IDENTIFICATION OF GENETIC DISTANCES AND HETEROTIC GROUPS OF INBRED MAIZE (*Zea mays*) LINES USING DNA FINGERPRINTING.

EZANNE SWANEPOEL

Submitted in fulfilment of the degree

Magister Scientiae Agriculturae

Faculty of Agriculture
Department of Plant Breeding
University of the Orange Free State
BLOEMFONTEIN

December 1999

Supervisor: Prof. M.T. Labuschagne
Co-Supervisor: Dr. C.D. Viljoen

Dedicated to
Jenny

PREFACE

The results presented in this thesis follow from the study which was carried out at the department of Plant Breeding, at University of the Orange Free State, Bloemfontein, under the supervision of Prof. M.T. Labuschagne and co-supervision of Dr. C.D. Viljoen.

The results presented here are original and have not been submitted at any other University.

Ezanne Swanepoel

ACKNOWLEDGEMENTS

I would like to thank the following people, organisations and institutions for their contribution towards the success of this thesis:

- Prof. Maryke Tine Labuschagne for the excellent supervision and continued support, motivation and enthusiasm.
- Dr. Chris Viljoen for excellent (co) supervision and valuable theoretical and practical input.
- Personnel of Department of Plant Breeding at UOFS for providing excellent support, encouragement and most of all their friendship.
- Department of Botany and Genetics for providing excellent research facilities.
- All personnel and students of Molecular and Genetics Laboratory at Department of Botany and Genetics for their outstanding technical support and assistance, research input, advice and their friendship.
- Molecular Biology Laboratory in the Department of Plant Breeding for providing research facilities for data interpretation.
- My best friend Jenny de Bruyn and her two daughters, Lorna Lee and Christal, for their encouragement, love, patience and support.
- My parents and two sisters, Celecia and Zanelle, for continued support and understanding.
- Advanta Africa Seeds for providing research funds.
- All the colleagues and friends that directly or indirectly have influenced my studies.
- God, who made all this possible.

OUTLINE

CONTENTS

LIST OF ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
bp	Base-pairs
BSA	Bulked Segregant Analysis
°C	degrees Celsius
CIAT	Centro Internacional de Agricultura Tropical
D	pairwise genetic distance
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotides
EDTA	Ethylenediaminetetra-acetic Acid
F	index of genetic similarity
kb	kilobases (1kb = 10^3 base-pairs)
L	litre
MAS	Marker-assisted Selection
$MgCl_2$	magnesium chloride
m/v	Mass per volume
mm	millimetre
mM	millimolar
NaCl	potassium chloride
NILs	Near-isogenic Lines
PAUP	Phylogenetic Analysis Using Parsimony
PCR	Polymerase Chain Reaction

PHYLIP	Phylogeny Inference Package
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
S	similarity coefficient
SCA	Specific Combining Ability
SSR	Short Sequence Repeats
STS	Sequence-tagged Site
TAE	Tris-acetate-EDTA buffer
Tris	Tris (hydroxymethyl) – aminomethane
UOFS	University of the Orange Free State
UPGMA	Unweighted Pair-group Mean Arithmetic
UV	Ultraviolet
v/v	Volume per volume
W	Watt

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Southern Africa is a developing region with a relatively low agronomic potential. Maize is currently the dominant crop and accounts for 79% of all grain production. The dominance of maize is highlighted by the fact that maize production for the world is 25% of the total cereal production compared to 79% for Southern Africa. There are a number of factors in Southern Africa, together with cereal science and technology that affect the maize industry. These factors include strong population growth, low living standards, rapid urbanisation, democratisation, free markets and improved scientific communication. Research needs for the maize market cover aspects such as the maize food market, the maize feed market and the industrial market. Cownie (1993) suggested that research in the following marketing areas should enjoy priority: small-scale maize producers, commercial maize producers and food requirements of the penurious sector.

Although never totally closed, the political developments in South Africa have opened the door for better scientific communication with neighbouring countries. The new dispensation in South Africa presents exciting possibilities. These developments can all contribute to cereal science in the quest to invigorate the maize giant of Southern Africa (Cownie, 1993).

Maize breeding has been effective in developing improved varieties and hybrids to meet the rapidly changing cultural and environmental conditions of the past century. The development of the commercial seed industry is testimony of successful breeding methods that have evolved for the economical production of high-quality hybrid seed that is accepted and sought after by the modern farmer (Hallauer and Miranda, 1988). White maize is the staple food of the majority of South Africans, with an estimated annual consumption of 3.0 metric tons (mt). A further 3.5 mt of yellow maize is used annually by various industries. High quality grain is in demand by both the consumers

and the industry and this compels farmers to produce grain of a given quality standard (Du Plessis, 1993).

The availability of elite, adapted varieties of different genetic backgrounds allows the farmer to hedge the risk of reduced performance by any one variety under unpredictable environmental stresses. Independent selection made in different locations, according to different breeding methodologies and criteria, utilising different germplasm, are all facets that have given rise to new genetic recombinants (Russell, 1974; Duvick, 1977; Castleberry *et al.*, 1984). Every year breeders of autogamous crops produce a multitude of potentially useful crosses and evaluate their varietal ability (Gallais, 1979) by testing either selfed progenies or doubled haploid lines in field experiments. If breeders could predict the prospect of crosses for line development before producing and testing lines derived from them in field trials, this would increase the efficiency of breeding programmes by concentrating the efforts on the most promising crosses (Bohn *et al.*, 1999).

One strategy used in breeding programmes to predict the prospects of crosses for line development, is based on the assumption that the specific combining ability (SCA) expressed by a hybrid is related to the genetic distance between its parental lines (Lee *et al.*, 1989). The ability to genetically define crop lines or breeding lines is useful for maintaining genetic purity, identifying proprietary germplasm and estimating genetic relationships. In breeding programmes, information on genetic relationships within species is used for organising germplasm collections, identifying heterotic groups and selecting breeding material (Prabhu *et al.*, 1997). DNA marker systems are useful tools for assessing genetic diversity between germplasm (Lee, 1995; Karp *et al.*, 1996). In comparison with morphological markers, DNA markers offer significant advantages with respect to increased numbers of loci detectable, overall phenotypic neutrality, and the ability to score the plant at any developmental stage (Prabhu *et al.*, 1997).

The aim of this study was to use AFLP fingerprinting for the identification and genetic distance analysis of a collection of maize inbred lines in the Advanta Africa genebanks. The knowledge of relative genetic variation and the genetic distance between maize inbred lines will be used to facilitate identification of diverse parents

in order to maximise the expression of heterosis in their potential progeny. To this end the inbred lines that have been used were assessed for genetic variation and then compared with morphological data.

CHAPTER 2

LITERATURE REVIEW

Introduction

Maize is well suited for genetic research as it is easy to grow, adapted to a wide range of environmental conditions and has a large number of distinct hereditary variations. Inbreeding or crossing is a simple and rapid procedure. Hundreds of kernels may be obtained on an ear from a single pollination and each kernel represents a distinct individual (Jugenheimer, 1976). The effective management of germplasm in terms of resources and biodiversity in breeding programmes, requires characterisation of material at both the agronomic (phenotypic) and genetic levels. Plant breeders focus on morphological traits to generate new lines and to meet the legal requirements for line registration, i.e. distinctness, uniformity and stability (Ellis *et al.*, 1997). Determining the relationship of the different lines at the DNA level, can however hold many advantages for the plant breeder, since it may increase the efficiency of breeding efforts to improve crop species (Barrett *et al.*, 1998). This may explain the reason for the development of all the different marker systems.

The first marker-assisted selection in breeding programmes was carried out using mutations at loci controlling plant morphology. Morphological markers have not, however, been extensively used by breeders, both because the mutations available are limited in number and most are not neutral in their phenotypic expression (Koebner *et al.*, 1994). Subjectivity in evaluating morphological characters and confinement of expression of some characters to a particular stage of development, such as flowering, presents further complications in line identification (Morell *et al.*, 1995).

The complications were overcome to some extent in the 1960's with the development of gel-based assays for proteins and other gene products. The electrophoretic or chromatographic differences between allelic gene products such as isozymes are

usually neutral in their effects on phenotype, i.e. they are not generally themselves the basis of recognisable variation at the whole plant level. Furthermore they are usually expressed co-dominantly, allowing the breeder to unequivocally distinguish between homozygous and heterozygous individuals for the expressed gene. Although isozyme analysis provided more markers than were available through morphology, they were still too few in number to use for line identification. In addition, because each isozyme system requires its own particular staining, and in some cases electrophoretic technique, their widespread use in breeding programmes has proven limited to a few, very specific applications (Koebner *et al.*, 1994).

An important breakthrough in the 1980's, was the development of RFLP's (Restriction Fragment Length Polymorphisms), in human genetics and later for plants. RFLP's are co-dominant, and are effectively unlimited in number since they rely on the ability to clone small (500-2000 nucleotide base pairs), random fragments from complex genomes (more than a billion nucleotide base pairs). The RFLP's are assayed by a common analytical technique, but require a relatively sophisticated assay environment and are time-consuming and expensive. In many applications of the technique there is a requirement for the use of radioactive assays, which present further problems for plant breeders, particularly those from developing countries. The late 1980's saw the emergence of the polymerase chain reaction (PCR) and with it a new generation of DNA marker techniques that are relatively simple and fast to use, lend themselves to automation and are often co-dominant (Koebner *et al.*, 1994).

DNA molecular markers have found wide scale application in agricultural breeding programmes through PCR-based technology (Koebner *et al.*, 1994). Of these techniques RAPD's (Random Amplified Polymorphic DNA markers) and AFLP's (Amplified Fragment Length Polymorphisms), have proved to be the most suitable for the identification of lines and varieties (Morell *et al.*, 1995). DNA fingerprinting has the potential to distinguish between distantly and closely related individuals (Nybom, 1994).

Morphological characteristics as markers

Morphological characters, for which the variant allelic phenotypes are sufficiently discrete to allow their segregation to be followed, are the easiest and generally most economical of all markers to assay. The use of morphological characters dates back to breeding and selection itself (Koebner *et al.*, 1994). Phenotypic identification of plants is based on morphological traits recorded in the field. The application of morphological characterisation has been used for different purposes. It has been used as a powerful tool in the classification of lines, to study taxonomic status, identification, determining of genetic variation and correlation of characteristics with agronomic potential (Centro Internacional de Agricultura Tropical (CIAT), 1993; Millan and Cubero, 1995).

There is a great range of phenotype variability among races, varieties, hybrids, and inbred lines of maize in terms of plant type, ear type, tassel type, and maturity. These phenotypic differences are relative to the germplasm used in breeding nurseries (Hallauer and Miranda, 1988). For example, Kiesselbach (1980) realized that the total number of leaves formed varies both between and within varieties. Hallauer and Miranda (1988) suggested the opportunities for intercrossing in maize are ample, because it is essentially 100% cross-pollinated through wind movement. Jugenheimer (1976) suggested that in maize the number and size of kernels contribute to grain yield and that the number of kernels are determined by the length of the ear, number of rows per ear, number of ears per plant, and number of plants per unit area. Although breeders and geneticists use morphological characteristics such as leaf and flower attributes to follow the segregation of genes in hybrids, most agronomic traits are not easily observed (Kochert, 1994).

Bonnett (1954, 1960) divided the development of the maize plant into the vegetative, transitional, reproductive, and seed stages. For these studies he used the methods of collection and classification of lines described by Wellhausen *et al.* (1952). These methods of classifying maize collections comprise four principal categories:

(1) vegetative characters of the plant which include response to altitude, height, total leaf number, number of leaves above the ear, width of ear-bearing leaf, venation-index and internode patterns;

(2) tassel characters which include tassel length, peduncle length, length of branching space, percentage of branching space, percentage of secondary and tertiary branches and the total number of branches;

(3) external and internal ear characters which include ear diameter, length, and row number, shank diameter and length, and denting, cob diameter, rachis diameter, cob/rachis index, rachilla length, rachilla/kernel index, glume/kernel index, cupule hairs, rachis flag, lower and upper glume traits, rachis induration and teosinte introgression;

(4) physiological, genetical, and cytological characters that include number of days from planting to anthesis, pubescence of leaf sheath, plant colour, midcob colour, chromosome knobs and B-chromosomes.

Morphological markers have been used successfully using the four principal categories of maize classification, to estimate the diversity of historically important sweetcorn inbreds (Gerdes and Tracy, 1994). For example, four independent populations comprising testcrosses of flint lines of maize were used by Lübberstedt *et al.* (1998) to select for forage traits based on morphological and agronomic traits (estimates of genotypic correlations among traits agreed well across the three validation populations). Zanoni and Dudly (1989) identified inbred lines useful for improving elite maize hybrids by evaluating them for grain yield, grain dry matter percentage, earliness of flowering and plant and ear height. Zehr *et al.* (1992) used a maize population as a donor of alleles for improvement of an elite hybrid (trait measurements like grain yield, grain moisture at harvest, final stand, number of stalk lodged plants, number of root lodged plants, plant and ear height and maturity were recorded). Mock and Frey (1970) measured association of meiotic stage with several morphological measurements including expanded leaf number, days after emergence, plant height, and stem diameter. Expanded leaf numbers proved to be a better index of meiosis than days after emergence, plant height, or stem diameter. However, use of the index is limited to genetically homogeneous populations.

Kuleshov (1933) classified maize by endosperm types in the following groups:

1. *Zea mays* indurata: flour
2. *Zea mays* indentata: dent
3. *Zea mays* everta: popcorn
4. *Zea mays* amylacea: flint
5. *Zea mays* saccharata: sweet
6. *Zea mays* amylea saccharata: starchy-sugary
7. *Zea mays* ceratina: waxy
8. *Zea mays* tunicata: pod

Classification of maize by endosperm types was a breakthrough for genetic research of maize in America (Anderson and Brown, 1952). These examples indicate that morphological markers have contributed to the improvement of maize in the past and that the technique has a role to play in the future, even if it is only a starting point in determining distinctiveness (Gerdes and Tracy, 1994).

The study of taxonomy and genetic relationships for identification of genotypes in many crops has become complicated due to a large number of spontaneous, as well as artificial, crosses. Certification of new lines is usually based on the genetic purity of a particular crop determined by botanical traits. However, morphology is often ambiguous and has limited use for line identification (Stegemann, 1984; Koebner *et al.*, 1994). Plant breeders face the difficult task of having to select for traits that are often under complex genetic control and subjected to environmental changes. Hybridisation and introgression make this problem even more difficult. Furthermore, many of these traits do not have the simple genetic control assumed in genetic models thus making analysis difficult (Liu and Furnier, 1993). Morphological characterisation requires mature plants, it usually displays dominant phenotype and there are too few available in any single species (Koebner *et al.*, 1994). The method involves a lengthy survey of plant growth that is labour intensive and timely (CIAT, 1993).

Advances in molecular biology have provided new methodologies, which extend the list of useful genetic markers (Paterson *et al.*, 1991). The first methods developed included protein techniques such as isozyme analysis (Bassiri, 1976; Arus *et al.*, 1982; Stegemann, 1984; Nienhuis *et al.*, 1995). Given a molecular map of sufficient density, it is possible to identify and locate short regions of the chromosome that account for much of the variation in yield or quality components of a cross. The identification and location of quantitative trait loci (QTL) in F2 populations of maize are also considered together with their manipulation using marker-facilitated procedures (Stuber *et al.*, 1990). However, these techniques were limited in use due to laboratory constraints (Hu and Quiros, 1991). Furthermore, isozymes often show low levels of polymorphism and problems with reproducibility arise due to tissue type and environmental conditions.

With the advance in molecular techniques, DNA based procedures have been proposed for line identification. DNA techniques have the advantage over conventional methods in that the composition of DNA is consistent in similar tissue types and is not affected by environmental changes (Beeching *et al.*, 1993). The development of DNA markers provides an opportunity to detect, monitor and manipulate genetic variation (Yamamoto *et al.*, 1994) more precisely than in the case of morphological and expressed phenotypic markers. These techniques include a variety of different methodologies commonly referred to as DNA fingerprinting (Nybom *et al.*, 1990). DNA molecular markers are potentially unlimited in number, are not affected by the environment and can be mapped on linkage maps (Soller and Beckmann, 1983; Winter and Kahl, 1995).

DNA-based molecular marker systems (DNA fingerprinting)

A number of diverse problems in plant breeding practice can be addressed via a genetic marker approach. The decision to exploit the possibilities opened up by the technology is more often influenced by economical or practical, rather than by technical considerations. Cost and labour efficiency has been steadily improving, and this trend can confidently be expected to continue (Koebner *et al.*, 1994).

Molecular markers provide a remarkable improvement in the efficiency and sophistication of plant breeding. It is now generally accepted that molecular markers represent the most significant advance in breeding technology that has occurred in the last few decades and is currently the most important application of molecular biology to plant breeding. The most common widespread application is indirect selection for linked traits, but other applications include accelerated backcrossing, pyramiding genes, identification of varieties and hybrids and detection of resistance genes (Langridge and Chalmers, 1998).

The application of DNA fingerprinting holds great potential in the identification of lines and species, and could help make breeding programmes more efficient through the detection of genetic linkages between DNA and important quantitative traits. The high variability of DNA fingerprinting described in humans, animals and plants allows the identification of different individual genotypes and species (Lin *et al.*, 1993). Thus molecular biological techniques can be employed in a variety of ways to facilitate crop improvement in the long term (Jones and Lindsey, 1990). DNA fingerprinting includes various techniques, used to differentiate between individuals at the species and subspecies levels (McClean *et al.*, 1994). There are two classes of DNA markers: those based on DNA-DNA hybridization as used in RFLP's and those based on the PCR amplification of genomic DNA sequences (Koebner *et al.*, 1994). With the exception of RFLP's, DNA markers lend themselves to automated analysis, which can provide the fast, high throughput desired for examining large numbers of samples that must typically be analysed in crop breeding (Perkin-Elmer, 1996). Detection frequency is always dependent on the level of genetic variance between the DNA templates or genomes compared (Sobral and Honeycutt, 1994b).

