The application of DNA fingerprinting and marker-assisted backcross selection in breeding for sunflower high oleic acid content lines

By

Tshediso Andrew Mokhele

Dissertation submitted in fulfilment of the requirements for the degree

MASTERS OF SCIENCE (BOTANY)

School of Environmental Science and Development: Botany
Faculty of Natural Science
North-West University (Potchefstroom Campus)

Supervisor
Dr. C.M.S. Mienie
Co-Supervisor
Prof. S. Barnard

2013

DECLARATION

I hereby declare that the work herein submitted as a dissertation for the degree of Master of Science in Botany is the result of my own investigation. The work by other authors that formed part of literature support has been duly acknowledged by the reference to the authors.

…………………………...   …………………

Tshediso Andrew Mokhele   Date

ACKNOWLEDGEMENTS

I owe my deepest gratitude to the following people and departments:

- My supervisor, Dr. C.M.S. Mienie and co-supervisor, Prof. S. Barnard, whose encouragement, guidance and support from the initial to the final end of this study enabled me to develop an understanding of the subject.
- Dr. P.J.A. van der Merwe and Mr. T. Matamela for providing sunflower seeds and their guidance throughout the breeding steps.
- The Agricultural Research Council–PDP for financial support and providing research facilities in making this study possible.
- The ARC-GCI biotechnology team (R. Terblanche and J. Gobiyeza) for their technical support in this study.
- My colleague, L.A. Madubanya for his interests in this study and his advice on statistical analysis.
- My †mother and father for their unconditional love and support throughout my life.
- The University of North-West bursary department in conjunction with NRF scare skills funding for additional funding.
- Last and more important, my Lord God who made this study possible.

Let us not become weary in doing good,
for at the proper time we will reap a harvest if we do not give up
Galatians 6:9

ABSTRACT

Sunflower (*Helianthus annuus* L.) high oleic acid content lines differ from conventional sunflower by an increase in oleic acid (C18:1) content of more than 60%. The current sunflower cultivars under production in South Africa are standard sunflower with high levels of linoleic acid (C18:2). The aim of this study was to improve the quality of oil produced by local sunflower germplasm with respect to oleic acid through employing a marker-assisted breeding technique to facilitate and speed up the recovery of the high oleic acid allele into the background of the recurrent parent genome. Eleven sunflower breeding genotypes with high and low oleic acid traits were obtained from the Agricultural Research Council-Grain Crops Institute (ARC-GCI) in Potchefstroom. The breeding genotypes were phenotypically characterised based on their oleic and linoleic acid levels using gas chromatography. Results demonstrated that the average mean of oleic and linoleic acid contents in high oleic acid genotypes were 72% and 17% respectively, while the average mean of oleic acid and linoleic acid contents in wild type lines were 33.5 % and 54 % respectively. These results indicated a perfect negative correlation between the amount of oleic and linoleic acids possessed in high and low oleic acid genotypes (R^2 = -99.16%). Sequence characterised amplified region (SCAR) markers were tested to ascertain if any of the ten available dominant FAD2-1 markers was segregating with the high oleic acid allele. Four dominant SCAR markers (FAD2-1F4/R1; FAD2-1F4/R2; FAD2-1F13/R1; FAD2-1F14/R2) were strongly associated with the high oleic acid trait ($P < 0.001$). With regard to the inheritance of the high oleic acid trait, 143 plants of the F_2 segregating population derived from a cross between the high oleic acid parent (AP901-95-3-4-1) and low oleic acid parent (H55-9-2-1-1) were genotyped with the four SCAR markers to determine the genetic state concerning the high oleic acid gene (*Ol*). Results from a Chi square analysis of the observed frequencies of each dominant FAD2-1 marker locus in 143 F_2 individuals indicated that the deviation from the expected ratio of 3:1 (high to low oleic acid) was not statistically significant ($P < 0.95$) from the observed segregation ratio. These results were consistent with the previous finding that an incomplete dominant gene governs sunflower high oleic acid. A multiplex assay of 78 Simple sequence repeat (SSR) markers was optimised and evaluated on 143 plants of the F_2 population to determine suitable SSR markers that can be used in a marker-assisted background selection. Only 14 markers were suitable for marker-assisted background selection based on their high polymorphic information content, allele frequency and maximum allele numbers. In conclusion, this study demonstrated the potential of using SSR and SCAR marker systems as a breeding tool to characterise and speed up the selection process in marker-assisted backcross breeding.

Keywords: Mapping population, marker-assisted backcross breeding, multiplex assay, SCAR markers, SSR markers, Sunflower high oleic acid.

OPSOMMING

Sonneblom (*Helianthus annuus* L.) hoë oleïensuur lyne verskil van konvensionele sonneblom deur ‘n verhoging van die oleïensuur (C18:1) inhoud na meer as 60%. Die huidige sonneblom kultivars wat in Suid-Afrika verbou word, is standaard sonneblom met hoë vlakke van linoleïensuur (C18:2). Die doel van die studie was om die kwaliteit van die olie van plaaslike sonneblom kiemplasma te verbeter in soverre dit die oleïensuur inhoud aangaan, deur gebruik te maak van merker-gesteunde seleksie om die inbouing van die hoë oleïensuur allele in die genetiese agtergrond van ‘n herhalende ouer te bespoedig. Elf sonneblom teellyne met hoë en lae oleïensuur vlakke is verkry van die Landbounavorsingsraad-Instituut vir Graangewasse (LNR-IGG) te Potchefstroom. Die teellyne is fenotipies gekarakteriseer gebaseer op hulle oleïensuur en linoleïensuur vlakke deur van gaschromatografie gebruik te maak. Die resultate het aangedui dat die gemiddelde oleïen- en linoleïensuur inhoud in hoë genotipies 72% en 17% respektiewelik was, terwyl die inhoud in lae genotipes gemiddeld 33.5% en 54% was. Hierdie resultate het ‘n perfekte negatiewe korrelasie tussen die oleïensuur en linoleïensuur vlakke aangedui (R^2 = -99.16%). SCAR merkers is getoets om vas te stel of enige van die tien beskikbare dominante FAD2-1 merkers saam met die hoë oleïensuur allele sou segregeer. Vier dominante SCAR merkers (FAD2-1F4/R1; FAD2-1F4/R2; FAD2-1F13/R1; FAD2-1F14/R2) het sterk geassosieer met die hoë oleïensuur eienskap ($P< 0.001$). Om die oorerwing van die eienskap te bepaal, is 143 plante van ‘n F_2 segregerende populasie, verkry vanuit ‘n kruising tussen ‘n hoë oleïensuur ouer (AP 901-95-3-4-1) en ‘n lae oleïensuur ouer (H55-9-2-1-1), met die vier SCAR merkers ge genotipeer. ‘n Chi-kwadraat toets van die waargenome frekwensie van die dominante FAD2-1 merker lokus in die 143 F_2 plante het aangedui dat die afwyking van die verwagte frekwensie nie statisties betekenisvol was nie ($P< 0.95$). Hierdie resultate het ooreengestem met ʼn vorige waarneming dat die hoë oleïensuur eienskap deur ‘n onvolledig dominante geen gekodeer word. ‘n Multipleks vingerafdruk tegniek met 78 SSR merkers is ge-optimiseer en ge evalueer in die F_2 generasie plante om te bepaal watter merkers vir merker-gesteunde agtergrond seleksie geskik sal wees. Slegs 14 SSR merkers was polimorfies tussen die teelouers met hoë polimorifiese inligting (PIC), hoë allele frekwensie en maksimum allele getalle. In gevolgtrekking het hierdie studie dus die potensiaal vir die gebruik van SSR en SCAR merkers as ‘n telingshulpmiddel om die seleksie proses te bespoedig tydens merker-gesteunde terugkruising, gekarakteriseer.

Sleutelwoorde: Kartering van die bevolking, merker bygestaande terugkruisteling, Multipleks toets, sonneblom hoë oleïensuur, SSR merkers.

LIST OF ABBREVIATIONS

A/A	homozygous alleles for parent A
A/B	heterozygous alleles for parent A and B
AFLP	amplified fragment length polymorphism
ARC-GCI	Agricultural Research Council-Grain Crops Institute
BC	Backcross
bp	base pair(s)
B/B	homozygous alleles for parent B
cM	centimorgan
CMS	cytoplasmic male sterile
CTAB	hexadecyltrimethylammonium bromide
dATP	2’-deoxyadenosine 5’-triphosphate
dCTP	2’-deoxycytidine 5’-triphosphate
dGTP	2’-deoxyguanosine 5’-triphosphate
DH	Doubled haploid
dNTP	2’-deoxynucleoside 5’-triphosphate
dTTP	2’-deoxythymidine 5’-triphosphate
EDTA	ethylenediaminetetraacetate
F	forward primer
F_1	first breeding population
F_2	second breeding population
BC_1F_1	first backcross population
BC_2F_1	second backcross population
FAD2-1	*fatty acid desaturase* gene
FAME	fatty acid methyl esters
FID	flame ionisation detector

GC	gas chromatography
GGT	graphical genotypes
ha	hectare(s)
HO	high oleic acid
INDEL	insertion-deletion
LO	low oleic acid
M	molecular weight marker
MAB	marker-assisted backcrossing
MAS	marker-assisted selection
N	negative
NIL	near isogenic line
P	positive
Parent A	high oleic/low linoleic acid parent
Parent B	low oleic/high linoleic acid parent
PCR	polymerase chain reaction
PDP	professional development program
PIC	polymorphic information content
PL	plant
QTL	quantitative trait loci
R	reverse primer
R^2	coefficient of determination
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RFU	relative fluorescence units
RIL	recombinant inbred line
RNA	ribonucleic acid
RPG	recurrent parent genome

SCAR	sequence characterised amplified region
SSR	Simple sequence repeats
STS	Sequence tagged site
TAE	Tris/Acetic acid/EDTA
Taq	*Thermus aquaticus*
TBE	Tris-borate/EDTA
TE	Tris/EDTA
TRAP	Target region amplification polymorphism
Tris	Tris(hydroxymethyl) aminomethane
U	Unit
μM	Micromolar
m/v	mass/volume
v/v	volume/volume
W	Watt(s)

TABLE OF CONTENTS

CHAPTER 4 EVALUATION OF AN F_2 POPULATION SEGREGATING FOR THE HIGH OLEIC ACID TRAIT USING A SCAR MARKER ASSAY…….. 28

.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 GENERAL INTRODUCTION AND PROBLEM STATEMENT

Cultivated sunflower (*Helianthus annuus* L.) is a major source of vegetable oil in the world and a widely grown ornamental crop. The crop is one of the 48 species in the *Helianthus* family which are indigenous to and distributed throughout North America. Globally, sunflower ranks fourth under production of vegetable oil, and South Africa is the world's twelfth largest producer of cultivated sunflower (Dredge, 2010, Tang *et al*., 2002). Sunflower seeds provide between 45% (g/w) of oil and about 40% of oil cake, which is processed to cooking oil and for animal feed respectively. The quality of sunflower oil is determined primarily by its constituent fatty acids. The major fatty acids of conventional sunflower oil are saturated and polyunsaturated fatty acids. The saturated fatty acid constitutes about 11% of both stearic and palmitic fatty acid while unsaturated fatty acid constitutes about 20% of oleic and 69% linoleic acids (Kusterer *et al*., 2004, Okuley *et al*., 1994, Pacureanu-Joita *et al*., 2005, Lacombe & Bervillé, 2001).

High oleic acid content sunflower cultivars differ from conventional sunflower cultivars by a significant increase in oleic (18:1) acid content of more than 60% (Lacombe & Bervillé, 2001, Pecureanu-Joita *et al.*, 2005). Currently, sunflower cultivars under production in South Africa are standard sunflower with high linoleic acid (18:2) content. High levels of linoleic acid in standard sunflower cultivars have undesirable traits since linoleic acid is highly unstable and easily oxidised to give off-flavour and rancidity to the oil resulting in short shelf life at storage (Hu *et al*., 2006, Pacureanu-Joita *et al*., 2005). Other possible health risks associated with high linoleic acid include metastasis in cancer patients.

The new market trends necessitate oil composition resistant to high temperature, with less saturated fatty acids and which can resist oxidation longer. High oleic acid sunflower oil increases the thermo stability of the oil, making it more suitable as cooking oil.

Solving the problem can however be achieved through releasing sunflower cultivars with high levels of oleic acid by means of a marker-assisted recurrent breeding programme. In marker-assisted recurrent breeding, the genetic diversity of selected parental lines could be determined and background markers employed to facilitate and speed up the recovery of the recurrent parent genome (Dehmer and Friedt, 1998, Kusterer *et al*., 2004, Okuley *et al*., 1994, Pacureanu-Joita *et al*., 2005, Severine & Bervillé, 2001).

1.2 AIM

The aim of the study was to apply molecular markers linked to the high oleic acid trait in sunflower breeding lines in an effort to facilitate marker-assisted backcross introgression of the high oleic acid trait into the background of Agricultural Research Council (ARC) sunflower germplasm.

1.3 OBJECTIVES

1. To evaluate molecular SCAR markers linked to the sunflower high oleic acid trait
2. To validate the selected SCAR markers through the evaluation of an F_2 population segregating for the high oleic acid trait
3. To apply the publicly available SCAR and SSR assays for the introgression of the high oleic acid trait into the background of the recurrent parent genome

A conventional backcross breeding method was used for high oleic acid gene introgression whereas molecular marker assays were used to facilitate and speed up the selection of positive crosses.

CHAPTER 2

LITERATURE REVIEW

2.1 Sunflower taxonomy

Globally, cultivated sunflower is an important source of edible oil. The genus (*Helianthus*) consists of 48 species of annuals and perennials with different ploïdy levels (Hvarleva *et al*., 2007). The most cultivated sunflower species in agriculture includes Jerusalem artichoke (*H. tuberosus*) grown on a limited basis for food and livestock feed in the United States of America (U.S.A). *Helianthus annuus* is a commercial sunflower, which is used as oilseed, bird feed, ornamental, or confectionery (Köhler and Friedt, 1999). The basic chromosome number for the *Helianthus* genus is 2n=34. Diploid, tetraploid and hexaploid species are known (Berglund, 2007).

2.2 Sunflower production

On a worldwide basis, sunflower is grown on about 20 million hectares of semi-arid regions and presented as one of the most important crops for oil production. The annual average grain yield in the world is around 1.1 tons/hectare, varying from 0.5 to 3.6 tons/hectare. Major producing countries of sunflower include Russia, Argentina, Eastern Europe, USA, France and Spain. These countries produce around 80% of the world's oilseed and non-oilseed sunflower (Berglund, 2007, Krizmanic *et al*., 2006, Putnam *et al*., 1990). In Africa, sunflower producing countries include South Africa, Zimbabwe, Sudan, Tanzania, Kenya, Zambia, Mozambique, Angola, and Malawi.

Sunflower is the third largest grain crop produced after maize and wheat in South Africa and is produced mostly in six out of the nine provinces; the Free State and North West provinces produce approximately 85% of sunflower seeds, with 15% being shared among Gauteng, Mpumalanga, Limpopo and Northern Cape Provinces. According to the report by Dredge (2010) total production of sunflower seeds in the 2008/09 growing season was about 900 300 tons on an area of 595 400 hectares and out of this volume, the Free State produced 470 250 tons and North West 300 000 tons. An area of 500 000 hectares of sunflower seeds is estimated for planting during the 2010/11 seasons.

Sunflower is generally considered a cross-pollinating crop and different types of breeding lines exist such as the A-line or cytoplasmic male sterile (CMS) lines. The CMS breeding lines resulted from an interaction between the organelles of the male genome that conditioned the failure to produce functional pollen. According to Venkana *et al*. (2008), the first CMS source discovered was *H. petiolaris* (PET1), discovered by Leclercq (1969). The genetic material of CMS lines was subsequently identified in 1970 by Kinman. In the cross to produce hybrid seeds, the A-line is used as a female parent and it contains the recessive fertility restorer (*rf1, rf2*) genes (Sleper & Poehlman, 2006). Many sunflower hybrids produced via CMS lines showed about 25% yield advantages over the open pollinated varieties and also showed improved disease resistance (Fick, 1984).

The second breeding line that exists in sunflower is the B-line or maintainer of the A-lines. The B-line possesses the same genotype as the A-line, but normal instead of sterile cytoplasm and recessive, non-restorer (*rf1, rf1*) genes like the A-lines. The R-line is used as a fertility restorer line in the cross with the A-line to produce hybrid seeds. The primary function of the R-line is to pollinate the A-line and restore fertility in the hybrid seeds. The R-line is characterised as having dominant genes (*Rf1 Rf1; Rf2 Rf2*) for fertility restoration (Sleper & Poehlman, 2006).

2.3 The genetics of sunflower

Due to the economic and ecological importance of wild and cultivated sunflower, *H. annuus* has been one of the primary models for genetic and genomic research in the Compositae (Knapp *et al*., 2001, Tang *et al*., 2002). The major objectives in sunflower breeding programmes are the development of new cultivars with high oil yield, yield stability and disease resistance. Extensive work has been done in breeding for high oleic acid content and other traits in sunflower. On the basis of the economic importance of the sunflower crop, the study of sunflower has led to the significant advancement in understanding the genetic determinants of economically important genes (Gentzbittel *et al*., 1998, Langar *et al*., 2003).

The development of sunflower oil with high oleic acid content was first reported by Soldatov (1976) by treating seeds from the M3 generation of VNIIMK8931 with dimethyl sulphate to obtain an induced mutation for high oleic acid content (Miller *et al*., 1987). On the basis of this source many breeders developed their own breeding lines using VNIIMK 8931 as genetic source for this trait (Pacureanu-Joita *et al*., 2005).

According to Pérez-Vich *et al*. (2002), the inheritance of sunflower high oleic acid trait was first reported by Fick (1984), in which a single partially dominant gene (*Ol*) was hypothesised as being directly involved in the enhancement of oleic acid. In addition to this gene, a study conducted by Urie (1985) showed that the gene controlling the high oleic acid trait was segregating with modifiers and that there was an unexplained reversal in the dominance of the gene. Miller *et al*. (1987) reported a second modifying gene designated (*Ml*) that segregated with the high oleic acid gene.

Two other genes which were also directly linked to the high oleic acid trait were later reported by Fernández-Martinez *et al*. (1989). These genes were named Ol_1, Ol_2, and Ol_3 and the phenotypes possessing these genes in dominant alleles were reported to have high oleic acid content, whereas phenotypes possessing two of these genes in dominance and one being homozygous recessive expressed intermediate levels of oleic acid.

Schuppert *et al*. (2006) reported that the Oleoyl-phosphatidyl choline desaturase (FAD2) enzyme was necessary for the synthesis of linoleic acid from oleic acid. According to Lacombe *et al.* (2001) and Garces & Mancha (1991), the accumulation of oleic acid was reported to be due to either high activity of 9-desaturase catalysing the desaturation of stearic acid to oleic acid, or lack of 12-desaturase activity catalysing the desaturation of oleic acid to linoleic acid.

2.4 Application of biotechnology techniques in cultivated sunflower

Genetic improvements of crop plants have traditionally been achieved through conventional plant breeding. Although generally slow and dependent on appropriate environmental conditions to identify and select plants with desirable phenotypes, traditional plant breeding continues to be the main means of crop improvement largely through exploitation of hybrid vigour and grain yield of many cultivated plant species (Jauhar, 2006). However, new technologies have introduced an additional means for improving sunflower yield and quality using molecular genetics.

Molecular markers have been looked upon as tools for a large number of applications ranging from localisation of a gene to improvement of plant varieties by marker-assisted selections. They have also become extremely popular markers for phylogenetic analysis adding new dimensions to the evolutionary theories. The development of these markers over the last two decades has provided easy, fast and automated assistance to scientists and breeders (Kaur *et al.*, 2009).

The aim of molecular genetics in sunflower breeding is to identify, isolate, amplify, and modify genes or other sequences of DNA and to combine and express the novel or modified sequence in new genotypes (Sarrafi & Gentzbittel, 2008).

2.4.1 Simple sequence repeat (SSR) or microsatellites

Microsatellites or simple sequence repeats are tandem repeated motifs of 1-6 base pair (bp) which have a frequent occurrence in all prokaryotic and eukaryotic genomes. SSRs are present in both coding and noncoding regions and are distributed throughout the nuclear genome (Kalia *et al.*, 2011, Zane *et al.*, 2002). In eukaryote species, $(AT)_n$ is the most common sequence followed by $(AG)_n$ or $(TC)_n$ (Powell *et al.*, 1996). Change in the number of repeats gives the length polymorphism which is revealed by designing primers for the sequences flanking the microsatellite motifs. The resultant sequence tagged microsatellite usually identifies a single locus which, because of the high mutation rate of SSRs, is often multi-allelic. Alleles which differ in single bp can be amplified through polymerase chain reaction (PCR) and resolved on agarose or denaturing polyacrylamide gels (Jones *et al.*, 1997).

SSRs are highly valued because they provide a co-dominant, high discriminatory power, a high information content arising from their multi-allelic nature, and highly reproducible markers that can be easily shared between different laboratories. Besides providing excellent molecular markers, SSR markers are useful for population genetics, variety identification and protection, monitoring of seed purity and hybrid quality, gene tagging, germplasm evaluation and phylogenetic studies (Jones *et al.*, 1997, Masi *et al.*, 2003, Morgante & Oliveiri, 1993).

Technical limitations that currently restrict the use of SSR markers include the cost and time consuming steps in identifying polymorphic loci and the possible scoring problems incurred when alleles differ only slightly in length, but once primers have been developed, the system becomes relatively inexpensive (Mitchelle *et al.*, 1997).

The widespread application of SSR markers in agricultural, biological and medical research has evolved into a high throughput DNA system such as an automated multiplex technique that involves the analysis of multiple loci in the same lane using different fluorescent labelled primer sets, to enable researchers to carry out large scale genetic mapping and population studies (Henegariu *et al.,* 1997, Tang *et al.,* 2002).

In a multiplex PCR assay markers labelled with the same fluorophore dye must be of different size, while markers that overlap in size must be distinguished using different coloured dyes to label the SSR marker (Blair *et al.,* 2002, Ziegel *et al.,* 1992). The different dye set colours such as 6-FAM (Blue), HEX (Green), and NED (Yellow) are commonly used to label one primer pair and are easily detected through an automated fluorescent detection system. However, amplification of multiple loci per reaction requires optimisation of PCR conditions in order to satisfactorily amplify all target amplicons. A prior knowledge about allele size range of each amplified locus is also required in order to rationally group primers into sets or panels (Masi *et al.,* 2003).

Multiplexing of SSR markers has been successfully applied in many areas of DNA testing. In agriculture, two to eleven loci PCR multiplexes have been developed for sets of four to 74 SSR marker loci in rapeseed (Mitchelle *et al.,* 1997, Tang *et al.,* 2002), peanuts (Hopkins *et al.,* 1999), maize (Gethi *et al.,* 2002) and soybean (Narvel *et al.,* 2000a, b). The first genetic map constructed for sunflower using a multiplex SSR assay was established by Tang *et al.* (2002) using recombinant inbred lines (RILs) from a cross between confectionery and oilseed fertility restorer lines (RHA280 X RHA 801) (Fick *et al.,* 1974, Yu *et al.,* 2003). Darvishzadeh *et al.* (2010) applied SSR multiplex in a diversity study to characterize the relationship among sunflower inbred lines. Blair *et al.* (2002) also applied the technique to determine the range of alleles found in a wide array of cultivated rice varieties and for fingerprinting and clustering of rice genotypes.

In biological and medical research, Fritz *et al.* (1996) identified mutations that affect cloned genes in zebrafish by analysing the genomic DNA of parthenogenetic haploid offspring of females carrying gamma-ray induced mutations. In a study conducted by Mehdipour *et al.* (2006) founder mutations of *BRCA1* and *BRCA2* genes associated with breast cancer were identified in Iranian breast cancer patients using a multiplex PCR assay. In a study conducted by Ashishi *et al.* (2003) a multiplex PCR assay was optimised in an effort to develop a diagnostic tool which can be used to detect human cytomegalovirus, human herpes virus type 6,7 from patient specimens.

2.4.2 Amplified fragment length polymorphism (AFLP)

Amplified fragment length polymorphism (AFLP) is a DNA-based fingerprinting technique developed by Zabeau & Vos (1993), which is based on PCR amplification of selected restriction fragments of a total digested genomic DNA. AFLP can be used as a tool for determining the identity of specific DNA samples, a source for genetic linkage maps or to identify molecular markers linked to phenotypic traits or genetic loci.

AFLP analysis involves the digestion of genomic DNA with rare and frequent cutter restriction endonucleases. Double stranded adaptors are ligated to the ends of the DNA fragments to generate template DNA for amplification. Thus, the sequence of the adaptors and the adjacent restriction sites serve as the primer binding sites for subsequent amplification of the restriction fragments by PCR. Selective nucleotides extending into the restriction fragments are added to the 3’ends of the PCR primers such that only a subset of the restriction fragments are recognised. Only restriction fragments in which the nucleotides flanking the restriction site match the selective nucleotides will be amplified. The subset of amplified fragments is then analysed by denaturing polyacrylamide gel electrophoresis or on an automated sequencer to generate the fingerprint (Hongtrakul *et al.,* 1997, Mullis & Faloona, 1987, Vos *et al.,* 1995).

The AFLP technique is highly accurate, reliable, does not require any DNA sequence information from the organism under study, and is reproducible from one laboratory to another. AFLP markers are distributed across the genome and have a high multiplex ratio (each band is assumed to come from a different area of the genome). As a result the technique provides a high throughput method, making it useful for large scale population studies (Kaur *et al.,* 2009).

Despite being a high throughput technique, there are features that limit the extensive use of AFLPs such being a dominant marker system, which does not differentiate homozygotes from heterozygotes. This reduces the accuracy of AFLP markers in population genetics analysis, genetic mapping, and marker-assisted selection. Other limitations include the use of polyacrylamide gels in combination with silver nitrate staining, radioactivity, or fluorescent methods of detection, which will be more expensive and laborious than agarose gels (Kaur *et al.,* 2009).

Since the first report on AFLP fingerprinting in cultivated sunflower was published (Hongtrakul *et al.,* 1997), several fingerprinting and mapping studies on sunflower and other crop species using AFLP are known. In a study conducted by Venkatesha *et al.* (2010), 22 dolichos bean accessions were characterised into different genetic groups. Using AFLP analysis, Špunarová *et al.* (2005) determined the diversity among spring barley breeding lines with a sufficient level of resistance towards *Fusarium* head blight.

