The development of a SCAR marker for the identification of the potato cultivars Astrid and Mnandi

by

Willem Sternberg Jansen van Rensburg

(9608445)

Dissertation submitted in fulfillment of the requirements for the degree

Master of Science

in

Biochemistry

in the

Faculty of Science

at the

Rand Afrikaans University

Supervisor: Prof. I. A. Dubery

May 1999

To suffer woes which hope thinks infinite,
To forgive wrongs darker then death or night,
To defy power which seems omnipotent,
Neither change, nor falter, nor repent:
This . . . . is to be
Good, great, joyous, beautiful and free,
This is alone life, joy, empire and victory.

from Shelley's Prometheus Unbound

Table of contents

Summary

Mnandi and Astrid are two commercially important potato cultivars in South Africa. These two cultivars are closely related and are morphological virtually identical. It is, however, necessary to be able to distinguish between these two cultivars, because each of these cultivars has certain desirable characteristics.

It was decided to use DNA markers, since DNA markers are not influenced by the environment and the polymerase chain reaction (PCR) based DNA markers are relatively easy, cheap and fast. It was decided to develop a sequenced characterized amplified region (SCAR) due to the problems with the reproducibility of random amplified polymorphic DNA (RAPDs). SCARs are derived from RAPD fragments by using the sequence of a RAPD derived fragment to design a set of new longer primers (usually 20-24mer) which are less sensitive to PCR conditions.

Ten commercial potato cultivars (Astrid, Mnandi. BP_1, Buffelspoort, VanderPlank, Up-to-Date, Hoëvelder, Hertha, Pimpernel and Agria) were used in this study. Commercially available RAPD primers (102) were evaluated to seek a polymorphism unique to either Mnandi or Astrid. Thirtyseven polymorphisms between Astrid and Mnandi were identified but only three were unique. The polymorphism obtained with OPH-15 was however, not reproducible. The polymorphisms obtained with UBC 509 and 582, corresponding to the presence in Mnandi of a 300 and 900 bp fragment respectively, were reproducible.

These two fragments, UBC 509^{300} and UBC 582^{900}, were cloned into the pMos*Blue* TA cloning vector and sequenced. The identity if the inserts in the recombinant plasmids were verified with PCR and Southern blotting. The sequences were used to develop two sets of SCAR primers, SCAR UBC 509^{300} and SCAR UBC 582^{900}.

The two SCAR primer pairs were then used in PCR reactions. The SCAR UBC 509^{300} primer pair amplified a fragment of 230 bp in both Astrid and Mnandi and a fragment of 260 bp in Mnandi. The polymorphism is thus retained and SCAR UBC 509^{300} can be used to distinguish between Astrid and Mnandi. The SCAR UBC 582^{900} primer pair amplify a fragment of 500 bp in both Astrid and Mnandi as well as some other longer fragments. It was not possible to regain a polymorphism by either elevating the annealing temperature or by digesting the amplification products with restriction enzymes. SCAR UBC 582^{900} could thus not be used to distinguish between Astrid and Mnandi.

Opsomming

Mnandi en Astrid is twee ekonomies belangrike aartappel kultivars in Suid Afrika. Die twee kultivars is naby verwant en lyk morfologies omtrent dieselfde. Tog besit elk spesifieke eienskappe en is dit belangrik om tussen die twee te onderskei.

Daar is besluit om gebruik te maak van DNS merkers, omdat hierdie merkers nie deur die omgewing beïnvloed word nie en omdat die polimerse ketting reaksie (PKR) gebaseerde merkers relatief goedkoop, vinnig en maklik is om te gebruik. Daar was verder besluit om 'n volgorde gekarakteriseerde geamplifiseerde area (VKAA) te ontwikkel, aangesien daar vrae bestaan oor die herhaalbaarheid van RAPDs. VKAAs word ontwikkel vanuit lukraak geamplifiseerde polimorfiese DNS (RAPD) merkers deur die RAPD geamplifiseerde fragment se DNS volgorde te bepaal en die volgorde te gebruik om twee langer (gewoonlik 20-24 mer) voorvoerders te ontwerp wat nie so sensitief vir PKR reaksie kondisies is nie.

Tien kommersiële aartappel cultivars (Astrid, Mnandi. BP_1, Buffelspoort, VanderPlank, Up-to-Date, Hoëvelder, Hertha, Pimpernel en Agria) was gebruik in die studie. Kommersiële RAPD voorvoerders (102) is geëvalueer vir polimorfismes uniek aan Astrid of Mnandi. Seweendertig polimorfismes is geïdentifiseer waarvan slegs drie uniek was. Die polimorfisme verkry met OPH15 was nie herhaalbaar nie, maar die polimorfisme verkry met UBC 509 en 582, ooreenstemmend met die teenwoordigheid van onderskeidelik 'n 300 en 900 bp fragment in Mnandi, was herhaalbaar.

Die twee fragmente, UBC 509^{300} en UBC 582^{900}, is gekloneer in die pMos*Blue* TA kloneringsvektor en die basisvolgorde van die gekloneerde fragmente is bepaal. Die identiteit van die gekloneerde fragmente is bevestig met PKR en Southern oordrag. Die DNS volgordes is gebruik om twee VKAA voorvoerder stelle, SCAR UBC 509^{300} en SCAR UBC 582^{900} te ontwerp.

Die twee VKAA voorvoerder stelle is in PKR reaksies gebruik. Die SCAR UBC 509^{300} voorvoerders amplifiseer 'n 230 bp fragment in beide Astrid en Mnandi en 'n 260 bp fragment in Mnandi. Die polimorfisme is dus behou en SCAR UBC 509^{300} kan dus gebruik word om te onderskei tussen Astrid en Mnandi. Die SCAR UBC 582^{900} voorvoerders amplifiseer 'n 500 bp fragment in beide Astrid en Mnandi asook langer fragmente. Die verhoging van die bindingstemperatuur sowel as die vertering van die amplifiserings produkte herwin nie 'n polimorfisme nie. SCAR UBC 582^{900} kan dus nie gebruik word om te onderskei tussen Astrid en Mnandi nie.

Acknowledgements

I would like to acknowledge the following institutions and people:

- Prof I Dubery for his guidance, encouragement and support
- The Agricultural Research Council for the use of their facilities and their financial support
- Potatoes South Africa (previously known as Potato Producers Organisation) for their financial support
- Cecilia McGregor and Melanie Arendse for their guidance and assistance
- Ria Greyling for her assistance and support
- Annemarie Ehlers and Marlyn Janse van Rensburg for the scanning of the Southern blots
- Marie McKenzie for her support, encouragement and jokes
- All my other colleagues for their support and encouragement
- My mother for her support and love - "Thankx Mom!"
- Thia and Paul Fenton for their support and encouragement
- Pieter and Andries for their support and encouragement
- All my friends for their support, jokes and laughter - Anna, Jaqui, Xavier and Hennie, Elizabeth, Robert and Henco, Berta and Gerhard, Schalk and Jan, The Duchess and Gerhard - Thankx guys
- All the other people who helped in small but significant ways
- And last for my Creator to whom I owe everything

List of abbreviations

Standard SI and Chemical abreviations are not included.

A	Adenine
AFLP	Amplified Fragment Length Polymorphism
ARC	Agriculture Research Council
AMP-PCR	Anchored microsatellite primed PCR
AP-PCR	Arbitrary Primed PCR
ASAP	Allele specific associated primers
b	Expected number of fragments per primer
bp	Base pair
BLAST	Basic local assignment search tool
c	Complexity of the genome of the organism
C	Cytosine
CAPS	Cleaved Amplified polymorphic sequences
cDNA	Complimentary DNA
CSPD	Disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)trichloro[3.3.1.1^{3,7}]decan}-4-yl) phenyl phosphate
DAF	DNA amplification finger printing
dATP	Deoxyadenine triphosphate
dCTP	Deoxycytosine triphosphate
DDRT PCR	Differential display reverse transcription PCR
DEM	Demeke primer
dGTP	Deoxyguanin triphosphate
dH_2O	Distilled water
DIG	Digoxygenin
DI/TRISEC	Di/Trinucleotide Sticky End Cloning
Dinuc	Dinucleotide
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide

dTTP	Deoxythymine triphosphate
DUS	Distinctness, Uniformity and Stability
F_2	Second Generation
G	guanine
IPTG	Isopropyl B-D-thiogalactopyranoside
kb	kilobasepairs
kcals	kilocalories
LB	Luria Broth
LMP	Low Melting Point
n	primer length in nucleotides
NCC	National Cultivar Collection
PBR	Plant Breeders Rights
PCR	Polymerase Chain Reaction
PPO	Potato Producers Organization
PSA	Potatoes South Africa
QTL	Quantitve Trait Loci
RAPD	Random Amplified Polymorphic DNA
RAPD-DGGE	RAPD Separated via Denaturating-Gradient-Gel Electrophoresis
RFLP	Restriction Fragment Length Polymorphism
rTaq	Recombinant *Taq* DNA polymerase
SAP	Specific Amplicon Polymorphism
SCAR	Sequenced Characterized Amplified Region
spp	species
STS	Sequence Tagged Site
T	thymine
T_ms	Melting temperature of dubblestranded DNA
U	Unit of enzyme
UBC	University of British Colombia

UPOV	International Society for the protection of Variety Rights
VNTR	Variable Number of Tandem Repeats

List of figures

List of tabels

Chapter 1

Introduction

1.1 The ARC-Roodeplaat potato genebank and the South African national cultivar collection (NCC).

The National Potato Genebank is situated at the Potato Section of the ARC-Roodeplaat Vegetable and Ornamental Plant Institute of the Agriculture Research Council (ARC). The main purpose of the potato genebank is to maintain and conserve potato germplasm. The potato germplasm conserved by the ARC-Roodeplaat Potato Genebank consists of South African bred, as well as foreign cultivars, wild species, dihaploid lines and the advanced breeding lines of the commercial breeding program of the ARC-Roodeplaat.

All the cultivars registered at the Directorate of Plant and Quality Control, including all the potato cultivars of commercial importance in South Africa, are maintained in a collection called the National Cultivar Collection (NCC) and consist of South African bred as well as foreign cultivars. New cultivars that are submitted for registration with the Directorate of Plant and Quality Control each year are added and maintained in the NCC, with permission from the owner of the cultivar rights.

Another duty of the potato genebank is to ensure that the NCC stays true to genotype. Cultivar descriptions based on the International Society for the protection of Variety Rights (UPOV) morphological descriptions are done on a regular base by the ARC-Roodeplaat and the Department of Plant and Quality Control. All the commercially important cultivars in the NCC are also characterised with RAPD and microsatelite technology (C McGregor personal communication). The cultivars maintained in the NCC are assessed annually to ensure that they are still true to genotype.

1.2 The need for cultivar identification

The accurate, fast, reliable and cost-effective identification of different potato cultivars is essential as each cultivar has specific qualities and characteristics. The identification of different potato cultivars is most frequently based on the assessment of morphological characteristics (Morrell 1995; Sosinski and Douches 1996). Morphological traits used for this morphological description are generally based on the set of 50 characteristics used by UPOV. These characteristics include colour and morphology of the flowers (Figure 1a), leaves (Figure 1b) and tubers (Figure 1d), the growth habit of the whole plant and the morphology of the sprouts from tubers sprouted under diffused light conditions (Figure 1c). These characteristics are listed in appendix 1 (UPOV; Jansen van Rensburg 1997). The morphological descriptions are simple, relatively easy to apply and need no equipment like thermocyclers or chemicals. The morphological descriptions are very effective but it has limitations. These limitations are subjectivity in the person performing the analyses and assessment of the characteristics, the influence of the environment and management practices on the characteristics. Only a limited number of morphological traits are fairly stable over all environments, and several months may be required to observe the distinguishing characteristics as the expressions of certain characteristics are confined to a particular stage of development, such as flowering. Field trails are necessary for the visual evaluation, these field trails are time consuming and subjective as mentioned earlier (Newbury and Ford-Lloyd 1993; Morell *et al.* 1995; Staub *et al.* 1996). It is furthermore very difficult to distinguish between closely related potato cultivars on morphological grounds alone, due to limited diversity in distinguishing characteristics. Some cultivars, like Mnandi and Astrid look morphologically almost identical and can only be distinguished by experts if the plants from both cultivars are grown next to each other under the same environmental conditions (A Visser and I Vorster personal communication). These limitations have led to the consideration of alternative techniques like cytogenetical analysis and molecular markers.

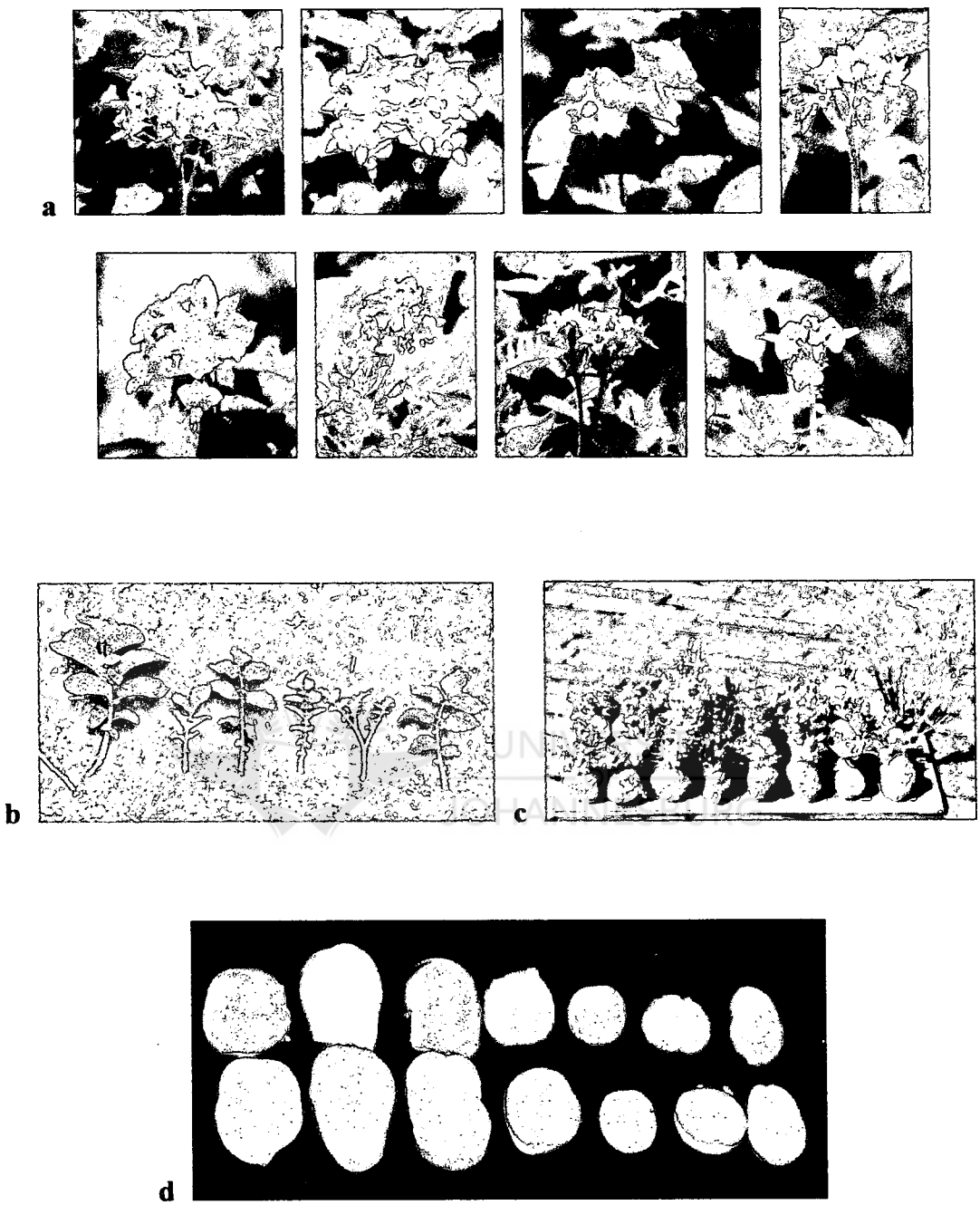


Figure 1.1: Characteristics used for the morphological description of potatoes. a) Flowers of eight different cultivars and breeding lines. Note the difference in colours and morphology of the flowers and inflorescence. b) Leaves of different cultivars and breeding lines showing difference in overall morphology as well as in the morphology of the leaflets. c) Diffused light sprouts of different cultivars. Note the marked variance between cultivars and the lack of variance within a cultivar. d) Tubers of various cultivars showing the difference in flesh colour and colour of the tuberskin.

1.3 Molecular markers

Walton (1993) describes molecular markers as any measurable chemical or molecular characteristic that is inherited in a simple Mendelian way. These markers are used to generate genetic fingerprints unique to a specific genotype. Also, for markers to have a value in fingerprinting, they must be reliable, practical and validated in independent laboratories (Smith and Smith 1992, in Xu *et al.* 1995), especially if they are to be used in patent right disclosures and in legal cases of varietal infringement. Isozyme analyses, restriction fragment length polymorphisms (RFLPs) and random amplified polymorphisms (RAPDs) are molecular markers commonly used to distinguish between different cultivars in a wide variety of crops and ornamentals.

1.3.1 Isozymes and Allozymes

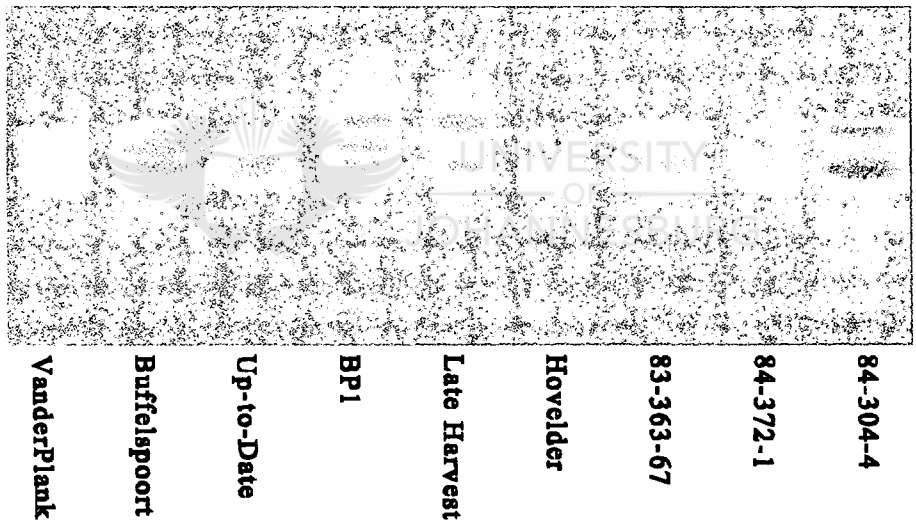


Figure 1.2: The banding pattern that illustrates the use of esterase to distinguish between nine potato cultivars. Note the difference in amount and relative position of bands (N Steyn personal communication).

The analyses of secondary metabolites and enzymes, such as **isozymes** and **allozymes**, have been used extensively to distinguish between different cultivars in various taxa (Chaparro *et al.* 1994; Becker *et al.* 1995; Maaß and Klaas 1995; Morell *et al.* 1995; Simonsen and Heenen 1995; Vicaro *et al.* 1995). Figure 1.2 shows a gel, using the esterase isozyme to distinguish between nine potato cultivars (N Steyn personal communication). However, the use of isozyme analyses is restricted to those

plant taxa that produce a suitable range of metabolites that can be rapidly analysed and which are able to distinguish between cultivars. A limited number of polymorphic loci are detected and visualised by this analysis. These secondary metabolites should ideally show limited response to environmental conditions or management practices, as well as type of tissue used (Demeke *et al.* 1993; Newbury and Ford-Lloyd 1993; Morell *et al.* 1995; Staub *et al.* 1996).

Molecular markers using genomic DNA are becoming very popular to distinguish between different plant cultivars because each individual's DNA is unique. The DNA markers allow direct comparison of the genetic material (genotype) and not the gene product or phenotype. Thus environmental factors, the type or developmental stage of tissue used, did not influence them because the same DNA is present in every living cell of the plant. It also takes considerably less time and is less labour intensive to perform than using morphological traits (Demeke *et al.* 1993; Newbury and Ford-Lloyd 1993; Morell *et al.* 1995).

1.3.2 Restriction fragment length polymorphisms

Restriction fragment length polymorphisms (RFLPs) were one of the earliest DNA profiling techniques used to study plant variation and it has been used extensively to distinguish between different cultivars. This technology uses restriction enzymes to digest DNA, extracted from plant cells, into small fragments of different lengths. These fragments are separated by gel electrophoresis, transferred to a membrane and detected by radioactive-probes that hybridise to specific DNA sequences. There is often a difference in the fragment lengths when two or more genotypes are compared, using the same DNA probe. This is called a polymorphism, which is the result of the loss of formation of a restriction site due to the occurrence of mutations and rearrangements in the DNA. These polymorphisms are inherited in a Mendelian fashion (Newbury and Ford-Lloyd 1993; Walton 1993; Morell *et al.* 1995, Staub *et al.* 1996). RFLPs have been used, with great success, to distinguish between different potato cultivars (Görg *et al.* 1992). However, RFLPs have various

disadvantages. It is time consuming and requires sometimes DNA sequence information as well as the development of suitable DNA probes. Radioisotopes are used and relative large amounts of DNA (5-10 μg) are needed (Tingey *et al.* 1993; Demeke *et al.* 1993; Rafalski and Tingey 1993).

1.3.3 Minisatellites or variable number of tandem repeats

This technique is similar to RFLPs except that the probes are designed to hybridise with repeated DNA sequences. In animals, DNA probes that hybridise to repetitive DNA sequences arranged in tandem repeats (0.1 - 20 kb long), known as "variable number of tandem repeats" (VNTRs) or "minisatellites", have proven highly informative (Beyerman *et al.* 1992; Morell *et al.* 1995). Nybom *et al.* (1990) showed VNTRs could be used to fingerprint plant cultivars. Perhaps the most promising plant DNA fingerprints are produced using oligonucleotide probes containing simple repetitive DNA patterns such as $(GATA)_n$, where n is 2-5.This approach has been used to produce species-specific fingerprints in a range of economical important genera including *Triticum*, *secale*, *Hordeum*, *Beta*, *Brassica* and *Nicotiana* (Beyerman *et al.* 1992, Morrell 1995).

1.3.4 The Polymerase Chain Reaction

The development of the polymerase chain reaction (PCR), by Kary B. Mullis in 1985, led to an evolution in molecular genetics and the development of a whole generation of new molecular markers (KB Mullis, US patent 4,683,195, July 1987; US patent 4,638,202 July 1987) (Innes and Gelfand 1990; Saiki *et al.* 1985).

PCR led to the development of a range of new molecular makers and is based on the exponential, enzymatic, *in vitro* amplification of DNA fragments. Starting from very low amounts of template DNA (mostly in the nanogram range) millions of copies of a target DNA fragment are produced and visualised by staining (mostly silver staining or ethidium bromide) or autoradiography. The PCR is characterised by its high speed, selectivity and sensitivity (Weising *et al.* 1995).

Three temperature-controlled steps can be discerned in a typical PCR (Figure 1.3). These three steps, **denaturating**, **annealing** and **elongation** are repeated in a series of 25 to 50 cycles. The PCR reaction mixes consist usually of the following ingredients:

1. A buffer, usually containing Tris-HCl, KCl and $MgCl_2$
2. A thermostable DNA polymerase which adds nucleotides to the 3' end of a primer annealed to single strand DNA
3. Four deoxynucleotides (dNTPs: dATP, dCTP, dGTP and dTTP)
4. One or two oligonucleotide primers
5. Template DNA

This reaction mix is overlaid with mineral oil to prevent evaporation and condensation of water onto the walls and sides of the tube (Weising *et al.* 1995).

The choice of primer influences the selectivity of the reaction. The primers are single stranded, synthetic DNA with sequences complementary to template sequence flanking the fragment(s) of DNA to be amplified. To allow for exponential amplification, the primers must anneal in opposite directions on the opposite strands, so that the 3' ends face the target sequence and each other (see 2 in Figure 1.3). The amplification is most efficient if the two primer binding sites are not further apart than about 4kb, although longer fragments can be obtained under optimal conditions (Weising *et al.* 1995).