Restriction Fragment Length Polymorphisms (RFLP) technique

RFLP (Restriction Fragment Length Polymorphism) analysis is based on the detection of differences in the length of restriction fragments generated by the complete digestion of genomic DNA with restriction endonucleases (Botstein *et al.* 1980; Tanksley *et al.*, 1989; Hille *et al.*, 1991). Fragment length polymorphism is generated when a particular recognition site of a restriction enzyme is absent in one individual

and present in another, resulting in differently sized restriction fragments at that locus. The polymorphic fragments are detected by resolving the DNA fragments using electrophoresis and detection with probes (Southern, 1975).

DNA probes are used to detect single or multiple RFLP loci. Single copy RFLP's have the advantage that their scoring is unambiguous and that allelism (co-dominance) can be determined. In contrast, RFLP probes, which give a multi-banded profile have the advantage that more than one locus can be detected simultaneously. In general, cDNA probes as well as many genomic DNA probes, produce single or low copy hybridisation profiles. Some genomic DNA probes can detect repetitive sequences, and therefore provide multi-banded patterns, which can find uses in fingerprinting or varietal identification application (Koebner *et al.*, 1994).

RFLP mapping is time consuming, costly and labour intensive. The technique is difficult in some species with large and complex genomes (Marsan *et al.*, 1993). The complexity of the technique for performing RFLP analysis, coupled with the widespread use of short-lived radio-isotopes has led to some limitation for routine application in large scale crop improvement programs (Yamamoto *et al.*, 1994). In addition, the RFLP technique requires a substantial amount of DNA (5-10 µg per lane) (Beeching *et al.*, 1995).

An advantage of the RFLP technique is that it generates a lot of detectable loci, which are insensitive to environmental factors and can be used at any developmental stage of the organism (Kelly, 1995). This has allowed the extensive use of RFLP analysis in genetic studies (O'Toole, 1989), exploring evolutionary relationships among different species and populations (Springer *et al.*, 1994; Eubanks, 1997), for line and genotypic identification (Ajmone-Marsan *et al.*, 1992), gene mapping (Louie *et al.*, 1991; Burstin *et al.*, 1994) and determining heterotic groups (Dudley *et al.*, 1991). The RFLP technique has been particularly useful in mapping species that display a high level of intraspecific variation. Several maps have been compiled for maize (*Zea mays*)(Burr *et al.*, 1983; Helentjaris *et al.*, 1986; Melchinger *et al.*, 1990; Smith and Smith, 1991; Boppenmaier *et al.*, 1992; Prabhu *et al.*, 1997; Gardiner *et al.*, 1993).

The future role of genetic engineering (gene identification, location, cloning and introduction into a target genome) will require the generation of an extensive information base for each trait. RFLP's have been suggested as molecular markers to facilitate improvement of agronomic traits in maize and the RFLP technique can be applied immediately (Lee *et al.*, 1989; O'Toole, 1989). Accordingly, RFLP's have been suggested as superior markers to protein based techniques for assessing diversity and relationships in maize breeding germplasm (Murray *et al.*, 1988; Walton and Helentjaris, 1987).

PCR-based technique

The polymerase chain reaction (PCR) based technology has produced a second generation of molecular markers (Mullis and Faloona, 1987; Saiki *et al.*, 1988; Ehrlich *et al.*,1991; Koebner *et al.*, 1994). Most PCR programs include an initial denaturation step of three to five min at 94-95°C that is intended to completely denature complex genomic DNA so that primers can anneal after cooling (Rolfs *et al.*, 1992). During the first few temperature cycles with genomic DNA as template, primers must perform a 'genomic screening' process. The most critical component for optimising the specificity of any PCR-based assay is the choice of the annealing temperature (Ruano *et al.*, 1991) until they find complementary annealing sites. In later cycles, denaturation temperatures may even be shortened in order to preserve enzymatic activity (Gelfand, 1989). Yu and Pauls (1992) concluded that the best results should be obtained by optimising for the shortest possible denaturing time. Too many cycles may result in primer depletion and subsequent priming by amplification products, which often leads to longer products and smears in the gel (Rolfs *et al.*, 1992). The last step of the PCR program should not be a denaturation step, because rapid cooling of single structures of even heteroduplexes, instead of reannealing to complimentary fragments. This may disturb electrophoretic product analysis (Rolfs *et al.*, 1992). The main advantage of this technique over RFLP analysis is its inherent simplistic analysis and the ability to conduct PCR tests with extremely small quantities of tissue for DNA extraction (Edwards *et al.*, 1991; Welsch *et al.*, 1991). On the other hand, PCR is limited in its usefulness because of

the time and expenses required to obtain the DNA sequence information required for primer design (Samec and Nasinec, 1995).

Random Amplified Polymorphic DNA markers (RAPD's) technique

The discovery that PCR with arbitrarily selected primers will amplify a specific set of arbitrarily distributed loci in any genome, laid the foundation for the high output of genetic markers that can be used for a variety of purposes (Sobral and Honeycutt, 1994a). Williams *et al.* (1990) was the first to use random primers in PCR on plant samples and suggested that the technique be named RAPD's. RAPD's are obtained by PCR amplification of DNA fragments from randomly distributed loci using single arbitrary primers. The arbitrary primers used in the technique are usually 10bp in length. Ideally they have a GC content of 50% to 80% and contain no palindromic sequences. The number of DNA fragments amplified is dependent on the sequence of the primer and the size of the genome being used as template. A single nucleotide change in a primer usually results in a completely different amplification profile although mismatch pairing can take place, at the first two nucleotides at the 5' end of the primer (Yu *et al.*, 1993; Hosaka and Hanneman, 1994).

Reaction conditions used in the RAPD procedure limit the size of the amplification to a range of between 100 and 3000bp. Only DNA sequences that are within subsequent primer recognition sequences in inverted orientation are amplified. Polymorphisms occur as a result of either single base pair changes, deletions of primer sites or insertions that increase the fragment length between primer sites. Size variants are only rarely detected in RAPD analysis and as a rule, individual amplification products represent one allele per locus (Devos and Gale, 1992; Waugh and Powell, 1992).

The frequency of detecting RAPD polymorphisms is generally expressed as the number of polymorphisms per RAPD primer used. The detection frequency is always dependent on the level of genetic variance between the DNA templates compared. Consequently, the detection frequencies of RAPD's for lower taxa such as lines and varieties tend to be lower than that reported between different species and genera (Sobral and Honeycutt, 1994a).

The PCR technique using random primers (RAPD's) has identified DNA polymorphisms useful for genetic mapping in a large variety of organisms (Nicholson *et al.*, 1993; Rowland and Levi, 1994; Wang *et al.*, 1996). Although technically very powerful, the use of arbitrary primers for genome mapping has the disadvantage of characterising DNA sequences of unknown function. Thus, there is no reason to anticipate that DNA fragments amplified by use of arbitrary primers will be enriched for either transcribed or promoter sequences that may be conserved in evolution. For these reasons, the arbitrarily primed PCR method was modified by using oligonucleotide primers derived from conserved promoter elements and protein motifs. Twenty-nine of these primers were tested individually and in pairwise combinations for their ability to amplify genomic DNA from fowls, maize, *Drosophila*, man and various inbred strains of laboratory mice and *Mus spretus*. Using recombinant inbred strains of mice, the chromosomal location of 27 polymorphic fragments in the mouse genome was determined. The results demonstrated that motif sequence-tagged PCR products are reliable markers for mapping the mouse genome, and that motif primers can also be used for genomic fingerprinting of divergent species (Birkenmeier *et al.*, 1992; Vaillancourt and Hanau, 1992).

Although the dominant nature of RAPD's can be seen as a limitation in their use for genetic mapping, there are specific cases where only dominant markers can be used. For example, genetic mapping of polyploid species and tree species that contain high amounts of DNA and long generation times have been brought out of a near standstill by the application of RAPD technology (Sobral and Honeycutt, 1994a). Using RAPD profiles it is also possible to reliably and unambiguously cluster lines into groupings (Vaillancourt and Hanau, 1992; Ting, 1994; Cao and Oard, 1997).

The success of the RAPD technique lies in the ability to generate a large number of molecular markers using low level technology, which is accessible to a variety of research environments. The advantage of using RAPD's include: (1) a universal set of primers can be used for all species; (2) no probe library, radioactivity, Southern transfers, or prior primer sequence data are required; (3) only the oligoprimer

sequence is needed to use the technique; (4) the process lends itself to automation, and (5) only small quantities of DNA are needed (Williams *et al.*, 1990).

The automation of the RAPD technique has realised the potential for routine use of molecular markers in breeding programmes, something that could not be possible with RFLP technology. Many crop species that were previously excluded from detailed mapping due to insufficient research funding, are now ideal candidates for genetic studies, because the RAPD technique does not require any previous molecular characterisation (Sobral and Honeycutt, 1994b). Unfortunately, the major problems experienced in the application of the technique, have been the lack of both reproducibility and polymorphism, which appear to be less severe in species with small genome sizes (Koebner *et al.*, 1994).

Microsatellite markers (SSR's)

Microsatellites or single sequence repeats (SSR's) are highly mutable loci, which are present at many sites throughout a genome. The flanking sequences at each of these sites are often unique. Specific primers can be designed according to the flanking sequences, which then result in single locus identification. Alleles which differ in length can be resolved using agarose gels or sequencing gels where single repeat differences can be resolved and all possible alleles detected. SSR's are highly informative because they are co-dominant (unlike RAPD's and AFLP's) and generally highly polymorphic (Jones *et al.*, 1997).

Unfortunately, SSR markers are time consuming and costly to develop in that the genomic regions carrying them must be identified and sequenced. Once the primers are developed, the technique is, however, one of the most informative marker systems available. Even between closely related individuals, the number of repeat units at a locus is highly variable (Mazur and Tingey, 1995). SSR's are used to cluster lines into groupings (Liu and Wu, 1998; Senior *et al.*, 1998). SSR's occur in many plant genomes including those of maize (Echeverrigaray *et al.*, 1996; Phelps *et al.*, 1996; Bird *et al.*, 1997), soybean (Akkaya *et al.*, 1992), rice (Wu and Tanksley, 1993), and barley (Saghai-Maroof *et al.*, 1994).

The Amplified Fragment Length Polymorphism (AFLP) technique

Amplified fragment length polymorphisms (AFLP's), developed by Zabeau and Vos (1993), is a reproducible, multiplex assay with the ability to generate large numbers of polymorphic genetic loci. AFLP's is a robust and rapid technique for determining large numbers of DNA polymorphisms and is being used extensively for genetic mapping and fingerprinting in plants. Use of this technique avoids problems, which may be encountered with reproducibility, and optimisation of reaction conditions when using arbitrarily primed PCR (Money *et al.*, 1996). This technology has been used to analyse diversity in many species (Mackill *et al.*, 1996; Maughan *et al.*, 1996; Paul *et al.*, 1996; Travis *et al.*, 1996). The generation of molecular marker profiles based on AFLP's allows a retrospective analysis of the consequences of breeding and selection on the production of new lines. In addition, it can facilitate the strategic planning of new breeding approaches based on combining and selecting new combinations of genotypes to maximize the rate of line improvement (Ellis *et al.*, 1997).

Principles of the AFLP technique

The AFLP technique is based on the selective PCR amplification of restriction fragments from an endonuclease digest of genomic DNA. The technique involves three steps: (1) restriction digestion of DNA and ligation of oligonucleotide adaptors, (2) pre-selective and selective amplification of restriction fragments, and (3) gel analysis of the amplified fragments (Vos *et al.*, 1995). PCR amplification of restriction fragments is achieved by using the ligated adaptors on to restriction ends as target sites for primer annealing. The pre-selective and selective amplification is achieved by using primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites (Fig. 2.1)(Vos *et al.*, 1995).

Two restriction enzymes, a rare cutter and a frequent cutter (Fig. 2.2) generate the restriction fragments for amplification. The rationale for using two restriction enzymes is the following: (1) The frequent cutter will generate small DNA fragments which will amplify well and are in the optimal size range for separation on denaturing gels (sequence gels). (2) The number of fragments to be amplified is reduced by using the rare cutter, since only the rare cutter/frequent cutter fragments are selectively amplified. (3) Using two different restriction endonuclease gives the greatest flexibility in selecting the number of fragments to be amplified. (4) Large numbers of different fingerprints can be generated by using selective primers (Vos *et al.*, 1995).

Resolving the amplified DNA fragments on standard sequencing gels allows for the detection of amplified fragment length polymorphisms (AFLP's). Similar to RAPD's, AFLP's require no prior sequence knowledge, and is able to detect a 10-fold number more of loci (20-100) than by RAPD analysis. Thus, using AFLP analysis it is possible to rapidly screen thousands of independent genetic loci. The ability to screen large numbers of loci has specific applications in linkage mapping studies when combined with near-isogenic lines (Muehlbauer *et al.*, 1988; Young *et al.*, 1988) or bulk segregant analysis (Michelmore *et al.*, 1991b).

The powerful nature of AFLP markers

The AFLP technique combines the reliability of the RFLP technique with the high throughput of PCR. In principle, AFLP's offer all the advantages expected of RFLP's and RAPD's, but in a robust PCR regime. However, the technology is largely untried and is as yet, technically demanding (Koebner *et al.*, 1994).

The AFLP approach is powerful because it requires no prior sequence characterisation of the target genome and is readily applicable to a wide variety of crops. Additionally, it is easily standardised and readily automated for high-throughput application. Researchers can analyse thousands of polymorphic data points each day (Perkin-Elmer, 1996). AFLP technology currently offers the fastest, most reproducible, and most cost-effective way to generate high-density genetic maps for

marker-assisted selection of desirable traits. It is also the ideal tool for determining varietal identity and assessing trueness to type (Perkin-Elmer, 1996).

Optimisation of the AFLP technique

DNA isolation procedures

Plants can be grown in a variety of environments and in different locations and still provide starting material for DNA isolation (Young, 1994). Any part of a plant can be used to extract DNA. The most common starting material is young leaves. They can be either fresh, lyophilised, dried in an oven or in some cases dried at room temperature (Kochert, 1994). Several methods for DNA extraction have been developed (John, 1992; Tai and Tanksley, 1990; Dellaporta *et al.*, 1983; Murray and Thompson, 1980). These methods all have the same goals of simplicity, speed, and utilisation of a small amount of starting material (Lamalay *et al.*, 1990). Simplicity and speed are absolutely essential for processing large numbers of individuals. It is an advantage if large quantities of starting material are hard to obtain (Young, 1994).

Yu and Pauls (1994) proposed a modification of the method described by Edwards *et al.* (1991). The procedure does not require the use of expensive or hazardous chemicals and uses only milligram quantities of leaf tissue. Approximately four to eight hours are needed to process 50 to 100 samples, and each sample contains enough DNA for at least 100 PCR's. An important modification on the method of Edwards *et al.* (1991) involves the addition of the extraction buffer. In addition, Yu and Pauls (1994) used a pestle that fits tightly into an eppendorf tube and that was attached to a motor revolving at approximately 500rpm. The purity of the DNA preparation was also increased by a centrifugation step after dissolving the DNA.

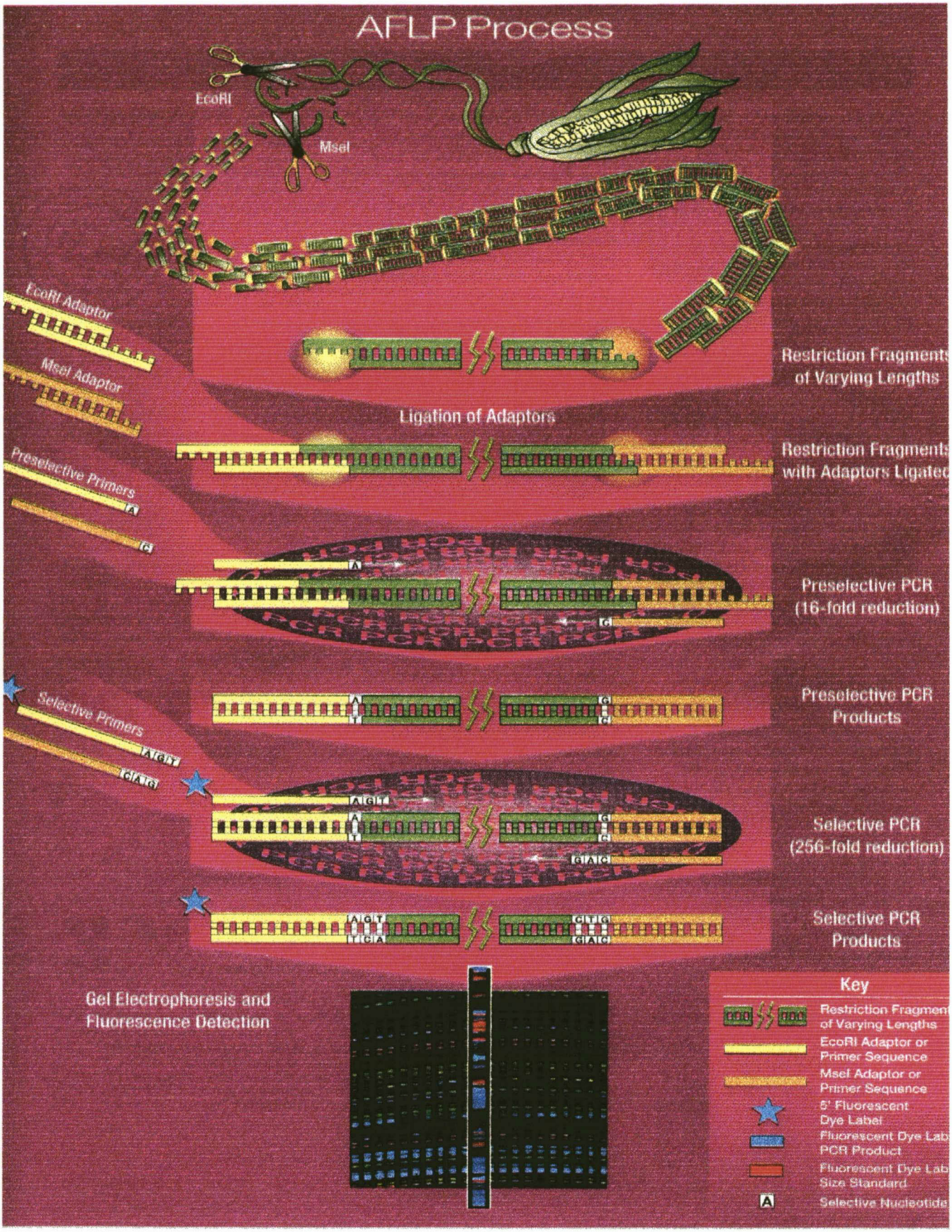


Fig. 2 1 The process of the Amplified Fragment Length Polymorphism (AFLP) technique (Perkin-Elmer, 1996).