2.4.3 Sequence characterised amplified region (SCAR) markers

SCAR markers are developed from other types of markers that are less reliable in order to be easily applicable in a breeding programme with high number of plants to be screened. AFLP and randomly amplified polymorphic DNA (RAPD) markers are from the most popular markers to be converted to SCARs. This process converts it from labour intensive, low throughput, non-repeatable to highly reliable, high throughput markers. SCAR markers linked to the high oleic acid trait in sunflower were developed by Schuppert *et al.* (2006) using a single forward primer (F4) complementary to intergenic DNA sequence in the mutant and reverse primers complementary to coding (R1 & R2), 3’ end upstream (R3), or noncoding (R4 & R2) DNA sequences. These markers were screened for length polymorphism among four wild type sunflower genotypes and four mutant lines using agarose gels and standard PCR. The DNA sequences downstream of *FAD2-1* were screened for DNA polymorphism by allele re-sequencing. The primer pair F3/R13 was used to amplify alleles from six wild types and the allele sequences were aligned using Clustal-W and screened for single nucleotide polymorphism. Nucleotide (ø) and haplotype diversity statistics were estimated using DNAsp (Rozas & Rozas, 1999, Rozas *et al.,* 2003). SCAR markers have already been developed and applied successfully in buck-wheat (All *et al*., 1999), eggplant (Bi-Hao *et al.,* 2009), brassica (Cheung *et al.,* 1997), soybean (Mienie *et al.,* 2002) and olive (Hernandez *et al.,* 2001).

2.4.4 Genetic linkage mapping

The development of molecular marker techniques have provided an additional tool to determine linkage maps which are used for the analysis of single or quantitative trait loci (QTL) in order to measure the relative positions of genes on chromosomes and distances between them (Sarrafi & Gentzbittel, 2008). During the past several years, the adoption of high throughput genotyping technologies has been paralleled by a substantial increase in the density and diversity of genetic markers. New genetic mapping algorithms are needed in order to efficiently process these large datasets and accurately construct high-density genetic maps. Genetic linkage maps are cornerstones of a wide spectrum of biotechnology applications, including marker-assisted breeding, association genetics, and map-assisted gene cloning (Wu *et al.,* 2008).

The construction of detailed genetic maps with high levels of genome coverage provides advantages in facilitating the introgression of desirable genes or QTL through marker-assisted selection. A framework for anchoring with the physical map based on chromosome translocation and DNA sequencing or other direct measures constitute the first step towards positional or map-based cloning of genes responsible for economically important traits (Mohan *et al.,* 1997, Semagn *et al.,* 2006, Vuysteke *et al.,* 1999).

An appropriate mapping population, suitable marker system and the software for analysis of data are the key requirements for a molecular mapping and breeding programme. The first step in producing a mapping population is the selection of two distinct parental lines that show significant polymorphism for the trait of interest at both the DNA sequence and phenotype level. Types of progenies often used for genetic mapping in self-pollinating species include: an F_2 population, backcross (BC) population, RILs, doubled haploid (DH) population, and near-isogenic lines (NILs) (Semagn *et al.,* 2006, Young *et al.,* 1988).

A F_2 mapping population is developed from selfing of F_1 individuals or inter-mating for cross-pollination species. The F_1 individuals are developed by crossing two parents that show significant polymorphism for the trait of interest at both DNA sequence and phenotypic level. The segregation ratio expected for dominant and co-dominant markers in a F_2 population is 3:1 and 1:2:1 respectively. However, F_2 populations are of limited use for fine mapping and QTL cannot be precisely mapped as each individual is genetically different and cannot be evaluated in replicated trials over locations and years, thus the effect of genotype by environment (G x E) interaction on the expression of QTL cannot be precisely estimated (Singh & Prasanna, www.scribd.com).

A backcross population is developed by crossing the F_1 population back to one of the parents (recipient or recurrent parent). With respect to molecular markers, a backcross with dominant parent (B_1) would segregate in a ratio of 1:1 for dominant and co-dominant markers, respectively and a backcross with recessive parent (B_2) or testcross would segregate in a ratio of 1:1 irrespective of the nature of the marker. However, the specific advantage of a backcross population is that the population can be further utilised for marker-assisted backcross breeding (Singh & Prasanna, www.scribd.com).

A RIL population is developed by single-seed selections from individual plants through selfing of each generation of a F_2 population. Single-seed descent is repeated for several generations until complete homozygosity is achieved and all seeds from an individual plant are bulked. This population of seed can then be grown to obtain a large quantity of seed of each individual line and each of these lines is fixed for many recombination events. RIL populations can be used to derive a map because it essentially is an eternal F_2 population with unlimited mapping possibilities. These lines are also powerful for analysing QTL because replicated trials can be analysed using identical genetic material. However, RIL populations require many seasons and generations to develop and developing RIL is relatively difficult in crops with high inbreeding depression (Burr & Burr, 1991, Keurentjies *et al.*, 2006).

The first consensus genetic linkage map constructed for cultivated sunflower was reported by Gentzbittel *et al.* (1994) in which a total of 180 restriction fragment length polymorphism (RFLP) probe-enzyme combinations were mapped on at least one of five segregating progenies (three F_2 and two BC_1 populations) revealing 237 loci that did not show any distortion of segregation. In a study conducted by Hu *et al.* (2004) a linkage map covering a total length of 1 140 centimorgan (cM) was constructed using a 129 individuals from a RIL population resulting from a cross between two sunflower inbred lines using 160 target region amplification polymorphism (TRAP) markers. Berry *et al.* (1995) reported a linkage map that consisted of 17 linkage groups of varying sizes and completeness from 289 individuals of a F_2 population derived from a cross between wild and cultivated sunflower inbred lines using about 213 RFLP probes (Tang *et al.,* 2002).

NIL populations are developed through backcrossing the F_1 plant to the recurrent parent until the progenies are almost identical to the recurrent parent. Plant breeders have often used NIL populations to transfer major genes between varieties by backcross breeding (Kaeppler *et al.,* 1993, Semagn *et al.,* 2006, Tanksley *et al.,* 1995). Near isogenic line populations are important genetic stocks for investigating the function and regulation of single and quantitative genes. Tsujimoto (2001) has developed NIL's through a recurrent backcrossing method to elucidate the effect of marker genes on plant morphology in the genetic background of the wheat cultivar Chinese spring.

DH populations are produced by a process of doubling gametes of a F_1 or F_2 population. A tissue culture technique is then used for regeneration after induction of chromosome doubling from a pollen grain or haploid embryos resulting from species crosses (Lefebvre *et al.,* 1995). Artificial production of DH is important in plant breeding since conventional inbreeding procedures take approximately six generations to achieve almost complete homozygosity, whereas DH's achieve it in one generation (Semagn *et al.,* 2006). According to Musial and Przywara (1998) the first report of the haploid plants was published by Blakeslee *et al.* (1922) on jimson weed (*Datura stramonium).*

2.4.5 Marker-assisted backcross (MAB) breeding

MAB is a molecular breeding tool that has been used by many plant breeders for the introgression of a target gene from a donor line into the genomic background of recipient lines. MAB provides great opportunities for precise transfer of a desirable donor segment by minimising the linkage drag into a recurrent parent (Semagn *et al.,* 2006, Singh *et al.,* 2011).

In MAB, parental lines differing in traits of interest are identified by molecular markers that co-segregate or are closely linked with the desired trait. A population of plants segregating for the trait of interest (mapping population) is developed and screened for the trait of interest (marker validation) (Hospital, 2005, Semagn *et al.,* 2006). The efficiency of MAB can be improved by employing two types of selection strategies i.e. foreground and background selection. In marker-assisted foreground selection, individual genotypes carrying the target allele are selected. This type of selection would be effective for the transfer of recessive genes since their classical transfer require additional recurrent selfing generations (Welz & Geiger, 2000). Marker-assisted background selection is based on the selection of individuals homozygous or heterozygous for the recurrent parent alleles at markers flanking the target allele or selecting individuals homozygous for recurrent parent alleles at all remaining markers on the same chromosome as the target allele, as well as the rest of the genome. This type of selection strategy is used extensively in the commercial hybrid maize breeding programmes for the transfer of herbicide tolerance or insect resistance genes (Hospital, 2005, Ragot *et al.*, 1995, Welz & Geiger, 2000).

MAB is particularly useful for obtaining improved germplasm lines and molecular markers allow more efficient selection and offer a mechanism to eliminate undesirable traits associated with hybridizing diverse genotypes. In backcross breeding, molecular markers are useful as they aid selection for traits that are not easily assayed in individual plants. Introduction of unwanted genes can be minimised and the time needed to obtain a plant with a high percentage of the original desirable genetic background can be substantially reduced (Frisch *et al.,* 1999, Hospital, 2005).

However, the success of MAB relies mostly on the distance between the marker and the target gene, the number of the target genes to be transferred, the genetic base of the trait, the number of individuals that can be analysed, the type of molecular markers to be used as well as the available technical facilities (Francia *et al.,* 2005, Semagn *et al.,* 2006, Weeden *et al.,* 1992).

In the current study, these techniques will be tested and applied in a practical breeding programme for the introgression of a high oleic acid trait into local sunflower germplasm containing low oleic acid. SCAR as well as SSR markers will be evaluated for MAB.

CHAPTER 3

EVALUATION OF SCAR MARKERS LINKED TO THE HIGH OLEIC ACID TRAIT IN SUNFLOWER

3.1 INTRODUCTION

Conventional breeding is time consuming and dependent on environmental conditions. Hence, breeders are extremely interested in new technologies that could make this procedure more efficient. Molecular markers offer such a possibility by adopting a wide range of novel approaches to improving the selection strategies in breeding of sunflower and other crop species (Korzun, 2002). Identification of molecular markers that should co-segregate or be closely linked with the desired trait is however a critical step for the success of MAB (Semagn *et al.,* 2006).

Different types of molecular markers such as RAPD (Dehmer & Friedt, 1998), AFLP (Hongtrakul *et al.,* 1997) and microsatellite (SSR) (Schuppert *et al.,* 2006, Tang *et al.,* 2002) have been identified and applied in breeding of sunflower and other crop species.

The most critical step involved in MAB is the selection of a suitable marker system and optimisation of the system on the crop involved. An effective marker system should be able to reveal maximum polymorphism between two inbred lines deferring in the trait of interest, have high throughput capacity, and preferably be a sequence based DNA marker (McGregor *et al.,* 2000). Schuppert *et al.* (2006) and Yu *et al.* (2003) have demonstrated that SSR based markers could be successfully employed to identify markers linked to the high oleic acid trait in sunflower.

Tanhuanpää *et al.* (1998) identified a RAPD marker linked to the QTL affecting oleic acid content in spring turnip rape (*Brassica rapa* subspecies *oleifera*). The marker was later converted into a SCAR. Schierholt *et al.* (2001) identified three AFLP markers linked to the high oleic acid mutation in winter oilseed rape (*B. napus* L.). Hongtrakul *et al.* (1997) have identified AFLP markers that can be used in DNA fingerprinting of sunflower germplasm to reveal genetic diversity among oilseed inbred lines. In the study conducted by Schuppert *et al.* (2006) SCAR markers diagnostic to the presence or absence of an *Ol* mutation (tandem *FAD2-1* repeats) were fully developed in sunflower high oleic acid breeding germplasm.

This chapter aims at testing the applicability of published SCAR markers to identify a suitable marker (linked marker) system that can be applied to facilitate and speed up the selection of high oleic acid content lines in local sunflower germplasm, as well as to characterise these lines based on their oleic and linoleic acid levels.

3.2 MATERIALS AND METHODS

3.2.1 Plant materials and DNA isolation

Sunflower breeding materials with high and low oleic acid content were received from Dr. P.J.A. van der Merwe at the Agricultural Research Council-Grain Crops Institute, Potchefstroom (ARC-GCI). Sunflower high oleic acid lines were used as male lines (pollen donors) while sunflower low oleic acid inbred lines were used as female lines (pollen receivers) (Table 3.1). A total of 50 seeds of both high and low oleic acid plants were planted in the greenhouse in 10 litre plastic pots in sterile soil. The greenhouse was maintained at 28^{o}C day/18^{o}C night temperatures. At seedling stage, young leaves were sampled for DNA extraction from 11 high and low oleic acid inbred lines.

Total genomic DNA was isolated using a modified Hexadecyl-trimethyl ammonium bromide (CTAB) DNA extraction method (Saghai Maroof *et al.*, 1984). A volume of 750 µl of 2X CTAB buffer [100 mM Tris.HCl pH 8.0, 1 400 mM NaCl, 20 mM Ethylene-diaminetetraacetic acid (EDTA), 2% (m/v) CTAB, 0.2% (v/v) B-mercaptoethanol] was transferred in each tube containing lyophilized leaf tissue and incubated in a 65^{o}C water bath for 60 minutes. After incubation, the mixture was allowed to cool to room temperature (25^{o}C) for 5 minutes. The suspension was extracted with 500 µl chloroform: iso-amylalcohol (24:1) (v/v) and centrifuged approximately at 7 826 gravitational force (g) for 10 minutes to allow phase separation. The aqueous top phase was transferred into a clean 1.5 ml tube, and 1 µl of 10 mg/ml *RNase*A enzyme was added into each tube followed by incubation at 37^{o}C for 30 minutes. After incubation, 500 µl of ice-cold isopropanol was transferred into each DNA tube, followed by gentle inversion for 5 minutes. DNA was stored at -20^{o}C for 20 minutes and centrifuged approximately at 11 269 g for 5 minutes.

Isopropanol was removed by draining the tubes upside down followed by the addition of 500 µl of ice cold 70% (v/v) ethanol. Tubes were centrifuged approximately at 11 269 g for 5 minutes. Ethanol was removed and tubes were left to stand open to air dry. DNA was resuspended in 200 µl 0.1 X TE buffer (pH 8.0) (1 mM Tris.HCl, pH 8.0, 0.1 mM EDTA). DNA stock was stored at -80^{o}C in an ultra-freezer.

Table 3.1 Sunflower high and low oleic acid genotypes used in the evaluation of molecular SCAR markers linked to the high oleic acid trait

Genotypes	Type
High oleic acid lines	
1. AP901-95-3-4-1(HO)	Maintainer (B-line)
2. AP901-95-3 (HO)	Maintainer (B-line)
3. AP901-56-1-1 (HO)	Maintainer (B-line)
4. AP901-94-3-2-2 (HO)	Maintainer (B-line)
Wild type lines (Low oleic acid)	
1. H55-9-2-1-1 (LO)	Cytoplasmic male sterile (A-line)
2. RK74-26-1-2-1 (LO)	Maintainer (B-line)
3. ND761-3-1-1 (LO)	Maintainer (B-line)
4. H55-9-2-1-1 (LO)	Maintainer (B-line)
5. AP991-76-1-4-1 (LO)	Maintainer (B-line)
6. AP991-76-1-4-1 (LO)	Cytoplasmic male sterile (A-line)
7. RK74-26-1-2-1 (LO)	Cytoplasmic male sterile (A-line)

3.2.2 Quantification and the quality of the extracted DNA samples

The absorbance (A_{260}) of extracted DNA was measured using a spectrophotometer (Nanodrop, Inqaba Biotech, South Africa) by loading 2 µl of each DNA sample on the spectrophotometer pedestal. Measurements were done at 260 and 280 nanometre (nm) wavelengths after which samples were diluted to a final concentration of 30 ng/µl with distilled water. A 0.8% (m/v) agarose gel was prepared by weighing out 1.2 g of Seakem agarose powder in a 250 ml Erlenmeyer flask containing 7.5 ml of 20X TAE buffer [40 mM Tris, Acetic acid, 10 mM EDTA pH 7.4]. The mixture was adjusted to 150 ml with distilled water and heated in a microwave oven for 5 minutes. The liquid gel was allowed to cool down to 50^oC before casting in a 150 ml casting tray containing a 22 tooth comb. The gel was allowed to set at room temperature for 30 minutes. A volume of 3 µl of each DNA sample was mixed with 3 µl of 15% (m/v) ficoll loading buffer (0.012% bromophenol blue, 15% ficoll, 30X gel-red) and loaded into each well of the gel. Electrophoresis was carried out at 80 volts (V) in a buffer tank containing 2 litres of 1X TAE running buffer for 1 hour.

3.2.3 Optimisation of SCAR markers

During the optimization process, SCAR markers were screened against eight sunflower genotypes. Primer sequences used in this study were obtained from Schuppert *et al*. (2006). The optimisation reactions were carried out in a 20 µl reaction volume containing: 30 ng of genomic DNA, 1X Go*Taq* flexi buffer (Promega USA), 2 mM $MgCl_2$, 250 µM of each dNTP (Promega, USA), 30 ng of each forward and reverse SCAR primers (Invitrogen, USA) (Table 3.2), 0.5 U Go*Taq* flexi DNA polymerase (Promega, USA). An alternative enzyme, SuperTherm *Taq* (Southern Cross Biotechnologies, Cape Town, South Africa) was also tested together with the buffer supplied by the manufacturer.

Reactions were performed using a thermo cycler multi-block system (Thermo Fisher Scientific, USA) and the following PCR gradient conditions were performed: Denaturing for 1 cycle at 94^oC for 30 seconds, followed by 35 cycles of 94^oC for 30 seconds and gradient annealing temperatures of 55.3^oC, 55.6^oC, 56.0^oC, 56.9^oC, 63.4^oC, 64.3^oC, 64.9^oC, 65.2^oC, for 30 seconds respectively, the primer extension temperature was at 72^oC for 30 seconds and final hold temperature was at 4^oC. A 2% (m/v) agarose gel was prepared by a method similar to the one described in section 3.2.2. Electrophoresis was carried out at 80 volts for 1.5 hours and the gel image was taken using a Bio-rad gel documentation system (Figure 3.1).

3.2.4 Sequence characterised amplified region (SCAR) analysis

SCAR analysis was performed under optimised PCR conditions as described in section 3.2.3 using Go*Taq* polymerase enzyme in 20 µl PCR reaction volumes. A 2% (m/v) agarose gel was prepared by a method similar to the one described in section 3.2.2. Electrophoresis was carried out using a method similar to the one described in section 3.2.2 (Figure 3.2).

Table 3.2 Sequences of FAD2-1 SCAR markers used to evaluate candidate markers linked to the high oleic acid trait in sunflower breeding genotypes (Schuppert *et al.*, 2006).

Forward primers	**Sequence (5’-3’)**	**Reverse primers**	**Sequence (5’-3’)**
FAD2-1 F4	GTAACGTCTGCGCGCTTGCAGACATCA	FAD2-1R1	GGTTTTGCATGAGGGACTCGA
FAD2-1 F13	TCAACAGCCTCTTCCTCCTCAG	FAD2-1R1	GGTTTTGCATGAGGGACTCGA
FAD2-1 F4	GTAACGTCTGCGCGCTTGCAGACATCA	FAD2-1R2	CCGATGTCGGACATGACTATC
FAD2-1 ZVG64	CTCATAACACAGATGGTGAAC	FAD2-1ZVG64	CATCATCGGTGCAACCAACA
FAD2-1 F5	GGGCAAAAACGCATTATGTC	FAD2-1R2	CCGATGTCGGACATGACTATC
FAD2-1 F14	AGGTTTTCGCAGGTCTAGGTTA	FAD2-1R13	AAGGAACGGAATGGACAGTGT
FAD2-1 F5	GGGCAAAAACGCATTATGTC	FAD2-1R1	GGTTTTGCATGAGGGACTCGA
FAD2-1 F13	TCAACAGCCTCTTCCTCCTCAG	FAD2-1R5	GTAGTTTTGGAAAGCTAGAGA
FAD2-1 F13	TCAACAGCCTCTTCCTCCTCAG	FAD2-1R13	AAGGAACGGAATGGACAGTGT
FAD2-1 F14	AGGTTTTCGCAGGTCTAGGTTA	FAD2-1R2	CCGATGTCGGACATGACTATC

3.2.5 Fatty acid analysis using gas chromatography

The oil of each of the 11 sunflower genotypes was extracted from physiologically matured seeds of greenhouse grown high oleic and low oleic acid inbred lines. Oil extraction and fatty acid analysis was done at ARC-Animal Production, Irene, Pretoria. Gas chromatography (GC) conditions were chosen according to a modified method similar to the one described by Schuppert *et al.* (2006), Tong *et al.* (2007) and Wan *et al.* (2007). A mass of 100 grams of seeds per sample was crushed with a pestle in a mortar containing 10 ml of hexane and incubated at room temperature for 10 minutes. A volume of 500 µl of this solution was evaporated to dryness at 50°C under a nitrogen stream. The lipids were solubilised in 100 µl of ethyl ether and converted to fatty acid methyl esters (FAME) by adding 100 µl of 0.1 M methanolic KOH, followed by incubating in a 50°C water bath for 5 minutes. The methylation reaction was stopped by the addition of 100 µl of 0.15 M of HCl. Oleic acid, linoleic acid and palmitic acid methyl esters were used as internal standards with the purity of 99% (as provided by ARC-Animal Production).

GC analysis was performed on an Agilent 6890 gas chromatograph equipped with a flame ionisation detector (FID). Fatty acid compounds were separated on a polar DB-23 HT capillary column (15 m x 0.25 mm i.d). The initial column temperature was set at 120°C and then programmed to 190°C at 2°C/min and maintained at 190°C for 8 minutes. The injector and detector temperatures were set at 250°C and the injector split ratio was set to 1:50. Helium was used as a carrier gas at a flow rate of 1 ml/min.

3.2.6 Statistical analysis of data

Data obtained from SCAR analysis were analysed using a general linear model of the STATGRAPHICS plus version 5.0 software to identify molecular markers linked to the high oleic acid trait in sunflower breeding lines. Molecular marker data was used as the independent variable and percentage fatty acid data as the dependent variable. Association between the DNA marker and high oleic acid trait was considered to be significant if the probability was less than 0.01. The coefficient of determination (R^2) was used as a measure of the magnitude of association.

Fatty acid concentrations of individual sunflower high and low oleic acid seeds were estimated from chromatograms using ChemStation software from Agilent. Calibration curves were calculated by plotting the peak area ratios of the different sunflower fatty acids to the internal standards against the concentrations of each fatty acid.

3.3 RESULTS

3.3.1 Plant material and DNA isolation

Eleven sunflower genomic DNA samples were extracted using a modified CTAB extraction protocol and high quality DNA was successfully obtained along with small traces of RNA. In a MAB process, high quality DNA always aid in the amplification of the targeted alleles and reduces non-specific binding of primers during PCR amplification. PCR conditions are reliant on the purity and the amount of DNA added together with other PCR components. Therefore, DNA concentrations were determined through measurement with a spectrophotometer. Results of the spectrophotometer indicated that high concentrations of DNA from each individual sunflower sample were obtained. For a successful PCR amplification, the optimal DNA concentration ranges should be between 30-100 ng.

3.3.2 Optimisation of SCAR markers

A SCAR marker (FAD2-1F4/R2) was optimised on 8 sunflower breeding genotypes using a temperature gradient ranging from 55.3^{o}C to 65.2^{o}C and two polymerase enzymes (Super*Therm* and Go*Taq* polymerase enzymes) respectively. This SCAR primer amplified a fragment size of 1 488 bp in both high and low oleic acid genotypes in a reaction containing Super*Therm* polymerase enzyme (Figure 3.1 a). In Figure 3.1 (b), the same fragment size of 1 488 bp was amplified only in high oleic acid genotypes in a reaction containing Go*Taq* polymerase enzyme. Both enzymes amplified fragments at all temperatures tested.

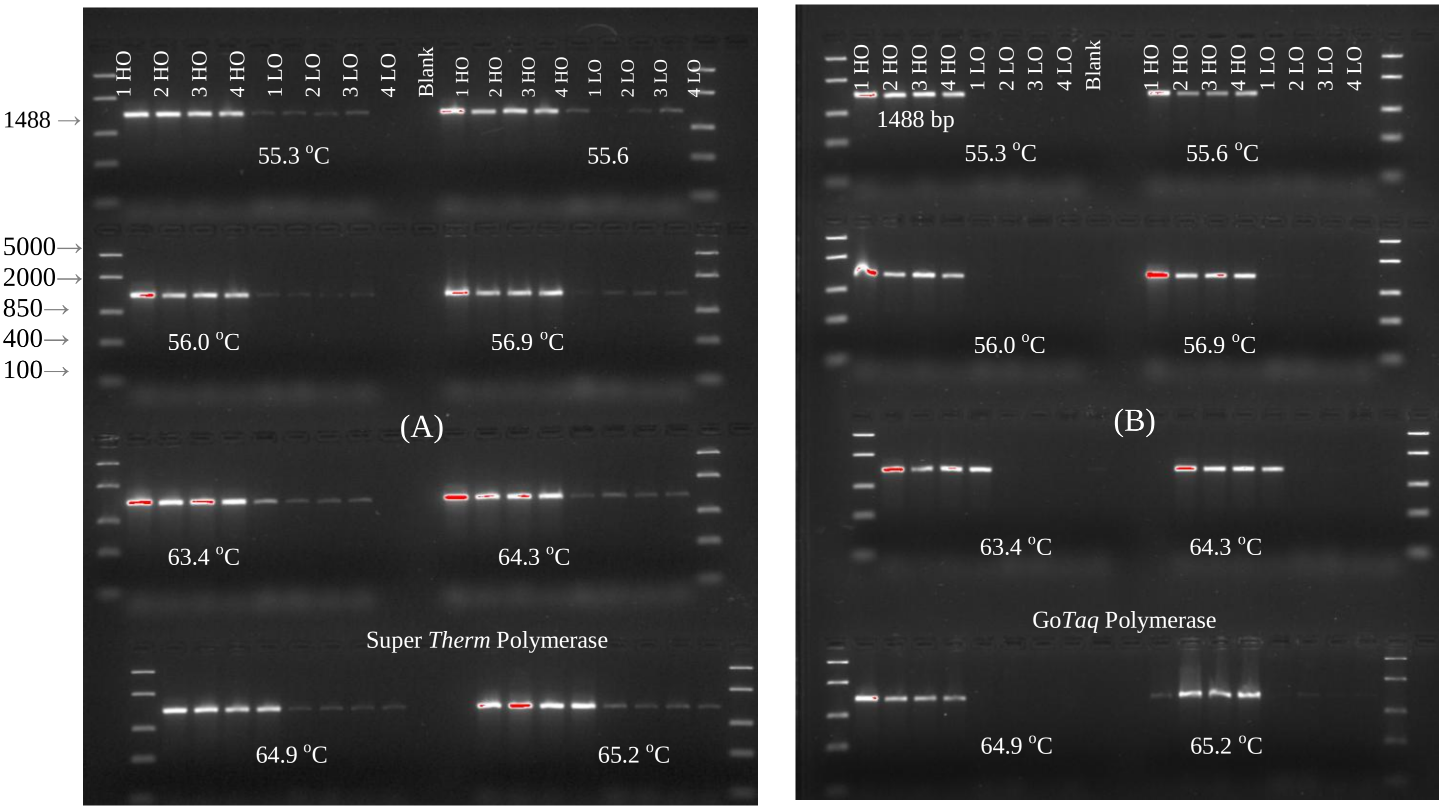


Figure 3.1 Optimisation of FAD2-1 SCAR markers under different PCR conditions (a) eight sunflower breeding genotypes were subjected under different gradient temperatures using Super*Therm* polymerase enzyme (b) the same breeding genotypes were optimized under the same temperatures but different enzyme (Go*Taq* polymerase).