The principle of the cycling reaction and exponential amplification is outlined in Figure 1.3. In the first step, the **denaturating step** (1 in Figure 1.3), the double stranded DNA is made single stranded by raising the temperature to ± 94 °C. In the second step, the **primer annealing step** (2 in Figure 1.3), the temperature is lowered to about 35 to 65 °C (depending on primer sequence and experimental strategy),

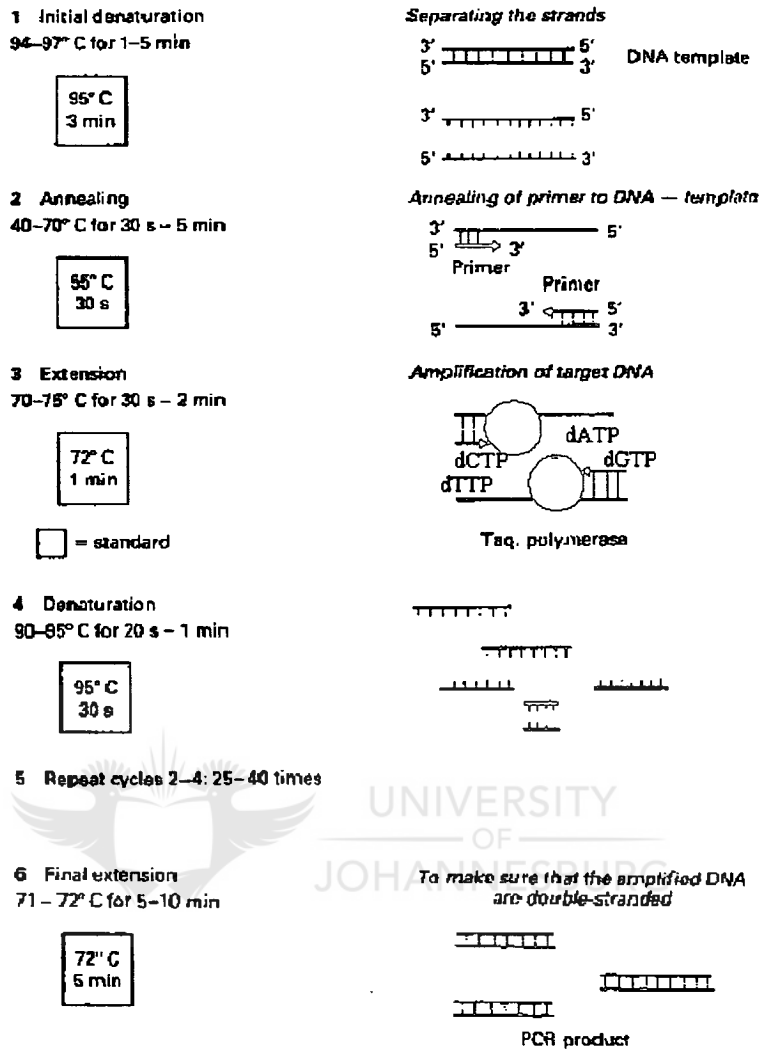


Figure 1.3: Schematic representation of the polymerase chain reaction. More detail in text (Weising *et al.* 1995)

allowing the primer to anneal to the single strand template DNA. Primers will preferably hybridize to binding site(s) that are identical or highly homologous to their nucleotide sequence although some mismatches, usually at the 5' end, are allowed. Gallego and Martinez (1996) found that the denaturating and annealing steps could be reduced from 1 min, as suggested by Williams *et al.* (1990), to 15 and 30 seconds respectively without a change in the profile. For the third step, the **elongation or extension step** (3 in Figure 1.3), the chosen temperature for the optimal activity of

the thermostable polymerase, i.e. usually 72 °C. The thermostable polymerase now extends the 3' ends of the primers, in the template DNA-primer hybrids, towards the other primer binding site. Since this whole process happens at both primer-annealing sites on both DNA strands, the target fragment between the two annealing sites is completely replicated (Weising *et al.* 1995).

In the second cycle, the resulting two double stranded DNAs are again denaturated (4 in Figure 1.3), and both the original strands and the product strands now act as template, thus four templates instead of two. Repeating these three-step cycles 25 to 50 times (5 in Figure 1.3) results in the exponential amplification of the target DNA between the two 5' ends of the two primer binding sites. Other longer products are also generated but since these longer fragments are only linearly amplified, they are negligible in the final product (Weising *et al.* 1995).

The specificity of the PCR is influenced by several parameters (Weising *et al.* 1995):

- The temperature profile of the thermocycler
- The annealing temperature which is usually as high as possible to prevent nonspecific amplification
- The activity, amount and type of thermostable polymerase
- The primer concentration
- The template DNA concentration
- the Mg^{2+} concentration
- Presence of additional chemicals like dimethylsulphoxide

One of the main reasons for the versatility of PCR is that almost any primer sequence and length can be chosen, depending on the purpose of the study. Two broad experimental applications can be identified in PCR. Two specific primers can amplify a particular piece of DNA. The primers for these specific PCRs are designed based on sequence information. This strategy is useful for the analysis of transferred genes in transgenic organisms or for gene analysis. On the other extreme, anonymous DNA

sequences can be amplified by **arbitrary primers**. This strategy is usually followed in genome mapping and DNA fingerprinting. Between these two extremes are many possible semispecific PCRs using **semispecific** primers (Weising *et al.* 1995).

1.3.5 Random Amplified Polymorphic DNA

Williams *et al.* (1990), Welsh and McCleland (1990) and Caetano-Anollés and *et al.* (1991) published another approach to PCR; the amplification by PCR of discrete loci or DNA fragments with single, arbitrarily chosen, synthetic oligonucleotide primers of known sequence. Amplification of these loci will proceed if these primers bind sufficiently close on opposite strands to allow amplification of the intervening DNA and a range of DNA fragments is amplified. No prior sequence information is required. Williams *et al.* (1990) proposed the term "random amplified polymorphic DNA", or RAPD as an acronym that is widely used nowadays.

1.3.6 Sequenced Characterized Amplified Regions

Another PCR-based marker, derived from RAPD markers, sequenced characterized amplified regions (SCARs) (Paran and Michelmore 1993) overcome some of the drawbacks of RAPDS. A SCAR is a genomic DNA fragment at a single genetically defined locus that is identified by PCR amplification using a pair of specific oligonucleotide primers. SCARs are derived by cloning and sequencing the two ends of RAPD amplification products. These sequences are then used to design sets of specific longer oligonucleotide primers. Amplification of target DNA using this set of primers will result in the reproducible amplification of specific loci when high annealing temperatures are used. (Paran and Michelmore 1993). SCARs are essentially similar to sequenced tagged sites (STS) proposed by Olson *et al.* (1989) according to Kesseli *et al.* (1993)

1.4 The aims of this Study

- to identify arbitrary RAPD primers which can be used for the identification of the two cultivars, Astrid and Mnandi and to identify polymorphisms unique to one of the two cultivars which can be used to develop SCAR markers (Chapter 2);
- to excise and purify the identified DNA fragment from agarose gel and clone the DNA fragment into a suitable vector in order to sequence the DNA fragment (Chapter 3);
- Design of SCAR primer sets (Chapter 4) and determine usability of the SCAR primer sets (Chapter 4).

Chapter 2

Random amplified polymorphic DNA analysis

2.1 Introduction

Standard PCR methodology forms the base of the amplification of anonymous polymorphic DNA fragments except that usually one primer is used. The term "random" in random amplified polymorphic DNA might be misleading. It might imply that the primers are chosen randomly which is not the case. Although primers are of arbitrary sequence, they are carefully chosen to have a GC content of ± 50% with no internal complementarity. The term might also imply that the technique is random, unreproducible and unreliable, which is definitely not the case, even if the technique is criticized by various authors (Morell *et al.* 1995).

Other protocols, closely related to RAPDs exist: "arbitrary primed PCR" (AP-PCR) described by Welsh and McCleland (1990), "DNA amplification finger printing" (DAF) described by Caetano-Anolles *et al.* (1991) and RAPDs separated via denaturating-gradient-gel electrophoresis (RAPD-DGGE) described by Dweikat *et al.* (1993). These approaches differ mainly in length and concentration of primer used in the amplification reactions and in the complexity of the generated profiles (Caetano-Anollés 1996). Highly complex fingerprints are generally produced when using high primer-to-template ratios (e.g. DAF; Caetano-Anollés *et al.* 1991) or when targeting abundant motifs either artificially produced in a template (e.g. amplified fragment length polymorphism or AFLP; Vos *et al.* 1995) or ubiquitous in a genome (e.g. anchored microsatelite primed PCR or AMP-PCR) or RNA population (e.g. differential display reverse transcription PCR or DDRT-PCR). Again, all these techniques are relatively fast and simple to perform, use only a few reagents and require equipment amendable to almost any laboratory, features which make them

especially useful to the novice or those with limited resources (Caetano-Anollés 1996). RAPDs became popular because of its low cost, simplicity and ease of use in a modestly equipped laboratory (Demeke *et al.* 1993; Newbury and Ford-Lloyd 1993; Rafalski and Tingey 1993; Walton 1993; Burger 1995; Weising *et al.* 1995).

Simultaneous use of two different primers is also possible and additional information is acquired (Caetano-Anollés *et al.* 1991; Welsh and McCleland 1990; Wiliams *et al.* 1990 and Wiliams *et al.* 1993). Although nothing is known about the identity and the sequence of the amplification product(s), its absence or presence can be highly informative for various purposes like evaluation of genetic relatedness and diversity as well as for identification. RAPD uses primers consisting of 10 nucleotides and a GC content of at least 50%. Primers with lower GC content usually do not yield amplification products, because the bonds between guanine and cytosine consist of three hydrogen bonds while the bonds between adenine and thymine consist of only two. A primer-DNA hybrid with a low GC content will probably not withstand the temperature, 72°C, at which DNA polymerisation take place and melt before the thermostable polymerase has started polymerisation. The choice of the nucleotide sequence is however still arbitrary. Thus, no prior knowledge of DNA sequences is needed and primers can be used universally for prokaryotes and eukaryotes. Very low amounts of DNA (0.1-400ng) are needed with RAPDs (Weising *et al.* 1995).

To obtain amplification with only one primer, two identical or highly similar target sequences in close vicinity of each other is needed. One site must be on the one strand and the other on the opposite strand, in the opposite direction with their 3' ends facing each other. The distance between these two sites should not exceed a few kilobasepairs, since smaller fragments are more efficiently amplified than the larger ones. The RAPD technique uses two random primers, of usually 10 base pairs, which are able to hybridise to opposite strands and flank the target DNA area to be amplified from the template DNA. These hybridisation events might occur at several

loci. When two hybridisation events occur close enough to one another the DNA segment between the two primers is enzymatically amplified by thermostable DNA polymerase. A repetitive series of cycles involving template denaturation, primer annealing and extension of annealed primers result in the exponential accumulation of amplified DNA fragments of various lengths. The ends of these fragments are defined by the 5' ends of the primers. These amplified fragments are separated by gel electrophoresis and visualised by staining with ethidium bromide to reveal the DNA fragment pattern. Polymorphisms are detected by either the presence or the absence of certain bands. Thus, these RAPD markers are dominant and are inherited in a simple Mendelian fashion (Walton 1993; Newbury and Ford-Lloyd 1993; Nybom 1994; Burger 1995, Boehringer Mannheim 1995; Weising *et al.* 1995) The use of RAPDs to distinguish between two genotypes is outlined in Figure 2.1.

Polymorphisms observed by the RAPD analysis may result from differences in the nucleotide sequence at the priming site or by structural rearrangements between or within the binding sites, like point mutations, insertions, deletions and inversions which may alter the length of the amplified fragment or prevent the successful amplification of target DNA (Williams *et al.* 1990; Tingey *et al.* 1993; Demeke *et al.* 1993; Newbury and Ford-Lloyd 1993; Paran and Michelmore 1993; Nybom 1994). RAPD amplification can be initiated from genomic sites that do not match the primer sequence perfectly. Sequencing the products of the RAPD analysis does not provide information on these mismatches within the RAPD primer sequences because the majority of templates will be identical in sequence to the primer after a few cycles of amplification (Paran and Michelmore 1993).

The number of fragments theoretically expected from one primer, annealing with 100% homology, can be calculated from the primer length and the complexity of the target genome, supposing that all nucleotides are present in equal proportions. Williams *et al.* (1993) gave the equation $b=(2000x4^{-2n})xC$, where b is the expected

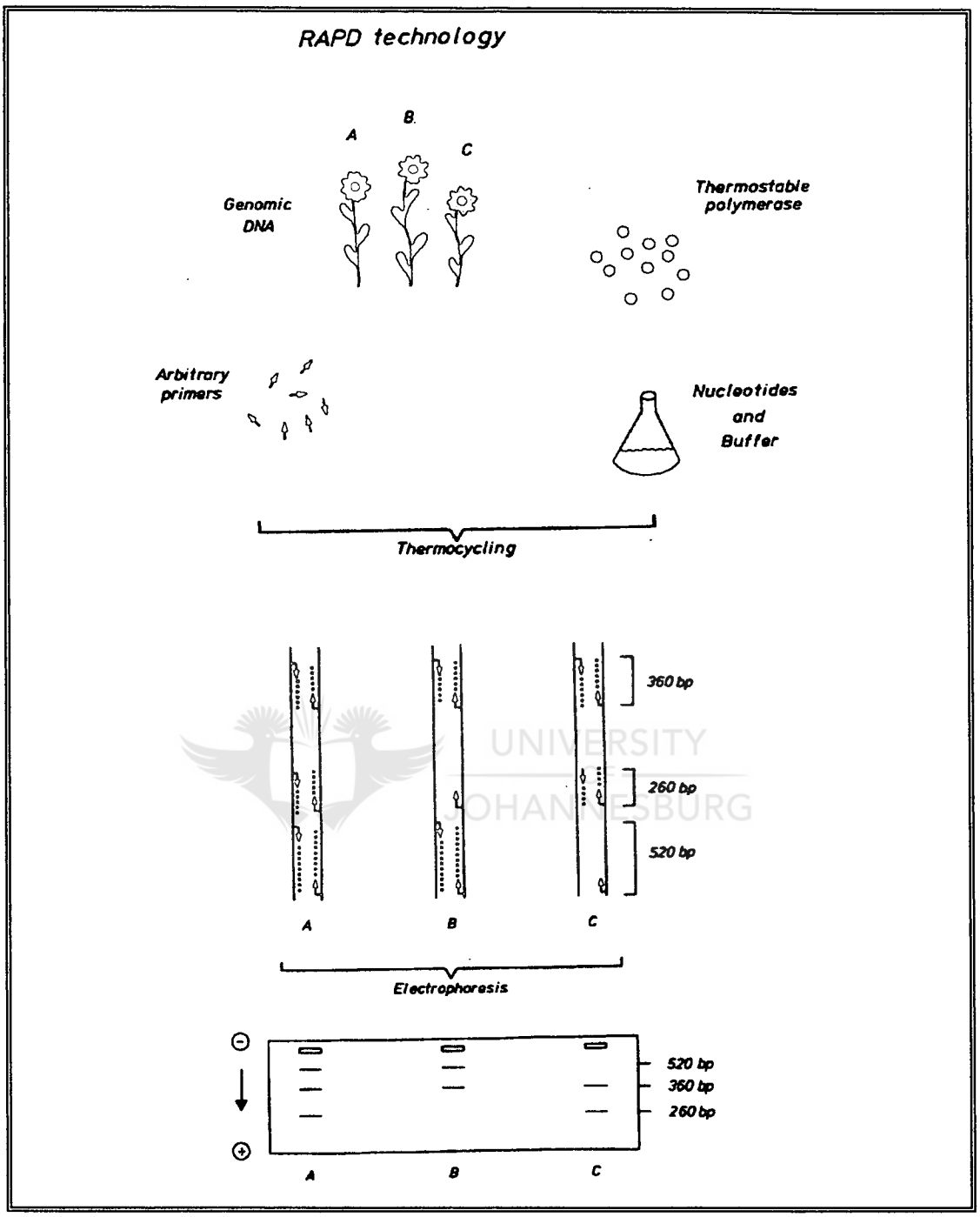


Figure 2.1: Strategy of RAPDs to distinguish between three genotypes. Genomic DNA of the three genotypes (A, B, and C), primers of arbitrary sequence, thermostable polymerase and a suitable buffer combined in reaction tubes and subjected to PCR. The arbitrary primers anneal to target sequences in the template DNA. Amplification takes place if two primers anneal in opposite directions and at suitable distance from each other. The different amplification products are separated by electrophoresis and visualised by ethidium bromide staining. The different genotypes have different RAPD profiles as seen in figure (Weising *et al.* 1995).

number of fragments per primer, n is the primer length in nucleotides, and C is the complexity of the genome of the organism, i.e. genome size en base pairs per haploid genome. There are, however, strong indications that the number of fragments per primer is largely independent of genome complexity. Plants with large genomes like onion and conifers do not exhibit more complex RAPD fragment patterns than plants with comparatively small genomes, like tomato and *Arabidopsis* (Carlson *et al.* 1991; Klein-Lankhorst *et al.* 1991 and Reiter *et al.* 1992; Wilkie *et al.* 1993 and Weising *et al.* 1995). The ploidy level of a plant does not appear to influence the number of fragments per primer (Wolf and Peters-van Rijn 1993). This obvious independence of the RAPD fragment number from genome size may be explained by mismatch and primer competition (Weising *et al.* 1995). The frequency of finding RAPD polymorphisms has been shown to be 0.3 polymorphisms per primer in *Arabidopsis thaliana*, 0.5 polymorphisms per primer in soybean, one polymorphisms per primer in corn, 2.5 polymorphisms per primer in *Neurospora crassa* (Tingey *et al.* 1993) and 5.5 polymorphisms per primer in potato (Milbourne *et. al.* 1997).

The ability of RAPDs to detect variation is quite variable depending on the taxa and primer used. Some primers result in monomorphic DNA fingerprints and others show a high level of polymorphism. Usually several primers are used consecutively in an investigation to obtain sufficient levels of variability (polymorphisms) (Nybom 1994). Using 16 potato cultivars and 14 primers, Milbourne and *et al.* (1997) found 114 scorable bands, an average of 8.36 per primer of which 65.8% were polymorphic. Thus, they were able to discriminate successfully among all 16 cultivars using 14 RAPD primers.

RAPDs have many advantages. RAPDs are relatively inexpensive compared to other methods, results can be obtained within a day and extremely small amounts of DNA (0.1 to 400 ng) are required. No DNA sequence information is required and large numbers of samples can be handled with ease. As with other DNA markers, the

analysis cannot be influenced by the environment or the physiological and developmental stage of the plant material. These advantages, together with the relative ease and speed of the technical procedure, have led to increased popularity of RAPDs in cultivar identification (Williams *et al.* 1990; Tingey *et al.* 1993; Demeke *et al.* 1993; Newbury and Ford-Lloyd 1993; Nybom 1994). According to Walton (1993) it also appears that more polymorphisms can be detected with RAPDs than with RFLPs.

Cultivar identification can be achieved with great accuracy from RAPDs, especially in materials characterised by high levels of genetic variation between cultivars and no variation within. This is the case in most vegetatively propagated cultivars from outcrossing species like fruit and berry crops as well as potatoes (Nybom 1994; Sosinski and Douches 1996). Sosinski and Douches (1996) are also of the opinion that, if it is assumed that no mechanical mixing occurs, the need for extensive intra cultivar sampling is eliminated when working with vegetatively propagated plants. However, they did suggest further study to take somaclonal variation into account, but they were able to distinguish Burbank from its russet skintype mutant, Russet Burbank. Demeke and co-workers (1993) were able to discriminate between clonal variants of Russet Burbank and Viking but not between those clonal variants from Norland, Superior and Norgold Russet. RAPDs are also able to distinguish between closely related varieties for example wheat varieties which presumably share more than 94% of their genomes (Dweikat *et al.* 1993).

RAPDs have been used in various dicotyledonous as well as monocotyledonous taxa for cultivar identification. A few examples are potatoes (Demeke *et al.* 1993; Hosaka *et al.* 1994; Demeke *et al.* 1996; Sosinski and Douches 1996; Milbourne *et al.* 1997), rye (Iqbal and Rayburn 1994), mango (Schnell *et al.* 1995), rice (Wang *et al.* 1994; Virk *et al.* 1995), blackcurrant (Lanham *et al.* 1995), *Musa spp.* (Howell *et al.* 1994), buffalo grass (Wu and Lin 1994), rabbiteye blueberry (Aruna *et al.* 1995), *Annona*

spp. (Ronning *et al.* 1995) and grapevine (Moreno *et al.* 1995). Obara-Okey and Kako (1998) are also of the opinion that RAPD might be useful to identify mislabelled *Cymbidium* germplasm. Table 2.1 summarises the outcome of RAPD studies for a range of species (modified from Morell *et al.* 1995).

Table 2.1: Selected example analyses of plant varieties using RAPD to show the wide application of RAPDs for cultivar identification (modified and extended from Morell *et al.* 1995).

Plant group	No. of Varieties	No. of Primers	No. of Bands	Reference
Allium cepa	7	6	nd	Wilkie *et al.* (1993)
Barley	17	8	nd	Francisco-Ortega *et al.* (1993)
Brassica oleracea cv Capitata	4	25	200	Kresovich *et al.* (1992)
B. oleracea cv Capitata	8	25	200	Kresovich *et al.* (1992)
Broccoli	14	4	37	Hu and Quiros (1991)
Buffalograss	25	2	32	Wu and Lin (1994)
Cauliflower	12	4	40	Hu and Quiros (1991)
Celery	19	28	309	Yang and Quiros (1991)
Cymbidium spp.	36	15	132	Obara-Okey and Kako (1998)
Oats	17	8	nd	Hu and Quiros (1991)
Papaya	10	11	102	Stiles *et al.* (1993)
Rose (*Rosa* spp.)	25	2	163	Gallego and Martinez (1996)
Solanum tuberosum	16	14	114	Milbourne *et al.* (1997)
Solanum tuberosum	28	12	128	Demeke *et al.* 1996
Solanum tuberosum	46	10	43	Sosinski and Douches (1996)
Theobroma cacoa	10	8	9	Wilde *et al.* (1992)
Theobroma cacoa	25	9	75	Russel *et al.* (1993)

RAPDs are used extensively and with an array of different aims in potatoes. RAPDs were used to distinguish between different cultivars by Demeke *et al.* (1993), Hosaka *et al.* (1994) and Demeke *et al.* (1996). The genetic divergence among potato

varieties with different levels of co-ancestry was studied by Demeke *et al.* (1996). Paz and Veilleux (1997) used RAPDs to estimate the genetic distance between *Solanum phureja* monoploids and diploid heterozygous pollinators. They come to the conclusion that RAPDs may facilitate the identification of diverse parents to maximise the expression of heterosis in *S. phureja* hybrids. Quiros et al. (1993), and Hosaka and Hanneman (1994) investigated the segregation of RAPD markers in diploid and tetraploid potatoes. RAPD markers were used by Xu *et al.* (1993) as well as Baird *et al.* (1992) to screen inter- and intra-specific somatic hybrids. Spooner *et al.* (1996) compare RAPDs and three other molecular markers, chloroplast DNA, isozymes and RFLPs, to measure the relationship among wild potato and relatives from the *Etuberosum* section of the genus, *Solanum.*

Marker-based cultivar identification techniques must be objective, reliable, cost effective, technically simple and rapid to be accepted by the seed industry and certification schemes (Sosinski and Douches 1996). However, Sosinsky and Douches (1996) are of the opinion that RAPD analysis, if it can be standardised across laboratories, might be used in combination with key morphological characteristics, like skin colour, or in conjunction with limited isozymes polymorphisms, to discriminate between different potato cultivars. Information obtained using the RAPD technique would be acceptable under the Australian Plant Breeders Rights (PBR) Act provided that the application furnished comparative data on the claimed novel variety and most similar varieties of common knowledge. The latter are standard requirements for all morphological, biochemical or DNA-based methods that seek to satisfy the distinctness, uniformity and stability (DUS) criteria for PBR registration (Morell *et al.* 1995).

RAPD analysis is one of the most commonly used DNA fingerprinting techniques and it has been applied successfully by scientists worldwide, but several properties of this technique have been criticized, including its reproducibility. Considerable care

has to be taken in maintaining consistent reaction conditions optimised for reproducible DNA amplification, due to the stochastic nature of the DNA amplification reaction. The template DNA concentration and quality (template-primer ratios), magnesium concentration, primer annealing temperature, primer length and primer base composition, all have an influence on the RAPD reaction (Kesseli *et al.* 1993; Tingey *et al.* 1993; Caetano-Anollés 1996; Gallego and Martinex 1996). Jones *et al.* (1997) found that it was difficult to reproduce RAPD results between nine European laboratories. Schierwater and Ender (1993) as well as Gallego and Martinez (1996) found that the source of the heat-stable DNA polymerase, used in the RAPD reaction, influenced the RAPD profile. Gallego and Martinez (1996) also found that the annealing temperature. as well as target DNA concentration. influence the RAPD profiles. There is some difficulty with the scoring of RAPD bands as well. Two bands/amplification products might have similar electrophoretic mobilities or may be overlapping. Another difficulty is that certain faint bands are amplified inconsistently in the F_2 offspring (Kesseli *et al.* 1993). Boiteux *et al.* (1999) also found that the DNA extraction method and tissue type used influenced the RAPD reactions.

Kesseli *et al.* (1993) summarised the disadvantages or shortcomings of RAPD analysis as follows:

1. Low reliability and repeatability
2. Dominant nature of alleles
3. Low allele number per locus
4. Lack of homology between related taxa
5. Lack of specificity to unique regions in the genome

All these limitations, sensitivity to reaction conditions, reproducibility and formation of artificial non-parental bands can be overcome to a certain extend by selecting the right polymerase, finding wide reproducible windows in yield, specificity and fidelity

of amplification of each reaction component, eliminating impurities in DNA preparations and increase fingerprint complexity to decrease competition artifacts. Non-parental bands are usually rare and result from heteroduplex molecules formed between allelic sequences that differ in length. The problem with band comigration could be controlled by increasing the resolving power of nucleic acid separation and/or confirming band identity using Southern or Northern hybridization (Caetano-Anollés 1996).

Although several properties of RAPDs have been criticized, it is still one of the most commonly used DNA fingerprinting techniques and it has been applied successfully by scientists worldwide. In this part of the study we aim to identify RAPDs which are able to discriminate between the potato cultivars, Astrid and Mnandi by showing a polymorphism unique to either Mnandi or Astrid. The nature of this polymorphism must be the presence of a fragment in only one of these two cultivars, because this fragment will be used for the development of a SCAR marker.

2.2 Material and Methods

2.2.1 Plant material

Astrid and Mnandi are two commercially important cultivars in South Africa. In the 1995 season, Astrid accounted for 3% of the seedpotatoes certified in the Potato Producers Organization's (PPO's) certification scheme, while Mnandi contributed 0.56% (PPO 1995). The figures for 1997 are 0.84% and 0.75% for Astrid and Mnandi respectively, see Figure 2.2 (PPO 1998). Being a young and upcoming cultivar, it is expected that Mnandi's economic importance will continue to grow in the future since it shows durable resistance against early and light blight and wind damage and good tolerance to heat and drought stress, characteristics which make Mnandi particularly attractive for the small scale farmers (Visser 1997; Visser 1999; A. Visser and I. Vorster, personal communications).