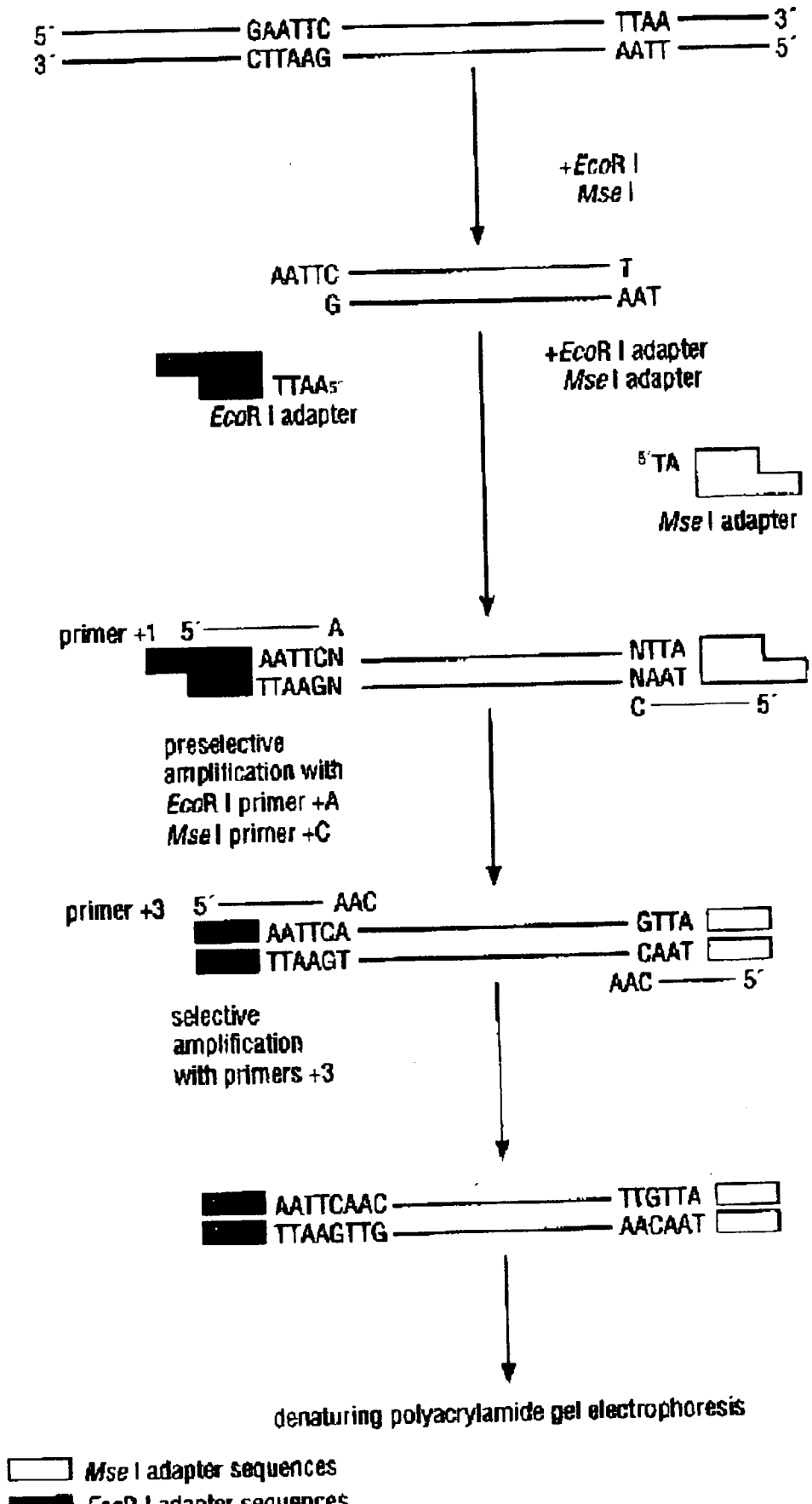


Fig. 2.2 Illustration of the principle of amplification of *EcoRI-MseI* fragments (GIBCO BRL, 1996).

Optimisation of the AFLP reaction

Several reaction components need to be optimised in order to give reproducible results. Reaction components that should be optimised include template, primer, $MgCl_2$, enzyme and dNTP concentration (Caetano-Anollés *et al.*, 1991). Optimisation of the PCR reaction usually relies on the sequential investigation of each reaction variable. This approach requires a large number of PCR reactions in order to include all the possible combinations of reaction conditions. However, in practice, the optimum reaction conditions are rarely identified. Cobb and Clarkson (1994) proposed a method in which the PCR reaction can be optimised using a modification of Taguchi optimisation methods.

AFLP primers and adaptors

AFLP primers consist of three parts, a core sequence, an enzyme specific sequence (EZ) and a selective extension (EXT) (Zabeau and Vos, 1993). AFLP-primers have enzyme-specific parts corresponding to the respective enzymes. The AFLP fingerprints show that large numbers of restriction fragments are amplified simultaneously. In general, the simultaneous amplification of DNA fragments using specific primer sets for each PCR fragments (multiple PCR) appears to be rather troublesome. Results suggest that multi-fragment amplification is efficient provided that all fragments use the same primer set for their amplification. This implies that the differences in amplification efficiency of DNA fragments in PCR, are mainly primer-associated and not fragment specific (Vos *et al.*, 1995).

DNA template preparation

In general, template DNA quantities can vary from picogram to nanogram amounts, depending on the genome size. However, if the template concentration is too high then the PCR amplification often results in a smear without distinct bands (Yu *et al.*, 1993). Hosaka and Hanneman (1994) demonstrated that certain amplified fragments show continuous increase or decrease in band intensity depending on template concentrations. This indicates that adjusting genomic DNA concentrations to the

same level in all samples may be an initial step in obtaining reproducible and comparable banding patterns.

Identification of AFLP markers linked to specific genes of interest

It is a common practice to determine mode of inheritance, of plant traits using classical genetics. In some instances the chromosomal location of the gene(s) coding for the traits is deduced by monosomic and telosomic analysis. However, for most traits, the location and characteristics of genes are unknown. Therefore, the identification of tightly linked molecular markers represents an important first step towards the molecular characterisation of plant genes (Yu and Pauls, 1994).

The identification of molecular markers for specific genes is usually a three step process that involves (1) assessing the mode of inheritance of the plant trait and the molecular marker, (2) verification of linkage between the marker and the trait through segregation analysis and (3) calculation of the recombination fraction and linkage distance (Yu and Pauls, 1994). The principle of Marker Assisted Selection (MAS) is, that indirect selection for a trait can be performed by selecting for a marker closely linked to a trait, rather than directly for the trait itself. Single traits can be targeted by the 'bulk segregant' method, where a segregating population is scored for the trait and DNA from individuals with two alternative bulked genotypes (Michelmore *et al.*, 1991a). The profiles of the two bulked genotypes are then compared with as many markers as possible. For traits with more complex genetic control an adequate complete genetic map is the primary prerequisite for successful MAS (Koebner *et al.*, 1994).

Although this study is concerned only with line identification, the applications of other molecular technologies will also be discussed, to give the broad picture of the uses of this technology.

The use of molecular markers for genetic analysis in crop plants.

Line identification

DNA fingerprints of closely related organisms are generally very similar. Differences in DNA sequence are observed as the presence or absence of bands. These differences are characteristic and heritable. The fingerprints can be used to establish the identities of individuals, or to assess the relatedness between individuals. The fingerprints can also be used as a source of DNA markers for genetic linkage maps or to identify markers linked to a specific locus (Perkin-Elmer, 1996). Examples cited include DNA fingerprinting of maize, developing herbicide and disease resistant crops, and the production of a biodegradable natural polymer by bacteria (Doyle, 1993). By using DNA markers, rather than morphological markers, breeders can select for desirable traits at the seedling stage, rather than at the adult plant stage. This can save the time of an entire growing season (Perkin-Elmer, 1996). The suitability of AFLP analysis for line identification, is demonstrated by the large number of reports recently published on the use of the technique for line identification in a variety of plant species, such as tomato, soyabean, brassicas, sunflower, pepper, sugar beet, lettuce (Perkin-Elmer, 1996) wheat (Donini *et al.*, 1997) and barley (Pakniyat *et al.*, 1997).

The use of near-isogenic lines to identify markers

Near isogenic lines (NIL's) are generated through repeated backcrossing to a target genotype and selection of the desirable trait in each backcross. The result of this process is the production of two genotypes that are in theory identical at all genetic loci except for the chromosomal region containing the gene of interest. The high probability that any polymorphism between NIL's will originate from the introgressed chromosomal segment provides a powerful approach towards identifying closely linked molecular markers (Kelly, 1995).

The size of the retained donor chromosome segment influences the efficiency of NIL analysis. The size of the donor segment specifies the maximum distance expected

between a marker identified by NIL analysis and the target gene. The shorter the donor segment, the tighter the required linkage of the marker. The size of the introgressed chromosome segment is influenced by the number of times the isogenic line was backcrossed to the recurrent parent, the position of the target gene on the chromosome and the size of the genome and target chromosome (Stam and Zeven, 1981; Zeven *et al.*, 1983). The AFLP technique has been used to generate molecular markers for specific genes of interest in a number of crop plants. These include wheat (Bohn *et al.*, 1999), barley (Ellis *et al.*, 1997), and soyabean (Maughan *et al.*, 1996).

The use of bulked segregant analysis (BSA) to identify markers

AFLP fingerprinting of bulked segregants/isogenic lines is an efficient means to identify introgression-specific DNA sequences, which can serve as a source of template for the production of STS markers. Particularly when numerous AFLP fragments are targeted, the bottleneck in the procedure is the identification of the 'correct' sequence among the heterogeneous population of clones retrieved from the amplification product of DNA excised from the gel (Koebner *et al.*, 1998).

The process of constructing NIL's is somewhat time-consuming and almost impossible in a number of vegetatively reproducing species. Michelmore *et al.* (1991b) developed an elegant method of creating 'in vitro near-isogenic lines'. The process is called bulked segregant analysis (BSA) and involves the bulking of F_2 individuals (segregants) from a cross between parents that are polarised for a specific trait. DNA from F_2 individuals, which represent the phenotypic extremes for a segregating trait, is bulked in two separate pools. The DNA pools are then screened for polymorphisms. It is assumed that the phenotypic distribution represent individuals that have opposing homozygous alleles for the trait in question. All other alleles are assumed to be in a heterozygous state because of the random distribution of these unselected alleles (Michelmore *et al.*, 1991b).

BSA can also be used to target multiple loci of highly heritable quantitative traits that are controlled by a few loci with a large phenotypic effect (Michelmore *et al.*, 1992). Yield and quality factors, as well as resistance to some biotic and abiotic stresses, are

usually controlled by polygenic systems. In addition to their biological and economical importance, these traits are very difficult to manipulate in breeding programmes due to a poor understanding of the genetic basis of their inheritance. Quantitatively determined traits are characterised by a continuous range of phenotypic variation in individuals at the extremes of the phenotypic distribution (Waugh and Powell, 1992).

BSA not only allows for rapid screening and enrichment for markers linked to specific traits, but allows the generation of genetic maps for the traits under investigation (Michelmore *et al.*, 1991b). This map can be refined by selection of appropriate progeny for further bulking and analysis, as well as individual progeny testing. These maps can then be incorporated into existing genetic maps by cross-mapping those regions in the standard mapping population (Michelmore *et al.*, 1992).

Crop plants that are of utmost importance include all the major cereals (rice, maize, wheat, barley, sorghum, millet, rye) and among arable crops, an ever-growing number of leguminous (soybean, pea, bean), solanaceous (potato, tomato, pepper), brassicaceous (oilseed rape) and other (sugar beet, lettuce) species (Mackill *et al.*, 1996; Paul *et al.*, 1996; Powell *et al.*, 1996; Tohme *et al.*, 1996). The incorporation of genes of agronomic importance into these plants depends critically on both the ability to identify precisely alternative alleles in a segregating population at the target locus, i.e. that the trait is qualitative, and that the population is adequately polymorphic with respect to the markers being used (Koebner *et al.*, 1994). The potential of MAS and BSA in agriculture are considerable. Selection of genotypes can be made at very early stages, and often at a juvenile age. This has implications for breeding programs, as only relatively small populations of segregating material need to be grown to maturity. Furthermore, the interaction between genotype and environment is irrelevant for those traits selected by MAS. This means that, in theory, generations can be advanced in off-season conditions or in locations where the environment would otherwise be a limitation. A further benefit is that individuals carrying a favorable recessive allele can be selected at each backcross generation, without the need for a progeny test, as is required in conventional backcross programs (Koebner *et al.*, 1994).

Parental selection and germplasm surveys

Parental selection is one of the most critical aspects of breeding. In general, the probability of achieving transgressive segregation from a cross (or maximum heterosis in a F1 hybrid) increases the more diverse the parents become. Genetic markers provide an objective method for assessing this divergence, since they can be used to sample polymorphisms across the entire genome, rather than just a few loci, as in the case of phenotypic comparisons. Similar considerations relate to germplasm surveys and the management of gene banks. Genetic markers may be useful in these contexts to minimise duplication, thereby allowing a more rational use of scarce germplasm storage capacity (Koebner *et al.*, 1994).

The coefficient of parentage provides a statistical expectation, based on pedigrees, of the parental contributions to a line or population. The coefficient of parentage is calculated from pedigree data but is subject to errors stemming from deviations between expected and actual parental contribution and incorrect pedigree information. This coefficient cannot be determined when pedigree records are unavailable. Molecular marker loci that are neutral and easily identified offer the opportunity to track parental genetic contributions directly. Furthermore, the potential abundance of molecular markers makes them ideal for use in determining genotype pedigree (Van Toai *et al.*, 1997).

Genetic distance analysis of molecular markers

Two approaches are generally used to deduce phylogenetic relationships from fingerprinting data. In the first approach, parsimony analysis, phenograms representing phylogenetic relationships are constructed on the basis of the lowest number of character state transformations that yields a particular phenogram. The software program PAUP (Phylogenetic Analysis Using Parsimony)(Swofford, 1991) and PHYLIP (Phylogeny Inference Package)(Felsenstein, 1993) is the most widely used for this purpose and calculates the most parsimonious tree directly from an input data matrix.

The second widely used approach involves the cluster analysis of pairwise genetic distances for the construction of dendograms. Pairwise genetic distances are calculated directly from an input data matrix containing presence (1) or absence (0) values for all markers. Each marker is considered a unique allele. Generally, the fraction of alleles shared between any two individuals is used to calculate a similarity coefficient (S) that is converted to a genetic distance value using either:

$$D = 1-S$$

or

$$D = -1n(S)$$

This process is repeated for all the possible pairwise groupings of individuals and the pairwise distance values tabled in a pairwise distance matrix. A number of formulas have been used to calculate pairwise distance matrixes from discrete data. The index of genetic similarity (F) of Nei and Li (1979), although initially developed for the analysis of RFLP data, is commonly used to analyse molecular data. The formula:

$$F = 2n_{xy}/(n_x + n_y)$$

where F is the ratio of shared bands/alleles between individuals x and y, $2n_{xy}$ is the number of shared alleles, and n_x and n_y are the total number of alleles observed in each of the individuals. The F value is used to calculate the distance value (D = 1-F) for every pairwise combination.

Development of genetic maps based on AFLP markers

Since the inception of the AFLP technique, the development of genetic maps for agriculturally important crops has become priority. Plant species that have already been mapped by this approach include potato (Rouppe van der Voort *et al.*, 1998), soybean (Keim *et al.*, 1997), *Astragalus cremnophylax* var. *cremnophylax* (a critically endangered plant)(Travis *et al.*, 1996.), barley (Simons *et al.*, 1997) and tomato (Thomas *et al.*, 1995). Many other crop plants are in the process of being mapped in

this way. The research that is being done underlines the importance of this technology in plant biotechnology.

Advantages of molecular markers over morphology

Molecular DNA markers such as AFLP's offer numerous advantages over markers traditionally used in plant mapping and selective breeding. In the past, breeders have selected among closely related strains on the basis of morphological markers that are readily observable and that are co-inherited with the desired trait. Because of the scarcity of such markers, this approach often fails to discriminate between closely related lines that differ for the trait of interest (Perkin-Elmer, 1996).

Even when a useful morphological marker is identified, its application in breeding programmes is time consuming and costly. Typically, an entire field of a particular genotype must be grown and analysed. The task of examining the myriad of individual plants is arduous and time-consuming (Perkin-Elmer, 1996). Consequently, breeders are abandoning this traditional approach in favor of faster, highly discriminating and less costly, approach of using molecular markers such as RFLP's, RAPD's, SSR's and AFLP's (Perkin-Elmer, 1996).

Comparison of major molecular systems

The marker systems that are currently available, and there relative advantages and disadvantages, are summarised in Table 2.1.

During the past few years, the use of polymerase chain reaction-based techniques has significantly increased the application of DNA markers to genotyping, genome mapping and phylogenetics. The advantage of these techniques over other DNA marker techniques includes the detection of a large number of polymorphisms from a single experiment, thus reducing expense (Van Toai *et al.*, 1997). AFLP and RAPD technology is PCR-based and has the advantage over RFLP's, as it is readily automatable. Unlike RAPD's, AFLP's have proven to be more reproducible.

Furthermore, AFLP analysis requires no prior sequence knowledge of the target genome and, therefore, has no prior characterisation costs (Perkin-Elmer, 1996).

