the variation in gall index in this study were linked to LG-F, with no markers on LG-D linked significantly ($P<0.05$) to the resistance trait.

The identification of the AFLP markers in this study was biased in 2 ways. The combined bulked segregant analysis and selective genotyping enriched the fraction tested for the area on the chromosome around the QTL with main effect on the resistance trait, as described above. This was clearly illustrated in the density of the markers assigned to LG-F. It could also be biased towards relatively AT-rich regions because the 2 enzymes used in the digestion of the genomic DNA, namely *EcoRI* (G↓AATTC) and *MseI* (T↓TAA) both target these regions.

Three of the AFLP markers were found to be linked in coupling phase and 4 in repulsion phase to the resistance to gall index in this population. Marker E-AAC/M-CAT1 accounted for the greatest variation in gall number (42%), and was linked in repulsion phase to the resistance trait. Of the 3 markers linked in coupling phase to gall index, marker E-ACC/M-CTC2 accounted for the largest variation in gall index (25%). Previous studies have shown the utility and application of both these types of markers in different selection populations in breeding programs. Homozygous resistant plants could be distinguished from heterozygous resistant plants by detecting the absence of the repulsion phase marker (BAI, *et al.* 1995). HALEY, *et al.* (1994) also found that the repulsion phase marker could provide greater selection efficiency than coupling phase markers. Their repulsion phase marker provided a greater proportion of homozygous resistant selections, and a lower proportion of both segregating and homozygous susceptible selections. Selection based on a repulsion phase marker could therefore be identical to selection based on a codominant locus such as most RFLP loci (HALEY, *et al.*, 1994). The utility of repulsion phase markers was greatest in MAS of homozygous resistant individuals in F_2 or later segregating generations

(JOHNSON, *et al.*, 1995). The selection approach must be determined by the nature of the population. Selection with a repulsion phase marker in a BC_nF_1 (Rr:rr) population of traditional backcross breeding will eliminate all progeny and will not be viable. The combined use of both markers, depending on the nature of the selection population, would therefore improve the efficiency of MAS.

The construction of a genetic map without chromosome-specific markers results at best in a map comprised of a number of linkage groups, not always corresponding to the chromosome number of the species under investigation. No specific chromosome can be attributed to any of these groups and neither can chromosome orientation be determined. Map development in soybean followed a long history covered extensively in chapter 1. The current soybean map includes 24 linkage groups with 8 of these having only 4 markers or less. This does not correspond to the chromosome number of 40, and illustrated the difficulty in determining accurate genetic maps, even in a crop like soybean, which is one of the oldest cultivated crops. The initial problems with mapping resulted from the low genetic variability, but with the currently evolving techniques like AFLP and DNA microarray systems, the map is bound to expand rapidly. KEIM, *et al.* (1997) differentiated 28 linkage groups in a study on AFLP marker data. The problem with correlation of marker data with the ISU-USDA map, is that the map is based on an extremely wide cross and there is a difference in number of polymorphic fragments, as well as identification of different alleles in other mapping populations with the same probes. The map construction is further complicated by duplication of chromosome segments which were retained on different chromosomes (SHOEMAKER, 1994). As a result, maps for specific linkage groups published in Soybase on the internet (http:\\129.186.26.94) sometimes seem to contain conflicting results. CREGAN *et al.* (1999) mapped a large number of SSR markers and integrated it with the existing maps, resulting in a consensus map comprising of 20 linkage groups, corresponding to the 20 pairs of soybean chromosomes. This is the most extensive integration of the classical, RFLP and SSR markers into one linkage map and possibly the first with real practical value in MAS.

The QTL for gall index resistance mapped between markers B212 and E-AAC/M-CAT1 (Figure 4.2). According to MAPMAKER-EXP analysis these two markers are only 2.4 cM apart, which means that the combined use of these two markers for MAS could be very effective. Another AFLP marker, E-ACC/M-CTC2, maps near B212 and the QTL, 3.8 cM from marker B212.

4.5 CONCLUSIONS

Eight markers linked significantly to resistance to *M. javanica* were successfully identified in the segregating population. An RFLP marker, B212, accounted for 62% of the variation in gall index values. An additional 7 AFLP markers were linked to the resistance trait, using a combination of BSA and selective genotyping methods. Three of the markers were linked in coupling phase to variation in gall index, with E-ACC/M-CTC2 accounting for 25% of the variation. Four AFLP markers were linked to susceptibility (repulsion phase), with E-AAC/M-CAT1 explaining 42% of variation in gall number. All markers linked to resistance were located on LG-F. The major QTL for resistance to gall formation was bracketed by two markers, namely B212 and E-AAC/M-CAT1. The combined use of these markers, of which one is codominant and the other linked in repulsion phase to the QTL, could be very effective in MAS. The selective process could be simplified further and made more economically viable with the conversion of these markers to sequence specific PCR-based markers (SCARs), which would be discussed in Chapter 5.

CHAPTER 5

CONVERSION OF PUTATIVE MARKERS TO SEQUENCE CHARACTERIZED AMPLIFIED REGIONS (SCARS)

5.1 INTRODUCTION

Breeding for disease resistance using marker-assisted selection (MAS) requires that (i) the resistance gene(s) be tagged by closely linked molecular markers, (ii) the linkage be stable across generations and populations and (iii) an efficient way of screening large populations for molecular markers be available (HITTALMANI, *et al.*, 1995). The first 2 prerequisites were attended to in the previous chapter (Chapter 4). On account of the third prerequisite, both RFLP and AFLP techniques are laborious and expensive, and unsuitable for accommodating large numbers of progeny testing early in a breeding program. The breeding objective is to maximize the response per unit cost, and the marker-assisted selection should be superior to phenotypic selection. Genotyping the marker loci should cost less than obtaining phenotypic measurements (XIE and XU, 1998). Converting a molecular marker to a PCR-based marker, could lower the cost and time of MAS substantially. A large number of individuals can also be screened in a relatively short time. SCARs are PCR-based markers that represent single, genetically defined loci, identified by amplification of genomic DNA with pairs of specific oligonucleotide primers (PARAN and MICHELMORE, 1993). The possibility of converting the molecular markers identified in Chapter 4 to SCAR markers is explored in this chapter.

5.2 MATERIALS AND METHODS

The development of SCARs consists of various steps, depending on the technique which identified the original polymorphism. SCARs are developed to create a faster, reliable test for the presence or absence of a specific polymorphism. The steps in development include (Figure 5.1): (I) isolation of a polymorphic fragment (from AFLP or RAPD products), (ii) cloning of the fragment and verification of the insert, (iii) sequencing of the insert or in the case of using the RFLP technique, sequencing of the RFLP probe detecting a polymorphism in the genomic DNA, (iv) design of oligonucleotide primers, and (v) verification by PCR of genomic DNA. (vi) In a case where no polymorphism is amplified, the PCR products can be digested with a range of restriction enzymes to screen for polymorphisms.

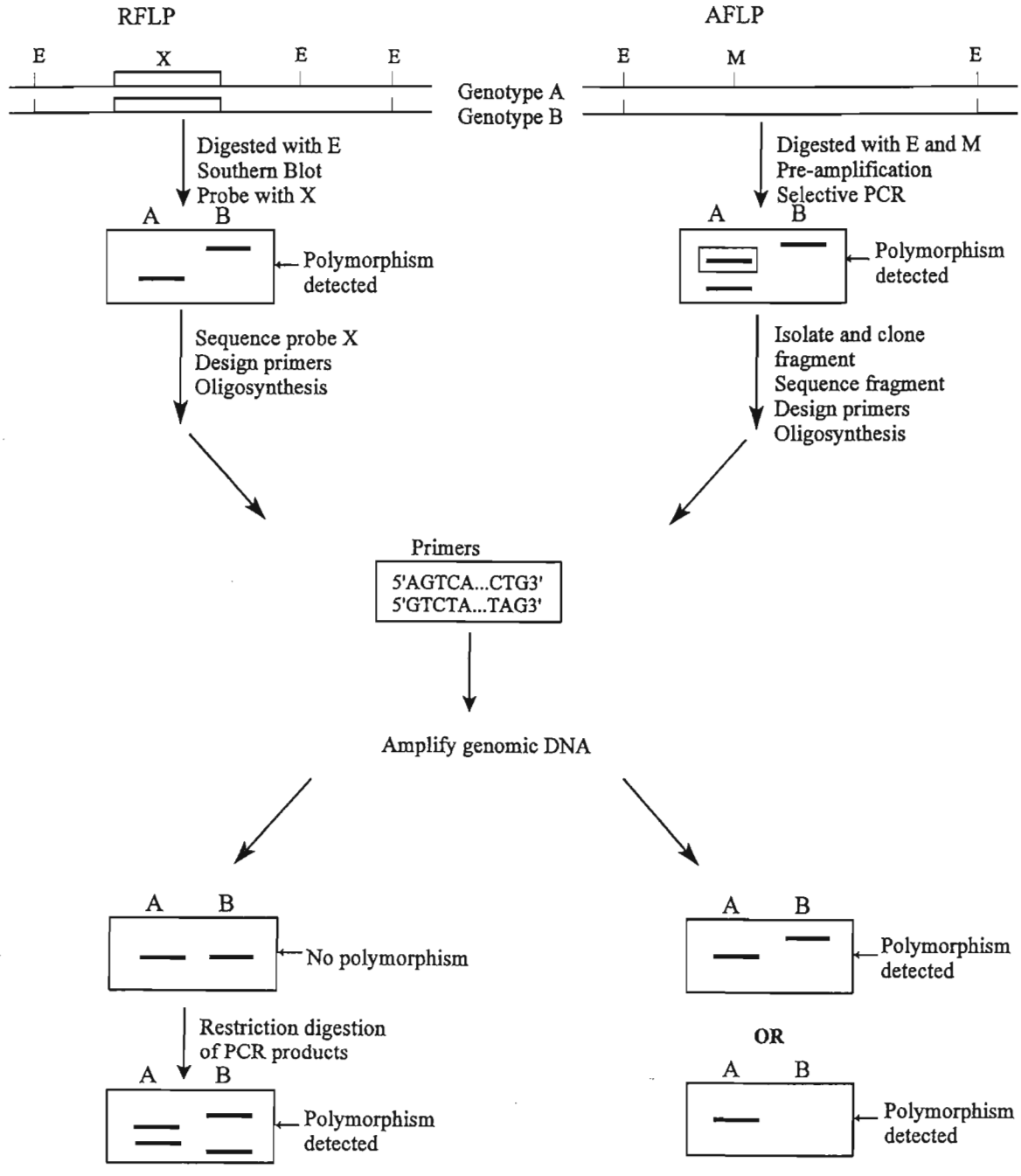


Figure 5.1 Protocol for the development of a SCAR.

5.2.1 *Materials*

The pGEM®T-Easy vector was obtained from Promega Corporation, Madison, WI, USA. Ingredients of LB-medium - Bacto-triptone, yeast extract and Bacto-agar were supplied by Difco Laboratories, Detroit, Michigan, USA. Ampicillin, IPTG (Isopropyl-ß-D-thiogalactopyranoside) and X-Gal (5-bromo-4-chloro-3-indolyl-ß-D-galactoside) were obtained from Roche Boehringer Mannheim, Randburg, South Africa. The pB212 probe was supplied by Biogenetic Services, Inc., Brookings, SD. Primers were synthesised by GibcoBRL - LifeTechnologies, Glasgow, United Kingdom. The Wizard® PCR preps DNA purification system and Wizard® Plus SV miniprep DNA purification system were obtained from Promega Corporation, Madison, WI, USA.

5.2.2 *Isolation and cloning of AFLP fragments*

The silver stained AFLP polyacrylamide gels (Chapter 4.2.2.2) were air-dried. The putative marker fragments were excised from the gel with a scalpel blade and the piece of gel rehydrated in 10 µL distilled water for easy removal from the glass plate (CHO, *et al.*, 1996). The piece of gel was transferred to a 0.5 mL Eppendorf tube and overlaid with 10 µL of distilled water, and stored at -20°C. The fragment was reamplified directly from the piece of gel without any purification, using the same conditions as for the original AFLP reaction. An aliquot of 1 µL was electrophoresed on a sequencing gel for determination of purity, with a possible second and third round of isolation and amplification where necessary.

The PCR products were purified using the Wizard® PCR preps DNA purification system (Promega Corporation) and cloned into the pGEM®T-Easy vector following the manufacturers recommendations. Ligation reactions were set up with a positive control containing control insert DNA and a background control containing no insert. Ligation were left at 4°C overnight to obtain the maximum number of transformants. The ligated plasmids were transformed into high efficiency competent JM109 bacterial cells, supplied by Promega, and plated onto selective LB-plates (10 g L^{-1} Bacto-triptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl, 15 g L^{-1} Bacto-agar containing 100 µg mL^{-1} ampicillin, 0.5 mM IPTG, 80 µg mL^{-1} X-Gal). The plates were incubated overnight at 37°C. White colonies were selected and cultured overnight in 100 µL LB-medium containing 100 µg mL^{-1} ampicillin. The cultured cells (4 µL) were tested for the presence of the desired insert DNA by using the bacterial cells directly as template in a PCR reaction. The PCR profile and reaction components were the same as for the original AFLP reaction. The size of the

PCR products were verified by electrophoresis of 1 µL samples on a denaturing polyacrylamide gel as described for the AFLP reactions (Chapter 4.2.2.2).

5.2.3 Sequencing of fragments and design of primers

The pB212 probe, cloned in a pBS vector, was used for the development of a SCAR (Figure 5.1). The probe was kindly sequenced by the Institute for Plant Biotechnology, University of Stellenbosch, South Africa, with the use of T3 (5'-ATT AAC CCT CAC TAA AG-3') and T7 (5'-AAT ACG ACT CAC TAT AG-3') promoter primers. Two 17-mer primers were designed from these data with a GC content of approximately 50%. Primers were synthesised by GibcoBRL.

Plasmids containing the cloned AFLP fragments were cultured overnight in 10 mL LB-medium containing 100 µg mL^{-1} ampicillin and purified using the Wizard® Plus SV miniprep DNA purification system (Promega Corporation) according to the manufacturers instructions. The relevant AFLP fragments, cloned into the pGEM®T-Easy vector, were sequenced by the Institute for Plant Biotechnology, University of Stellenbosch, South Africa, with the use of M13 sequencing primers. All sequences contained the *EcoRI* adapter at one end and the *MseI* adapter at the other end. Two 17-mer oligonucleotides (Table 5.1) internal to the 5' and 3' ends of the fragment were designed using the NetPrimer program of PREMIER Biosoft International (www.PremierBiosoft.com). Primers were synthesised by GibcoBRL.

Table 5.1 SCAR primers designed for AFLP fragments

FRAGMENT			FORWARD PRIMER	REVERSE PRIMER
NAME	SOURCE	SIZE(BP)	5'-3'	5'-3'
SOJA-1	E-AAC/M-CAC5	177	TGAGATACTTAGAGATG	CAAAAAGTTTCACAAGA
SOJA-3	E-AAG/M-CTA5	371	ATAGCCAATAGAAAACA	ATGCCTATCTACTAACG
SOJA-4	E-AAG/M-CTA6	369	GTCTATGTACTAACCGA	GTTCGAATTGGCTTGTC
SOJA-6	E-ACC/M-CTC2	244	CATGGGCCATCCTAGAG	TTGTACCAAATCAGCTC
SOJA-7	E-AAC/M-CAT1	246	TTTGAGATCACTTGGCT	GATCCTAAATCACCTAA
SOJA-9	E-AGC/M-CTG5	83	GTAGGAGAGGAAAGACC	GCAAATGAAGGAAGGCA

5.2.4 *Verification by PCR analysis*

Amplification conditions for the designed primers were determined empirically, varying annealing temperature and $MgCl_2$ concentration. Approximately 500 ng genomic DNA was amplified with the SCAR primers and the products electrophoresed on 1.5% to 2% (m/v) agarose gels.

The PCR reaction with the designed B212 primers was optimized for $MgCl_2$ concentration (between 1.5 mM and 5 mM), DNA concentration (100 ng - 1 000 ng), Taq DNA polymerase (1-2 U) and annealing temperature (50°C, 55°C, 60°C, 62°C and 65°C). Reactions were carried out in 25 µL containing 250-500 ng genomic DNA, 3 mM $MgCl_2$, 0.2 µM each of primers B212F (5'-AGT CTT TGT CGC CGC AT-3') and B212R (5'-GCC TCA GGC ATT TGG TC-3'), 100 µM each of dATP, dCTP, dGTP and dTTP, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1% (v/v) Triton X-100, 1.5 U Taq DNA polymerase. Amplification was done in a Hybaid Thermal Cycler (Hybaid Limited, United Kingdom) with a denaturation step at 94°C for 5 min, followed by 40 cycles, each cycle consisting of 94°C for 1 min, 60°C for 1 min and 72°C for 1.5 min. A final elongation step of 5 min at 72°C was included in the program.

Primers were designed for 6 cloned AFLP fragments (Table 5.1). The amplification reactions were optimized for annealing temperature (40°C, 45°C, 50°C, 55°C and 60°C) and $MgCl_2$ (2 mM and 3 mM) concentration. Reactions were carried out in 25 µL containing approximately 500 ng genomic DNA, 2 or 3 mM $MgCl_2$, 0.4 µM each of the forward and reverse primers, 100 µM each of dATP, dCTP, dGTP and dTTP, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1% (v/v) Triton X-100, 1 U Taq DNA polymerase. Amplification was done in a Hybaid Thermal Cycler (Hybaid Limited, United Kingdom) with a denaturation step at 94°C for 5 min, followed by 40 cycles, each cycle consisting of 94°C for 1 min, x°C for 1 min and 72°C for 1.5 min. A final elongation step of 5 min at 72°C was included in the program.

5.2.5 *Restriction digestion of PCR products*

The B212 SCAR primers did not produce linked polymorphisms between the parent genomic DNA and 20 µL of the PCR products were subjected to digestion with a wide range of restriction enzymes - *AccI, AluI, BamHI, BglII, ClaI, DraI, EcoRI, EcoRV, EclXI, HaeIII, HindIII, KpnI, MspI, MvnI, NaeI, PstI, PvuII, RsaI, SacI, SalI, Sau3A, ScaI, SfuI, SmaI, SpeI, SspI, StuI, StyI, XbaI and XhoI* at 37°C; *BssHII, BstXI* and *BstEII* at 50°C; *TaqI* at 65°C. The restricted fragments were electrophoresed in 2% (m/v) agarose gels

at 80 V for 2.5 h or alternatively in 3% (m/v) Metaphor agarose gels (0.5 x TAE running buffer) at 80 V for 5 h.

5.2.6 Verification by Southern analysis

Southern analysis was conducted with the cloned AFLP fragments as probes on genomic DNA of the parents. DNA from individuals from the F_2 population was hybridised to fragments which were successfully converted to SCARs. Samples of DNA (10 µg) were digested with *EcoRI* and treated as in Chapter 3.2.3.2. The cloned AFLP fragments were labelled as described in Chapter 3.2.3.2(iii), using the appropriate AFLP primers and used for hybridisation with the genomic DNA blots (Chapter 3.2.3.2(iv)) at 40-45°C. Detection was done as described in Chapter 3.2.3.2(v). These analyses gave an indication of the copy number and heritability of the fragments.

5.3 DEVELOPMENT OF A SCAR FROM AN RFLP PROBE (B212)

5.3.1 Sequencing data

The probe insert of pB212 was sequenced from both ends and new primers designed near the ends to amplify the fragment bracketed between them (Figure 5.2). The PstI sites which were used for cloning of the cDNA probe, were sequenced accurately and the ends of the cDNA could be identified. Two primers of 17 nucleotides each were designed with a GC-content between 50 and 60%. The T_m values of the primers differed by only 2°C (64 vs. 66°C). No palindromic sequences were found within the primers, although some cross-dimers could be expected. The positions of the oligonucleotides are indicated with a double line in Figure 5.2. A total of 1781 bp of the probe fragment were sequenced.

5.3.2 PCR and restriction analysis results

A polymorphism (600 bp) was amplified between Gazelle and Prima at an annealing temperature of 60°C with a Mg^{2+} ion concentration of 3 mM (Figure 5.3). The enzyme concentration did not have a visible effect on the banding pattern. Two additional fragments were amplified in both parent plants. PCR analysis of 50 F_2 plants revealed that the $B212_{600}$ fragment was not linked to variation in gall index values (P=0.406).

```
                    PstI
                5'..TGCA ↓ GGGA  GCCMATTCAC  MAGGCTGGCC  MAAGTCTTTG  TCCGCATCGC   50
GCAAGTGGCG  CGTGCAGGCC  ACGCACGGTC  ACCACAATTA  TTGCTATGAC  GAGGATGAGG  110
AGCATYATGA  GCCCCATAAT  CATCTCCATG  TTCAACCCAA  AATCCACACT  CTCCTCCATA  170
CTCATCAATA  TAGGTGGGTG  TCACACTCTT  TCTAACTAAC  TACTAGAACC  CTCCATGTTT  230
TTCTTTCCAT  CATGAAGTTA  TTATTAACAA  AAAAAGAAGA  AGCTATACTT  TACGATGTCA  290
TAGCCATTTC  AAAACAATAA  ATGAATAACT  AATTTTTGTA  GGTGCATGAA  ATACTTTCAT  350
CATGTGGTTT  ATGATTAAGT  TGAAGTGGAG  GGAATGGTAT  GATGTTGAAA  TTGTGGTGGG  410
AAGAGTAAAG  TTTGNTGAAT  AGTTTTGATT  AAATTTTTTA  TGCTTATATT  ATAATATGCT  470
TTTTAGNTAG  TTTGATGATA  AAGACATTGA  GCAAATAAGT  TAAATTTGAG  TTTTTTTGTT  530
AGGGAAGGAG  AATAAATATA  ATGGACATAT  TNTNATCAAA  GNCAACTATC  TTCTTAAAAG  590
GAATAAGGAG  GGTTTCCAAA  AATTCTGCTC  TTCTGAATTG  NAAAGGNGAG  CTTTWAGAAG  650
ATACGCACAT  GTAATTATTG  GTTNAAAAAA  TNGSGNNTTA  AAATTATTAT  AAATWTTNAA  710
YGCCTMTTTN  TAGGTAAAAT  TTGAAGNTAG  GTNSCAATCA  TGGGCMATTN  AGNATWCCTT  770
TATTAAATTA  NTCCCGAAAA  WACAMCTWTT  TNYTAMCTTT  TWRAAAAAGG  GGGGAAACCC  830
TTTTTTNAAT  ACYYNMMANG  GGNAAAAYGG  NTWNTTTWTT  TNGCMWNKKN  AAGNCNNTTT  890
TTTAAANANM  YNCNGGGGNN  AAANAAANGG  GGGGNYYYN  ----------  TTTTTCCCCC  940
CACCCTTTAA  TTACTAATAA  CCCNTTTGGA  ANATTTTTTT  WAAAAKSTTT  TCCAAAKGAG  1000
NSRCTGKTAA  ACAYCAAAAA  CTMATGGGGC  CCCCACTAGG  NTGGNTNTAT  GGTTTTGAGT  1060
TGANACCAMC  ANCATTYCAT  TGTKGAANNT  ACTAGACTCT  CTTTTTTTCC  KGCCTTTTTG  1120
TKGCTTCAAA  AGTTRACACT  TYCTATTCCC  ACATCATTGT  TAAANCGAAA  TTTAATTGCA  1180
ACAGATTTTT  GACAAGNCAG  CGATCATAGA  TGAAATTGAT  GCCAAGACGA  AGACAAGAGA  1240
CCCTTGGAGG  CTATGTTCAG  TAACACAAGT  TGAGGAAGTG  AAGCTTGTCC  TTCGTTTGAT  1300
TCCAATATGG  CTAAGTTGCT  TAATGTTCAC  TGTTGTACAA  GCTCAAGTGC  ACACATTTTT  1360
CATCAAGCAA  GGTGCCACAA  TGGTACGTAC  CATAGGACCA  CATTTCCAAG  TCCCACCAGC  1420
ATCGCTCCAA  GGCCTAGTTG  GAGTCACAAT  ACTCTTTGCT  GTGCCATTCT  ATGACCGTGT  1480
CTTTGTGCCA  CTAGCAAGGA  AAATCACAGG  GAAACCCACT  GGGATAACAG  TGCTACAAAG  1540
AATTGGGGTA  GGACTTTTCT  TGTCAATCCT  TAACATGGTT  GTGTCAGCAC  TTGTGGAGGC  1600
CAAAAGGGTT  GGTGTTGCAA  AAGAGAGTGG  CCTAATTGAT  GACCCAAAAG  CAGTGTTACC  1660
AATCAGCATT  TGGTGGCTGC  TTCCTCAGTA  CATGATCACT  GGGATCTCTG  ATGCATTCAC  1720
AATTGTGGGG  CTACAAGAGT  TGTTTTATGAC  CAAATGCCT  GAGGCACTTA  GGAGTTTGGG  1780
GGCTGCA ↓ GGT  3'                                                      1787
```

Figure 5.2 Sequences of the ends of the cDNA probe B212. The positions of the forward and reverse primers are indicated with double lines. Downward arrows indicate *PstI* sites.