3.3.3 SCAR assay

A total of ten insertion-deletion SCAR markers (Table 3.2) were screened against eleven sunflower breeding genotypes to reveal polymorphisms between high and low oleic acid genotypes. Of the ten SCAR markers screened, four markers (FAD2-1 F4/R1, FAD2-1 F4/R2, FAD2-1 F13/R1 and FAD2-1 F14/R2) revealed a dominant polymorphism between high and low oleic acid genotypes and were linked in coupling phase to the presence of the *Ol* allele determining high oleic acid content (Figure 3.2). These four polymorphic fragments were approximately 787, 1 488, 1 758 and 2 000 bp respectively. However, no DNA amplification was observed with markers FAD2-1 ZVG 64 and FAD2-1 F5R1 for both high and low oleic acid genotypes. Primers FAD2-1 F5/R2, FAD2-1 F13/R13, FAD2-1 F14/R13 and FAD2-1 F13/R5 amplified fragment sizes of 1572, 1572, 787 and 850 bp respectively from both high and low oleic acid genotypes (Figure 3.2). These markers were however, not polymorphic and could not differentiate between high and low oleic acid genotypes.

According to regression analyses, four of the polymorphic fragments revealed by SCAR markers FAD2-1 F4/R1, FAD2-1 F4/R2, FAD2-1 F13/R1 and FAD2-1 F14/R2 showed a strong association with the high oleic acid trait with $P < 0.001$. A statistical P-value of 0.0403 was observed from markers FAD2-1F5/R2; FAD2-1F13/R13; FAD2-1 F13/R5 and FAD2-1 F14/R13 and linked at 39% (R^2) with a P-value only significant to the 0.05 level (Table 3.3).

Table 3.3 Statistical analysis of polymorphic SCAR markers to determine the magnitude of association between the high oleic acid gene and each polymorphic marker

INDEL markers	**R^2 Linkage**	**P-value**
FAD2-1F4/R1	100	0.000**
FAD2-1F4/R2	100	0.000**
FAD2-1F13/R1	100	0.000**
FAD2-1F14/R2	100	0.000**
FAD2-1ZVG64	-	-
FAD2-1F5/R2	39	0.0403*
FAD2-1F13/R13	39	0.0403*
FAD2-1F13/R5	39	0.0403*
FAD2-1F14/R13	39	0.0403*
FAD2-1F5R1	-	-

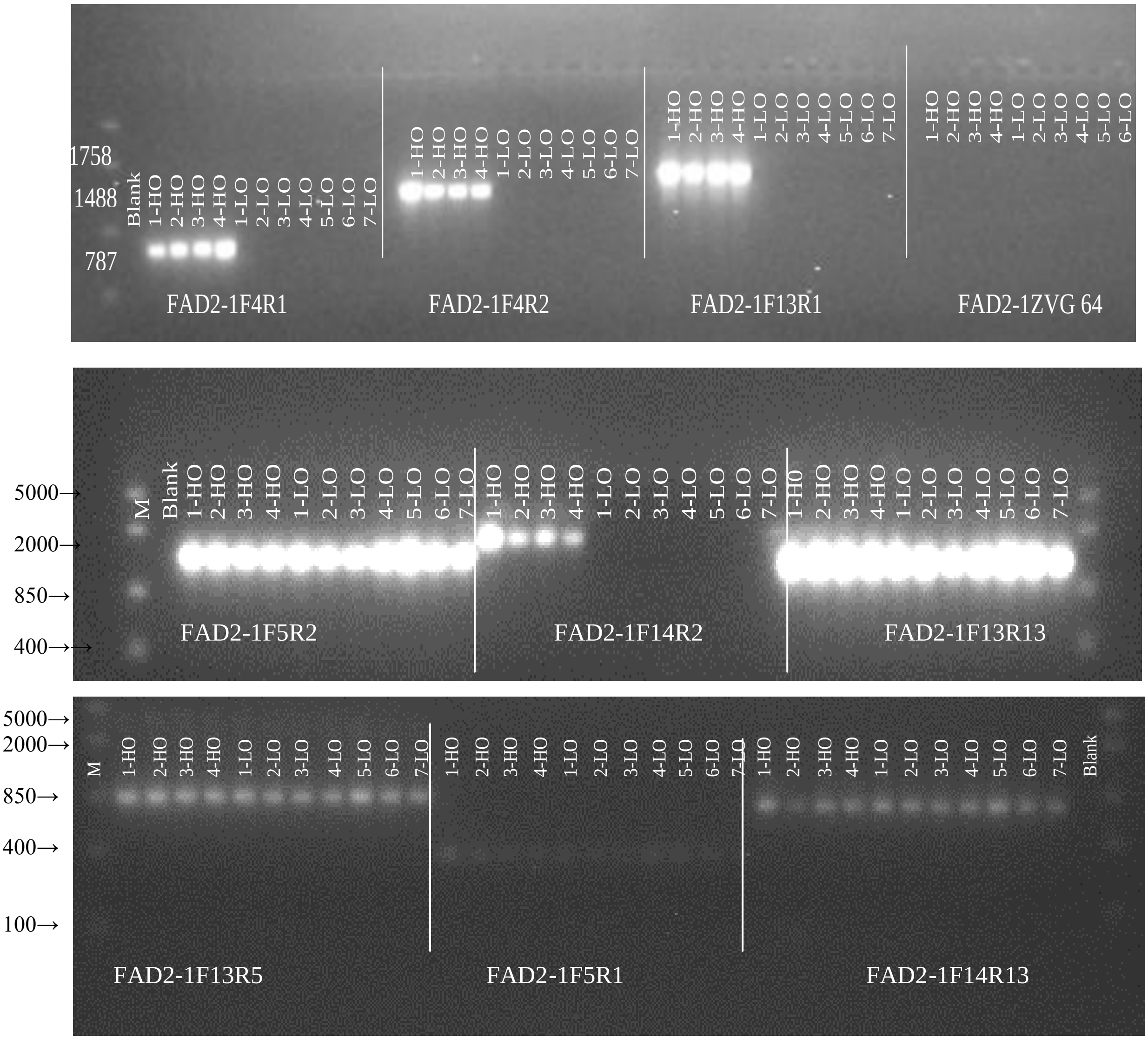


Figure 3.2 Screening of different SCAR primers on a 2% (m/v) agarose gel in an effort to identify putative markers associated to the high oleic acid trait on 11 sunflower high and low oleic acid breeding genotypes.

3.3.4 Fatty acid analysis

The fatty acid composition of high oleic acid and high linoleic acid genotypes were profiled using gas chromatography to determine phenotypically if there was a linkage between the high oleic acid trait and the polymorphic markers among breeding genotypes. The average mean oleic acid content of high oleic acid lines was 72%, while the average mean oleic acid content of high linoleic acid lines was 31.4%. These results indicated a significant difference (R^2= -99.16%) of oleic acid content between donors and recipient breeding genotypes (Figure 3.3).

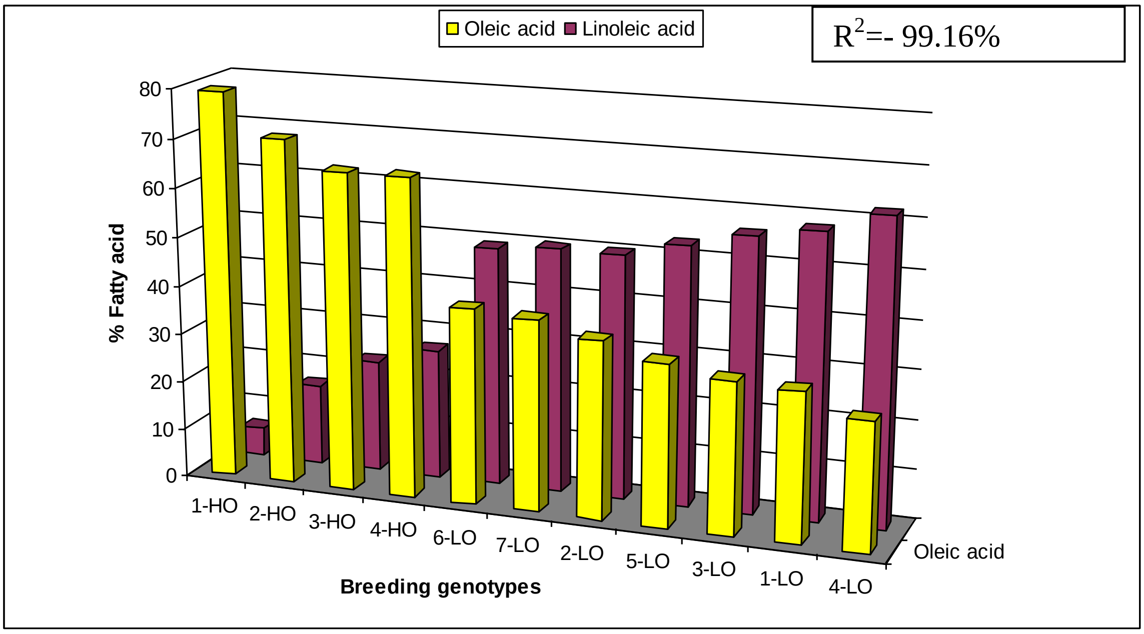


Figure 3.3 The distribution of oleic and linoleic acids in percentage fatty acids as analysed by gas chromatography of the 11 sunflower breeding genotypes listed in Table 3.1.

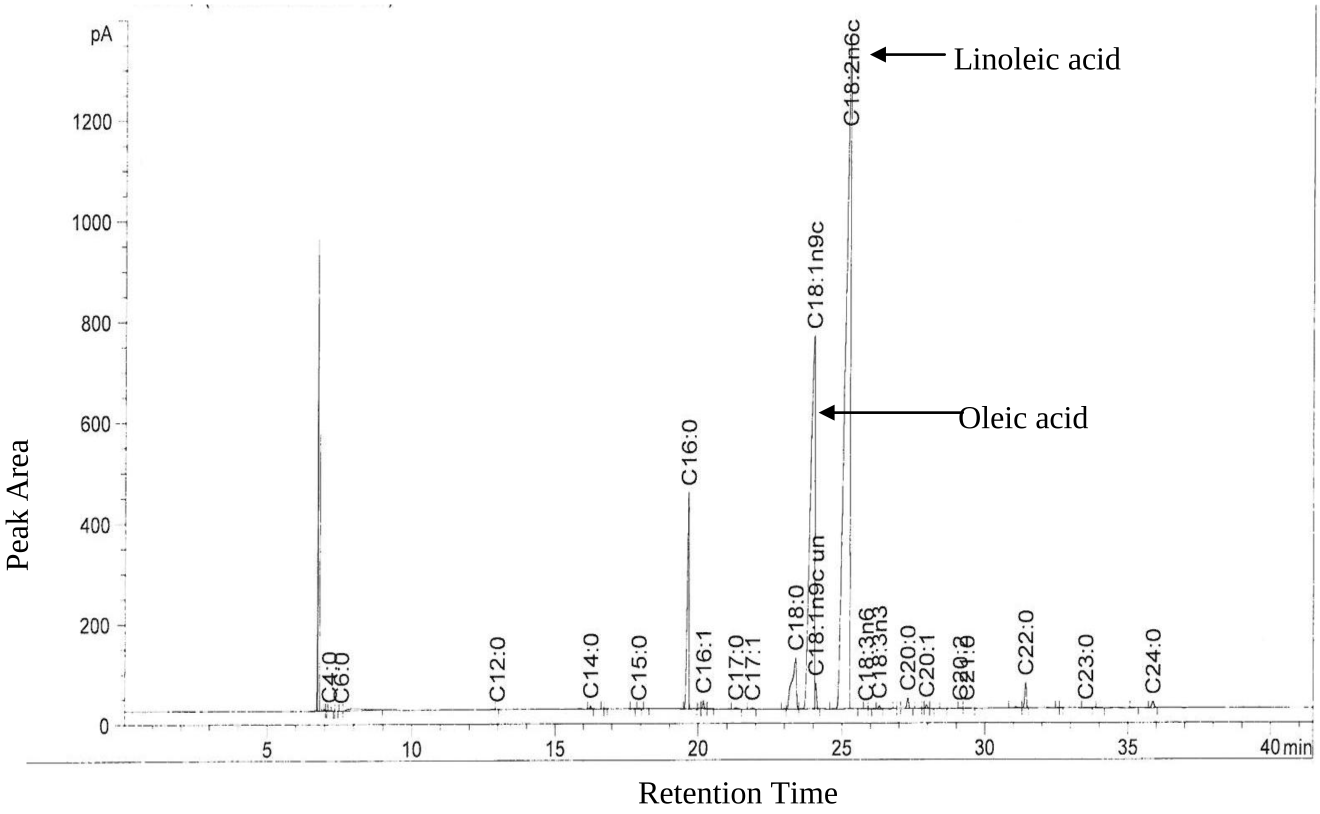


Figure 3.4 The representative high linoleic acid chromatogram during the analysis of low oleic acid sample. Different fatty acid methyl esters with their respective peak areas and retention times are shown.

Different fatty acid methyl esters of sunflower seeds were profiled using gas chromatography to identify peaks having the same retention times as that of the internal standards (results not shown for internal standards) (Figure 3.4). From the results presented in Figure 3.4, it was found that linoleic acid was the main compound with a retention time of 25 minutes followed by an oleic acid having a retention time of 24 minutes. The retention times of these two fatty acids correspond with the retention time of their internal standards respectively. Other fatty acid methyl esters were present in trace amounts.

3.4 DISCUSSION

In an effort to characterise sunflower breeding genotypes based on oleic and linoleic acid levels using a DNA based marker system (SCAR) and chromatographic technique, several dominant polymorphisms were discovered at a molecular level and facilitated the selection of the *Ol* gene among different breeding genotypes.

3.4.1 Optimisation of SCAR markers

A PCR based assay such as SCAR requires a significant optimisation for efficiency and sensitivity of the PCR method that will provide essential guidance on how to troubleshoot an inefficient reaction. In this study, the optimisation experiment was conducted to determine the optimum annealing temperatures and the suitable enzyme for the amplification of SCAR markers in cultivated sunflower. Results showed that temperatures between 55.3^{o}C–65.2^{o}C were optimal for the amplification of SCAR markers on cultivated sunflower. These results however, indicated that the SCAR marker could be amplified over wide ranges of annealing temperatures.

Optimisation of PCR conditions using Super*Therm* polymerase enzyme has resulted in the amplification of a 1 488 bp fragment in both high and low oleic acid genotypes. Since these SCAR markers have a dominant reaction (able to amplify only positive donor lines), this enzyme was however, not usable and therefore not considered for use in identification of markers linked to high oleic acid genotypes due to nonspecific amplification by the enzyme. Results obtained when using Go*Taq* polymerase showed a dominant reaction by only amplifying a 1 488 bp fragment in high oleic acid genotypes. These results are supported by results obtained by Schuppert *et al*. (2006).

3.4.2 SCAR analysis

In an effort to differentiate between mutant homozygous high oleic acid donors (*Ol Ol*) and heterozygous (*Ol ol*) low oleic acid recipient inbred lines (Miller *et al.*, 1987, Schuppert *et al.*, 2006) ten SCAR markers were employed. SCAR analysis of 11 sunflower genotypes indicated that four markers (FAD2-1 F4/R1, FAD2-1 F4/R2, FAD2-1 F14/R2 and FAD2-1 F13/R1) successfully detected a polymorphism between sunflower high and low oleic acid genotypes. Results obtained in this study correspond with previous findings that insertion-deletion markers are diagnostic for the *Ol* mutation (presence or absence of tandem *FAD2-1* repeats) (Schuppert *et al.*, 2006). Based on the SCAR results, these polymorphic markers are applicable in evaluating the genetic state of sunflower breeding lines concerning the presence or absence of the *Ol* gene in the individual plants and will be potentially useful in MAB selection aimed at introgressing the *Ol* allele into elite sunflower cultivars.

The analysis of fatty acid composition of individual donor and recipient seeds showed a considerable variation for oleic and linoleic acids, which were strongly and negatively correlated ($R^2 = -99.16$, $P = 0.001$). This indicated that the relative proportions of oleic and linoleic acids are under the control of one genetic system (Lacombe & Bervillé 2001, Miller *et al*., 1987).

Studies of the inheritance of high oleic acid/low linoleic acid content in sunflower have indicated that several genes are involved, each with a small effect. A single partially dominant gene designated *Ol* was found to control the high oleic acid content (Fick, 1984). Urie (1984) however, reported a complete dominant gene and later reported the existence of reversal in dominance and modifying genes (Rojas-Barros *et al.*, 2005). In a study conducted by Miller *et al*. (1987) a second locus designated (*Ml),* whose recessive allele (*mlml)* was necessary for the expression of the high oleic acid trait was postulated. Finally, Fernández-Martinez *et al*. (1989) identified three complementary dominant genes (*Ol1, Ol2 and Ol3*) controlling the high oleic trait in sunflower seed oil.

However, due to the dominant state of the SCAR markers which makes it difficult to differentiate between mutant homozygote and heterozygote alleles (Schuppert *et al.*, 2006) it is foreseen that still another approach would be needed to develop more high density co-dominant markers to aid in marker-assisted breeding of high oleic acid genotypes. Tang *et* *al.* (2002) screened 1 000 SSR markers. Of these total, 579 markers were found to be polymorphic among four sunflower inbred lines. In a study conducted by Yu *et al.* (2003) 72 INDEL markers were screened between confectionery and oilseed genotypes and only 17 markers were found to be polymorphic.

3.4.3 Fatty acid analysis

To assay fatty acid methyl esters from sunflower with good resolution and at a reasonable elution time, GC was shown to be a reliable and the most accurate technique. The amount and type of free fatty acids can be determined if the GC analysis is carried out post derivatisation of fatty acids to methyl esters (Tong *et al.,* 2007, Wan *et al.,* 2007). Results of GC analysis had confirmed that the 11 sunflower breeding genotypes differed in the quantity of the main unsaturated fatty acid under study and that there is a perfect association between molecular markers and the amount of oleic acid presented by high oleic acid genotypes.

In conclusion, the 11 sunflower breeding genotypes were successfully characterised based on their oleic and linoleic acid levels and (SCAR markers) putatively linked to the high oleic acid trait were identified. For this means, the applicability and validity of these markers will be proven in segregating populations derived from breeding genotypes used in this study.

CHAPTER 4

EVALUATION OF AN F_2 POPULATION SEGREGATING FOR THE HIGH OLEIC ACID TRAIT USING A SCAR MARKER ASSAY

4.1 INTRODUCTION

The development and application of molecular technologies in plant genomics provided the impetus for the intensive genetic mapping studies that have been conducted since 1980. By correlating the pattern of inheritance of the phenotypic trait score with those of individually mapped genetic markers, many monogenic and polygenic characters have been located. Progress has been made in mapping and tagging of many agriculturally important genes with molecular markers which forms the foundation of marker-assisted selection in crop plants (Mohan *et al.,* 1997).

Selection of suitable mapping populations for a marker system is the key requirement for a molecular breeding programme and for studying the inheritance of the gene controlling the trait of interest. Different types of mapping populations exist in a breeding programme and their application depend mostly on the objective of experiments, availability of markers and the known molecular map (Singh & Prasanna, www.scribd.com). High throughput genetic mapping techniques using multiplexed SSR and other DNA sequenced-tagged site (STS) markers have been extensively used for the construction of both whole genome (Tang *et al.,* 2002, Yu *et al.,* 2003) and high resolution maps around loci that control agronomically important traits.

SCAR markers are more reproducible and easier to manipulate in marker-assisted selection programmes. SCAR markers can provide a valid tool for the accurate assessment of genotypes at the linked locus and are useful to screen the whole F_2 population without losing genetic information. In a co-dominant marker system, homozygous and heterozygous plants can be identified in a segregating population by discriminating between different alleles (Bi-Hao *et al.,* 2009). In this study, a F_2 segregating population is used to study the inheritance of the gene controlling the high oleic acid trait because it requires less time to develop therefore, this chapter aimed at validating the existing SCAR marker system optimised in Chapter 3 through evaluating the F_2 population to confirm the inheritance of the high oleic acid trait and for further mapping of the trait. These markers will be applied in the backcross breeding programme.

4.2 MATERIALS AND METHODS

4.2.1 Production of F_1 populations and DNA extraction

Sunflower breeding materials with high and low oleic acid content were obtained from Dr. P.J.A. van der Merwe at the ARC-GCI, Potchefstroom. A total of four different F_1 populations were developed using four different sets of parental lines. AP901-95-3-4-1 (HO) and H55-9-2-1-1 (LO) inbred lines were selected for crosses and used in the production of a F_2 segregating population based on the early flowering trait. The inbred line crosses summarised in Table 4.1 were done as follows: A total of 20 seeds of both high and low oleic acid plants were planted in the greenhouse in 10 litre plastic pots in sterile soil. The greenhouse was maintained at 28^oC day and 18^oC night temperatures. At flowering stage crosses were made between high and low oleic acid inbred lines. Pollen was harvested from male lines and stored at -20^oC and anthers of female lines were removed using sterile tweezers before sunrise. Pollination was conducted seven hours later when stigmas of female lines were fully developed to receive pollen. The pollination process was repeated daily until the stigma production cycle was complete. F_1 seeds of each population were harvested at maturity stage. Only the F_1 population derived from a cross between AP901-95-3-4-1 (HO) and H55-9-2-1-1 (LO) inbred lines were allowed to self-pollinate to produce 143 plants of the F_2 segregating population. At seedling stage, leaf samples were harvested from these 143 plants of the F_2 population and total genomic DNA was extracted through employing a CTAB extraction protocol (Saghai-Maroof *et al.,* 1984) similar to the one described in section 3.2.1. Quantification and the determination of the quality of DNA were assayed by methods described in section 3.2.2.

Table 4.1. Sunflower inbred lines used for crosses in the greenhouse for the production of four F_1 populations. SCARs were used as foreground markers for the selection of individual plants of the F_1 populations possessing the high oleic acid allele.

F_1 Pedigree	**Recurrent parent**	**Markers used for selection**
AP901-94-3-2-2 X AP991-76-1-4-1	AP991-76-1-4-1	FAD2-1 F4/R1
AP901-56-1-1 X ND761-3-1-1	ND761-3-1-1	FAD2-1 F4/R2
AP901-56-1-1 X RK74-26-1-2-1	RK74-26-1-2-1	FAD2-1 F13/R1
AP901-95-3-4-1 X H55-9-2-1-1	H55-9-2-1-1	FAD2-1 F14/R2

4.2.2 Genotyping of a F_2 population using dominant SCAR markers

The authenticity of the F_2 population to exhibit the expected segregating ratio was verified using SCAR markers (Schuppert *et al.*, 2006) (Invitrogen, Scientific Group, Carlsbad USA), which generated informative polymorphisms in section 3.3.2. PCR of the DNA from each of the 143 plants of the F_2 population was performed in a 20 µl reaction volume with dominant SCAR markers. The PCR conditions were similar to the procedure described in section 3.2.4. PCR products were run on a 2% (m/v) agarose gels (Seakem, USA) using the method similar to the one described in section 3.2.2. The image of the gel was captured using a gel doc system (BIO-RAD Laboratories, Hercules, California, USA) as described in section 3.2.2 (Figure 4.1).

4.2.3 Analysis of data

A Chi-square test (Pérez-Vich, 2000) was applied to determine the significance of deviations from the expected 3:1 segregation ratio for dominant SCAR markers in 143 plants of the F_2 segregating population derived from AP901-95-3-4-1(HO) and H55-9-2-1-1 (LO) inbred lines.

4.3 RESULTS

4.3.1 Genotyping of 143 plants of the F_2 segregating population using SCAR markers

One hundred and forty three (143) F_2 plants derived from a cross between AP901-95-3-4-1 (HO) X H55-9-2-1-1 (LO) were screened against four dominant polymorphic SCAR markers (FAD2-1F4/R1, FAD2-1F4/R2, FAD2-1F13/R1 and FAD2-1F14/R2) to evaluate the genetic state concerning the high oleic acid gene (*Ol*) of an individual plant as well as to validate the applicability and reliability of each polymorphic marker. These polymorphic markers amplified fragment sizes of 787, 1 488, 1 758 and 2 000 bp respectively, in a coupling phase linkage (Figure 4.1 a, b, c and d).

Of the 143 F_2 plants screened with marker FAD2-1 F4/R1, the 787 bp fragment was amplified in coupling phase in 107 plants and 36 plants showed no amplification. Marker FAD2-1 F4/R2 amplified the 1 488 bp in 109 out of 143 individuals of the F_2 population. Marker FAD2-1 F13/R1 amplified the 1 758 bp fragment in 108 out of 143 F_2 plants and FAD2-1 F14/R2 marker amplified the 2 000 bp fragment in 110 individuals. The Chi-square test of the observed frequencies of each marker locus in 143 plants of the F_2 population indicated that deviation from the expected ratio of 3:1 was not significant (Table 4.3).

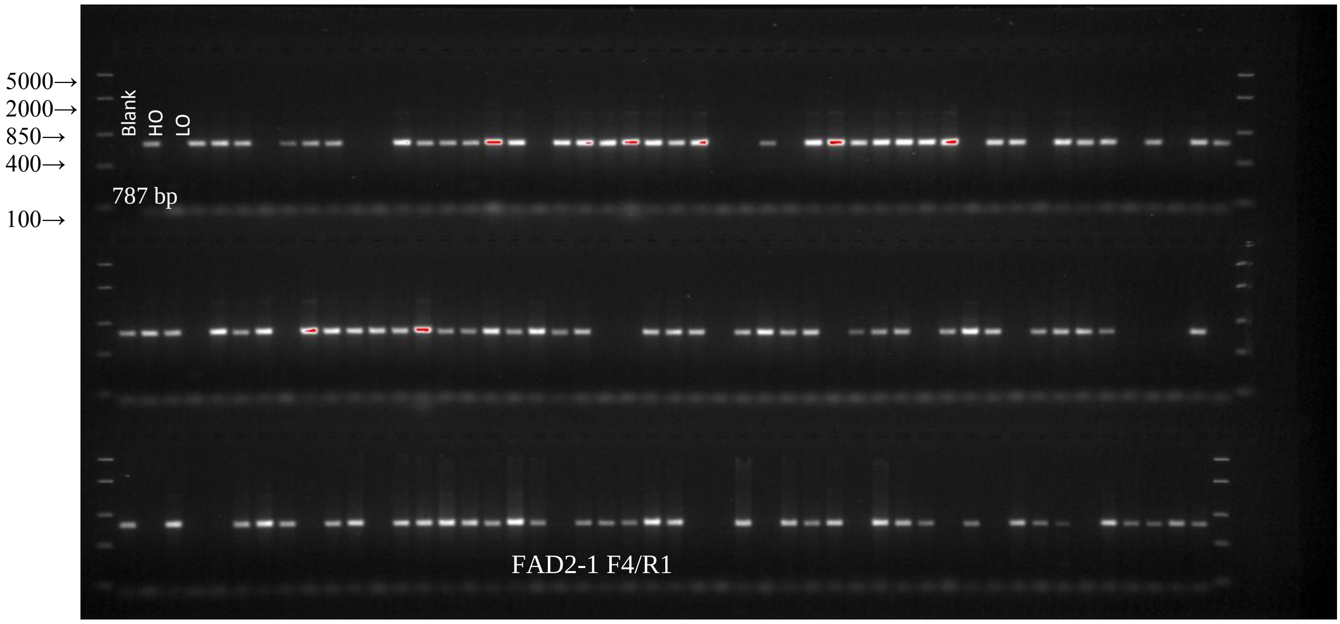


Figure 4.1 (a) Genotyping of 143 plants from an F_2 mapping population resulting from a cross between AP901-95-3-4-1 and H55-9-2-1-1 using dominant SCAR marker FAD2-1F4/R1.