Table 2.1 Comparison of major marker systems (Langridge and Chalmers, 1998)

Marker system	Loci/ assay	DNA amount	Time/ assay	% bands poly	Informative value	Comments
RFLP	3	5.0μg	5 days	10	0.3	Dominant, reliable
SSR	1	0.2μg	5 hrs	10	0.1	Dominant, reliable
AFLP	50	0.2μg	1 day	10	5.0	Mostly recessive, reliable
RAPD	10	0.2μg	5 hrs	20	2.0	Recessive, less reliable

The AFLP approach provides 10-fold more markers on average than do the other molecular approaches. This translates into 10-100 times denser linkage maps. AFLP's also tend to be more informative than RFLP's, RAPD's, or SSR's (Table 2.2). AFLP markers offer considerable advantages in terms of both the number of informative data points generated per sample reaction and the number of markers placed in a genome (Perkin-Elmer, 1996).

Using the AFLP approach allows the specific amplification of numbers of restriction fragments. However, the number of fragments that can be analysed simultaneously, is dependent on the resolving power of the detection system. The AFLP technique provides a novel and very powerful DNA fingerprinting technique for DNA's of any origin or complexity (Vos *et al.*, 1995).

AFLP's were initially difficult to perform by most groups, whose first attempts produced profiles in which over 50% of bands were not reproducible. However, with greater familiarity, these problems were resolved and the AFLP profiles subsequently obtained showed extremely high reproducibility. This confirms previous reports on the reproducibility of AFLP profiles (Jones *et al.*, 1997).

The data in Table 2.1 represents averages for these marker systems screened against commercial varieties. Although Table 2.1 suggests that RAPD markers are the most

efficient, the poor reliability of these markers and the inability to transfer them between crosses has greatly limited their use (Langridge and Chalmers, 1998). The crucial question in all aspects of marker implementation is cost. The ramifications of this question are the value of the trait or traits being monitored, the availability and speed of alternative screening techniques and the likely segregation pattern of the trait. It is not feasible to use markers to select complex traits that are determined by five or more genes unless special strategies are introduced to keep the effective population size small. The efficient implementation of markers will depend upon devising new breeding strategies. The implementation of markers in breeding programmes does not simply lie in replacing existing procedures but devising new breeding strategies (Langridge and Chalmers, 1998).

CHAPTER 3

MORPHOLOGICAL CHARACTERISATION OF MAIZE INBRED LINES.

Abstract

Sixty eight maize (*Zea mays* L.) inbred lines (both flint and dent) from the Advanta Africa genebank, representing five countries, were evaluated for genetic diversity to select germplasm that can be used in breeding programmes to improve hybrid development. The objective of this study was to evaluate the inbred lines, using 25 morphological characters. The morphological characters were coded into a discrete binary data matrix based on the presence and absence of character states. The inbred lines were grouped, using UPGMA and Neighbour joining, into clusters. The results indicated that the grouping of inbred lines are correlated with origin. Morphology was found to be a costly method to determine genetic relatedness.

Introduction

Germplasm curators as well as plant breeders have an interest in quantification and classification of genetic diversity. In germplasm collections, such information is used to designate core collections and enhance efficiency of collection management and utilisation (Brown *et al.*, 1987). Van Beuningen and Busch (1997), two U.S. maize (*Zea mays* L.) breeders, developed a new classification system using heterotic grouping. This method of classifying maize collections has made a significant contribution to maize breeding.

The recognition and exploitation of heterotic patterns among genetically divergent groups of germplasm are fundamental in hybrid breeding (Hallauer *et al.*, 1988).

Classification of elite germplasm into heterotic groups and assignment of new inbred lines to establish heterotic groups are major decisions in any maize breeding programme. For these reasons maize breeders have a keen interest in the characterisation of genetic diversity among and within existing heterotic groups as well as in the relationships between current and historically important lines (Messmer *et al.*, 1992).

More recently, an interest in the quantification of genetic diversity of pure line crops has developed. Transgressive segregation may be more likely to occur when parents in a cross are less similar, allowing different favourable alleles to be combined in the offspring (Cowen and Frey, 1987). The perspective that hybrid lines and consequently the detection of heterotic patterns may become important for grain crops adds to the importance of quantification of genetic distances and relatedness patterns in these crops (Cregan and Busch, 1978; Cox and Murphy, 1990).

Traditionally, distance estimation and classification were based entirely on morphological markers and quantitative traits (Goodman, 1972). Distance estimations and classification was greatly enhanced in the last decade with the introduction of novel genetic marker systems (Keim *et al.*, 1997). Recent studies have focused on genetic markers in the form of gene products such as isozymes (Messmer *et al.*, 1991) and direct DNA markers such as Restriction Fragment Length Polymorphisms (RFLP's) (Tanksley, 1983), Random Amplified Polymorphic DNA (RAPD's) (Wang *et al.*, 1996), microsatellites (SSR's) (Phelps *et al.*, 1996) and Amplified Fragment Length Polymorphisms (AFLP's) (Ellis *et al.*, 1997).

An assumption underlying the use of phenotypic similarity estimates based on quantitative and qualitative traits is that such estimates are an accurate reflection of genotypic similarity (Van Beuningen and Busch, 1997). Morphological characterisation is useful in breeding programmes because: (i) existing databases on germplasm collection or breeders crossing block entries can often be used for genetic analysis, (ii) statistical procedures for processing morphological characterisation are readily available, (iii) morphological information adds to an understanding of ideotype-performance relations and (iv) explanations of heterosis may be enhanced by adding morphological measures of distance as another independent variable (Cox and

Murphy, 1990). Therefore, the grouping of lines based on morphological similarities will enable breeders to make better decisions about which crosses will generate higher combining ability. The reason is that the further the genetic distance between two individuals, the more likely to expect a better combining ability.

The aim of this study was to determine clustering patterns of a large collection of maize inbred lines on the basis of morphology and plant development. Furthermore, to evaluate the relationship between distance based on quantitative and qualitative morphological traits by dividing them in heterotic groups for later molecular genetic analysis. The data would then be a starting point for breeders to do better planning for field trials to get the maximum benefit from the growing season.

Materials and Methods

Lines

The plant material used in this study was obtained from the Advanta group (Table 3.1). Out of a total of 68 lines studied (both flint and dent), 36 lines were from South Africa (local), three from Zimbabwe (CIMMYT), three from Argentina, 12 from Garst (USA) and six from Thailand. Each lines name and country of origin is summarised in Table 3.1. (Arbitrary codes were assigned to lines to protect intellectual property rights and confidentiality.) Seed was obtained from the Advanta grain germplasm depository at Petit, South Africa and represents a broad range of diversity in origin. All 68 inbred lines are currently used in the maize-breeding programme at Bapsfontein, South Africa, with the aim of making crosses between lines, which have the best combining ability. Morphological data from this material was collected at Bapsfontein Research Station during the 1998/1999 seasons.

Geographical location and field design

The average minimum temperature of the area is 8.9°C and the average maximum temperature is 28.7°C. The relative humidity is approximately 60%. The seasonal rainfall is approximately 700mm, lasting from September to April and is inconsistent.

It is classified as a summer rainfall- and semi-arid area (Weather Bureau, JHB International).

Table 3.1 Maize lines used for morphological characterisation

Nº	Lines	Origin	Nº	Lines	Origin
1	YO12	Argentina	35	YM2	Local
2	YO13	Argentina	36	YM3	Local
3	YO14	Argentina	37	YI1	Local
4	YS1	Brazil (tropical)	38	YI2	Local
5	WO8	Brazil (tropical)	39	YI3	Local
6	WO9	Brazil (tropical)	40	YR1	Local
7	WO10	Brazil (tropical)	41	YS3	Local
8	WO11	Brazil (tropical)	42	YF4	Local
9	WO5	(CIMMYT) Zimbabwe	43	WO1	Local
10	WO6	(CIMMYT) Zimbabwe	44	WF1	Local
11	WO7	(CIMMYT) Zimbabwe	45	WF2	Local
12	YM4	GARST (USA)	46	WI2	Local
13	YL2	GARST (USA)	47	WP3	Local
14	YL3	GARST (USA)	48	WK1	Local
15	YR2	GARST (USA)	49	WK2	Local
16	YR3	GARST (USA)	50	YO8	Local
17	YR5	GARST (USA)	51	YO9	Local
18	YL1	GARST (USA)	52	YO10	Local
19	YL4	GARST (USA)	53	YO16	Local
20	YL5	GARST (USA)	54	YO17	Local
21	YO11	GARST (USA)	55	WM1	Local
22	WR1	GARST (USA)	56	WM2	Local
23	WR2	GARST (USA)	57	WO3	Local
24	YO1	Local	58	WI1	Local
25	YO2	Local	59	WP1	Local
26	YO3	Local	60	WP2	Local
27	YF1	Local	61	WO2	Local
28	YF2	Local	62	WO4	Local
29	YF3	Local	63	YR4	Thailand
30	YO4	Local	64	YR6	Thailand
31	YO5	Local	65	YR7	Thailand
32	YO6	Local	66	YS2	Thailand
33	YO7	Local	67	YS4	Thailand
34	YM1	Local	68	YO15	Thailand

Y = yellow, W = white, M = M heterotic group, F = F heterotic group, I = I137TN, R = Reid, L = Lancaster, S = Suwan, P = Potch Pearl, K = K64R, O = Other heterotic groups

The collection was planted at Bapsfontein Research Station in the last week of October and harvested in May of the following year. Each plot consisted of a single

5.7m long row, with 0.96m-row spacing. Plots were planted at 30 kernels per plot row and thinned to 20 plants per plot. This produced an average plant density of approximately 30 000 plants per ha. Two guard rows were planted around the trials, but not between plots. Weeds, insects, and foliar diseases were controlled as follows: Lamba-cyhalothrin (Pyretroid)(6g $aiha^{-1}$)(insecticide), Sulcotrione, Atrazine and other Triazines (200g $aiha^{-1}$, 465,6g $aiha^{-1}$, 14,4g $aiha^{-1}$)(herbicide) and Acetochlor (Chloroacetanilide)(840g $aiha^{-1}$)(herbicide) were applied at planting, followed by 6g $aiha^{-1}$ Lamba-cyhalothrin (Pyretroid) and 200g $aiha^{-1}$ Sulcotrione after planting. Asetochlor (Chloroacetanilide)(840g $aiha^{-1}$)(herbicide) was applied once after planting. Supplementary irrigation was applied during the growing season.

Morphological parameters

A total of 25 morphological and developmental characters were scored for all entries during the growing season. The plants were maintained and self-pollinated. The data was collected from five individual plants from the middle of each row and the averages were determined. Three to four months after planting the following traits were measured:

1. Plant height (m): Height from the base to the lowest tassel branch determined from five plants that were visually judged to represent an average of each plot.
2. Ear height (m): Height from the base to the ear node determined from five plants that were visually judged to represent an average of each plot.
3. Ear prolificacy: The number of ears per plant, to represent an average of the five middle plants of the plot. (With a plant density of 30000 and the narrow row width, ear prolificacy may have been underestimated.)
4. Tillers: The presence or absence of tillers in the plot.
5. Tassel colour: The observation of colour variations ranging from cream to purple.
6. Silk colour: The observation of colour variations ranging from white to purple.
7. Flowering date: The number of days from planting to 50% pollen shed.
8. Leaf type: The direction (upwards or hanging downwards) of leaf growth.
9. Rust: Scoring on a scale of 0-9, where 0 is susceptible and 9 is resistant.
10. Stalkborer: Scoring on a scale of 0-9, where 0 is susceptible and 9 is resistant.

11. Germination (%): The number of plants that germinated, as a percentage of the number of kernels planted.

At harvest time, plant and ear characteristics were recorded as follows:

12. Percentage root lodging: The number of plants leaning more than 30° from the vertical at harvest divided by the total stand of each plot.
13. Mass per ear (g): The average mass per ear of the selected ears.
14. Ear length (cm): The length from the base to the tip of the ear, determined from five plants to represent the average of each plot.
15. Ear width (cm): The width in the middle of the ear, determined from five plants to represent the average of each plot.
16. Kernel colour: The intensity of white or yellow respectively.
17. Kernel rows: The number of kernel rows at the base of the ear, determined from five plants to represent the average of each plot.
18. Diplodia ear rot: The presence or absence of Diplodia ear rot.
19. Low ear: If the height from the base to the ear node is low enough to cause ear damage by rodents, birds etc., or not.
20. Wide inter row spacing: The presence or absence of wide kernel rows.
21. Kernel abortion: The presence or absence of aborted kernels.
22. Tapering ears: The presence or absence of an ear that becomes narrower from the base to the tip of the ear.
23. Grain type: Flint or dent type.
24. Cap: Presence or absence of a cap (starchy endosperm).
25. Cob colour: The observation of colour variations ranging from white to purple.

Computer analysis

For qualitative characters, presence (0) or absence (1) was scored and summarised in an input matrix (Table 3.6). For quantitative characters, averages were calculated and individual values scored for being above (1) or below (0) the average. All subsequent calculations were performed using the computer programme PHYLIP (Phylogeny Inference Package)(Felsenstein, 1993).

The index of genetic similarity (F) of Nei and Li (1979) was used to calculate pairwise genetic distance (D) for all lines:

$$F = 2N_{xy}/ (N_x + N_y)$$

$$D = 1-F$$

where N_{xy} = the number of shared traits between any two lines x and y and N_x and N_y are the number of traits for lines x and y, respectively (Wang and Tanksley, 1989). Unweighted Pair Group Mean Arithmetic Analysis (UPGMA)(Sneath and Sokal, 1973) and Neighbour joining methods were used for cluster analysis of the pairwise distance matrix representing the distances among the maize lines, based on the similarity matrix. To estimate the reliability of the analysis, a thousand bootstrap replicates were performed and a consensus tree was drawn.

Results

Morphological data

Morphological data are provided in Tables:

Table 3.2 Morphological traits scored three to four months after planting, for white inbred lines

No	Line	Plant height (m)	Ear height (m)	Ear prolificacy	Tillers	Tassel color	Silk color	FD (days)	Leaf Type	Rust	Stalk borer	Germination %
1	WM1	1.39	0.50	2-3	0	w/p	p	96	upright	8	7	56.67
2	WM2	1.83	0.65	2	0	cream/p	p	101	floppy	7	8	56.67
3	WF1	1.70	0.55	1-2	0	cream/p	w	105	floppy	7	8	73.33
4	WF2	1.15	0.38	1-2	0	pur	w	94	upright	7	8	56.67
5	WI1	1.67	0.50	1	0	o/y	w	102	upright	7	8	26.67
6	WI2	1.40	0.44	1	0	p	p/w	97	upright	7	9	60.00
7	WP1	1.60	0.80	2	0	w	w	109	upright	7	8	50.00
8	WP2	1.60	0.73	1-2	0	p	w	111	upright	8	8	60.00
9	WP3	2.03	0.75	1	0	w/p	w	103	upright	6	8	80.00
10	WR1	1.67	0.65	1	0	o/p/w	w/p/pur	79	upright	6	9	63.33
11	WR2	1.64	0.60	1	0	w	w	79	upright	5	8	80.00
12	WK1	1.41	0.50	1-2	0	w	w	103	upright	8	8	63.33
13	WK2	1.68	0.60	2	0	w/p	w/p	106	upright	8	8	46.67
14	WO1	1.50	0.50	1-2	0	w	w/p	97	floppy	7	7	83.33
15	WO2	1.83	0.60	2	0	p/pur	w	102	upright	6	8	40.00
16	WO3	1.87	0.80	2-3	0	p	p	113	floppy	6	8	66.67
17	WO4	1.67	0.60	2	0	w/o	w/p	100	floppy	8	7	60.00
18	WO5	1.50	0.57	0-1	0	o/p	w/p	123	upright	8	8	80.00
19	WO6	1.54	0.60	1	0	w	pur	105	upright	8	9	76.67
20	WO7	1.40	0.70	1	0	w	w	102	upright	7	8	93.33
21	WO8	1.50	0.65	1-2	0	p/w		115	floppy	6	8	36.67
22	WO9	1.60	0.70	1-2	0	w	w	94	floppy	5	9	80.00
23	WO10	2.30	0.90	1	0	w	p	109	floppy	7	8	86.67
24	WO11	1.67	0.66	1-2	0	o/p	p/pur	111	upright	8	8	46.67

FD = flowering date, w/p = white/pink, pur = purple, o/y = orange/ yellow, p = pink, w = white, p/pur = pink/purple

Table 3.3 Morphological traits scored three to four months after planting, for yellow inbred lines