Restriction analysis of PCR products with a wide range of enzymes produced informative digestion patterns with *HindIII*, *TaqI*, *AluI* and *DraI* between Prima and Gazelle. Digestion of PCR products from 50 F_2 plants with *AluI* produced a highly complex pattern with no single fragment linked significantly to

variation in gall index (P>0.05). The patterns could be classified as similar to parent A, parent B or heterozygous genotypes. These patterns, scored as A, B or H, also showed no significant linkage to gall index values (P>0.05). Digestion patterns of PCR products with *HindIII*, *TaqI* and *DraI* did not reveal any significantly linked fragments (P>0.05).

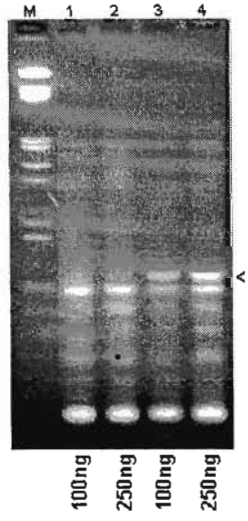


Figure 5.3 PCR analysis of parent genomic DNA with B212-primers. M: λDNA digested with *EcoRI* and *HindIII*. 1,2: Gazelle. 3,4: Prima.

5.4 DEVELOPMENT OF SCARS FROM AFLP FRAGMENTS

5.4.1 Sequencing data

Seven AFLP markers, identified in Chapter 4.3.2.1, which were significantly linked to gall index variation, were cloned and sequenced. Sequencing data of six of the markers is given in Figure 5.4.

pSOJA1 : AAC-CAC5 (177bp)

GATT**GACTGC**	**GTACCAATTC**	**AAC**TGTTGAG	ATACTTAGAG	ATGAAGGTTA	27
CTAACTGACG	CATGGTTAAG	TTG ACAACTC	TATGAATCTC	TACTTCCAAT	
TATATAGAAT	AAGCTTTGAA	GAAACAAGAC	ACGAATCACC	TATGTGAATC	77
ATATATCTTA	TTCGAAACTT	CTTTGTTCTG	TGCTTAGTGG	ATACACTTAG	
ATTCTTTCAT	TATTTCTTGT	GAAACTTTTT	GTAAATTCTT	GTAAAGTAAA	127
TAAGAAAGTA	ATAAAGAACA	CTTTGAAAAA	CATTTAAGAA	CATTTCATTT	

GATACAAAGC TTTCAAAACG CCTTATATAC CTTGAGAGAA AAAACTAAAA 177
CTATGTTTCG AAAGTTTTGC GGAATATATG GAACTCTCTT TTTTGATTTT

GTGTTACTCA GGACTCATCA ATCACTAGTG AATT
CACAATGAGT **CCTGAGTAGT** TAGTGATCAC TTAA

pSOJA3 : AAG-CTA5 (371bp)
TTCGATT**GAT** **GAGTCCTGAG** **TAACTA**GTTA GAATTGGCTT GTTTACAGCT 24
AAGCTAACTA CTCAGGACTC ATTGATCAAT CTTAACCGAA CAAATGTCGA

TTGCACGTTT GATTTAGATA GCCAATAGAA AACAATTTTT TTTCATCTGA 74
AACGTGCAAA CTAAATCTAT CGGTTATCTT TTGTTAAAAA AAAGTAGACT

GACATGTCTG CTTATGTGGT GGACATACAT CCTCATTCAT TGCACACGCA 124
CTGTACAGAC GAATACACCA CCTGTATGTA GGAGTAAGTA ACGTGTGCGT

AGAGACAAAA GGAGAGCCAT ATGGAAAATA TTTTTTTTTT TTTTTGGTAT 174
TCTCTGTTTT CCTCTCGGTA TACCTTTTAT AAAAAAAAAA AAAAACCATA

AAGCCGAAAA CATAGACTCA GACATACAAC AGATACTGTA CGAGTACACA 224
TTCGGCTTTT GTATCTGAGT CTGTATGTTG TCTATGACAT GCTCATGTGT

TGTGACAGCA ATGACACCAA CGCCAAATCC ACTAAGCTCC CCCCATCATA 274
ACACTGTCGT TACTGTGGTT GCGGTTTAGG TGATTCGAGG GGGGTAGTAT

CAAAGCAAAT ATGAGGGGCC AAAAATTAGA CGCCTTTGCT GCTGCTGCTC 324
GTTTCGTTTA TACTCCCCGG TTTTTAATCT GCGGAAACGA CGACGACGAG

TATGCAACAT GAACCTAGAA ATCGTTAGTA GATAGGCATG GATGGAT**CTT** 371
ATACGTTGTA CTTGGATCTT TAGCAATCAT CTATCCGTAC CTACCTAGAA

GAATTGGTAC **CAGTC**AATC ACTAGTGAAT
CTTAACCATG CGTCAGTTAG TGATCACTTA

pSOJA4 : AAG-CTA6 (369bp)
GGCGGCCGCG GGAATTCGAT T **GACTGCGTA** **CCAATTCAAG** ATCCATTCAT 10
CCGCCGGCGC CCTTAAGCTA A CTGACGCAT GGTTAAGTTC TAGGTAAGTA

GTCTATGTAC TAACCGA TTT CTAGGCTCAT GTTGCATAGA GCAGCAGCAG 60
CAGATACATG ATTGGCTAAA GATCCGAGTA CAACGTATCT CGTCGTCGTC

CAAAGGCGTC TAATTTTTGG CCCCTCATAT TTGCTCTGTA TGATGGGGGG 110
GTTTCCGCAG ATTAAAAACC GGGGAGTATA AACGAGACAT ACTACCCCCC

AGCTTAGTGG ATTTGGCGTT GGTGTCATTG CTGTCACATG TGTACTCGTA 160
TCGAATCACC TAAACCGCAA CCACAGTAAC GACAGTGTAC ACATGAGCAT

CAGTATCTGT TGTATGTCTG AGTCTATGTT TTCGGCTTGT ACCAAAAAAA 210
GTCATAGACA ACATACAGAC TCAGATACAA AAGCCGAACA TGGTTTTTTT

AAGGATATTT TCCATATGGC TCTCCTTTTG TCTCTTGCAT GTGCAATGAA 260
TTCCTATAAA AGGTATACCG AGAGGAAAAC AGAGAACGTA CACGTTACTT

TGGGGATGTA TGTCCACCAC ATAAGCAGAC ATGTCTCATA TGAAAAAAAA 310
ACCCCTACAT ACAGGTGGTG TATTCGTCTG TACAGAGTAT ACTTTTTTTT

TTGTCTTCTA TTGGCTATCT AAATCAAACG TGCAAAGTTG TAGACAAGCC 360
AACAGAAGAT AACCGATAGA TTTAGTTTGC AGCTTTCAAC AT CTGTTCGG

AATTCGAAC T **AGTTACTCAG** **GACTCATC** AA TCACTAGTGA ATT 369
TTAAGCTTG A TCAATGAGTC CTGAGTAGTT AGTGATCACT ATT

pSOJA6 : ACC-CTC2 (244bp)

GGCGGCCGCG	GGAATTCGAT	T **GACTGCGTA**	**CCAATTCACC**	ATCATTACG C	10
CCGCCGGCGC	CCTTAAGCTA	ACTGACGCAT	GGTTAAGTGG	TAGTAATGCG	
ATGGGCCATC	CTAGAG AAAA	GCAATTATAA	CACTGAGATA	TTATGGAACA	60
TACCCGGTAG	GATCTCTTTT	GCTTAATATT	GTGACTCTAT	AATACCTTGT	
CATGGAAAGT	GTACTCACAT	TGTGAAATAT	GTAGCGTAKA	CACTCTGGCC	110
GTACCTTTCA	CATGAGTGTA	ACACTTTATA	CATCGCAT?T	GTGAGACCGG	
TGAAACGAAC	ATTGGACGCT	TCACCGTAAA	CTATGTTTTC	GTCCCAAGCA	160
ACTTTGCTTG	TAACCTGCGA	AGTGGCATTT	GATACAAAAG	CAGGGTTCGT	
CTTGCACCAT	GATTTGKAAT	TCTTGAAGAG	ATTTTCCATC	GCTGAATCAA	210
GAACGTGGTA	CTAAACMTTA	AGAACTTCTC	TAAAAGGTAG	CGACTTAGTT	
CAGCCCAATC	ATCGAGCTGA	TTTGGTACAA	CCGA **GAGTTA**	**CTCAGGACTC**	244
GTCGGGTTAG	TAG CTCGACT	AAACCATGTT	GGCTCTCAAT	GAGTCCTGAG	
ATCAATCACT	AGTGAATTCG	CGG			
TAGTTAGTGA	TCACTTAAGC	GCC			

pSOJA7 : AAC-CAT1 (246bp)

GATT**GACTGC**	**GTACCAATTC**	**AAC**TTTGAGA	TCACTTGGCT	TGATAGGAGA	27
CTAACTGACT	CATGGTTAAG	TTGAAACTCT	AGTGAACCGA	ACTATCCTCT	
TCGATTGTTT	TAGATCCCAA	ATCTTGATGT	TTCTTTCCTC	CCTTCCACTG	77
AGCTAACAAA	ATCTAGGGTT	TAGAACTACA	AAGAAAGGAG	GGAAGGTGAC	
TAGCTGTTCA	AAAACTTTAC	AGATAAAGCT	TGTGATAATT	TCTGTTTGTA	127
ATCGACAAGT	TTTTGAAATG	TCTATTTCGA	ACACTATTAA	AGACAAACAT	
AAATCACAGT	AACAAGAAGA	TTTTACCATA	AATGAAGTTG	TGAACAATAT	177
TTTAGTGTCA	TTGTTCTTCT	AAAATGGTAT	TTACTTCAAC	ACTTGTTATA	
ATCCTACATC	ATGATATTTT	TATGCAATCA	AAGAGAATTA	TTAGGTGATT	227
TAGGATGTAG	TACTATAAAA	ATACGTTAGT	TTCTCTTAAT	AATCCACTAA	
TAGGATCAGT	ACATCATTTA	TGTTACTCAG	GACTCATCAA	TCACTAGTGA	246
ATCCTAGTCA	TGTAGTAAAT	**ACAATGAGTC**	**CTGAGTAG**TT	AGTGATCACT	

PSOJA9 : AGC-CTG5 (83bp)

ATTCGATT**GA**	**CTGCGTACCA**	**ATTCAGC**TAA	GCTACAT GTA	AGAGAGGAAT	23
TAAGCTAACT	GACGCATGGT	TAAGTCGATT	CGATGTACAT	TCTCTCCTTA	
GACCAGGCAT	CAAAATGCGA	CTTCATGCAA	GGAATGTACT	GCCTTCCTTC	73
CTGGTCCGTA	GTTTTACGCT	GAAGTACGTT	CCTTACATGA	CGGAAGGAAG	
ATTTGCTACC	CAGTTACTCA	GGACTCATCA	ATCACTAGTG	AATTCGCGGC	83
TAAACGATGG	**GTCAATGAGT**	**CCTGAGTAG**T	TAGTGATCAC	TTAAGCGCCG	

Figure 5.4 Sequencing data of cloned AFLP fragments. *EcoRI* and *MseI* primer sequences are indicated in bold type. SCAR primers are double underlined.

Fragment E-AGC/M-CTG6 was not cloned successfully.

5.4.2 *PCR of individual F_2 plants*

After optimization of the reaction conditions with genomic DNA from the two parents, the segregation of the polymorphisms was analysed in the F_2 progeny. pSOJA1 derived primers (Table 5.1) did not amplify the expected polymorphism in the range of 100 bp between Prima and Gazelle at the optimum reaction conditions (annealing temperature of 40°C and 3 mM $MgCl_2$). Three polymorphisms were amplified between 1 000 bp and 1 300 bp. Two of these fragments were present in Gazelle and one in Prima. Reactions performed with 38 F_2 plants indicated no significant linkage of either of these fragments with variation in gall index ($P>0.05$). Banding patterns scored as similar to parent A, parent B or heterozygous, also did not reveal any significant linkage. No polymorphisms linked to gall index was amplified at higher temperatures.

Amplification of fragments with primers derived from pSOJA3 (Table 5.1) occurred at temperatures between 40°C and 55°C at 2 mM or 3 mM $MgCl_2$. A total of 5 fragments were amplified with polymorphic fragments of >1 000 bp at 50°C and 45°C. These fragments were not significantly linked to gall index variation ($P>0.05$). A ±500 bp fragment was polymorphic between Prima and Gazelle at 55°C (3 mM $MgCl_2$), but was not linked significantly to the resistance trait. Similarly, pSOJA4 primers amplified 4 fragments at 40°C (3 mM $MgCl_2$), with two polymorphic fragments between the two parents. Analysis of the progeny did not reveal any significant linkage of these fragments with the resistance trait ($P>0.05$).

Amplification of genomic DNA of the parents and progeny with primers of pSOJA6 revealed a codominant segregation pattern linked significantly ($P=0.0001$) to gall index variation (Figure 5.5). A total of 5 fragments were amplified at 55°C (2 mM $MgCl_2$), of which two were monomorphic and three polymorphic between the two parents. The three polymorphic fragments were all smaller than 560 bp. Two of these amplified in Gazelle and one in Prima, and segregated according to a codominant pattern. The codominant SCAR fragments explained 41% of variation in gall index in the mapping population, and segregated in a 1:2:1 relationship.

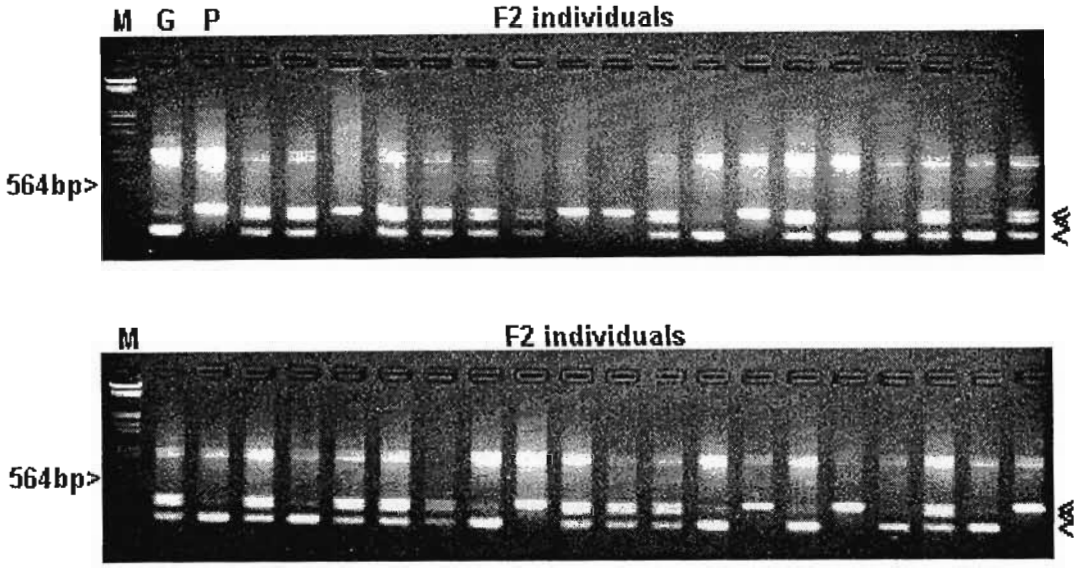


Figure 5.5: Amplification of genomic DNA from parents, Gazelle (G) and Prima (P), and 38 F_2 progeny with pSOJA6 primers. M=Lambda DNA *EcoRI/HindIII*.

A total of 9 fragments were amplified with the sequence specific pSOJA7 primers (Table 5.1). Three polymorphisms were amplified between Gazelle and Prima at an annealing temperature of 45°C with 3 mM $MgCl_2$ (Figure 5.6A). Two bands between 500 and 1 500 bp segregated in the $_2$F plants, but no significant linkage with gall index variation could be established (P>0.05). A polymorphic band at the predicted 240 bp showed significant linkage with gall index variation (P=0.000) and explained 42% of variability in the trait. The fragment was present in Prima and therefore linked in repulsion. The band did not segregate in a 1:1 relationship, and it can therefore not be concluded that it is a single locus - a postulate that is also observed in the multiple bands amplified by the sequence specific primers.

Six fragments were amplified with the designed primers for pSOJA9 at 50°C (2 mM $MgCl_2$), with a polymorphism between Gazelle and Prima at <500 bp. The polymorphism segregated in the F_2 progeny, but was not linked significantly to gall index variation (P>0.05).

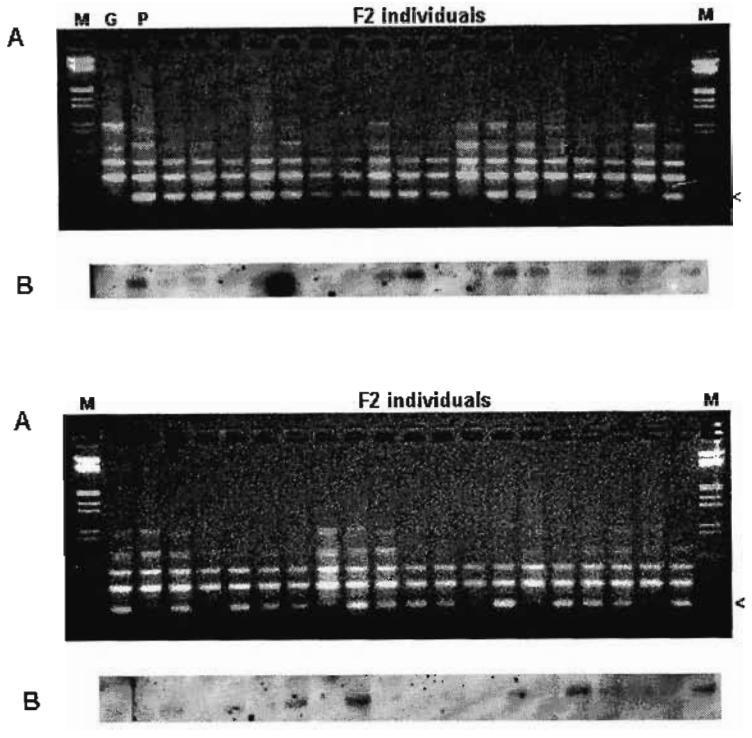


Figure 5.6: (A) Amplification of genomic DNA from parents, Gazelle (G) and Prima (P), and 38 progeny with pSOJA7 primers. M=Lambda DNA *EcoRI/HindIII*.
(B) Segregation of pSOJA7 in an RFLP blot with *EcoRI* digested genomic DNA of Gazelle (G), Prima (P) and 38 F_2 progeny.

5.4.3 Verification by Southern analysis

The cloned pSOJA6 and pSOJA7 fragments were labelled with DIG and used as probes to hybridize to *EcoRI* digested genomic DNA of the parent and F_2 progeny plants. Fragment pSOJA6 hybridised to 6 distinct bands, with two fragments (5 800 bp and 950 bp) polymorphic between the two parents. These two fragments displayed codominant Mendelian segregation between the F_2 progeny (results not shown). Three distinct bands could be distinguished using the pSOJA7 fragment as probe, with one band (3 700 bp) polymorphic between the parents, and present only in Prima (dominant polymorphism) (Figure 5.6B). This fragment segregated between the F_2 progeny with exactly the same pattern as the PCR product.

pSOJA1 gave rise to 20 distinct bands, whereas pSOJA3 and pSOJA4 led to 4 to 9 bands with a background smear. pSOJA9 gave rise to 3 distinct bands on the Southern Blot. No polymorphisms were identified by any of these fragments.

5.5 DISCUSSION

Attempts were made for the development of SCAR markers from both the RFLP probe (B212) and several AFLP markers. The multiple fragments amplified with the newly designed sequence specific B212 primers indicated that the primers were not specific to one locus. The polymorphic $B212_{600}$ fragment was much smaller than the expected ±2 000 bp and did not originate from the B212 locus on LG-F. As expected, this fragment did not cosegregate with gall index variation. The ±2 000 bp fragment amplified was not polymorphic between the parents, and restriction digestion of the PCR products did not reveal significant linkage to the variation in gall index. A possible explanation for the failure of the sequence specific primers to amplify a polymorphic fragment linked to resistance, lies in the origin of the probe itself. The RFLP probe B212 was cloned from a cDNA library. The primers could therefore bracket a piece of DNA including intron sequences, which could render the fragment too long to be optimally amplified by Promega Taq DNA polymerase. The multiple fragments amplified mean that the primers are not specific for this locus and the shorter fragments could possibly compete for hybridisation to the primers. One solution would be the designing and testing of multiple primer pairs, bracketing shorter fragments.