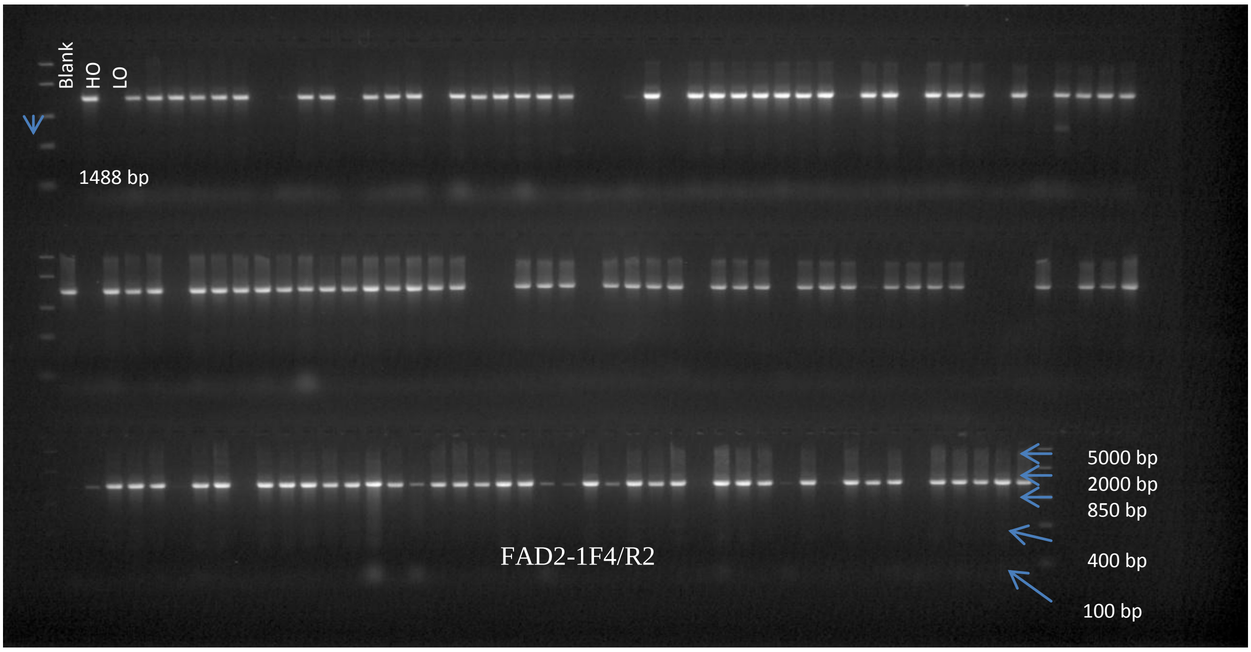


Figure 4.1 (b) Genotyping of 143 plants from an F_2 mapping population resulting from a cross between AP901-95-3-4-1 and H55-9-2-1-1 using dominant SCAR marker FAD2-1F4/R2

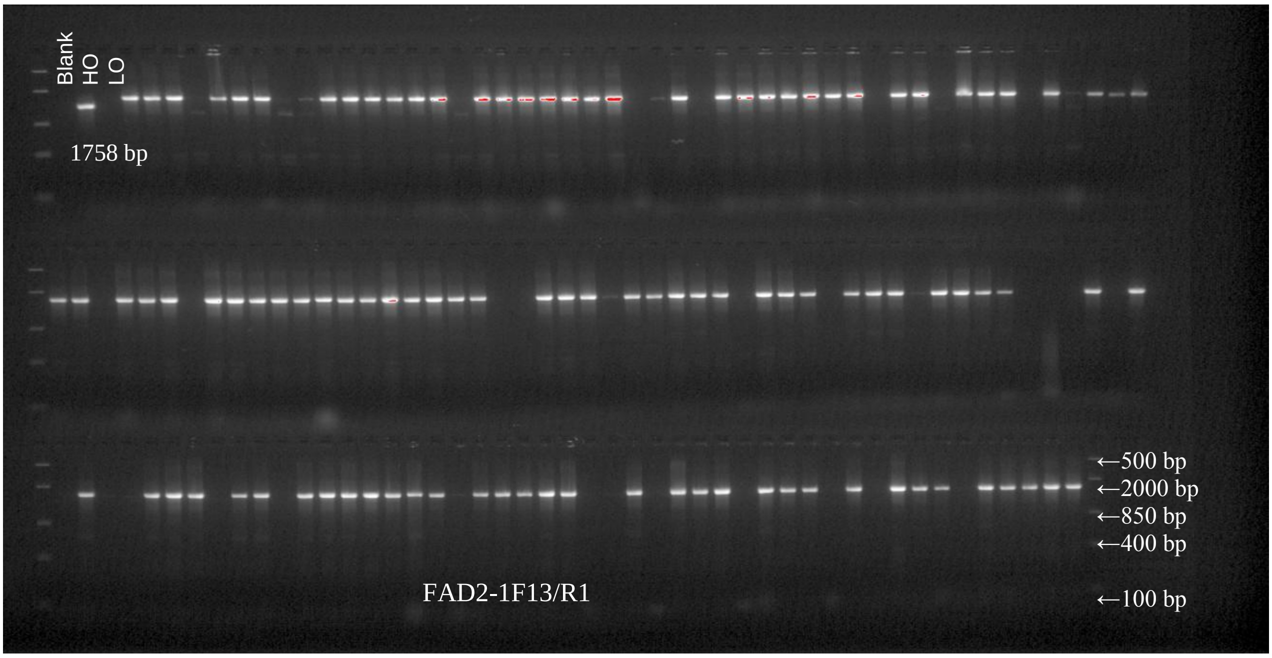


Figure 4.1 (c) Genotyping of 143 plants from an F_2 mapping population resulting from a cross between AP901-95-3-4-1 and H55-9-2-1-1 using dominant SCAR marker FAD2-1F13/R1.

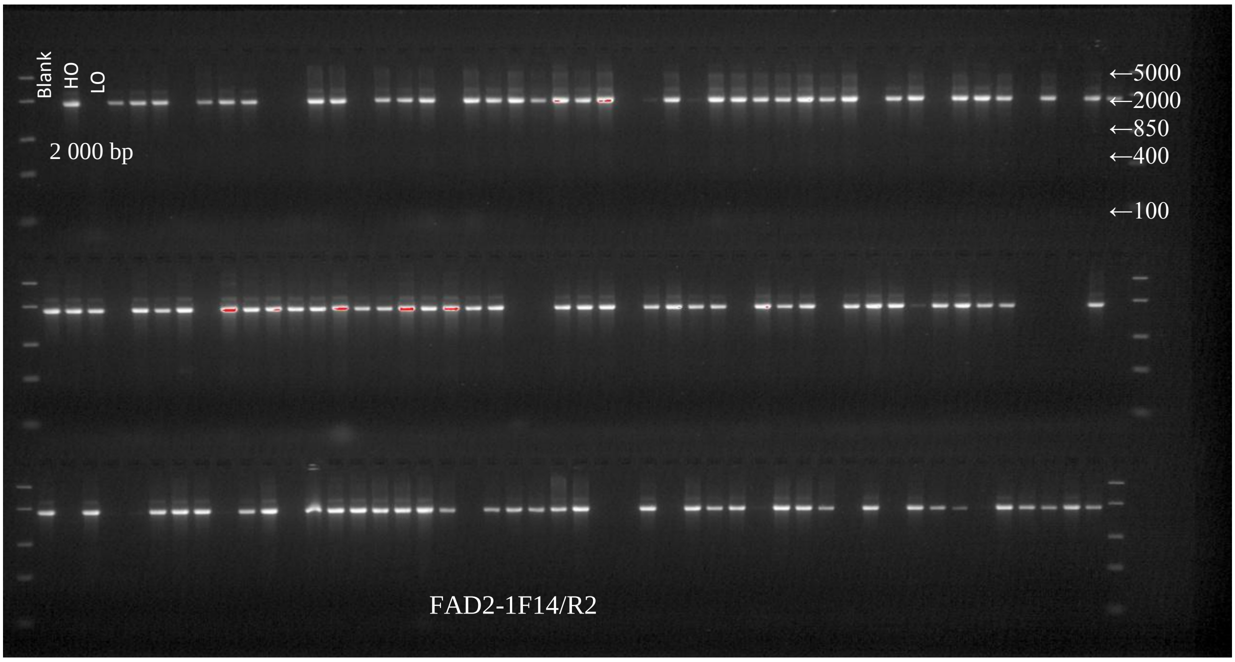


Figure 4.1 (d) Genotyping of 143 plants from an F_2 mapping population resulting from a cross between AP901-95-3-4-1 and H55-9-2-1-1 using dominant SCAR marker FAD2-1F14/R2.

Table 4.2 Chi square test for the analysis of four dominant SCAR markers to evaluate the segregation ratio in 143 plants of the F_2 population derived from a cross between AP901-95-3-4-1 (HO) and H55-9-2-1-1 (LO).

Pedigree	**Marker**	**Observed ratio Number of F_2 plants**			**Expected ratio**	**Chi^2**	**P value**
		Positive		**Negative**			
AP901 (HO) X H55 (LO)	FAD2-1F4/R1	107	:	36	3:1	0.001	<0.95
AP901 (HO) X H55 (LO)	FAD2-1F4/R2	109	:	34	3:1	0.12	>0.80
			:				
AP901 (HO) X H55 (LO)	FAD2-1F13/R1	108	:	35	3:1	0.66	<0.30
AP901 (HO) X H55 (LO)	FAD2-1F14/R2	110	:	33	3:1	0.30	<0.50

4.4 DISCUSSION

The authenticity of 143 individual F_2 plants derived from a cross between a high oleic acid parent (AP 901-95-3-4-1) and low oleic acid parent (H 55-9-2-1-1) was verified using SCAR analysis. The inheritance pattern of the gene responsible for the expression of high oleic acid (*Ol*) as observed with dominant SCAR markers (FAD2-1F4/R1, FAD2-1F4/R2, FAD2-1F13/R1, and FAD2-1F14/R2) confirmed that each of the DNA fragments amplified by these markers inherited at least one allele from the high oleic acid male parent. The inheritance pattern, furthermore confirmed that the F_2 individuals were progeny from true F_1 hybrids of the cross and did not result from self-pollination of the female parent. A chi-square test of the observed frequencies of each marker locus in 143 plants of the F_2 population indicated that deviation from the expected ratio of 3:1 (high to low oleic acid) was not statistically significant from the expected ratio (Table 4.3), and the observed marker segregation ratio of 3:1 for the F_2 population was consistent with the previous findings that an incomplete dominant gene governs the sunflower high oleic acid trait (Fernández-Martinez *et al.*, 1989, Fick, 1984, Urie, 1985). However, due to financial constraints, analysis of the fatty acid composition of the F_2 population was not carried out to support results obtained at genotypic level as well as to verify the maternal inheritance of the high oleic acid trait as was hypothesised by Miller *et al*. (1987).

SCAR markers linked to the sunflower (*Ol*) locus provides a useful approach for early identification of lines carrying the high oleic acid allele. Sunflower plants carrying an incomplete dominant mutation of the *Ol* gene produce a seed oil with high oleic acid content, which is correlated with the reduced expression of a seed-specific oleoyl-phosphatidyl choline desaturase enzyme (FAD2-1) in the development of seeds of sunflower (Schuppert *et al.*, 2006). The *Ol* locus was mapped by Schuppert *et al.* (2006) between FAD2-1 SCAR markers at a genetic distance of 62.1 cM on a carrier chromosome and found to co-segregate with co-dominant SSR markers (CRT20, CRT25 and ORS333). These markers will be potentially useful in a marker-assisted selection programme aimed at introgressing *Ol* alleles into the background of the sunflower cultivars.

Studies of the inheritance of the oil quality in other oilseed crops have been carried out using different types of marker systems with each having the potential to reveal the genetic state of the gene influencing the trait. In a study conducted by Hamdan *et al.* (2008), molecular markers closely linked to the gene (*Li*) controlling high linoleic acid content and gene for controlling nuclear male sterility (*Ms*) were developed using the RAPD technique. These markers were further converted into SCAR markers which then revealed a recessive allele at a single locus in safflower (Futehally & Knowles, 1981).

In a further investigation of the inheritance studies in oilseed crops, Pérez-Vich *et al.* (2002) studied the genetic relationship between the loci controlling the high stearic and high oleic acid trait in a sunflower F_2 segregating population. The results showed that the oleic acid content was assigned to two phenotypic classes: low-intermediate and four of the six F_2 populations analysed satisfactorily fitted a 1:3 ratio. The stearic acid content was assigned to two phenotypic classes: low-intermediate and all F_2 populations analysed were adjusted to a 15:1 ratio (low: intermediate).

In conclusion, the inheritance of the high oleic acid trait was successfully determined at a genotypic level using a SCAR marker assay. The inheritance of the high oleic acid trait indicated that the segregation fitted a model of 3:1 ratio with a partial dominance for the high oleic acid gene. However, the application of this marker system will be carried out through introgression of the high oleic acid allele in the background of the recurrent parent genome. The backcross programme can be shortened by selecting the individual plants in each generation with the highest content, i.e. the most alleles from the recurrent parent. Application of SCAR markers together with SSRs as foreground and background markers respectively, for the diagnosis of high oleic acid allele and the selection of the recurrent parent genome in backcross breeding programme will be examined in Chapter 5.

CHAPTER 5

APPLICATION OF SCAR MARKER AND MICROSATELLITE ASSAYS FOR THE INTROGRESSION OF THE HIGH OLEIC ACID TRAIT INTO THE BACKGROUND OF THE RECURRENT PARENT GENOME

5.1 INTRODUCTION

The development of molecular markers and genetic maps assist in understanding of the genetic base of economically important traits and facilitates plant breeding via MAS and MAB (Yue *et al.,* 2009). MAB provides an important tool to plant breeders to practice indirect selection of agronomically important traits and improve the efficiency of plant breeding through precise transfer of genomic regions of interest and accelerate the recovery of the recurrent parent genome (Ibitoye & Akin-Idowu, 2010). In backcross breeding markers can be used to tag the target gene in a process known as foreground selection or be used to select the genetic background of the recurrent parent in a process known as background selection (Hospital, 2002).

The improvement of sunflower breeding genotypes through an introgression of the high oleic acid gene in the background of the recurrent parent genome requires a method that can speed up the recovery of the recurrent parent's genome and will help to reduce the genetic background of the donor parent's genome in the progeny. MAB offer such a possibility by employing markers that are tightly linked to the gene controlling the trait (Frisch *et al.,* 1999). The efficiency of MAB rely mostly on a genetic map having enough uniformly spaced polymorphic markers to precisely locate the desired gene, a close linkage between the marker and the gene of interest as well as an adequate recombination between the marker and the rest of the genome (Ibitoye & Akin-Idowu, 2010, Semagn *et al.,* 2006).

In this study, a conventional breeding technique was used to introgress one major high oleic acid gene into four sunflower recurrent parents from three high oleic acid donor lines differing in the level of oleic acid content. At each backcross generation, SCAR markers closely linked to the high oleic acid allele were used to select plants possessing the high oleic acid gene during foreground selection. During background selection, SSR markers (multiplex assay) polymorphic between the donor and recurrent parent were used to select plants that had the maximum recurrent parent genome contribution in order to achieve NILs (progenies almost identical to the recurrent parent). Molecular SCAR markers used in this study were already evaluated and tested for validity in Chapters 3 and 4, respectively.

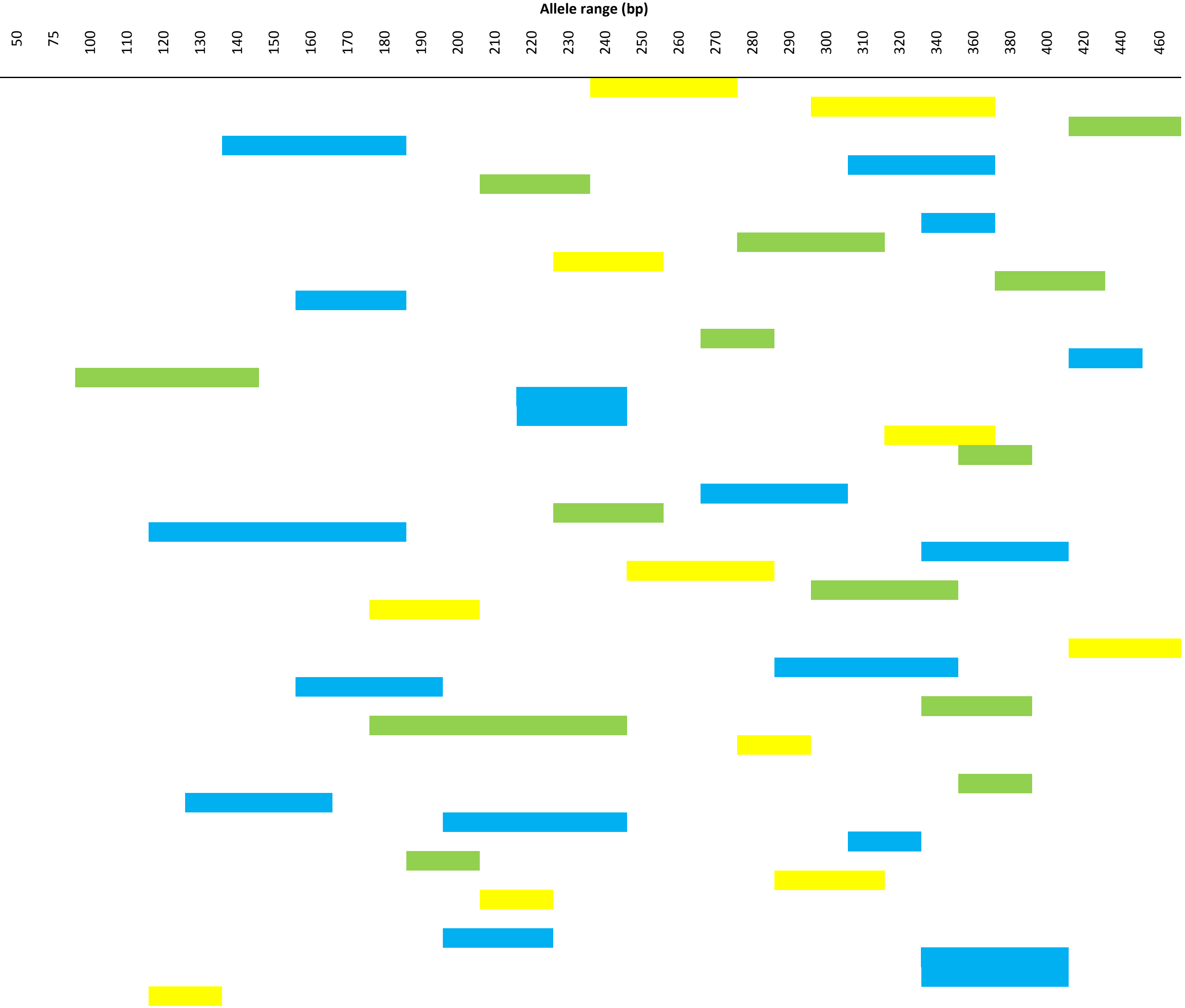
5.2 MATERIALS AND METHODS

5.2.1 Multiplex genotyping assay to optimise and evaluate suitable SSRs to be used as background markers in marker-assisted backcross breeding

The 143 plants of the F_2 segregating population used in this study was developed and quantified in section 4.2.1. Seventy eight fluorescently-labelled SSR markers (Applied BioSystems, Warrington, UK) published by Tang *et al*. (2002) were optimised for parents of the F_2 populations derived from crosses as described in section 4.2 in multiplex sets to identify background SSR markers that can be used to estimate the amount of recurrent parent genome contribution in marker-assisted background selection (Table 5.1). These SSR markers were selected from 1 089 public SSR markers on the basis of their polymorphic information content (PIC), proximity to known sunflower high oleic acid loci, map positions and distribution throughout the sunflower genome (Blair *et al.*, 2002, Tang *et al.*, 2002, Yu *et al.*, 2003). Between five and seven-plexes of SSR markers labelled with three different fluorophores (6-FAM, NED and VIC) were grouped into 14 panels (Table 5.1). SSR markers were sorted by allele-length range and combined so as to minimise co-migration of identically labelled non-allelic bands. Multiplex PCR analysis was performed in a 10 µl reaction mix containing 30 ng genomic DNA, 6 µl of 2X multiplex PCR master mix (Qiagen, Valencia, USA) containing HotStar*Taq* DNA polymerase, multiplex PCR buffer, 2 mM $MgCl_2$, 250 mM of each dNTP and 0.05-0.15 µM of each of the fluorescent primers. Reactions were performed using a thermal cycler multi-block system (Thermo Fisher Scientific, USA) and a touchdown cycling programme, similar to the one used by Tang *et al.* (2002), was used: Denaturing for 1 cycle at 94°C for 30 seconds, followed by 7 cycles of 94°C for 30 seconds, 64°C for 30 seconds and 72°C for 30 seconds, with a temperature decrease of 1°C each cycle. The third stage was performed for 32 cycles at 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 1 minute and 30 seconds, with a final stage of 1 cycle at 72°C for 30 minutes and a hold temperature of 4°C. A volume of 30 µl of Gene scan 500 ROX internal size standard (Applied BioSystems, USA) was mixed with 1 ml Hi-Di formamide (Applied BioSystems, USA) in a 1.5 ml micro-centrifuge tube. Nine µl of the mixture was transferred to a 384 well microtitre plate containing 2 µl of PCR products. Samples were denatured at 94^{o}C for 5 minutes, chilled on ice for 5 to 10 minutes and centrifuged briefly. A multiplex genotyping assay similar to the one described by Tang *et al.* (2002) was performed on an ABI Prism3130XL genetic analyser (Applied BioSystems) by electrophoresis of the fluorescently labelled amplicons under the following run conditions: Oven temperature at 60^{o}C, injection voltage at 1.2 kV, injection time of 23 seconds and run voltage of 15 kV. A matrix filter set (G5) configured for five dye sets (DS-33) (ROX-Red, 6-FAM-Blue, HEX-Green, NED-Yellow and LIZ-Orange) were used to assay labelled amplicons.