No	Line	Plant height (m)	Ear height (m)	Ear prolificacy	Tillers	Tassel color	Silk color	FD (days)	Leaf Type	Rust	Stalk borer	Germination %
1	YM1	1.50	0.50	2	0	w	w/p	98	floppy	8	8	43.3
2	YM2	1.55	0.70	2-3	0	w	p	89	floppy	7	6	73.3
3	YM3	1.66	0.59	2	1	o	w/p	90	floppy	8	5	53.3
4	YM4	1.56	0.50	1	0	w	w/p	89	upright	7	9	76.7
5	YF1	1.43	0.41	1-2	0	w	w	94	floppy	8	8	73.3
6	YF2	1.17	0.57	2	0	w/y	w	79	floppy	7	6	73.3
7	YF3	1.22	0.34	1-2	0	w	w	94	floppy	8	8	60.0
8	YF4	1.76	0.53	2	0	o	w	93	floppy	8	7	63.3
9	YI1	1.30	0.42	1-2	1	w	w	80	upright	8	6	56.7
10	YI2	1.39	0.40	1	0	w	w	89	upright	8	9	73.3
11	YI3	0.80	0.20	1	0	w	w	97	upright	8	8	70.0
12	YR1	1.38	0.40	1	0	y	w	89	floppy	6	5	43.3
13	YR2	1.74	0.61	1-2	0	w	w/p	85	upright	8	6	73.3
14	YR3	1.25	0.40	1	0	o/p	w/p	82	upright	5	7	70.0
15	YR4	1.30	0.50	1	0	o/y	w/p	86	upright	7	8	76.7
16	YR5	1.37	0.48	1	0	w	w	79	upright	5	6	63.3
17	YR6	1.40	0.40	1	1	w	w/p	89	floppy	6	7	70.0
18	YR7	1.55	0.50	1-2	0	w	w/p/pur	78	upright	4	7	33.3
19	YL1	1.44	0.60	1	0	w	w	78	upright	7	8	53.3
20	YL2	1.33	0.49	1	0	w	w	82	floppy	8	8	56.7
21	YL3	1.55	0.55	1	0	w/p	w/p	78	upright	7	8	46.7
22	YL4	1.38	0.50	2	0	y/o	w/p	79	upright	7	7	73.3
23	YL5	1.30	0.60	1-2	0	w/p	w/p	85	upright	7	7	46.7
24	YS1	1.50	0.50	1	0	o/y	w	95	upright	8	8	63.3
25	YS2	1.28	0.50	1	0	w	w/p	84	upright	6	8	66.7
26	YS3	1.00	0.38	1	0	w	w	95	upright	8	7	46.7
27	YS4	1.20	0.50	2	0	w	w	89	floppy	7	7	53.3
28	YO1	1.47	0.48	1	0	w	w	91	floppy	7	8	83.3
29	YO2	1.34	0.40	1	0	y/o	w	89	upright	6	8	40.0
30	YO3	1.53	0.64	1	0	o	w	79	upright	7	6	73.3
31	YO4	1.46	0.60	1	0	w/y	w/p/pur	80	upright	7	8	73.3
32	YO5	1.44	0.55	1	1	w	w	74	upright	4	6	53.3
33	YO6	1.40	0.45	1	0	w	w	91	upright	8	8	76.7
34	YO7	1.62	0.64	1	0	w	w	94	floppy	8	7	63.3
35	YO8	1.21	0.40	1	0	w/y	w/p	94	upright	8	8	76.7
36	YO9	1.60	0.60	1	0	o	w/p	93	floppy	8	7	73.3
37	YO10	1.40	0.45	1	1	o	w	85	floppy	7	8	53.3
38	YO11	1.35	0.38	1	0	w	w	77	upright	6	7	50.0
39	YO12	1.18	0.48	1	0	w	w	85	upright	6	8	70.0
40	YO13	1.44	0.60	1	0	pur	pur	84	upright	6	8	50.0
41	YO14	1.34	0.40	1	0	w/o	w/p	84	upright	8	8	60.0
42	YO15	1.04	0.30	1	0	w	w/p	93	upright	6	8	63.3
43	YO16	1.92	0.80	2 (big)	0	w	w/p	99	floppy	7	7	20.0
44	YO17	1.53	0.50	1	0	w	w	93	floppy	7	7	66.7

FD = flowering date, w/p = white/pink, pur = purple, o/y = orange/ yellow, p = pink, w = white, p/pur = pink/purple

Table 3.4 Morphological traits scored at harvest, for white inbred lines

No	Line	% Root lodging	Mass per ear (g)	Ear length (cm)	Ear width (cm)	Kernel color	Kernel rows	Diplodia ear rot	Low ear	Wide inter row	Kernel abortion	Tapering ears	Grain type	Cap	Cob color
1	WM1	0	42	150	45	off white	10	0	0	0	0	0	dent	0	white
2	WM2	0	87	160	50	lemon	14	1	1	1	1	1	flint/dent	0	white
3	WF1	0	62	120	40	off white	10	1	0	0	0	0	flint	0	white
4	WF2	0	32	140	45	cream	10	0	1	1	0	1	dent	1	white
5	WI1	0	54	165	45	lemon	12	0	1	0	1	0	dent	0	white
6	WI2	0	31	145	45	off white	10	0	0	0	1	0	dent	1	white
7	WP1	0	19	100	35	white	10	1	1	0	1	1	flint	0	white
8	WP2	0	19	100	40	off white	10	0	0	0	1	0	flint	0	white
9	WP3	0	23	120	45	opaque	10	1	0	1	0	0	dent	0	white
10	WR1	0	100	140	50	white	16	1	0	0	1	0	dent	1	white
11	WR2	0	42	150	45	white	10	0	1	0	0	0	dent	1	white
12	WK1	5	60	145	50	off white	12	1	0	1	1	0	dent	0	white
13	WK2	0	86	180	45	white	12	0	0	0	1	0	dent	0	white
14	WO1	0	64	145	50	white	12	0	0	0	0	1	flint	0	white
15	WO2	0	99	190	50	lemon	12	1	0	1	1	0	flint/dent	0	white
16	WO3	0	29	130	45	off white	12	0	0	0	1	1	dent	0	white
17	WO4	0	39	120	45	white	10	0	0	0	0	0	flint	0	white
18	WO5	0	35	160	50	white	10	1	0	1	1	1	flint	0	white
19	WO6	0	94	165	50	off white	12	0	1	1	1	0	dent	1	white
20	WO7	0	53	150	50	off white	14	1	0	0	1	1	flint	0	white
21	WO8	0	61	140	40	white	14	1	1	0	1	0	flint/dent	1	white
22	WO9	0	32	85	40	white	12	0	0	1	0	0	dent	1	white
23	WO10	0	97	180	50	white	14	0	0	0	1	0	flint/dent	1	white
24	WO11	0	50	155	45	white	14	0	0	0	1	0	dent	1	white

Table 3.5 Morphological traits scored at harvest, for yellow inbred lines

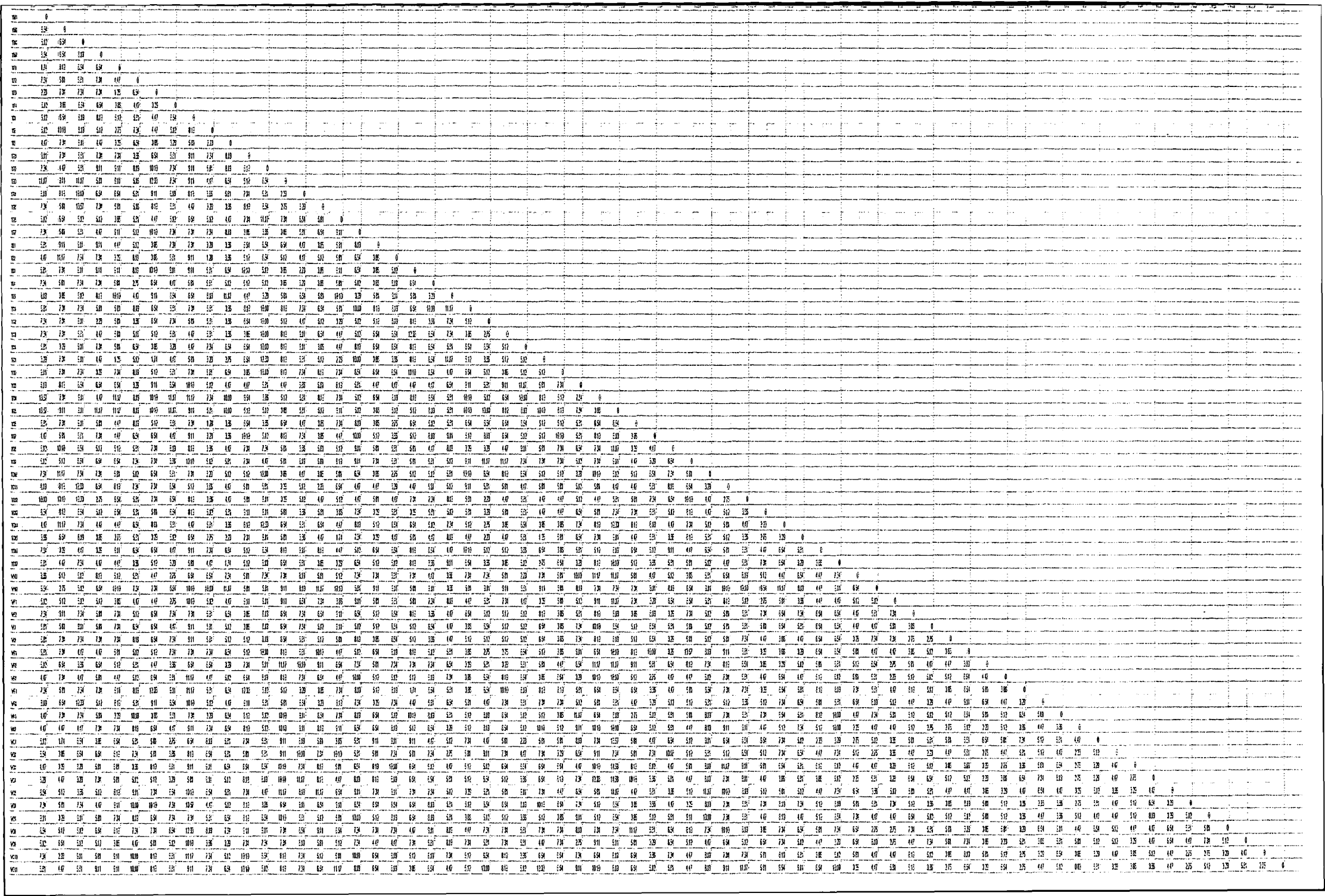
No	Line	% Root lodging	Mass per ear (g)	Ear length (cm)	Ear width (cm)	Kernel color	Kernel rows	Diplodia ear rot	Low ear	Wide inter row	Kernel abortion	Tapering ears	Grain type	Cap	Cob color
1	YM1	8	30	60	45	yellow	12	1	0	0	0	0	dent	0	white
2	YM2	0	48	135	50	pale yellow	14	0	0	0	1	1	flint	0	white
3	YM3	0	20	120	40	orange	12	1	0	1	1	1	dent	0	white
4	YM4	13	45	120	35	orange	10	0	1	1	0	1	flint	0	white
5	YF1	9	36	125	40	yellow	12	1	0	0	0	0	dent	0	white
6	YF2	0	22	125	35	yellow	10	0	1	0	0	1	dent	0	white
7	YF3	28	24	135	30	yellow	10	1	0	0	0	0	flint/dent	0	white
8	YF4	0	56	140	40	yellow	12	0	0	0	0	0	flint	0	white
9	YI1	6	38	110	45	yellow	14	0	0	0	0	1	flint	0	white
10	YI2	0	58	100	40	yellow	10	1	0	1	0	0	dent	1	white
11	YI3	0	16	110	40	yellow	12	0	0	0	0	0	flint/dent	1	white
12	YR1	0	29	150	35	light yellow	12	0	1	1	1	1	dent	0	red
13	YR2	0	61	135	45	dark yellow	12	1	0	1	1	1	dent	1	red
14	YR3	0	42	110	40	yellow	14	0	1	1	0	0	dent	1	red
15	YR4	0	42	130	35	yellow	14	1	0	0	0	1	dent	1	pink
16	YR5	0	50	105	45	yellow	14	0	0	0	0	0	dent	1	white
17	YR6	0	36	100	40	orange	10	1	0	0	0	0	flint	0	white
18	YR7	0	47	120	45	yellow	12	0	1	1	1	1	dent	0	red
19	YL1	0	28	135	40	yellow	10	1	0	0	0	0	dent	1	white
20	YL2	0	78	130	40	yellow	10	1	0	1	0	0	dent	1	red
21	YL3	0	49	100	50	yellow	14	1	0	0	0	0	dent	1	red
22	YL4	0	26	160	40	yellow	12	0	0	0	0	1	dent	1	white
23	YL5	0	62	150	45	yellow	12	0	0	1	1	1	dent	0	white
24	YS1	0	25	130	45	orange	16	1	1	0	1	0	flint	0	white
25	YS2	0	33	110	35	orange	12	0	1	0	0	1	flint	0	white
26	YS3	0	34	115	35	orange	10	0	1	0	1	1	flint	0	white
27	YS4	0	46	140	55	orange	16	0	0	0	0	0	flint	0	white
28	YO1	12	16	95	40	yellow	12	1	0	0	0	0	flint/dent	0	white
29	YO2	17	45	140	40	orange	12	0	0	1	1	1	flint	0	white
30	YO3	0	20	125	35	yellow	10	0	0	1	0	1	dent	1	red
31	YO4	5	44	150	40	orange	12	1	1	1	1	1	flint	1	red
32	YO5	0	75	145	40	yellow	14	1	0	1	1	0	dent	1	pink
33	YO6	0	71	170	50	yellow	12	1	0	1	0	1	dent	1	white
34	YO7	0	45	130	45	orange	12	1	0	0	0	0	dent	1	white
35	YO8	0	24	95	45	orange	12	1	1	0	0	1	dent	1	white
36	YO9	5	55	135	45	light yellow	10	0	0	1	0	0	dent	1	white
37	YO10	0	48	125	40	yellow	12	0	0	1	0	0	dent	1	white
38	YO11	0	42	125	45	yellow	10	0	0	1	0	0	dent/flint	1	red
39	YO12	0	50	125	40	orange	12	0	1	1	0	0	flint	0	red
40	YO13	0	57	130	40	orange	10	1	0	0	0	0	flint	0	white
41	YO14	0	32	115	35	orange	12	0	0	0	0	0	flint	0	white
42	YO15	0	40	100	35	dark yellow	10	1	0	0	0	0	flint/dent	0	white
43	YO16	0	74	185	40	orange	12	0	0	1	0	1	flint	0	white
44	YO17	0	30	130	40	dark yellow	12	0	0	0	0	1	dent/flint	1	white

Phenotypic analysis

The data matrix of morphological markers scored for genetic distance analysis is provided in Table 3.6 and 3.7.

Table 3.6 Input data matrix used for UPGMA and Neighbour Joining analysis of 25 morphological markers

YM1	11111211212222222221212222	YO8	21122222211221222211212211
YM2	21111211122211121121121121	YO9	11111222111212211121221221
YM3	21121221111212111222211222	YO10	22112122211112211222222122
YM4	11121212222211212122222111	YO11	22212112221212211221221112
YF1	11112111212222222222212221	YO12	22222112222221221212222111
YF2	22112121112222122212221121	YO13	22122222122212222222212112
YF3	11112111222222222122212222	YO14	12122222222222222222222212
YF4	21111121222212222122221221	YO15	21112212222222222222212111
YI1	12112111222121222211212121	YO16	21121212122211211222211122
YI2	21112112211212212222212211	YO17	21111112221222122222221121
YI3	21111112221222222222222211	WM1	21122221212212221212212222
YR1	21212122212222111112221122	WM2	21111221122211111111112122
YR2	22211211111211111212112211	WF1	21121121122212222222212121
YR3	22212222211212212121211111	WF2	21122121211221211111222112
YR4	22212222211212122222112111	WI1	21111122212212121111222112
YR5	22112112211212222211211111	WI2	21122222211222121212222112
YR6	21122212222212222222211121	WP1	21122221222221122122122112
YR7	22211211212211111212221112	WP2	21121121122222122222222212
YL1	22112112111222222122212112	WP3	21121122112222221221212121
YL2	21212112211212212222212222	WR1	22121222111212122211112111
YL3	22211222111212222211112112	WR2	22121112111212222111222111
YL4	22112221211222122122221111	WK1	11122111212221211121212211
YL5	22112221112211111212211112	WK2	21121221112212121212222212
YS1	21121122222222122111122211	WO1	21121211222211222121221121
YS2	22122212222222122212222111	WO2	21111121122212111121222112
YS3	21122112222221122122221212	WO3	21121221112221122221222121
YS4	21122111222212222121121122	WO4	21121221122222222221221222
YO1	11112112222222222222212121	WO5	21121222122221111212212211
YO2	11122122222211111222222112	WO6	21121212111212111111222211
YO3	22211122111222121222222121	WO7	21122112122211112121112111
YO4	12222212121211111122211111	WO8	21121221121212212212112122
YO5	22212112111112111112111112	WO9	21121111111222221222222121
YO6	21112112211211211212122211	WO10	21121212121212121212122121
YO7	21121112111212222212211221	WO11	21121221111212121212122212



The resulting dissimilarities were subjected to clustering by the genetic distances. The values on the phenograms represent the genetic distances among different inbred lines. To estimate the reliability or effectiveness of the analysis, a thousand bootstrap replicates were performed. A thousand trees were produced and the computer programme determined a consensus tree (Fig 3.3 and 3.4). The consensus tree gives a good idea of how effective the total analysis is. The numbers at the forks indicate the number of times the group consisting of the species which are to the right of that fork occurred among the trees, out of 1000 trees.

Based on the Neighbour joining dendogram and consensus tree, the lines were grouped into 21 major clusters (Fig. 3.1 and 3.3). Clusters from different trees were consolidated and those occurring in both trees were used. When comparing the dendogram (using Neighbour joining) and Consensus (using Neighbour joining) generated trees it is evident that these trees are similar.

In cluster I, there were two subgroups. Subgroup a consisted of lines YO13 and YO14. Subgroup b contained lines YR4, YR3, YL3 and WR1. In the subgroup b, lines YR4 and YR3 were closer related to each other, and YL3 and WR1 were closer related to each other. In the dendogram the two subgroups were separated and in the consensus tree the two subgroups were grouped together. In general the genotypes of this cluster had pink or purple tassel and silk colour, an upright leaf type and ears longer than 130mm.

Cluster II had five lines: YF2, YL4, YL1, YO17 and YI3. This cluster was divided into two subgroups. Subgroup a consisted of lines YF2, YL4 and YL1, where YF2 and YL4 were closer related to each other than YL1. Subgroup b consisted of lines YO17 and YI3. In the dendogram the two subgroups were separated and in the consensus tree the two subgroups were grouped together. These lines had some common characters, they had a white cob colour, no kernel abortion or wide inter row spacing, and they had more than 12 kernel rows per ear.