Specific oligonucleotide primers were designed for 6 AFLP fragments after sequence analysis. All of the fragment sequences analysed were very rich in adenine and thymine, which could be explained by the choice of enzymes for digestion, namely *EcoRI* (G↓AATTC) and *MseI* (T↓TAA), both targeting AT-rich regions of the genome.

Cloning of AFLP markers has been described in only a few reports (SHAN, *et al.*, 1999, MEKSEM, *et al.*, 1995, CHO, *et al.*, 1996, QU, *et al.*, 1998). Amplified products of most of the primer sets (developed for AFLP fragments) tested by SHAN, *et al.* (1999) were not polymorphic, even after digestion with a battery of restriction enzymes. MEKSEM *et al.* (1995) also did not obtain polymorphic fragments after development of sequence specific primers for AFLP fragments, but found linked polymorphisms after restriction digestion of PCR products. CHO *et al.* (1996) successfully cloned AFLP fragments giving single

copy sequences which displayed normal Mendelian segregation in a mapping population as determined with Southern analysis.

The two AFLP markers mapping closely to and bracketing the resistance trait (Chapter 4.4), E-ACC/M-CTC2 and E-AAC/M-CAT1, were successfully converted to SCARs (SOJA6 and SOJA7). The SOJA6 marker amplified three polymorphic fragments which segregated in a codominant manner in an F_2 population according to a Mendelian segregation pattern. The marker explained 41% of gall index variation. The dominant SOJA7 marker was linked in repulsion phase and was shown to explain 42% of variation in gall index. Although a single locus is probably not amplified by the sequence specific primers, the 240 bp fragment cosegregated with gall index variation. Southern analysis of genomic DNA of the parents and progeny with the cloned SOJA7 fragment displayed segregation of a single dominant marker. These markers can thus be employed for MAS in a breeding population, with SOJA6 distinguishing between homozygotic and heterozygotic progeny, and SOJA7 selecting against homozygous resistant plants, which can be highly efficient in segregating populations.

The inheritance of resistance to root-knot nematode in soybean is quantitative and has a moderate to high heritability (TAMULONIS, *et al.*, 1997b). The objective of this study was the development of a practically applicable marker in the soybean breeding program, and thus focussed on the main QTL affecting nematode resistance. Both markers identified accounted for 41-42% of gall index variation individually, which was of the same magnitude as the heritability estimate for other soybean populations (h^2 = 0.48-0.76) (TAMULONIS *et al.*, 1997b). The theoretical limit of percent genotypic variation that can be accounted for is 0.48-0.76 (heritability estimate). The SCAR markers explained 88% of the lower estimate (0.48) and the RFLP marker explained 62%, which accounts for 100% of the average heritability estimate (0.62). Additional QTLs would likely be of minor importance in terms of its contribution to explaining variation in gall index.

5.6 CONCLUSIONS

Two AFLP markers bracketing the resistance trait were both successfully converted to SCAR markers. The coupling phase marker, E-ACC/M-CTC2, were converted to a codominant SCAR marker (SOJA6), explaining 41% of variation in gall index in the mapping population. The repulsion phase AFLP marker (E-AAC/M-CAT1) was successfully converted to a dominant SCAR marker (SOJA7). This represents the first

report of the development of PCR-based sequence specific markers linked to resistance to *M. javanica* in soybean. The use of these markers in a breeding program can lead to highly efficient selection of homozygous resistant individuals. The SCAR markers explained 41% and 42% respectively of variation in gall index (Chapter 4.4). An RFLP marker B212 was closely linked to resistance and explained 62% of the variation in the mapping population. The development of a SCAR marker from the RFLP probe B212 was as yet unsuccessful.

CONCLUSIONS

Breeding for nematode resistance in soybean is a lengthy and tedious process which is largely dependent on environmental factors. Almost no resistance to *M. javanica*, the most widespread nematode in South Africa, is present in local soybean cultivars. An efficient and economically viable selection method had to be developed for breeding of urgently needed new varieties with resistance to the nematode pest. Although a genetic marker was identified in an American soybean population (TAMULONIS, *et al.*, 1997b), this marker has several disadvantages. The marker accounted for 46% of variation in gall number in the screening population, but the applicability of the marker in foreign genetic material was not established and had to be evaluated in South African breeding material. Secondly, it is an RFLP marker, which is very reliable, but tedious and impractical for use in screening large populations. A PCR based marker is preferable as this lends itself to easy, inexpensive and time saving screening, which can be fully automated.

Resistance to *M. javanica* was identified in the Zimbabwe cultivar, Gazelle, and used in the development of a mapping population for identification of molecular markers linked to the trait. Three different molecular techniques (RAPD, RFLP and AFLP) were used for identification of markers linked to the resistance trait. It was found that the RAPD technique was not suitable for the identification of markers in individuals in segregating populations. This was due to the fact that individual plant variation was detected, i.e. the technique was too sensitive (Chapter 3), and in this case, unreliable. However, this problem can be overcome by pooling DNA samples of a number of individuals, for example when varieties are being identified by fingerprinting. It is evident that the pooling of DNA samples from individuals allows the easy and very rapid identification of varieties (Chapter 2). In addition, this technique could constitute for marker identification if individual variation is minimized (removed) through bulking techniques such as BSA, provided that reliable linkage can be established with phenotypic variation, or with the use of NILs.

Two phenotypic parameters were evaluated (Chapter 3), i.e. reproductive ability of nematodes (Rf) and gall index of plant roots, both displaying continuous patterns of distribution. The gall index values could be classified in three groups with a 1:2:1 phenotypic ratio. This could indicate a partially dominant inheritance pattern with at least one gene involved in this particular cross. Closer examination of the Rf-values of the F_1 and F_2 plants revealed that the F_1 plants had intermediate resistance, suggesting that this data could fit a hypothesis of 2 genes contributing to the resistance trait. Thus, whether it is considered one gene with incomplete dominance or 2 genes, depends largely on the level of analysis. These findings

are in accordance with results obtained by other groups. Resistance to *M. javanica* in three American genotypes, Gordon, PI80466 and PI230977 was found to be quantitatively inherited (LUZZI, *et al.*, 1995b). LUZZI, *et al.* (1995b) used more than one cross and therefore could make a more accurate suggestion on the inheritance of the trait in different plants, possibly different genes in the different breeding lines.

Variation in the relationship between the two different phenotypic parameters (gall index and Rf-value) for individual plants was observed (Chapter 3). This could be explained by two possibilities: (1) The inheritance of resistance to gall formation could be independent of the resistance to support nematode reproduction, or (2) it could be due to experimental variability in measurements between individual plants. It was clear that the method of determination of reproductive ability of the nematodes inherently lends itself to experimental variation between duplicate samples from the large variability observed within the control populations. Less variability was observed between replicates evaluated with gall index measurements, where the evaluation was based on relative values, rather than exact numbers. In view of these findings, it was concluded that the use of determination of reproductive ability for individual plants held a greater risk of experimental error in determination of plant phenotype and therefore relative gall index values were used in subsequent experiments.

RFLP and AFLP analysis were further explored in linkage analysis with variation in gall index values of a segregating F_2 population (Chapter 4). The population was screened with 16 RFLP probes for cosegregation with the resistance trait. The RFLP patterns obtained were compared to published results for linkage of markers to the public soybean molecular map. RFLP marker B212 linked closely to the resistance trait and explained 62% of variation in gall index values, giving a tighter linkage as in the original mapping population (46%) as found by TAMULONIS, *et al.* (1997b), suggesting that it could probably be the same gene or at least a gene located in the same region on Linkage Group F (LG-F). The other marker found by TAMULONIS, *et al.* (1997b), A725-2, situated on LG-D1, and which accounted for only a minor part (13%) of gall variation, was not polymorphic between Gazelle and Prima. None of the other markers which mapped to LG-D were linked to the resistance trait. According to the various maps for LG-D/LG-D1 published on the internet (http://probe.nalusda.gov:8000/plant/ aboutsoybase.html), the distance between these markers (K411, B142-2 as anchor markers) could be between 150 cM and 450 cM, which would explain the lack of close linkage, especially if the QTL is situated on the far side of A725-2. As the main QTL with large effect (62%) was identified on LG-F, it was decided to concentrate on this region for the development of a marker system which would be easily and economically applicable in marker assisted selection in a breeding program. As it was found that this marker was situated in a cluster of several other

disease resistance loci (TAMULONIS, *et al.*, 1997b), a marker in this region could possibly also find application in detection of these loci.

A three step design was followed for identification of linkage of AFLP markers to the resistance trait, in which a process of elimination was used to minimize the samples to be analysed. The strategy included bulked segregant analysis and selective genotyping (Chapter 4). Two DNA pools were constructed from the two extremes of the population, using two criteria - phenotype (gall index) and genotype according to RFLP probe B212. The DNA pools were screened with 64 AFLP primer combinations, followed by genotyping of the individual plants included in the bulks. Polymorphic markers were linked to gall index variation with analysis of variance, and markers displaying significant linkage were tested for segregation in the larger F_2 population. This strategy meant that the identification of AFLP markers in this study was biased in two ways - the combined bulked segregant analysis and selective genotyping enriched the fraction tested for the area on the chromosome around the QTL with main effect on the resistance trait. This was clearly illustrated in the density of the markers assigned to LG-F (Chapter 4). Secondly, it could also be biased towards relatively AT-rich regions due to the two enzymes used in the digestion of the genomic DNA, namely *EcoRI* (G↓AATTC) and *MseI* (T↓TAA), both targeting these regions.

A linkage map was constructed with the limited amount of data available which spanned a total of 74.8 cM, including 64 RFLP and AFLP markers converging in nine linkage groups, with 34 unlinked markers (Chapter 4). Seven AFLP markers, identified by analysis of variance as closely linked to resistance to gall formation, were grouped together with marker B212, which could be anchored on LG-F. Four of the markers were linked in repulsion and three in coupling phase. The total distance covered by the markers on LG-F was 19.0 cM. Marker B212 was flanked by AFLP markers on both sides, with marker E-AAC/M-CAT1 (284 bp) mapping the closest to B212 at 2.4 cM. The other 6 markers mapped to the opposite side of B212 with marker E-ACC/M-CTC2 (281 bp) at a distance of 3.8 cM. Three of the markers were completely linked to each other, i.e. distance 0 cM. Marker E-AAG/M-CTA5 (409 bp) mapped the furthest away from B212 at 16.6 cM. As only one known RFLP marker was polymorphic which could be linked to LG-F, the orientation of this map on the classical map could not be established.

The AFLP markers were cloned and sequenced and markers E-AAC/M-CAT1 and E-ACC/M-CTC2 were successfully converted to SCAR markers (Chapter 5). E-ACC/M-CTC2 was converted to a codominant SCAR marker (SOJA6) which accounted for 41% of gall index variation. E-AAC/M-CAT1 converted to a dominant SCAR marker (SOJA7) explaining 42% of variability in phenotype. This represents the first

report of the development of PCR-based sequence specific markers linked to resistance to *M. javanica* in soybean. The main QTL for gall index resistance mapped between markers B212 and E-AAC/M-CAT1 (SOJA7), with these two markers 2.4 cM apart. Marker E-ACC/M-CTC2 (SOJA6) mapped 3.8 cM from B212, with the resistance trait bracketed between markers SOJA6 and SOJA7.

The inheritance of resistance to root-knot nematode in soybean was determined to have a moderate to high heritability (h^2=0.48-0.76) (TAMULONIS, *et al.*, 1997b). The two SCAR markers accounted for 41-42% of gall index variation and explained 88% of the lower estimate (0.48), and the RFLP marker (B212) accounted for 100% of the average of the heritability estimate (0.62). Additional QTLs would likely be of minor importance in its contribution to explaining variation in gall index.

The utility and application of these markers would have to be validated in the current soybean breeding program. Previous studies have shown the application of both repulsion and coupling-phase markers or codominant markers in different selection populations in breeding programs. The selection approach must be determined by the nature of the population. In MAS of individuals in F_2 or later segregating generations, the homozygous resistant plants could be distinguished from heterozygous resistant plants by detecting the absence of a repulsion-phase marker (BAI, *et al.* 1995). HALEY, *et al.* (1994) also found that the marker could provide greater selection efficiency than coupling-phase markers and provided a greater proportion of homozygous resistant selections, with a lower proportion of both segregating and homozygous susceptible selections. Selection based on a repulsion-phase marker could therefore be identical to selection based on a codominant locus such as most RFLP loci (HALEY, *et al.*, 1994). Selection with a repulsion-phase marker in a BC_nF_1 (Rr:rr) population of traditional backcross breeding will however, eliminate all progeny and will not be viable.

As the coupling-phase marker E-ACC/M-CTC2 converted to a codominant SCAR marker (SOJA6), the problems with applicability were overcome and it can be applied in any selection population. The combined use of both markers, bracketing the resistance trait and minimizing the possibility of cross-overs in this genomic region, can lead to highly efficient selection of homozygous resistant individuals. The presence or absence of both SCAR markers can be determined by a simple PCR reaction and an efficient and economically viable technique for MAS of resistance to *M. javanica* was therefore successfully developed in this study.

APPENDICES

APPENDIX A

Phenotypical data of population PG3-1, PG3-2

	ROOT MASS*	PODS*	SEEDS*	EGG MASSES*	EGGS	RF	GALL INDEX*
GAZELLE BB	52.47	55			2300	0.46	1
GAZELLE CC	86.92	42			1225	0.25	1
GAZELLE DD	69.27	43			2683	0.54	1
GAZELLE EE	61.29	40			2042	0.41	1
GAZELLE FF	145.93	40			11550	2.31	1
GAZELLE II	67.68	53			8050	1.61	1
GAZELLE JJ	135.14	46			2888	0.58	1
GAZELLE KK	81.98	45			4620	0.92	1
GAZELLE LL	57.19	59			4060	0.81	1
GAZELLE-A	203.25	2		9	5863	1.17	1
GAZELLE-AA	120.66	31			5040	1.01	1
GAZELLE-B	203.89	6		4	1925	0.39	1
GAZELLE-C	163.24	14		3	7000	1.40	1
GAZELLE-D	216.92	12		0	3080	0.62	1
GAZELLE-E	111.47	3			10938	2.19	1
GAZELLE-F	103	13			2485	0.50	1
GAZELLE-G	92.74	13			8400	1.68	1
GAZELLE-H	117.17	13			1330	0.27	1
						AVG 0.95 ±0.63	
PRIMA-A	131.2	5		>100	391806	78.36	9
PRIMA-B	161.34	14		82	275450	55.09	9
PRIMA-C	158.48	24		>100	367150	73.43	9
PRIMA-D	164.43	7		>100	416023	83.20	9
PRIMA-E	141.55	52		>100	138950	27.79	9
PRIMA-F	134.28	22		>100	35000	7.00	9
PRIMA-G	149.58	18		>100	106050	21.21	9
PRIMA-H	93.53	33			69125	13.83	9
PRIMA-I	82.98	33			84000	16.80	9
PRIMA-J	92.34	23			107800	21.56	9
PRIMA-K	104.66	17			122150	24.43	9
PRIMA-L	78.45	19			70700	14.14	9
						AVG 36.40 ±26.76	
PG3-1-2	85.37	23		>100	81550	16.31	
PG3-1-3	124.2	29			101500	20.30	5
PG3-1-5	88.86	31			82600	16.52	9
PG3-1-6	97.87	29		>100	217700	43.54	
PG3-1-7	130.77	31		>100	290150	58.03	
PG3-1-8	77.12	20		>100	143850	28.77	
PG3-1-9	76.65	36			114450	22.89	3
PG3-1-10	113.06	27		>100	78400	15.68	7
PG3-1-11	75.82	31		>100	185500	37.10	
PG3-1-12	21.11	10		37	10430	2.09	
PG3-1-13	118.13	36		>100	84350	16.87	
PG3-1-14	100.13	31		>100	157850	31.57	
PG3-1-15	116.4	30		>100	150500	30.10	
PG3-1-16	117.05	22		>100	163800	32.76	
PG3-1-17	87.9	26		>100	183750	36.75	
PG3-1-18	135.36	35		>100	61600	12.32	
PG3-1-19	108.8	32		>100	393313	78.66	
PG3-1-20	113.35	27		>100	231350	46.27	
PG3-1-21	121.27	30		>100	95900	19.18	
PG3-1-22	109.87	26		>100	88900	17.78	
PG3-1-25	107.12	40		>100	20125	4.03	
PG3-1-31	79.64	22		>100	124950	24.99	
PG3-1-32	87.41	31		>100	33250	6.65	
PG3-1-33	120.26	39		>100	85400	17.08	

	ROOT MASS*	PODS*	SEEDS*	EGG MASSES*	EGGS	RF	GALL INDEX*
PG3-1-34	149.6	32		>100	38150	7.63	
PG3-1-35	122.92	24		>100	36400	7.28	
PG3-1-37	166.71	32		>100	189000	37.80	
PG3-1-38	119.28	33		>100	102900	20.58	
PG3-1-41	120.16	21		>100	148750	29.75	
PG3-1-42	143.09	26		>100	100450	20.09	
PG3-1-44	65.47	27			34825	6.97	7
PG3-1-45	143.36	31		>100	110600	22.12	
PG3-1-46	166.56	28		>100	45850	9.17	
PG3-1-49	176.37	33		>100	77350	15.47	
PG3-1-50	154.01	35		>100	132650	26.53	
PG3-1-53	69.8	34			27475	5.50	6
PG3-1-54	64.55	44			16975	3.40	5
PG3-1-55	94.10	44			9917	1.98	7
PG3-1-57	88.47	44			14817	2.96	7
PG3-1-58	40.08	37			16450	3.29	8
PG3-1-59	66.91	42			6125	1.23	9
PG3-1-61	68.72	34			16683	3.34	6
PG3-1-65	89.91	15			77000	15.40	3
PG3-1-67	36.84	32			17325	3.47	4
PG3-1-69	66.02	44			40950	8.19	7
PG3-1-70	48.22	41			71400	14.28	4
PG3-1-71	79.90				45850	9.17	5
PG3-1-72	57.82	44			76650	15.33	5
PG3-1-74	60.58	3			13860	2.77	7
PG3-1-75	115.38	35			43750	8.75	5
PG3-1-76	77.98	31			53550	10.71	3
PG3-1-77	111.10	40			64050	12.81	6
PG3-1-78	120.32	30			57225	11.45	6
PG3-1-79	119.28	47			62300	12.46	6
PG3-1-80	105.21	44			51800	10.36	5
PG3-1-81	32.99	37		24	4130	0.83	2
PG3-1-83	54.39	46			20825	4.17	3
PG3-1-84	36.47	37			26950	5.39	4
PG3-1-85	33.96	35			14875	2.98	3
PG3-1-86	65.01	18			84350	16.87	5
PG3-1-88	57.97	42			38500	7.70	2
PG3-1-89	76.68	5			65100	13.02	5
PG3-1-90		9		118	16100	3.22	2
PG3-1-91	97.03	19			93100	18.62	8
PG3-1-92	70.85	2			45500	9.10	4
PG3-1-93	65.60	4			98700	19.74	7
PG3-1-94	77.47	38			52850	10.57	6
PG3-2-1	140.34	28		>100	50400	10.08	
PG3-2-2	169.13	31		>100	26425	5.29	
PG3-2-3	114.18	16		>100	27300	5.46	
PG3-2-4	126.71	31		>100	46900	9.38	
PG3-2-7	141.56	34		>100	166250	33.25	
PG3-2-8	173.71	34		96	47250	9.45	
PG3-2-10	125.79	23		>100	57400	11.48	
PG3-2-11	111.69	16		>100	21000	4.20	
PG3-2-12	89.50	28		>100	40250	8.05	
PG3-2-13	68.50	42		85	35350	7.07	
PG3-2-14	139.72	30		>100	129850	25.97	
PG3-2-15	151.80	43		>100	89600	17.92	
PG3-2-16	70.66	32			36050	7.21	3
PG3-2-20	106.29	36		>100	82600	16.52	
PG3-2-21	153.15	36		>100	227850	45.57	
PG3-2-22	74.55	43		>100	84700	16.94	
PG3-2-30	122.53	32		>100	80850	16.17	

	ROOT MASS*	PODS*	SEEDS*	EGG MASSES*	EGGS	RF	GALL INDEX*
PG3-2-32	181.23	11		>100	389375	77.88	
PG3-2-34	160.40	29		>100	87850	17.57	
PG3-2-35	209.27	17		>100	217350	43.47	
PG3-2-38	45.52	26			13738	2.75	8
PG3-2-39	213.22	30		>100	224700	44.94	
PG3-2-40	132.82	14		>100	170450	34.09	
PG3-2-41	132.74	20		>100	148750	29.75	
PG3-2-42	47.41	30			23100	4.62	3
PG3-2-45	207.71	29		>100	203350	40.67	
PG3-2-49	205.22	26		>100	222950	44.59	
PG3-2-51	38.95	45			38500	7.70	4
PG3-2-52	53.03	45			32550	6.51	4
PG3-2-53	43.63	43			43400	8.68	3
PG3-2-54	27.43	33			16625	3.33	3
PG3-2-55	45.56	44			113050	22.61	6
PG3-2-56		36		72	16800	3.36	2
PG3-2-57	51.70	30			88200	17.64	4
PG3-2-58	86.99	32			63350	12.67	6
PG3-2-60		3		214	48650	9.73	2
PG3-2-61	37.10	39			24850	4.97	3
PG3-2-62	86.52	3			143150	28.63	7
PG3-2-63	46.94	18			50750	10.15	4
PG3-2-64	96.01	36			3617	0.72	3
PG3-2-65	63.87	34		36	29750	5.95	2
PG3-2-66	134.67	41			44450	8.89	8
PG3-2-67	42.52	41		68	18900	3.78	2
PG3-2-68	25.39	37		58	7933	1.59	2
PG3-2-69	27.16	41			12775	2.56	3
PG3-2-70	18.74	12			31850	6.37	8
PG3-2-71	47.13	24			38850	7.77	4
PG3-2-72	73.11	40			33775	6.76	4
PG3-2-73	40.27	11			9217	1.84	3
PG3-2-74	13.89	18			2310	0.46	4
PG3-2-75	39.53	10			42350	8.47	3
PG3-2-77	90.81	1			85050	17.01	4
PG3-2-78	104.37	8			41650	8.33	8
PG3-2-79	87.80	15			69650	13.93	8
PG3-2-80	35.26	22			13038	2.61	3
PG3-2-82	52.16	25			10150	2.03	8
PG3-2-87	34.28	42			30800	6.16	7
PG3-2-88	30.82	43			16100	3.22	7
PG3-2-93	22.64	40			14000	2.80	9
PG3-2-99	110.97	33			57400	11.48	5
PG3-2-100	18.61	42			11550	2.31	6
						AVG	**15.31 ±14.50**

*Blank spaces indicate missing values.