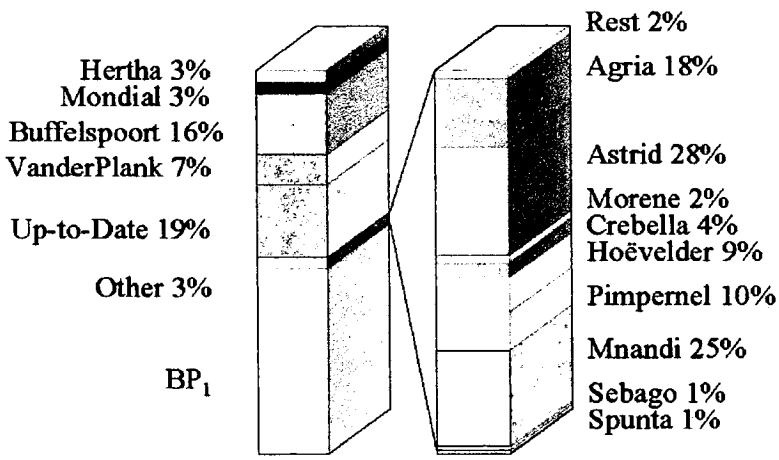


Figure 2.2: The relative marketshare of different registered cultivars in South Africa, illustrating the importance of Astrid and Mnandi (PSA 1998).

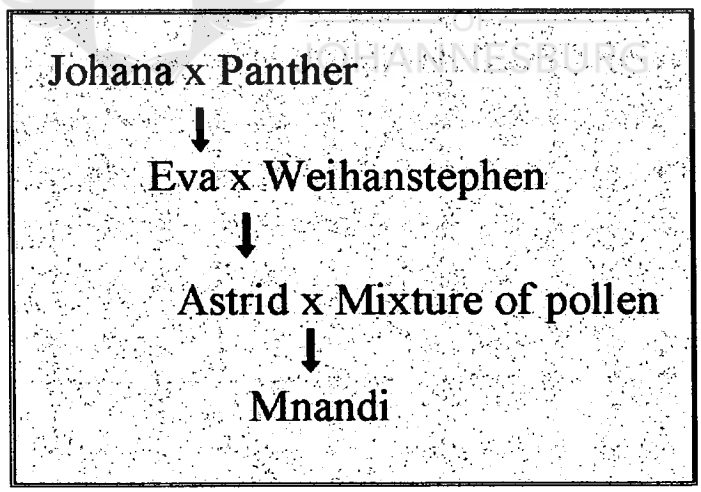


Figure 2.3: The breeding history of Astrid and Mnandi.

The two cultivars are genetically closely related, Astrid being the maternal parent of Mnandi (see Figure 2.3), and morphological very similar - which makes identification on morphological grounds very difficult. The fraternal parent is unfortunately unknown, because the initial pollination was made as part of a policross (A. Visser,

personal communication). Both Mnandi and Astrid are maintained *in vitro* in the Potato Genebank of the ARC-Roodeplaat. The *in vitro* material is cultured on MS media (Murashigi and Schoog 1962) in growth rooms at 25 ± 2°C (Jansen van Rensburg 1995).

Ten of the economically most important cultivars (including Astrid and Mnandi) were used in the screening of RAPD primers. Their RAPD profiles were taken into account in the choice of the RAPD products to clone. These cultivars are Mnandi, Astrid, VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel and Agria; the relative economic importance of these cultivars are seen in Figure 2.2.

2.2.2 DNA Extraction

DNA was extracted from plant material using an adaptation of the method developed by Edwards *et al.* (1991). Leaf material (approximately 1 mg), from *in vitro* plantlets, were collected in clean microfuge tubes. A pinch of carborundum powder was added and the leaf material was grounded thoroughly in the microfuge tube using **clean** glass grinders. 400µl prewarmed (60°C) Supaquick buffer (200 mM Tris-Cl; 250 mM NaCl; 25 mM EDTA; 0.5 % SDS) was added and gently mixed. The microfuge tubes were incubated at 60°C in a waterbath or heating block for 10 minutes. Equal volumes of chloroform:isoamyl alcohol (24:1) were added and mixed for approximately 5 minutes by inverting the tubes several times. The tubes were not vortexed. The microfuge tubes were then centrifuged at 10 000 x g for 10 minutes at room temperature. 350 µl of the supernatant (aqueous phase) was carefully transferred into clean microfuge tubes. Ice-cold isopropanol (0.6 volumes) were added and gently mixed by inverting. The microfuge tubes were left at -20°C for at least 30 minutes to allow for the precipitation of the DNA. The microfuge tubes were centrifuged at 10 000 x g for 10 minutes (4°C) and decanted carefully to drain. The DNA pellets were washed by adding 500 µl of ice-cold 70 % ethanol. The microfuge tubes were centrifuged again at 10 000 x g for 10 minutes (4°C) and decanted

carefully to drain. The DNA pellets were resuspended in 100 µl ddH_2O and shaken gently. The DNA was left overnight at 4°C to dissolve thoroughly and stored in a freezer at -20°C. The DNA concentrations were determined using a Sequoia-Turner fluorometer. The DNA samples were compared to a sample of 100 $ng.\mu l^{-1}$ calf thymus DNA. The DNA was diluted in ddH_2O to a final concentration of 2 $ng.\mu l^{-1}$.

2.2.3 RAPD analyses

Decamer primers provided by Operon Technologies (OPA and OPH), Almeda, CA, and the University of British Colombia (UBC), Vancouver, Canada, as well as four primers described by Demeke (1993), originally also obtained from the UBC, was used in the analyses (Table 2.2.) Amplification was carried out in 10 µl reaction

Table 2.2: The decamer primers used in this study.

OPA	OPH	Demeke[1]	UBC		
OPA-02	OPH-01	DEM 116	UBC 501	UBC 531	UBC 559
OPA-03	OPH-02	DEM 131	UBC 502	UBC 533	UBC 560
OPA-04	OPH-03	DEM 153	UBC 503	UBC 534	UBC 561
OPA-05	OPH-07	DEM 184	UBC 504	UBC 537	UBC 562
OPA-07	OPH-10		UBC 505	UBC 538	UBC 563
OPA-08	OPH-11		UBC 506	UBC 539	UBC 564
OPA-10	OPH-15		UBC 508	UBC 540	UBC 565
OPA-11	OPH-17		UBC 509	UBC 541	UBC 566
OPA-12	OPH-18		UBC 510	UBC 542	UBC 567
OPA-14	OPH-19		UBC 511	UBC 543	UBC 568
OPA-15			UBC 512	UBC 544	UBC 569
OPA-17			UBC 513	UBC 545	UBC 570
OPA-18			UBC 514	UBC 546	UBC 571
			UBC 515	UBC 547	UBC 580
			UBC 516	UBC 548	UBC 581
			UBC 517	UBC 549	UBC 582
			UBC 518	UBC 550	UBC 583
			UBC 519	UBC 551	UBC 584
			UBC 520	UBC 552	UBC 585
			UBC 521	UBC 553	UBC 586
			UBC 522	UBC 554	UBC 587
			UBC 523	UBC 555	UBC 588
			UBC 525	UBC 556	UBC 589
			UBC 528	UBC 557	UBC 590
			UBC 529	UBC 558	UBC 591

[1] - Demeke *et al.* 1993

volumes, containing 10 ng of template DNA, reaction buffer (Dynazyme buffer; 10 mM Tris-HCl, pH 8.8, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100), 0.5 U/µl enzyme (Dynazyme™), 1 µM decamer primer and 0.1 mM of each dNTP. The reaction volumes were overlaid with 20 µl of mineral oil to prevent evaporation during the PCR process. DNA amplification was conducted on a Hybaid thermal cycler programmed for 35 cycles of 20 sec denaturating at 94.5°C, 30 sec annealing at 37°C and 45 sec elongation at 72°C, with the initial denaturation of 90 sec at 94.5°C, and a final extension of 3 min at 72°C. The final extension time was later lengthened to 5 min to ensure easier cloning of RAPD fragments (see section 3.3.2). A negative control reaction (with no DNA added) was included in each run to check for DNA contamination of reagents and reaction mixes. The amplification products were resolved electrophoretically in a 2 % MetaPhor™ agarose gel, stained with ethidium bromide and photographed under UV light. Molecular marker VI (Boehringer, 152 bp - 2176 bp) was used to determine fragment sizes.

The RAPD analyses were conducted with DNA from Astrid and Mnandi until polymorphisms between the two were detected. The primers detecting the polymorphism were then used in RAPD reaction using Astrid and Mnandi as well as the eight other cultivars, VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel and Agria. This was done to detect polymorphisms unique to either Astrid or Mnandi.

2.3 Results

2.3.1 DNA extraction

Good quality DNA from all the cultivars were obtained. The average of three readings from the fluorometer was taken as the concentration of the sample. The concentrations range from 107 $ng.ul^{-1}$ to 5.5 $ng.ul^{-1}$, these values correspond to yields of 10 700 - 550 ng DNA per gram leaf material, see Table 2.3.

Table 2.3: Concentration of DNA samples after Supaquick DNA extractions.

Cultivar	DNA Yield (ng per mg leaf material	[DNA] $ng.ul^{-1}$
Mnandi	10 700	107.0
Astrid	550	5.5
VanderPlank	12 200	122.0
Up-to-Date	6 500	65.0
BP_1	6 500	65.0
Buffelspoort	4 150	41.5
Hoëvelder	4 100	41.0
Hertha	6 150	61.5
Pimpernel	1 150	11.5
Agria	3 150	31.5

2.3.2 RAPD reactions and profiles

A total of 102 decamer primers were tested (Table 2.1). The profiles obtained from the RAPD reactions could be classified into three groups (Table 2.4):

- those without any polymorphisms between Mnandi and Astrid (Table 2.4 and Figure 2.4a)
- those showing a polymorphism between Astrid and Mnandi, but the polymorphism was not unique (for the presence of a fragment) to either Mnandi or Astrid (Table 2.4 and Figure 2.4b)
- those showing an unique polymorphism to Mnandi (Table 2.4 and Figures 2.6 a and b)

Two unique polymorphisms were identified, both unique to Mnandi. The polymorphism obtained with OPH-15 (AATGGCGCAG) was the presence of an ±1100 bp fragment amplified only in Astrid (Figure 2.5a). This polymorphism was, however not repeatable as seen in Figure 2.5b The band, which was present in the first gel, was possibly due to mispriming which took place. The polymorphism obtained with DEM184 (CAAACGGCAC) was for the absence of a ± 150bp fragment in Astrid, see Figure 2.4b.

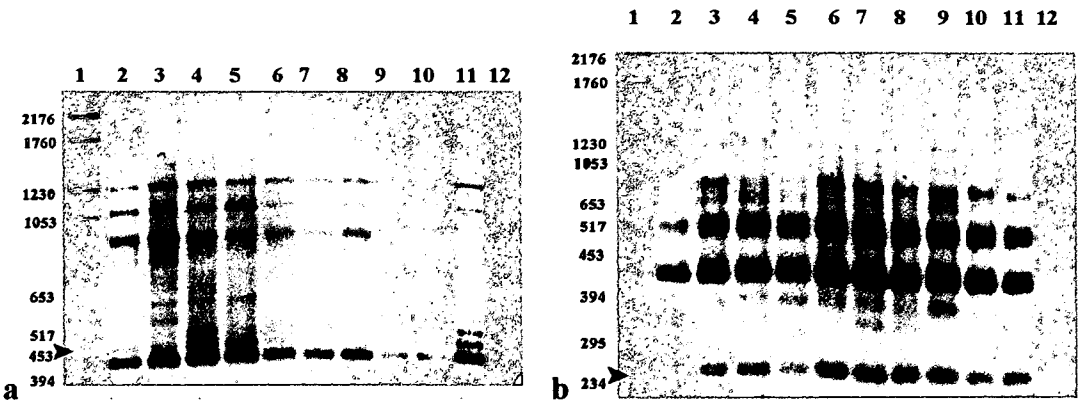


Figure 2.4: RAPD profiles obtained with the reactions conditions as described in the text. Lane 1 is molecular marker VI, lane 2 is Astrid, lane 3 is Mnandi and lanes 4 to 12 is VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel and Agria and a H_2O control respectively. (a) RAPD profile obtained with primer OPA-18. No polymorphism between Astrid and Mnandi is observed. Polymorphisms between the other cultivars are observed. An arrow indicates one of these polymorphisms. (b). RAPD profile obtained with primer Dem 184. A polymorphism of about 250 bp is observed between Astrid and Mnandi (indicated by the arrow). This polymorphism is however not unique for the presence of a fragment.

Table 2.4: Distinction between the different primers on grounds of polymorphisms.

No polymorphisms between Mnandi and Astrid					
OPA-02	OPH-03	UBC504	UBC521	UBC542	UBC565
OPA-03	OPH-07	UBC506	UBC522	UBC545	UBC566
OPA-05	OPH-11	UBC508	UBC523	UBC547	UBC569
OPA-12	OPH-19	UBC511	UBC528	UBC548	UBC571
OPA- 14	DEM116	UBC512	UBC531	UBC553	UBC580
OPA-15	DEM131	UBC513	UBC533	UBC555	UBC583
OPA-17	DEM153	UBC514	UBC534	UBC557	UBC584
OPA-18	UBC501	UBC515	UBC537	UBC558	UBC589
OPH-01	UBC502	UBC517	UBC538	UBC559	UBC590
OPH-02	UBC503	UBC520	UBC541	UBC563	
Polymorphisms between Mnandi and Astrid, but not unique					
OPA-04	OPH-10	UBC518	UBC546	UBC560	UBC585
OPA-07	OPH-15[1]	UBC519	UBC550	UBC564	UBC586
OPA-08	OPH-17	UBC525	UBC551	UBC567	UBC587
OPA-10	UBC505	UBC529	UBC552	UBC568	UBC591
OPA-11	UBC510	UBC543	UBC554	UBC570	*DEM184*[2]
OPA-17	UBC516	UBC544	UBC556	UBC581	
Polymorphisms unique to either Mnandi or Astrid					
OPH-15[1]	UBC509	UBC582			

[1] - The polymorphism obtained with OPH-15 was not reproducible (Figure 2.5 a and b).

[2] - The polymorphism obtained with DEM184 was the absence of a band for Astrid (Figure 2.6).

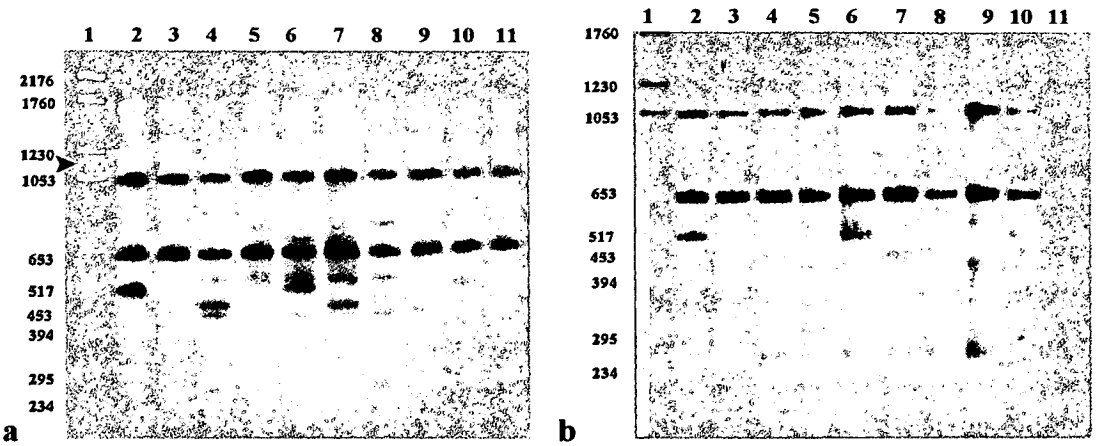


Figure 2.5: RAPD profiles obtained with the reactions conditions as described in the text. Lane 1 is molecular marker VI, lane 2 is Astrid, lane 3 isMnandi and lanes 4 to 12 is VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel and Agria and a H_2O control respectively. (a) RAPD profile obtained with primer OPH-15. A polymorphism, of about 1100 bp, unique to Astrid observed (indicated by the arrow in a) This polymorphism is not reproducible if the same primer and reaction conditions were used, as seen in (b). The polymorphic fragment of about 500 bp in Astrid was present in some of the other cultivars as well..

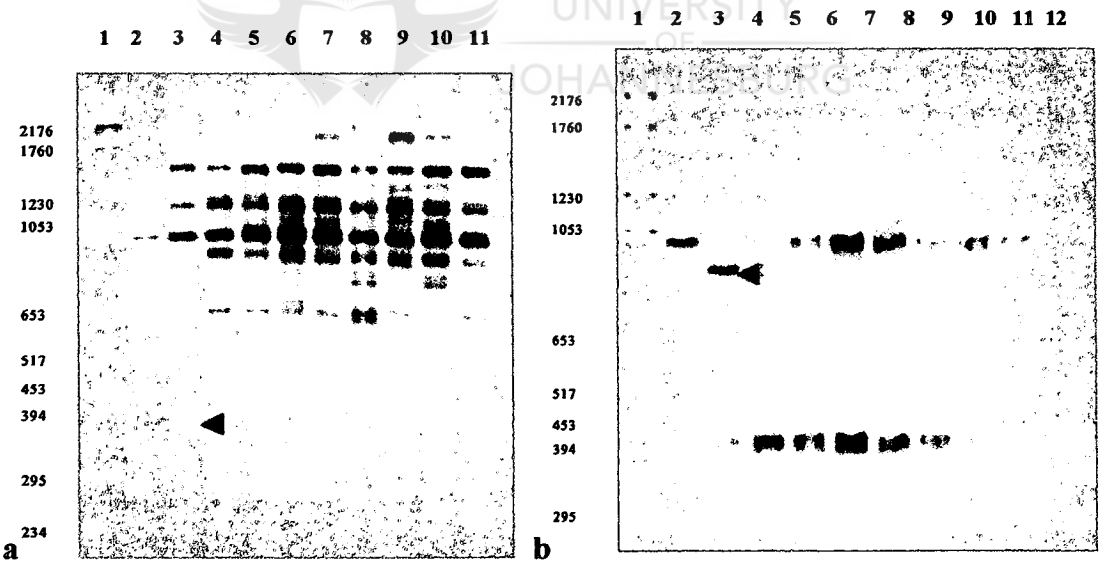


Figure 2.6: Unique polymorphism for Mnandi obtained with the reactions conditions as described in the text. Lane 1 is molecular marker VI, lane 2 is Astrid, lane 3 is Mnandi and lanes 4 to 12 is VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel and Agria and a H_2O control respectively. (a). RAPD profile obtained with primer UBC509 and the reactions conditions as described in the text. A unique polymorphism between Astrid and Mnandi, of about 300bp, is observed. (indicated by the arrow) (b). RAPD profile obtained with primer UBC582. A unique polymorphism of about 900bp is observed between Astrid and Mnandi (indicated by the arrow).

The two polymorphisms, unique to Mnandi, were obtained with UBC509 (ACAGAGACTG) and UBC582 (GGATAGACG). UBC509 amplified a ± 300 bp fragment in Mnandi alone (Figure 2.6a). UBC 582 amplified a ± 900 bp fragment only in Mnandi (Figure 2.6b). It was decided to use the polymorphisms obtained with UBC509 and UBC582 to develop two sets of SCAR primers to distinguish between Astrid and Mnandi. These polymorphisms were named $UBC509^{300}$ and $UBC582^{900}$ respectively.

Chapter 3

Cloning and sequencing of the RAPD product

3.1 Introduction

The two polymorphisms, UBC509^{300} and UBC582^{900}, were identified and had to be cloned into a suitable vector. Some expected and unexpected problems are encountered when cloning amplified DNA into a vector. These result from the final state of the PCR product:

- The primers used in the PCR reactions are mostly synthetic oligonucleotides with 5'-OH ends. Thus, the PCR product is normally dephosphorylated (Fabry and Dietmaier 1995).
- When double-stranded synthesis in the last PCR cycle is allowed to completion, the PCR-fragments are **blunt ended** if special thermostable DNA polymerases containing a proofreading activity are used, e.g. *Pwo* DNA polymerase, *Pfu* DNA polymerase, and *Vent*® DNA polymerase, New England Biolabs (Lohff and Cease 1992) or contain **one-nucleotide single strand overhangs**, if the "conventional" *Taq* and Dynazyme®, Finnzymes OY, DNA polymerases are used, because this enzyme exhibits a terminal transferase activity which adds a single nucleotide (mainly dATP) to the 3' ends of the DNA (Clark 1988).

These features of the PCR product cause several difficulties during cloning and various approaches can be followed to clone these PCR products.

The routine, but tedious, way to clone PCR fragments involves blunt-end ligation. This includes various enzymatic steps (Fabry and Dietmaier 1995):

- Phosphorylation of the PCR fragment by T4 polinucleotide kinase.
- Fragment ends (one-nucleotide single strand overhangs) must be made flush with the 3'-5' exonuclease activity of T4 polymerase or DNA polymerase I.

- The cloning vector must be dephosphorylated to prevent self-ligation.

The cloning efficiency of blunt end ligation is however much lower than with sticky end cloning, even if large amounts of T4 ligase is used. No opportunity to direct the orientation of the fragment within the vector exists (Fabry and Dietmaier 1995). Other methods have recently been used to clone PCR amplified DNA fragments.

***In vivo* cloning** exploits an intrinsic enzymatic activity of *E. coli*. PCR products with long terminal sequences identical to that of a linearised vector are generated and the PCR fragments and linearised vector are simultaneously introduced into an *E. coli* strain possessing high recombination activity (Oliner *et al.* 1990). The PCR products and vector are subsequently connected, by homologous recombination, to form circular plasmids. Despite a high cloning efficiency reported, this is not suitable for routine practice because the primers must contain additional 50 - 90 nucleotides at their 5' ends (Fabry and Dietmaier 1995). This method is totally impractical for cloning RAPD fragments due to the length of the final primers.

Turbo cloning exploits the *lox*/Cre recombination system of bacteriophage P1[2]. A linearised, dephosphorylated vector containing a *lox* sequence near one end is blunt-end ligated with the PCR fragments under conditions (15% polyethylene glycol; PEG) that favour an intermolecular reaction ("macro-molecular crowding"). Subsequent addition of the Cre protein to the mixture will cut all those constructs in which a PCR fragment is flanked by two properly orientated vectors creating a single circular recombinant plasmid ready for transformation. This method requires many enzymatic steps, a special enzyme and a very accurate ligation procedure to avoid side reactions. In addition, even though 15% PEG may dramatically increase blunt end ligation conditions and macro-molecular crowding, it is still necessary to remove the 3' adenosine nucleotide generated by the heat stable polymerase (*Taq* or Dynazyme) as in normal blunt-end cloning (Fabry and Dietmaier 1995).

Ligase-independent cloning of PCR amplified fragments completely avoid the use of enzymatic ligation. Relatively large, complementary, single strand overhangs between the vector and amplified fragments are generated which provide by subsequent hybridisation nicked circular molecules that are competent for *E. coli* transformation. The *E. coli* host repair the nicks *in vivo*, after introduction, with its repair enzymes. Several specific protocols for ligase-independent cloning have been published by Aslanidis and de Jong (1990), Hsiao (1993) and Schuldiner *et al.* (1990). All of these protocols require the addition of 12 or more nucleotides at the 5' ends of each amplification primer, rendering these methods impractical for the cloning of RAPD fragments (Fabry and Dietmaier 1995).

The **introduction of restriction sites** at the 5' ends of the two amplification primers and the generation of sticky ends by the subsequent digestion of the amplified DNA fragments with the appropriate restriction endonuclease (Scharf *et. al.* 1986) also improve cloning efficiency. However, various restriction enzymes do not recognise sites that occur at the ends of linear double stranded DNA (Fabry and Dietmaier 1995). This method is not suitable for cloning RAPD fragment, except if the 10mer RAPD primer contain a restriction site.

UDG cloning make use of PCR primers that contain "Uracil DNA Glycosylase (UDG) cloning sites" during amplification. These sites are introduced by incorporation of several dUMP residues within the 5' end of the primers. The fragments are incubated with UDG after amplification in order to remove the uracil. Base pairing is disrupted at the 5' end leaving 3' free to anneal directly to complementary vector ends. This method is reported to be efficient with DNA amounts of 1-10 pg. (Buchman *et al.* 1992 in Fabry and Dietmaier 1995). A disadvantage is that it is necessary to synthesise primers with twelve additional 5' nucleotides including four or five unusual dUMP residues (Fabry and Dietmaier

1995). Thus, this method of cloning is impractical for the cloning of RAPD fragments if commercial 10mer RAPD fragments were used.

Directional cloning using exonuclease III (Kaluz *et al.* 1992) is based on generating two different short 5' cohesive ends of PCR amplified DNA by time and temperature-controlled exonuclease III digestion. These ends are complementary to normal vector sticky ends produced by restriction digests. The procedure requires kinasing of PCR primers that contain only four additional nucleotides. The yield of recombinant clones (<100/experiment), however, may be suitable for routine cloning of most PCR amplified DNA, but is insufficient for more subtle procedures like generating PCR-based libraries (Fabry and Dietmaier 1995). It is possible to use this method to clone RADP fragments if the correct primers were used.