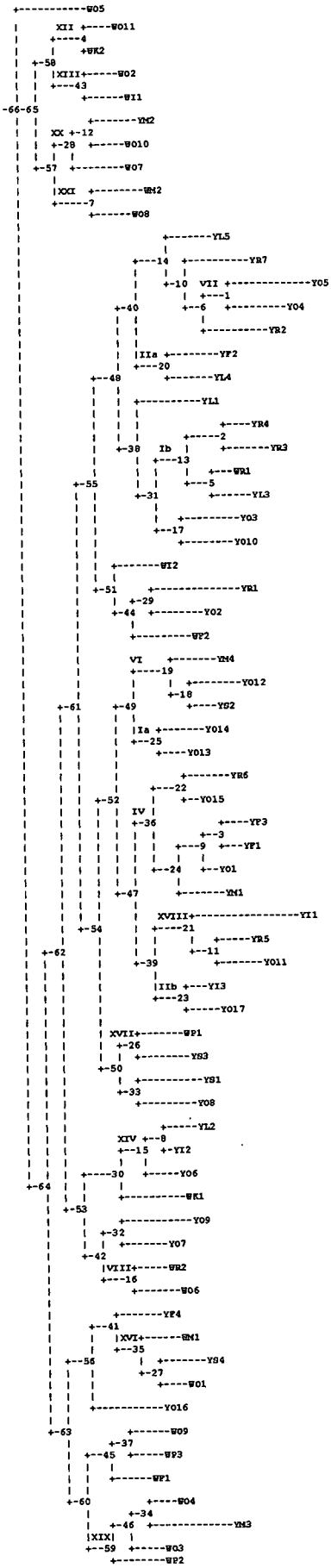


Fig. 3.1 Dendogram generated by Neighbour joining analysis from morphological data for the 68 different maize lines.

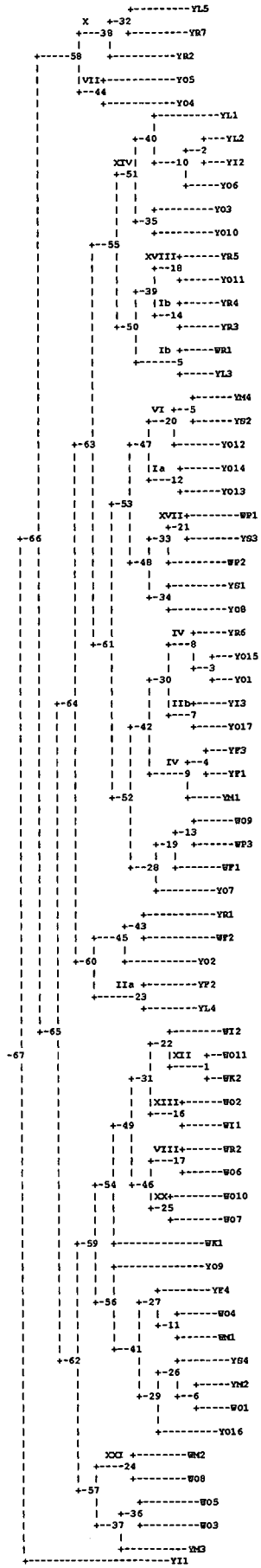


Fig. 3.2 Dendogram generated by UPGMA analysis from morphological data for the 68 different maize lines.

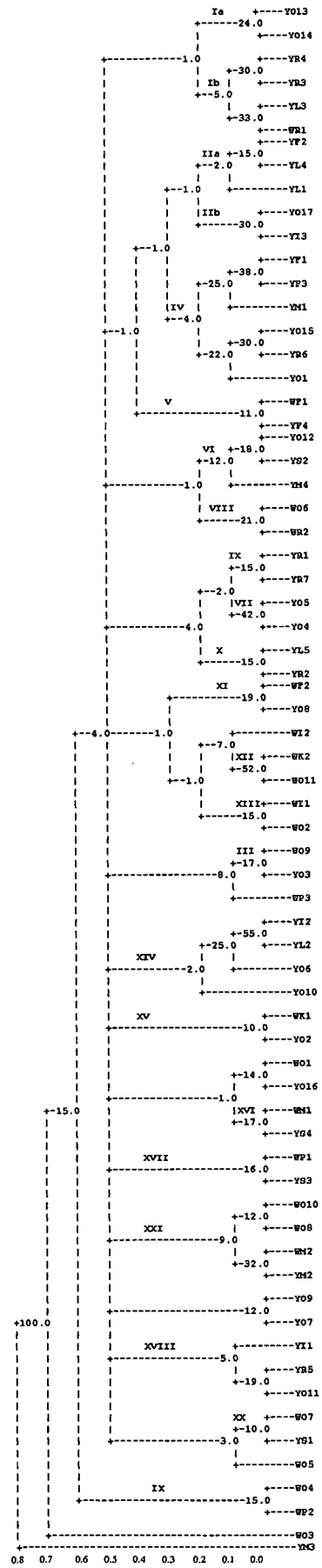


Fig. 3.3 Consensus tree generated by Neighbour joining analysis from morphological data for the 68 different maize inbred lines.

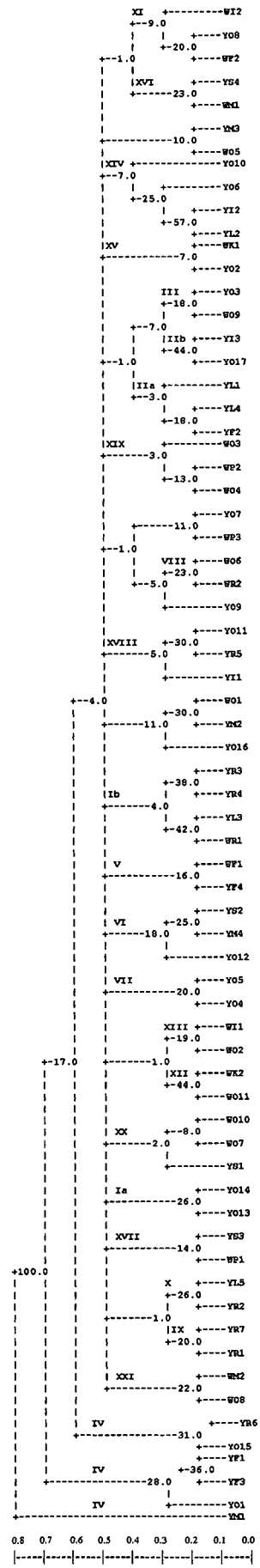


Fig. 3.4 Consensus tree generated by UPGMA from morphological data for the 68 different maize inbred lines.

In the consensus tree cluster III consisted of lines WO9 and YO3, but the lines were not grouped together in the dendogram. Common features have been observed within each cluster group. The genotypes of this cluster possessed wide inter row spacing, pink silk colour, dent grain type, good germination, no kernel abortion and floppy leaf type.

Cluster IV consisted of six lines: YF1, YF3, YM1, YO15, YR6 and YO1. In both trees these lines were grouped together, but the subgroups in the cluster were grouped differently. This was the biggest cluster in both the trees. These lines had some common characteristics, they had a white cob colour, no kernel abortion or wide inter row spacing, and they had more than 12 kernel rows per ear.

Cluster V consisted of lines WF1 and YF4 in the consensus tree. In the dendogram the lines were not grouped together. These lines had some common characters, they had a white cob colour, no kernel abortion or wide inter row spacing, and they had more than 12 kernel rows per ear.

Cluster VI comprised the lines YM4, YO12 and YS2. Line YO12 and YS2 were the closest related within this cluster. The lines in this cluster both had upright leaf type, there were more than 12 kernel rows per ear, orange grain colour and good germination.

Cluster VII comprised lines YO4 and YO5. In the consensus tree this cluster was closely related to cluster IX and in the dendogram this cluster was closely related to cluster X. Common features that have been observed within the cluster are kernel abortion, wide inter row spacing and susceptibility to stalk borer.

Cluster VIII consisted of two lines, WO6 and WR2. In the consensus tree this cluster was closely related to cluster VI and in the dendogram this cluster was closely related to lines YO9 and YO7. The lines in this cluster both had upright leaf type, there were more than 12 kernel rows per ear, orange grain colour and good germination.

In the consensus tree cluster IX consisted of two lines (YR1 and YR7). In the dendogram these lines were not grouped together. In general the genotypes of this cluster had common traits such as kernel abortion, less kernel rows per ear and susceptibility to stalk borer.

Cluster X consisted of only two lines, YL5 and YR2. In the dendogram these two lines were also grouped together, but not alone like in the consensus tree. The genotypes were characterised by kernel abortion, wide inter row spacing and susceptibility to stalk borer.

Cluster XI consisted of three lines in the consensus tree: WF2, WI2 and YO8. In the dendogram YO8 was not included and the other two lines were not grouped alone. The major features of this cluster were white cob colour, no root lodging, no tillers, upright leaf and purple tassel colour.

Cluster XII consisted of two lines: WO11 and WK2. In the consensus tree this cluster was closely related to cluster XI and XIII and in the dendogram this cluster was closely related to cluster XIII, XX and XXI. The genotypes were characterised by white cob colour, no root lodging, no tillers, upright leaf and purple tassel colour.

Cluster XIII comprised of the lines WI1 and WO2. Common features were observed within this cluster group. The genotypes possessed white cob colour, no root lodging, no tillers, upright leaf and purple tassel colour. Cluster XII and XIII were closely related to each other.

In cluster XIV there were four lines in the consensus tree, YO10, YO6, YL2 and YI2. Line YL2 and YI2 were the closest related in this cluster. In the dendogram YO10 was excluded from the cluster. The genotypes of this cluster possessed pink silk colour, presence of a cap, less than 12 kernel rows per ear, a dent grain type, less kernel rows per ear and no kernel abortion.

In the consensus tree cluster XV comprised two lines, WK1 and YO2. The lines were however not grouped together in the dendogram. The genotypes were characterised by upright leaf type, root lodging, white cob colour, ear length less than 130 mm and wide inter row spacing.

In cluster XVI there were only two lines, WM1 and YS4, in the consensus tree. In the dendogram a third line, WO1, was included in the cluster. The genotypes had ear length longer than 130 mm, floppy leaf type, stalk borer susceptibility, Diplodia ear rot tolerance and no cap.

Cluster XVII consisted of two lines, YS3 and WP1. In the consensus tree this cluster was grouped on its own and in the dendogram this cluster was closely related to lines YS1 and YO8. The genotypes were characterised by flint grain type, upright leaf type, less than 12 kernel rows per ear and tapering ears.

Cluster XVIII comprised of three lines, YI1, YR5 and YO11, in both trees. Line YR5 and YO11 were closer related to each other than YI1. The genotypes were characterised by purple silk and tassel colour, upright leaf type, an ear width less than 40 mm, an ear length more than 130 mm, no kernel abortion and Diplodia ear rot tolerance.

In cluster XIX there were four lines, WO4, WO3, WP2 and YM3. In the consensus tree WO3 and YM3 had less resolution and were not as closely related to WP2 and WO4 as in the dendogram. The genotypes were characterised by weight per ear of more than 40 g, prolificacy, flint grain type, more than 12 kernel ears per ear and rust tolerance.

Cluster XX had two lines, WO7 and YS1, in the consensus tree, but they were not grouped together in the dendogram. The genotypes of this cluster possessed tapering ears, good germination, upright leaf type, stalk borer tolerance, flint grain type and susceptibility to Diplodia ear rot.

i 151 165 42

Cluster XXI comprised of two lines, WO8 and WM2 in both trees. In the consensus tree this cluster was closely related to lines YM2 and WO10 and in the dendogram this cluster was closely related to cluster XX and line YM2. The major features of this cluster were flint grain type, high plant height, no tillers, kernel abortion and floppy leaf type. Clusters Ia, IIb, III, V, VII, VIII, IX, X, XII, XIII, XV, XVII, XX and XXI all only had two lines.

The clustering position of lines WO1, YO16, WO10, YM2, YO9, YO7, WO5 and YM3 were uncertain. The clusters' composition indicated absence of relation between geographic and genetic diversity as estimated from morphological attributes. Therefore the genetic relationship between the above mentioned lines are also uncertain.

Based on the UPGMA dendogram and consensus tree, the lines were grouped into 21 major clusters (Fig. 3.2 and 3.4). Clusters from different trees were consolidated and those occurring in both trees were used. When comparing the dendogram (using UPGMA) and Consensus (using UPGMA) generated trees it is evident that these trees are similar.

In cluster I, there were two subgroups. Subgroup a was composed of lines YO13 and YO14. The subgroup b consisted of lines YR3, YR4, YL3 and WR1. In both trees the subgroups were not grouped together. In general the genotypes of this cluster had pink or purple tassel and silk colour, an upright leaf type and ears longer than 130 mm.

Cluster II had five lines: YF2, YL4, YL1, YO17 and YI3. This cluster was divided into two subgroups. Subgroup a consisted of lines YF2, YL4 and YL1, where YF2 and YL4 were closer related to each other than YL1. Subgroup b consisted of lines YO17 and YI3. In the dendogram subgroups a was grouped with lines YO2, WF2 and WR1 and subgroup b was grouped with cluster IV. In the consensus tree the two subgroups were grouped together with cluster III. These lines had some common characters, they had a white cob colour, no kernel abortion or wide inter row spacing, and they had more than 12 kernel rows per ear.

Cluster III consisted of lines WO9 and YO3 in the consensus tree, but the lines were not grouped together in the dendogram. Common features have been observed within each cluster group. The genotypes of this cluster possessed pink silk colour, dent grain type, good germination, no kernel abortion and floppy leaf type.

Cluster IV consisted of six lines: YF1, YF3, YM1, YO15, YR6 and YO1. In both trees these lines were grouped together. However, in the dendogram the cluster included subgroup IIb in the cluster. This was the biggest cluster in both the trees. These lines had some common characters, they had a white cob colour, no kernel abortion or wide inter row spacing, and they had more than 12 kernel rows per ear.

Cluster V consisted of lines WF1 and YF4 in the consensus tree. In the dendogram the lines were not grouped together. These lines had some common characteristics, they had a white cob colour, no kernel abortion or wide inter row spacing, and they had more than 12 kernel rows per ear.

Cluster VI comprised of the lines YM4, YO12 and YS2. Line YM4 and YS2 were the closer related to each other than YO12. The lines in this cluster had upright leaf type, tolerance to stalk borer and Diplodia ear rot, there were more than 12 kernel rows per ear, orange grain colour and good germination.

Cluster VII comprised of lines YO4 and YO5. In the dendogram this cluster was closely related to cluster X. Common features that have been observed within the cluster are kernel abortion, wide inter row spacing and susceptibility to stalk borer.

Cluster VIII consisted of two lines, WO6 and WR2. In the consensus tree this cluster was closely related to lines YO9, YO7 and WP3 and in the dendogram this cluster was closely related to cluster XX. The lines in this cluster had upright leaf type, tolerance to stalk borer and Diplodia ear rot, there were more than 12 kernel rows per ear, orange grain colour and good germination.

Cluster IX consisted of two lines: YR1 and YR7. In the consensus tree this cluster is closely related to cluster X. In the dendogram these lines were not grouped together, YR7 was grouped in cluster X and YR1 was grouped with subgroup IIa. In general the genotypes of this cluster had common traits such as kernel abortion, wide inter row spacing and susceptibility to stalk borer.

Cluster X consisted of two lines, YL5 and YR2. In the dendogram YR7 was also included. The genotypes were characterised by kernel abortion, wide inter row spacing and susceptibility to stalk borer.

Cluster XI consisted of three lines in the consensus tree: WF2, WI2 and YO8. Lines YO8 and WF2 were closer related to each other than WI2. In the consensus tree this cluster is closely related to cluster XVI. In the dendogram the lines were not grouped together. The major features of this cluster were white cob colour, no root lodging, no tillers, upright leaf and purple tassel colour.

Cluster XII consisted of two lines: WO11 and WK2. In both trees this cluster was closely related to cluster XIII. The genotypes were characterised by white cob colour, no root lodging, no tillers, upright leaf and purple tassel colour.

Cluster XIII comprised of the lines WI1 and WO2. Common features were observed within this cluster group. The genotypes possessed white cob colour, no root lodging, no tillers, upright leaf and purple tassel colour. Cluster XII and XIII were closely related to each other.

In cluster XIV there were four lines in the consensus tree, YO10, YO6, YL2 and YI2. Line YL2 and YI2 were the closest related in this cluster. In the dendogram YO3 was included in the cluster and YL1 was closely related to this cluster. The genotypes of this cluster possessed pink silk colour, presence of a cap, less than 12 kernel rows per ear, a dent grain type, no kernel abortion.

Cluster XV comprised of two lines, WK1 and YO2, in the consensus tree. The lines were however not grouped together in the dendogram. The genotypes were characterised by upright leaf type, root lodging, white cob colour, ear length less than 130 mm and wide inter row spacing.

In cluster XVI there were two lines: WM1 and YS4. In the consensus tree the cluster is closely related to cluster XI. In the dendogram the cluster included YF4, WO4, YM2, WO1 and YO16. The genotypes had ear length longer than 130 mm, floppy leaf type, stalk borer susceptibility, Diplodia ear rot tolerance and no cap.

Cluster XVII consisted of two lines, YS3 and WP1. In the consensus tree this cluster was grouped on its own and in the dendogram this cluster was closely related to line WP2. The genotypes were characterised by flint grain type, upright leaf type, less than 12 kernel rows per ear and tapering ears.

Cluster XVIII comprised of three lines, YI1, YR5 and YO11, in both trees. Line YR5 and YO11 were closer related to each other than YI1. The genotypes were characterised by purple silk and tassel colour, upright leaf type, an ear width less than 40 mm, an ear length more than 130 mm, no kernel abortion and Diplodia ear rot tolerance.

In cluster XIX there were three lines, WO4, WO3 and WP2, in the consensus tree. In the dendogram these lines were not grouped together. The genotypes were characterised by mass per ear of more than 40 g, prolificacy, flint grain type, more than 12 kernel ears per ear and rust tolerance.

Cluster XX had three lines, WO7, WO10 and YS1, in the consensus tree. There were only two lines, WO7 and WO10, grouped together in the dendogram. The genotypes of this cluster possessed tapering ears, good germination, upright leaf type, stalk borer tolerance, flint grain type and susceptibility to Diplodia ear rot.

Cluster XXI comprised of two lines, WO8 and WM2 in both trees. In the consensus tree this cluster was grouped on its own and in the dendogram this cluster was closely related to lines WO5, WO3 and YM3. The major features of this cluster were flint grain type, high plant height, no tillers, kernel abortion and floppy leaf type. Clusters Ia, IIb, III, V, VII, VIII, IX, X, XII, XIII, XV, XVI, XVII and XXI all only had two lines.