APPENDIX B

Phenotypical data of population GP20-2

PLANT	GROWTH	PUBESC	PODS	SEED	ROOT MASS (g)	EGGS	RF	GALL INDEX	EGGS/g ROOT
GP20-2-1	I	B	99	178	129.97	6300	0.63	1	48
GP20-2-2	D	B	36	68	127.27	45150	4.52	9	355
GP20-2-3	I	B	30	53	169.19	45850	4.59	9	271
GP20-2-4	I	G	83	170	138.02	29400	2.94	3	213
GP20-2-5	I	B	33	68	111.24	243250	24.33	9	2,187
GP20-2-6	I	B	89	171	93.39	97300	9.73	5	1,042
GP20-2-7	I	B	59	102	66.44	136719	13.67	9	2,058
GP20-2-8	I	B	17	29	110.33	49700	4.97	9	450
GP20-2-9	D	G	28	42	105.83	105000	10.50	9	992
GP20-2-10	I	B	24	41	114.61	58450	5.85	9	510
GP20-2-12	I	B	30	50	172.07	11900	1.19	1	69
GP20-2-13	D	B	25	51	98.21	32200	3.22	9	328
GP20-2-14	D	B	24	39	89.12	32550	3.26	5	365
GP20-2-15	I	B	64	107	64.67	40950	4.10	1	633
GP20-2-16	I	B	32	64	132.32	42350	4.24	9	320
GP20-2-17	D	B	28	60	81.02	28000	2.80	7	346
GP20-2-18	D	G	24	41	116.2	90300	9.03	1	777
GP20-2-19	D	G	40	80	94.42	31500	3.15	1	334
GP20-2-20	I	B	21	40	94.58	23042	2.30	5	244
GP20-2-21	I	B	41	84	94.37	95550	9.56	3	1,013
GP20-2-22	I	B	30	54	117.66	50050	5.01	1	425
GP20-2-23	I	G	22	46	93.84	108150	10.82	7	1,152
GP20-2-24	D	B	51	106	89.14	41300	4.13	3	463
GP20-2-25	D	G	33	57	131.7	31850	3.19	1	242
GP20-2-26	I	B	29	53	108.63	128333	12.83	7	1,181
GP20-2-27	I	G	24	58	145.09	40600	4.06	7	280
GP20-2-28	I	B	27	47	116.76	25200	2.52	7	216
GP20-2-29	I	B	24	41	96.24	51450	5.15	3	535
GP20-2-30	D	G	26	49	90.11	27300	2.73	9	303
GP20-2-31	D	B	79	149	102.02	140000	14.00	5	1,372
GP20-2-32	I	B	24	49	85.67	80625	8.06	7	941
GP20-2-33	I	B	20	37	142.4	32426	3.24	5	228
GP20-2-34	I	G	36	55	134.41	8050	0.81	3	60
GP20-2-35	I	B	29	52	135.77	61250	6.13	9	451
GP20-2-36	I	B	23	46	110.84	27650	2.77	1	249
GP20-2-37	I	B	24	54	138.77	54250	5.43	7	391
GP20-2-38	I	B	18	36	75.63	110600	11.06	7	1,462
GP20-2-39	I	G	27	51	141.96	40250	4.03	5	284
GP20-2-40	I	G	24	53	129.69	13650	1.37	3	105
GP20-2-41	I	B	94	203	85.8	14700	1.47	5	171
GP20-2-42	I	B	26	48	77.64	28700	2.87	1	370
GP20-2-43	I	B	55	107	66.23	15050	1.51	5	227
GP20-2-44	I	B	19	36	94.81	57050	5.71	9	602
GP20-2-45	D	B	29	53	86.64	37100	3.71	3	428
GP20-2-46	I	B	24	46	65.47	24500	2.45	5	374
GP20-2-47	I	G	22	40	110.81	28000	2.80	7	253
GP20-2-48	I	B	22	44	130.34	133700	13.37	9	1,026
GP20-2-49	I	B	24	48	102.31	29050	2.91	3	284
GP20-2-50	D	G	25	43	105.07	111300	11.13	9	1,059
GP20-2-51	I	B	17	34	88.12	90650	9.07	5	1,029
GP20-2-52	I	B	25	47	140.67	16100	1.61	5	114
GP20-2-53	I	G	25	44	70.94	9450	0.95	1	133
GP20-2-54	I	B	32	64	112.11	5950	0.60	1	53

PLANT	GROWTH	PUBESC	PODS	SEED	ROOT MASS (g)	EGGS	RF	GALL INDEX	EGGS/g ROOT
GP20-2-55	I	G	26	46	62.91	25550	2.56	5	406
GP20-2-56	I	B	28	53	116.7	82250	8.23	7	705
GP20-2-57	I	B	25	58	93.72	50050	5.01	9	534
GP20-2-58	I	B	27	54	127.91	7000	0.70	1	55
GP20-2-59	D	B	30	49	121.5	33250	3.33	9	274
GP20-2-61	I	B	22	47	88.66	29050	2.91	5	328
GP20-2-62			28	49	108.36	17500	1.75	5	161
GP20-2-63	I	B	26	51	108.9	25900	2.59	5	238
GP20-2-64	D	B	22	41	66.15	101500	10.15	9	1,534
GP20-2-65	I	G	29	51	66.68	14350	1.44	5	215
GP20-2-66	I	B	14	21	116.72	238438	23.84	9	2,043
GP20-2-67	I	B	22	43	80.28	37800	3.78	5	471
GP20-2-68	I	G	28	53	103.59	8750	0.88	1	84
GP20-2-69	D	B	21	37	76.08	31150	3.12	7	409
GP20-2-70	I	B	31	57	71.4	35700	3.57	1	500
GP20-2-71	I	G	24	51	119.96	45500	4.55	9	379
GP20-2-72	D	B	23	48	61.1	9450	0.95	1	155
GP20-2-73			20	30	97.23	124950	12.50	9	1,285
GP20-2-74	I	B	23	43	92.3	67550	6.76	9	732
GP20-2-75	I	B	96	191	86.23	17150	1.72	5	199
GP20-2-76	I	B	24	35	55.62	31500	3.15	1	566
GP20-2-78	I	G	28	58	82.94	30800	3.08	3	371
GP20-2-79	D	B	25	57	91.46	20300	2.03	3	222
GP20-2-80	D	B	29	55	18.84	17850	1.79	7	947
GP20-2-81	I	G	25	38	127.29	277550	27.76	7	2,180
GP20-2-82	I	B	28	50	75.24	24150	2.42	5	321
GP20-2-83	D	G	31	51	88.1	17850	1.79	7	203
GP20-2-84			28	42	52.46	13650	1.37	3	260
GP20-2-87	I	B	27	49	53.69	47600	4.76	3	887
GP20-2-89	D	B	35	59	91.65	9800	0.98	5	107
						AVG.	5.32±5.20		
GAZELLE 1	D	G	53	100	113.09	9450	0.95	1	84
GAZELLE 3	D	G	31	55	131.55	14000	1.40	3	106
GAZELLE 4	D	G	29	55	169.05	10150	1.02	3	60
GAZELLE 5	D	G	25	42	111.93	7000	0.70	1	63
GAZELLE 6	D	G	105	209	179.56	2100	0.21	3	12
GAZELLE 7	D	G	26	48	121.27	9800	0.98	3	81
GAZELLE 12	D	G	39	79	122.84	12600	1.26	3	103
GAZELLE 9	D	G	23	39	100.69	10850	1.09	5	108
GAZELLE 10	D	G	19	41	102.97	5950	0.60	1	58
						AVG.	0.91 ±0.34		
PRIMA 1	I	B	20	43	79.83	137499	13.75	9	1,722
PRIMA 4	I	B	77	155	122.9	48650	4.87	7	396
PRIMA 5	I	B	76	155	143.54	52150	5.22	5	363
PRIMA 6	I	B	19	32	84.78	131250	13.13	9	1,548
PRIMA 8	I	B	19	34	45.43	112656	11.27	9	2,480
PRIMA 11	I	B	27	56	103.85	215833	21.58	9	2,078
PRIMA 12	I	B	17	38	91.6	256667	25.67	9	2,802
						AVG.	13.64±7.19		

APPENDIX C

RFLP analysis of GP20-2 population with probe B212

PLANT	GALL INDEX	GENOTYPE	PLANT	GALL INDEX	GENOTYPE
Gazelle	1	A	GP20-2-32	7	H
Prima	9	B	GP20-2-33	5	H
F1	-	H	GP20-2-34	3	A
GP20-2-1	1	H	GP20-2-35	9	B
GP20-2-2	9	B	GP20-2-36	1	A
GP20-2-3	9	B	GP20-2-37	7	B
GP20-2-4	3	H	GP20-2-38	7	H
GP20-2-5	9	H	GP20-2-39	5	H
GP20-2-6	5	H	GP20-2-40	3	A
GP20-2-7	9	H	GP20-2-41	5	B
GP20-2-8	9	B	GP20-2-42	1	A
GP20-2-9	9	B	GP20-2-43	5	H
GP20-2-10	9	H	GP20-2-44	9	B
GP20-2-12	1	A	GP20-2-45	3	A
GP20-2-13	9	B	GP20-2-46	5	H
GP20-2-14	5	H	GP20-2-47	7	B
GP20-2-15	1	A	GP20-2-48	9	B
GP20-2-16	9	H	GP20-2-49	3	H
GP20-2-17	7	H	GP20-2-50	9	B
GP20-2-18	1	H	GP20-2-51	5	H
GP20-2-19	1	A	GP20-2-52	5	H
GP20-2-20	5	H	GP20-2-53	1	A
GP20-2-21	3	H	GP20-2-54	1	A
GP20-2-22	1	A	GP20-2-76	1	A
GP20-2-23	7	H	GP20-2-78	3	H
GP20-2-24	3	H	GP20-2-79	3	H
GP20-2-25	1	A	GP20-2-80	7	H
GP20-2-26	7	H	GP20-2-81	7	H
GP20-2-27	7	H	GP20-2-82	5	H
GP20-2-28	7	H	GP20-2-83	7	B
GP20-2-29	3	A			
GP20-2-30	9	B			
GP20-2-31	5	H			

APPENDIX D

AFLP analysis of bulk samples

PRIMER COMBINATION	TOTAL NO OF FRAGMENTS	NO. OF POLYMORPHIC FRAGMENTS	
		GAZELLE AND PRIMA	INFORMATIVE IN 2 BULKS*
E-AAC M-CAA	98	6	1(C)
E-AAC M-CAC	85	9	2(R)+1(Q)
E-AAC M-CAG	52	4	1(C)
E-AAC M-CAT	92	10	1(R)
E-AAC M-CTA	76	9	3(Q)
E-AAC M-CTC	90	7	-
E-AAC M-CTG	59	6	-
E-AAC M-CTT	61	3	-
E-AAG M-CAA	86	8	-
E-AAG M-CAC	61	3	-
E-AAG M-CAG	64	5	-
E-AAG M-CAT	69	2	-
E-AAG M-CTA	70	9	1(C)+4(Q)
E-AAG M-CTC	58	7	-
E-AAG M-CTG	59	7	-
E-AAG M-CTT	93	4	1(C)
E-ACA M-CAA	67	6	-
E-ACA M-CAC	57	7	-
E-ACA M-CAG	56	8	1(C)
E-ACA M-CAT	75	9	2(C)+3(Q)
E-ACA M-CTA	56	2	-
E-ACA M-CTC	34	3	1(Q)
E-ACA M-CTG	40	5	-
E-ACA M-CTT	41	1	-
E-ACC M-CAA	82	10	1(Q)
E-ACC M-CAC	37	1	-
E-ACC M-CAG	61	4	-
E-ACC M-CAT	66	5	-
E-ACC M-CTA	41	3	-
E-ACC M-CTC	34	6	1(C)+1(Q)
E-ACC M-CTG	37	3	-
E-ACC M-CTT	62	7	1(C)+1(Q)
E-ACG M-CAA	53	4	-
E-ACG M-CAC	77	9	2(Q)
E-ACG M-CAG	20	3	-

PRIMER COMBINATION	TOTAL NO OF FRAGMENTS	NO. OF POLYMORPHIC FRAGMENTS	
		GAZELLE AND PRIMA	INFORMATIVE IN 2 BULKS*
E-ACG M-CAT	41	5	-
E-ACG M-CTA	73	12	3(Q)
E-ACG M-CTC	51	2	-
E-ACG M-CTG	73	2	1(C)
E-ACG M-CTT	58	7	-
E-ACT M-CAA	80	7	-
E-ACT M-CAC	21	4	-
E-ACT M-CAG	64	10	-
E-ACT M-CAT	70	13	-
E-ACT M-CTA	33	1	-
E-ACT M-CTC	24	1	-
E-ACT M-CTG	37	2	-
E-ACT M-CTT	46	4	3(Q)
E-AGC M-CAA	58	3	-
E-AGC M-CAC	51	7	-
E-AGC M-CAG	88	9	1(C)
E-AGC M-CAT	66	7	1(C)+1(Q)
E-AGC M-CTA	57	7	-
E-AGC M-CTC	61	8	-
E-AGC M-CTG	55	6	1(C)+1(R)
E-AGC M-CTT	56	3	-
E-AGG M-CAA	69	11	1(R)
E-AGG M-CAC	53	9	1(C)+2(Q)
E-AGG M-CAG	45	3	-
E-AGG M-CAT	69	7	-
E-AGG M-CTA	48	7	1(R)
E-AGG M-CTC	62	5	-
E-AGG M-CTG	46	2	-
E-AGG M-CTT	90	11	1(C)
TOTAL	3814	377	15(C)+6(R)+26(Q)=47

*C: Coupling phase; R: Repulsion phase; Q:Quantitative difference

APPENDIX E

AFLP data of individual plants from the 2 bulks

	POLYMORPHIC FRAGMENTS				PARENTS		INDIVIDUAL PLANTS									
							RESISTANT					SUSCEPTIBLE				
PRIMER	TOTAL	G[1]&P[2]	R[3]&S[4]	NO	G[1]	P[2]	12	15	19	22	25	2	3	8	9	13	R:S RATIO[5]
AAC-CAA	110	11	+	1	1	0	1	1	1	0	1	1	0	1	0	1	4:3
			-	2	0	1	1	0	0	0	1	1	0	1	0	0	2:2
			-	3	1	0	0	1	1	1	1	1	1	0	1	0	4:3
			-	4	1	0	0	0	1	1	1	0	0	1	1	0	3:2
AAC-CAC	76	8	-	1	1	0	1	1	1	1	1	1	1	0	1	0	5:3
			+	2	0	1	0	0	1	0	0	0	1	1	1	1	1:4*
			-	3	1	0	1	1	1	1	1	1	1	1	0	1	5:4
			-	4	0	1	1	1	1	1	1	1	1	1	1	1	5:5
			+	5	0	1	0	0	0	0	0	1	1	1	1	1	0:5*
			Q	6	0	1	1	0	1	1	1	1	0	0	1	1	4:3
			-	7	1	0	1	1	1	1	1	1	0	1	0	1	5:3
			?	8	1	0	0	1	0	0	0	1	0	1	1	1	1:4
AAC-CAT	82	14	+	1	0	1	0	0	0	0	0	1	1	1	1	1	0:5*
			-	2	0	1	1	1	0	0	1	1	1	1	1	1	3:5
			-	3	1	0	0	0	1	1	1	0	0	1	1	0	3:2
			-	4	1	0	1	1	0	0	1	1	1	1	1	0	3:4
			?	5	1	0	0	0	0	0	0	0	0	0	0	0	-
			-	6	0	1	0	1	0	1	0	1	1	0	1	0	2:3
			-	7	1	0	1	0	1	0	1	1	0	1	1	1	3:4
			-	8	0	1	1	1	1	1	1	1	1	1	1	1	5:5
			-	9	1	0	1	1	1	1	1	1	0	1	1	1	5:4
			-	10	1	0	1	1	0	1	0	0	1	1	1	1	3:4
			-	11	1	0	1	0	1	0	1	1	1	1	1	1	3:5
			-	12	1	0	0	0	1	1	1	0	0	1	1	0	3:2
			?	13	1	0	0	0	0	0	0	0	0	0	0	0	-
			?	14	0	1	1	0	1	1	1	1	1	0	0	1	4:3
AAG-CTA	66	14	-	1	0	1	1	1	0	1	-	1	1	0	0	-	3:2
			-	2	0	1	1	1	0	1	0	1	1	1	1	-	3:4
			-	3	0	1	0	1	1	1	-	1	1	0	0	1	3:3
			-	4	1	0	1	1	1	1	-	1	1	1	1	0	4:4
			+	5	1	0	1	1	1	1	1	1	0	0	0	0	5:1co*
			+	6	0	1	0	0	0	0	0	1	1	1	1	1	0:5co
			-	7	1	0	0	0	1	1	1	0	0	1	1	0	3:2
			-	8	1	0	1	1	1	1	0	1	0	1	0	0	4:2
			Q	11	0	1	1	1	0	1	0	1	1	1	0	1	3:4
			Q	12	1	0	1	1	1	1	1	1	0	1	1	1	5:4
			Q	13	0	1	1	0	-	1	1	1	1	0	0	-	3:2
			Q	14	1	0	0	0	1	1	1	0	0	-	1	-	3:1
ACA-CAG	42	6	-	1	0	1	1	1	1	0	1	0	0	0	1	1	4:2
			-	2	0	1	0	1	1	1	0	1	1	1	1	1	3:5
			-	3	1	0	1	1	1	1	1	1	1	1	1	1	5:5
			-	4	0	1	1	1	1	1	1	1	1	1	1	1	5:5
			-	5	0	1	1	0	1	1	0	0	1	1	1	0	3:3
			+	6	1	0	-	1	1	0	1	1	0	1	0	1	3:3
ACA-CAT	90	7	Q	1	0	1	1	1	0	1	0	1	1	1	1	1	3:5
			Q	2	1	0	1	1	0	1	1	1	0	0	1	1	4:3
			Q	3	0	1	1	1	0	1	0	1	1	1	1	1	3:5
			+	4	1	0	1	1	1	1	1	1	0	1	1	0	5:3
			+	5	1	0	1	1	1	1	1	0	1	1	1	1	5:4
			-	6	0	1	1	1	1	1	0	1	1	1	1	0	4:4
			-	7	1	0	0	1	0	1	1	1	0	0	1	1	3:3
AAC-CAG	40	5	-	1	0	1	1	1	1	1	-	1	1	1	0	-	4:3
			-	2	0	1	1	1	1	1	-	1	0	1	1	1	4:4
			-	3	1	0	1	1	1	1	-	1	1	0	1	-	4:3
			+	5	1	0	0	0	1	1	-	0	0	1	1	-	2:2
ACA-CTC	54	5	Q	1	1	0	0	0	1	1	-	0	0	1	1	1	2:3

	POLYMORPHIC FRAGMENTS				PARENTS		INDIVIDUAL PLANTS										
							RESISTANT					SUSCEPTIBLE					
PRIMER	TOTAL	G[1]&P[2]	R[3]&S[4]	NO	G[1]	P[2]	12	15	19	22	25	2	3	8	9	13	R:S RATIO[5]
			-	2	0	1	1	1	0	1	-	0	0	1	1	-	3:2
			-	3	1	0	1	1	1	1	-	1	0	1	1	-	4:3
			-	4	0	1	0	1	1	1	-	1	1	1	1	-	3:4
			?	5	0	1	1	0	0	1	-	1	1	1	1	-	2:4
ACG-CAC	62	7	Q	1	1	0	0	1	1	1	1	0	0	0	0	0	4:0*
			-	2	0	1	0	0	1	0	0	1	0	1	0	1	1:3
			-	3	0	1	1	0	1	1	0	1	1	1	1	1	3:5
			-	4	1	0	0	1	1	0	1	1	1	1	1	0	3:4
			-	5	0	1	1	1	1	1	1	0	1	1	1	1	5:4
			Q	6	1	0	0	0	1	1	1	0	0	1	1	0	3:2
			-	7	0	1	0	1	1	1	1	1	1	1	1	1	4:5
ACC-CAA	78	11	Q	1	1	0	1	1	1	0	1	1	1	0	1	0	4:3
			-	2	0	1	0	1	1	1	1	1	1	1	1	0	4:4
			-	3	1	0	1	0	1	1	1	1	1	1	1	0	co4:4
			-	4	0	1	0	1	1	0	1	1	1	1	0	1	co3:4
			-	5	1	0	1	1	1	1	0	1	0	1	0	1	4:3
			-	6	0	1	1	0	1	0	1	1	1	0	1	0	3:3
			-	7	1	0	1	1	1	1	1	1	1	1	1	1	5:5
			-	8	0	1	0	0	0	0	0	0	0	0	0	0	-
			-	9	0	1	1	1	1	0	0	1	1	1	1	1	3:5
			-	10	1	0	0	1	1	1	1	1	1	0	1	0	4:3
			-	11	1	0	1	1	1	1	1	1	0	1	1	1	5:4
ACC-CTT	62	7	-	1	1	0	1	0	1	0	1	1	1	1	1	1	3:5
			Q	2	1	0	1	1	0	1	0	1	0	0	0	1	3:2
			-	3	1	0	1	1	1	1	1	1	1	0	1	1	5:4
			-	4	0	1	0	1	1	1	1	1	1	1	1	1	4:5
			-	5	0	1	1	1	0	1	0	1	1	1	1	1	3:5
			-	6	0	1	0	1	0	1	1	1	1	1	1	1	3:5
			+	7	1	0	1	1	1	1	0	0	0	1	0	1	4:2
AGC-CAT	60	7	-	1	1	0	1	1	1	1	0	1	0	1	0	1	4:3
			+	2	1	0	0	0	1	1	1	0	0	1	1	0	3:2
			-	3	1	0	0	1	0	1	1	0	1	1	1	1	3:4
			Q	4	1	0	1	1	1	1	0	0	0	1	0	1	4:2
			-	5	0	1	1	1	1	1	0	0	1	1	1	1	4:4
			-	6	1	0	1	1	-	1	1	1	1	1	1	1	4:5
			-	7	0	1	1	1	0	1	1	0	1	1	1	0	4:3
ACG-CTA	68	9	-	1	0	1	1	1	0	1	0	1	1	0	0	1	3:3
			-	2	0	1	1	1	0	1	0	1	1	0	0	1	3:3
			-	3	0	1	1	1	0	1	1	0	1	1	1	1	4:4
			-	4	0	1	1	1	1	1	1	1	1	1	1	1	5:5
			Q	5	0	1	1	1	0	0	1	1	1	1	1	1	3:5
			-	6	0	1	1	1	1	1	1	1	1	1	1	1	5:5
			-	7	1	0	1	1	1	0	1	1	1	0	1	0	4:3
			Q	8	1	0	1	0	1	1	1	1	0	1	0	1	4:3
			Q	9	0	1	0	1	1	1	0	1	1	1	1	1	3:5
ACG-CTG	81	2	-	1	1	0	0	1	1	1	1	0	1	0	1	0	4:2
			+	2	1	0	1	1	1	1	0	0	0	1	1	1	4:3
AGG-CAA	71	7	+	1	0	1	1	1	1	0	0	1	1	1	0	1	3:4
			-	2	0	1	1	1	0	1	0	1	1	1	1	1	3:5
			-	3	1	0	1	1	0	1	0	1	0	0	0	1	3:2
			-	4	0	1	1	1	0	1	1	0	1	1	0	0	4:2
			-	5	0	1	1	1	1	1	1	1	1	1	1	0	5:4
			-	6	1	0	1	1	1	1	0	1	0	1	0	0	4:2
			-	7	0	1	0	1	1	1	1	0	1	1	0	1	4:3
AGG-CAC	64	8	?	1	1	0	1	0	1	0	0	0	1	0	1	1	2:3
			Q	2	0	1	0	0	1	1	1	1	1	1	1	1	3:5co
			+	3	1	0	1	1	0	0	0	0	0	0	0	0	2:0co