DI/TRISEC (Di/Trinucleotide sticky end cloning) described by (Dietmaier *et al.* 1993) overcome some of the problems mentioned above. It additionally facilitates directional cloning. The DI/TRISEC method is simple, inexpensive in routine practise and can be used with any heat stable polymerase without special chemicals or vectors (Fabry and Dietmaier 1995). The DI/TRISEC method relies on the generation of complementary di- or trinucleotide sticky ends on the linearised vector and the PCR-amplified DNA. The vector ends are generated by Klenow polymerase after restriction digestion that produce 5' overhangs. The PCR fragment ends are trimmed to short complementary ends by the controlled 3'-5' exonuclease activity of T4 DNA polymerase in the presence of "stopping" nucleotides (Dietmaier and Fabry 1995).

The so-called **T-vectors or TA-cloning** improved the cloning efficiency of PCR and RAPD fragments. It is reported that TA-cloning is more than 50 times more efficient than blunt-end ligation (Clark 1988; Halton and Graham 1991, Mead *et al.* 1991 and Zhou *et al.* 1995). This system overcomes, and utilises, the problem arising from the terminal transferase activity of the heat stable polymerase. The T-vectors have 3'

overhang nucleotide ends, each with a single thymidine residue, which is complementary to the A-overhang ends produced by the polymerase. No common available restriction enzyme produce directly single thymidine overhangs, therefore, is it necessary to produce these overhangs by the addition of a single thymidine residue to the 3' ends of the vector by either a heat stable polymerase, e.g. *Taq* or Dynazyme (Marchuk *et. al.* 1991) or a terminal transferase in the presence of dideoxythymidine triphosphate (Holton and Graham 1991). The restriction endonuclease *Xcm* I does provide single 3' thymidine nucleotide ends, but the enzyme has a very unusual recognition site, thus, it is necessary to insert synthetic linkers into the cloning site of the vector before it is possible to use this enzyme (Kovalic *et. al.* 1991). However, it is not possible to ligate PCR products, produced by heat stable polymerase which lacks the terminal transferase activity, e.g. *Pwo* DNA polymerase, *Pfu* DNA polymerase, Vent DNA polymerase, into T-vectors (Clark 1988, Marchuk 1991, Holton and Graham 1991, Anonymous 1994, Fabry and Dietmaier 1995). However, Zhou *et al.* (1995) adapted the TA-cloning method in such a manner that it is possible to clone and subclone any DNA fragment. The purified DNA fragments are treated with *Taq* DNA polymerase for 30 minutes, at 72°C, in a reaction in the presence of $MgCl_2$ and dATPs before ligation.

The T-vectors or TA-cloning is being used frequently for the cloning of RAPD fragments. Lorenz, *et al.* (1997) used a TA-cloning system to clone RAPD fragments amplified from sugar beet mitochondrial DNA. Francis *et al.* (1995) used it to clone a RAPD fragment specific for rye chromosome arm of the 1BL.1RS translocation. Rameau *et al.* (1998) use the TA-cloning system to clone two RAPD markers in pea linked to *rms3*, a one of three *ramosus* genes, and *dna*, a gene conferring flowering response to photoperiod, in order to develop SCAR markers for these two genes. The TA-cloning system was also used by Barret *et al.* (1998) and Lahogue *et al.* (1998) in the development of their respective SCAR markers. It was decided to use the TA-

cloning system because it was used with success by other authors to clone RAPD fragments.

3.2 Materials and Methods

3.2.1 Amplification of RAPD products

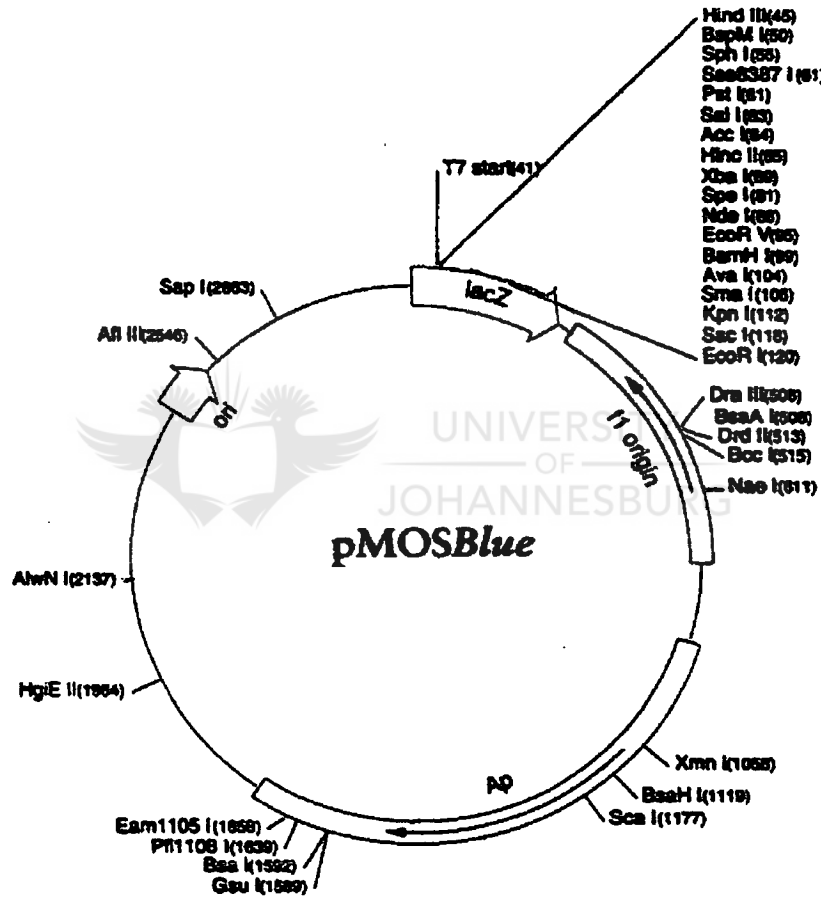


Figure 1: Vector map of the pMOS*Blue* vector (Anonymous 1994)

The fragments of the two polymorphisms, $UBC509^{300}$ and $UBC582^{900}$, were amplified in order to get a high enough concentration of the two fragments to clone into the pMOS*Blue* T-vector (Amersham Life Science). Two approaches were followed in order to get a high enough concentration of DNA for cloning. 40 standard

10μl RAPD reactions, using genomic DNA from Mnandi (see Materials and methods, Chapter 2) were amplified along with one Astrid and a water control with primers UBC509 and UBC582 respectively. The Mnandi RAPD products were precipitated with sodium acetate as described in the GELase manual (Anonymous 1998). These products were separated on a 1.5% low melting point agarose gel (SeaPlaque).

3.2.2 Purification of PCR products

The fragments were purified from the low melting point agarose by digesting the low melting point agarose with the GELase^{TM} enzyme (Epicentre technologies) as described in the high activity protocol (Anonymous 1998).

Bands containing fragments of the correct size were cut out of the LMP gels under long-wavelength UV light to minimise damage to the ethidium-bromide stained DNA. The gel slices were weighed and soaked in a 1X GELase buffer for one hour. The excess buffer was carefully removed. The gel slices were thoroughly melted at 70°C (3 min per 200 mg gel slice). The molten gel slices were equilibrate at 45°C for 3 min. GELase enzyme, 1 unit per 600 mg of 1.0% LMP-agarose gel, was added to the molten gel and incubated for 1 hour at 45 °C. The DNA fragments were then precipitated with ethanol and ammonium acetate (Anonymous 1995). The precipitated DNA fragments were resuspended in 20 μl ddH_2O and stored in a freezer at -20 °C.

The concentration of the DNA fragments to be cloned was increased by using DNA from RAPD fragments, which was excised and purified from the gels, as the template for another PCR reaction as done by Lorenz, *et al.* (1997) and Barret *et al.* (1998).

3.2.3 Cloning of PCR products

The two fragments UBC509^{300} and UBC582^{900}, (about 300bp and 900bp respectively) were cloned into the pMOS*Blue* T-vector (Amersham Life Science) exploiting the template independent activity of the thermostable polymerase, Dynazyme, which preferentially ad a single adenosine to the 3' end of double stranded DNA following the protocol given by the supplier (Clark 1988, Anonymous 1994).

10μl ligation reactions were set up, each ligation reaction contained 1X ligase buffer, 5 mM DTT, 0.5 mM ATP, 50ng pMOS*Blue* T-vector, 2 U T4 DNA ligase, 2.0 μl diluted fragment, obtained from the GELase purified fragments and nuclease free water. The ligation mixes were gently stirred and incubated in a waterbath, at 16°C overnight before transformation.

3.5 μl of the ligation mixes were then added to 20 μl competent *E. coli* cells. The mixtures were put on ice for 30 minutes. The ligation mixtures were heat shocked for 3 min at 37°C and placed back on ice for 5 min. 80 μl SOC medium were added to each tube. The mixtures were then incubated at 37°C for one hour. The cells were plated out on Luria agar plates which contained 50mg.ml^{-1} ampicillin and surface coated with 35 μl X-gal (50 mg.ml^{-1}) and 20 μl IPTG (100 mM) half an hour before plating out of the transformed cells. The plated cells were left to incubate overnight at 37 °C. This cloning system uses the blue/white screening method. The positive colonies are white because the insert was inserted in the *lac Z* gene, thus deactivating it. The cells are no longer able to convert the X-gal into the blue product, resulting in white colonies.

3.2.4 Verification of cloned products

3.2.4.1 PCR with specific primers

The white colonies were screened for inserts of the right size by direct PCR of the white colonies, with two primers, T7 and U19mer, designed specifically for the pMos*Blue* and related vectors. White colonies, approximately 1 mm in diameter were picked up using a sterile toothpick. Replica colonies were made by touching the toothpick onto another plate. The colonies were then transferred to Eppendorf tubes containing 50μl sterile water. The tubes were vortexed to disperse the cells. The cells were lysed and the DNases were denaturated by boiling for 5 min. The tubes were centrifuged at 12 000 g for 1 min to remove the cell debris. 10μl of the supernatant was transferred to clean tubes and left on ice untill use. The PCR reaction mixes contain 10X Dynazyme buffer, 0.25 mM dNTPs , 5 pM T7 primer, 5 pM U 19mer primer, 1.25 U Dynazyme enzyme and ddH_2O. 5 μl of the DNA from the white colonies was added as template.

The reaction volumes were overlaid with 20 µl of mineral oil to prevent evaporation during the PCR process. DNA amplification was conducted in a Hybaid thermal cycler programmed for 35 cycles of 1 min denaturating at 94.5°C, 1 min annealing at 55°C and 2 min elongation at 72°C, with the initial denaturation of 1 min at 94.5°C, and a final extension of 5 min at 72°C. The amplification products were resolved electrophoretically in a 1% agarose gel, stained with ethidium bromide and photographed under UV light. Boehringer Mannheim molecular weight marker VI (152 bp - 2176 bp) was used to determine fragment lengths.

Positive colonies were grown on LB liquid broth containing 50 $mg.ml^{-1}$ ampicillin and stored on beads (Davies Diagnostics) at 70°C.

3.2.4.2 Southern blotting

The recombinant plasmids were Southern blotted to the corresponding RAPD profiles, using fragments amplified from the recombinant plasmids as probes, to ensure that the correct fragment was cloned. The Southern blotting and detection was done as described in the DIG Systems user's guide (Van Miltenburg *et al.* 1995).

DIG labelled probes were produced by PCR, with the T7 and U19 primers, using the recombinant plasmids as template. The T7 and U19 primers amplified the inserted fragments as well as a part of the vector. The PCR reaction mixes contain 10X Dynazyme buffer, 0.25 mM dATP, 0.25 mM dCTP, 0.25 mM dGTP, 0.1625 mM dTTP, 0.0875 mM alkali labile DIG-11-dUTP, 5 pM T7 primer, 5 pM U 19mer primer, 1.25 U Dynazyme enzyme and ddH_2O. 5 µl of template DNA (recombinant plasmids, $2ng.100\mu l^{-1}$). The mixes were briefly centrifuge and overlaid with mineral oil. The PCR programme consisted of an initial denaturation of 1 min at 94.5°C. followed by 35 cycles of 1 min denaturating at 94.5°C, 1 min annealing at 55°C and 2 min elongation at 72°C. A final extension time of 5 min at 72°C were allowed. The PCR products were separated on a 1% agarose gel along with unlabelled PCR reactions and, DIG labelled, Boehringer Mannheim marker VI and low DNA mass ladder to determine the concentration of the probe.

Efficiency of incorporation of the DIG-11-dUTP was determined by constructing a dry blot (Hoisting *et al.* 1994). A piece of glass plate, larger than the size of the gel, was laid on a level bench. Five layers of blotting paper, soaked wet in transfer buffer (20X SSC: 3M NaCl; 300mM sodium citrate, pH 7), slightly larger than the size of the gel, were placed on the glass plate. The gel was then placed, face down, on blotting paper followed by a piece of moist blotting membrane (MSI, Micron Separation Inc.), all bubbles between the two layers were removed. A thin dry filter paper, the same size as the membrane, was placed on top, followed by a small stack of dry paper towels. A weight was placed on the construct and was left overnight.

The detection was started by washing the membrane for 5 min in 2xSSC at room temperature before baking it at 90°C for 60 min. The baked membrane was rinsed in washing buffer (100mM maleic acid, 150mM NaCl; pH 7.5; 0.3% Tween20). The membrane was put in blocking solution (1% blocking reagent for nucleic acid in maleic acid buffer containing 150mM NaCl; pH 7.5) at room temperature for ten minutes followed by anti-DIG solution (1:10 000 diluted anti-digoxygenin-AP in blocking solution) at room temperature for 10 min. The membrane was washed for 5 min in washing buffer (100mM maleic acid, 150mM NaCl; pH 7.5; 0.3% Tween20) at room temperature and rinsed in detection buffer (100mM Tris-HCl; 100mM NaCl; pH 9.5) at room temperature. The membrane was placed in CSPD (diluted 1:100 in detection buffer) in the dark at room temperature for 10 min. The membrane was then rinsed again in detection buffer at room temperature before it was sealed in a plastic bag. An X-ray film was exposed to it for 90 min (Van Miltenburg *et al.* 1995).

RAPD reactions were performed with UBC 509 and UBC 582 respectively with genomic DNA from Astrid, Mnandi, BP_1, Buffelspoort, Up-to-Date, VanderPlank, Hoëvelder, Hertha, Pimpernel, Agria and dH_2O control reaction. 30 µl of these RAPD products were separated on a 2% agarose gel along with the DIG labelled PCR products of the appropriate purified recombinant plasmids amplified with the T7 and U19 primers. Unlabelled PCR products amplified with the same primers were loaded as well as DIG labelled Boehringer marker VI. The gels were stained with ethidium bromide.

The gels were then depurinated by submerging it in a depurination solution (250 mM HCl) for 10 min at room temperature on a shaker followed by a rinse in dH_2O. The gels were denaturated by two washes in a denaturation solution (0.5N NaOH; 1.5M NaCl) for 15 min at room temperature followed by a rinse in H_2O. The gel was neutralised by two washes in neutralising solution (0.5 M Tris-HCl, pH 7.5; 3M NaCl) for 15 min at room temperature (Van Miltenburg *et al.* 1995).

The Southern blot was set up by placing a glass plate on top of a plastic support in a shallow tray. A filter paper wick was placed on top of the glass plate and the gel was placed, face down, onto filter paper. The blotting membrane was placed on top, and all bubbles between the gel and the membrane were removed. Two thin pieces of dry filter paper, same size as membrane, were placed on top of the membrane followed by a stack of dry paper towels. Another glass plate was placed on top followed by a weight. Enough transfer buffer (20x SSC: 3M NaCl; 300mM sodium citrate, pH 7) was added to the tray to ensure that the buffer level stayed high during the entire blotting process. The construct was left overnight to allow for the capillary transfer of the DNA to the membrane.

The membrane was removed the following morning and immediately placed in 2x SSC. Excess agar particles were gently rubbed off from the membrane with a gloved hand. The membrane was washed in washing buffer for 15 min at room temperature on the shaker and dripped dry until moist after which it was baked at 95°C for 120 min (Van Miltenburg *et al.* 1995).

The membrane was then rolled, DNA side inwards, and placed in a hybridisation bottle. Prehybridising was done for 3-4h at 42°C in DIG Easy Hyb solution (20ml.100cm^{-2}). The probe was meanwhile denaturated by heating to 95°C in a boiling waterbath after which it was quenched on ice. 100 µl TE (10mM Tris-HCl; 1mM EDTA; pH 8.0) was added to increase the volume. The prehybridisation solution was removed and the hybridisation solution, 25ng.ml^{-1} probe diluted in DIG Easy Hyb solution (3.5ml.10cm^{-2}), was added. Hybridisation was done overnight at 42°C. The hybridisation solution was discarded the next morning and the membrane was removed from the hybridisation bottles. The membrane was washed twice for 5 min, on a shaker, in 2x wash solution (2xSSC; 1%SDS) at room temp, followed by two washes for 15min, on a shaker, in 0.5x wash solution (0.5xSSC; 0.1%SDS) at 68°C (Van Miltenburg *et al.* 1995).

Detection was performed by rinsing the membrane in washing buffer for 1 min at room temperature. The membrane was then washed in a blocking solution for 30 min at room temp. Antibody solution was prepared by diluting anti-digoxygenin-AP 1:10,000 in blocking solution. The membrane was incubated for 30 min in the anti-DIG solution at room temp. The membrane was washed twice for 15 min in wash buffer at room temperature after which it was equilibrated for 2 min in detecting buffer at room temperature. The CSPD was diluted 1:100 in detection buffer and the membrane was washed in the diluted CSPD for 10 min in the dark at room temperature followed by a rinse in detection buffer at room temperature. The membrane was slowly removed from the detection buffer, allowing the solution to drip off, and placed DNA side down on a plastic bag which was then sealed. An X-ray film was exposed to the membrane for 12h at room temperature (Van Miltenburg *et al.* 1995).

3.2.5 Sequencing

Two colonies, one from UBC 509^{300} and UBC 582^{900} each, were plated on Luria Bertani agar plates, containing 50 $mg.ml^{-1}$ ampicillin, for sequencing. The sequencing of the inserts was carried out with automated sequencing by the Department of Microbiology of the University of Cape Town.

3.3 Results

3.3.1 Amplification and cloning of the PCR product

The fragments of the two polymorphisms, UBC 509^{300} and UBC 582^{900}, were amplified by PCR as described. The initial attempt to clone the fragments was unsuccessful. The TA-cloning efficiency of PCR amplified fragments decrease with time, possibly due to the loss of the 3' overhang (Invitrogen Corporation, Eucaryotic TA cloning kit, Instruction Manual, Version 1.0. San Diego, CA. in Zhou *et al.* 1995). Furthermore, according to Li and Guy (1996), longer final extension times improves efficiency of cloning with TA vector systems, suggesting that the

polymerase needs more time to add the template independent A to the 3' end. Thus, the final extension time of the RAPD reactions was extended to 5 min if the products were to be used for cloning. Barret *et al.* (1998) said that ultra-violet light might also affect the quality of the A overhangs and hence the cloning efficiency.

The second attempt at cloning was successful. Blue, white as well as bullseye colonies were present on the plates. White colonies were picked up and transferred to Luria-agar plates in order to verify if the cells from the colonies were transformed with the correct length inserts.

3.3.2 Verification of cloned products

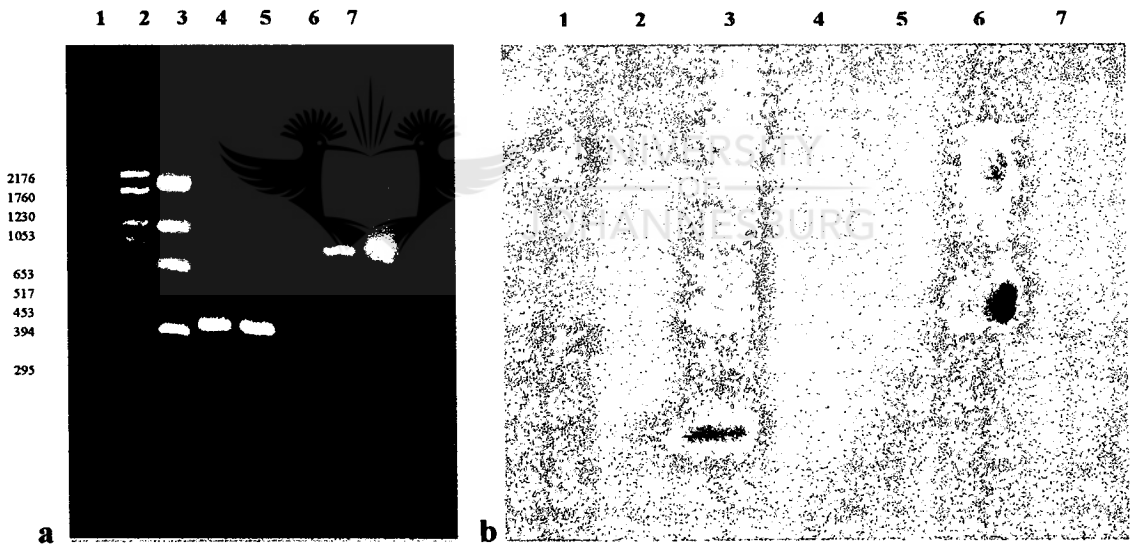


Figure 3.2: Dig labelling of the probes by PCR. with primers T7 and U19 using the recombinant plasmid as template. (a) Shows the gel stained with ethidium bromide. Lane 1 is the DIG labelled marker VI, lane 2 the low DNA mass ladder, lane 3 and 4 the labelled (probe) and unlabelled UBC 509^{300} fragment and lane 6 and 7 the labelled (probe) and unlabelled UBC 582^{900} fragment. Lane 5 is empty (b) Southern blot of the same gel. The labelling reaction was successful as illustrated by the detection of only the labelled probes.

Single fragments were amplified by the specific primers in both cases. A fragment of approximately 400 bp was amplified with the UBC 509^{300} colonies and a fragment of

approximately 1,000 bp was amplified with the UBC 582^{900} colonies. These fragment lengths correspond to the lengths of the initial RAPD fragments plus the sequences between the insertion sites and primer recognition sites in the vector.

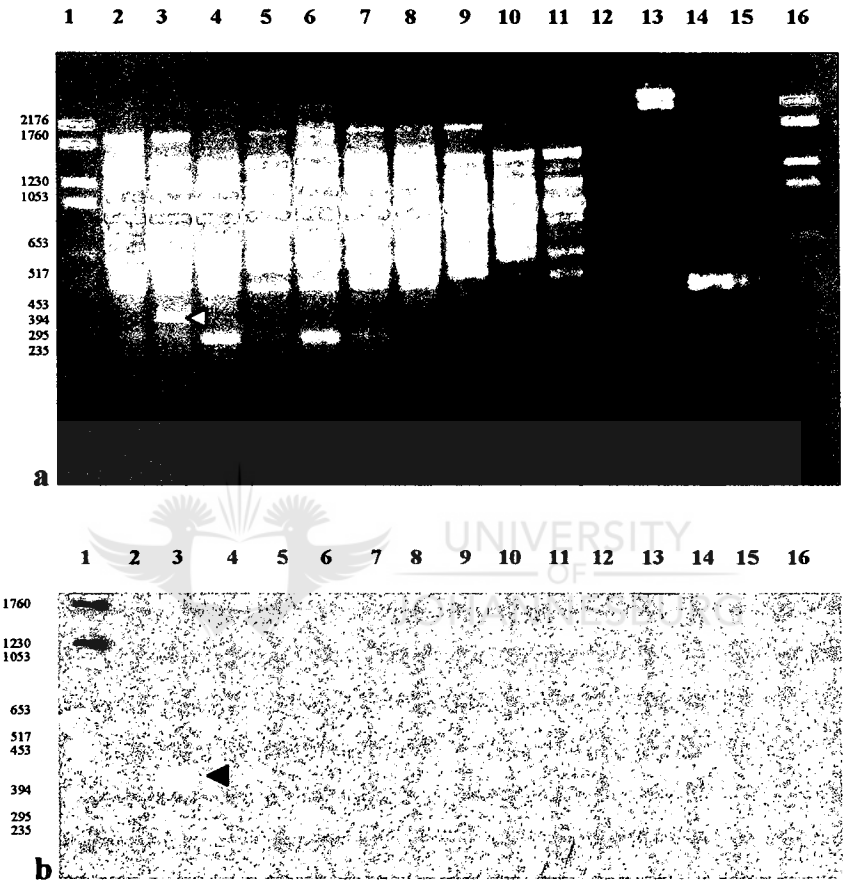


Figure 3.3: Southern blots of the RAPD profiles obtained with UBC 509. (a). Ethidium bromide stained gel. The unique polymorphism is indicated by an arrow. Lane 1 is DIG labelled marker VI, lane 2 to 12 is Astrid, Mnandi, VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel, Agria and the water control respectively. Line 13 is *Eco*RI digested recombinant plasmid (partially digested), and line 14 and 15 is the labelled and unlabelled probe respectively. Lane 16 is again marker VI. b. Southern blot of the RAPD profile. The probe hybridised with the unique fragment (indicated by an arrow) confirming that the correct fragment was cloned.

The labelling of the probe with DIG was successful as seen in Figure 3.2a and b. The concentration of the labelled probe was 12.0 $ng.\mu l^{-1}$ for the probe UBC 509^{300} and 8.0

ng.μl^{-1} for the probe UBC 582^{900} as determined by comparing the intensity of the amplified fragments to that of the low DNA mass ladder (Figure 3.2a).

The RAPD profiles and the corresponding Southern blots are given in Figures 3.3 and 3.4.