The position of lines YM3, WO5, YO7, WP3, YO9, YM2, WO1 and YO16 was placed out of any cluster. The clusters' composition was indicating absence of relation between geographic and genetic diversity as estimated from morphological attributes. Therefore the genetic relationship between the above mentioned lines are also uncertain.

If the Neighbour joining and UPGMA consensus trees are compared with each other, there were a few differences. In the Neighbour joining tree lines YF1 and YF3 clustered close toYM1, but in the UPGMA tree they clustered close to YO1 (in neighbour joining tree lines YR6 and YO15 clustered close to YO1). If the grouping of lines YO12, YS3 and YM4 were compared in the two trees, YO12 and YS2 were closest related of the three lines in the Neighbour joining tree, and in the UPGMA tree YS2 and YM4 were closest related of the three. Lines YO3 and WO9 were closest related to WP3 in the Neighbour joining tree, and closest related to YI3 and YO17 in the UPGMA tree. In the Neighbour joining treeYM3 was the most distant line in the tree, and in the UPGMA tree YM1 was the most distant related to other lines. It is clear that the differences between the trees are minor. Common features were observed within each cluster group. If the grouping pattern in the two different analysis methods were compared, there were similar patterns observed, indicating increased confidence in the constructed trees.

Discussion and conclusion

Morphological characters have proven to be a useful tool in the *Zea mays* classification (Whythe and Gevers, 1988; Gerdes and Tracy, 1994; Lübberstedt *et al*, 1998). Looking at the Neighbour joining-, as well as the UPGMA dendograms and cluster analysis, the lines were grouped very similarly. Therefore the morphological characterisation method has gained confidence and certain patterns of grouping were identified.

There was a correlation of origin with grouping of inbred lines throughout the consensus trees. There was a tendency of American inbred lines grouping together (Cluster I, X and XVIII), local inbred lines grouping together (Cluster II, IV, V, VII, XI, XV, XVII AND XIX), Thailand inbred lines grouping together (Cluster IV) and Argentinean inbred lines grouping together (Cluster I). Although the inbred lines differ in origin, it was revealed in the cluster patterns that the material in the Advanta group has been interbred to a large extent, therefore a lot of overlapping of groups was found.

The maize lines included in this study displayed high levels of polymorphism. Calculation of pairwise genetic distances (Table 3.7) revealed an average distance of 8.08% (1.28 to 14.88) between the maize lines. These results suggest that the genetic distance data based on morphological characterisation of maize lines can be a useful tool when selecting the most suitable parental maize lines used to construct populations and hybrids. However, this is not the only criterion for the selection of parental maize lines used to construct populations and hybrids. Plant breeders generally prefer to use recurring parents that have the highest number of desirable traits (Hallauer, 1987).

A problem that has been experienced is that in general when breeders make selections in field breeding programmes, they are based on intimate knowledge of the line, as well as selective characteristics. However, the history of the inbred lines used in this study, was unknown and it was not a natural population where selection occurred naturally, over a long period of time. This group was merely a selection of commercial strains that were determined by their ability to cross and yield. These

lines were selected purely on the basis of their importance to the maize breeding programme of Advanta Africa Seeds in South Africa. The aim of the genetic distance analysis was to determine the levels of genetic variability that exists between a group of potential recurrent or donor parents. There was a wide variability in the morphological characteristics indicating a high degree of interspecific hybridisation, assuming different characters could distinguish the genotypes. In this study all the characters contributed equally, there were no specific major features used to group the lines.

Salhuana *et al.* (1998) evaluated nearly 12 000 maize accessions in 12 countries to identify the heterotic group in which the accessions belonged, using morphological characterisation. Varying responses of the accessions to different environmental conditions were recognised, and five homologous areas were defined according to elevation and latitude. Therefore environment interaction also plays an important role in determining the expression of the characteristics of the maize plant. Reports from Goodman and Brown (1988), Smith and Smith (1989), Lee (1995) and Lübberstedt *et al.* (1998) have shown that characterisation of maize germplasm accessions should be based on descriptors that are not influenced by the environment. For this reason it is also problematic to use morphological characterisation.

Reports on morphological characterisation of maize in Africa are fairly scarce, because evaluation of maize lines in extensive field tests is the most costly and time consuming part in modern hybrid breeding of grain and forage maize (Boppenmaier *et al.*, 1992). Therefore, maize breeders often need to make a "by-chance" selection in breeding programmes, based solely on prior experience. However, the approach used in this study makes it possible for the breeder to evaluate morphological characters in a scientific manner to determine which inbred lines crossed will give good combining ability. Using morphological characters in maize breeding will remain priority especially when the characteristics are sought after in new lines.

In conclusion, the results of this study revealed that definite groups could be defined based on morphological data. These groupings are based on similarity in morphological characters. Therefore breeding lines that cluster together would not result in good combining ability if crossed and should not be used to make crosses.

Maize lines should have the largest possible genetic distance between them, even if based on morphological characters, to yield the best hybrids. Measurements of genetic distance facilitate identification of diverse parents in order to maximise the expression of heterosis (Smith *et al.*, 1990). The statistical analysis of morphological characters will allow the breeder to make objected choices for the selection of parental material to produce new hybrids. When looking at improving the variation within populations, in other words the source of parental material, inbred lines from the same groupings would be crossed with each other. This would enable the breeder to increase variability within the parental material. The dendograms in this study showed significant grouping, but more objective methods are needed to confirm the results. These results will however contribute to the breeding programme, since it will make the possibility of higher combining ability in crosses, as well as improved populations, higher.

CHAPTER 4

IDENTIFICATION AND GENETIC DISTANCE ANALYSIS OF MAIZE INBRED LINES USING AFLP FINGERPRINTING

Abstract

In maize (*Zea mays* L.) breeding programmes it is important to determine the genetic relatedness between inbreds, as the amount of relatedness will directly affect the amount of heterosis expressed in the resulting hybrids. The objective of this research was to utilise AFLP's (Amplified Fragment Length Polymorphic markers) to characterise and determine groupings of 50 maize inbred lines (both flint and dent) from the Advanta Africa genebank (representing five countries). The AFLP data was coded using binary character states and analysed using PHYLIP (Phylogeny Inference Package) Version 3.5p (Felsenstein, 1993). AFLP fingerprinting of maize inbred lines allowed the efficient detection of large numbers of polymorphic amplified fragments. The inbred lines were clustered into two major clusters. Both clusters were further divided into two subgroups, which were divided into a number of minor groups. Correlation was found using different methods of analyses. These results have identified genetically distant related lines, which if used in breeding programmes will increase the expression of heterosis in hybrids. The incorporation of molecular markers in a breeding programme has the potential to reduce cost and time.

Introduction

The development of techniques for analysis and quantification of genetic relatedness has generated a variety of new applications that can be used to complement classical plant breeding techniques. A variety of DNA fingerprinting techniques is presently available, most of which use PCR for detection of fragments. The identification of lines, estimation of genetic distance, phylogenetic analysis, genome mapping and gene tagging in particular, have benefited from the

application of DNA fingerprinting techniques (Yu *et al.*, 1993; Nybom, 1994; Vos *et al.*, 1995).

Amplified fragment length polymorphism's (AFLP's) is a PCR-based method which first involves restriction digestion of the genomic DNA. Adapters are ligated to the ends of the restriction fragments and pre-selection is performed with two AFLP primers having a single selective nucleotide, followed by a second amplification using primers having longer selective extensions (Zabeau and Vos, 1993; Vos *et al.*, 1995). The resulting amplification products are separated on a sequencing gel and can be visualised using radioactive or fluorescent labelling. Studies suggest that AFLP's are comparable to RFLP's (Restriction Fragment Length Polymorphisms) in terms of reproducibility and should therefore be highly suited to use in experiments (Jones *et al.*, 1997).

AFLP analysis has found widespread application in agriculture (Donini *et al.*, 1997; Pakniyat *et al.*, 1997). It is more cost-effective, reproducible and readily automatable for high marker output than RFLP's and RAPD's (Perkin-Elmer, 1996; Van Toai *et al.*, 1997; Barrett and Kidwell, 1998). AFLP analysis tends to produce more informative data points and does not have some of the limitations associated with RFLP's, RAPD's and SSR's (Keim *et al.*, 1997). The AFLP technique does not require sequence data or access to probe libraries and is, therefore, accessible to research environments previously restricted by these requirements (Mazur and Tingey, 1995).

The accurate identification of plant breeding material is extremely important, not only for protection of breeder's rights, but also to drive plant-breeding programmes. DNA-based identification methods have proved useful in cases where line identification is based on mature plant characteristics such as flowering or fruit ripening and where growth conditions may influence plant qualities (Morell *et al.*, 1995). AFLP analysis has been used to establish phylogenetic relationships among species, subspecies and lines (Pakniyat *et al.*, 1997) and to measure genetic variation in populations (Barrett and Kidwell, 1998) and species (Van Toai *et al.*, 1997).

Since morphological variation among maize lines is not very distinct, phenotypic identification based on morphological traits is difficult. AFLP's have been used as an alternative for line identification in some crop species. The aim of this study was to use AFLP fingerprinting to identify and determine the genetic distance between inbred maize lines from the Advanta Africa genebank. All 68 maize inbred lines are currently in use in maize breeding programmes at Bapsfontein Research Station, South Africa, with the aim to make crosses with the best combining ability. Morphological data from this material was collected during the 1998/1999 season.

Material and Methods

Plant material

A total of 50 maize inbred lines were used in this study. The material was obtained from the Advanta Africa genebank (Building number 9, Old Trafford, Corner of Leith and Trichardt Road, Bartlett, Boksburg). The maize inbred lines used in this study are summarised in Table 4.1.

The germplasm used included lines from Africa (local), Argentina, Zimbabwe, Garst (USA) and Thailand. Six individuals (three per pot) of each inbred line were grown under glasshouse conditions at University of the Orange Free State (UOFS), in Bloemfontein, South Africa. Growth conditions were set at 20°C night temperature and 26°C day temperature.

Table 4.1 Maize lines used for AFLP analysis.

Nº	Lines	Origin	Nº	Lines	Origin
1	YO12	Argentina	26	YM1	Local
2	YO13	Argentina	27	YM2	Local
3	WO8	Brazil (tropical)	28	YM3	Local
4	WO9	Brazil (tropical)	29	YI2	Local
5	WO10	Brazil (tropical)	30	YS3	Local
6	WO11	Brazil (tropical)	31	YF4	Local
7	WO5	(CIMMYT) Zimbabwe	32	WO1	Local
8	WO6	(CIMMYT) Zimbabwe	33	WF1	Local
9	WO7	(CIMMYT) Zimbabwe	34	WF2	Local
10	YM4	GARST (USA)	35	WI2	Local
11	YL2	GARST (USA)	36	WP3	Local
12	YR5	GARST (USA)	37	WK1	Local
13	YL1	GARST (USA)	38	WK2	Local
14	YL4	GARST (USA)	39	YO9	Local
15	YL5	GARST (USA)	40	YO10	Local
16	YO11	GARST (USA)	41	WM1	Local
17	WR1	GARST (USA)	42	WM2	Local
18	WR2	GARST (USA)	43	WO3	Local
19	YO1	Local	44	WI1	Local
20	YO2	Local	45	WP1	Local
21	YF1	Local	46	WP2	Local
22	YF2	Local	47	WO2	Local
23	YF3	Local	48	WO4	Local
24	YO4	Local	49	YR6	Thailand
25	YO6	Local	50	YS2	Thailand

DNA extraction

DNA was extracted from young fresh leaves using the modified DNA isolation method described by Dellaporta *et al.* (1983). Plant material was homogenised using a mortar and pestle. The leaf material was frozen in liquid nitrogen. The macerated tissue was transferred to a 50 ml centrifuge tube containing 5 ml extraction buffer (100 mM Tris-HCl, 500 mM NaCl, 50 mM EDTA, 1.25% (m/v) SDS, 20mM $Na_2S_2O_5$, pH 8.0) and mixed vigorously. The homogenate was incubated at 65°C for 20 minutes, with periodic shaking every ten minutes. Thereafter chloroform-isoamylalcohol extraction was performed using 5 ml chloroform-isoamylasetate (24:1). Centrifugation of the extraction was performed at 10 000 rpm for ten minutes at 20°C. After three or four extractions the DNA was precipitated using isopropanol (1:1 v/v). The DNA was spooled and washed in 70% Ethanol. The DNA pellet was resuspended in one ml of sterile water (Sabax, Non Pyrogenic). The DNA concentration was determined spectrophotometrically. The DNA samples were diluted to a final concentration of 250 ng/µl and aliquoted for storage at -20°C.

Primers

AFLP reactions employed two oligonucleotide primers, one corresponding to the PstI-ends and one corresponding to the MseI-ends. The PstI-primer was fluorescently labelled. DNA was digested with MseI (frequent 4-base cutter) and PstI (rare 6-base cutter) as described by Vos *et al.* (1995). Three primer pairs (Life Technologies Ltd., Scotland, UK) (Table 4.2) were used for AFLP analysis of the maize inbred lines (Vos *et al.*, 1995; Keim *et al.*, 1997). Oligonucleotide sequences used for adapters and primers are listed in Table 4.2.

Table 4.2 Oligonucleotide sequences used in AFLP analysis.

Primers	
MseI:	5'-GAC GAT GAG TCC TGA G-3'
PstI:	5'- CTC GTA GAC TGC GTA CAT GCA-3'
Adapters	
MseAd1:	5'-GAC GAT GAG TCC TGA-3'
MseAd2:	3'-TAC TCA GGA CTC AT-5'
PstAd1:	5'-CTC GTA GAC TGC GTA CAT GCA-3'
PstAd2:	3'-CAT CTG ACG CAT GT-5'
PstP1:	5'-GAC TGC GTA CAT GCA C-3'
PstP2:	5'-GAC TGC GTA CAT GCA A-3'
Primers labeled with 6-FAM labels:	
PstA1:	5'-GAC TGC GTA CAT GCA ATA-3'
PstA2:	5'-GAC TGC GTA CAT GCA CTC-3'
PstA3:	5'-GAC TGC GTA CAT GCA AGG-3'
Primers labeled with HEX labels:	
PstA4:	5'-GAC TGC GTA CAT GCA ACC-3'
PstA5:	5'-GAC TGC GTA CAT GCA CAA-3'
PstA6:	5'-GAC TGC GTA CAT GCA CGA-3'

Pre-amplification and selective amplification

After adapter ligation, non-selective amplification (pre-amplification) of DNA fragments was performed using non-selective (zero base extension) primers. The PCR products of the pre-amplification reaction were then diluted and used as template for the second AFLP reaction using primers both having three selective nucleotides. Selective amplification was carried out using various primer combinations of primer MseI and primers PstI (2-4 base pair extension) (Table 4.2) (Zabeau and Vos, 1993; Vos *et al.*, 1995).

Pre-selective amplification reactions were as follows: 20 cycles of 30 sec at 94°C, one minute at 56°C followed by one minute at 72°C. Selective amplification reactions were as follows: 36 cycles of 30 sec at 94°C, 30 sec at 65°C reduced to 56°C using a ramp of –0.7°C per cycle, followed by one minute at 72°C. All amplification reactions were performed in a Hybaid thermocycler (Vos *et al.*, 1995).

Gel analysis

Amplification products through the AFLP process were resolved by electrophoresis on 2 % agarose gels (Seakem LE, FMC) in a TAE buffer (40 mM Tris-acetate, one mM EDTA, pH 8.0). Electrophoresis was performed at constant power, 110 W, for two hours. Amplification products were detected under UV light after staining with ethidium bromide (0.5 μg/ml) (Vos *et al.*, 1995). The AFLP profiles were resolved using an ABI377 automatic sequencer.

Data collection and analysis

AFLP fragment data for three primer pairs was coded using a binary system and summarised in a data matrix (Appendix, Table 8.1). The PHYLIP (Phylogeny Inference Package) Version 3.5p (Felsenstein, 1993) was used to determine the genetic distances and to group the inbred lines into clusters. Kimura (Table 4.3) and Jukes and Cantor (Table 4.4) methods were used to calculate the genetic distances between the inbred lines. The number of polymorphic bands was determined (Barrett and Kidwell, 1998). The index of genetic similarities (F) of Nei and Li (1979) was used to calculate pairwise genetic distances (D) for all the lines:

$$F = 2N_{xy}/(N_x N_y)$$

$$D = 1 - F$$

Where N_{xy} = the number of shared bands between any two lines x and y and N_x and N_y are the number of bands for line x and y, respectively (Wang and Tanksley, 1989). Neighbour joining was used in cluster analysis of the pairwise distance matrix. Support for clusters was evaluated by 1000 bootstrap replicates using

Kimura and Jukes and Cantor algorithms with Neighbour joining analysis to produce consensus trees (Barrett and Kidwell, 1998). The DNA parsimony algorithm version 3.5p was also used to evaluate 1000 bootstrap replicates (Fig. 4.10).

Results

Quality of the genomic DNA

The DNA quality was determined spectrophotometrically, by taking readings at 260nm and 280nm. The 260: 280 ratio was 1.8 or higher and the average DNA yield was 1095.5 µg/µl. The quality of the DNA was good.

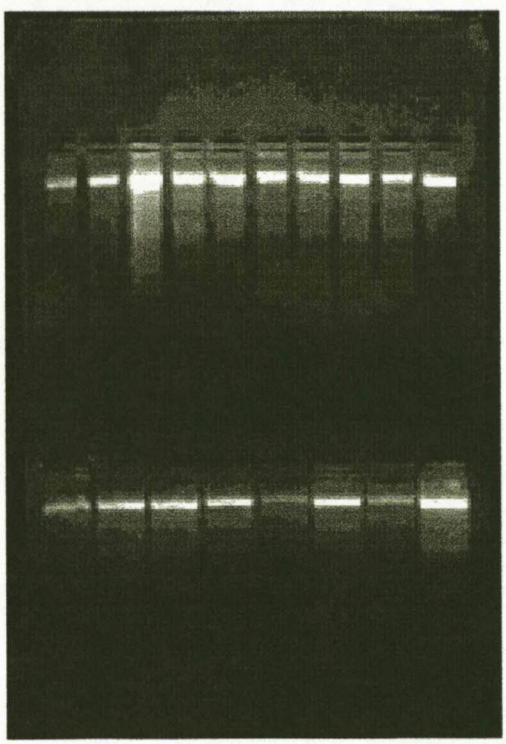


Fig 4.1 Gel illustrating the presence or absence of DNA after DNA extraction.