Table 5.1 The names, colour, repeat, allele length and linkage groups of 78 ORS multiplex SSR markers screened for polymorphism among parents and 143 plants of the F_2 segregating population of a cross between AP901-95-3-4-1 (HO) and H55-9-2-1-1 (LO) inbred lines.

anels	Marker name	Label	Colour	Linkage group	Allele Length (bp)	Repeat
Panel 1	ORS 543	NED	Yellow	1	244-268	$(CT)_{23}$
	ORS 342	NED	Yellow	2	308-341	$(GT)_{10}$
	ORS 1222	HEX	Green	3	436-442	$(CT)_{11}$
	ORS 533	6FAM	Blue	5	144-174	$(CT)_{17}$
	ORS 456	6FAM	Blue	8	310-322	$(GT)_{10}$
	ORS 297	VIC	Green	17	217-229	$(GT)_{12}$
Panel 2	ORS 674	6FAM	Blue	4	345-358	$(CT)_{10}$
	ORS 844	HEX	Green	9	285-309	$(AC)_{17}$
	ORS 621	NED	Yellow	11	234-251	$(CT)_{14}$
	ORS 810	HEX	Green	12	399-406	$(CT)_{20}$
	ORS 687	6FAM	Blue	15	165-171	$(CT)_{14}$
Panel 3	ORS 483	HEX	Green	6	270-272	$(GT)_{15}$
	ORS 1178	6FAM	Blue	7	420-431	$(CT)_{21}$
	ORS 595	HEX	Green	10	105-148	
	ORS 457	6FAM	Blue	11	225-231	$(AC)_{12}(CT)_6$
	ORS 630	NED	Yellow	13	339-349	$(CT)_{15}$
	ORS 1248	VIC	Green	14	372-376	$(AG)_{17}$
Panel 4	ORS 665	6FAM	Blue	3	272-295	$(AG)_{15}$
	ORS 505	HEX	Green	5	235-249	$(AC)_{10}$
	ORS 70	6FAM	Blue	8	125-176	
	ORS 1146	6FAM	Blue	11	346-383	$(AG)_{15}$
	ORS 1180	NED	Yellow	14	250-280	$(CT)_{12}$
	ORS 899	HEX	Green	16	301-322	$(AC)_{19}$
	ORS1245	NED	Yellow	17	184-196	$(AG)_{20}$
Panel 5	ORS 837	NED	Yellow	1	430-442	$(GT)_6$
	ORS 716	6FAM	Blue	1	299-323	$(AG)_{24}$
	ORS 229	6FAM	Blue	2	167-185	
	ORS 695	HEX	Green	4	356-370	$(AG)_{16}$
	ORS 1265	HEX	Green	9	189-236	$(CT)_{18}$
	ORS 1085	NED	Yellow	12	280-283	$(CT)_{13}$
Panel 6	ORS 230	HEX	Green	6	360-366	
	ORS 16	6FAM	Blue	7	135-152	
	ORS 613	6FAM	Blue	6	204-234	$(AG)_{16}$
	ORS 650	6FAM	Blue	6	415-416	$(CT)_8$
	ORS 316	HEX	Green	13	181-191	$(AC)_{11}$
	ORS 1079	NED	Yellow	14	359-398	$(CT)_{26}$
	ORS 857	NED	Yellow	15	212-218	$(AC)_{13}$
	ORS 925	6FAM	Blue	2	201-215	$(AC)_{16}$
	ORS 949	6FAM	Blue	3	359-398	$(GT)_{10}(AT)_5$
	ORS 309	NED	Yellow	4	121-130	$(A)_{19}$



Panel 7	ORS 894	NED	Yellow	8	253-263	$(GT)_{14}$
	ORS 887	VIC	Green	9	245-253	$(AC)_{11}$
	ORS 307	HEX	Green	14	114-136	$(AT)_{8}(GT)_{12}$
	ORS 885	HEX	Green	16	339-342	$(AC)_{6}$
Panel 8	ORS 610	6FAM	Blue	1	145-159	$(AG)_{14}$
	ORS 878	NED	Yellow	10	194-214	$(AC)_{11}$
	ORS 1227	NED	Yellow	11	351-357	$(CT)_{25}$
	ORS 7	VIC	Green	15	267-273	
	ORS 561	6FAM	Blue	17	362-434	$(AG)_{20}$
Panel 9	ORS 1024	VIC	Green	5	216-234	$(AG)_{22}$
	ORS 1041	NED	Yellow	7	276-284	$(CT)_{19}$
	ORS 938	VIC	Green	9	312-322	$(GT)_{20}$
	ORS 437	6FAM	Blue	10	334-348	$(AC)_{13}$
Panel 10	ORS 1161	6FAM	Blue	8	226-235	$(CT)_{17}$
	ORS 778	NED	Yellow	12	380-384	$(AG)_{11}$
	ORS 1179	NED	Yellow	13	317-320	$(AG)_{18}$
	ORS 832	6FAM	Blue	14	355-362	$(GT)_{9}$
Panel 11	ORS 1036	NED	Yellow	3	245-255	$(CT)_{23}$
	ORS 428	6FAM	Blue	9	217-225	$(GT)_{12}$
	ORS 691	NED	Yellow	10	451-466	$(CT)_{19}$
	ORS 750	VIC	Green	16	328-344	$(AG)_{19}$
Panel 12	ORS 371	VIC	Green	1	255-263	$(GT)_{7}$
	ORS 423	NED	Yellow	2	238-244	$(GT)_{18}$
	ORS 963	6FAM	Blue	4	333-341	$(GT)_{10}$
	ORS 966	NED	Yellow	7	370-375	$(GT)_{9}$
	ORS 666	6FAM	Blue	11	249-251	$(ACC)_{5}$
	ORS 407	VIC	Green	16	430-466	$(GT)_{13}$
Panel 13	ORS 1114	VIC	Green	3	242-256	$(CT)_{13}$
	ORS 774	6FAM	Blue	5	296-312	$(AG)_{17}$
	ORS 166	NED	Yellow	8	334-345	$(GT)_{7}$
	ORS 442	VIC	Green	9	391-409	$(GT)_{8}(GA)_{7}$
	ORS 502	NED	Yellow	12	95-119	$(AACAC)_{7}$
	ORS 1030	6FAM	Blue	13	434-437	$(GGT)_{5}$
Panel 14	ORS 1065	NED	Yellow	2	275-297	$(CT)_{12}$
	ORS 761	HEX	Green	12	349-342	$(AG)_{14}$
	ORS 735	6FAM	Blue	17	359-373	$(AG)_{17}$
	ORS 668	VIC	Green	15	158-160	$(AG)_{6}$
	ORS 534	6FAM	Blue	13	251-257	$(CT)_{17}$

5.2.2 Marker-assisted foreground selection of the F_1 populations

DNA of four F_1 populations derived from high and low oleic acid inbred lines used in this study were quantified in section 4.2.1. The SCARs were used as foreground markers for the diagnosis of plants having the high oleic acid allele (Schuppert *et al.,* 2006). PCR reactions were performed using a touchdown protocol similar to the one described in section 5.2.1. Gel electrophoresis and gel imaging were done as described in sections 3.2.2 and 3.2.3.

5.2.3 Production and marker-assisted foreground and background selection of BC_1F_1 populations

Seeds of each of the four recurrent parents and their respective F_1 seeds that were developed in section 4.2.1 were planted in 10 litre plastic pots containing sterile soil. At flowering stage, pollen was harvested from plants of the F_1 generations positive for the high oleic acid allele and backcrossed to the recurrent parents using the breeding method and greenhouse conditions similar to that described in section 4.2.1. At maturity, seeds of BC_1F_1 individuals from each recurrent parent were harvested and planted for selection of the high oleic acid trait. Selection of plants of BC_1F_1 populations carrying the high oleic acid allele and possessing maximum recurrent parent genome contribution was done using methods similar to that explained in sections 5.2.1 and 5.2.2.

5.2.4 Production and marker-assisted selection of the second backcross generation (BC_2 F_1)

Following selection of positive BC_1F_1 plants for the high oleic acid allele, the resulting BC_1 F_1 positive plants were allowed to grow in the greenhouse under similar greenhouse conditions specified in section 5.2.2. Recurrent parents were planted approximately 14 days apart to allow synchronisation. During the flowering period, pollen was harvested from BC_1 F_1 plants respectively, and pollinated on the recurrent parents using the method similar to the one explained in section 5.2.3. At maturity, BC_2F_1 seeds were harvested from each recurrent parent and planted for DNA extraction and subsequently selected for high oleic acid trait using only foreground markers (SCAR markers) due to insufficient population sizes achieved in the second backcrossing and insufficient background markers available for recurrent parent genome recovery.

5.2.5 Data analysis

The data obtained from screening of the 143 plants of the F_2 mapping population using co-dominant multiplex markers was first analysed using Genemapper software version 4 (Applied BioSystems). The genotype plot for each marker was examined and the size of alleles associated with a marker was determined with maximum of two alleles per marker. The software, Graphical genotypes (GGT) version 2.0 (www.plantbreeding.wur.nl) was used for genetic map construction of 78 SSR markers to illustrate the location of markers.

The percentage recovery of the recurrent parent’s genome contribution from plants of the BC_1F_1 populations was calculated using a descriptive calculation method in which the number of heterozygous alleles that exist between individual plants of the population was divided by the total number of background markers screened, using the following formula:

$G = (A/B)/N*100\%$
A/B = Heterozygous alleles
N = Total number of polymorphic markers used

5.3 RESULTS

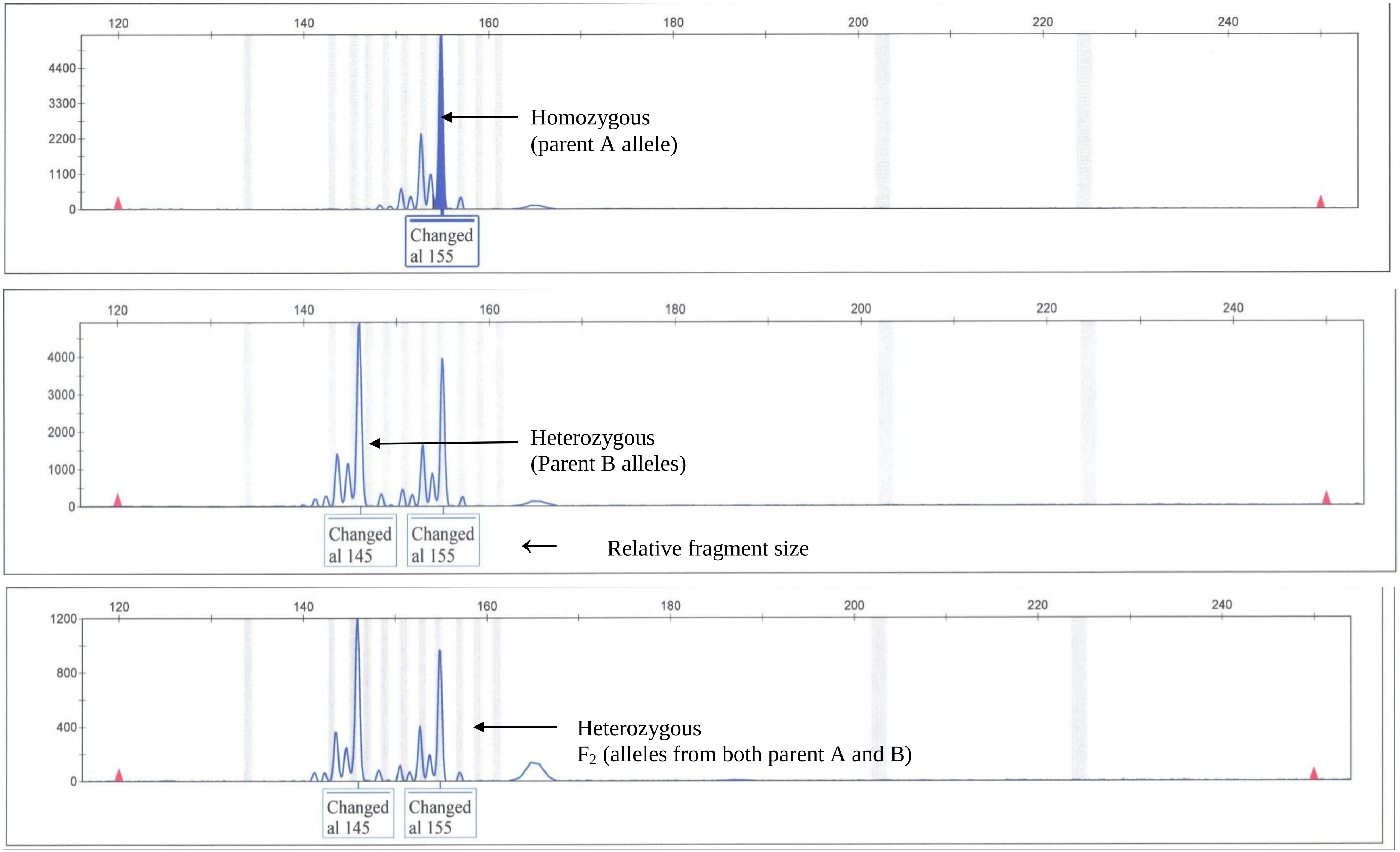
5.3.1 Multiplex genotyping assay to optimise and evaluate suitable SSRs to be used as background markers in marker-assisted backcross breeding

Seventy eight fluorescent-labelled SSR markers published by Tang *et al.* (2002) were optimised for the screening of 143 plants of the F_2 segregating population derived from a cross between H55-9-2-1-1 (LO) and AP901-95-3-4-1 (HO) in multiplex sets to identify background SSR markers that could be used to estimate the amount of recurrent parent genome contribution in marker-assisted background selection. Of the 78 SSR markers screened, 15 markers were polymorphic among the two inbred lines with eight displaying co-dominant reactions (ORS483, ORS543, ORS630, ORS1265, ORS316, ORS309, ORS885 and ORS925) and seven exhibiting dominant reactions (ORS1178, ORS621, ORS674, ORS810, ORS595, ORS230 and ORS613), nine markers were monomorphic, 19 markers amplified more than one allele but were not polymorphic between the two control parents (non-informative markers) and 35 had null alleles or did not amplify (Table 5.2). The minimum and maximum allele number observed between polymorphic SSR markers was 2.0 and 6.0, and a total of 47 alleles were produced. The average mean of allele numbers for polymorphic markers was 3.0. The minimum and maximum PIC values of polymorphic markers were 0.350 and 0.580. The minimum and maximum gene diversity of polymorphic markers was between 0.460 and 0.650 and the total average mean of gene diversity of all markers was 0.270.

Table 5.2 The characteristics of polymorphic SSR markers used in the screening of 143 plants of the F_2 population derived from a cross between high and low oleic acid inbred lines.

Polymorphic markers	Allele Frequency	Allele Number	Gene diversity	Heterozygosity	Polymorphic information content
ORS543	0.5000	3.0000	0.5068	0.4345	0.3852
ORS621	0.5000	3.0000	0.5068	0.4483	0.3852
ORS674	0.5207	4.0000	0.5121	0.4414	0.3944
ORS810	0.4444	5.0000	0.6431	0.4931	0.5754
ORS1178	0.5793	6.0000	0.5097	0.6276	0.4057
ORS483	0.5207	3.0000	0.5057	0.4966	0.3846
ORS595	0.5345	3.0000	0.5160	0.5724	0.4026
ORS630	0.5243	2.0000	0.4988	0.5486	0.3744
ORS1265	0.5207	3.0000	0.5057	0.5172	0.3846
ORS316	0.6034	2.0000	0.4786	0.6828	0.3641
ORS230	0.5243	2.0000	0.4988	0.4931	0.3744
ORS617	0.5035	2.0000	0.5000	0.5315	0.3750
ORS309	0.4236	4.0000	0.6529	0.7847	0.5820
ORS885	0.6329	2.0000	0.4647	0.5944	0.3567
ORS925	0.5175	2.0000	0.4994	0.4615	0.3747

Amongst 143 individual plants of the F_2 population screened with SSR markers listed in Table 5.1, DNA fragments of 126 individuals of the F_2 population were successfully amplified during a multiplex assay. GeneScan ROX-500 size standard as a molecular weight marker estimated the sizes of each of the amplified alleles, the minimum allele size was 50 bp and the maximum allele size was 500 bp. As an example, of the 126 individuals screened with marker ORS621, 38 individuals had an allele size of 155 bp (homozygous to the high oleic acid parent), 38 individuals had an allele size of 145 bp (homozygous to the low oleic acid parent) and 50 individuals had both the parent alleles (heterozygous for 155 bp and 145 bp alleles) (Figure 5.1 a). The results indicated in Figure 5.1 (b) represent one of 9 monomorphic markers screened for polymorphism in the F_2 population. The allele size of 361 bp was observed in both parental lines as well as in 143 plants of the F_2 population when ORS687 marker was screened.



Homozygous
(parent A allele)

Heterozygous
(Parent B alleles)

Relative fragment size

Heterozygous
F_2 (alleles from both parent A and B)

Figure 5.1 (a) Electropherogram showing the analysis of multiplexed labelled microsatellite PCR products. Peaks represent microsatellite alleles of both parental lines and heterozygous F_2 individuals screened with ORS621.

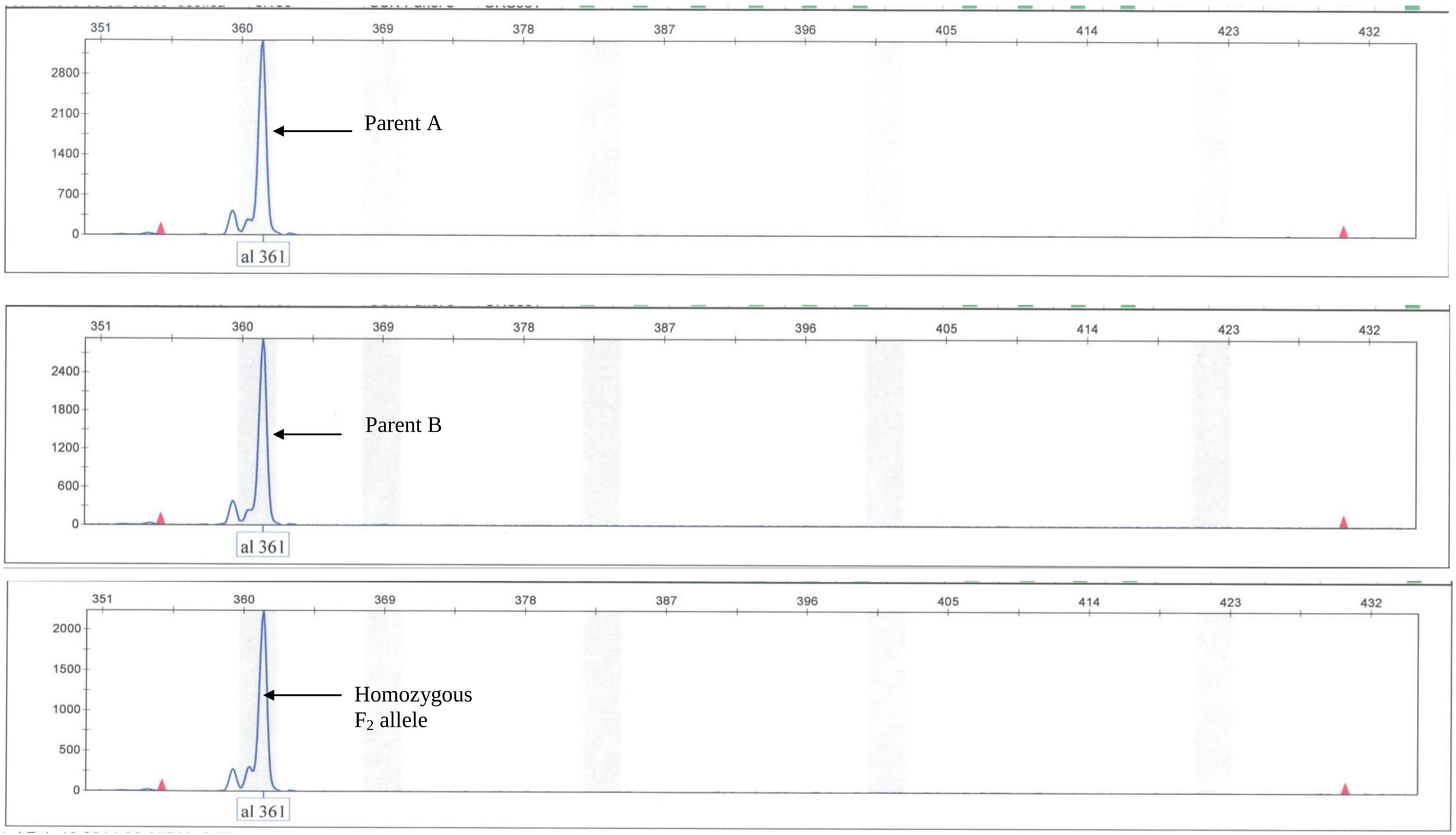


Figure 5.1 (b) Electropherogram showing the analysis of ORS687 labelled microsatellite PCR products. Peaks represent microsatellite alleles of both parental lines and homozygous F_2 individual.

5.3.2 Marker-assisted foreground selection of the F_1 populations

A total of 50 plants of the F_1 generation derived from AP901-94-3-2-1(HO) and AP991-76-1-4-1 (LO) were screened with FAD2-1F4/R1 SCAR marker to select plants possessing the high oleic acid allele. Of the total, 23 plants were positive for the high oleic acid allele (Figure 5.2 a). In a population derived from AP901-56-1-1 (HO) and ND761-3-1-1 (LO), out of 41 plants screened with FAD2-1F4/R2, only 10 showed clear bands associated with the high oleic acid allele (Figure 5.2 b). In another F_1 population derived from a cross between AP901-56-1-1 (HO) and RK74-26-1-2-1, 64 plants were screened with FAD2-1 F14/R2 SCAR marker, out of which, 18 plants amplified with clear bands corresponding with that of the high oleic acid control (Figure 5.2 c). The last F_1 population derived from a cross between AP901-95-3-4-1 (HO) and H55 -9-2-1-1 (LO) had a total of 40 plants screened with SCAR marker (FAD2-1F13/R1), of which only 10 plants were selected for backcross (Figure 5.2 d).

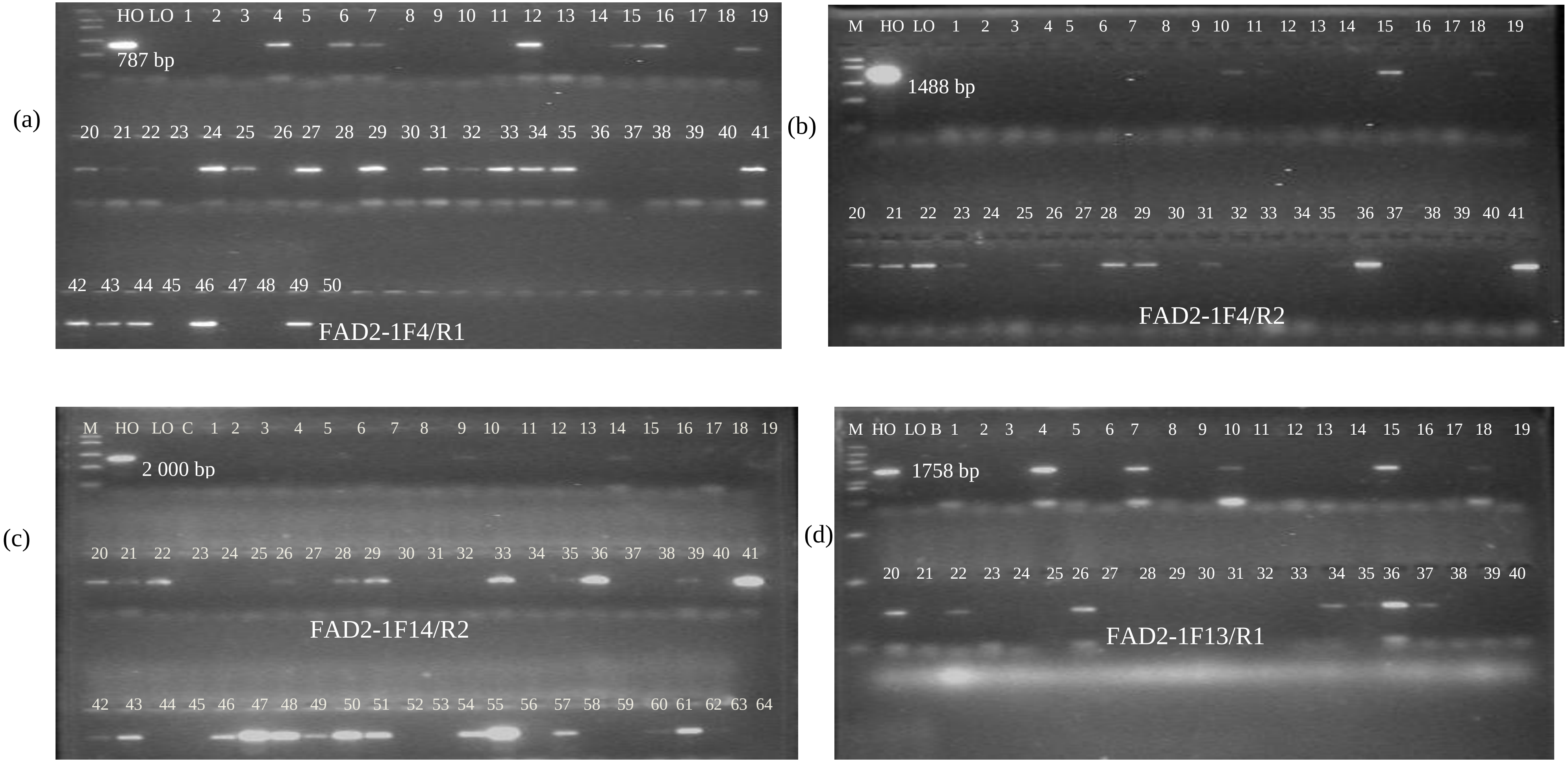


Figure 5.2 Amplified DNA of the four F_1 populations screened for the presence of high oleic acid allele using four linked foreground SCAR markers.

5.3.3 Construction of genetic linkage map for cultivated sunflower

A composite sunflower genetic linkage map of 3 275.7 cM long was constructed from 78 ORSSSR marker loci based on the consensus linkage map constructed by Tang *et al*., 2003.

These 78 markers were screened for polymorphism in the current study between the two parents and 143 plants of the F_2 population. Only 15 of these markers were polymorphic in the segregating F_2 population. Figure 5.3 illustrates the location of the SSR markers used as a guide for selecting background markers in marker-assisted introgression. The 78 SSR markers are distributed evenly over the 17 chromosomes. Due to high genetic similarities that existed between the high and low oleic acid parents, only 15 polymorphic markers could be used for background selection. These markers were located on the following linkage groups: ORS543 was mapped on linkage group 1, ORS925 on linkage group 2, ORS309 and ORS674 on linkage group 4, ORS483 and ORS230 on linkage group 6, ORS1178 on linkage group 7, ORS1265 on linkage group 9, ORS595 on linkage group 10, ORS621 on linkage group 11, ORS810 on linkage group 12, ORS630 and ORS316 on linkage group 13 and ORS885 on linkage group 16. None of these polymorphic markers was located on high oleic acid carrier chromosome 14.

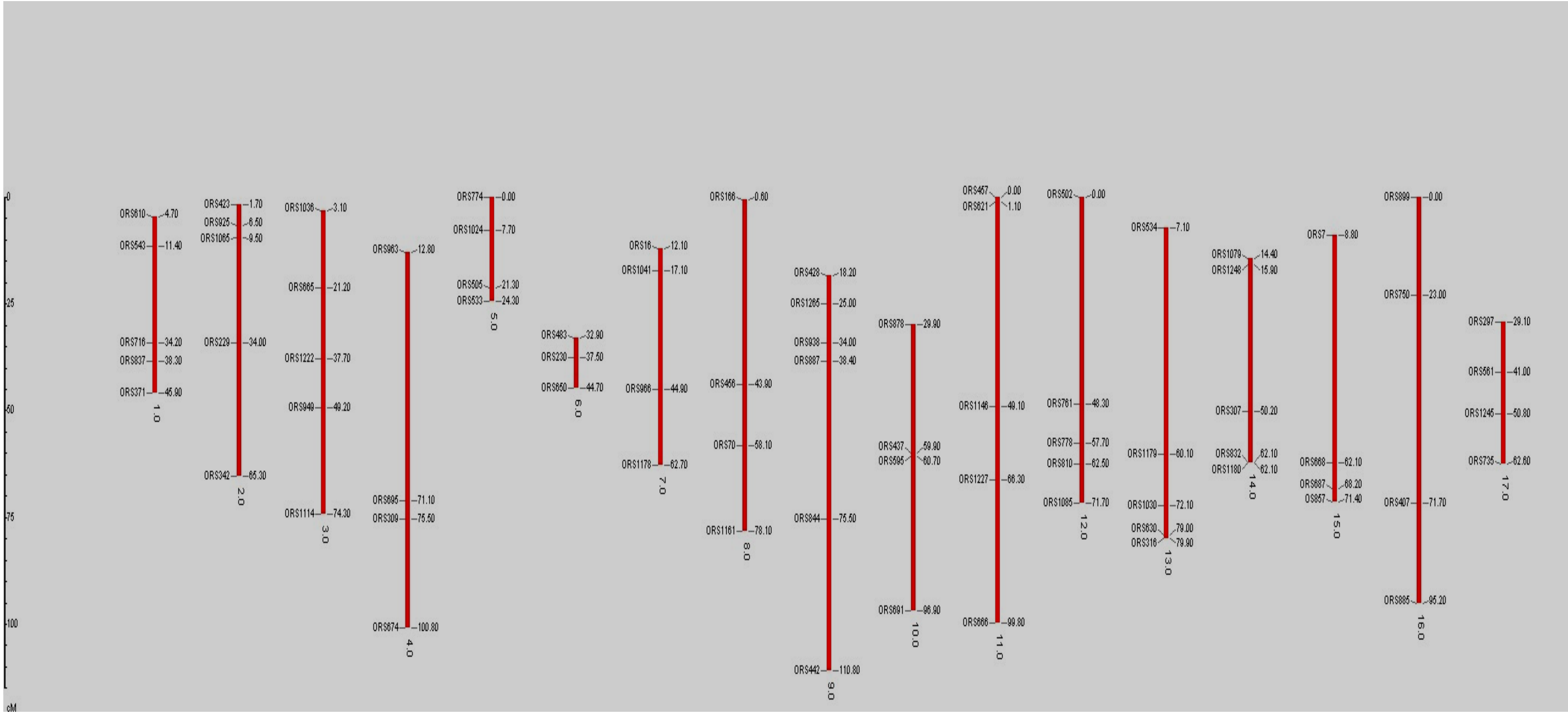


Figure 5.3. A composite genetic linkage map of sunflower showing the positions of 78 ORS SSR markers amplified by 14 PCR multiplex sets (Adapted from Tang *et al*. (2003)).