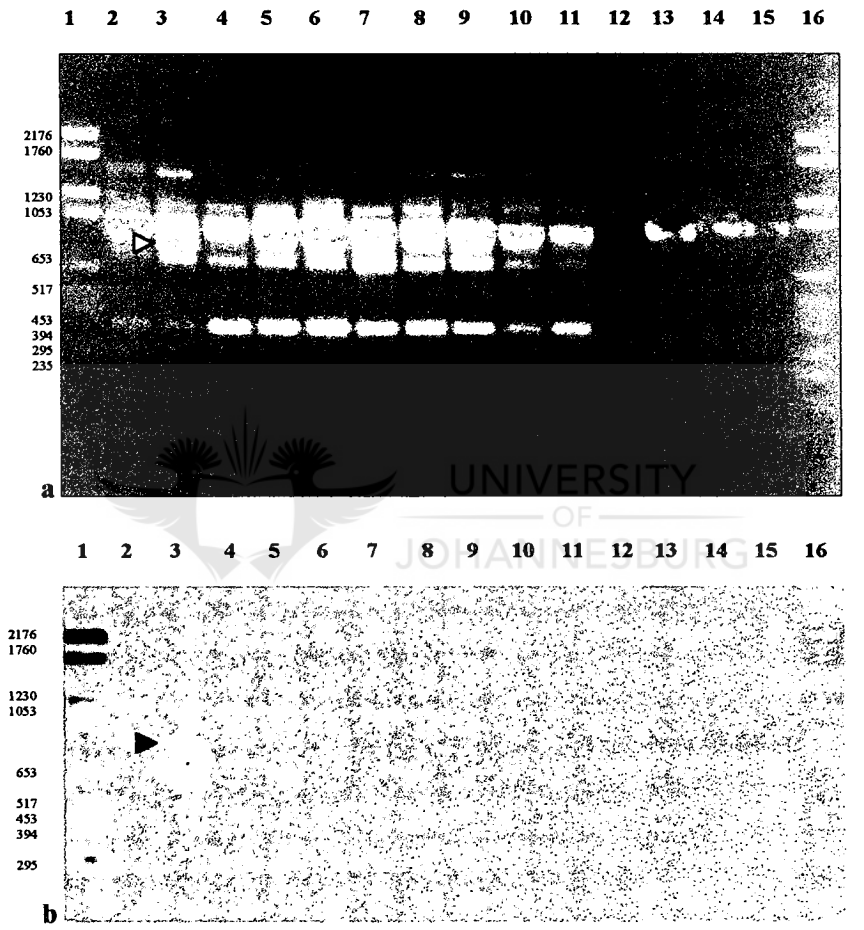


Figure 3.4: Southern blots of the RAPD profiles obtained with UBC 582. (a). Ethidium bromide stained gel. The unique polymorphism is indicated by an arrow. Lane 1 is DIG labelled marker VI, lane 2 to 12 is Astrid, Mnandi, VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel, Agria and the water control respectively. Line 13 is *Eco*RI digested recombinant plasmid (partially digested as seen on the Southern blot), and line 14 and 15 is the labelled and unlabelled probe respectively. Lane 16 is molecular marker VI again. (b). Southern blot of the RAPD profile. Several homologous fragments exist in the profile. The fragment homologous to the unique fragment is indicated by an arrow.

The fragment UBC 509^{300} is unique to Mnandi as illustrated in Figure 3.3b. The DIG labelled UBC 509^{300} amplified from the cloned plasmids hybridise only to the corresponding fragment on the RAPD profile, thus the correct fragment was cloned. No other homologous sequences of the same length were present in the RAPD profile.

Fragment UBC 582^{900} seem to be unique on the RAPD profile but it showed extensive homology to other fragments on the RAPD profile. These might all be different alleles of the same locus.

3.2.4 Sequencing

The sequence for the 300 bp fragment (UBC 509^{300}) amplified with UBC 509 is given underneath. The binding sites for UBC 509 (ACAGAGACTG) are highlighted in red.

ACAGAGACTGGAGGGATAGTGGTGCTACAGAGACTGGAGAACTCATTG
GTGGTGTGTCAGTTGCATGTCTTGGAAGTCCATCCGCCCGTTGGAATAAA
GGATATTCAGTGGACTGTTTTGCAGGTAAATCATGAAGACCTGCATATGT
TTTAGCAAAATATATTCAAATTGCATTACATCCATTCATTGAATAAACATT
CAACATATTTATTGGTTTAGGTTGGGAACCTTTGCCTCTAACTTGCTTTGA
AGAGGTGGAAGAGGCATTAAAGATATTCAATTGAGCAGTCTCTGT

Two binding sites exist for the primer UBC 509 in the 5' end of the sequence. The difference in length between the products is only 26 bp. The electrophoretic mobility of both bands would be almost be the same resulting in only one band, that might be more diffused.

A BLAST search (Altschul *et al.* 1997) showed no extensive homology to other sequences in their data bases (Appendix 3) apart from 87% (74/78) homology to EST246148 tomato ovary, TAMU *Lycopersicon esculentum* cDNA clone

cLED13D22, mRNA sequence. The other homologies included homologue to *Drosophila melanogaster* DNA sequence (P1 DS07070 (D331))(20/20 homology), *Caenorhabditis elegans* cosmid F55D10 (20/20 homology), *Pyrococcus horikoshii* OT3 genomic DNA (23/24) Human factor XIII b subunit gene 20/20) and other mentioned in appendix 3.

The 900 bp fragment (UBC 509^{300}) amplified with UBC 509 was sequenced in both directions. The two sequences were aligned with the aid of Genepro Sequence Analysis Software (Riverside Scientific Enterprises). The aligned sequences are given in appendix 2. The aligned sequences were compared and the corrected sequence is given underneath. The binding sites for UBC 582 (GGTATAGACG) are highlighted in red.

GGTATAGACGGCTTGTTGGAAAGTTGAATTATCTCACAGTGACTAGACCTGACATCTCTT
TTCCTGTAAGTgTTGTAAGTCAGTTTATGTTCCCCTTGTGATAGTCATTGGGAAGCAGTTGT
TCGTATTCTGCGATATATAAAGTCAGCTCCCGGTAAAGGACTACTCTTTGAGGATCAAGG
CCATGAGCATATCATTGGATATATAGACGCTGATTGGGCAAGATCACCCTCTGACAGACG
TTCGACATCCGGATATTGCGTTTTAGTAGGAGGTAATTTGGTGTCGTGGAAGAGTAAGAA
AAAGCaTgTGGTTGCTCGATCTAGTGCAGAATCAGAATATCGAGCAATGGCGACAACAAC
TTGTGAGCtAGTTTGGATCAAACAGTTGCTGGGAGAACTAAAATTTGGCAAAATTGATCAG
ATGGAACTTATGTGTGATAATCAAGCGGCTCTTCATATCGCATCAAATCCAGTGTTTCATG
AAAGGACTAAGCACATTGAGATTGACTGTCACTTTGTCAGAGAAAAGATACTCTCGGGAG
ATATTGTTACGAAATTTGTGAAGTCAAATGATCAGCTTGCAGATATCTTCACCAAGTCCCT
CACATGTCCTCGTATTAATTACATCTGTAACAAGCTAGGTACATATGATTTGTATGCACCA
GCTTGAGGGGGAGTGTTAGAATAGAAATAAATATTTCCtACTTAGAATAGGAATAGGAAT
AGGAATAGGGAATTAAGGAATAGGGAATcCtaaAGTTGGGAAAAGGGATTCCAATGGTAA
GTGGTCcAATAAAATAAGGGGgcctCAtATGTAACtTaTTTTACCtTTaCACaAATTTCCAATAaT
TTATTTTAATTTCCCCCCCAAGGATaCCATTTaTTataAGCGTCTATACC

A BLAST search (Altschul *et al.* 1997) showed xtensive homology to other sequences in their date bases (Appendix 4). Homologies found include *Lycopersicon esculentum* cv Red River unknown sequence PCR random amplified RAPD band 10 (41/43), and *L. esculentum* Le-hsf8 gene for heat stress transcription factor 8 (56/65).

Others include *Arabidopsis thaliana* DNA chromosome 3, BAC clone F18B3 (ESSA project) (42/48) (87%), and *Pisum sativum* retrotransposon (28/30).

In both cases (UBC 582^{900} and UBC 509^{300}) the sequence obtained show homology to *L. esculentum*, a close relative of potato. The fragment UBC 882^{900} might be a conserved sequence in the Solanaceae family, if one take into account that the probe (figure 3.4b) binds to various fragments in all the cultivars in the RAPD profile, and that the sequence of the fragment show homology to sequences in *L. esculentum*.

Chapter 4

SCAR design and analysis

4.1 Introduction

SCAR (sequence characterised amplified region) markers, derived from RAPD markers, overcome some of the drawbacks of RAPDs. A SCAR is a genomic DNA fragment at a single genetically defined locus that is identified by PCR amplification using a pair of specific oligonucleotide primers. SCARs are derived by cloning and sequencing the two ends of RAPD amplification products. These sequences are then used to design sets of longer oligonucleotide primers, usually 20 -24 base pairs long. Amplification of target DNA using this set of primers will result in the reproducible amplification of specific loci when high annealing temperatures are used. These SCARs are advantageous over RAPD markers as they detect only a single locus, the amplification reaction is less sensitive to reaction conditions, and they can potentially be converted into more informative codominant markers (Paran and Michelmore 1993, Naqvi *et al.* 1996 and Barret *et al.* 1998). Horejsi *et al.* (1999) found that the SCAR markers were less influenced by $MgCl_2$ concentration than the corresponding RAPD markers and that the SCAR markers were not influenced by either the source of *Taq* polymerase or the type of thermocycler used. They found however that some SCAR markers were more sensitive to reaction impurities than the corresponding RAPD markers and that the SCAR markers were less readily visualised at lower template concentrations (1 and 2 ng).

The conversion of the dominant RAPD marker to a co-dominant SCAR marker, which is able to distinguish heterozygous plants in a polymorphic segregating population, is the biggest advantage of the SCAR markers according to Barret *et al.* (1998). Lahogue *et al.* (1998) are also of the opinion that their SCC8 SCAR marker is more specific and simpler than the OPC08-1020 RAPD marker for seedlessness in grapevine. SCARs are simple to generate but require some investment of time and

labour. SCARs can be used in conjunction with RAPDs to provide multi-allelic codominant markers. Adam-Blondon *et al.* (1994) concluded that both RAPD and SCAR analysis provide identical conclusions (demonstrating that the observed polymorphism observed with the two techniques correspond effectively to the same locus), but they are of the opinion that the SCAR was much easier to score than the RAPD. Barret *et al.* (1998) are also of the opinion that the SCAR markers are more reliable and that the SCAR markers are frequently more rapidly amplified. However, Horejsi *et al.* (1999) found that the efficiency of converting RAPD markers to SCAR markers is low. They were able to convert 64% of specific RAPD markers, in cucumber, into SCAR markers and only 11% of these reproduced the original RAPD polymorphism. Some SCAR primer pairs even produce multiple polymorphic products.

Certain considerations must be taken in account when the SCAR primers are designed. The primers must ideally have a GC content of at least 50% and a random base distribution. The calculated T_ms for the primer pair should be balanced and, depending on the application, T_ms between 55°C and 80°C are desirable. Complementarity on the 3' ends of primer pairs should be avoided as this could promote the formation of primer-dimer artefacts and thus the yield of the desired product reduces. Primers with long stretches (runs) of polypurines of polypirimidines, or unusual sequences should preferably be avoided. Runs (three or more) of G or C at the 3' and of the primers should be avoided as it might promote mispriming in GC rich areas. Furthermore, primer sequences should be checked against each other for complimentary, especially for overlaps at the 3' end. Complimentary and 3' overlaps promote the formation of "primer-dimer" artefacts. Sequences with significant secondary structures, such as palindromic sequences, in the primer itself should be avoided as this might lead to a reduction in yield as well (Innes and Gelfand 1990, Saiki 1990).

If these longer SCAR primers are used in PCR reactions there are a few possible outcomes. Single bands of the same or different size as the progenitor RAPD fragment might be amplified, only in the genotypes positive for the progenitor RAPD band, thus only one allele is detected. The RAPD polymorphism is retained and these SCAR markers could be scored as dominant markers (Paran and Michelmore 1993; Xu *et al.* 1995; Horejsi *et al.* 1999).

In some instances alleles or fragments were amplified in genotypes which are positive and negative for the progenitor RAPD polymorphism. These fragments might be of the same size or they might differ in size. If there is a length polymorphism between the alleles, the progenitor dominant RAPD marker is converted into a co-dominant SCAR marker (Paran and Michelmore 1993). In the cases where the progenitor RAPD polymorphism is lost, the SCAR primer amplified the same size fragment in all the genotypes because the SCAR primers are able to find binding sites in both genotypes (Paran and Michelmore 1993; Xu *et al.* 1995; Dax *et al.* 1998 and Lahogue *et al.* 1998; Horejsi *et al.* 1999). It is possible to restore the polymorphism. The amplified SCAR fragments may be digested with certain restriction enzymes like *Hae*III, *Rsa*I, *Hind*III or *Alu*I. This allows the detection of restriction site differences within the amplified SCAR fragment and cuts the amplified SCAR fragments into smaller fragments in order to detect small fragment length differences (Paran and Michelmore 1993; Adam-Blondon *et al.* 1994; Barret *et al.* 1998; Dax *et al.* 1998, Lahogua *et al.* 1998 and Rameau *et al.* 1998; Horejsi *et al.* 1999). This is called cleaved amplified polymorphic sequences or CAPS (Staub *et al.* 1996) or a specific amplicon polymorphism (SAP) by Hittalmani *et al.* (1995). Hittalmani *et al.* (1995) restored the polymorphism by cleavage of the amplification products with *HaeIII*. Lahogue *et al.* 1998 restored the polymorphism of their SCC8 SCAR marker, for seedlessness in grapevine, by digesting the amplification product with the restriction enzyme *Bgl*II, thus revealing two allelic forms $SCC8^{+}$ and $scc8^{-}$. Thus, the SCC8 SCAR marker was a co-dominant marker. Roupe van der Voort *et al.* (1999) used CAPS to compare the *Solanum bulbocastanum* genome to the *S.*

tuberosum genome in order to identify PCR based markers linked to the *Rmc1* gene for root-knot nematode resistance.

The polymorphic nature of the newly derived SCAR markers could also be restored by elevating the annealing temperature during amplification. This may reveal different levels of mismatches within a primer binding sequence and no amplification of an allele occurs in the genotype negative for the progenitor RAPD marker if the annealing temperature is elevated. The SCAR marker can be scored as a dominant marker (Paran and Michelmore 1993; Horejsi *et al.* 1999). The amplification of alleles from genotypes positive and negative for the progenitor RAPD polymorphism, in the instance when the SCAR primers contain the original RAPD primer sequences as well, indicated that the original RAPD polymorphism was, most probably, caused by mismatches in one or a few nucleotides in the RAPD priming site. These mismatches were tolerated by the longer SCAR primers resulting in the amplification of both lines (Paran and Michelmore 1993; Xu *et al.* 1995; Horejsi *et al.* 1999). Adding 10 to 20 bases to the 3' terminus of the 10-mer primers in a non-polymorphic region may have provided enough homology to overcome the original mismatch and to permit annealing and amplification.

Single base differences between two alleles might be identified by various other methods such as denaturating gradient gel electrophoresis or ribonuclease cleavage (Paran and Michelmore 1993; Barret *et al.* 1998). The alternate alleles may also be sequenced to identify diverged regions, for which it may be possible to design allele specific primers, or to reveal restriction site polymorphisms between the alternate alleles (Paran and Michelmore 1993).

Additional bands might also be amplified if these SCAR primers are used in amplification reactions (Adam-Blondon *et al.* 1994; Xu *et al.* 1995, Barret *et al.* 1998, Lahogue *et al.* 1998). Adam-Blondon *et al.* (1994) found that an additional band, other than the markerband, was present if they used their SCH20 SCAR primer

pair. This additional band might represent another amplified locus. Dax *et al.* (1998) found that an additional, slightly longer band, was amplified in all the genotypes, when using their SCAR primer, this band does not hybridise with the RAPD marker.

It is also possible to restore the polymorphic nature of the SCAR marker by using genetically more diverged lines, not used in the initial development of the SCAR marker. Thus a different allele might be amplified in this "new" genotype (Paran and Michelmore 1993).

The advantages of SCAR markers are summarised by Weising *et al.* (1995) and Lahogue *et al.* (1998) as follows:

1. reproducibility - since long primers complementary to specific genomic loci are used, stringent PCR conditions can be applied which exclude competition between primer binding sites, resulting in reliable and reproducible bands which are less sensitive to reaction conditions. McDermot *et al.* (1994) are also of the opinion that a set of standard SCAR primers are more amenable than RAPDs for comparison of work between laboratories;
2. mapping studies - the codominant SCAR markers are more informative for genetic mapping than the dominant RAPDs;
3. map-based cloning - ordinary RAPD fragments often contain interspersed repetitive DNA and thus cannot be used as probes for identifying a clone of interest, here the SCAR primer can be used to screen pooled genomic libraries by PCR;
4. physical mapping - since the SCAR markers are essentially similar to the "sequence tagged sites" (STS) proposed by Olsen *et al.* (1989) as landmarks in physical mapping of the human genome, they might also serve as anchoring points between physical and genetic maps;
5. locus specificity - the reproducible amplification of defined genomic regions allows comparative mapping (as has been done by RFLPs) or homology studies between related species.

SCAR markers linked to powdery mildew resistance genes in barley germplasm (McDermott *et al.* 1994), downy mildew resistance genes in lettuce (Paran and Michelmore 1993), anthracnose resistant gene, *Are*, in common bean (Adam-Blondon *et al.* 1994), the rice blast resistance gene, *Pi-2(t)*, in rice (Hittalmani *et al.* 1995) and the TMV resistance gene $Tm2^2$ in tomatoes (Dax *et al.* 1998) have been developed. SCARS have also been used to fingerprint grape (*Vitis*) rootstocks (Xu *et al.* 1995). Francis *et al.* (1995) converted a RAPD marker specific for the rye chromosome arm of the 1BL.1RS translocation, common in high yielding bread wheat varieties, into a probe which is used to detect the presence of rye chromatin by *in situ* hybridisation and Southern blotting. Rameau *et al.* (1998) developed two SCAR markers that are linked to *rms3*, one of three *ramosus* genes, and *dne*, a gene conferring flowering response to photoperiod, in pea. A SCAR marker linked to seedlessness in grapevine was developed by Lahogue *et al.* (1998). Barret *et al.* (1998) developed a SCAR marker for the molecular tagging of the dwarf BREIZH (*Bzh*) gene in *Brassica napus* L. Horejsi *et al.* (1999) converted several RAPD markers in cucumber into SCAR markers in order to establish the efficiency of the conversion of RAPD markers into SCAR markers. Naqvi *et al.* (1996) developed a SCAR marker for the indirect selection of blast resistance in rice.

SCARs can be derived not only from RAPDs but also from cloned sequences like RFLP probes. SCARs derived from RFLP probes will be identical to STS's (Olsen *et al.* 1989) but may not be useful as PCR-based markers. The differences underlying the RFLPs are often outside the region of hybridisation with the probe and these polymorphisms will not be detected by primers from the two ends of the RFLP probe. Thus, it may be difficult to identify polymorphisms within the regions amplified. cDNA sequences may also be less useful than genomic clones as a source of SCARs because of the absence of introns in cDNA. If the two primer sites flank an intron, the intervening fragment may be to big to be amplified or if the sequence of an individual SCAR primer is interrupted by an intron, primer binding will be obstructed.

Furthermore, coding regions may be more conserved than the random sequences identified from RAPDs. Niewohner *et al.* (1995) converted RFLP probes for loci *CP56* and *CP113* in potato, into allele specific markers. They are of the opinion that the same approach is also applicable for other RFLP markers. Livneh *et al.* (1992) converted an RFLP marker into a PCR primer for the detection of purity in hybrid pepper cultivars. Lee and Penner (1997) converted RFLP markers, linked to the QTL for malting quality in barley, to allele specific amplicons by a two-step sequencing approach. First they sequenced the RFLP probes used. These sequences were then used to amplify locus specific fragments. These fragments were sequenced and these sequences were then used to develop allele specific primers. Lee and Penner are however of the opinion that the cross-applicability of the RFLP marker does not necessarily imply the cross-applicability of the allele specific amplicon. A lack of cross-applicability due to recombination between the gene controlling the trait and the marker used for marker assisted selection for the trait would be the same for both markers. Both RFLP and allele specific amplicon marker systems may be based on as little as a single base difference, therefore both marker systems are equally susceptible to a loss of cross-applicability due to mutation. Mutation events that affect the restriction site of an RFLP would not necessarily affect the priming site for the allele specific amplicon. Thus, the cross-applicability of the allele specific amplicon must be verified on parental material, independently of the RFLP data for each new cross analysed. These are also essentially SCARs.

Closely related to SCARs is "allele specific associated primers" (ASAPs) which is specific primers that are used in a fluorescence based procedure to amplify DNA template in microtitre plates to generate a single DNA fragment, per primer used, at stringent annealing temperatures. The DNA fragment is present in only the genotype possessing the appropriate allele and thus eliminate the need do separate the amplification products by electrophoresis. This approach decreases the time consumed and increases the reliability of large scale screening with PCR (Gu *et al.* 1995 and Staub *et al.* 1996).

The two SCAR primer pairs, based on the sequences of UBC 509^{300} and UBC 582^{900}, are designed and evaluated in this part of the study.

4.2 Materials and methods

4.2.1 Design of the SCAR primers

Two sets of SCAR primers were designed with the aid of the Pirmer Designer 3 for Windows (Scientific and Educational Software). The guidelines mentioned earlier were followed in designing the primes. The primer pairs suggested by Primer Designer for fragment UBC 509^{300} is given and analysed in Table 4.1, 4.2 and 4.3.

Table 4.1:The primer pairs suggested by Primer Designer for fragment UBC 509^{300}

Rank	Position		GC%		T_m °C[1]		Product		
	A	B	A	B	A	B	Length	GC%	T_m °C
1	15	267	50	50	63	73	253	40	83
2	15	268	50	50	63	72	254	39	83

[1] - T_m °C as calculated by Primer Designer

Table 4.2: A detailed analysis of the first pair of primers for UBC 509^{300}:

Primer Balance	**A1**	**B1**	**Comment**
Length	20	20	
% GC	50	50	
T_m °C (Primer Designer)	63	73	10 °C difference
Primer Cautions			
Stability (kcals)	2.9	2.2	
Runs of bases	2	2	
Repeats (dinuc)	2	2	
Hairpins	none	none	
Primer Dimers	A1:A1	B1:B1	A1:B1
3' end matches	2	2	2
Any adj homol bases	3	2	4

Primer A1 is 5'-GATAGTGGTGCTACAGAGAC-3' and primer B1 5'-ATGCC TCTTCCACCTCTTCA-3'

Table 4.3: A detailed analysis of the second pair of primers for UBC 509^{300}:

Primer Balance	A2	B2	Comment
Length	20	20	
% GC	50	50	
T_m °C (Primer Designer)	63	72	9 °C difference
Primer Cautions			
Stability (kcals)	2.9	2.6	
Runs of bases	2	2	
Repeats (dinuc)	2	2	
Hairpins	none	none	
Primer Dimers	A2:A2	B2:B2	A2:B2
3' end matches	2	1	2
Any adj homol bases	3	2	4

Primer A2 is 5'-GATAGTGGTGCTACAGAGAC-3' (the same as primer A1 of the first set) and B2 is 5'-AATGCCTCTTCCACCTCTTC-3'

It is however recommended that a G or C is present at the 3' end of the primer. The primers designed by Primer Designer were altered, based on the sequence of the fragment, to see if it was possible to meets this recommendation. The primer sequences were then as followed:

Sequence SCAR UBC 509^{300}a: 5'-CAGTTGCATGTCTTGGAAGTC-3'
Sequence SCAR UBC 509^{300}b: 5'-ATGCCTCTTCCACCTCTTCA-3'

A detailed analysis of the final primer set SCAR UBC 509^{300}a and b is given in Table 4.4 and 4.5.

Table 4.4: A detailed analysis of primer SCAR UBC 509^{300}a:

Criteria	PCR Primer	This Primer	Meets Criteria
% GC	Range 50-60	47	Yes
T_m °C[1]	Range 50-80	72	Yes
T_m °C[2]	Range 50-80	57.9	Yes
T_m °C[3]	Range 50-80	62	Yes
Stability	>= 2.0 kcals 5'vs3'	2.5	Yes
Hairpins	Annealing 55 °C	None	Yes
3' Dimers	< 3 matches 3'end	2	Yes
Dimers-Any	< 7 adj homol bases	4	Yes
Runs	< 3 base runs	2	Yes
Repeats	< 3 dinuc repeats	1	Yes

[1] - T_m as calculated by the Primer Designer software.

[2] - T_m as calculated on the oligo synthesis report.

[3] - T_m as calculated by 2(A+T)+4(G+C)

Sequence: 5'-CAGTTGCATGTCTTGGAAGTC-3' and composition: A:4 C:4 G:6 T:7

Table 4.5: A detailed analysis of primer SCAR UBC 509^{300}b:

Criteria	PCR Primer	This Primer	Meets Criteria
% GC	Range 50-60	50	Yes
T_m °C[1]	Range 50-80	73	Yes
T_m °C[2]	Range 50-80	57.3	Yes
T_m °C[3]	Range 50-80	60	Yes
Stability	>= 2.0 kcals 5'vs3'	2.2	Yes
Hairpins	Annealing 55°C	None	Yes
3' Dimers	< 3 matches 3'end	2	Yes
Dimers-Any	< 7 adj homol bases		Yes
Runs	< 3 base runs	2	Yes
Repeats	< 3 dinuc repeats	2	Yes

[1] - T_m as calculated by the Primer Designer software.

[2] - T_m as calculated on the oligo synthesis report.