	POLYMORPHIC FRAGMENTS				PARENTS		INDIVIDUAL PLANTS										
							RESISTANT					SUSCEPTIBLE					
PRIMER	TOTAL	G[1]&P[2]	R[3]&S[4]	NO	G[1]	P[2]	12	15	19	22	25	2	3	8	9	13	R:S RATIO[5]
			Q	4	1	0	1	1	1	1	1	1	1	1	1	0	5:4
			-	5	1	0	0	1	1	1	1	1	1	0	1	0	4:3
			-	6	0	1	1	1	0	0	1	1	1	1	1	1	3:5
			-	7	1	0	1	1	1	1	1	1	1	1	1	1	5:5
			-	8	1	0	1	1	0	1	0	1	0	0	0	1	3:2
AGG-CTA	41	7	+	1	0	1	1	1	0	1	0	1	1	0	0	1	3:3
			-	2	0	1	1	1	0	1	1	1	1	1	0	1	4:4
			-	3	1	0	0	1	1	0	1	1	1	1	1	1	3:5
			-	4	0	1	1	1	1	1	1	0	1	1	1	1	5:4
			-	5	1	0	1	1	0	1	0	1	0	0	0	1	3:2
			-	6	1	0	1	1	0	0	0	1	0	0	0	0	2:1
			-	7	0	1	1	1	0	1	1	0	1	1	1	0	4:3
AGG-CTT	85	7	-	1	1	0	1	0	1	1	1	1	0	1	1	1	4:4
			+	2	1	0	1	0	0	1	0	1	1	1	1	0	2:4
			-	3	1	0	1	0	1	1	1	0	1	1	1	1	4:4
			-	4	1	0	1	0	1	1	1	0	1	1	1	1	4:4
			-	5	1	0	1	1	1	1	1	1	1	1	1	1	5:5
			-	6	0	1	0	1	1	1	1	1	1	1	1	1	4:5
			-	7	1	0	0	1	1	1	1	0	0	1	0	1	4:2
AAC-CTA	56	5	-	1	0	1	1	1	0	1	0	1	1	0	0	1	3:3
			Q	2	0	1	1	0	0	1	0	1	1	0	1	1	2:4
			-	3	1	0	0	0	1	1	1	0	0	1	1	0	3:2
			Q	4	1	0	1	0	1	1	0	1	1	0	1	1	3:4
			Q	5	1	0	1	0	1	0	0	1	1	1	1	0	2:4
AAG-CTT	93	4	-	1	0	1	0	1	1	1	1	1	1	1	-	-	4:3
			-	2	1	0	1	0	1	0	1	1	1	1	0	-	3:3
			-	3	0	1	1	1	1	1	1	1	1	1	1	-	5:4
			+	4	1	0	1	-	1	0	1	1	1	0	0	-	3:2
			-	5	1	0	1	0	1	1	1	0	1	1	1	-	4:3
			-	6	1	0	0	0	1	1	1	0	0	-	-	-	-
			-	7	1	0	0	-	1	1	1	0	1	1	1	-	3:3
ACT-CTT	46	4	-	1	1	0	1	1	1	1	1	1	1	1	1	1	-
			Q	2	0	1	1	1	1	1	1	1	1	1	1	1	-
			Q	3	0	1	1	1	1	0	0	0	1	1	1	1	3:4
			Q	4	0	1	1	1	1	1	1	1	1	1	1	1	-
AGC-CAG	88	9	-	1	0	1	1	-	1	1	1	1	1	-	1	1	-
			-	2	0	1	1	-	1	1	1	1	1	1	0	1	4:4
			-	3	0	1	1	-	1	1	1	1	1	1	1	1	4:5
			-	4	1	0	1	-	1	0	1	1	1	1	1	1	3:5
			+	5	1	0	1	1	1	0	1	1	0	1	0	1	4:3
			-	7	1	0	1	-	1	1	1	1	1	0	1	0	4:3
			-	8	1	0	1	1	1	1	1	1	1	1	1	1	-
			-	9	1	0	0	-	1	1	1	1	1	0	1	1	3:4
AGC-CTG	55	6	-	1	0	1	1	1	1	1	1	1	1	1	0	1	5:4
			-	2	0	1	1	1	1	1	1	1	1	1	1	1	-
			-	3	0	1	1	1	1	1	1	1	1	1	1	1	-
			+	5	0	1	0	0	0	0	0	1	1	1	1	1	0:5co
			+	6	1	0	1	1	1	1	1	1	0	0	0	1	5:2co
ACC-CTC	34	6	-	1	0	1	1	-	-	1	0	1	0	1	-	1	-
			+	2	1	0	1	1	1	1	1	1	0	0	0	0	5:1
			Q	3	1	0	1	1	0	1	1	0	1	1	0	0	4:2
			-	4	0	1	1	0	0	1	1	1	0	1	0	0	3:2
			-	5	1	0	1	1	1	1	1	1	1	1	1	1	-
			-	6	0	1	0	1	1	1	1	1	1	1	0	0	4:3
TOTAL	**1604**	**176**	**47**														

[1]:G : Gazelle; [2]:P : Prima; [3]:R : Resistant bulk; [4]:S : Susceptible bulk

[5]:R:S ratio: Ratio of fragments present in each bulk.

APPENDIX F

Statistical analysis of AFLP fragments

MARKER	BP	BULKS	LINKAGE**	GI(10 PLANTS)		GI(50 PLANTS)		log_{10}RF(10 PLANTS)		log_{10}RF(50 PLANTS)		Segregation	
				R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	χ_2	P
AACCAA1		+	C	0.00	0.545	ND	ND	25.97	0.075	ND	ND		
AACCAA2		-	R	0.00	1.00	ND	ND	5.60	0.251	ND	ND		
AACCAA3		-	C	0.00	0.545	ND	ND	13.95	0.155	ND	ND		
AACCAA4		-	C	0.00	0.580	ND	ND	5.75	0.248	ND	ND		
AACCAC1	577	-	C	15.62	0.141	0.00	0.660	0.00	0.957	4.64	0.214	3.77	0.05
AACCAC2	277	+	R	28.00	0.067	0.00	0.638	3.12	0.289	0.04	0.898	0.44	0.50
AACCAC3	271	-	C	0.00	0.347	3.22	0.116	**32.71**	**0.049***	1.95	0.343	1.00	0.32
AACCAC4	234	-	R	0.00	1.000	0.00	0.366	ND	ND	0.00	0.838	1.78	0.18
AACCAC5	**208**	**+**	**R**	**100.00**	**0.000****	**29.55**	**0.000****	16.99	0.130	4.81	0.073	0.44	0.50
AACCAC6	**156**	**K**	**R**	0.00	0.544	2.12	0.162	0.00	0.608	**7.66**	**0.032***	0.11	0.74
AACCAC7	151	-	C	15.62	0.141	0.00	0.354	21.73	0.098	1.90	0.173	0.00	1.00
AACCAC8	148	-	C	28.00	0.067	2.20	0.158	12.18	0.172	4.60	0.077	4.60	0.74
AACCAG1		-	R	0.00	0.356	ND	ND	28.45	0.099	ND	ND		
AACCAG2		-	R	0.00	0.407	ND	ND	0.00	0.821	ND	ND		
AACCAG3		-	C	0.00	0.356	ND	ND	0.00	0.769	ND	ND		
AACCAG5		+	C	0.00	1.000	ND	ND	8.14	0.250	ND	ND		
AACCAT15				ND	ND	ND	ND	ND	ND	0.00	0.550		
AACCAT1	**284**	**+**	**R**	**100.00**	**0.000****	**41.70**	**0.000****	16.99	0.130	**11.31**	**0.014***	0.72	0.39
AACCAT2	218	-	R	15.62	0.141	0.00	0.389	0.00	0.972	0.00	0.338	0.01	0.93
AACCAT3	216	-	C	0.00	0.579	0.00	0.690	5.75	0.249	0.00	0.997	0.44	0.50
AACCAT4	215	-	C	0.00	0.544	ND	ND	0.00	0.844	0.00	0.821		
AACCAT6	196	-	R	0.00	0.579	0.97	0.247	29.13	0.062	0.00	0.568	2.47	0.12
AACCAT7	195	-	C	0.00	0.544	4.89	0.091	0.00	0.601	0.00	0.786	4.30	0.04
AACCAT8	176	-	R	ND	ND	1.34	0.207	ND	ND	0.00	0.352	1.78	0.18
AACCAT9	172	-	C	0.00	0.346	0.00	0.739	0.00	0.780	0.29	0.291	0.55	0.46
AACCAT10	170	-	C	0.00	0.545	0.00	0.996	0.00	0.744	0.00	0.498	0.01	0.93
AACCAT11			C	15.62	0.141	ND	ND	0.00	0.699	0.00	0.998		
AACCAT12	158	-	C	0.00	0.579	0.00	0.495	5.75	0.249	0.00	0.568	0.24	0.62
AACCAT14	126	-	R	0.00	0.545	0.00	0.456	19.57	0.112	0.00	0.518	0.06	0.80
AACCTA1		-	R	0.00	1.000	ND	ND	0.00	0.376	ND	ND		
AACCTA2		K	R	6.25	0.241	ND	ND	0.00	0.877	ND	ND		
AACCTA3		-	C	0.00	0.579	ND	ND	5.75	0.248	ND	ND		
AACCTA4		K	C	0.00	0.545	ND	ND	0.00	0.930	ND	ND		
AACCTA5		K	C	6.25	0.241	ND	ND	0.00	0.903	ND	ND		
AAGCTA1	755	-	R	0.00	0.537	3.92	0.099	0.86	0.343	0.27	0.295	1.42	0.23
AAGCTA2	694	-	R	11.83	0.193	0.00	0.484	0.00	0.549	0.00	0.732	2.35	0.13
AAGCTA3	478	-	R	0.00	0.684	0.10	0.312	0.00	0.966	0.00	0.947	2.05	0.15
AAGCTA4	475	-	C	0.00	0.407	0.00	0.424	0.00	0.713	0.00	0.495	0.18	0.67
AAGCTA5	**409**	**+**	**Cco**	**62.50**	**0.004****	**22.99**	**0.0003****	11.97	0.174	0.00	0.704	0.11	0.74
AAGCTA6	**405**	**+**	**Rco**	**100.00**	**0.000****	**29.74**	**0.000****	16.99	0.130	**7.98**	**0.029***	0.44	0.50

MARKER	BP	BULKS	LINKAGE**	GI(10 PLANTS)		GI(50 PLANTS)		log10RF(10 PLANTS)		log10RF(50 PLANTS)		Segregation	
				R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	χ^2	P
AAGCTA7	394	-	C	0.00	0.579	0.00	0.755	5.75	0.249	0.00	0.354	0.35	0.56
AAGCTA8	387	-	C	6.25	0.241	0.00	0.323	0.00	0.413	0.00	0.586	0.11	0.74
AAGCTA10	198	-		ND	ND	ND	ND	ND	ND	ND	ND		
AAGCTA11	163	K	R	0.00	0.544	ND	ND	0.00	0.511	ND	ND		
AAGCTA12	162	K	C	0.00	0.346	ND	ND	0.00	0.780	ND	ND		
AAGCTA13	131	K	R	0.00	0.537	ND	ND	13.19	0.201	ND	ND		
AAGCTA14	130	K	C	0.00	0.537	ND	ND	0.00	0.376	ND	ND		
AAGCTA15		-		ND	ND	ND	ND	ND	ND	ND	ND		
AAGCTT1		-	R	0.00	0.481	0.00	0.926	**83.14**	**0.001****	0.00	0.497	1.72	0.19
AAGCTT2		-	C	0.00	0.684	2.26	0.179	17.39	0.145	0.00	0.527	3.77	0.05
AAGCTT3		-	R	ND	ND	0.00	0.586	ND	ND	0.00	0.623	1.03	0.31
AAGCTT4		+	C	0.00	0.537	0.00	0.696	30.25	0.091	2.76	0.155	6.53	0.01
AAGCTT5		-	C	0.00	0.878	ND	ND	0.00	0.856	ND	ND		
AAGCTT6		-	C	16.00	0.203	ND	ND	0.00	0.725	ND	ND		
AAGCTT7		-	C	0.00	1.000	0.00	0.873	18.16	0.161	0.00	0.839	0.79	0.37
AAGCTT8		-		ND	ND	4.38	0.138	ND	ND	**15.42**	**0.018***		
ACACAG1		-	R	6.25	0.241	ND	ND	0.00	0.386				
ACACAG2		-	R	15.63	0.141	ND	ND	**38.69**	**0.032***				
ACACAG5		-	R	0.00	1.000	ND	ND	0.00	0.815				
ACACAG6		+	C	0.00	0.606	ND	ND	**70.57**	**0.006****				
ACACAT1		K	R	15.62	0.141	ND	ND	0.00	0.570	ND	ND		
ACACAT2		K	C	0.00	0.544	ND	ND	0.00	0.838	ND	ND		
ACACAT3		K	R	15.62	0.141	ND	ND	0.00	0.570	ND	ND		
ACACAT4		+	C	15.62	0.141	ND	ND	0.00	0.959	ND	ND		
ACACAT5		+	C	0.00	0.346	ND	ND	0.00	0.801	ND	ND		
ACACAT6		-	R	0.00	1.000	ND	ND	0.00	0.591	ND	ND		
ACACAT7		-	C	0.00	1.000	ND	ND	5.51	0.252	ND	ND		
ACACTC1		K	C	0.00	0.789	ND	ND	2.88	0.303	ND	ND		
ACACTC2		-	R	0.00	0.537	ND	ND	0.00	0.941	ND	ND		
ACACTC3		-	C	0.00	0.356	ND	ND	0.00	0.866	ND	ND		
ACACTC4		-	R	0.00	0.356	ND	ND	**63.00**	**0.011***	ND	ND		
ACACTC5		-	R	22.22	0.134	ND	ND	0.00	0.742	ND	ND		
ACCCAA1		K	C	0.00	0.544	ND	ND	0.00	0.740	ND	ND		
ACCCAA2		-	R	0.00	1.000	ND	ND	**38.01**	**0.034***	ND	ND		
ACCCAA3		-	Cco	0.00	1.000	ND	ND	0.00	0.841	ND	ND		
ACCCAA4		-	Rco	0.00	0.545	ND	ND	0.00	0.964	ND	ND		
ACCCAA5		-	C	0.00	0.545	ND	ND	5.02	0.259	ND	ND		
ACCCAA6		-	R	0.00	1.000	ND	ND	0.00	0.719	ND	ND		
ACCCAA9		-	R	15.62	0.141	ND	ND	0.00	0.958	ND	ND		

MARKER	BP	BULKS	LINKAGE**	GI(10 PLANTS)		GI(50 PLANTS)		log10RF(10 PLANTS)		log10RF(50 PLANTS)		Segregation	
				R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	χ^2	P
ACCCAA10		-	C	0.00	0.545	ND	ND	13.59	0.155	ND	ND		
ACCCAA11		-	C	0.00	0.346	ND	ND	0.00	0.780	ND	ND		
ACCCTC1		-	R	0.00	0.846	0.00	0.799	0.00	0.793	ND	ND	0.88	0.35
ACCCTC2	**281**	**+**	**C**	**62.5**	**0.004****	**24.52**	**0.0003****	11.97	0.174	0.00	0.920	0.26	0.61
ACCCTC3		K	C	6.25	0.242	0.00	0.669	0.00	0.432	1.42	0.206	1.42	0.23
ACCCTC4		-	R	0.00	0.579	0.00	0.626	0.00	0.403	0.00	0.907	0.90	0.34
ACCCTC6		-	R	0.00	0.545	ND	ND	0.00	0.639	ND	ND		
ACCCTT1		-	C	15.62	0.141	ND	ND	0.00	0.699	ND	ND		
ACCCTT2		K	C	0.00	0.579	ND	ND	2.81	0.294	ND	ND		
ACCCTT3		-	C	0.00	0.346	ND	ND	0.00	0.673	ND	ND		
ACCCTT4		-	R	0.00	0.346	ND	ND	**53.66**	**0.009****	ND	ND		
ACCCTT5		-	R	15.62	0.141	ND	ND	0.00	0.570	ND	ND		
ACCCTT6		-	R	15.62	0.141	ND	ND	**39.62**	**0.030***	ND	ND		
ACCCTT7		+	C	6.25	0.241	ND	ND	7.85	0.220	ND	ND		
ACGCAC1	449	K	C	**62.5**	**0.004****			0.00	0.884				
ACGCAC2	436	-	R	6.25	0.241			0.00	0.969				
ACGCAC3	404	-	R	15.63	0.141	0.48	0.272	0.00	0.831	0.00	0.804	2.16	0.14
ACGCAC4	233	-	C	0.00	0.545	0.00	0.496	13.59	0.159	2.20	0.154	0.24	0.62
ACGCAC5	232	-	R	0.00	0.347	0.00	0.619	0.00	0.801	0.00	0.643	0.67	0.41
ACGCAC6	139	K	C	0.00	0.579	0.00	0.789	5.75	0.248	0.00	0.995	0.24	0.62
ACGCAC7	132	-	R	0.00	0.346	0.00	0.498	**53.66**	**0.009****	0.00	0.778	0.33	0.56
ACGCTA1		-	R	0.00	1.000	ND	ND	0.00	0.376	ND	ND		
ACGCTA2		-	R	0.00	1.000	ND	ND	0.00	0.376	ND	ND		
ACGCTA3		-	R	0.00	1.000	ND	ND	0.00	0.919	ND	ND		
ACGCTA5		K	R	15.62	0.141	ND	ND	0.00	0.972	ND	ND		
ACGCTA7		-	C	0.00	0.545	ND	ND	0.00	0.740	ND	ND		
ACGCTA8		K	Cco	0.00	0.545	ND	ND	15.67	0.141	ND	ND		
ACGCTA9		K	Rco	15.62	0.141	ND	ND	**38.69**	**0.032***	ND	ND		
ACGCTG1		-	C	6.25	0.241	ND	ND	5.27	0.255	ND	ND		
ACGCTG2		+	C	0.00	0.545	ND	ND	0.00	0.913	ND	ND		
ACTCTT3		K	R	0.00	0.545	ND	ND	0.00	0.833	ND	ND		
AGCCAG2		-	R	0.00	0.467	0.00	0.643	32.72	0.063	0.27	0.294	1.20	0.27
AGCCAG4		-	C	3.57	0.292	ND	ND	0.00	0.677	ND	ND		
AGCCAG5	313	+	C	0.00	0.545	0.00	0.340	25.97	0.076	1.82	0.178	5.94	0.01
AGCCAG7		-	C	11.84	0.193	ND	ND	0.00	0.949	ND	ND		
AGCCAG9		-	C	0.00	0.878	ND	ND	10.15	0.210	ND	ND		
AGCCAT1		-	C	0.00	0.545	ND	ND	5.02	0.259	ND	ND		
AGCCAT2		+	C	0.00	0.579	ND	ND	5.75	0.248	ND	ND		
AGCCAT3		-	C	0.00	0.545	ND	ND	19.59	0.111	ND	ND		

MARKER	BP	BULKS	LINKAGE**	GI(10 PLANTS)		GI(50 PLANTS)		log10RF(10 PLANTS)		log10RF(50 PLANTS)		Segregation	
				R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	R_2(%)*	P	χ_2	P
AGCCAT4		K	C	6.25	0.241	ND	ND	7.85	0.220	ND	ND		
AGCCAT5		-	R	0.00	1.000	ND	ND	0.00	0.933	ND	ND		
AGCCAT7		-	R	0.00	0.545	ND	ND	0.00	0.852	ND	ND		
AGCCTG1	391	-	R	0.00	0.347	ND	ND	**32.70**	**0.049***	ND	ND		
AGCCTG5	**132**	**+**	**Rco**	**100.00**	**0.000****	**29.68**	**0.000****	16.99	0.130	**5.98**	**0.048***	0.67	0.41
AGCCTG6	**131**	**+**	**Cco**	**35.71**	**0.040***	**20.72**	**0.0005****	25.53	0.078	0.00	0.722	1.42	0.23
AGGCAA1		+	R	0.00	0.545	ND	ND	8.45	0.213	ND	ND		
AGGCAA2		-	R	15.62	0.141	ND	ND	0.00	0.570	ND	ND		
AGGCAA3		-	C	0.00	0.579	ND	ND	2.81	0.294	ND	ND		
AGGCAA4		-	R	6.25	0.241	ND	ND	0.00	0.432	ND	ND		
AGGCAA5		-	R	0.00	0.346	ND	ND	0.00	0.727	ND	ND		
AGGCAA6		-	C	6.25	0.241	ND	ND	0.00	0.413	ND	ND		
AGGCAA7		-	R	0.00	0.545	ND	ND	0.00	0.941	ND	ND		
AGGCAC1		-	C	0.00	0.579	ND	ND	0.00	0.620	ND	ND		
AGGCAC2		K	Rco	15.62	0.141	ND	ND	21.96	0.097	ND	ND		
AGGCAC3		+	Cco	15.62	0.141	ND	ND	21.96	0.097	ND	ND		
AGGCAC4		K	C	0.00	0.346	ND	ND	0.00	0.727	ND	ND		
AGGCAC5		-	C	0.00	0.545	ND	ND	13.95	0.155	ND	ND		
AGGCAC6		-	R	15.62	0.141	ND	ND	0.00	0.972	ND	ND		
AGGCAC8		-	C	0.00	0.579	ND	ND	2.81	0.294	ND	ND		
AGGCTA1		+	R	0.00	1.000	ND	ND	0.00	0.376	ND	ND		
AGGCTA2		-	R	0.00	1.000	ND	ND	2.89	0.293	ND	ND		
AGGCTA3		-	C	15.62	0.141	ND	ND	10.97	0.184	ND	ND		
AGGCTA4		-	R	0.00	0.346	ND	ND	0.00	0.801	ND	ND		
AGGCTA5		-	C	0.00	0.579	ND	ND	2.81	0.294	ND	ND		
AGGCTA6		-	C	0.00	0.545	ND	ND	7.64	0.223	ND	ND		
AGGCTA7		-	R	0.00	0.545	ND	ND	0.00	0.752	ND	ND		
AGGCTT1		-	C	0.00	1.000	ND	ND	0.00	0.788	ND	ND		
AGGCTT2		+	C	6.25	0.241	ND	ND	0.00	0.528	ND	ND		
AGGCTT3		-	C	0.00	1.000	ND	ND	0.00	0.804	ND	ND		
AGGCTT4		-	C	0.00	1.000	ND	ND	0.00	0.804	ND	ND		
AGGCTT6		-	R	0.00	0.346	ND	ND	**53.66**	**0.009****	ND	ND		
AGGCTT7		-	C	6.25	0.241	ND	ND	0.00	0.919	ND	ND		

*R^2 adjusted for degrees of freedom.