5.3.4 Marker-assisted foreground and background selection of BC_1F_1 populations

A marker-assisted backcross technique was used for the introgression of the high oleic acid gene into the background of each of the four recurrent parents using FAD2-1F4/R1 as a foreground marker to select plants of the BC_1F_1 populations possessing the high oleic acid allele. Polymorphic SSRs were used to select individual plants with approximately 75% resemblance to the recurrent parent so that they can be used as pollen donors in the second backcross (Collard *et al.*, 2005). In a backcross using ND761-3-1-1 as recurrent parent, 40 plants of the BC_1F_1 were screened with FAD2-1F4/R1 marker as foreground selection. Of the total, DNA of 18 plants were amplified and proved to be positive for the high oleic acid allele (Figure 5.4 a). For background selection, 14 SSR markers polymorphic between high and low oleic acid parents were used to genotype the 40 individual plants and identify the plant having the highest percentage of the recurrent parent genome contribution using the formula described in section 5.2.5. Table 5.3 indicates a representative example for the percentage of recurrent parent genome contribution of the 40 plants of the BC_1F_1 population obtained when genotyped with 14 background SSR markers. The minimum and maximum percentage contribution of the recurrent parent genome observed in 40 individuals of the BC_1F_1 population was between 14% and 64%. Plants 1.5, 1.28, and 1.37 having the highest recurrent parent genome contribution and positive for high oleic acid trait, were backcrossed to recurrent parents to generate BC_2F_1. In a backcross using AP991-76-1-4-1 as recurrent parent, 41 plants were screened and only 14 plants were positive for the high oleic acid allele (Figure 5.4 b). In another backcross using H55-9-2-1-1 as recurrent parent, 18 plants of the BC_1F_1 population were screened and all tested positive for the high oleic acid allele (Figure 5.4 c). In the last backcross using RK74-26-1-2-1 as recurrent parent, out of 42 plants screened for the high oleic acid trait, only seven were positive for the high oleic acid allele (Figure 5.4 d).

The results of the other three BC_1F_1 populations for background markers indicated that plants 1.22 and 1.28 had 64% and 57% respectively, of the recurrent parent genome contributions from RK74-26 as recurrent parent even though plant 1.28 did not tested positive for high oleic acid trait (Appendix A 1). In a population from H55-9-2-1-1 as recurrent parent, plants 1.7, 1.10 and 1.18 had each 50% of the recurrent parent genome contribution (Appendix A 2). A population of AP991-76-1-4-1 as recurrent parents had

64%, 57%, 57% and 57% recurrent parent genome contribution from plants 1.28, 1.36, 1.37 and 1.41 (Appendix A 3) respectively.

5.3.5 Marker-assisted foreground selection of BC_2F_1 populations

With regard to the second backcrossing between BC_1F_1 and ND761-3-1-1, 16 plants of this cross were screened for the high oleic acid allele using a SCAR marker (FAD2-1F4/R2), DNA fragments of 13 plants were successfully amplified and displayed a fragment size of 1 488 bp that corresponded with the high oleic acid control (Figure 5.5 a). Other plants screened for the high oleic acid allele were derived from a cross between H55-9-2-1-1 and $BC_1F_1$and the fragment amplified in 27 of the plants using a SCAR marker FAD2-1F4/R2 from a total of 34 plants (Figure 5.5 b). A total of 18 plants were screened from a cross between AP991-76-1-4-1 and BC_1F_1 using a marker (FAD2-1F4/R1) and eight plants displayed successful amplification with a fragment size of 787 bp (Figure 5.5 c). The result of the last cross between RK74-26-1-2-1 and BC_1F_1 indicated that DNA of 11 plants contained the desired allele using marker FAD2-1F4/R2 from a total of 19 plants of the population (Figure 5.5 d). Selection was never done with background markers due to insufficient population sizes obtained as explained in section 5.2.4.

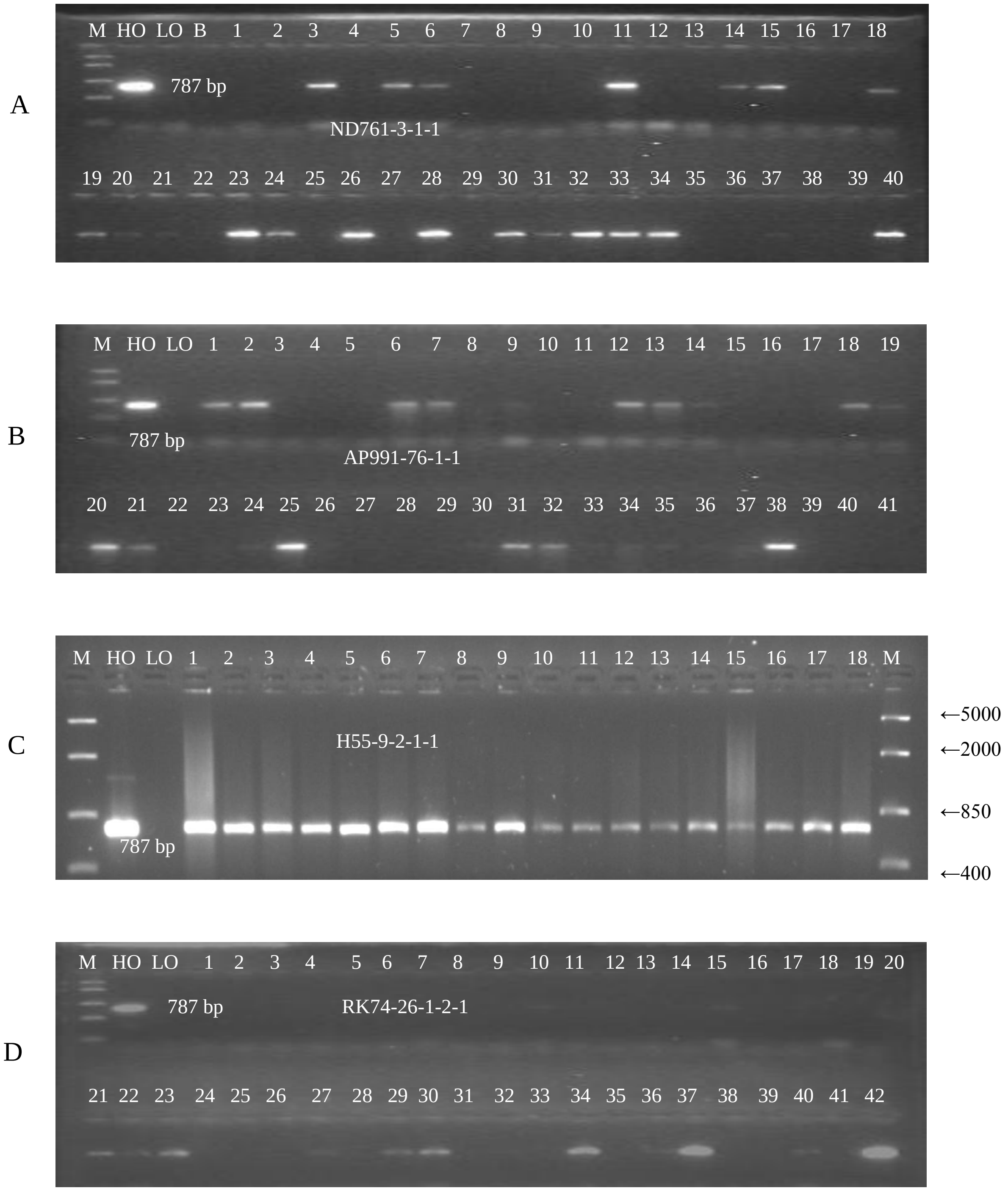


Figure 5.4 Amplified DNA of the first backcross (BC_1F_1) individuals screened for the presence of high oleic acid allele using FAD2-1 F4/R1 SCAR marker.

Table 5.3 The percentage of recurrent parent genome contribution after the first backcrossing of 40 BC_1F_1 derived from a cross between ND761-3 (LO) and AP 901-95-3-4-1 (HO) inbred lines. Fourteen polymorphic SSRs were used as background markers

BC_1F_1	**ORS543**	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% Recovery**
Parent (HO)	A/B	A/B	A/B	A/B	A/A	A/A	A/B	A/A	A/B	A/A	A/A	A/B	A/A	A/B	
Parent (LO)	B/B	B/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	A/B	B/B	
1.1	B/B	A/B	B/B	A/B	A/B	B/B	A/B	A/B	B/B	B/B	A/B	A/A	A/B	B/B	50
1.2	B/B	B/B	A/B	A/B	A/B	B/B	B/B	A/B	A/B	A/B	B/B	B/B	A/B	B/B	50
1.3	A/B	B/B	B/B	B/B	A/B	A/B	A/A	A/A	B/B	B/B	B/B	A/B	A/B	B/B	36
1.4	A/B	A/A	B/B	B/B	A/B	A/B	A/B	A/B	A/A	B/B	A/B	B/B	B/B	A/B	50
1.5	**A/B**	**B/B**	**A/B**	**A/A**	**A/B**	**B/B**	**A/B**	**A/B**	**A/A**	**A/A**	**A/B**	**A/B**	**A/B**	**A/A**	**57**
1.6	A/B	A/B	A/A	A/B	B/B	A/B	B/B	A/B	B/B	A/A	A/A	B/B	A/B	A/B	50
1.7	A/A	B/B	B/B	B/B	A/A	B/B	A/A	A/B	B/B	B/B	B/B	A/A	A/B	B/B	14
1.8	A/A	B/B	A/A	A/B	A/B	B/B	A/A	A/A	B/B	B/B	A/B	B/B	B/B	B/B	21
1.9	**B/B**	**A/B**	**A/B**	**A/B**	**A/B**	**A/A**	**A/B**	**A/B**	**A/B**	**A/A**	**A/B**	**B/B**	**A/A**	**A/B**	**64**
1.10	B/B	B/B	A/B	A/A	A/A	?/?	B/B	A/B	B/B	A/B	B/B	B/B	?/?	?/?	21
1.11	A/A	B/B	B/B	A/A	B/B	A/A	B/B	B/B	B/B	A/B	A/A	A/A	A/B	A/B	21
1.12	B/B	A/A	A/B	B/B	A/B	B/B	B/B	B/B	A/A	B/B	B/B	A/B	A/B	A/B	36
1.13	B/B	B/B	B/B	B/B	A/B	A/B	A/A	A/A	A/A	A/A	A/B	B/B	A/B	B/B	30
1.14	B/B	B/B	A/A	B/B	A/A	A/B	B/B	A/B	A/A	B/B	A/B	B/B	A/B	A/B	36
1.15	A/A	A/B	B/B	A/A	A/A	B/B	A/B	A/B	A/B	B/B	B/B	A/B	A/B	A/B	50
1.16	B/B	A/A	A/B	B/B	B/B	B/B	B/B	A/B	B/B	A/A	A/A	B/B	A/B	B/B	21
1.17	B/B	B/B	B/B	A/B	A/B	A/B	A/A	B/B	B/B	B/B	A/A	B/B	A/B	A/A	29
1.18	A/B	A/A	A/B	A/A	A/B	B/B	A/A	A/B	A/B	A/A	B/B	B/B	A/B	B/B	43
1.19	A/A	B/B	A/A	A/B	A/B	B/B	B/B	A/B	B/B	B/B	B/B	A/B	A/B	B/B	36
1.20	B/B	A/B	B/B	A/B	B/B	A/A	A/B	A/B	B/B	A/A	A/B	B/B	A/A	A/A	36
1.21	B/B	B/B	A/A	B/B	A/B	A/A	B/B	A/A	B/B	A/B	A/B	B/B	A/B	B/B	29

Table 5.3 (continued)

BC_1F_1	**ORS543**	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% Recovery**
1.22	A/B	B/B	A/A	B/B	B/B	B/B	B/B	A/B	A/B	A/A	A/B	B/B	A/B	B/B	36
1.23	A/A	A/B	B/B	A/A	A/A	B/B	A/A	A/B	A/A	B/B	A/B	A/B	B/B	B/B	29
1.24	B/B	B/B	0	B/B	A/A	A/B	A/A	A/B	B/B	B/B	A/B	A/B	B/B	A/A	36
1.25	A/B	B/B	B/B	A/A	A/B	A/B	B/B	A/B	B/B	A/A	B/B	A/A	A/B	B/B	36
1.26	B/B	A/A	B/B	B/B	A/B	B/B	A/B	A/A	A/B	B/B	A/B	A/B	A/A	A/B	36
1.27	A/B	B/B	A/B	A/A	A/A	B/B	B/B	A/B	B/B	B/B	A/B	A/B	A/A	B/B	36
1.28	**B/B**	**A/B**	**B/B**	**B/B**	**A/B**	**A/B**	**A/B**	**A/B**	**A/B**	**A/A**	**A/B**	**A/B**	**A/A**	**A/B**	**64**
1.29	A/B	B/B	A/B	A/A	A/B	A/B	B/B	A/A	B/B	A/B	A/B	B/B	A/B	A/A	50
1.30	A/A	A/B	A/A	A/B	A/A	B/B	B/B	A/B	A/A	A/A	A/B	A/A	A/B	B/B	36
1.31	B/B	B/B	B/B	B/B	A/A	B/B	A/B	A/B	?/?	A/B	A/B	B/B	B/B	B/B	29
1.32	B/B	A/A	B/B	A/B	A/B	A/A	B/B	A/B	A/A	B/B	A/B	B/B	B/B	A/A	29
1.33	A/A	A/A	?/?	B/B	A/B	A/A	A/A	A/B	B/B	B/B	A/A	A/B	A/A	A/A	21
1.34	B/B	A/B	A/A	A/A	B/B	B/B	A/A	A/A	B/B	B/B	A/A	A/A	A/B	B/B	14
1.35	B/B	B/B	B/B	B/B	A/B	B/B	A/A	A/A	A/B	A/A	B/B	A/B	A/B	B/B	29
1.36	**B/B**	**A/A**	**A/B**	**A/B**	**A/B**	**A/B**	**B/B**	**A/B**	**A/A**	**A/B**	**A/B**	**A/A**	**A/B**	**B/B**	**57**
1.37	**A/B**	**B/B**	**A/A**	**A/B**	**A/A**	**A/B**	**A/B**	**A/B**	**A/A**	**A/A**	**A/B**	**A/B**	**A/B**	**A/A**	**57**
1.38	B/B	A/B	A/B	A/B	B/B	B/B	A/B	B/B	B/B	A/A	B/B	A/B	A/A	A/A	36
1.39	A/B	B/B	B/B	A/B	A/B	B/B	B/B	B/B	A/B	B/B	A/B	B/B	A/B	A/B	50
1.40	B/B	B/B	B/B	B/B	A/B	A/B	B/B	A/B	B/B	B/B	A/B	B/B	A/B	A/B	43

A= homozygous allele for high oleic acid parent

B= homozygous allele for low oleic acid parent

A/B= heterozygous alleles for both parents

?/? = data missing

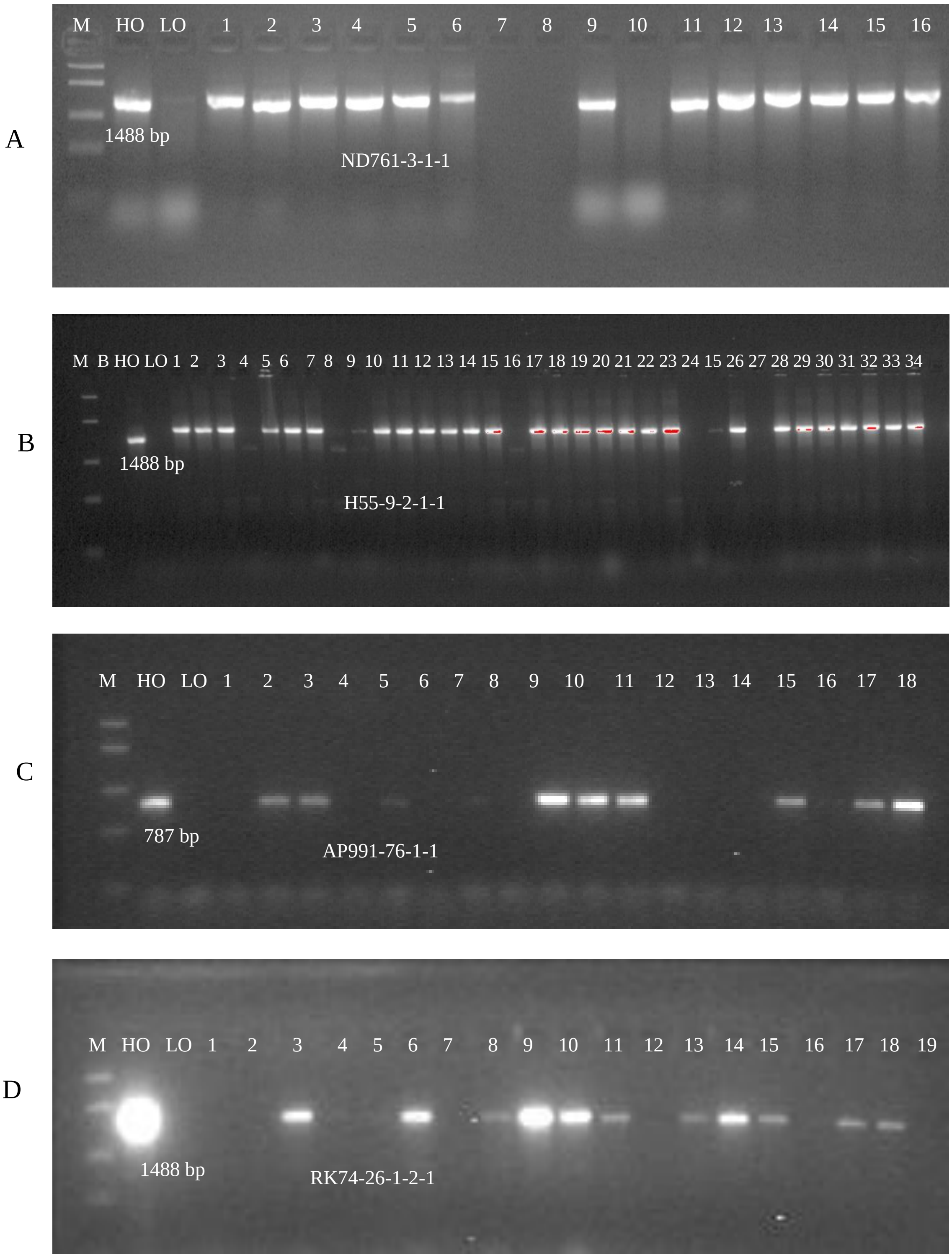


Figure 5.5 Amplified DNA of the second backcross (BC_2F_1) individual screened for the presence of the high oleic acid allele using SCAR markers FAD2-1F4/R1 and FAD2-1F4R2.

5.4 DISCUSSION

Marker-assisted breeding has provided an opportunity as a tool for discovering and tagging alleles and genes, and this tool can enhance the efficiency of breeding programmes through selection of target traits by indirectly using molecular markers that are closely linked to the underlying genes. This represents a substantial saving in time compared to conventional backcross breeding (Xu & Crouch, 2008).

5.4.1 Multiplex genotyping assay to identify background SSR markers in a F_2 population

The importance of identifying suitable SSR markers for background selection is to help accelerate the return to recurrent parent genome outside the target gene so as to reduce the length of the intact chromosomal segment of donor type dragged around the target gene on the carrier chromosome (reduce linkage drag) (Semagn *et al.*, 2006). In this study, 78 SSR markers were screened for polymorphism between 143 plants of the F_2 segregating population derived from a cross between AP901-56-1-1(HO) and H55-9-2-1-1(LO) in an effort to identify the most suitable background markers that can be used in marker-assisted background selection. Of the total, 15 SSR markers were found to be polymorphic between high and low oleic acid parents. In order to improve the pace and the precision of backcross breeding, an adequate number of uniformly-spaced polymorphic markers are critical to accurately recover high percentage of recurrent parent genome (Semagn *et al.*, 2006, Xu & Crouch, 2008). For this reason, 15 markers are however not sufficient to recover a high percentage of the recurrent parent genome. Nine SSR markers were monomorphic between high and low oleic acid parents (share the same allele). Monomorphic markers are known to have only one allele and are genotypically not beneficial since they do not possess the capacity to distinguish between two contrasting parents at a particular locus. Due to this limitation, monomorphic markers could not be used further for marker-assisted background selection. A total of 35 markers had produced null alleles, and these were confirmed through repeated PCR amplification in order to avoid the possibility of failed PCR reactions. Null alleles are non-amplified alleles that can sometimes segregate with another allele resulting in an apparent homozygote. Null alleles sometimes occur as a result of the mutation in the flanking regions preventing one or both of the primers from binding (Masi *et al.,* 2003, Smulders *et al.,* 1997). Nineteen markers had more than one allele but were not polymorphic between the two control parents (non-informative markers).

With regard to the selection of suitable SSR markers which can be used as candidates in marker-assisted background selection, Table 5.2 summarised the characteristics of 43 of these SSR markers in the evaluation of the genetic diversity that exist between 143 plants of the F_2 population. From 78 SSR markers screened, only 15 (19%) were polymorphic and the total average allele number was 3.1, which is much closer to the mean number of alleles per locus reported in other studies for inbred lines and hybrids (Tang & Knapp, 2003, Yu *et al.,* 2002). The average allele numbers for monomorphic and non-informative markers were 1.0 and 2.5 respectively. This is much lower than the average allele number for polymorphic markers. In a study carried out by Tang and Knapp (2003) 3.5 alleles was reported on average per locus when 122 microsatellite markers were used to genotype 9 elite confectionery and oilseed sunflower inbred lines (Darvishzadeh *et al.,* 2010). The high number of alleles for a marker indicates the existence of larger genetic diversity among the genotype under study (Agram & Tuinstra, 2003). In this study only four polymorphic markers (ORS674, ORS309 ORS810 and ORS1178) had 4, 4, 5 and 6 alleles, respectively. Contrary to the number of alleles influencing the genetic diversity, Priolli *et al.* (2002) reported that a higher mutation rate in the SSR loci may be the reason for observing larger number of alleles in a particular marker.

The average mean allele frequency for monomorphic, polymorphic and non-informative markers was 1.00, 0.52 and 0.84 respectively. Priolli *et al.* (2002) reported that markers with the lower frequency of the predominant allele have more differentiation ability than other markers. As an example, marker ORS309 had an allele frequency of 0.424, which is much lower than for example the allele frequency of ORS844 (1.00). The average mean of PIC value observed between monomorphic, polymorphic and non-informative markers was 0.00, 0.40 and 0.17 respectively. According to Junjian *et al*. (2002) PIC is regarded as one of the important features of molecular markers that can be used to evaluate the differentiation ability of markers. In this study, the polymorphic markers had the highest PIC value that is closest to one reported by Darvishzadeh *et al*. (2010) indicating the capacity of being effectively used in a genetic diversity study of sunflower.

5.4.2 Selection of plants of the F_1 populations for the high oleic acid allele using SCAR markers

A total of four F_1 sunflower breeding populations were developed through incorporation of a high oleic allele from a donor to a recurrent parent by means of crosses. Selection for the trait was done by evaluating individual plants of each of the four F_1 populations by screening with SCAR markers that co-segregate with the high oleic acid allele. The results indicated a successful introgression of the allele thereby amplifying specific bands corresponding with the high oleic acid parents respectively the fragment sizes for each SCAR marker is an indication of a true allele amplified during PCR assay.

5.4.3 Marker-assisted backcross selection of a BC_1F_1 population

In a foreground selection of individual plants of the F_1BC_1 population possessing the high oleic acid allele, an insufficient number of individual plants of the BC_1F_1 population were successfully selected as having the high oleic acid allele using SCARs as foreground markers due to self-pollination of recurrent parents during crosses. For background selection, 14 SSR markers identified 5 plants of the population as having between 57% and 64% of the recurrent parent genome contribution (Table 5.3) as per the analysis. These results indicate a variation from the theoretically expected 75% contribution from the recurrent parent genome to the BC_1F_1 plants and all 40 plants of the BC_1F_1 population had a recurrent parent genome contribution that was less than the expected 75% (Collard *et al.*, 2005). Table 5.4 summarizes an example of the expected percentage of recurrent parent genome contribution in each generation after backcrossing when marker-assisted breeding is applied. It is apparent that using a conventional breeding method requires up to 6 backcrosses to fully recover the recurrent parent genome while marker-assisted breeding needs only three backcrosses to achieve around 93.8% of the recurrent parent genome. Similar results were found for plants of the other three BC_1F_1 populations between AP991-76-1-4-1, RK74-26-1-2-1 and H55-9-2-1-1 as recurrent parents (Appendix A). This suggests that a large number of individuals in a population is required in order to efficiently recover the maximum recurrent parent genome contribution and because background selection was not performed with sufficient and tightly linked or flanking markers to the target gene, more and high density markers are essential for recovery of recurrent parent genome characteristics. In a study conducted by Sundaram *et al.* (2008), BC_1F_1 plants that were less than the expected 75% were reported in an introgression of bacterial blight resistance genes on an elite indica rice variety. In the study, they hypothesised a pull in *xa21, xa13* and *xa5* genes which favoured the inheritance of additional unlinked loci from the donor genome in BC_1F_1 plants. The small number of background markers limits the genome coverage and the ability to discover the maximum recovery of the recurrent parent genome contribution in backcrossing cycles.

Table 5.4 Percentage of the recurrent parent genome after backcrossing (Collard *et al.*, 2005).

Generation	**Expected % contribution of RPG**
BC_1	75.0
BC_2	87.5
BC_3	93.8
BC_4	96.9
BC_5	98.4
BC_6	99.2

RPG= Recurrent parent genome

In conclusion, the multiplex assay could not be used to its full potential because of low polymorphism between the parents used in the study. However, it was found to be a very useful technique for genotyping of an F_2 segregating population and BC_1F_1 backcrossing population in an effort to identify candidate markers that could be used in a marker-assisted background selection. This would result in the identification of individual plants that have the maximum recurrent parent genome contribution. According to Collard *et al.* (2005), selection of these plants would speed up the backcrossing process by a least two generations (Table 5.4).