[3] - T_m as calculated by 2(A+T)+4(G+C)

Sequence: 5'-ATGCCTCTTCCACCTCTTCA-3' and composition: A:3 C:9 G:1 T:7

The binding position of the RAPD primer in the fragment is highlighted in red and the position of the SCAR primers in the fragment sequence is highlighted in purple:

ACAGAGACTGGAGGGATAGTGGTGCTACAGAGACTGGAGAACTCATTGGTGGTGTG
TCAGTTGCATGTCTTGGAAGTCCATCCGCCCGTTGGAATAAAGGATATTCAGTGGAC
TGTTTTGCAGGTAAATCATGAAGACCTGCATATGTTTTAGCAAAATATATTCAAATTGC
ATTACATCCATTCATTGAATAAACATTCAACATATTTATTGGTTTAGGTTGGGAACCTT
TGCCTCTAACTTGCTTTGAAGAGGTGGAAGAGGCATTAAAGATATTCAATTGAGCAG
TCTCTGT

The primer pairs suggested by Primer Designer for fragment UBC 582^{900} is given and analysed in Table 4.6, 4.7 and 4.8.

Table 4.6: The primer pairs suggested for 582^{900} by Primer Designer:

Rank	Position		GC%		T_m °C		Product		
	A	B	A	B	A	B	Length	GC%	T_m °C
1	174	674	50	50	74	73	501	41	85
2=	1	454	50	50	71	72	451	42	85
2=	2	462	45	45	69	70	461	42	85
2=	2	456	45	45	69	70	455	42	85
5	1	463	50	50	71	75	463	42	85
6=	1	456	50	45	71	70	456	42	85
6=	1	462	50	45	71	70	462	42	85
6=	176	674	45	50	71	73	499	41	85
9	2	454	45	50	69	75	453	42	85
10	2	463	45	50	69	75	462	42	85

It was decided to analyse the two pairs that give the longest product. The analysis of these two primer pairs are given in Table 4.7 en 4.8.

Table 4.7: A detailed analysis of the first pair of primers for UBC 582^{900}:

Primer Balance	A1	B1	Comment
Length	20	20	
% GC	50	50	
T_m °C	74	73	1 °C difference
Primer Cautions			
Stability (kcals)	3.3	3.1	
Runs of bases	2	2	
Repeats (dinuc)	1	1	
Hairpins	none	none	
Primer Dimers	A1:A1	B1:B1	A1:B1
3' end matches	2	1	2
Any adj homol bases	4	4	4

Primer A1 (bp174): 5'-GGATCAAGGCCATGAGCATA-3' and primer B1 (bp674C): 5'-CCTCAAGCTGGTGCATACAA-3'

Table 4.8: A detailed analysis of the second pair of primers for UBC 582^{900}:

Primer Balance	A2	B2	Comment
Length	20	20	
% GC	45	50	5% difference
T_m °C	71	73	9 °C difference
Primer Cautions			
Stability (kcals)	5.3	3.1	
Runs of bases	2	2	
Repeats (dinuc)	2	1	
Hairpins	none	none	
Primer Dimers	A2:A2	B2:B2	A2:B2
3' end matches	2	1	1
Any adj homol bases	4	4	4

Primer A2: 5'-ATCAAGGCCATGAGCATATC-3' and primer B2: 5'-CCTCAAGCTGGTGCATACAA-3' (The same as primer B1 of first pair)

No primer that met all the criteria as well as the recommendation of a C or G at the 3' end could be designed. It was decided to choose primer pair 1 (Table 4.7) which gives the longest amplification product.

SCAR UBC 582^{900}a: 5'-GGATCAAGGCCATGAGCATA-3'

SCAR UBC 582^{900}b: 5'-CCTCAAGCTGGTGCATACAA-3'

A detailed analysis of these two primers is given in Table 4.9 and 4.10.

Table 4.9: A detailed analysis of primer SCAR UBC 582^{900}a:

Criteria	PCR Primer	This Primer	Meets Criteria
% GC	Range 50-60	50	Yes
T_m °C[1]	Range 50-80	74	Yes
T_m °C[2]	Range 50-80	57.3	Yes
T_m °C[3]	Range 50-80	60	Yes
Stability	>= 2.0 kcals 5'vs3'	3.3	Yes
Hairpins	Annealing 55 °C	None	Yes
3' Dimers	< 3 matches 3'end		Yes
Dimers-Any	< 7 adj homol bases		Yes
Runs	< 3 base runs		Yes
Repeats	< 3 dinuc repeats		Yes

[1] - T_m as calculated by the Primer Designer software.

[2] - T_m as calculated on the oligo synthesis report.

[3] - T_m as calculated by 2(A+T)+4(G+C)

Sequence: 5'-GGATCAAGGCCATGAGCATA-3'

Composition: A:7 C:4 G:6 T:3

Table 4.10: A detailed analysis of primer SCAR UBC 582^{900}b:

Criteria	PCR Primer	This Primer	Meets Criteria
% GC	Range 50-60		Yes
T_m °C[1]	Range 50-80	73	Yes
T_m °C[2]	Range 50-80	57.3	Yes
T_m °C[3]	Range 50-80	60	Yes
Stability	>= 2.0 kcals 5'vs3'	3.1	Yes
Hairpins	Annealing 55°C	None	Yes
3' Dimers	< 3 matches 3'end		Yes
Dimers-Any	< 7 adj homol bases		Yes
Runs	< 3 base runs	2	Yes
Repeats	< 3 dinuc repeats	1	Yes

[1] - T_m as calculated by the Primer Designer software, [2] - T_m as calculated on the oligo synthesis report, [3] - T_m as calculated by 2(A+T).4(G+C)

Sequence: 5'-CCTCAAGCTGGTGCATACAA-3' and composition: A:6 C:6 G:4 T:4

The binding position of the RAPD primer in the fragment is highlighted in red and the position of the SCAR primers in the fragment sequence is highlighted in purple:

GGTATAGACGGCTTGTTGGAAAGTTGAATTATCTCACAGTGACTAGACCTGACATCTC
TTTTCCTGTAAGTgTTGTAAGTCAGTTTATGTTCCCCTTGTGATAGTCATTGGGAAGCA
GTTGTTCGTATTCTGCGATATATAAAGTCAGCTCCCGGTAAAGGACTACTCTTTGAGGA
TCAAGGCCATGAGCATATCATTGGATATATAGACGCTGATTGGGCAAGATCACCCTCT
GACAGACGTTCGACATCCGGATATTGCGTTTTAGTAGGAGGTAATTTGGTGTCGTGGA
AGAGTAAGAAAAAGCaTgTGGTTGCTCGATCTAGTGCAGAATCAGAATATCGAGCAATG
GCGACAACAACTTGTGAGCtAGTTTGGATCAAACAGTTGCTGGGAGAACTAAAATTTGG
CAAAATTGATCAGATGGAACTTATGTGTGATAATCAAGCGGCTCTTCATATCGCATCAA
ATCCAGTGTTTCATGAAAGGACTAAGCACATTGAGATTGACTGTCACTTTGTCAGAGAA
AAGATACTCTCGGGAGATATTGTTACGAAATTTGTGAAGTCAAATGATCAGCTTGCAG
ATATCTTCACCAAGTCCCTCACATGTCCTCGTATTAATTACATCTGTAACAAGCTAGGT
ACATATGATTTGTATGCACCAGCTTGAGGGGGAGTGTTAGAATAGAAATAAATATTT
CCtACTTAGAATAGGAATAGGAATAGGAATAGGGAATTAAGGAATAGGGAATcCtaaAGT
TGGGAAAAGGGATTCCAATGGTAAGTGGTCcAATAAAATAAGGGGgcctCAtATGTAACtT
aTTTTACCtTTaCACaAATTTCCAATAaTTTATTTTAATTTCCCCCCCAAGGATaCCATTTaT
TataAGCGTCTATACC

4.2.2 PCR with the SCAR primer pairs

PCR reactions were carried out with the two SCAR primer pairs. The 10 µl reaction mix for one reaction contained 2.05 µl ddH_2O, 1.25 µl 10X Buffer (100mM Tris-HCl, pH 8.3, 500mM KCl), 3.5 mM $MgCl_2$, 0.4 mM dNTPs, 0.24 mM of each primer and 0.75 U *rTaq*. 10ng template DNA was used in the reactions.

The reaction volumes were overlaid with 20 µl of mineral oil to prevent evaporation during the PCR process. DNA amplification was conducted in a PTC100 (MR Research) thermal cycler programmed for a initial denaturation cycle of 10 min, 35 cycles of 30 sec denaturating at 95°C, 45 sec annealing at a specific temperature and 120 sec elongation at 72°C, and a final extension of 7 min at 72°C. The amplification products were resolved electrophoretically in a 1 % agarose gel, stained with ethidium bromide and photographed under UV light. Boehringer Mannheim molecular weight marker VI (152 bp - 2176 bp) was used in all the samples to determine fragment lengths. Southern blots of the SCAR profiles were also done as described in section 3.2.4.2.

4.2.3 Restriction enzyme digests

If no polymorphisms were observed when using the SCAR primers it is often possible to regain the polymorphisms by digesting the amplification product with restriction enzymes. The amplification products of the two sets of SCAR primers were digested with restriction enzymes. The digests were carried out in 10 µl reaction volumes and 5 µl of the amplification products were used for these restriction enzyme digests. The reaction mixes for these digests contain 0.1 U restriction enzyme, 1.0 µl appropriate buffer, 3.9 µl ddH_2O and 5.0 µl of the amplification products.

The amplification products were digested at 37°C, unless otherwise specified for the specific restriction enzyme. The products were resolved electrophoretically in a 2 %

MetaPhor agarose gel, stained with ethidium bromide and photographed under UV light. Boehringer Mannheim molecular weight marker VI (152 bp - 2176 bp) was used in all the samples to determine fragment lengths of the fragments. .

4.3 Results

4.3.1 Amplification with the SCAR primers

SCAR primer set UBC 509a+b amplified a fragment of about 230 bp as well as a slightly longer fragment of about 260 (Figure 4.1) in Mnandi. The 260 bp fragment corresponds with the predicted length of 253 (Chapter 4.2.1). The polymorphism was retained. The 230 bp fragment was, however, amplified in Astrid as well as all the other cultivars. These fragments might be amplified because there is other recognition sites in the genome is well. Some longer unspecific fragments were amplified, when the annealing temperature was 55°C, in both Astrid and Mnandi. The annealing temperature was elevated to eliminate the nonspecific longer bands which were amplified as well but some of these bands still amplify at 65°C as seen in Figure 4.1a. The probe hybridised to the corresponding 260 bp fragment as well as to the longer fragments amplified in Mnandi (Figure 4.1b). Hybridisation also occur in all the cultivars.

SCAR primer set UBC 582a+b amplified a fragment of about 500 bp, as well as some longer fragments, at an annealing temperature of 55 °C. The longer SCAR primers might be able to withstand mismatches at the priming sites better than the short RAPD primers. The 500 bp fragment corresponds with the predicted length of 501bp (Chapter 4.2.1), in both Astrid and Mnandi. However, the polymorphism was lost. Two possible approaches to regain the polymorphism were to elevate the annealing temperature and/or digest the amplification products with restriction enzymes. A higher annealing temperature would also eliminate the nonspecific longer bands, which were amplified as well. Some of the nonspecific amplification products were eliminated but the same length fragment was still amplified in both Astrid and

Mnandi as well as in all the other cultivars as well as seen in Figure 4.2. Hybridisation also occured to different fragments in all the cultivars.

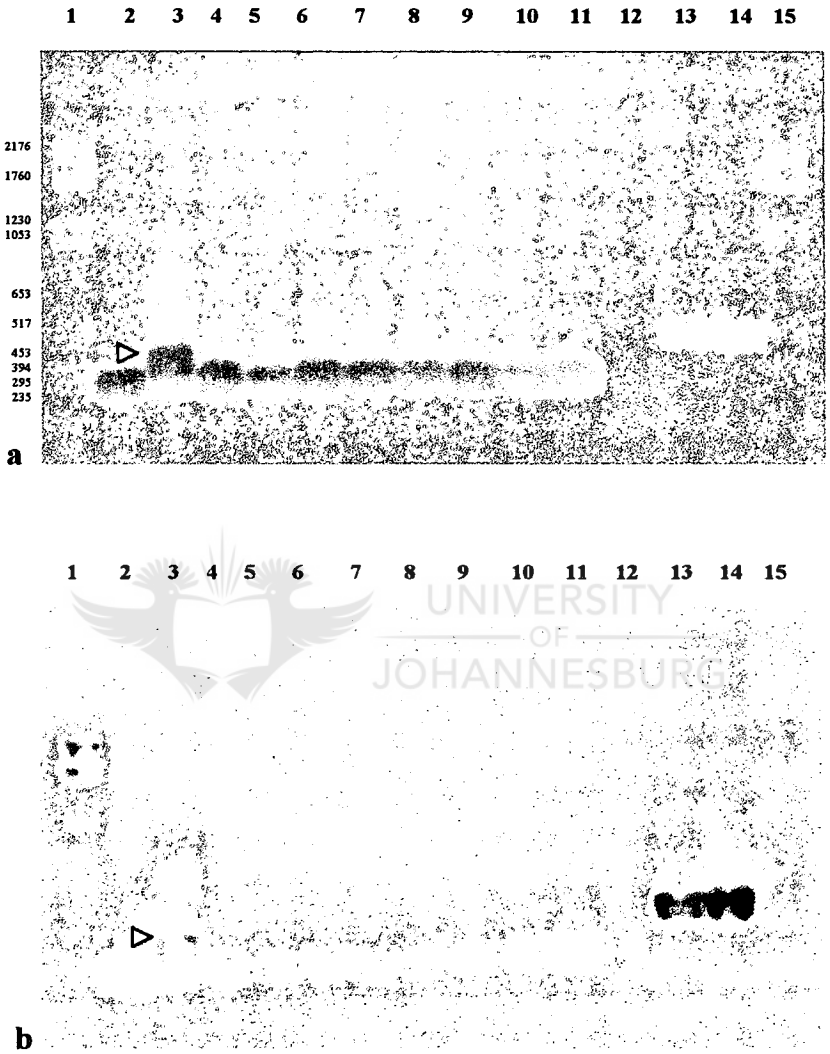


Figure 4.1: Amplification of genomic DNA with SCAR primers SCAR UBC 509a+b at a annealing temperature of 65°C. (a) Ethidium bromide stained gel of the SCAR profile.The polymorphic 260 bp fragment, unique to Mnandi, is indicated by an arrow. The nonspecific fragments could be seen above the unique fragment. Lane 1 and 15 marker VI, lanes 2 to 12 are Astrid, Mnandi, VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel, Agria and a water control respectively. Lanes 13 and 14 are the labelled and unlabelled probes respectively. (b) Southern blot of the same gel. The probe hybridised to the 260 bp fragment (indicated by an arrow) as well as the other longer fragments.

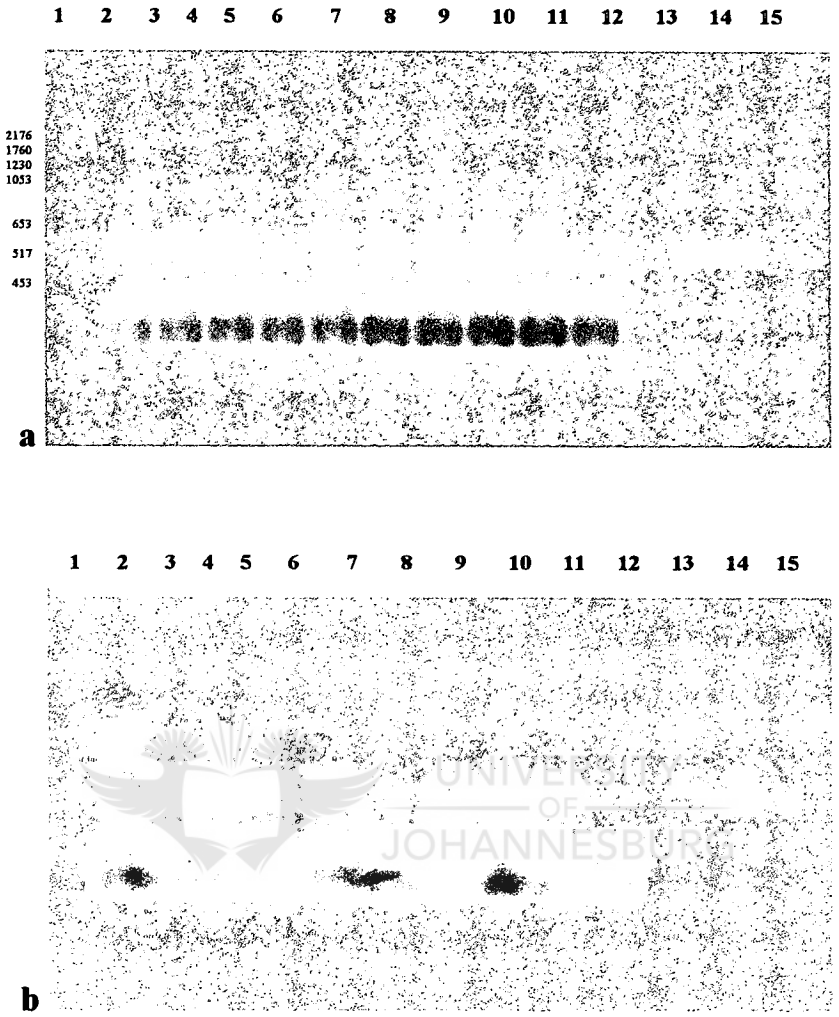


Figure 4.2: Amplification of genomic DNA with SCAR primers SCAR UBC 582a+b at an annealing temperature of 65°C. (a) Ethidium bromide stained gel of the SCAR profile. Various fragments were amplified. Lane 1 and 15 marker VI, lanes 2 to 12 are Astrid, Mnandi, VanderPlank, Up-to-Date, BP_1, Buffelspoort, Hoëvelder, Hertha, Pimpernel, Agria and a water control respectively. Lane 13 and 14 are the labelled and unlabelled probe respectively. (b) Southern blot of the same gel. Hybridisation occurred to all the amplified bands.

4.3.2 Restriction enzyme digest of Amplified products

Digests with 17 enzymes were tried. From these only *Eco*RV , *Ssp*I, *Stu*I and *Rsa*I (Table 4.11) digest the amplification products. The products were however the same for both Mnandi and Astrid.

Table 4.11: Restriction enzymes that cut SCAR amplification product

Enzyme	SCAR UBC 582a+b	SCAR UBC 509a+b
*Alu*I	-	-
*Aoc*I	-	-
*Bam*II	-	-
*Bgl*II		-
*Eco*RV	✂	-
*Hae*III	-	-
*Hind*III	-	-
*Nco*I	-	-
*Nsi*I	-	-
*Pst*I	-	-
*Pvu*I	-	-
*Rsa*I,	✂	-
*Sac*I	-	-
*Sal*I	-	-
*Ssp*I	✂	-
*Stu*I	✂	-
*Xba*I	-	-

The SCAR UBC 582^{900} could not be converted into a CAPS as well. It was thus impossible to regain the polymorphism and the SCAR UBC582^{900} could not be used to identify Mnandi from Astrid.

Chapter 5

Discussion

The development of SCAR markers gives us a potential powerful tool for identifying certain genotypes, ie. cultivars or resistant genotypes. The SCAR marker's biggest and most powerful usage will be in the field of marker assisted selection. The SCAR marker can give a quick and easy "yes or no" answer to the question if the specific genotype includes the specific marker and thus the gene it is linked to. The use of SCAR markers in a marker assisted breeding program will speed up the selection process and more genotypes could be screened in a shorter period and at an earlier stage.

The use of SCAR markers in cultivar identification would however be limited. The SCAR marker would be able to identify a specific cultivar. No more information would be obtained except if the marker is linked to a specific gene. Another problem of using a SCAR marker to identify a cultivar is that the marker is inherited by some of its offspring. However, due to the growing economical importance of Mnandi, it was decided to develop a SCAR marker for the identification of Mnandi, because the SCAR marker is less sensitive to variation in reaction conditions as discussed in chapter 4. The results obtained from the SCAR marker would be more reproducible between laboratories. However this SCAR marker could however only be used to identify Mnandi from Astrid.

Some of the difficulty of converting a RAPD marker into a SCAR marker is illustrated in this study. Although the RAPD marker was unique to Mnandi, amplification occurs in both Astrid and Mnandi, when the longer SCAR primers were used. No length polymorphism between the amplification products from Astrid and Mnandi, exists when the primers derived from the UBC 582^{900} fragment (SCAR

UBC 582^{900} a:5'-GGATCAAGGCCATGAGCATA-3' and SCAR UBC 582^{900} b:5'-CCTCAAGCTGGTGCATACAA-3') were used in a PCR reaction. It was also not possible to regain a polymorphism by elevating the annealing temperature, nor was it possible to regain a polymorphism by digesting the amplification products with restriction enzymes. The conversion of the RAPD marker UBC 582^{900} to a SCAR marker was unsuccessful.

Amplification occurs in both Astrid and Mnandi when the SCAR UBC 509^{300} primers (SCAR UBC 509^{300}a:5'-CAGTTGCATGTCTTGGAAGTC-3' and SCAR UBC 509^{300}b:5'-ATGCCTCTTCCACCTCTTCA-3') were used and a length polymorphism was detected. A slightly longer fragment was amplified in Mnandi along with a fragment which was amplified in both Astrid and Mnandi. The RAPD marker, UBC 509^{300}, was successfully converted into a SCAR marker, SCAR UBC 509^{300} and could be used to distinguish Mnandi from Astrid. The reaction conditions for using the SCAR primer are given in chapter 4.

Adam-Blondon *et al.* (1994); Xu *et al.* (1995), Barret *et al.* (1998) and Lahogue *et al.* (1998) all found the amplification of additional bands with their SCAR markers. These additional bands might be the amplification products of other loci (Adam-Blondon *et al.* 1994) or it might be amplification artefacts or heteroduplexes formed between different amplification products (Ayliffe *et al.* 1994).

The developed SCAR marker, SCAR UBC 509^{300}, can now be used by the ARC-Roodeplaat and the potato industry to distinguish Mnandi from Astrid. This SCAR marker can be used in combination with the existing DNA markers as well as morphological markers. The advantage is that this marker would be able to distinguish between Mnandi and Astrid more rapidly, as discussed previously.

This marker could also be used to try and identify the possible paternal father of Mnandi. Since it contributed durable field late blight resistance to the genotype of Mnandi. Segregation studies of "Mnandi x Astrid cross" or "Mnandi x light blight susceptible genotype" progeny must however be done to determine if SCAR UBC 509^{300} is linked to the late blight resistance genes.

Chapter 6
References

Adam-Blonden AF, M. Sévignac H. Bannerot and M Dron. 1994. SCAR, RAPD and RFLP markers linked to dominant gene (*Are*) conferring resistance to anthracnose in common bean. **Theoretical and Applied Genetics 88**:865-870

Altschul SF, TL Madden, AA. Schäffer, J Zhang, Z Zhang, W Miller and DJ Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research** 25:3389-3402

Anonymous. 1998. GELase™ Agarose Gel-Digesting Preparation Product information, Epicentre Technologies, Madison, Wisconson, USA

Anonymous. 1994. pMOS*Blue* T-vector kit, Amersham Life Science, Little Chalfont, Buckinghamshire, England

Aruna M, ME Austin and P Ozias-Akins. 1995. Randomly amplified polymorphic DNA fingerprinting for identifying Rabbiteye Blueberry (*Vaccinium ashei* Reade) cultivars. **Journal of the American Society for Horticultural Science 120**: 710-713

Aslanidis C and PJ de Jong, 1990, Ligation-independaent cloning of PCR products (LIC-PCR). **Nucleic Acids Research 18:**6069-6074

Baird E, S Cooper-Bland, R Waugh, M DeMaine and W Powell. 1992. Molecular characterisation of inter- and intra-specific somatic hybrids of potato using randomly amplified polymorphic DNA (RAPD) markers. **Molecular and General Genetics 233**:469-475

Barret P, R Delourme, F Foisset and M Ranard. 1998. Development of a SCAR (sequence characterised amplified region) marker for molecular tagging of the dwarf BREIZH (*Bzh*) gene in *Brassica napus* L. **Theoretical and Applied Genetics 97:**828-833

Becker HC, GM Engqvist, B Karlsson. 1995. Comparison of rapeseed cultivars and resynthesised lines based on allozyme and RFLP markers. **Theoretical and Applied Genetics 91**: 62-67

Beckman JS and M Soller. 1990. Towards a unified approach to the genetic mapping of eukaryotes based non sequence-tagged microsatellite sites. **Bio/Technology 8**:930-932

Beyerman B, P Nurnberg, A Weihe, M. Meixner, JT Epplen and T Borner. 1992. Fingerprinting plant genomes with oligonucleotide probes specific for simple repetitive DNA sequences. **Theoretical and Applied Genetics 83:**691-694

Boehringer Mannheim. 1995. Boehringer Mannheim PCR Application Manual. Mannheim, Germany.

Boiteux LS, MEN Fonseca and PW Simon. 1999. Effects of plant tissue and DNA purification method on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. **Journal of the American Society for Horticultural Science 124**:32-38

Burger A. 1995. UNESCO/BAC BETCEN Molecular marker course manual. ARC-Roodeplaat, South Africa. 13-17 November 1995.