**C: Coupling phase; R: Repulsion phase

APPENDIX G

Complete RFLP data for 50 F2 progeny

	G	P	F1	12	15	19	22	25	2	3	8	9	13	1	4	5	6	7	10	14	16	17	18	20	21	23	24	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
GI	1	9		1	1	1	1	1	9	9	9	9	9	1	3	9	5	9	9	5	9	7	1	5	3	7	3	7	7	7	3	9	5	7	5	3	9	1	7	7	5	3	5	1	5	9	3	5	7	9	3	9	5
B212	A	B	H	A	A	A	A	A	B	B	B	B	B	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	A	B	H	H	H	A	B	A	B	H	H	A	B	A	H	B	A	H	B	B	H	B	H
A063	A	B	H	A	H	B	A	B	B	A	H	A	H	H	H	H	B	H	H	A	H	A	H	A	H	B	H	H	B	B	A	A	B	H	H	H	A	H	H	H	B	B	H	A	H	H	A	A	A	B	H	A	H
K007	A	B	H	A	A	H	H	A	A	A	B	H	H	A	H	B	H	A	H	H	A	B	H	H	H	H	A	H	H	A	H	A	H	B	A	H	A	H	H	A	H	H	H	H	H	B	A	H	A	H	B	H	A
A006	A	B	H	B	A	H	H	H	B	A	A	H	H	A	A	H	B	H	B	A	B	H	H	H	H	B	B	H	H	B	A	B	H	H	H	A	H	H	H	B	A	B	H	H	B	H	B	H	A	B	B	B	H
A808	A	B	H	B	B	D	D	D	B	B	D	D	D	D	D	D	D	D	D	D	D	D	D	B	D	B	D	D	B	B	D	B	D	D	D	D	B	D	B	D	D	D	D	B	D	D	D	B	D	B	B	D
A946	A	B	H	H	B	A			A	H	B	A	H	B	B	A	A	B	B	H	B	A	A	B				B	H	A	H		H	H	H	H	H	B	B	A				B		H		H	H		H	H	B
B142-1	A	B	H	H	B	A	H	H	B	H	B	H	B	H	H	H	H	B	A	H	H	A	A	B	H	B	A	B	B	A	B	H	H	B	H	A	B	H	H	H	B	H	B	A	B	H	H	H	A	H	H	B	A
B142-2	A	B	H	B	H	H	B	H	A	H	H	B	B	B	H	H	H	B	A	H	A	B	A	B	H	H	H	H	B	B	H	H	A	B	A	B	H	H	A	H	H	H	B	B	B	A	H	H	H	H	A	H	H
A878	A	B	H	A	H	H	H	A	B	H	H	H	H	A	A	H	B	H	H	A	H	H	B	H	A	H	B	A	H	B	A	A	A	A	A	H	H	H	H	A	H	B	B	H	A	B	B	B	B	H	B	H	H
A685	A	B	H	H	B	H	H	H	A	B	A	B	B	A	H	A	A	A	H	H	B	H	H	B	H	H	H	B	B	H	B	H	H	H	H	H	B	H	B	H	B	H	B	H	B	H	B	H	H	H	H	B	H
K411	A	B	H	B	H	H	B	H	A	H	H	B	B	B	H	H	H	B	A	H	A	B	A	B	H	H	H	H	B	B	H	H	A	B	A	B	H	H	A	H	H	H	B	B	B	A	H	H	H	H	A	H	H
K258	A	B	H	A	H	H	H	H	H	B	H	B	H	B	H	B	A	H	H	H	H	H	B	H		H	B	H	A	H	H	H	H	H	H	H	H	B	A	H	A	B	B	A	H	H	H	A	H	H	B	B	H	B
K636	A	B	H	H	A	H	B	A	A	A	A	A	A	B	H	H	A	A	A	B	B	H	B	B	A	B	B	H	A	H	H	B	B	H	B	B	A	H	H	H	B	B	B	H	B	H	H	H	H	H	A	B	H
K647	A	B	H	H	A	H			H	H	B	A	H	H	B	H	A	B	A	B	H	B	A	H								H										H	A	B		H	H		H	A		H	A	B
K494	A	B	H	B	H	H	B	H	A	H	H	B	B	B	H	H	H	B	A	H	A	B	A	B	H	H	H	H	B	B	H	H	A	B	H	B	H	B	H	H	A	H	B	H	B	B	B	H	H	H	H	H	H
K387	A	B	H	H	H	H	B	A	H	H	B	H	H	A	H	H	A	H	A	H	A	B	B	B	H	H	B	H	B	H	B	A	A	H	H	H	A	A	H	B	A	H	H	H	B	B	B	A	H	H	A	H	B
B122	A	B	H	H	B	H	A	B	H	H	H	H	B	H	A	H	B	B	B	H	B	H	B	B	A	A	B	H	A	H	H	B	A	H	H	B	A	B	H	H	B	H	B	H	B	H	H	H	A	A	B	A	H

Complete AFLP data for 50 F2 progeny

	G	P	F1	12	15	19	22	25	2	3	8	9	13	1	4	5	6	7	10	14	16	17	18	20	21	23	24	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
AACCAC1	1	0	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0	-	-	0	1	1	1	1	1	1	0	0	1	1	1	0	1	0	1	0	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AACCAC1				D	D	D	D	D	D	D	B	D	B	B	B	D	B	-	-	B	D	D	D	D	D	D	B	B	D	D	D	B	D	B	D	B	B	B	B	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AACCAC2	0	1	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	-	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	-	-	-	-	-	-
AACCAC2				A	A	C	A	A	A	C	C	C	C	C	C	C	C	-	A	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	A	C	C	C	C	C	A	C	C	A	C	C	C	-	-	-	-	-	-
AACCAC3	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	-	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	-
AACCAC3				D	D	D	D	D	D	D	D	B	D	D	D	D	D	B	-	D	B	D	B	D	D	D	D	D	B	D	B	D	D	D	D	D	B	D	D	B	D	D	B	D	D	D	D	D	D	D	D	D	-
AACCAC4	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	-	0	1	1	1	1	0	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	-
AACCAC4				C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	-	A	C	C	C	C	A	C	C	C	C	A	A	C	C	C	A	C	C	C	C	C	A	C	C	C	A	C	C	C	C	C	C	C	-
AACCAC5	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	-	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	1	1	1	-
AACCAC5				A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	-	C	C	A	C	C	C	C	C	C	C	C	A	C	C	C	C	A	C	A	C	A	C	A	C	A	C	C	A	C	C	C	C	C	-
AACCAC6	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0	1	1	-	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	-
AACCAC6				C	A	C	C	C	C	A	A	C	C	A	A	C	C	-	C	A	C	C	C	C	C	C	C	C	C	C	A	C	C	C	A	A	C	A	C	C	A	C	C	C	C	C	C	C	A	C	C	C	-
AACCAC7	1	0	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	-	0	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	1	0	0	1	-
AACCAC7				D	D	D	D	D	D	B	D	B	D	B	D	B	D	-	B	D	D	D	B	D	D	D	B	B	D	D	D	D	D	D	D	D	B	D	B	D	B	B	D	D	D	D	D	D	D	B	B	D	-
AACCAC8	1	0	1	0	1	0	0	0	1	0	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	-
AACCAC8				B	D	B	B	B	D	B	D	D	D	D	D	D	D	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	B	B	D	B	D	D	D	B	D	D	D	D	D	D	D	D	D	D	-
AACCAT1	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	-	-	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	-	-	1	0	1	1	1	0	1	0	0	1	0	1	-	1	1	1	1
AACCAT1				A	A	A	A	A	C	C	C	C	C	C	C	C	C	-	-	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	-	-	C	A	C	C	C	A	C	A	A	C	A	C	-	C	C	C	C

	G	P	F1	12	15	19	22	25	2	3	8	9	13	1	4	5	6	7	10	14	16	17	18	20	21	23	24	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
CAT10	1	0	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	-	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
CAT10				D	D	B	D	B	B	D	D	D	D	D	D	D	B	D	-	D	D	D	D	D	D	B	D	D	B	B	D	D	B	D	D	D	D	D	D	D	D	B	B	D	D	D	D	D	D	D	B	D	D	D
CAT12	1	0	1	0	0	1	1	1	0	0	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	1	0	1	1	0	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1
CAT12				B	B	D	D	D	B	B	D	D	B	D	D	B	D	D	D	B	D	D	D	D	B	D	B	D	D	B	B	D	B	D	D	D	D	B	D	D	D	D	D	D	D	D	D	D	B	D	B	D	D
CAT14	0	1	1	1	0	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	1	1	1
CAT14				C	A	C	C	C	C	C	A	A	C	A	C	C	C	C	C	C	C	A	C	C	C	C	A	C	A	C	C	C	C	C	A	C	C	C	A	C	C	C	A	A	C	C	C	A	C	C	C
CAT15	0	1	1	-	-	-	-	-	-	-	-	-	-	1	1	1	1	-	-	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	-	1	0	0	1
CAT15				-	-	-	-	-	-	-	-	-	-	C	C	C	C	-	-	A	C	C	C	A	C	C	C	C	A	C	C	A	C	C	C	C	C	C	C	A	C	C	A	C	C	C	A	C	-	C	A	A	C
CAT2	0	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	0	1	1	1	0	1	0	0
CAT2				C	C	A	A	C	C	C	C	C	C	A	A	C	C	C	C	C	C	C	C	A	C	A	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	A	A	A	C	A	C	C	C	A	C	A	A
CAT3	1	0	1	0	0	1	1	1	0	0	1	1	0	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	-	1	1	1	-	1	0	1	1
CAT3				B	B	D	D	D	B	B	D	D	B	D	D	B	D	B	D	B	D	D	D	D	B	D	B	D	D	B	B	D	D	D	D	D	D	D	B	D	D	D	-	D	D	D	-	D	B	D	D
CAT4	1	0	1	1	1	0	0	1	1	1	1	1	0	1	0	0	0	1	0	1	1	1	1	0	1	1	0	1	0	1	0	0	1	1	1	1	0	1	1	0	1	0	0	0	-	0	1	0	-	1	1	0	1
CAT4				D	D	B	B	D	D	D	D	D	B	D	B	B	B	B	D	B	D	D	D	B	D	D	B	D	B	B	D	D	D	D	B	D	D	B	D	D	B	D	B	B	B	-	B	D	B	-	D	D	B	D
CAT6	0	1	1	0	1	0	1	0	1	1	0	1	0	1	1	1	0	1	-	-	1	1	0	1	0	1	1	1	1	-	-	-	0	0	0	-	-	-	1	0	0	0	0	1	1	1	1	-	1	-	0	0	0	0
CAT6				A	C	A	C	A	C	C	A	C	A	C	C	A	C	-	-	C	C	A	C	A	C	C	C	C	-	-	-	A	A	A	-	-	-	C	A	A	A	A	C	C	C	C	-	C	-	A	A	A	A
CAT7	1	0	1	1	0	1	0	1	1	0	1	1	1	0	1	1	1	-	-	1	1	1	0	1	0	0	0	0	-	-	-	1	1	1	-	-	-	1	1	1	1	1	1	0	0	1	0	-	0	-	1	1	1	1
CAT7				D	B	D	B	D	D	B	D	D	D	B	D	D	D	-	-	D	D	D	B	D	B	B	B	B	-	-	-	D	D	D	-	-	-	D	D	D	D	D	B	B	D	B	-	B	-	D	D	D	D
CAT8	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	-	-	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	0	0
CAT8				C	C	C	C	C	C	C	C	C	C	C	A	C	C	-	-	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	A	C	A	C	C	C	C	A	C	A	C	A	C	A	C	A	A
CAT9	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	-	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
CAT9				D	D	D	D	D	D	B	D	D	D	B	D	B	B	B	D	-	D	D	D	D	D	B	D	D	D	D	D	D	D	D	B	D	B	D	D	D	D	D	D	B	D	D	D	D	D	D	D	D	D
CTA1	0	1	1	1	1	0	1	-	1	1	0	0	-	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	-	0	1	1	-	0	0	0	0
CTA1				C	C	A	C	-	C	C	A	A	-	C	A	C	C	A	C	C	C	C	C	C	C	C	A	A	C	C	A	C	C	C	C	C	C	C	C	A	A	C	C	-	A	C	C	-	A	A	A	A
CTA2	0	1	1	1	1	0	1	0	1	1	1	1	-	1	1	0	1	0	0	0	0	1	1	1	0	1	1	1	0	0	1	1	1	0	1	0	-	1	1	1	1	1	0	0	-	1	1	0	-	0	0	0	0
CTA2				C	C	A	C	A	C	C	C	C	-	C	C	A	C	A	A	A	A	A	C	C	C	A	C	C	C	A	A	C	C	C	A	C	A	-	C	C	C	C	C	A	A	-	C	C	A	-	A	A	A	A
CTA3	0	1	1	0	1	1	1	-	1	1	0	0	1	1	1	1	1	1	1	0	0	1	0	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	1	-	1	1	1	-	0	1	1	1
CTA3				A	C	C	C	-	C	C	A	A	C	C	C	C	C	C	C	A	A	C	A	C	C	A	A	C	C	C	C	C	A	A	A	C	A	A	A	C	C	C	C	A	C	-	C	C	C	-	A	C	C	C
CTA4	1	0	1	1	1	1	1	-	1	1	1	1	0	0	1	0	1	0	1	1	0	1	1	0	1	1	1	1	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	-	0	0	1	-	1	0	1	1
CTA4				D	D	D	D	-	D	D	D	D	B	B	D	B	D	B	D	D	B	D	D	B	D	D	D	D	B	B	B	D	D	D	B	D	D	D	D	D	D	D	D	D	-	B	B	D	-	D	B	D	D
CTA5	1	0	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	0	1	0	1	0	1	0	1	0	1	-	0	1	1	-	1	1	0	1
CTA5				D	D	D	D	D	D	B	B	B	B	D	D	D	D	D	D	D	D	D	D	D	D	B	D	D	B	D	D	B	D	D	B	D	B	D	B	D	B	D	B	D	-	B	D	D	-	D	D	B	D
CTA6	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	-	1	0	1	-	1	1	1	1
CTA6				A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	A	C	C	C	C	A	C	A	C	A	C	A	C	A	-	C	A	C	-	C	C	C	C
CTA7	1	0	1	0	0	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	-	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	-	1	1	1	-	1	0	1	1
CTA7				B	B	D	D	D	B	B	D	D	B	D	D	D	D	B	D	D	D	D	D	D	D	D	-	D	D	B	B	D	D	D	D	D	D	B	D	D	D	D	D	D	-	D	D	D	-	D	B	D	D
CTA8	1	0	1	1	1	1	1	0	1	0	1	0	0	0	1	1	1	1	0	1	0	1	0	1	1	0	-	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	-	1	1	1	-	1	1	0	1
CTA8				D	D	D	D	B	D	B	D	B	B	B	D	D	D	D	D	B	D	B	D	B	D	D	B	-	D	B	D	D	B	D	D	D	D	D	D	D	D	D	D	B	B	-	D	D	D	-	D	D	B	D
CTT1	0	1	1	0	1	1	1	1	1	1	1	-	-	1	1	1	0	-	-	1	1	1	1	0	1	1	-	1	0	0	1	1	-	-	1	1	-	1	1	1	0	1	1	1	-	1	1	1	-	1	1	-	-
CTT1				A	C	C	C	C	C	C	C	-	-	C	C	C	A	-	-	C	C	C	C	A	C	C	-	C	A	A	C	C	-	-	C	C	-	C	C	C	A	C	C	C	-	C	C	C	-	C	C	-	-
CTT2	1	0	1	1	0	1	0	1	1	1	1	0	-	0	1	1	0	-	-	1	1	0	1	1	0	0	-	1	0	1	1	1	-	-	-	1	-	0	0	1	1	1	1	0	1	1	0	1	-	1	1	-	-

	G	P	F1	12	15	19	22	25	2	3	8	9	13	1	4	5	6	7	10	14	16	17	18	20	21	23	24	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
GCTT2				D	B	D	B	D	D	D	D	B	-	B	D	D	B	-	-	D	D	B	D	D	B	B	-	D	B	D	D	D	-	-	-	D	-	B	B	D	D	D	D	B	D	D	B	D	-	D	D	-	-
GCTT3	0	1	1	1	1	1	1	1	1	1	1	1	-	0	1	0	1	-	-	1	1	1	1	0	0	1	-	1	0	1	1	1	-	-	-	1	-	0	1	1	1	1	1	1	0	1	1	1	-	1	1	-	-
GCTT3				C	C	C	C	C	C	C	C	C	-	A	C	A	C	-	-	C	C	C	C	A	A	C	-	C	A	C	C	C	-	-	-	C	-	A	C	C	C	C	C	C	A	C	C	C	-	C	C	-	-
GCTT4	1	0	1	1	-	1	0	1	1	1	0	0	-	1	0	1	0	-	-	1	0	1	1	0	0	1	-	0	0	1	0	0	-	0	0	1	-	0	1	1	1	1	1	1	1	1	0	1	-	1	1	-	-
GCTT4				D	-	D	B	D	D	D	B	B	-	D	B	D	B	-	-	D	B	D	D	B	B	D	-	B	B	D	B	B	-	B	B	D	-	B	D	D	D	D	D	D	D	D	B	D	-	D	D	-	-
GCTT7	1	0	1	0	-	1	1	1	0	1	1	1	-	1	1	0	1	-	-	1	0	1	0	1	1	1	1	1	-	1	1	0	-	0	1	1	-	0	1	0	0	-	0	0	1	1	0	1	-	1	0	-	-
GCTT7				B	-	D	D	D	B	D	D	D	-	D	D	B	D	-	-	D	B	D	B	D	D	D	D	D	-	D	D	B	-	B	D	D	-	B	D	B	B	-	B	B	D	D	B	D	-	D	B	-	-
GCTC1	0	1	1	1	-	-	1	0	1	0	1	-	1	1	1	1	1	-	-	0	1	1	1	0	1	1	1	1	0	1	1	-	1	1	1	1	1	1	1	0	1	1	0	1	-	1	-	1	-	1	-	-	-
GCTC1				C	-	-	C	A	C	A	C	-	C	C	C	C	C	-	-	A	C	C	C	A	C	C	C	C	A	C	C	-	C	C	C	C	C	C	C	A	C	C	A	C	-	C	-	C	-	C	-	-	-
GCTC2	1	0	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	0	1	1	0	1	1	-	1	-	-	-
GCTC2				D	D	D	D	D	D	B	B	B	B	D	D	D	D	D	-	D	D	D	D	D	D	D	D	D	D	D	D	B	D	D	D	D	B	D	B	D	D	D	B	D	D	B	D	D	-	D	-	-	-
GCTC3	1	0	1	1	1	0	1	1	1	0	0	1	1	0	1	0	1	0	-	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	-	1	-	0	-
GCTC3				D	D	B	D	D	B	D	D	B	B	D	D	B	D	B	-	D	D	D	D	B	B	D	B	B	D	D	D	D	D	D	D	D	B	D	D	B	D	B	B	D	D	D	D	-	D	-	B	-
GCTC4	0	1	1	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	-	0	1	0	0	1	1	1	0	1	1	0	0	1	1	1	0	1	-	0	1	1	1	1	1	1	1	1	1	0	-	0	-	1	-
GCTC4				C	A	A	C	C	C	A	C	A	A	C	C	C	C	C	-	A	C	A	A	C	C	C	A	C	C	A	A	C	C	C	A	C	-	A	C	C	C	C	C	C	C	C	C	A	-	A	-	C	-
GCTC5	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	-	1	-	1	-
GCTC5				D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	-	D	D	D	D	D	D	D	D	D	D	D	-	D	-	D	-
GCAC3	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	0	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1	1
GCAC3				C	A	C	C	A	C	C	C	C	C	C	C	C	C	C	A	C	C	A	C	A	C	A	A	C	C	C	A	C	C	C	C	C	C	C	A	A	A	C	A	A	C	A	C	C	A	C	C	C	C	C
GCAC4	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1
GCAC4				B	D	D	B	D	D	D	D	D	B	D	D	D	D	B	D	B	D	B	D	D	D	D	B	D	B	D	D	D	D	B	D	B	D	B	D	D	D	D	B	D	B	B	D	D	D	D	B	B	D	D	D
GCAC5	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1
GCAC5				C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	A	C	C	C	A	C	C	C	C	C	C	C	A	C	A	C	A	C	A	C	C	C	C	C	C	A	C	C	C	C	C	A	C	C	C	C
GCAC6	1	0	1	0	0	1	1	1	0	0	1	1	0	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1
GCAC6				B	B	D	D	D	B	B	D	D	B	D	D	B	D	B	D	B	D	D	D	D	B	D	B	D	D	B	B	D	D	D	D	D	D	B	D	D	D	D	D	D	D	D	D	B	D	B	D	D
GCAC7	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	0	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0
GCAC7				A	C	C	C	C	C	C	C	C	C	C	C	C	A	A	A	C	C	C	C	A	C	C	A	C	A	A	C	A	C	A	C	C	C	C	C	C	A	C	C	C	A	C	C	C	C	C	C	C	A
CCAG2	0	1	1	1	-	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	0	1	1	0	0	1	1	1	0	0	0	-	0	1	1	0	0	1	1	0	1	1	1	-	1	1	1	0
CCAG2				C	-	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	A	A	A	C	A	C	C	A	A	C	C	C	A	A	A	-	A	C	C	A	A	C	C	A	C	C	C	-	C	C	C	A
CCAG5	1	0	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	0	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0	1	1	-	1	1	1	1	1	1	1	1	1	0	1	-	1	0	1	1
CCAG5				D	D	D	B	D	D	B	D	B	D	B	D	B	B	D	B	B	B	B	D	D	B	B	D	D	B	D	B	B	B	B	D	D	-	D	D	D	D	D	D	D	D	D	B	D	-	D	B	D	D
CCTG5	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	1	1	1	1	1
CCTG5				A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	A	C	C	C	C	A	C	A	C	A	C	A	C	A	C	C	A	C	C	C	C	C	C	C	C
CCTG6	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	0	1	1	0	1	1	0	1	1	0	1
CCTG6				D	D	D	D	D	D	B	B	B	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	B	D	D	D	D	B	D	B	D	D	D	B	D	D	B	D	D	B	D	D	B	D

REFERENCES

ADAM-BLONDON, A.F., SÉVIGNAC, M., BANNEROT, H. and DRON, M. 1994. SCAR, RAPD and RFLP markers linked to a dominant gene (*Are*) conferring resistance to anthracnose in common bean. *Theoretical and Applied Genetics* 88:865-870.