CHAPTER 6

GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS

The main aim of this study was to improve the quality of oil on local sunflower germplasm with special reference to oleic acid through employing marker-assisted breeding to facilitate and speed up the recovery of the high oleic acid allele into the background of the recurrent parent genome. This study demonstrated the importance of using a SCAR technique as a molecular marker system in evaluating different sunflower breeding genotypes for the presence of the high oleic acid allele. Information was gained on which SCAR marker co-segregated or was closely linked with the high oleic acid allele and suitable for use as foreground markers in the marker-assisted introgression process. Furthermore, information regarding the genetic state of the gene influencing the high oleic acid trait was reported. These results are not unique and were also reported by previous studies as discussed in section 4.4. A composite genetic map of 3 275.7 cM was constructed from 78 SSR markers as a guide for selecting background markers in marker-assisted introgression. These markers were successfully optimised and validated on 143 plants of an F_2 segregating population. Only fourteen of these markers revealed the capacity to be used in a marker-assisted background selection due to their high polymorphic information content and allele frequency.

6.1 Characterisation of sunflower breeding genotypes

The first critical step in a successful breeding programme of a crop species depends largely on the presence of genetic diversity in the germplasm and knowledge about the characteristics of the genotypes being investigated. Therefore, the methods that evaluate and identify the genotypes more precisely are preferred by plant breeders (Moghaddam *et al.,* 2009). In this study, it was demonstrated that the high oleic acid sunflower genotypes differed significantly (R^2 = -99.16%) from the wild types (recurrent parents) with regard to the level of their oleic and linoleic acid content. This was demonstrated through phenotypic analysis of their fatty acid profiles using gas chromatography. However, fatty acid analysis using gas chromatography coupled with a FID always leads to the destruction of compounds being investigated, limiting further analysis of the same samples. A non-destructive method is however, recommended for fatty acid analysis so that the same seed samples can be used further in a marker-assisted selection. One of the most practical applications of the DNA based markers in a plant breeding programme is the ability to select phenotypic traits using markers tightly linked to the gene controlling the trait. In this study, molecular SCAR markers such as FAD2-1 were used for such application and four of these markers were reported as tightly linked to the high oleic acid allele in high oleic acid genotypes. The high oleic acid trait in cultivated sunflower is controlled by a chemically induced, incomplete dominant mutation (*Ol*) gene and is negatively correlated with the expression of linoleic acid influenced by the *FAD2-1* gene. *FAD2-1* co-segregates with the *Ol* gene and is duplicated in high oleic acid (mutant) strains (Lacombe & Bervillé, 2001, Schuppert *et al.,* 2006). Schuppert *et al.* (2006) have developed and applied these SCAR markers in an effort to diagnose for the presence or absence of the *Ol* gene mutation (presence or absence of tandem *FAD2-1* repeats) between wild and mutant sunflower genotypes. SCAR markers used in this study have successfully characterised the sunflower breeding genotypes by tagging the allele associated with the high oleic acid trait.

6.2 Inheritance of the high oleic acid gene

With regards to the status of the gene controlling high oleic acid trait in cultivated sunflower, four putatively linked FAD2-1 SCAR markers were successfully validated through the evaluation of 143 plants of the F_2 segregating population derived from a cross between AP901-95-3-4-1 (HO) and H55-9-2-1-1 (LO). Information around the status of the gene controlling high oleic acid was achieved, which confirmed that the high oleic acid trait is governed by a single incomplete dominant gene and because the SCAR markers are inherited in a dominant state, the F_2 plants should segregate in a 3:1 (high oleic acid : low oleic acid) ratio. Due to financial implications, fatty acid analysis could not be determined to completely validate the SCAR markers at phenotypic level. One of the limiting factors however inherent to the SCAR markers was the inability to distinguish heterozygous from homozygous high oleic acid plants at F_2 stage due to their dominant state. This would limit their application in marker-assisted selection or would require a progeny testing prior to use. Several genetic studies relating to the number of genes involved in the high oleic acid trait have already been conducted with contradicting findings (Lacombe & Bervillé, 2001). In a study conducted by Urie (1985) the existence of reversal in dominance and modifying genes were reported. Later Miller *et al.* (1987) reported a second recessive locus (*Ml*) that is necessary in the expression of oleic acid.

6.3 Marker-assisted backcross breeding

Marker-assisted breeding has been greatly used by breeders to tag and map agriculturally important genes through employing molecular markers to precisely transfer the genomic regions of interest and accelerate the recovery of the recurrent parent genome (Ibitoye & Akin-Idowe, 2010). In this study, the four SCAR markers (FAD2-1F4/R1, FAD2-1F4/R2, FAD2-1F13/R1, and FAD2-1F14/R2) linked to the high oleic acid trait were successfully employed as foreground markers for selecting plants of the F_1, BC_1F_1 and BC_2F_1 populations for the high oleic acid allele, respectively. Different fragment sizes of bands were obtained at each backcross stage that corresponded directly to the fragment sizes of their respective high oleic acid donor parents. However, the fragment sizes of the four SCAR markers were slightly different from the one obtained by Schuppert *et al.* (2006). This could be due to the differences that exist between the sunflower genetic backgrounds utilised. For a marker to be useful in breeding programmes, they should reveal polymorphism in different populations derived from a wide range of different parental populations (Langridge *et al.*, 2001).

Information about the genetic diversity and relationships among the breeding lines and varieties is not only limited to germplasm conservation and inbred line identification, but also more significant for selection of parental lines for hybrid production and can assist in selecting breeding progeny carrying desirable alleles. Microsatellite or SSRs constitute the current marker system of choice for increasing the efficiency of conventional breeding (Darvishzadeh *et al.*, 2010, Ibitoye & Akin-Idowu, 2010). In this study, 78 SSR markers were successfully optimised on parents of the F_2 segregating population in an effort to select the most suitable markers to be used in a marker-assisted background selection process. Fifteen of these markers were successfully selected based on their characteristics of high polymorphic information content, low frequency of the predominant allele and the number of alleles they possessed, as outlined in section 5.4.2. Molecular markers intended for marker-assisted background selection are selected based on these attributes. In addition, their association with the candidate gene and distribution throughout the sunflower genome are also important features of a suitable marker selected for background selection. Fourteen of the fifteen markers were applied successfully in the background selection. It was found that five plants (Table 5.6) had a greater number of heterozygous alleles (A/B) for the recurrent parent than the rest of the plants in the BC_1F_1 population, indicating a recovery of 57%, 64%, 64%, 57% and 57% respectively of the recurrent parent genome. These numbers are not close to the expected values presented in Table 5.4 with respect to the recovery of the recurrent parent genome at BC_1F_1. Similar results were also observed with three other populations at BC_1F_1 backcrosses (Appendix A). Furthermore, none of the background markers screened was tightly linked to the high oleic acid loci (linkage group 14) (Figure 5.4). It is however clear from this study that background selection with limited and unlinked number of polymorphic SSR markers in conjunction with two backcrosses is not sufficient to recover the yield and quality characteristics of the recurrent parent. Unfortunately the breeding project for the incorporation of high oleic acid gene into the background of the productive sunflower recurrent parent have not been finished due to a limited number of plants in the breeding population obtained at each backcross cycle and insufficient number of polymorphic background markers obtained during the optimization step. Furthermore, information around the fatty acid content at each breeding cycle was not obtained to complement the genotypic data as a result of financial constraints.

In conclusion, this study provided preliminary data of the potential of using SCARs and microsatellite assays as foreground and background markers in an attempt to incorporate a trait of high oleic acid into the background of well adapted and high yielding sunflower varieties that have been grown by farmers over the years on account of their unique characteristics. The results obtained in this study showed that molecular markers could be used to facilitate and speed up the recovery of the recurrent parent genome that is much faster than the conventional breeding programme.

6.4 Recommendations and prospect for future research

Information gained in this study could be used to develop a better understanding around the principles of marker-assisted breeding versus conventional breeding in sunflower and which breeding tools are more efficient in characterising the breeding genotypes. It is highly recommended that pure breeding genotypes (Inbred lines) are used in MAB as this will help to save time and cost needed to purify them. In MAB of the sunflower high oleic acid trait, phenotypic data (fatty acid profiles) will always form an important part of the results in order to know whether the high oleic acid gene is maternally influenced. This will however, be tested at each breeding population developed. Molecular markers tightly linked to the gene of interest are highly recommended.

REFERENCES

AGRAM, H.A. & TUINSTRA, M.R. 2003 Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African Journal of Biotechnology*, 2:334-340.

ALL, J., NAGANO, M., WOO, S.H., CAMPBELL, C. & ADACHI, T. 1999. Development of the SCAR markers linked to the S^{th} gene in buck-wheat. *Fagopyrum*, 16:19-22.

ASHISHI, M.A., KLAPPER, P.E. & COOPER, R.J. 2003. Detection of human cytomegalovirus, human herpesvirus type 6 and human herpesvirus type 7 in urine specimens by multiplex PCR. *Journal of Infection*, 47:59-64.

BERGLUND, D.R. 2007. Sunflower production. Extension service. North Dakota, state University. [Online]. Available at: <http://sunflowernsa.com/upload/resources/540/sustainerbility_-use.html.> Accessed: 04/04/2010.

BERRY, S.T., LEON, A.J., HANFREY, C.C., CHALLIS, P., BURKHOLZ, A., BARNES, S.R., RUFENER, G.K., LEE, M. & CALIGARI, P.D.S. 1995. Molecular marker analysis of *Helianthus annuus L*. 2. Construction of an RFLP linkage map for cultivated sunflower. *Theoretical and Applied Genetics*, 91:195-199.

BI-HAO, C., JIAN-JUN, L., YONG, W. & GUO-JU, C. 2009. Inheritance and identification of SCAR marker linked to bacterial wilt-resistance in eggplant. *African Journal of Biotechnology*, 8(20):5201-5207.

BLAIR, M.W., HEDETALE, V. & McCOUCH, S.R. 2002. Fluorescent-labeled microsatellite panels useful for detecting allelic diversity in cultivated rice (*Oryza sativa L.*). *Theoretical and Applied Genetics*, 105:449-457.

BLAIR, M.W. & McCOUCH, V.H.S.R. 2002. Fluorescent-labelled microsatellite panels useful for detecting allelic diversity in cultivated rice *(Oryza sativa* L.*). Theoretical and Applied Genetics*, 105:449-457.

BLAKESLEE, A.F., BELLING, J., FARNHAM, M.E. & BERGNER, A.D. 1922. A haploid mutant in the jimson weed *datura stramonium*. *Science,* 55:646-647.

BURR, B. & BURR, F.A. 1991. Recombinant inbred lines for molecular mapping in maize. *Theoretical and Applied Genetics,* 85:55-60.

CHEUNG, W.Y., CHAMPAGNE, G., HUBERT, N. & LANDRY, B.S. 1997. Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. *Theoretical and Applied Genetics,* 94:569-582.

COLLARD, B.C.Y., JAHUFER, M.Z.Z., BROUWER, J.B. & PANG, E.C.K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica,* 142:169-196.

DARVISHZADEH, R., AZIZI, M., HATAMI-MALEKI, H., BERNOUSI, I., MANDOULAKANI, B.A., JAFARI, M. & SARRAFI, A. 2010. Molecular characterization and similarity relationship among sunflower (*Helianthus annuus* L.) inbred lines using some mapped simple sequence repeats. *African Journal of Biotechnology,* 9:7280-7288.

DEHMER, K.J. & FRIEDT, W. 1998. Development of molecular markers for high oleic acid content in sunflower (*Helianthus annuus* L.). *Industrial Crops and Products,* 7:311-315.

DREDGE, R.D. 2010. Crop estimate committee, Private Bag x 246, Pretoria, South Africa. [Online]. Available at: http://www.sagis.org.za>. Accessed: 13/06/2010.

FERNÀNDEZ-MARTINEZ, J., JIMENEZ, A., DOMINGUEZ, J., GARCIA, J.M., GARCÉS, R. & MANCHA, M. 1989. Genetic analysis of the high oleic acid content in cultivated sunflower (*Helianthus annuus* L.). *Euphytica,* 41:39-51.

FICK, G.N. 1984. Inheritance of high oleic acid in the seed oil of sunflower. *Proceedings of Sunflower Research Workshop,* Bismark.ND, 9-10.

FICK, G.N., ZIMMER, D.E. & KINMAN, M.L. 1974. Registration of six sunflower parental lines. *Crop Science,* 14:9-12.

FRANCIA, E., TACCONI, G., CROSATTI, C., BARABASCHI, D., BULGARELLI, D., DALL'AGLIO, E. & VALE, G. 2005. Marker-assisted selection in crop plants. *Plant Cell Tissue Organ cult,* 82:317-342.

FRISCH, M., BOHN, M. & MELCHINGER, A.E. 1999. Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Science,* 39:1295-1301.

FRITZ, A., ROZOWSKI, M., WALKER, C. & WESTERFIELD, M. 1996. Identification of selected gamma-ray induced deficiencies in zebrafish using multiplex chain reaction. *Genetics,* 144:1735-1745.

FUTEHALLY, S. & KNOWLES, P.F. 1981. Inheritance of very high levels of linoleic acid in an introduction of safflower (*Carthamus tinctorius L.*) from Portugal. In: KNOWLES, P.F. (ed) *Proceedings of the first international Safflower Conference,* Davis, CA, USA.

GARCES, R. & MANCHA, M. 1991. *In vitro* oleate desaturase in developing sunflower seeds. *Phytochemistry,* 30:2127-2130.

GENTZBITTEL, L., MOUZEYAR, S., BADAOUI, S., MESTRIES, E., VEAR, F., DE LABROUHE, D.T. & NICOLAS, P. 1998. Cloning of molecular markers in disease resistance in sunflower, *Helianthus annuus* L. *Theoretical and Applied Genetics,* 96:519-525.

GENTZBITTEL, L., VEAR, F., ZANG, Y.-X., BERVILLÉ, A. & NICOLAS, P. 1994. Development of a consensus linkage map of cultivated sunflower (*Helianthus annuus* L.). *Theoretical and Applied Genetics*, 90:1079-1086.

GETHI, J.G., LABATE, J.A., LAMKEY, K.R., SMITH, M.E. & KRESOVICH, S. 2002. SSR variation in important U.S. maize inbred lines. *Crop Science*, 42:951-957.

GRAPHICAL GENOTYPE VERSION 2.0. [Online]. Available at: <www.plantbreeding.wur.nl/uk/software.ggt.html. >.Accessed: 16/05/2011.

HAMDAN, Y.A.S., VELASCO, L. & PĖREZ-VICH, B. 2008. Development of SCAR markers linked to male sterility and very high linoleic acid content in safflower. *Molecular Breeding*, 22:385-393.

HENEGARIU, O., HEEREMA, N.A., DLOUHY, S.R., VANCE, G.H. & VOGT, P.H. 1997. Multiplex PCR: Critical parameters and step-by-step protocol. *Bio-Techniques*, 23:504-511.

HERNANDEZ, P., DE LA ROSA, R., RALLO, L., DORADO, G. & MARTINA, A. 2001. Development of SCAR markers in olive (*Olea europaea*) by direct sequencing of RAPD product: application in olive germplasm evaluation and mapping. *Theoretical and Applied Genetics*, 103:788-791.

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

HOPKINS, M.S., CASA, A.M., WANG, T., MITCHELL, S.E., DEAN, R.E., KOCHERT, G.D. & KRESOVICH, S. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Science*, 39:1243-1247.

HOSPITAL, F. 2002. A program for the numerical optimization of population sizes in marker-assisted backcross programs. *Heredity,* 93:383-384.

HOSPITAL, F. 2005. Selection in backcross programmes. *Philosophical Transactions of the Royal Society,* 360:1503-1511.

HU, J., CHEN, J., BERVILLÉ, A. & VICK, B.A. 2004. High potential of TRAP markers in sunflower genome mapping. In: SEILER, G.J. (ed.) *Proceedings of the 16th International Sunflower Conference*. North Dakota: 665-671.

HU, X., SULLIVAN-GILBERT, M., GUPTA, M. & THOMPSON, S.A. 2006. Mapping of the loci controlling oleic and linolenic acid contents and development of *fad2* and *fad3* allele-specific markers in canola (*Brassica napus L.*). *Theoretical and Applied Genetics,* 113:497-507.

HVARLEVA, T., BAKALOVA, A., CHEPINSKI, I., HRISTOVA-CHERBADJI, M., HRISTOV, M. & ATANASOV, A. 2007 Characterization of Bulgarian sunflower cultivars and inbred lines with microsatellite markers. *Biotechnology and Biotechnology Equipment*, 21:408-412.

IBITOYE, D.O. & AKIN-IDOWU, P.E. 2010. Marker-assisted selection (MAS): A fast track to increase genetic gain in horticultural crop breeding. *African Journal of Biotechnology,* 9(52):8889-8895.

JAUHAR, P.P. 2006. Modern biotechnology and the future of plant breeding. *Science in Action.*

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

JUNJIAN, N., COLOWIT, P.M. & MACKILL, D. 2002. Evaluation of genetic diversity in rice subspecies by microsatellite markers. *Crop Science*, 42:601-607.

KAEPPLER, S.M., PHILIPS, R.L. & KIM, T.S. 1993. Use of near-isogenic lines derived by backcrossing or selfing to map quantitative traits. *Theoretical and Applied Genetics,* 87:233-237.

KALIA, K.R., RAI, M.K., KALIA, S., SINGH, R. & DHAWAN, A.K. 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica,* 177:309-334.

KAUR, B., WADHWANI, C., CHARAYA, P. & MALIK, C.P. 2009. Markers used in diversity analysis. In: *Crop breeding and biotechnology,* India: Aavishka: 65-102.

KEURENTJIES, J.B., BENTSINK, L., ALONSO-BLANCO, C., HANHART, C.J., BLANKESTIJN-DE VRIES, H., EFFGEN, S., VREUGDENHIL, D. & KOORNNEEF, M. 2006. Development of a Near-Isogenic line population of *Arabidopsis thaliana* and comparison of mapping power with a Recombinant Inbred Line Population. *Genetics,* 175:891-905.

KINMAN, M.L. 1970. New development in USADA and State experimental station sunflower breeding programmes. *Proceedings of 4th International sunflower conference,* 181-183.

KNAPP, S.J., BERRY, S. & RIESEBERG, L.H. 2001. Genetic mapping in sunflowers. In: PHILLIPS, R.L. & VASI, I.K. (ed) *DNA markers in plants. Kluwer,* Dordrecht, the Netherlands, 379-403.

KÖHLER, H. & FRIEDT, W. 1999. Genetic variability as identified by AP-PCR and reaction to midstem infection of *Sclerotinia sclerotiorum* in interspecific sunflower (*Helianthus annuus* L.) hybrid progenies. *Crop Science,* 39:1456-1463.

KORZUN, V. 2002. Molecular markers and their applications in cereal breeding. *Journal of Cell Molecular Biology*, 7:811-820.

KRIZMANIC, M., MIJIC, A., LIOVIC, I., BILANDZIC, M. & DUVNJAK, T. 2006. Sunflower breeding at the agricultural institute Osijek. *Helia*, 29:153-158.

KUSTERER, B., ROZYNEK, B., BRAHM, L., PRÜFE, M., TZIGOS, S., HORN, R. & FRIEDT, W. 2004. Construction of genetic map and localization of major traits in sunflower (*Helianthus annuus* L.). *Helia,* 40:15-24.

LACOMBE, S. & BERVILLÉ, A. 2001. A dominant mutation for high oleic acid content in sunflower (*Helianthus annuus* L.) seed oil is genetically linked to a single oleate-desaturase RFLP locus. *Molecular Breeding*, 8:129-137.

LACOMBE, S., KAAN, F., LEGER, S. & BERVILLĚ, A. 2001. An *oleate desaturase* and a suppressor loci direct high oleic acid content of sunflower (*Helianthus annuus* L.) oil in the pervenets mutant. *C.R Acod. Science (Paris),* 324:1-7.

LANGAR, K., LORIEUX, M., DESMARAIS, E., GRIVEAU, Y., GENTZBITTEL, L. & BERVILLĚ, A. 2003. Combined mapping of DALP and AFLP markers in cultivated sunflower using F9 recombinant inbred lines. *Theoretical and Applied Genetics*, 106:1068-1074.

LANGRIDGE, P., LAGUDAH, E., HOLTON, T., APPELS, R., SHARP, P. & CHALMERS, K. 2001. Trends in genetic and genome analysis in wheat. *Australian Journal of Agriculture*, 52:1043-1077.

LECLERCQ, P. 1969. Une sterilite male cytoplasmique chez le tournesol. *Annals Amelior des Plantes,* 19(2):99-106.

LEFEBVRE, V., PALLOIX, A., CARANATA, C. & POCHARD, E. 1995. Construction of an intraspecific linkage map of pepper using molecular markers and double-haploid progenies. *Genetics,* 38:112-121.

MASI, P., ZEULI, P.L.S. & DONINI, P. 2003. Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.). *Molecular Breeding,* 11:303-313.

MCGREGOR, C.E., LAMBERT, C.A., GREYLING, M.M., LOUW, J.H. & WARNICH, L. 2000. A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. *Euphytica,* 113:135-144.

MEHDIPOUR, P., HOSSEINI-ASI, S., SAVABI-E, A., HABIBI, H., ALVANDI, E. & ATRI, M. 2006. Low frequency of 185 deIAG founder mutation of *BRCA1* gene in Iranian breast cancer patients. *Journal of Cancer Molecules,* 2:123-127.

MIENIE, C.M.S., FOURIE, H., SMIT, M.A., VAN STADEN, J. & BOTHA, F.C. 2002. Identification of AFLP markers in soybean linked to resistance to *Meloidogyne javanica* and conversion to sequence characterized amplified regions (SCARs). *Plant Growth Regulation,* 37:157-166.

MILLER, J.F., ZIMMERMAN, D.C. & VICK, B.A. 1987. Genetic control of high oleic acid content in sunflower oil. *Crop Science,* 27:923-926.

MITCHELLE, S.E., KRESOVICH, S., JESTER, C.A, HERNANDEZ, C.J. & SZEWC-McFADDEN, A.K. 1997. Application of Multiplex PCR fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. *Crop Science,* 37:617-624.

MOGHADDAM, M., MOHAMMADI, S.A., MOHEBALIPOUR, N., TOORCHI, M., AHARIZAD, S. & JAVIDFAR, F. 2009. Assessment of genetic diversity in rapeseed cultivars as revealed by RAPD and microsatellite markers. *African Journal of Biotechnology,* 8(14): 3160-3167.

MOHAN, M., NAIR, S., BHAGWAT, A., KRISHNA, T.G. & YANO, M. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding,* 3:87-109.

MORGANTE, M. & OLIVIERI, A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *Journal of Plant Science,* 3:175-182.

MULLIS, K.B. & FALOONA, F.A. 1987. Specific synthesis of DNA *in vitro* via polymerase catalysed chain reaction. *Methods in Enzymology,* 155:335-339.

MUSIAL, K. & PRZYWARA, L. 1998. Influence of irradiated pollen on embryo and endosperm development in kiwifruit. *Annals of Botany,* 82:747-756.

NARVEL, J.M., CHU, W.C., FEHR, W.R., CREGAN, P.B. & SHOEMAKER, R.C. 2000a. Development of multiple sets of simple sequence repeat DNA marker recovering the soybean genome. *Molecular Breeding,* 6:175-183.

NARVEL, J.M., FEHR, W.R., CHU, W.C., GRANT, D. & SHOEMAKER, R.C. 2000b. Simple sequence repeat diversity among soybean plant introductions and elite genotypes. *Crop Science,* 40:1452-1458.

OKULEY, J., LIGHTNER, J., FELDMANN, K., YADAV, N., LARK, E. & BROWSE, J. 1994. *Arabidopsis FAD2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant cell,* 6:147-158.

PACUREANU-JOITA, M., STANCIU, D., PETCU, E., RARANCIUC, S. & SOREGA, I. 2005. Sunflower genotypes with high oleic acid content. *Romanian Agricultural Research,* 22:23-26.

PẺREZ-VICH, B., FERNẢNDEZ-MARTINEZ, J.M., GRONDONA, M., KNAPP, S.J., & BERRY, S.T. 2002. *Stearoyl-ACP* and *oleoyl-PC* desaturase genes cosegregate with quantitative trait loci underlying high stearic and high oleic acid mutant phenotypes in sunflower. *Theoretical and Applied Genetics,* 104:338-349.

POWEL, W., MACHRAY, G.C. & PROVAN, J. 1996. Polymorphism revealed by simple sequence repeats. *Trends in Plant Science,* 1:215-222.

PRIOLLI, R.H.G., JUNIOR, C.T.M., ARANTES, N.E. & CONTEL, E.P.B. 2002. Characterization of Brazilian soybean cultivars using microsatellite markers. *Genetics and Molecular Biology,* 25:185-193.

PUTNAM, D.H., OPLINER, E.S., HICKS, D.R., DURGAN, B.R., NOETZEL, D.M., MERONUCK, R.A., DOLL, J.D. & SCHULTE, E.E. 1990. Sunflower. Alternative field crop manual. [Online]. Available at: <http www.hort.purdue.edu/newcrop/sunflower.html>. Accessed: 04/04/2010.

RAGOT, M., BIASIOLLI, M.F. & GAY, G. 1995. Marker-assisted backcrossing: a practical example. *Les Colloques,* 72:45-56.

ROJAS-BARROS, P., DE HARO, A. & FERNÁNDEZ-MARTÍNEZ, J.M. 2005. Inheritance of high oleic/low ricinoleic acid content in the seed oil of caster mutant *OLE-1*. *Crop Science,* 45:157-162.

ROZAS, J. & ROZAS, R. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics,* 15:174-175.

ROZAS, J., SÅNCHEZ-DELBARRIO, J.C., MESSEGUER, X. & ROZAS, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics,* 19: 2496-2497.

SAGHAI-MAROOF, M.A., SOLIMAN, K.M., JORGENSEN, R.A. & ALLARD, R.W. 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences USA,* 81:8014-8019.

SARRAFI, A. & GENTZBITTEL, L. 2008 Genomics as efficient tools: example sunflower breeding. In: Nagata, T., Lorz, H. & Widholm, J.M. (ed.)

Molecular marker systems in plant breeding and crop improvement. New York: *Springer-Verlag*, 107-120.

SCHIERHOLT, A., RÜCKER, B. & BECKER, H.C. 2001. Inheritance of high oleic acid mutations in winter oilseed rape (*Brassica napus* L.). *Crop Science,* 41:1444-1449.