Caetano-Anollés G. 1996. Scanning nucleic acids by *in vitro* amplification: New developments and applications. **Nature Biotechnology 14:**1668-1674

Caetano-Anollés G, BJ Bassam and PM Gresshoff. 1991. DNA Amplification Fingerprinting: A strategy for genome analysis. **Plant Molecular Biology Reporter 9:**294-307

Carlson JE, LK Tulsieram, JC Glaubitz, VWK Luk, C Kauffeldt. 1991. Segregation op random amplified DNA markers in F1 progeny conifers. **Theoretical and Applied Genetics 83:**194-200

Chaparro JX, DJ Werner, D O'Malley and RR Sederoff. 1994. Targeted mapping and linkage analysis of morphological isozyme and RAPD markers in peach. **Theoretical and Applied Genetics 78**:805-815

Clark JM, 1988, Novel non-template nucleotide addition reaction reaction catalysed by prokaryotic and eucaryotic DNA polymerases. **Nucleic Acids Research 16:**9677-9686

Cregan PB. 1992. Simple sequence repeat DNA length polymorphisms. **Probe 2:**18-22

Dax E, O Livneh, E Aliskevicius, O Edelbaum, N Kedar, N Gavish, J Milo, F Geffen, A Blummenthal, HD Rabinowich and I Sela. 1998. A SCAR marker linked to the ToMV resistance gene, *Tm2*^{2} in tomato. **Euphytica 101:**73-77

Demeke T, DR Lynch, LM Kawchuk, GC Kozub and JD Amstrong. 1996. Genetic diversity of potato determined by random amplified polymorphic DNA analysis. **Plant Cell Reports 15**:662-667

Demeke T, LM Kawchuk and DR Lynch. 1993. Identification of potato cultivars and clonal variants by random amplified polymorphic DNA analysis. **American Potato Journal 70:**561-570

Dietmaier W and S Fabry, 1995, Protocol: Di-/Trinucleotide Sticky End Cloning (DI/TRISEC). in **PCR Aplication Manual** Boehringer Mannheim, Manheim Germany, pp136-140

Dietmaier W, S Fabry and R Schmitt. 1993. DISEC-TRISEC: di- and tri-nucleotide -stickey-end cloning of PCR amlpified DNA. **Nucleic Acids Research 21:**3603-3604

Dweikat I, S Mackenzie, M Levy and H Ohm. 1993. Pedigree assessment using RAPD-DGGE in cereal crop species. **Theoretical and Applied Genetics 85:**797-505

Edwards K, C Johnson and C Thompson. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. **Nucleic Acids Research 19**:1349

Fabry S and W Dietmaier, 1995, Review: PCR and Cloning. in **PCR Aplication Manual** Boehringer Mannheim, Manheim Germany, pp133-135

Francis HA, AR Leitch and RMD Koebner. 1995. Conversion of a RAPD-generated PCR product, containing a novel dispersed repetitive element, into a fast and robust assay for the presence of rye chromatin in wheat. **Theoretical and Applied Genetics 90:**6363

Francisco-Ortega J, HJ Newburry and BV Ford-Lloyd. 1993. Numerical analyses of RAPD data highlight the origin of tagasate (*Chamaecytisus proliferus* ssp, palmensis) in the Canary Islands. **Theoretical and Applied Genetics 87:**264-270

Gallego FJ and I Martinez. 1996. Molecular typing of rose cultivars using RAPDs. **Journal of Horticultural Science 71:**901-908

Görg R, U Schachtschabel, E Ritter, F Salamini and C Gebhardt. 1992. Discrimination among 136 tetraploid potato varieties by fingerprints using highly polymorphic DNA markers. **Crop Science 32**:815-819

Gu WK, NF Weeden, J Yu, and DH Wallace. 1995. Large-scale cost-effective screening of PCR products in marker-assisted selection applications. **Theoretical and Applied Genetics 91:**465-470

Hittalmani S, MR Fooland, T Mew, RL Rodriguez and N Huang. 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

Hoistings D, K Khairallah and D Conzalez-de-Leon. 1994. Laboratory protocols: CIMMYT applied molecular genetics laboratory. Second edition. Mexico DF:CIMMYT

Holton TA and MW Graham, 1991, A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. **Nucleic Acids Research 19:**1156

Horejsi T, JM Box and JE Staub. 1999. Efficiency of randomly amplified polymorphic DNA to sequence characterised amplified region marker conversion and their comparative polymerase chain reaction sensitivity in cucumber. **Journal of the American Society for Horticultural Science 124**:128-135

Hosaka K, M Mori and K Ogawa. 1994. Genetic relationship of Japanese potato cultivars assessed by RAPD analysis. **American Potato Journal 71**:535-546

Hosaka K, RE Hanneman. 1994. Random amplified polymorphic DNA markers detected in a segregating hybrid population of *Solanum chacoense* X *S. phureja*. **Japanese Journal of Genetics 69**:53-66

Howel EC, HJ Newbury, RL Swennen, LA Withers and BV Ford-Lloyd. 1994. The use of RAPD for identifying and classifying *Musa* germplasm. **Genome 37**:328-332

Hsiao KC, 1993, Exonuclease III induced ligase-free directional subcloning of PCR products. **Nucleic Acids Research 21:**5528-5529

Hu J and CF and Quiros. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. **Plant Cell Reports 10:**505-511

Innes MA and DH Gelfand. 1990. Optimisation of PCRs. In **PCR Protocols. A guide to methods and applications.** edited by Innes MA, DG Gelfand, JJ Sninsky and TJ White. Academic Press, Inc., San Diego, California, USA

Iqbal MJ and AL Rayburn. 1994. Stability of RAPD markers for determining cultivar specific DNA profiles in rye (*Secale cereale* L.). **Euphytica 75**: 215-220

Jansen van Rensburg WS. 1995, *In vitro* maintenance of potato germplasm. **Roodeplaat Bulletin 44:**16-17

Jansen van Rensburg WS. 1997. Identification of Potato cultivars. In **Potato Production in South Africa, with the emphasis on KwaZulu Natal**. edited by Lizel Urquart.

Jones CJ, KJ Edwards, S Castaglione, MO Winfield, F Sala, C van de Wiel, G Bredemeijer, B Vosman, M Matthes, A Daly, R Brettschneider, P Bettini, M Buiatti, E Maestri, A Malcevschi, N Marmiroli, R Aert, G Volckaert, J Rueda, R Linacero, A Vasquez and A Karp. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. **Molecular Breeding 3:**381-390

Kaluz S, K Kölbe and KBM Reid, 1992, Directional cloning of PCR products using exonuclease III. **Nucleic Acids Research 20:**4369-4370

Kesseli RV, I Paran and RW Michelmore. 1993. Efficient mapping of specifically targeted genomic regions and the tagging of these regions with reliable PCR-based genetic markers. In **Proceedings of the Symposium: Application of RAPD Technology to Plant Breeding**, ed. M. Neff. ASHS Publishers. St. Paul. pp 31-36

Klein-Lankhorst RM, A Vermunt, R Weide, T Liharska and P Zabel. 1991. Isolation of molecular markers for tomato (*L, esculentum*) using random amplified polymorphic DNA (RAPDs). **Theoretical and Applied Genetics 83:**108-144

Kovalic D, JH Kwak and B Weisblum, 1991, General method of direct cloning of DNA fragments generated by the polymerase chain reaction. **Nucleic Acids Research 19:**4560

Kresovich S, JGK Williams JR McFErson, JR Routman and BA Schaal. 1992. Characterisation of genetic identities and relationships of *Brassica olaracea* L. via a random amplified polymorphic DNA assay. **Theoretical and Applied Genetics 85:**190-196

Lahogue F, P This and A Bouquet. 1998. Identification of a codominant SCAR marker linked to seedlessness in grapevine. **Theoretical and Applied Genetics 97:**916-928

Lanham PG, RM Brennan, C Hackett, RJ McNicol. 1995. RAPD fingerprinting of blackcurrant (*Ribes nigrum* L) cultivars. **Theoretical and Applied Genetics 90**:166-172

Lee SJ and GA Penner. 1997. The conversion of RFLP markers to allele specific amplicons linked to QTLs governing malting quality in barley. **Molecular Breeding 3:**457-462

Li Q and CL Guy. 1996. Prolonged final extension time increase cloning efficiency of PCR products. **BioTechniques 21:**192-196

Livneh O, E Vardi, Y Stram, O Edelbaum and I Sela. 1992. The conversion of RFLP assay into PCR for the determination of purity in a hybrid pepper cultivar. **Euphytica 62**: 97-102

Lohff CJ and KB Cease, 1992, PCR using a thermostable polymerase with 3' to 5' exonuclease activity generates blunt products suitable for direct cloning. **Nucleic Acids Research 20:144**

Lorenz M, A Weihe and T Börner, 1997, Cloning and Sequencing of RAPD fragments amplified from mitochondrial DNA of male-sterile cytoplasm of sugar beet (*Beta vulgaris* L.) **Theoretical and Applied Genetics 94:**273-278

Maaß HI, M Klaas. 1995. Infra specific differentiation of garlic (*Allium sativum* L) by isozyme and RAPD markers. **Theoretical and Applied Genetics 91**:89-97

Marchuk D, M Drumm, A Saulino and FS Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. **Nucleic Acids Research 19:**1154

McDermott JM, U Brändle, F Dutly, UA Haemmerli, S Keller, KE Müller and MS Wolfe. 1994. Genetic variation in powdery mildew of barley: Development of RAPD, SCAR, and VNTR markers. **Phytopathology 84**:1316-1321

Mead DA, NK Prey, C Herrnstadt, RA Marcil and LM Smith. 1991. A universal method for the direct cloning of PCR amplified nucleic acid. **Bio/Technology 9:**657-663

Milbourne D, R Meyer, JE Bradshaw, E Baird, N Bonar, J Provan, W Powel and R Waugh. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationship in cultivated potato. **Molecular Breeding 3:**127-136

Morell MK, R Peakall, R Appels, LR Preston and HL Lloyd. 1995. DNA profiling techniques for plant variety identification. **Australian Journal of Experimental Agriculture 45:**807-819

Moreno S, Y Gogorcena, JM Ortiz. 1995. The use of RAPD markers for identification of cultivated grapevine (*Vitis vinifera* L). **Scientia Horticulturae 62**:237-243

Murashigi T and F Schoog. 1962.A revised medium for rapid growth and bio assays with tobacco tissue cultures. **Physiologia Plantarum 15:**473-497

Naqvi NI and BB Chattoo, 1996, Development of a sequence characterized ampified region (SCAR) based indirect selection method for a dominant blast-resistance gene in rice. **Genome 39**:26-30

Newbury HJ and BV Ford-Lloyd. 1993. The use of RAPD for assessing variation in plants. **Plant Growth Regulation 12**:43-51

Niewöhner J, F Salamini, C Gebhardt. 1995. Development of PCR assays diagnostic for RFLP marker alleles closely linked to alleles *Gro1* and *H1*, conferring resistance to the root cyst nematode *Globodera rostochiensis* in potato. **Molecular Breeding 1**:65-78

Nybom H.. 1994. DNA Fingerprinting - A Useful tool in fruit Breeding. **Euphytica 77**:59-64

Nybom H, SH Rogstad and BA Schaal. 1990. Genetic variation detected by use of the M13 "DNA fingerprint" probe in *Malus, Prunus* and *Rubus* (Rosaceae). **Theoretical and Applied Genetics 79:**153-156

Obara-Okey P and S Kako. 1998. Genetic diversity and identification of *Cymbidium* cultivars as measured by random amplified polymorphic DNA (RAPD) markers. **Euphytica 99:**95-101

Oliner JD, KW Kinzler and B Vogelstein,1993, *In vivo* cloning of PCR products in *E coli*. **Nucleic Acids Research 21:**5192-5197

Olson M, L Hood, C Cantor and D Botstein. 1989. A common language for mapping of the human genome. **Science 245**:1434-1435

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

Paz MM and RE Veilleux. 1997. Genetic diversity based on Random Amplified Polymorphic DNA (RAPD) and its relationship with the performance of diploid potato hybrids. **Journal of the American Society for Horticulture Science 122:**740-744

PSA Annual Report 1998 Potatoes South Africa annual report for the year 1997/1998. Pretoria, South Africa

Quiros CF, A Ceada, A Georgescu and J Hu. 1993. Use of RAPD markers in potato genetics: Segregations in diploid and tetraploid families. **American Potato Journal 70**:35-42

Rafalski JA and SV Tingey. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatelites and machines. **Trends in Genetics 9**:275-280

Rameau C, D Dénoue, F Fraval, K Haurogné, J Josserand, V Laucou, S Batge and IC Murfet. 1998. Genetic mapping in pea. 2. Identification of RAPD and SCAR markers linked to genes affecting plant architecture. **Theoretical and Applied Genetics 97:**916-928

Reiter RS, JGK Wiliams, KA Feldman, JA Rafalski, SV Tingey and PA Scholnic. 1992. Global and local genome mapping in *Arabisopsis thaliana* by using recombinant inbred lines and random polymorphic DNAs, **Proceedings of the National Academy of Science in U.S.A. 89:**1477-1481

Ronning CM and RJ Schnell. 1995. Using randomly amplified polymorphic DNA markers to identify *Annona* cultivars. **Journal of the American Society for Horticultural Science 120**:726-729

Roupe van der Voort JNAM, GJW Janssen, H Overmars, PM van Zandvoort, A van Norel, OE Scholten, R Janssen and J Bakker, 1999, Development of a PCR-based selection assay for root-knot nematode resistance (*Rmc1*) by comparative analysis of the *Solanum bulbocastanum* and *S. tuberosum* genome. **Euphytica** 106:**187-195**

Russel J, F Hosein, E Johnson, R Waugh and W Powell.1993. Genetic differentiation of cocoa (*Theobroma cocao* L.) populations revealed by RAPD analysis. **Molecular Ecology 2:**89-97

Saiki RK, S Scharf, F Faloona, KB Mullis, GT Horn, HA Erlich and N Arnheim. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. **Science 230:**1350-1354

Saiki RK. 1990. Amplification of genomic DNA. In **PCR Protocols. A guide to methods and applications.** edited by Innes MA, DG Gelfand, JJ Sninsky and TJ White. Academic Press, Inc., San Diego, California, USA

Scharf SJ, GT Horn and AH Erlich, 1986, Direct cloning and sequence analysis of enzymatically amplified genomic sequences . **Science 233:**1067-1078

Schierwater B and A Ender. 1993, Different thermostable DNA polymerase may amplify different RAPD products. **Nucleic Acids Research 21**:4647-4648

Schnell RJ, CM Ronning, RJ Knight. 1995. Identification of cultivars and validation of genetic relationships in *Mangifera indica* L using RAPD markers. **Theoretical and Applied Genetics 90**:269-274

Schuldiner AR, LA Schott and J Roth, 1990, PCR induced (ligase free) subclonign: a rapid, relaible method to suclone polymerase chain reaction (PCR) products. **Nucleic Acids Research 18:**1920

Simonsen W, WK Heneen. 1995. Inheritance of isozymes in *Brassica campestris* L. and genetic divergence among different species of Brassiceae. **Theoretical and Applied Genetics 91**:353-360

Sosinski B. and DS Douches. 1996. Using polymerse chain reaction-based DNA amplification to fingerprint North American potato cultivars. **HortScience 31**:130-133

Spooner DM, J Tivang, J Nienhuis, JT Miller, DS Douches and A Contreras-M. 1996. Comparison of four molecular markers in measuring relationships among the wild potato relatives *Solanum* section *Etuberosum* (subgenus *Potato*). **Theoretical and Applied Genetics 92:**532-540

Staub JE, FC Serquen and M Gupta. 1996. Genetic markers, map construction, and their application in plant breeding. **HortScience 31**:792-141

Stiles JI, C Lemme, S Sondhur, MB Morshidi and R Manshardt. 1993. Using random amplified polymorphic DNA for evaluating papaya cultivars. **Theoretical and Applied Genetics 85:**697-701

Tingey SV, JA Rafalski and JGK Williams. 1993. Genetic analysis with RAPD markers. In **Proceedings of the Symposium: Application of RAPD Technology to Plant Breeding**, ed. M. Neff. ASHS Publishers. St Paul. pp 3-8

Van Miltenburg R, B Ruger, S Grunwald-Janho, M Leons and C Schroder. 1995. The DIG system user's guide for filter hybridisation. Mannheim, Germany

Vicario F, GG Vendramin, P Rossi, P Lio, R Giannini. 1995. Allozyme, chloroplast DNA and RAPD markers for determining genetic relationships between *Abies alba* and the relic population of *Abies nebrodensis*. **Theoretical and Applied Genetics 90**:1012-1018

Virk PS, HJ Newbury, MT Jackson, BV Ford-Lloyd. 1995. The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. **Theoretical and Applied Genetics 90**:1049-1055

Visser A. 1997. Potato cultivars and planting times. In **Potato Production in South Africa, with the emphasis on KwaZulu Natal**. Edited by Lizel Urquart.

Visser A. 1999. Aartappel cultivars en plantye. in **Aartappelverbouing in Suid Africa met die klem op die Vrystaat en Suid Kaap**. edited by W Jansen van Rensburg and A Mcleod

Vos P, R Hogers, M Bleeker, M Riejans, T van de Lee, M Hornes *et al* 1995, AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research 23:**4407-4414

Walton M. 1993. Molecular markers: Which one to use? **Seed World 131**:22-29

Wang G, S Castiglione, J Zhang, R Fu, F Ma, W Li, Y Sun, and F Sala. 1994. Hybrid rice (*Oryza sativa* L.): Identification and Parentage determination by RAPD fingerprinting. **Plant Cell Reports 14**:112-115

Weising K, H Nybom, K Wolff and W Meyer. 1995. DNA fingerprinting in plants and fungi. CRC Press, Inc. Boca Raton.

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

Wilde J, R Waugh and W Powel. 1992. Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. **Theoretical and Applied Genetics 83:**871-877

Wilkie S, PG Isaac and RJ Slater. 1993. Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Alium*. **Theoretical and Applied Genetics 86:**497-504

Williams JGK, AR Kubelik, KJ Livak, JA Rafalski and SV Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Res. 18**:6231-6235

Williams JGK, MK Hanafey, JA Rafalski and SV Tingey. 1993. Genetic analysis using random amplified polymorphic markers. **Methods in Enzymology 218:**704-740

Wolf K and J Peters-van Rijn. 1993. Rapid detection of genetic variability in chrysanthemum (*Dendranthema grandiflora* Tzvelev) using random primers. **Heredity 71**: 335-341

Wu L and H Lin. 1994. Identifying Buffalograss [*Buchloe dactyloides* (Nutt) Engelm.] cultivar breeding lines using random amplified polymorphic DNA

(RAPD) markers. **Journal of the American Society for Horticultural Science 119**:126-130

Xu Y, DJ Wilson, S Arulsekar and AT Bakalinsky. 1995. Sequence specific polymerase chain-reaction markers derived from randomly amplified polymorphic DNA markers for fingerprinting Grape (*Vitis*) rootstocks. **Journal of the American Society for Horticultural Science 120**:714-720

Xu Y, MS Clark and E Pehu. 1993. Use of RAPD markers to screen somatic hybrids between *Solanum tuberosum* and *S. brevidens*. **Plant Cell Reports 12**:107-109

Yang X and C Quiros. 1991. Identification and classification of Celery cultivars with RAPD markers. **Theoretical and Applied Genetics 86:**205-212

Zhou M-Y, SE Clark and CE Gomes-Sanchez. 1995. Universal cloning method by TA strategy, **BioTechniques 19**:34-35

Appendix 1
Upov creteria for morphological cultivar identification

Criteria for the morphological descriptions of potatoes as used by UPOV and the Department of Plant and Quality Control (N Steyn and T Barnard personal communication).

UPOV DESCRIPTION OF CULTIVAR/LINE:

1.Lightsprouts: Size	**1**	**Small** **Medium** **Large**	**3** **5** **7**
2. Lightsprouts: Shape	**1**	**Spherical** **Ovoid** **Conical** **Broad Cylindrical** **Narrow Cylindrical**	**1** **2** **3** **4** **5**
3. Lightsprouts: Anthocyanin Coloration of base	**1**	**Red-Violet** **Blue-Violet**	**1** **2**
4. Lightsprouts: Intensity of Anthocyanin Coloration of base	**1**	**Very weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
5. Lightsprouts: Pubescence of Base	**1**	**Very Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
6. Lightsprouts: Size of Tip	**1**	**Very Small** **Small** **Medium** **Large** **Very Large**	**1** **3** **5** **7** **9**
7. Lightsprouts: Habit of Tip	**1**	**Closed** **Medium** **Open**	**3** **5** **7**
8. Lightsprouts: Intensity of Anthocyanin Coloration of Tip	**1**	**Very Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
9. Lightsprouts: Pubescence of Tip	**1**	**Absent or Very weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**

10. Lightsprouts: Number of Roottips	**1**	**Few** **Medium** **Many**	**3** **5** **7**
11. Lightsprouts: Protrusion of Lenticels	**1**	**Weak** **Medium** **Strong**	**3** **5** **7**
12. Lightsprouts: Length of Lateral Shoots	**1**	**Short** **Medium** **Long**	**3** **5** **7**
13. Plant: Height	**2**	**Very Short** **Short** **Medium** **Tall** **Very Tall**	**1** **3** **5** **7** **9**
14. Plant: Type	**2**	**Stem-type** **Intermediate-Type** **Leaf-Type**	**1** **2** **3**
15. Plant: Growth Habit	**2**	**Erect** **Semi-Erect** **Spreading**	**3** **5** **7**
16. Stem: Thickness of Main Stem	**2**	**Thin** **Medium** **Thick**	**3** **5** **7**
17. Stem: Extension of Anthocyanin Coloration	**2**	**Absent or Very Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
18. Leaf: Size	**2**	**Very Small** **Small** **Medium** **Large** **Very Large**	**1** **3** **5** **7** **9**
19. Leaf: Silhouette	**2**	**Closed** **Medium** **Open**	**3** **5** **7**
20. Leaf: Intensity of Green Colour	**2**	**Light** **Medium** **Dark**	**3** **5** **7**
21. Leaf: Extension of Anthocyanin Coloration of Midrib	**2**	**Absent or Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
22. Leaflet: Size	**2**	**Very Small** **Small** **Medium** **Large**	**1** **3** **5** **7**

		Very Large	9
23. Leaflet: Width	2	Narrow Medium Broad	3 5 7
24. Leaflet Frequency of Coalescence	2	Low Medium High	3 5 7
25. Leaflet: Waviness of Margin	2	None or Very Weak Weak Medium Strong Very Strong	1 3 5 7 9
26. Leaflet: Depth of Veins	2	Shallow Medium Deep	3 5 7
27. Leaflet: Anthocyanin Pigmentation of Blade of the Young Leaflets at Apical Rosette	2	Absent Present	3 9
28. Leaflet: Glossiness of the Upperside	2	Dull Medium Glossy	3 5 7
29. Leaf Midrib: Frequency of Secondary Leaflets	2	Nil or Very Low Low Medium High Very High	1 3 5 7 9
30. Terminal Leaflet: Frequency of Secondary Leaflets	2	Nil or Very Low Low Medium High Very High	1 3 5 7 9
31. Lateral Leaflet: Frequency of Secondary Leaflets	2	Nil or Very Low Low Medium High Very High	1 3 5 7 9
32. Lateral Leaflet: Size of Secondary Leaflet	2	Small Medium Large	3 5 7
33. Inflorescence: Size	3	Small Medium Large	3 5 7
34. Inflorescence: Anthocyanin coloration of peduncle	3	Absent or Very Weak Weak Medium Strong Very Strong	1 3 5 7 9

35. Plant: Frequency of Flowers	**3**	**Nil or Very Low** **Low** **Medium** **High** **Very High**	**1** **3** **5** **7** **9**
36. Flower: Anthocyanin Coloration of Bud	**3**	**Absent or Very Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
37. Flower Corolla: Size	**3**	**Very Small** **Small** **Medium** **Large** **Very Large**	**1** **3** **5** **7** **9**
38. Flower Corolla: Colour of Inner Side	**3**	**White** **Red-Violet** **Blue-Violet**	**1** **2** **3**
39. Flower Corolla: Intensity of Anthocyanin Coloration of Innerside in Coloured flowers	**3**	**Very Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**
40. Flower Corolla: Intensity of Anthocyanin Coloration of Innerside in White flowers	**3**	**Absent** **Present**	**1** **9**
41. Flower Corolla: Size of White Tips in Coloured Flower	**3**	**Small** **Medium** **Large**	**3** **5** **7**
42. Plant: Frequency of Fruits	**4**	**Absent or Very Few** **Few** **Medium** **Many** **Very Many**	**1** **3** **5** **7** **9**
43. Plant: Time of Maturity	**4**	**Very Early** **Early** **Medium** **Late** **Very Late**	**1** **3** **5** **7** **9**
44. Tuber: Shape **(100 x Length / Width)**	**5**	**Round (< 109)** **Short-Oval (110-129)** **Oval (130-149)** **Long-Oval (150-169)** **Long (170-199)** **Very Long (> 200)**	**1** **2** **3** **4** **5** **6**
45. Tuber: Depth of Eyes	**5**	**Very Shallow** **Shallow** **Medium** **Deep** **Very deep**	**1** **3** **5** **7** **9**

46. Tuber: Smoothness of Skin	**5**	**Smooth** **Medium** **Rough**	**3** **5** **7**
47. Tuber: Colour of Skin	**5**	**Yellow** **Red** **Blue** **Red parti-coloured** **Blue parti-colored**	**1** **2** **3** **4** **5**
48. Tuber: Colour of Base of Eye	**5**	**Yellow** **Red** **Blue**	**1** **2** **3**
49. Tuber: Colour of Flesh	**5**	**White** **Cream** **Light Yellow** **Yellow** **Dark Yellow**	**1** **2** **3** **4** **5**
50. Tuber: Anthocyanin Coloration of Skin in Reaction to Light (Only Yellow Skinned Varieties) **10 days Daylight or 150 hours Artificial light**	**5**	**Absent or Very Weak** **Weak** **Medium** **Strong** **Very Strong**	**1** **3** **5** **7** **9**