AKKAYA, M.S., BHAGWAT, A.A. and CREGAN, P.B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 32:1131-1139.

AKKAYA, M.S., SHOEMAKER, R.C., SPECHT, J.E., BHAGWAT, A.A. and CREGAN, P.B. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. *Crop Science* 35:1439-1445.

ALI, S., MÜLLER, C.R. and EPPLEN, J.T. 1986. DNA fingerprinting by oligonucleotide probes specific for simple repeats. *Human Genetics* 74:239-243.

APUYA, N.R., FRAZIER, B.L., KEIM, P., ROTH, E.J. and LARK, K.G. 1988. Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merrill. *Theoretical and Applied Genetics* 75: 889-901.

BAI, D., REELEDER, R. and BRANDLE, J.E. 1995. Identification of two RAPD markers tightly linked with the *Nicotiana debneyi* gene for resistance to black root rot of tobacco. *Theoretical and Applied Genetics* 91:1184-1189.

BARRETT, B.A., KIDWELL, K.K. and FOX, P.N. 1998. Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. *Crop Science* 38:1271-1278.

BEAUMONT, V.H., MANTET, J., ROCHEFORD, T.R. and WIDHOLM, J.M. 1996. Comparison of RAPD and RFLP markers for mapping F_2 generations in maize (*Zea mays* L.). *Theoretical and Applied Genetics* 93:606-612.

BECKER, J., VOS, P., KUIPER, M., SALAMINI, F. and HEUN, M. 1995. Combined mapping of AFLP and RFLP markers in barley. *Molecular and General Genetics* 249:65-73.

BIERWERTH, S., KAHL, G., WEIGAND, F. and WEISING, K. 1992. Oligonucleotide fingerprinting of plant and fungal genomes: A comparison of radioactive, colorigenic and chemiluminescent detection methods. *Electrophoresis* 13:115-122.

BOERMA, H.R. and MIAN, M.A.R. 1999. Soybean quantitative trait loci and marker assisted breeding. World Soybean research conference VI, Chicago, Illinois, USA. August 4-7.

BONIERBALE, M.W., PLAISTED, R.L. and TANKSLEY, S.D. 1988. RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120:1095-1103.

BOONE, W.E., STAVELY, J.R. and WEEDEN, N.F. 1999. Development of a sequence-tagged site (STS) marker for *Ur-11*, a gene conferring resistance to the bean rust fungus, *Uromyces appendiculatus*. *Annual Report of the Bean Improvement Corporation* 42:33-34.

BOTSTEIN, D., WHITE, R.L., SKOLNICK, M. and DAVIS, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32:314-331.

BOUTIN, S., ANSARI, H., CONCIBIDO, V., DENNY, R., ORF, J. and YOUNG, N. 1992. RFLP analysis of cyst nematode resistance in soybeans. *Soybean Genetics Newsletter* 19:123-127.

BRUMMER, E.C., GRAEF, G.L., ORF, J., WILCOX, J.R. and SHOEMAKER, R.C. 1997. Mapping QTL for seed protein and oil content in eight soybean populations. *Crop Science* 37:370-378.

BUCHKO, J. and KLASSEN, G.R. 1990. Detection of length heterogeneity in the ribosomal DNA of *Pythium ultimum* by PCR amplification of the intergenic region. *Current Genetics* 18:203-205.

BURR, B. 1994. Some concepts and new methods for molecular mapping in plants. In: DNA-Based Markers in Plants. R.L. Phillips and I.K. Vasil (eds), 1-7.

BURTON, J.W. 1997. Soyabean (*Glycine max* (L.) Merr.) *Field Crops Research* 53: 171-186.

BYRUM, J.R., KIMBIRAUSKAS, P.M., SHOEMAKER, R.C. and DIERS, B.W. 1993. Identification of a RAPD marker linked to the RPS_4 gene in soybean [*Glycine max* (L.) Merr.]. *Soybean Genetics Newsletter* 20:112-117.

CAETANO-ANOLLÉS, G., BASSAM, B.J. and GRESSHOFF, P.M. 1991. DNA amplification fingerprinting using short arbitrary oligonucleotide primers. *Bio/Technology* (New York) 9:553-557.

CAETANO-ANOLLÉS, G., BASSAM, B.J. and GRESSHOFF, P.M. 1993a. Enhanced detection of polymorphic DNA : identification of markers tightly linked to the supernodulation locus in soybean by multiple arbitrary amplicon profiling of endonuclease digested DNA. *Soybean Genetics Newsletter* 20:173-179.

CAETANO-ANOLLÉS, G., BASSAM, B.J. and GRESSHOFF, P.M., 1993b. 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. *Molecular and General Genetics* 241: 57-64.

CANCILLA, M.R., POWELL, I.B., HILLIER, A.J. and DAVIDSON, B.E. 1992. Rapid genomic fingerprinting of *Lactococcus lactis* strains by arbitrarily primed polymerase chain reaction with ^{32}P and fluorescent labels. *Applied and Environmental Microbiology* 58(5):1772-1775.

CAP, G.B., ROBERTS, P.A. and THOMASON, I.J. 1993. Inheritance of heat-stable resistance to *Meloidogyne incognita* in *Lycopersicon peruvianum* and its relationship to the *Mi* gene. *Theoretical and Applied Genetics* 85:777-783.

CERVERA, M.T., GUSMAO, J., STEENACKERS, M., PELEMAN, J., STORME, V., VANDEN BROECK, A., VAN MONTAGU, M. and BOERJAN, W. 1996. Identification of AFLP molecular markers for resistance against *Melampsora larici-populina* in *Populus*. *Theoretical and Applied Genetics* 93:733-737.

CHAPARRO, J.X., WERNER D.J., O'MALLEY, D. and SEDEROFF, R.R. 1994. Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. *Theoretical and Applied Genetics* 87:805-815.

CHARTERS, Y.M., ROBERTSON, A., WILKINSON, M.J. and RAMSAY, G.1996. PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *Oleifera*) using 5'-anchored simple sequence repeat (SSR) primers.

CHASE, C.D., ORTEGA, V.M. and VALLEJOS, C.E. 1991. DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. *Theoretical and Applied Genetics* 81:806-811.

CHEN, L.-F.O., KUO, H.-Y., CHEN, M.-H., LAI, K.-N. and CHEN, S.-C.G. 1997. Reproducibility of the differential amplification between leaf and root DNAs in soybean revealed by RAPD markers. *Theoretical and Applied Genetics* 95:1033-1043.

CHITWOOD, B.G. 1949. Root-knot nematodes. I. A revision of the genus Meloidogyne Goeldi 1887. Proceedings of the Helminthological Society Washington 16:90-104.

CHO, Y.G., BLAIR, W.M., PANAUD, O. and McCOUCH, S.R. 1996. Cloning and mapping of variety-specific rice genomic DNA sequences: amplified fragment length polymorphisms (AFLP) from silver-stained polyacrylamide gels. *Genome* 39:373-378.

CLOSE, P.S., SHOEMAKER, R.C. and KEIM, P. 1989. Distribution of restriction site polymorphism within the chloroplast genome of the genus *Glycine*, subgenus *Soja*. *Theoretical and Applied Genetics* 77: 768-776.

COE, E.H. and GARDINER, J.M. 1994. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp.240-245.

CONCIBIDO, VC, DENNY, RL, BOUTIN, SR, HAUTEA, R, ORF, JH and YOUNG ND. 1994. DNA marker analysis of loci underlying resistance to soybean cyst nematode (*Heterodera glycines* Ichinohe). *Crop Science* 34:240-246.

CONCIBIDO, V.C., DENNY, R.L., LANGE, D.A., ORF, J.H. and YOUNG, N.D. 1996a. RFLP mapping and marker-assisted selection of soybean cyst nematode resistance in PI209332. *Crop Science* 36:1643-1650.

CONCIBIDO, V.C., YOUNG, N.D., LANGE, D.A., DENNY, R.L., DANESH, D. and ORF, J.H. 1996b. Targeted comparative genome analysis and qualitative mapping of a major partial-reistance gene to the soybean cyst nematode. *Theoretical and Applied Genetics* 93:234-241.

CUBETA, M.A., ECHANDI, E., ABERNETHY, T. and VILGALYS, R. 1991. Characterization of anastomosis groups of binucleate *Rhizoctonia* species using restriction analysis of an amplified ribosomal RNA gene. *Phytopathology* 81(11):1395-1400.

DÁVILA, J.A., LOARCE, Y. and FERRER, E. 1999. Molecular characterization and genetic mapping of random amplified microsatellite polymorphism in barley. *Theoretical and Applied Genetics* 98:265-273.

DAX, E., LIVNEH, O., ALISKEVICIUS, E., EDELBAUM, O., KEDAR, N., GAVISH, N., MILO, J., GEFFEN, F., BLUMENTHAL, A., RABINOWICH, H.D. and SELA, I. 1998. A SCAR marker linked to the ToMV resistance gene, *Tm2*^{2}, in tomato. *Euphytica* 101:73-77.

DELANNAY, X., RODGERS, D.M. and PALMER, R.G. 1983. Relative genetic contributions among ancestral lines to North American soybean cultivars. *Crop Science* 23:944-949.

DELLAPORTA, S.L., WOOD, J. and HICKS, J.B. 1983. A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter* 1:19-20.

DE O. ZIMMERMAN, M.J. 1993. Genome structure: uses of molecular tools. Proceedings of the Second International Scientific Meeting: *Phaseolus* Beans Advanced Biotechnology Research Network, CIAT. Cali-Colombia, 7-10 September.

DERISI, J.L., IYER, V.R. and BROWN, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.

DEVOS, K.M. and GALE, M.D. 1992. The use of random amplified polymorphic DNA markers in wheat. *Theoretical and Applied Genetics* 84:567-572.

DIERS, B.W., KEIM, P., FEHR, W.R. and SHOEMAKER, R.C. 1992a. RFLP analysis of soybean seed protein and oil content. *Theoretical and Applied Genetics* 83:608-612.

DIERS, B.W., MANSUR, L., IMSANDE, J. and SHOEMAKER, R.C. 1992b. Mapping Phytphtora resistance loci in soybean with restriction fragment length polymorphism markers. *Crop Science* 32:377-383.

DIWÀN, N. and CREGAN, P.B. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics* 95:723-733.

DOERGE, R.W., ZENG, Z-B. and WEIR, B.S. 1994. Statistical issues in the search for genes affecting quantitative traits in populations.In:Analysis of molecular marker data. Joint plant breeding symposia series. 5-6 August.p15-26.

DONINI, P., ELIAS, M.L., BOUGOURD, S.M. and KOEBNER, R.M.D. 1997. AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. *Genome* 40:521-526.

DOONG, J.Y.H. and KIANG, Y.T. 1987. Cultivar identification by isozyme analysis. *Soybean Genetics Newsletter* 14: 187-227.

DOYLE, J.J., 1988. 5S ribosomal gene variation in the soybean and its progenitor. *Theoretical and Applied Genetics* 75: 621-624.

DOYLE, J.J. and BEACHY, R.N., 1985. Ribosomal gene variation in soybean (*Glycine*) and its relatives. *Theoretical and Applied Genetics* 70: 369-376.

DROPKIN, V.H. 1969. The necrotic reaction of tomatoes and other hosts resistant to *Meloidogyne*: reversal by temperature. *Phytopathology* 59:1632-1637.

DROPKIN, V.H. and NELSON, P.E. 1960. The histopathology of root-knot nematode infections in soybeans. *Phytopathology* 50:442-447.

DWEIKAT, I., MACKENZIE, S., LEVY, M. and OHM, H. 1993. Pedigree assessment using RAPD-DGGE in cereal crop species. *Theoretical and Applied Genetics* 85:497-505.

DWEIKAT, I., OHM, H., PATTERSON, F. and CAMBRON, S. 1997. Identification of RAPD markers for 11 Hessian fly resistance gene in wheat. *Theoretical and Applied Genetics* 94:419-423.

EDWARDS, K., JOHNSON, C. and THOMPSON, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* 19(6):1349.

EUJAYL, I., BAUM, M., ERSKINE, W., PEHU, E. and MUEHLBAUER, F.J. 1997. The use of RAPD markers for lentil genetic mapping and the evaluation of distorted F_2 segregation. *Euphytica* 96:405-412.

FICK, G.N. 1978. Breeding and Genetics. In: Sunflower science and technology. Carter, J.F. (ed.) Agronomy 19: 279-338.

FIEDLER, J., BUFLER, G. and BANGERTH, F. 1998. Genetic relationships of avocado (*Persea americana* Mill.) using RAPD markers. *Euphytica* 101:249-255.

FODOR, S.P.A. 1997. Massively parallel genomics. *Science* 277:393-394.

FOURIE, H., McDONALD, A.H. and LOOTS, G.C. 1999. Host suitability of South African commercial soybean cultivars to two root-knot nematode species and races. In press: *African Plant Protection*.

GAIOTTO, F.A., BRAMUCCI, M. and GRATTAPALIA, D. 1997. Estimation of oucrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. *Theoretical and Applied Genetics* 95:842-849.

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.

GIOVANNONI, J.J., WING, R.A., GANAL, M.W. and TANKSLEY, S.D. 1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research* 19:6553-6558.

GRABAU, E.A., DAVIS, W.H., PHELPS, N.D. and GENGENBACH, B.G. 1992. Classification of soybean cultivars based on mitochondrial DNA restriction fragment length polymorphisms. *Crop Science,* 32: 271-274.

GUILLIN, E., BAUM, B.R. and MECHANDA, S. 1998. Development of an identification scheme for Canadian registered oat cultivars using RAPDs. *Canadian Journal of Plant Science* 78:605-610.

GURLEY, W.B., HEPBURN, A.G. and KEY, J.L. 1979. Sequence organization of the soybean genome. *Biochemica Biophysica Acta* 561:167-183.

HALEY, S.D., AFANADOR, L. and KELLY, J.D. 1994. Selection for monogenic pest resistance with coupling- and repulsion-phase RAPD markers. *Crop Science* 34:1061-1066.

HALEY, S.D., MIKLAS, P.N., STAVELY, J.R., BYRUM, J. and KELLY, J.D.1993. Identification of RAPD markers linked to a major rust resistance gene block in common bean. *Theoretical and Applied Genetics* 86:505-512.

HALWARD, T., STALKER, H.T. and KOCHERT, G. 1994. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp.246-260.

HAN, T.H., VAN ECK, H.J., DE JEU, M.J. and JACOBSEN, E. 1999. Optimization of AFLP fingerprinting of organisms with a large-sized genome: a study on *Alstroemeria* spp. *Theoretical and Applied Genetics* 98:465-471.

HANSON, W.D. 1959. Minimum family sizes for the planning of genetic experiments. *Agronomy Journal* 51(11):711-715.

HARTL, L., WEISS, H., STEPHAN, U., ZELLER, F.J. and JAHOOR, A. 1995. Molecular identification of powdery mildew resistance genes in common wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 90:601-606.

HAYMES, K.M., HENKEN, B., DAVIS, T.M. and VAN DE WEG, W.E. 1997. Identification of RAPD markers linked to a *Phytophthora fragaria* resistance gene (*Rpf1*) in the cultivated strawberry. *Theoretical and Applied Genetics* 94:1097-1101.

HE, S., OHM, H. and MACKENZIE, S. 1992. Detection of DNA sequence polymorphisms among wheat variaties. *Theoretical and Applied Genetics* 84:573-578.

HELENTJARIS, T. 1987. A genetic linkage map for maize based on RFLPs. *Trends in Genetics* 3:217-221.

HELENTJARIS, T., KING, G., SLOCUM, M., SIEDENSTRANG, C. and WEGMAN, S. 1985. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant Molecular Biology* 5:109-118.

HERMAN, M., HUSSEY, R.S. and BOERMA, H.R. 1991. Penetration and development of *Meloidogyne incognita* on roots of resistant soybean genotypes. *Journal of Nematology* 23(2):155-161.

HITTALMANI, S., FOOLAD, M.R., MEW, T., RODRIGUEZ, R.L. and HUANG, N. 1995. Development of a PCR-based marker to identify rice blast resistance gene, *Pi-2(t)*, in a segregating population. *Theoretical and Applied Genetics* 91:9-14.

HONGTRAKUL, V., HUESTIS, G.M. and KNAPP, S.J. 1997. Amplified fragment length polymorphisms as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. *Theoretical and Applied Genetics* 95:400-407.

HU, X.Y., OHM, H.W. and DWEIKAT, I. 1997. Identification of RAPD markers linked to the gene *PM1* for resistance to powdery mildew in wheat. *Theoretical and Applied Genetics* 94:832-840.

HUANG, N., ANGELES, E.R., DOMINGO, J., MAGPANTAY, G., SINGH, S., ZHANG, G., KUMARAVADIVEL, N., BENNETT, J. and KHUSH, G.S. 1997. Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. *Theoretical and Applied Genetics* 95:313-320.

HUSSEY, R.S. and BOERMA, H.R. 1981. A greenhouse screening procedure for root-knot nematode resistance in soybeans. *Crop Science* 21:794-796.

IMSANDE, M., GRANT, D. and SHOEMAKER, R.C. 1998. QTL in SoyBase: A new perspective. *Soybean Genetics Newsletter* 25:146-148.

JEFFREYS, A.J., WILSON, V. and THEIN, S.L. 1985. Individual-specific "fingerprints" of human DNA. *Nature* 316:76-79.

JOHNSON, E., MIKLAS, P.N., STAVELY, J.R. and MARTINEZ-CRUZADO J.C. 1995. Coupling- and repulsion-phase RAPDs for marker-assisted selection of PI 181996 rust resistance in common bean. *Theoretical and Applied Genetics* 90:659-664.

JONES, C.J., EDWARDS, K.J., CASTAGLIONE, S., WINFIELD, M.O., SALA, F., VAN DE WIEL, C., BREDEMEIJER, G., VOSMAN, B., MATTHES, M., DALY, A., BRETTSCHNEIDER, R., BETTINI, P., BUIATTI, M., MAESTRI, E., MALCEVSCHI, A., MARMIROLI, N., AERT, R., VOLCKAERT, G., RUEDA, J., LINACERO, R., VAZQUEZ, A. and KARP, A. 1997. Reprodicibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3:381-390.

KAZAN, K., MANNERS, J.M. and CAMERON, D.F. 1993a. Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus *Stylosanthes*. *Genome* 36:50-56.

KAZAN, K., MANNERS, J.M. and CAMERON, D.F. 1993b. Genetic variation in agronomically important species of *Stylosanthes* determined using random amplified polymorphic DNA markers. *Theoretical and Applied Genetics* 85: 882-888.

KEIM, P., DIERS, B.W., OLSON, T.C. and SHOEMAKER, R.C. 1990a. RFLP mapping in the soybean: association between marker loci and variation in quantitative traits. *Genetics* 126:735-742.

KEIM, P., DIERS, B.W. and SHOEMAKER, R.C. 1990b. Genetic analysis of soybean hard seededness with molecular markers. *Theoretical and Applied Genetics* 79:465-469.

KEIM, P., SCHUPP, J.M., TRAVIS, S.E., CLAYTON, K., ZHU, T., SHI, L., FERREIRA, A. and WEBB, D.M. 1997. A high-density soybean genetic map based on AFLP markers. *Crop Science* 37:537-543.

KEIM, P. and SHOEMAKER, R.C. 1988. Construction of a random recombinant DNA library that is primarily single copy sequences. *Soybean Genetics Newsletter* 15:147-148.

KEIM, P., SHOEMAKER, R.C. and PALMER, R.G. 1989. Restriction fragment length polymorphism diversity in soybean. *Theoretical and Applied Genetics* 77: 786-792.

KIANG, Y.T. and CHIANG, Y.C. 1985. Genetic linkage of biochemical and morphological traits in soybean. *Crop Science* 25:1025-1027.

KIANG, Y.T. and GORMAN, M.B. 1983. Soybean. In: S.D. Tanksley and T.J. Orton (Editors), Isozymes in Plant Genetics and Breeding, Part B. Elsevier Science Publishers, Amsterdam, 295-328 pp.

KINLOCH, RA. 1974. Response of soybean cultivars to nematicidal treatments of soil infested with *Meloidogyne incognita*. *Journal of Nematology* 6:7-11.

KLEYNHANS, K.P.N. 1991. The root-knot nematodes of South Africa. Technical communication - Department of Agricultural Development, no.231.

KNAPP. S.J. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. *Crop Science* 38:1164-1174.

KOCHERT, G. 1994. RFLP technology. In: DNA-Based Markers in Plants. R.L. Phillips and I.K. Vasil (eds), 8-38.

KOLLER, B., LEHMANN, A., MCDERMOTT, J.M. and GESSLER, C. 1993. Identification of apple cultivars using RAPD markers. *Theoretical and Applied Genetics* 85: 901-904.

LANDER, E.S. and BOTSTEIN, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.

LANDRY, B.S., KESSELI, R.V., FARRARA, B. and MICHELMORE, R.W. 1987. A genetic map of lettuce (*lactuca sativa* L.) With restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. *Genetics* 116:331-337.

LARK, K.G., EVANS, J., BASHA, F., BOGDEN, R., COPELAND, R., ELLISON, R., HORNE, D., LEE, K., MCDONALD, K., PIERSON, J., SCHUSTER, W., WILHELM, P. and YU, Y. 1992. Molecular phylogeny as a genetic tool for soybean breeding. *Soybean Genetics Newsletter* 19: 174-181.

LARK, K.G., ORF, J. and MANSUR, L.M. 1994. Epistatic expression of quantitative trait loci (QTL) in soybean [*Glycine max* (L.) Merr.] determined by QTL association with RFLP alleles. *Theoretical and Applied Genetics* 88:486-489.

LARK, K.G., WEISEMAN, J.M., MATTHEWS, B.F., PALMER, R., CHASE, K. and MACALMA, T. 1993. A genetic map of soybean (*Glycine max* L.) Using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir 1'. *Theoretical and Applied Genetics* 86:901-906.

LARSEN, A.L. and BENSON, W.C. 1970. Variety-specific variants of oxidative enzymes from soybean seed. *Crop Science* 10: 493-495.

LAW, J.R., DONINI, P., KOEBNER, R.M.D., REEVES, J.C. and COOKE, R.J. 1998. DNA profiling and plant variety registration. III: The statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. *Euphytica* 102:335-342.

LAWSON, W.R., GOULTER, K.C., HENRY, R.J., KONG, G.A. and KOCHMAN, J.K. 1996. RAPD markers for a sunflower rust resistance gene. *Australian Journal for Agricultural Research* 47:395-401.

LEE, M. 1995. DNA markers and plant breeding programs. *Advances in Agronomy* 55:265-344.

LEE, S.H., BAILEY, M.A., MIAN, M.A.R., CARTER, T.E., ASHLEY, D.A., HUSSEY, R.S., PARROTT, W.A. and BOERMA, H.R. 1996. Molecular markers associated with soybean plant height, lodging, and maturity across locations. *Crop Science* 36:728-735.