SCHUPPERT, G.F., TANG, S., SLABAUGH, M.B. & KNAPP, S.J. 2006. The sunflower high-oleic mutant *Ol* carries variable tandem repeats of FAD2-1, a seed-specific *oleoyl-phosphatidyl choline desaturase*. *Molecular Breeding,* 17:241-256.

SEMAGN, K., BJØRNSTAD, Å. & NDJIONDJOP, M.N. 2006. Progress and prospect of marker-assisted backcrossing as a tool in crop breeding programs. *African Journal of Biotechnology,* 5:2588-2603.

SINGH, A.K., GOPALAKRISHNAN, S., SINGH, N.K., SHARMA, T.R., NAGARAJAN, M., VINOD, K.K., DEVINDER, S., SINGH, U.D., SUBHASH, C., ATWAL, S.S., RAKESH, S., VIKAS, K.S., RANJITH, K.E., ATUL, S., DEEPTI, A., APURVA, K., SHEEL, Y., NITIKA, G., ASHUTOSH, S., ASIF, B.S., AVITA, S. & BALRAM, M. 2011. Marker assisted selection: a paradigm shift in Basmati breeding. *Indian Journal of Genetics,* 71:1-9.

SINGH, A.K. & PRASANNA, B.M. Unknown. Molecular mapping in crop Plants: Development and characterization of mapping populations. [Online]. Available at:<www.scribd.com/doc/6229849/mapping-population>. Accessed 16/02/2011.

SLEPER, D.A. & POEHLMAN, J.M. 2006. Breeding field crops. In: SLEPER, D.A. & POEHLMAN, J.M (ed). *Breeding hybrid cultivars.* Iowa: Blackwell:171-184.

SMULDERS, M.J.M., BREDEMEIJER, G., RUS-KORTEKAAS, W., ARENS, P. & VOSMAN, B. 1997. Use of short microsatellites from data base sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical and Applied Genetics,* 97:264-271.

SOLDATOV, K.I. 1976. Chemical mutagenesis in sunflower breeding. *Proceedings of 7th International Sunflower Conference,* 352-357.

ŠPUNAROVÁ, M., OVESNÁ, J., TVARŮŽEK, L., KUČERA, L., ŠPUNAR, J. & HOLLEROVÁ, I. 2005. The use of molecular markers for characterisation of spring barley for breeding to *Fusarium* head blight resistance. *Journal of Plant, Soil and Environment,* 51:483-490.

SUNDARAM, R.M., VISHNUPRIYA, M.R., BIRADAR, S.K., LAHA, G., REDDY, G.A., RANI, N.S., SARMA, N.P. & SONTI, R.V. 2008. Marker-assisted introgression of bacterial blight resistance in *Samba mahsuri*, an elite indica rice variety. *Euphytica,* 160:411-422.

TANG, S., KISHORE, V.K. & KNAPP, S.J. 2003. PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. *Theoretical and Applied Genetics,* 107:6-19.

TANG, S. & KNAPP, S.J. 2003. Microsatellites uncover extraordinary molecular genetic diversity in Native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics,* 106:990-1003.

TANG, S., YU, J.K., SLABAUGH, M.B, SHINTANI, D.K. & KNAPP, S.J. 2002. Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics,* 105:1124-1136.

TANHUANPÄÄ, P., VILKKI, J. & VIHINEN, M. 1998. Mapping and cloning of *FAD2* gene to develop allele-specific PCR for oleic acid in spring turnip rape (*Brassica rapa* subspecies. *oleifera*). *Molecular Breeding,* 4:543-550.

TANKSLEY, S.D., GANAL, M.W. & MARTIN, G.B. 1995. Chromosome landing: A paradigm for map-based gene cloning in plants with large genomes. *Trends in Genetics*, 11:63-68.

TONG, L., ZHANG, L., YU, S., CHEN, X. & BI, K. 2007. Analysis of the fatty acids from *Periploca sepium* by GC-MS and GC-FID. *Asian Journal of Traditional Medicines*, 2:110-114.

TSUJIMOTO, H. 2001. Production of Near-Isogenic lines and marked Monosomic lines in common wheat (*Triticum aestivum*) c.v. Chinese spring. *Heredity*, 92:254-259.

URIE, A.L. 1984. Inheritance of very high oleic acid content in sunflower. *Proceedings of Sunflower Research Workshop, Bismark. ND*, 9-10.

URIE, A.L. 1985. Inheritance of the high oleic acid in sunflower. *Crop Science*, 25:986-989.

VENKANA, V., REDDY, D.L. & RANGANATHA, A.R.G. 2008. Identification of restorers and maintainers for different CMS sources in sunflower using new inbred lines. *Helia*, 31(49):65-70.

VENKATESHA, S.C., RAMANJINI GONDA, P.H., GANAPATHY, K.N., BYRE GONDA, M., RAMACHANDRA, R., GIRISH, G., CHANNAMALLIKARJUNA, V., SHANTALA, L. & GONDA, T.K.S. 2010. Genetic fingerprinting in Dolichos bean using AFLP markers and morphological traits. *International Journal of Biotechnology & Biochemistry*, 6:395-404.

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

VUYSTEKE, M., MANK, R., ANTONISE, R., BASTIAANS, E., SENIOR, M.L. & STUBER, C.W. 1999. Two high density AFLP linkage maps of *Zea mays* L. analysis of distribution of AFLP markers. *Theoretical and Applied Genetics,* 99:921-935.

WAN, P.J., DOWD, M.K., THOMAS, A.E. & BUTLER, B.H. 2007. Trimethylsilyl derivatization/Gas Chromatography as a method to determine the free fatty acid content of vegetable oils. *American Oil Chemistry Society,* 84:701-708.

WEEDEN, N.R., MUEHLBAUER, F.J. & LADIZINSKY, G. 1992. Extensive conservation of linkage relationships between pea and lentil genetic maps. *Heredity,* 83:123-129.

WELZ, H.G. & GEIGER, H.H. 2000. Principles of marker-assisted selection. In: HAUSSMANN, B.I.G., GEIGER, H.H., HESS, D.E., HASH, C.T. & BRAMEL-COX, P. (ed.). *Application of molecular markers in plant breeding*. *Training manual for a seminar.* IITA, Nigeria, Ibadan, 63-69.

WU, Y., BHAT, P.R., CLOSE, T.J. & LONARDI, S. 2008. Efficient and accurate contraction of genetic linkage maps from the minimum spanning tree of graph. *PLoS Genetics,* 4*:*e1000212.

XU, Y. & CROUCH, J.H. 2008. Marker-Assisted selection in plant breeding: From publications to practice. *Crop Science,* 48:391-407.

YOUNG, N.D., ZAMIR, D., GANAL, M.W. & TANKSLEY, S.D. 1988. Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *tm-2a* gene in tomato. *Genetics,* 120:579-585.

YU, J., TANG, S., SLABAUGH, M.B., HEESACKER, A., COLE, G., HERRING, M., SOPER, J., HAN, F., CHU, W., WEBB, D.M., THOMPSON, L., EDWARDS, K.J., BERRY, S., LEON, A.J., GRONDONA, M., OLUNGU, C., MAES, N. & KNAPP, S.J. 2003. Towards a saturated molecular genetic linkage map for cultivated sunflower. *Crop Science,* 43:367-387.

YU, J.K., MANGOR, J., THOMPSON, L., EDWARDS, K.J., SLABAUGH, M.B. & KNAPP, S.J. 2002 Allelic diversity of simple sequence repeats among elite inbred lines of cultivated sunflower. *Genome,* 45:652-660.

YUE, B., CAI, X., YUAN, W., VICK, B.A. & HU, J. 2009. Mapping the quantitative trait loci (QTL) controlling seed morphology in sunflower (*Helianthus annuus* L.). *Helia,* 32(50):17-36.

ZABEAU, M. & VOS, P. 1993. Selective restriction amplification: A general method for DNA fingerprinting. *European Patent Publication,* 92402629 (Publication number: EP0534858A1).

ZANE, L., BARGELLONI, L. & PATARNELLO, T. 2002. Strategies for micro-satellite isolation: a review. *Molecular Ecology,* 11:1-16.

ZIEGEL, J.S., SU, Y., CORCORAN, K.P., NIE, L., MAYNARD, P.E., HOFF, L.B., McBRIDE, L.J., KRONICK, M.N. & DIEHL, S.R. 1992. Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics,* 14:1026-1031.

APPENDIX A

Table A1 The percentage of recurrent parent genome contribution after the first backcrossing of 42 plants of the BC_1F_1 population derived from a cross between AP901-56-1-1 (HO) and RK74-26-1-2-1 (LO) inbred lines. Fourteen polymorphic SSRs were used as background markers

	ORS543	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% recurrent parent**
Parent (HO)	**A/A**	**A/B**	**A/B**	**A/B**	**A/A**	**A/A**	**A/B**	**A/A**	**A/B**	**A/A**	**A/A**	**A/B**	**A/A**	**A/B**	
Parent (LO)	**B/B**	**B/B**	**B/B**	**B/B**	**A/B**	**B/B**	**B/B**	**A/B**	**B/B**	**B/B**	**A/B**	**B/B**	**A/B**	**B/B**	
1.1	B/B	B/B	B/B	A/A	A/A	B/B	A/B	A/B	A/B	B/B	B/B	A/B	A/A	B/B	29
1.2	B/B	A/B	A/B	B/B	B/B	B/B	B/B	A/B	B/B	A/A	A/A	B/B	A/B	A/A	29
1.3	B/B	A/B	B/B	A/B	A/B	A/B	A/A	B/B	B/B	B/B	A/A	B/B	A/B	A/A	36
1.4	A/A	A/A	A/B	A/A	A/B	B/B	A/A	A/B	A/B	A/A	B/B	B/B	A/B	A/A	36
1.5	B/B	A/A	A/A	A/B	A/B	B/B	B/B	A/A	A/A	B/B	B/B	A/B	A/B	B/B	29
1.6	B/B	B/B	B/B	A/B	B/B	A/A	A/B	A/B	B/B	A/A	A/B	B/B	A/A	A/B	36
1.7	A/B	B/B	A/A	B/B	A/B	A/A	B/B	A/A	B/B	A/B	A/B	B/B	A/B	A/A	36
1.8	A/A	A/B	A/A	B/B	B/B	B/B	B/B	A/A	A/B	A/A	A/B	B/B	A/B	B/B	29
1.9	B/B	A/B	B/B	A/A	A/A	B/B	A/A	A/B	A/A	B/B	A/B	A/B	B/B	B/B	29
1.10	B/B	B/B	A/B	B/B	A/A	A/B	A/A	A/A	B/B	B/B	A/B	A/B	B/B	A/A	29
1.11	A/B	B/B	B/B	A/A	A/B	A/B	B/B	A/A	B/B	A/A	B/B	A/A	A/B	A/A	29
1.12	A/A	B/B	B/B	B/B	A/B	B/B	A/B	A/A	A/B	B/B	A/B	A/B	A/A	B/B	36

Table A1 (continued)

BC$_1$F$_1$	ORS543	ORS674	ORS621	ORS810	ORS483	ORS1178	ORS595	ORS630	ORS1265	ORS230	ORS316	ORS925	ORS309	ORS887	% recurrent parent
1.13	B/B	A/A	A/B	A/A	A/A	B/B	B/B	A/A	B/B	B/B	A/B	A/B	A/A	B/B	21
1.14	A/B	A/A	B/B	B/B	A/B	A/B	A/B	A/A	A/B	A/A	A/B	A/B	A/A	B/B	50
1.15	B/B	A/A	A/B	A/A	A/B	A/B	B/B	A/A	B/B	A/B	B/B	B/B	A/B	A/A	36
1.16	A/B	A/A	A/A	A/B	A/A	B/B	B/B	A/B	A/A	A/A	A/B	A/A	A/B	A/A	36
1.17	B/B	A/B	B/B	B/B	A/A	B/B	A/B	A/B	B/B	A/B	A/B	B/B	B/B	A/B	36
1.18	A/B	A/B	B/B	A/B	A/A	A/A	B/B	A/B	A/A	B/B	A/B	B/B	B/B	A/B	43
1.19	A/A	A/B	?/?	B/B	A/B	?/?	A/B	A/B	?/?	B/B	?/?	A/B	A/A	B/B	36
1.20	B/B	B/B	A/A	A/A	B/B	B/B	A/A	A/A	B/B	B/B	A/A	A/A	A/B	B/B	7
1.21	B/B	B/B	B/B	B/B	A/B	B/B	A/A	A/A	A/B	A/A	B/B	A/B	A/B	B/B	29
1.22	**A/A**	**B/B**	**A/B**	**A/B**	**A/B**	**A/B**	**B/B**	**A/B**	**B/B**	**A/B**	**A/B**	**A/A**	**A/B**	**A/B**	**64**
1.23	B/B	A/A	A/A	A/B	A/A	A/B	B/B	A/B	A/A	A/A	A/B	A/B	A/B	A/A	43
1.24	B/B	A/A	A/B	A/B	B/B	B/B	B/B	B/B	B/B	A/A	B/B	A/B	A/A	A/B	29
1.25	A/B	A/B	B/B	A/B	A/B	B/B	B/B	B/B	A/B	B/B	A/B	B/B	A/B	B/B	43
1.26	B/B	A/B	B/B	B/B	A/B	A/B	B/B	A/B	B/B	B/B	A/B	B/B	A/B	A/A	43
1.27	B/B	A/B	B/B	A/B	A/B	B/B	A/B	A/B	B/B	B/B	A/B	A/A	A/B	B/B	50
1.28	**A/B**	**B/B**	**A/B**	**A/B**	**A/B**	**B/B**	**B/B**	**A/B**	**A/B**	**A/B**	**A/A**	**B/B**	**A/B**	**B/B**	**57**

Table A1 (continued)

BC_1F_1	ORS543	ORS674	ORS621	ORS810	ORS483	ORS1178	ORS595	ORS630	ORS1265	ORS230	ORS316	ORS925	ORS309	ORS887	% recurrent parent
1.29	B/B	B/B	B/B	B/B	B/B	A/B	A/A	A/A	B/B	B/B	B/B	A/B	A/B	A/B	29
1.30	A/B	A/A	B/B	B/B	A/B	A/B	A/B	A/B	A/A	B/B	A/B	B/B	B/B	A/B	50
1.31	B/B	A/A	A/B	A/A	A/B	B/B	B/B	A/B	B/B	A/A	A/B	B/B	A/B	B/B	36
1.32	A/B	A/B	A/A	A/B	B/B	A/B	B/B	A/B	B/B	A/A	A/A	B/B	A/B	A/B	50
1.33	A/A	A/B	B/B	B/B	A/A	B/B	A/A	A/B	B/B	B/B	B/B	A/A	A/B	A/B	29
1.34	B/B	A/B	A/A	A/B	A/B	B/B	A/A	A/A	B/B	B/B	A/B	A/A	B/B	B/B	29
1.35	B/B	A/A	B/B	B/B	A/B	A/A	A/B	A/B	A/B	A/A	A/B	B/B	A/A	A/A	29
1.36	A/A	A/A	A/B	A/A	B/B	B/B	B/B	A/B	B/B	A/B	B/B	B/B	A/A	B/B	21
1.37	B/B	B/B	B/B	A/A	B/B	A/A	B/B	B/B	B/B	A/B	A/A	A/A	A/B	B/B	14
1.38	B/B	B/B	A/B	B/B	A/B	B/B	B/B	B/B	A/A	B/B	B/B	A/B	A/B	A/A	29
1.39	B/B	B/B	B/B	B/B	A/B	A/B	A/A	A/A	A/A	A/A	A/B	B/B	A/B	B/B	29
1.40	A/B	A/A	A/A	B/B	A/A	A/B	B/B	A/B	A/A	B/B	A/B	B/B	A/B	B/B	36
1.41	B/B	A/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	B/B	A/A	A/A	A/A	?/?	21
1.42	A/B	A/B	B/B	A/A	A/B	B/B	B/B	A/B	B/B	B/B	B/B	B/B	A/A	?/?	29

Table A2 The percentage of recurrent parent genome contribution after the first backcrossing of 18 plants of the BC_1F_1 population derived from a cross between AP901-56-1-1 and H55-9-2-1-1 inbred lines. Fourteen polymorphic SSRs were used as background markers.

	ORS543	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% recurrent parent**
Parent (HO)	A/B	A/B	A/B	A/B	A/A	A/A	A/B	A/A	A/B	A/A	A/A	A/B	A/A	A/B	
Parent (LO)	B/B	B/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	A/B	B/B	
1.1	B/B	A/B	B/B	A/A	B/B	B/B	B/B	A/B	A/A	A/A	A/B	B/B	A/B	B/B	29
1.2	B/B	B/B	A/A	B/B	A/B	A/B	A/B	A/B	B/B	B/B	B/B	A/B	A/B	B/B	43
1.3	A/B	A/B	A/A	A/B	B/B	B/B	B/B	A/A	A/B	A/A	A/A	B/B	B/B	B/B	29
1.4	A/A	B/B	B/B	A/A	A/A	B/B	A/B	A/B	A/B	A/B	B/B	B/B	A/A	A/A	29
1.5	A/B	A/A	A/B	A/B	A/A	A/A	B/B	A/B	B/B	A/A	A/B	B/B	A/B	A/A	43
1.6	B/B	B/B	B/B	A/B	?/?	B/B	B/B	B/B	B/B	B/B	A/B	A/B	A/B	A/B	36
1.7	**B/B**	**A/B**	**B/B**	**B/B**	**A/B**	**B/B**	**A/B**	**A/B**	**A/B**	**B/B**	**B/B**	**A/B**	**A/B**	**B/B**	**50**
1.8	A/A	B/B	A/B	B/B	A/A	B/B	B/B	A/B	A/A	A/A	A/A	A/A	A/B	A/B	29
1.9	A/A	A/A	B/B	A/A	A/B	A/B	A/A	A/B	B/B	B/B	A/A	A/B	A/B	A/A	36
1.10	A/B	B/B	A/B	B/B	A/B	A/B	A/B	B/B	B/B	B/B	B/B	A/B	A/B	B/B	50
1.11	A/A	A/B	A/A	A/A	A/A	B/B	A/B	A/A	A/B	A/A	B/B	A/B	A/B	B/B	36
1.12	B/B	B/B	B/B	A/B	A/A	B/B	B/B	A/A	A/A	A/B	A/B	B/B	A/B	A/A	29
1.13	A/A	A/B	B/B	A/B	A/A	A/A	A/A	A/B	B/B	A/A	A/B	A/A	A/B	A/A	36
1.14	B/B	B/B	A/A	B/B	A/B	B/B	A/B	A/A	B/B	A/B	A/B	B/B	A/B	B/B	36

Table A2 (continued)

BC_1F_1	**ORS543**	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% recurrent parent**
1.15	B/B	A/B	A/A	B/B	B/B	B/B	B/B	A/B	A/B	B/B	A/B	B/B	A/A	B/B	29
1.16	?/?	B/B	B/B	A/A	A/B	A/A	B/B	A/B	A/A	B/B	A/B	A/B	A/B	A/A	36
1.17	B/B	A/B	A/B	B/B	A/B	B/B	B/B	A/A	A/B	A/A	B/B	A/A	A/A	A/A	29
1.18	**B/B**	**B/B**	**B/B**	**A/A**	**A/B**	**B/B**	**A/B**	**A/B**	**B/B**	**A/B**	**A/B**	**A/B**	**B/B**	**A/B**	**50**

Table A3 The percentage of recurrent parent genome contribution after the first backcrossing of 18 plants of the BC_1F_1 population derived from a cross between AP901-56-1-1 and H55-9-2-1-1 inbred lines. Fourteen polymorphic SSRs were used as background markers.

	ORS543	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% recurrent parent**
Parent (HO)	A/B	A/B	A/B	A/B	A/A	A/A	A/B	A/A	A/B	A/A	A/A	A/B	A/A	A/B	
Parent (LO)	B/B	A/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	B/B	A/B	B/B	A/B	B/B	
1.1	A/B	A/B	A/A	B/B	A/A	A/B	B/B	B/B	A/B	B/B	A/B	A/A	A/B	B/B	43
1.2	A/B	B/B	B/B	A/A	A/A	A/B	A/A	A/A	A/A	A/B	A/A	B/B	A/B	B/B	29
1.3	A/B	A/A	B/B	A/B	A/B	A/A	B/B	A/B	A/B	B/B	B/B	A/B	A/B	B/B	50
1.4	A/B	B/B	A/A	B/B	A/B	A/B	B/B	A/B	A/A	?/?	A/B	B/B	B/B	A/B	43
1.5	A/B	A/B	B/B	B/B	A/B	B/B	A/B	A/B	A/A	A/A	A/B	B/B	A/B	A/A	50
1.6	B/B	A/B	B/B	B/B	B/B	A/B	A/A	A/A	B/B	A/A	A/A	B/B	A/B	A/B	29
1.7	A/A	B/B	B/B	B/B	A/A	B/B	B/B	A/A	A/B	B/B	B/B	A/A	A/B	B/B	14
1.8	A/A	B/B	A/A	A/B	A/B	B/B	B/B	A/B	A/B	B/B	A/B	A/A	B/B	A/A	36
1.9	B/B	B/B	B/B	B/B	A/B	A/A	B/B	A/B	A/B	A/A	A/B	B/B	A/A	B/B	29
1.10	B/B	B/B	B/B	B/B	A/B	B/B	?/?	A/B	B/B	A/B	B/B	B/B	A/B	B/B	29
1.11	A/A	B/B	A/A	A/B	B/B	B/B	A/A	B/B	B/B	A/B	A/A	A/A	A/B	A/B	29
1.12	B/B	A/A	B/B	B/B	A/B	B/B	A/B	B/B	A/A	B/B	B/B	A/B	A/B	A/B	36
1.13	B/B	B/B	B/B	B/B	A/B	A/B	B/B	A/B	B/B	A/A	A/B	B/B	A/B	B/B	36
1.14	B/B	B/B	A/B	B/B	A/A	A/B	B/B	A/B	A/A	B/B	A/B	B/B	A/B	A/B	43
1.15	A/A	A/B	B/B	A/A	A/A	B/B	A/B	A/B	A/B	B/B	B/B	A/B	A/B	A/B	50

Table A3 (continued)

BC$_1$F$_1$	**ORS543**	**ORS674**	**ORS621**	**ORS810**	**ORS483**	**ORS1178**	**ORS595**	**ORS630**	**ORS1265**	**ORS230**	**ORS316**	**ORS925**	**ORS309**	**ORS887**	**% recurrent parent**
1.16	B/B	B/B	A/B	B/B	B/B	B/B	B/B	A/B	B/B	A/A	A/A	B/B	A/B	B/B	21
1.17	B/B	B/B	B/B	A/B	A/B	A/B	B/B	B/B	B/B	B/B	A/A	B/B	A/B	A/A	29
1.18	A/B	A/A	A/B	A/A	A/B	B/B	A/A	?/?	A/B	A/A	A/B	B/B	A/B	B/B	43
1.19	A/A	B/B	A/A	A/B	A/B	B/B	B/B	A/B	A/A	B/B	B/B	A/B	A/B	B/B	36
1.20	A/B	A/B	B/B	A/B	B/B	A/A	A/B	A/B	B/B	A/A	A/B	B/B	A/A	A/A	43
1.21	B/B	B/B	A/A	B/B	B/B	B/B	B/B	A/A	B/B	A/B	A/B	B/B	A/B	B/B	21
1.22	B/B	B/B	A/A	B/B	B/B	B/B	B/B	A/B	A/B	A/A	A/B	B/B	A/B	B/B	29
1.23	A/A	A/B	B/B	A/A	A/A	B/B	A/B	A/B	A/A	B/B	A/B	A/B	B/B	B/B	36
1.24	B/B	B/B	A/B	B/B	A/A	A/B	B/B	A/B	B/B	B/B	A/B	A/B	B/B	A/A	36
1.25	A/B	B/B	B/B	A/A	A/B	A/B	B/B	A/B	B/B	A/A	B/B	A/A	A/B	A/A	36
1.26	B/B	A/A	B/B	B/B	A/B	B/B	A/B	A/A	A/B	B/B	A/B	A/B	A/A	A/B	43
1.27	B/B	B/B	A/B	?/?	A/A	B/B	B/B	A/B	B/B	B/B	A/B	A/B	A/A	B/B	29
1.28	**B/B**	**A/B**	**B/B**	**B/B**	**A/B**	**A/B**	**A/B**	**A/B**	**A/B**	**A/A**	**A/B**	**A/B**	**A/A**	**A/B**	**64**
1.29	A/B	B/B	A/B	A/A	A/B	B/B	B/B	A/A	B/B	A/B	B/B	B/B	A/B	A/A	36
1.30	A/A	A/B	A/A	A/B	A/A	B/B	B/B	A/B	A/A	A/A	A/B	A/A	A/B	B/B	36
1.31	B/B	B/B	B/B	B/B	A/A	B/B	A/B	A/B	B/B	A/B	A/B	B/B	B/B	B/B	29
1.32	B/B	A/A	B/B	A/B	A/A	B/B	B/B	A/B	A/A	B/B	A/B	B/B	B/B	A/A	21
1.33	A/A	A/A	A/B	B/B	A/B	A/A	A/A	A/B	B/B	B/B	A/A	A/B	A/A	A/A	29

Table A3 (continued)

BC$_1$F$_1$	ORS543	ORS674	ORS621	ORS810	ORS483	ORS1178	ORS595	ORS630	ORS1265	ORS230	ORS316	ORS925	ORS309	ORS887	% recurrent parent
1.34	B/B	A/B	A/A	A/A	B/B	B/B	B/B	A/A	B/B	B/B	A/A	A/A	A/B	B/B	14
1.35	B/B	B/B	B/B	B/B	A/B	B/B	A/A	A/A	A/B	A/A	A/B	A/B	A/B	B/B	36
1.36	**B/B**	**A/A**	**A/B**	**A/B**	**A/B**	**A/B**	**B/B**	**A/B**	**A/A**	**A/B**	**A/B**	**A/A**	**A/B**	**B/B**	**57**
1.37	**A/B**	**B/B**	**A/A**	**A/B**	**B/B**	**A/B**	**A/B**	**A/B**	**A/A**	**A/A**	**A/B**	**A/B**	**A/B**	**A/A**	**57**
1.38	B/B	A/B	A/B	A/B	B/B	B/B	A/B	B/B	B/B	A/A	B/B	A/B	A/A	A/A	36
1.39	A/B	B/B	B/B	A/B	A/B	B/B	B/B	B/B	A/B	B/B	A/B	B/B	A/B	A/B	50
1.40	B/B	B/B	B/B	B/B	A/B	A/B	B/B	A/B	B/B	B/B	A/B	B/B	A/B	A/B	43
1.41	**A/B**	**A/B**	**A/B**	**A/B**	**A/A**	**A/A**	**A/B**	**A/A**	**A/B**	**A/A**	**A/A**	**A/B**	**A/A**	**A/B**	**57**