Appendix 2
Aligned sequences for fragment UBC 582^{900}

The two sequences, obtained for the fragment UBC 582^{900} with the two sequencing primers, aligned with the aid of Genepro.

```
Seq 1   ubi582(T7)
Seq 2   INV Ubi582(U19) starting and ending at primer sequence

Seq 1   GG--TATAGACGT---AT--AATG---TTCTT--GTGAGAAAT--AAATAA--TATTGV-     43
        **  **    * *   *   ****   * ***  * * *****  ******  *****
Seq 2   GGTATARDRMC-TBBDAAWAAATGGYATCCTTGGGGGGGAAATTAAAATAAAWTATTGGA     59

Seq 1   -ATT-GTGTAT--GT-AAATA-GTTACATTGAGA---CCC--TA-TTTATA-GA-CACT-     88
         *** *** *   ** *** * ****** *  *    ***  ** *****  ** ****
Seq 2   AATTKGTGKAAWGGTAAAAWARGTTACA-TWTGRRRMCCCCTTATTTTATTRGACCACTT    118

Seq 1   A-CATTGGAAT-CCTTTT-CCAACT--AGGATT-CCTATT-CTTA--TTCCTATTCCTAT    139
        * ********* ****** ******   * *** ****** ****  * ***********
Seq 2   ACCATTGGAATCCCTTTTCCCAACTWRRGRATTCCCTATTCCTTAATTCCCTATTCCTAT    178

Seq 1   TCCTATTCCTATTCTAAGTAGGAAATATTTATTTCTATTCTAACACTCCCCCTCAAGCTG    199
        ******************* ****************************************
Seq 2   TCCTATTCCTATTCTAAGTRGGAAATATTTATTTCTATTCTAACACTCCCCCTCAAGCTG    238

Seq 1   GTGCATACAAATCATATGTACCTAGCTTGTTACAGATGTAATTAATACGAGGACATGTGA    259
        ************************************************************
Seq 2   GTGCATACAAATCATATGTACCTAGCTTGTTACAGATGTAATTAATACGAGGACATGTGA    298

Seq 1   GGGACTTGGTGAAGATATCTGCAAGCTGATCATTTGACTTCACAAATTTCGTAACAATAT    319
        ************************************************************
Seq 2   GGGACTTGGTGAAGATATCTGCAAGCTGATCATTTGACTTCACAAATTTCGTAACAATAT    358

Seq 1   CTCCCGAGAGTATCTYTTCTCTGACAAAGTGACAGTCAATCTCAATGTGCTTAGTCCTTT    379
        *************** ******************************************
Seq 2   CTCCCGAGAGTATCTTTTCTCTGACAAAGTGACAGTCAATCTCAATGTGCTTAGTCCTTT    418

Seq 1   CATGAAACACTGGATTTGATGCGATATGAAGAGCCGCTTGATTATCACACATAAGTTCCA    439
        ************************************************************
Seq 2   CATGAAACACTGGATTTGATGCGATATGAAGAGCCGCTTGATTATCACACATAAGTTCCA    478

Seq 1   TCTGATCAATTTTGCCAAATTT-AGTTCTCCCAGCAACTGTTTGATCCAAACTAGCTCMM    498
        ********************** ****************************** ****
Seq 2   TCTGATCAATTTTGCCAAATTTTAGTTCTCCCAGCAACTGTTTGATCCAAACTRGCTCAC    538

Seq 1   MAAGTTGTTGTCGCCATTGCBCGATATTCTGAGTTCTGCRCTAGATCGVGCGAA-CACAT    557
         ****************** *********** ****** ******** ** ** ** *
Seq 2   -AAGTTGTTGTCGCCATTGCTCGATATTCTGA-TTCTGCACTAGATCGAGC-AACCASAD    595

Seq 1   TCTTTTTCTYAMAC--CYTCCACGAGACCMMAATTACCYCCYACCTAAAAACGCGAWATT    615
         ******** *  *    ******* ***  ******* ** ** ***** *** * **
Seq 2   GCTTTTTCTTACTCT---TCCACGACACCA-AATTACCTCCTAC-TAAAA-CGCAATAT-    648

Seq 1   CCGGGTGTCGAACGTCTSTCAGAGGGGKGATCTTGCCCAATCAGCGKCTATATATCCCMM    675
        **** ************ ********  ****************** **********
Seq 2   CCGGATGTCGAACGTCTGTCAGAGGGT-GATCTTGCCCAATCAGCGTCTATATATCCAA-    706
```

```
Seq 1  TGGTRTGCBCATGGGCSTTGATCCTCAAA-AGRRGTAGTCCYTTTWYCGGGGGGCTGRCT   734
       ** * *** ***** * ************ **   ****** ***  **** * *** **
Seq 2  TGATATGCTCATGGCC-TTGATCCTCAAAGAG---TAGTCC-TTTACCGGGAG-CTGACT   760

Seq 1  TTATATATCGSVGARTACGAGCAACTGBTTTCCCAMTG-CGCWTC-CCAAMGGGGGGMMT   792
       **********  ** ***** ****** ** **** ** *   ** * ** ****    *
Seq 2  TTATATATCGCAGAATACGAACAACTGCTT-CCCAATGACTA-TCAC-AA-GGGGAACAT   816

Seq 1  AA-CTGGCTTHHAC---MCCTTTCGGGAAA-GAG-TGTCCCGG-CNYTGYCCC        838
       ** *** ***  **    * ** * ***** *** ****  ** *   * * *
Seq 2  AAACTGACTT--ACAABAC-TTACAGGAAAAGAGATGTCA-GGTCTA-GTCACTGTGAGA   871

Seq 2  TAATTCAACTTTCCAACAAGCCGTCTATACC                               902

Seq 1       1-838
Seq 2       1-902
Aligned     893
Matches     737
Mismatches  156
Homology    82%
```

Appendix 3

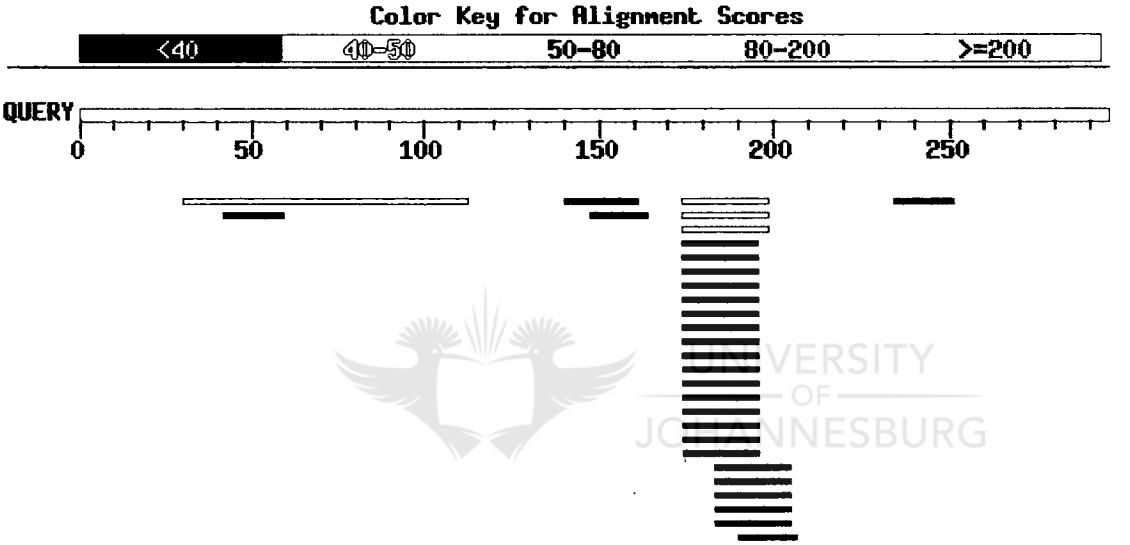
Result of the blast search for homologies with fragment UBC 509^{300}. Results is given as printed out from website. **http://www.ncbi.nlm.nih.gov/blast/blast.cgi**. The "dbest" database was used.

Commencing search, please wait for results.

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

If you have any problems or questions with the results of this search please refer to the **BLAST FAQs**

Distribution of 30 Blast Hits on the Query Sequence



Sequences producing significant alignments:	Score (bits)	E Value
gb|AI487826.1|AI487826 EST246148 tomato ovary, TAMU Lycoper...	68	1e-09
dbj|AV024796.1|AV024796 AV024796 Mus musculus adult C57BL/6...	42	0.061
dbj|AV163022.1|AV163022 AV163022 Mus musculus head C57BL/6J...	42	0.061
gb|W90954|W90954 mf84g07.r1 Soares mouse embryo NbME13.5 14...	42	0.061
gb|AA994297|AA994297 ou50c04.s1 NCI_CGAP_Br2 Homo sapiens c...	36	3.8
gb|R26716|R26716 yh51c03.s1 Homo sapiens cDNA clone 133252 3'.	36	3.8
gb|AI802848.1|AI802848 tt57g04.x1 NCI_CGAP_HSC4 Homo sapien...	36	3.8
gb|H09303|H09303 yl95a01.s1 Homo sapiens cDNA clone 45786 3'.	36	3.8
gb|AI537778.1|AI537778 tp29b08.x1 NCI_CGAP_Gas4 Homo sapien...	36	3.8
gb|AA759307|AA759307 ah90h08.s1 Soares NFL T GBC S1 Homo sa...	36	3.8
gb|AA541560|AA541560 ni81g07.s1 NCI_CGAP_Pr12 Homo sapiens ...	36	3.8
gb|AI689206.1|AI689206 tx12d10.x1 NCI_CGAP_Ut4 Homo sapiens...	36	3.8
gb|AI423715.1|AI423715 tf34c05.x1 NCI_CGAP_Brn23 Homo sapie...	36	3.8
gb|AI811764.1|AI811764 tw75g12.x1 NCI_CGAP_Ut3 Homo sapiens...	36	3.8
gb|AI713179.1|AI713179 UI-R-Y0-aaz-b-08-0-UI.s1 UI-R-Y0 Rat...	36	3.8
gb|AI923944.1|AI923944 wn77h06.x1 NCI_CGAP_Ut1 Homo sapiens...	36	3.8
gb|AI421255.1|AI421255 tf14d07.x1 NCI_CGAP_Brn23 Homo sapie...	36	3.8

gb|AA528511|AA528511 nf01b03.s1 NCI_CGAP_Kid1 Homo sapiens ... 36 3.8
gb|AI774618.1|AI774618 EST255718 tomato resistant, Cornell ... 36 3.8
gb|AI694871.1|AI694871 we52f09.x1 NCI_CGAP_Co3 Homo sapiens... 36 3.8
gb|AI232088|AI232088 EST228776 Normalized rat kidney, Bento... 36 3.8
gb|AI043521|AI043521 L30-524T3 Ice plant Lambda Uni-Zap XR ... 36 3.8
gb|AI206227|AI206227 qr27g11.x1 NCI_CGAP_GC6 Homo sapiens c... 36 3.8
gb|AI370863|AI370863 ta63h06.x1 Soares_total_fetus_Nb2HF8_9... 36 3.8
gb|AA187189|AA187189 zp60e05.s1 Stratagene endothelial cell... 36 3.8
gb|AI459455|AI459455 ar81d10.x1 Barstead colon HPLRB7 Homo ... 36 3.8
gb|AA446591|AA446591 zw84e06.s1 Soares total fetus Nb2HF8 9... 36 3.8
gb|H17623|H17623 ym40b06.s1 Homo sapiens cDNA clone 50591 3'. 36 3.8
gb|AI753762.1|AI753762 cr14h04.x1 Jia bone marrow stroma Ho... 36 3.8
gb|AI094717|AI094717 qa08e03.x1 NCI_CGAP_Brn23 Homo sapiens... 36 3.8

gb|AI487826.1|AI487826 EST246148 tomato ovary, TAMU Lycopersicon esculentum cDNA clone cLED13D22, mRNA sequence
Length = 521

Score = 67.9 bits (34), Expect = 1e-09
Identities = 74/85 (87%), Gaps = 3/85 (3%)
Strand = Plus / Plus

```
Query: 31  agactggagaactcattggtggtgtgtc--agttgcatgtcttggaagtccatccgcccg 88
           ||||||||||||||| ||| |||| |   ||||||||||||||||||| ||||||||
Sbjct: 438 agactggagaactcatcggttgtgtatgcaagttgcatgtcttggaagtcaatccgccca 497


Query: 89  ttggaataaaggatattcagtggac 113
           |||||| | ||||| ||||||||||
Sbjct: 498 ttggaaaacaggat-ttcagtggac 521
```


dbj|AV024796.1|AV024796 AV024796 Mus musculus adult C57BL/6J lung Mus musculus cDNA clone 1200007E20, mRNA sequence
Length = 146

Score = 42.1 bits (21), Expect = 0.061
Identities = 24/25 (96%)
Strand = Plus / Plus

```
Query: 175 tacatccattcattgaataaacatt 199
           ||||||| ||||||||||||||||
Sbjct: 110 tacatcctttcattgaataaacatt 134
```

dbj|AV163022.1|AV163022 AV163022 Mus musculus head C57BL/6J 13-day embryo Mus musculus cDNA clone 3110006K15, mRNA sequence
Length = 120

Score = 42.1 bits (21), Expect = 0.061
Identities = 24/25 (96%)
Strand = Plus / Plus

```
Query: 175 tacatccattcattgaataaacatt 199
           ||||||| ||||||||||||||||
Sbjct: 95  tacatcctttcattgaataaacatt 119
```

gb|W90954|W90954 mf84g07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 421020 5'
Length = 453

Score = 42.1 bits (21), Expect = 0.061
Identities = 24/25 (96%)
Strand = Plus / Minus

```
Query: 175 tacatccattcattgaataaacatt 199
           ||||||| ||||||||||||||||
Sbjct: 70  tacatcctttcattgaataaacatt 46
```

gb|AA994297|AA994297 ou50c04.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1631238 3'
Length = 369

Score = 36.2 bits (18), Expect = 3.8
Identities = 21/22 (95%)
Strand = Plus / Minus

```
Query: 175 tacatccattcattgaataaac 196
           ||||||| ||||||||||||||
Sbjct: 53  tacatcctttcattgaataaac 32
```

gb|R26716|R26716 yh51c03.s1 Homo sapiens cDNA clone 133252 3'.
Length = 432

Score = 36.2 bits (18), Expect = 3.8
Identities = 21/22 (95%)
Strand = Plus / Minus

```
Query: 175 tacatccattcattgaataaac 196
           ||||||| ||||||||||||||
Sbjct: 38  tacatcctttcattgaataaac 17
```

gb|AI802848.1|AI802848 tt57g04.x1 NCI_CGAP_HSC4 Homo sapiens cDNA clone IMAGE:2244918 3', mRNA sequence
Length = 220

Score = 36.2 bits (18), Expect = 3.8
Identities = 21/22 (95%)
Strand = Plus / Plus

```
Query: 175 tacatccattcattgaataaac 196
           ||||||| ||||||||||||||
Sbjct: 144 tacatcctttcattgaataaac 165
```

gb|H09303|H09303 yl95a01.s1 Homo sapiens cDNA clone 45786 3'.
Length = 503

Score = 36.2 bits (18), Expect = 3.8
Identities = 21/22 (95%)
Strand = Plus / Minus

```
Query: 184 tcattgaataaacattcaacat 205
           ||||| ||||||||||||||||
```

```
Sbjct: 196 tcatttaataaacattcaacat 175
```

gb|AI537778.1|AI537778 tp29b08.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2189175 3',
mRNA sequence
Length = 305

Score = 36.2 bits (18), Expect = 3.8
Identities = 21/22 (95%)
Strand = Plus / Minus

```
Query: 184 tcattgaataaacattcaacat 205
           ||||| ||||||||||||||||
Sbjct: 242 tcatttaataaacattcaacat 221
```

gb|AA759307|AA759307 ah90h08.s1 Soares NFL T GBC S1 Homo sapiens cDNA clone 1326399 3'
Length = 482

Score = 36.2 bits (18), Expect = 3.8
Identities = 21/22 (95%)
Strand = Plus / Minus

```
Query: 175 tacatccattcattgaataaac 196
           ||||||| ||||||||||||||
Sbjct: 37  tacatcctttcattgaataaac 16
```

Appendix 4

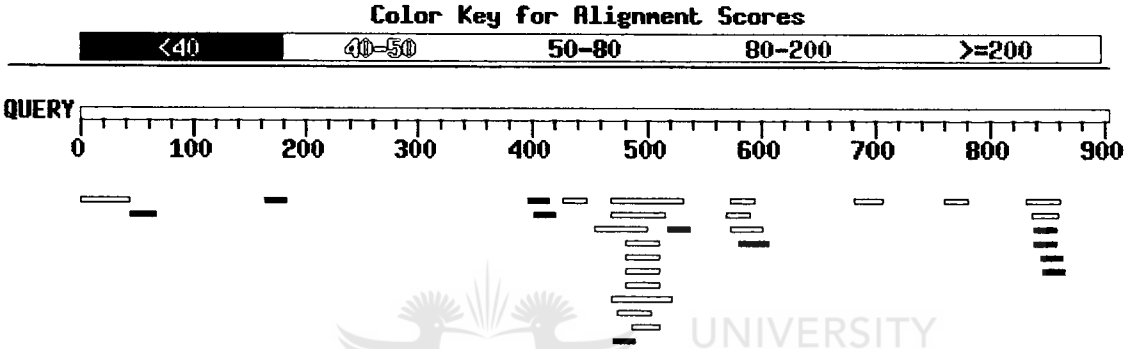
Result of the blast search for homologies with fragment UBC 582^{900}. Results is given as printed out from website **http://www.ncbi.nlm.nih.gov/blast/blast.cgi**. The "dbest" database was used.

Commencing search, please wait for results.

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

If you have any problems or questions with the results of this search please refer to the **BLAST FAQs**

Distribution of 31 Blast Hits on the Query Sequence



Sequences producing significant alignments:	Score (bits)	E Value
emb|AJ223825|LEAJ3825 Lycopersicon esculentum cv Red River ...	70	1e-09
emb|X67599|LEHSF8 L.esculentum Le-hsf8 gene for heat stress...	58	4e-06
emb|AL049862.1|ATF18B3 Arabidopsis thaliana DNA chromosome ...	48	0.004
gb|AC006219.8|ATAC006219 Arabidopsis thaliana chromosome II...	46	0.015
emb|AJ243356.1|PSA243356 Pisum sativum retrotransposon PDR1...	44	0.059
emb|AJ243357.1|PHU243357 Pisum humile retrotransposon PDR1,...	44	0.059
emb|AJ243358.1|PEL243358 Pisum elatius retrotransposon PDR1...	44	0.059
emb|X66399|PSRETRO P.sativum retrotransposon	44	0.059
gb|AC005149|AC005149 Drosophila melanogaster DNA sequence (...	44	0.059
gb|AC004247|AC004247 Drosophila melanogaster (P1 DS07773 (D...	42	0.23
gb|AC005825.2|ATAC005825 Arabidopsis thaliana chromosome II...	42	0.23
emb|X68786|BJCANRD B.juncea Xle7-2EB gene	42	0.23
gb|U12980|SCU12980 Saccharomyces cerevisiae chromosome I le...	40	0.92
emb|Z81102.1|CEM02B1 Caenorhabditis elegans cosmid M02B1, c...	40	0.92
gb|M92280|PPVPOLPRO Plum pox potyvirus polyprotein gene, co...	40	0.92
gb|AF024497|CELZC308 Caenorhabditis elegans cosmid ZC308	40	0.92
emb|AL078469.2|ATF25O24 Arabidopsis thaliana DNA chromosome...	40	0.92
emb|AJ243355.1|PSA243355 Pisum sativum retrotransposon PDR1...	40	0.92
gb|AC005199|AC005199 Homo sapiens chromosome 16, cosmid clo...	40	0.92
gb|AC003988|AC003988 Human PAC clone DJ1171L06 from 7q21, c...	38	3.6
gb|AC007385.3|AC007385 Homo sapiens clone NH0325M10, comple...	38	3.6
emb|AL022727|HS80I19 Human DNA sequence from clone 80I19 on...	38	3.6
gb|AF072732|AF072732 Lycopersicon esculentum polygalacturon...	38	3.6

emb|AL009179|HS97D16 Homo sapiens DNA sequence from PAC 97D...	38	3.6
emb|AL034417.14|HS215D11 Human DNA sequence from clone 215D...	38	3.6
dbj|AB008772|AB008772 Triticum aestivum retrotransposon Tar...	38	3.6
gb|AF016429|CELT21H3 Caenorhabditis elegans cosmid T21H3	38	3.6
gb|K01580|DRORGM101 d.melanogaster 28s rrna 5' region, 5 kb...	38	3.6
gb|AC005255|AC005255 Homo sapiens chromosome 19, CIT-HSP-14...	38	3.6
emb|X51968|DMRER1DM Drosophila melanogaster DNA for type I ...	38	3.6

emb|AJ223825|LEAJ3825 Lycopersicon esculentum cv Red River unknown sequence PCR random
amplified RAPD band 10
Length = 472

Score = 69.9 bits (35), Expect = 1e-09
Identities = 41/43 (95%)
Strand = Plus / Minus

```
Query: 1  ggtatagacggcttgttggaaagttgaattatctcacagtgac 43
          |||| ||||| ||||||||||||||||||||||||||||||||
Sbjct: 45 ggtacagacgacttgttggaaagttgaattatctcacagtgac 3
```

emb|X67599|LEHSF8 L.esculentum Le-hsf8 gene for heat stress transcription factor 8
Length = 9005

Score = 58.0 bits (29), Expect = 4e-06
Identities = 56/65 (86%)
Strand = Plus / Minus

```
Query: 469  aatccagtgtttcatgaaaggactaagcacattgagattgactgtcactttgtcagagaa 528
            ||||| |||||||| ||||| || || ||||| || |||||||| |||||| ||||||||
Sbjct: 8516 aatcctgtgtttcacgaaagaacaaaacacatagatattgactgccactttatcagagaa 8457

Query: 529  aagat 533
            |||||
Sbjct: 8456 aagat 8452
```

emb|AL049862.1|ATF18B3 Arabidopsis thaliana DNA chromosome 3, BAC clone F18B3
Length = 89469

Score = 48.1 bits (24), Expect = 0.004
Identities = 42/48 (87%)
Strand = Plus / Plus

```
Query: 469   aatccagtgtttcatgaaaggactaagcacattgagattgactgtcac 516
             |||||||| |||||||||||||| ||| |||| ||||| || ||||||
Sbjct: 13667 aatccagtctttcatgaaaggaccaagtacatagagatagattgtcac 13714
```

gb|AC006219.8|ATAC006219 Arabidopsis thaliana chromosome II BAC T14C8 genomic sequence, complete
sequence
Length = 82532

Score = 46.1 bits (23), Expect = 0.015
Identities = 29/31 (93%)
Strand = Plus / Plus

```
Query: 455   ttcatatcgcatcaaatccagtgtttcatga 485
             ||||||||||| |||| ||||||||||||||
Sbjct: 13059 ttcatatcgcaacaaacccagtgtttcatga 13089

 Score = 40.1 bits (20), Expect = 0.92
 Identities = 26/28 (92%)
 Strand = Plus / Minus

Query: 473   cagtgtttcatgaaaggactaagcacat 500
             ||||||| |||||||||||||| |||||
Sbjct: 29415 cagtgttccatgaaaggactaaacacat 29388
```

emb|AJ243356.1|PSA243356 Pisum sativum retrotransposon PDR1, PCR product of rearranged copy,
strain JI2547
Length = 1410

```
 Score = 44.1 bits (22), Expect = 0.059
 Identities = 28/30 (93%)
 Strand = Plus / Plus

Query: 481  catgaaaggactaagcacattgagattgac 510
             ||||| |||||||| ||||||||||||||
Sbjct: 1233 catgacaggactaaacacattgagattgac 1262
```

emb|AJ243357.1|PHU243357 Pisum humile retrotransposon PDR1, PCR product of rearranged copy,
strain JI1794
Length = 1443

```
 Score = 44.1 bits (22), Expect = 0.059
 Identities = 28/30 (93%)
 Strand = Plus / Plus

Query: 481  catgaaaggactaagcacattgagattgac 510
             ||||| |||||||| ||||||||||||||
Sbjct: 1266 catgacaggactaaacacattgagattgac 1295
```

emb|AJ243358.1|PEL243358 Pisum elatius retrotransposon PDR1, PCR product of rearranged copy,
strain JI254
Length = 2099

```
 Score = 44.1 bits (22), Expect = 0.059
 Identities = 28/30 (93%)
 Strand = Plus / Plus

Query: 481  catgaaaggactaagcacattgagattgac 510
             ||||| |||||||| ||||||||||||||
Sbjct: 1922 catgacaggactaaacacattgagattgac 1951
```

emb|X66399|PSRETRO P.sativum retrotransposon
Length = 3925

```
 Score = 44.1 bits (22), Expect = 0.059
```

```
 Identities = 28/30 (93%)
 Strand = Plus / Plus

Query: 481  catgaaaggactaagcacattgagattgac 510
            ||||| |||||||| ||||||||||||||
Sbjct: 3571 catgacaggactaaacacattgagattgac 3600
```

gb|AC005149|AC005149 Drosophila melanogaster DNA sequence (P1s DS03364 (D237) and DS02911
(D145)), complete sequence [Drosophila melanogaster]
Length = 98942

```
 Score = 44.1 bits (22), Expect = 0.059
 Identities = 28/30 (93%)
 Strand = Plus / Plus

Query: 832   tttacacaaatttccaataatttattttaa 861
              ||||||||||||||| | ||||||||||||
Sbjct: 74622 tttacacaaatttcaagtaatttattttaa 74651
```