LINCOLN, S., DALY, M. and LANDER, E. 1992. Mapping genes controlling quantitative traits with MAPMAKER/QTL 1.1. Whitehead Institute Technical Report 2nd edition.

LIU, C.J. 1997. Geographical distribution of genetic variation in *Stylosanthes scabra* revealed by RAPD analysis.

LU, J., KNOX, M.R., AMBROSE, M.J., BROWN, J.K.M. and ELLIS, T.H.N. 1996. Comparative analysis of genetic diversity in pea assessed by RFLP- and PCR-based methods. *Theoretical and Applied Genetics* 93(7):1103-1111

LUO, Z.W. and KEARSEY, M.J. 1989. Maximum likelihood estimation of linkage between a marker gene and a quantitative locus. *Heredity* 63:401-408.

LUZZI, B.M., BOERMA, H.R. and HUSSEY, R.S. 1987. Resistance to three species of root-knot nematode in soybean. *Crop Science* 27:258-262.

LUZZI, B.M., BOERMA, H.R. and HUSSEY, R.S. 1994a. Inheritance of resistance to the Southern root-knot nematode in soybean. *Crop Science* 34:1240-1243.

LUZZI, B.M., BOERMA, H.R. and HUSSEY, R.S. 1994b. A gene for resistance to the southern root-knot nematode in soybean. *Journal of Heredity* 85:484-486.

LUZZI, B.M., BOERMA, H.R. and HUSSEY, R.S. 1995a. Inheritance of resistance to the peanut root-knot nematode in soybean. *Crop Science* 35:50-53.

LUZZI, B.M., TAMULONIS, J.P., HUSSEY, R.S. and BOERMA, H.R. 1995b. Inheritance of resistance to the Javanese root-knot nematode in soybean. *Crop Science* 35:1372-1375.

MACKILL, D.J. 1995. Classifying Japonica rice cultivars with RAPD markers. *Crop Science* 35:889-894.

MAHESWARAN, M., SUBUDHI, P.K., NANDI, S., XU, J.C., PARCO, A., YANG, D.C. and HUANG, N. 1997. Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. *Theoretical and Applied Genetics* 94:39-45.

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.). *Theoretical and Applied Genetics* 86:907-913.

MARTIN, G.B., WILLIAMS, J.G.K. and TANKSLEY, S.D. 1991. Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. *Proceedings of the National Academy of Sciences, USA* 88:2336-2340.

MAUGHAN, P.J., SAGHAI MAROOF, M.A., BUSS, G.R. and HUESTIS, G.M. 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. *Theoretical and Applied Genetics* 93:392-401.

McMULLEN, M.D., JONES, M.W., SIMCOX, K.D. and LOUIE, R. 1994. Three genetic loci control resistance to wheat streak mosaic virus in the maize inbred Pa405. *Molecular Plant-Microbe Interactions* 7(6):708-712.

MEKSEM, K., LEISTER, D., PELEMAN, J., ZABEAU, M., SALAMINI, F. and GEBHARDT, C. 1995. A high-resolution map of the vicinity of the *R1* locus on chromosome V of potato based on RFLP and AFLP markers. *Molecular and General Genetics* 249:74-81.

MELOTTO, M. and KELLY, J.D. 1998. SCAR markers linked to major disease resistance genes in common bean. *Annual Report of the Bean Improvement Corporation* 41:64-65.

MESSEGUER, M.G., DE VICENTE, M.C., YOUNG, N.D., BOLKAN, H. and TANKSLEY, S.D. 1991. High resolution RFLP map around the root knot nematode resistance gene (*Mi*) in tomato. *Theoretical and Applied Genetics* 82:529-536.

MIAN, M.A.R., BAILEY, M.A., ASHLEY, D.A., WELLS, R., CARTER, T.E., PARROTT, W.A. and BOERMA, H.R. 1996. Molecular markers associated with water use efficiency and leaf ash in soybean. *Crop Science* 36:1252-1257.

MIAN, M.A.R., SHIPE, E.R., ALVERNAZ, J., MUELLER, J.D., ASHLEY, D.A. and BOERMA, H.R. 1997. QTLs associated with chlorimuron ethyl sensitivity in soybean: Effects on seed yield and related traits. *Theoretical and Applied Genetics* 94:971-974.

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 using segregating populations. *Proceedings of the National Academy of Sciences USA* 88:9828-9832.

MIENIE, C.M.S., SMIT, M.A. and PRETORIUS, P.J. 1995. Use of random amplified polymorphic DNA for identification of South African soybean cultivars. *Field Crops Research* 43:43-49.

MIKLAS, P.N., JOHNSON, E., STONE, V., BEAVER, J.S., MONTOYA, C. and ZAPATA, M. 1996. Selective mapping of QTL conditioning disease resistance in common bean. *Crop Science* 36:1344-1351.

MIKLAS, P.N., STAVELY, J.R. and KELLY, J.D. 1993. Identification and potential use of a molecular marker for rust resistance in common bean. *Theoretical and Applied Genetics* 85:745-749.

MONEY, T., READER, S., QU, L.J., DUNFORD, R.P. and MOORE, G. 1996. AFLP-based mRNA fingerprinting. *Nucleic Acids Research* 24:2616-2617.

MUDGE, J., CREGAN, P.B., KENWORTHY, J.P., KENWORTHY, W.J., ORF, J.H. and YOUNG, N.D. 1997. Two microsatellite markers that flank the major soybean cyst nematode resistance locus. *Crop Science* 37:1611-1615.

MUEHLBAUER, G.J., STASWICK, P.E., SPECHT, J.E., GRAEF, G.L. SHOEMAKER, R.C. and KEIM,P. 1991. RFLP mapping using near-isogenic lines in the soybean [*Glycine max* (L.) Merr.]. *Theoretical and Applied Genetics* 81:189-198.

MULROONEY, R.P. 1986. Soybean disease loss estimate for Southern United States in 1984. *Plant Disease* 70:893.

NANDI, S., SUBUDHI, P.K., SENADHIRA, D., MANIGBAS, N.L., SEN-MANDI, S. and HUANG, N. 1997. Mapping QTLs for submergence tolerance in rice by AFLP analysis and selective genotyping. *Molecular and General Genetics* 255:1-8.

NEI, M. and LI, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: 5269-5273.

NICESE, F.P., HORMAZA, J.I. and McGRANAHAN, G.H. 1998. Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. *Euphytica* 101:199-206.

NIEBEL, A., DE ALMEIDA, E.J., TIRÉ, C., ENGLER, G., VAN MONTAGU, M. and GHEYSEN, G. 1993. Induction patterns of an extensin gene in tobacco upon nematode infection. *Plant Cell* 5:1697-1710.

PADILLA, J.E., CAETANO-ANOLLÉS, G. and GRESSHOFF, P.M. 1994. Isolation and cloning of AFLPs generated by tec-MAAP and tightly linked to the *nts* locus in soybean. *Soybean Genetics Newsletter.* May. pp285-288.

PAIVA, E., VILARINHOS, A.D., DEBARROS, E.G., SEDIYAMA, C.S. and MOREIRA, M. A. 1994. Use of the random amplified polymorphic DNA technique to characterize soybean (*Glycine max*(L) Merrill) genotypes. *Revista Brasileira De Genetica,* 17: 287-290.

PAKNIYAT, H., POWELL, W., BAIRD, E., HANDLEY, L.L., ROBINSON, D., SCRIMGEOUR, C.M., NEVO, E., HACKETT, C.A., CALILGARI, P.D.S. and FORSTER, B.P. 1997. AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. *Genome* 40:332-341.

PALMER, R.G. and KILEN, T.C. 1987. Qualitative genetics and cytogenetics. P.135-209. *In* Wilcox, J.R. (Ed.) Soybeans: Improvement, production and uses. 2nd edition of Agronomy Monograph 16. ASA, CSSA, and SSSA, Madison, WI.

PARAN, I. and MICHELMORE, R.W. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics* 85:985-993.

PATERSON, A.H. 1996. Comparative mapping sows seeds of progress. *Probe - Newsletter for the USDA Plant Genome Research Program* 6:1-3.

PATERSON, A.H., LANDER, E.S., HEWITT, J.D., PETERSON, S., LINCOLN, S.E. and TANKSLEY, S.D. 1988. Resolution of quantitative traits into Mendelian factore by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721-726.

PATERSON, A.H., TANKSLEY, S.D. and SORRELLS, M.E. 1991. DNA markers in plant improvement. *Advances in Agronomy* 46:39-90.

PAUL, S., WACHIRA, F.N., POWELL, W. and WAUGH, R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. *Theoretical and Applied Genetics* 94:255-263.

POWELL, W., MORGANTE, M., ANDRE, C., HANAFEY, M., VOGEL, J., TINGEY, S. and RAFALSKI, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2:225-238.

PRABHU, R.R. and GRESSHOFF, P.M. 1994. Inheritance of polymorphic markers generated by DNA amplification fingerprinting and their use as genetic markers in soybean. *Plant Molecular Biology* 26: 105-116.

PRABHU, P.R., WEBB, D., JESSEN, H., LUK, S., SMITH, S. and GRESSHOFF, P.M. 1997. Genetic relatedness among soybean genotypes using DNA amplification fingerprinting (DAF), RFLP, and pedigree. *Crop Science* 37:1590-1595.

PREVOST, A. and WILKINSON, M.J. 1999. A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics* 98:107-112.

QIU, B.X., ARELLI, P.R. and SLEPER, D.A. 1999. RFLP markers associated with soybean cyst nematode resistance and seed composition in a 'Peking' x 'Essex' population. *Theoretical and Applied Genetics* 98:356-364.

QU, L.J., FOOTE, T.N., ROBERTS, M.A., ARAGON-ALCAIDE, L., SNAPE, J.W. and MOORE, G. 1998. A simple PCR-based method for scoring the *ph1b* deletion in wheat. *Theoretical and Applied Genetics* 96:371-375.

QUILLET, M.C., MADJIDIAN, N., GRIVEAU, Y., SERIEYS, H., TERSAC, M., LORIEUX, M. and BERVILLÈ, A. 1995. Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* sect. *Helianthus*. *Theoretical and Applied Genetics* 91:1195-1202.

RAPER, C.D. and KRAMER, P.J. 1987. Stress physiology. In: Wilcox, J.R. (Ed.). Soybeans: Improvement, production and uses, 2nd edition - Agronomy Monograph no. 16 p608.

REITER, R.S., WILLIAMS, J.G.K., FELDMAN, 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. *Proceedings of the National Academy of Sciences, USA* 89:1477-1481.

RIEKERT, H.F. 1995. An adapted method for extraction of root-knot nematode eggs from maize root samples. *African Plant Protection* 1:41-43.

RIGGS, R.D. and SCHMITT, D.P. 1987. Nematodes. In: Wilcox, J.R. (Ed.) Soybeans: Improvement, production and uses, 2nd edition - Agronomy Monograph no. 16.

RITTER, E., GEBHARDT, C. and SALAMINI, F. 1990. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. *Genetics* 125:645-654.

ROA, A.C., MAYA, .M.M, DUQUE, M.C., TOHME, J., ALLEM, A.C. and BONIERBALE, M.W. 1997. AFLP analysis of relationships among cassava and other *Manihot* species. *Theoretical and Applied Genetics* 95:741-750.

ROBERTS, P.A. and THOMASON, I.J. 1986. Variability in the reproduction of isolates of *Meloidogyne incognita* and *M. javanica* on resistant tomato genotypes. *Plant Disease* 70:547-551.

ROGERS, S.O. and BENDICH, A.J. 1988. Extraction of DNA from plant tissues. *Plant Molecular Biology Manual* A6:1-10. Kluwer Academic Publishers, Dordrecht.

ROUPPE VAN DER VOORT, J.N.A.M., VAN ZANDVOORT, P., VAN ECK, H.J., FOLKERTSMA, R.T., HUTTEN, R.C.B., DRAAISTRA, J., GOMMERS, F.J., JACOBSEN, E., HELDER, J. and BAKKER, J. 1997a. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. *Molecular and General Genetics* 255:438-447.

ROUPPE VAN DER VOORT, J., WOLTERS, P., FOLKERTSMA, R., HUTTEN, R., VAN ZANDVOORT, P., VINKE, H., KANYUKA, K., BENDAHMANE, A., JACOBSEN, E., JANSSEN, R. and BAKKER, J. 1997b. Mapping of the cyst nematode resistance locus *Gpa2* in potato using a strategy based on comigrating AFLP markers. *Theoretical and Applied Genetics* 95:874-880.

SAIKI, R.K., GELFAND, D.H., STOFFEL, S., SCHARF, S.J., HIGUCHI, R., HORN, G.T., MULLIS, K.B. and ERLICH, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.

SAM, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: *Joinmap*. *Plant Journal* 3:739-744.

SARTORATO, A., NIETSCHE, S., BARROS, E.G. and MOREIRA, M.A. 1999. SCAR marker linked to angular leaf spot resistance gene in common bean. *Annual Report of the Bean Improvement Corporation* 42:23-24.

SAS Institute. 1992. SAS system for Windows 3.10, Release 6.08. SAS Institute Inc, Cary, N.C.

SASSER, J.N. 1977. Worldwide dissemination and importance of the root-knot nematodes, *Meloidogyne* spp. *Journal of Nematology* 9(1):26-29.

SCHACHERMAYR, G., SIEDLER, H., GALE, M.D., WINZELER, H., WINZELER, M. and KELLER, B. 1994. Identification and localization of molecular markers linked to the *Lr9* leaf rust resistance gene of wheat. *Theoretical and Applied Genetics* 88:110-115.

SCHÄFER, R., ZISCHLER, H., BIRSNER, U., BECKER, A. and EPPLEN, J.T. 1988. Optimized oligonucleotide probes for DNA fingerprinting. *Electrophoresis* 9:369-374.

SCHNEIDER, K., ACOSTA-GALLEGOS, J., IBARRA-PEREZ, F. and KELLY, J.D. 1996. Application of marker assisted selection to improving a quantitative trait in common bean. *Annual Report of the Bean Improvement Corporation* 39:61-62.

SCHUT, J.W., QI, X. and STAM, P. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theoretical and Applied Genetics* 95:1161-1168.

SENIOR, M.L., MURPHY, J.P., GOODMAN, M.M. and STUBER, C.W. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Science* 38:1088-1098.

SHALON, D., SMITH, J.S. and BROWN, P.O. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Research* 6:639-645.

SHAN, X., BLAKE, T.K. and TALBERT, L.E. 1999. Concersion of AFLP markers to sequence-specific PCR markers in barley and wheat. *Theoretical and Applied Genetics* 98:1072-1078.

SHARMA, S.K., KNOX, M.R. and ELLIS, T.H.N. 1996. AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theoretical and Applied Genetics* 93:751-758.

SHOEMAKER, R.C. 1994. RFLP map of soybean. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp. 299-309.

SIJMONS, P.C. 1993. Plant-nematode interactions. *Plant Molecular Biology* 23:917-931.

SKORUPSKA, H.T., SHOEMAKER, R.C., WARNER, A., SHIPE, E.R. and BRIDGES, W.C. 1993. Restriction fragment length polymorphism in soybean germplasm of the southern USA. *Crop Science* 33:1169-1176.

SKROCH, P. and NIENHUIS, J. 1995. Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. *Theoretical and Applied Genetics* 91:1086-1091.

SKROCH, P., TIVANG, J. and NIENHUIS, J. 1991. A new PCR based DNA polymorphism assay. *Annual Report of the Bean Improvement Corporation* 34:v-x.

SOLLER, M., BRODY, T. and GENIZI, A. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theoretical and Applied Genetics* 47:35-39.

SONG, K.M., SUZUKI, J.Y., SLOCUM, M.K., WILLIAMS, P.H. and OSBORN, T.C. 1991. A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. *Theoretical and Applied Genetics* 82:296-304.

SOUTHERN, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98:503-517.

STAM, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: *Joinmap. Plant Journal* 3:739-744.

STILES, J.I., LEMME, C., SONDUR, S., MORSHIDI, M.B. and MANSHARDT, R. 1993. Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. *Theoretical and Applied Genetics* 85: 697-701.

TALBERT, L.E., BLAKE, N.K., CHEE, P.W., BLAKE, T.K. and MAGYAR, G.M. 1994. Evaluation of "sequence-tagged-site" PCR products as molecular markers in wheat. *Theoretical and Applied Genetics* 87:789-794.

TAMULONIS, J.P., LUZZI, B.M., HUSSEY, R.S., PARROTT, W.A., and BOERMA, H.R. 1997a. RFLP mapping of resistance to Southern root-knot nematode in soybean. *Crop Science* 37:1903-1909.

TAMULONIS, J.P., LUZZI, B.M., HUSSEY, R.S., PARROTT, W.A. and BOERMA, H.R. 1997b. DNA markers associated with resistance to Javanese root-knot nematode in soybean. *Crop Science* 37:783-788.

TAMULONIS, J.P., LUZZI, B.M., HUSSEY, R.S., PARROTT, W.A. and BOERMA, H.R. 1997c. DNA marker analysis of loci conferring resistance to peanut root-knot nematode in soybean. *Theoretical and Applied Genetics* 95:664-670.

TANKSLEY, S.D. 1994a. Tomato molecular map. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp.310-326.

TANKSLEY, S.D. 1994b. Rice molecular map. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp.286-290.

TANKSLEY, S.D., BERNATZKY, R., LAPITAN, N.L. and PRINCE, J.P. 1988. Conservation of gene repertoire but not gene order in pepper and tomato. *Proceedings of the National Academy of Sciences USA* 85:6419-6423.

TANKSLEY, S.D., YOUNG, N.D., PATERSON, A.H. and BONIERBALE, M.W. 1989. RFLP mapping in plant breeding: New tools for an old science. *Bio/Technology* 7:257-264.

TAYLOR, A.L. and SASSER, J.N. 1978. Biology, identification and control of root-knot nematodes (*Meloidogyne species*). A cooperative publication of the Department of Plant pathology North Carolina State University and the United States Agency for International Development. 111p.

THOMPSON, J.A., NELSON, R.L. and VODKIN, L.O. 1998. Identification of diverse soybean germplasm using RAPD markers. *Crop Science* 38:1348-1355.

TIVANG, J., NIENHUIS, J. and SMITH, O.S. 1992. Sampling variance for molecular marker data. *Annual Report of the Bean Improvement Corporation* 35:66-67.

TOHME, J., GONZALEZ, D.O., BEEBE, S. and DUQUE, M.C. 1996. AFLP analysis of gene pools of a wild bean core collection. *Crop Science* 36:1375-1384.

TSUMURA, Y. and TOMARU, N. 1999. Genetic diversity of *Cryptomeria japonica* using co-dominant DNA markers based on sequenced-tagged sites. *Theoretical and Applied Genetics* 98:396-404.

VALLEJOS, C.E. 1994. *Phaseolus vulgaris*: the common bean. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp.261-270.

VANTOAI, T.T., PENG, J. and ST.MARTIN, S.K. 1997. Using AFLP markers to determine the genomic contribution of parents to populations. *Crop Science* 37:1370-1373.

VEREMIS, J.C., VAN HEUSDEN, A.W. and ROBERTS, P.A. 1999. Mapping a novel heat-stable resistance to *Meloidogyne* in *Lycopersicon peruvianum*. *Theoretical and Applied Genetics* 98:274-280.

VOORRIPS, R.E., JONGERIUS, M.C. and KANNE, H.J. 1997. Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of doubled haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. *Theoretical and Applied Genetics* 94:75-82.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., VAN DE LEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M. and ZABEAU, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23(21):4407-4414.

WANG, M.L., ATKINSON, M.D., CHINOY, C.N., DEVOS, K.M., HARCOURT, R.L., LIU, C.J., ROGERS, W.J. and GALE, M.D. 1991. RFLP-based genetic map of rye (*Secale cereale* L.) Chromosome *1R*. *Theoretical and Applied Genetics* 82:174-178.

WANG, G.-L. and PATERSON, A.H. 1994. Assessment of DNA pooling strategies for mapping of QTLs. *Theoretical and Applied Genetics* 88:355-361.

WANG, Z.Y. and TANKSLEY, S.D. 1989. Restriction fragment length polymorphisms in *Oryza sativa* L. *Genome* 32: 1113-1118.

WANG, Y.-H., THOMAS, C.E. and DEAN, R.A. 1997. A genetic map of melon (*Cucumis melo* L.) based on amplified fragment length polymorphism (AFLP) markers. *Theoretical and Applied Genetics* 95:791-798.

WELLER, J.I. 1986. Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. *Biometrics* 42:627-640.

WELSH, J. and McCLELLAND, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18:7213-7218.

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. *Nucleic Acids Research* 18:6531-6535.

WILLIAMS, M.N.V., PANDE, N., NAIR, S., MOHAN, M. and BENNETT, J. 1991. Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa L.*) genomic DNA. *Theoretical and Applied Genetics* 82:489-498.

WILLIAMSON, V.M. and HUSSEY, R.S. 1996. Nematode pathogenesis and resistance in plants. *The Plant Cell* 8:1735-1745.

XIE, C. and XU, S. 1998. Strategies of marker-aided recurrent selection. *Crop Science* 38:1526-1535.

YANAGISAWA, T., HAYASHI, M., HIRAI, A. and HARADA, K. 1994. DNA fingerprinting in soybean (*Glycine max* (L.) Merrill) with oligonucleotide probes for simple repetitive sequences. *Euphytica* 80:129-136.

YANG, X. and QUIROS, C. 1993. Identification and classification of celery cultivars with RAPD markers. *Theoretical and Applied Genetics*, 86: 205-212.

YOUNG, N.D. 1994. Constructing a plant genetic linkage map with DNA markers. In: DNA-Based Markers in Plants. Phillips, R.L. and Vasil, I.K. (eds.)pp.39-57.

YOUNG, R.A. and KELLY, J.D. 1996. Gene pyramiding using marker assisted selection for stable resistance to bean anthracnose. *Annual Report of theBean Improvement Corporation* 39:57-58.

YU, K., PARK, S.J. and POYSA, V. 1998. Abundance and variation of microsatellite DNA sequences in common beans. *Annual Report of the Bean Improvement Corporation* 41:45-46.

YU, K., PARK, S.J. and POYSA, V. 1999. Application of molecular markers to assist selection of common beans resistant to common bacterial blight (*Xanthomonas campestris pv. Phaseoli*). *Annual Report of the Bean Improvement Corporation* 42:3-4.

ZABEAU, M. and VOS, P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application 92402629.7 (Publ. No. 0 534 858 A1).

ZHANG, Z., CHEN, S. and GAI, J. 1998. A codominant SCAR marker linked to SMV resistance gene *Rsa*. *Soybean Genetics Newsletter* 25:27-28.

ZIETKIEWICZ, E., RAFALSKI, A. and LABUDA, D. 1994. Genome fingerprinting by simple sequence repeats (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20:176-183.

ZINK, D. and NAGL, W. 1996. Microsatellite-primed PCR in different *Phaseolus* species. *Annual Report of the Bean Improvement Corporation* 39:243-244.