Identification of maize inbred lines

A line fingerprinting profile was compiled using unambiguously scorable AFLP polymorphism's from the least number of primers allowing differentiation of all the lines in the study. The fingerprinting profile was used to construct a molecular identification key that displayed the sequence and combination of AFLP primers needed to identify each of the lines in the study. Analysis of the 50 *Zea mays* accessions with three AFLP primer pairs (Table 4.1) identified a total of 107 fragments. Within an individual plant the three primer pairs each detected different AFLP profiles between 0.5 to 2.0 kilobases (kb). Of these, all 107 bands were polymorphic (no monomorphic bands were determined). The average number of bands detected per primer for all plants was 17.8. This also reflects the high genetic similarity among the maize lines and necessitated the use of more than one AFLP marker to distinguish any of the maize lines. Polymorphic loci that distinguish all the lines used in the analysis allowed the lowest number of primers for line identification.

Genetic distance analysis of maize inbred lines

Pairwise genetic distances (Table 4.3 and 4.4) based on the F statistic of Nei and Li (1979), revealed an average distance of 4.38% (ranging from 0.09 to 8.67) for all lines. The lines were originally from different sources and countries of origin (Table 3.1). The biggest pairwise distance (8.67%) was observed between line WO2 and WP1. This is further illustrated by the fact that Neighbour joining cluster analysis grouped them separately (Fig. 4.3). The smallest pairwise distance was observed between WO1 and WO2, as well as YS2 and YO10. These calculations were based on the presence or absence of specific AFLP markers among the lines (Appendix, Table 8.1).

The maize inbred lines were grouped into groups (major clusters), which in turn were subdivided into subgroups and these subgroups were again subdivided into minor groups. When looking at the clustering pattern, among the maize inbred lines from group I, the average distance between them was 4.39% (ranging from 0.1 to 8.67). The lines from group II had an average distance between them of

Table 4.3 Pairwise distance matrix using Kimura based on the F statistic of Nei and Li (1979).

	WM1	WM2	WF1	WF2	WH	WM	WP1	WP3	WR1	WR2	WK1	WK2	WO1	WO2	WO3	WO4	WO5	WO6	WO7	WO8	WO9	WO11	YOH	YM2	YM3	YM4	YF2	YF3	YF4	WP2	WO10	YF1	YO	YO5	YO6	YL1	YL2	YL4	YL5	YS2	YS3	YO1	YO2	YO4	YO6	YO9	YO10	YO11	YO12	YO13
WM1	0																																																		
WM2	4.39	0																																																		
WF1	4.69	4.69	0																																																	
WF2	4.99	4.11	3.83	0																																															
WH	3.83	3.55	4.11	4.39	0																																														
WM	3.97	4.25	4.54	3.97	3.69	0																																													
WP1	4.25	4.54	4.54	4.54	4.25	5.03	0																																												
WP3	3.83	4.11	3.29	3.83	3.25	3.69	5.15	0																																										
WR1	3.97	3.43	4.54	3.43	3.69	4.11	3.83	3.97	0																																										
WR2	4.39	4.39	3.55	4.69	3.83	3.69	5.15	3.29	3.69	0																																									
WK1	4.54	3.16	3.97	3.69	3.69	3.29	4.11	3.43	3.55	3.83	0																																								
WK2	4.25	4.54	5.79	5.45	4.25	5.03	4.99	4.84	4.99	4.54	4.11	0																																								
WO1	4.39	4.69	4.69	4.99	4.39	4.84	6.13	3.55	4.25	3.29	4.25	3.43	0																																							
WO2	4.39	4.69	4.69	4.99	4.39	4.84	6.13	3.55	4.25	3.29	4.25	3.43	0.01	0																																						
WO3	3.83	3.55	3.29	3.83	3.83	3.97	4.25	2.79	3.83	3.29	3.97	4.84	4.11	4.11	0																																					
WO4	4.25	3.43	3.69	4.84	2.66	4.69	4.69	3.16	4.11	3.83	3.04	4.39	3.97	3.97	3.16	0																																				
WO5	3.83	3.83	4.39	4.99	3.29	5.15	4.84	3.04	4.25	3.83	3.97	4.25	4.39	4.39	2.54	2.66	0																																			
WO6	4.11	4.11	3.83	4.11	3.04	3.97	4.25	3.29	3.16	3.55	3.97	3.43	3.55	3.55	3.83	2.91	3.29	0																																		
WO7	3.83	4.11	3.83	4.69	3.83	4.25	3.97	3.04	3.69	3.04	3.97	4.25	3.55	3.55	3.83	2.91	3.04	2.07	0																																	
WO8	2.79	3.04	4.11	4.11	2.54	3.16	4.25	3.29	3.16	2.54	2.91	3.43	3.29	3.29	3.29	2.19	3.04	2.54	2.07	0																																
WO9	3.97	3.69	4.54	4.54	2.91	4.11	3.29	3.69	3.29	3.69	3.83	3.55	3.97	3.97	2.91	3.55	2.42	2.91	3.43	2.66	0																																
WO11	3.83	4.11	4.69	4.69	4.11	4.25	5.45	3.55	4.54	3.83	3.97	4.25	3.55	3.55	3.29	3.97	3.04	3.04	3.55	3.04	3.65	0																															
YOH	4.54	4.25	4.54	5.79	2.66	4.99	5.63	4.25	4.39	3.69	4.99	4.99	4.84	4.84	4.54	2.79	3.69	3.69	3.69	2.19	3.69	4.25	0																															
YM2	3.83	3.55	3.29	3.29	2.79	3.97	3.97	1.84	3.43	3.04	3.16	3.83	3.83	3.83	2.54	2.66	2.79	2.54	3.04	2.54	2.42	3.29	3.16	0																														
YM3	4.54	4.25	3.16	3.97	3.97	3.55	4.99	3.97	4.11	3.43	4.11	4.39	4.25	4.25	2.91	3.83	2.91	3.43	3.83	3.43	3.83	4.84	4.69	3.16	0																													
YM4	3.43	2.91	3.97	4.84	3.69	3.04	3.29	4.25	4.11	3.69	3.55	4.69	4.54	4.54	3.16	4.11	4.25	4.25	3.16	3.16	3.04	4.25	4.99	3.43	4.11	0																													
YF2	3.83	3.83	4.99	4.39	2.79	3.69	3.43	3.83	3.16	3.29	2.91	3.69	3.55	3.55	3.29	3.43	3.29	2.79	3.55	2.30	2.66	4.11	3.97	3.04	3.69	3.69	0																													
YF3	3.04	4.39	4.11	4.69	3.04	3.97	5.45	3.55	4.54	3.29	4.25	4.25	3.83	3.83	2.79	3.43	2.79	3.29	3.83	2.79	3.43	3.29	3.83	2.91	3.97	4.11	0																													
YF4	3.16	3.16	3.97	3.83	2.66	3.04	4.39	3.43	3.55	3.43	3.04	4.69	3.97	3.97	3.43	3.29	3.43	3.16	3.16	2.19	2.79	3.16	3.55	3.16	3.29	3.29	2.91	2.91	0																													
WP2	4.11	3.55	4.69	4.69	3.29	3.43	5.15	3.04	4.54	2.79	2.91	5.15	4.11	4.11	3.29	2.66	3.29	4.11	3.55	3.04	3.97	3.55	4.25	3.29	3.97	3.97	2.79	4.11	3.16	0																													
WO10	4.69	3.55	3.04	4.39	3.55	3.69	4.25	3.55	3.97	3.83	3.69	5.45	3.83	3.83	4.11	4.25	4.39	3.55	3.29	3.04	3.69	4.39	3.97	3.04	3.97	4.25	3.04	4.69	3.16	3.83	0																													
YF1	2.79	2.54	3.55	3.55	2.30	2.66	3.69	2.30	2.91	2.54	2.42	3.16	2.54	2.54	2.07	2.66	2.30	2.54	2.54	1.84	1.73	2.54	3.16	1.62	2.66	1.96	2.07	2.79	1.73	2.54	2.79	0																													
YO	2.79	2.30	3.29	3.04	2.07	2.91	3.43	2.07	2.42	2.54	2.19	2.91	2.79	2.79	2.07	2.19	2.30	2.07	2.30	1.40	1.51	2.54	2.91	1.40	2.42	2.19	1.84	2.54	1.73	2.54	2.30	0.55	0																													
YO5	2.54	2.54	4.11	4.11	2.54	3.16	3.43	3.04	2.91	3.29	2.91	3.43	3.55	3.55	2.54	2.91	2.54	2.79	2.79	1.84	1.96	3.29	2.91	2.07	2.91	1.96	2.07	3.29	2.19	3.55	3.04	0.98	0.98	0																													
YO6	3.97	5.15	3.69	4.25	3.16	4.39	4.11	2.91	4.39	4.54	3.83	4.69	4.25	4.25	3.97	3.55	3.69	3.16	3.69	3.43	3.29	4.84	4.39	2.91	3.04	3.53	3.16	2.91	3.29	4.54	3.97	3.16	2.66	3.16	0																													
YL1	2.66	2.42	3.43	3.16	2.19	3.04	3.29	2.19	2.54	2.66	2.07	3.04	2.91	2.91	2.19	2.07	2.19	2.19	2.42	1.51	1.62	2.42	3.04	1.51	2.54	2.30	1.96	2.66	1.84	2.42	2.42	0.68	0.68	1.08	2.79	0																													
YL2	3.16	3.43	3.16	3.69	2.66	3.55	4.39	3.69	2.79	3.43	3.55	4.39	4.25	4.25	2.19	3.04	2.91	3.16	3.43	2.19	2.54	3.43	3.55	3.16	2.79	3.29	3.16	2.19	3.04	3.43	3.16	2.19	1.73	2.42	3.29	1.84	0																													
YL4	2.79	2.54	3.29	3.29	2.30	2.91	3.69	2.30	2.66	2.54	2.19	2.91	2.79	2.79	2.07	2.19	2.07	2.07	2.07	1.40	1.51	2.30	2.66	1.40	2.42	2.19	2.07	2.54	1.73	2.54	2.54	0.38	0.38	0.98	2.91	0.48	1.96	0																													
YL5	3.55	3.04	3.55	4.39	3.83	4.54	4.54	4.11	4.54	4.99	4.54	5.15	5.63	5.63	3.55	3.43	3.29	3.29	3.83	3.04	3.43	4.11	3.43	3.29	2.91	3.69	4.11	3.55	2.91	4.69	4.39	3.04	2.54	3.04	2.91	2.66	3.16	2.54	0																													
YS2	2.54	2.30	3.29	3.04	1.84	2.91	3.43	2.07	2.42	2.30	2.19	2.66	2.54	2.54	2.07	1.96	2.07	1.84	2.07	1.19	1.51	2.54	2.66	1.19	2.42	2.19	1.62	2.30	1.73	2.30	2.30	0.58	0.19	0.98	2.42	0.28	1.73	0.38	2.54	0																													
YS3	3.16	3.16	4.25	3.16	2.66	3.04	5.30	2.42	3.56	3.97	2.79	3.56	3.43	3.43	3.43	2.79	3.16	2.91	3.16	2.19	3.04	3.69	3.56	2.19	3.29	3.04	2.66	3.43	2.54	3.16	3.43	1.73	1.51	1.96	3.04	1.62	2.79	1.73	3.43	1.30	0																													
YO1	3.29	3.04	2.79	4.11	2.79	3.43	4.25	2.79	2.91	3.29	3.43	4.54	3.83	3.83	3.04	2.66	2.79	2.30	2.79	2.30	3.16	3.56	3.16	2.30	2.91	3.16	3.29	2.79	2.91	3.29	3.04	1.84	1.84	2.30	3.68	1.73	2.42	1.84	3.29	1.62	2.91	0																													
YO2	3.68	3.43	4.25	3.97	2.42	3.56	4.11	2.91	2.79	2.42	2.79	3.04	3.43	3.43	2.66	2.54	2.42	2.42	3.16	1.96	2.07	3.16	3.29	1.96	3.04	3.04	1.96	2.91	2.54	2.19	3.16	1.30	1.09	1.96	3.25	1.19	2.30	1.30	3.43	0.86	2.30	2.19	0																													
YO4	2.66	2.42	3.16	2.91	1.96	2.79	3.56	2.19	2.30	2.19	2.07	2.79	2.66	2.66	2.19	1.84	2.19	1.96	2.19	1.09	1.62	2.42	2.54	1.30	2.54	2.30	1.73	2.19	1.84	2.42	2.42	0.68	0.28	1.09	2.54	0.38	1.62	0.48	2.66	0.08	1.49	1.51	0.78	0																													
YO6	2.42	2.42	3.43	3.16	1.96	3.04	3.29	2.19	2.54	2.42	2.07	2.79	2.66	2.66	2.19	1.84	1.96	1.96	2.19	1.30	1.62	2.42	2.79	1.30	2.54	2.30	1.73	2.42	1.84	2.19	2.42	0.68	0.28	1.09	2.54	0.19	1.84	0.48	2.66	0.08	1.49	1.51	0.98	0.19	0																													
YO9	2.79	2.30	3.29	3.29	2.07	2.91	3.16	2.30	2.66	2.54	2.19	2.91	2.54	2.54	2.07	2.19	2.30	2.07	2.30	1.40	1.30	2.54	2.91	1.40	2.42	2.19	1.62	2.54	1.73	2.54	2.07	0.53	0.19	0.98	2.66	0.28	1.73	0.38	2.54	0.19	1.51	1.84	1.09	0.28	0.28	0																													
YO10	2.54	2.30	3.29	3.04	1.84	2.91	3.43	2.07	2.42	2.30	2.19	2.66	2.54	2.54	2.07	1.96	2.07	1.84	2.07	1.19	1.51	2.54	2.66	1.19	2.42	2.19	1.62	2.30	1.73	2.30	2.30	0.58	0.19	0.98	2.42	0.28	1.73	0.38	2.54	0.01	1.30	1.62	0.88	0.09	0.09	0.19	0																													
YO11	3.97	4.84	5.45	4.54	3.97	4.11	4.99	5.15	5.30	4.54	3.29	4.99	5.15	5.15	4.84	4.39	4.54	4.25	4.84	2.91	3.56	6.13	4.11	3.69	3.83	4.11	3.43	4.25	3.83	4.25	4.54	3.43	3.16	3.69	4.11	3.25	3.83	3.16	4.25	2.91	4.11	4.25	3.29	3.04	3.04	3.16	2.91	0																													
YO12	2.79	2.54	3.29	3.29	2.07	2.66	3.43	2.30	2.91	2.30	2.19	2.91	2.79	2.79	1.84	2.42	2.07	2.30	2.30	1.62	1.51	2.30	2.91	1.40	2.42	1.96	1.84	2.54	1.73	2.30	2.54	0.19	0.38	0.98	2.91	0.48	1.96	0.19	2.79	0.38	1.73	1.84	1.09	0.48	0.48	0.38	0.38	3.16	0																												
YO13	3.04	2.30	3.29	3.29	2.07	3.16	3.43	2.54	2.66	2.54	2.42	2.91	3.04	3.04	2.30	2.19	2.30	2.07	2.30	1.40	1.30	2.30	2.66	1.19	2.66	2.42	2.07	2.79	1.73	2.79	2.30	0.78	0.38	1.19	2.91	0.48	1.96	0.58	2.54	0.38	1.73	2.07	1.30	0.48	0.48	0.38	0.38	3.16	0.58	0																												

	VAH1	VAH2	VF1	VF2	VH1	VH2	VP1	VP3	VR1	VR2	VW1	VW2	VW3	VW4	VW5	VW6	VW7	VW8	VW9	WO11	YAH	YM2	YM3	YM4	YT2	YT3	YT4	YW2	WG10	YF1	YI2	YR5	YR6	YL1	YL2	YL4	YL5	YS2	YS3	YO1	YO2	YO4	YO6	YO8	YO10	YO11	YO12	YO13
VAH1	0																																																
VAH2	5.56	0																																															
VF1	6.04	6.04	0																																														
VF2	6.04	5.11	4.88	0																																													
VH1	4.88	4.29	5.11	5.55	0																																											
VH2	4.90	5.33	5.80	4.90	4.49	0																																										
VP1	5.33	5.80	5.80	5.80	5.33	7.70	0																																									
VP3	4.88	5.11	3.91	4.88	3.91	4.49	6.83	0																																								
VR1	4.90	4.10	5.80	4.10	4.49	5.11	4.88	4.90	0																																							
VR2	6.55	5.55	4.29	6.04	4.88	4.49	6.83	3.91	4.49	0																																						
VW1	5.80	3.73	4.90	4.49	4.49	3.91	5.11	4.10	4.29	4.49	0																																					
VW2	5.33	6.29	8.01	7.49	5.33	7.70	6.55	6.29	6.55	5.80	5.11	0																																				
VW3	5.56	6.04	6.04	6.55	5.55	6.29	8.67	4.29	5.33	3.91	5.33	4.10	0																																			
VW4	5.55	6.04	6.04	6.55	5.55	6.29	8.67	4.29	5.33	3.91	5.33	4.10	0.10	0																																		
VW5	4.88	4.22	3.91	4.88	4.88	4.90	5.33	3.22	4.49	3.91	4.90	6.29	5.11	5.11	0																																		
VW6	5.33	4.10	4.49	6.29	3.06	6.04	6.04	3.73	5.11	4.49	3.55	5.55	4.90	4.90	3.73	0																																	
VW7	4.88	4.88	5.55	6.55	3.91	6.83	6.29	3.55	5.33	4.88	4.90	5.33	5.55	5.55	2.90	3.06	0																																
VW8	5.11	5.11	4.88	5.11	3.55	4.90	5.33	3.91	3.73	4.29	4.90	4.10	4.29	4.29	4.88	3.33	3.91	0																																
VW9	4.88	5.11	4.88	6.04	4.88	5.33	4.90	3.55	4.49	3.55	4.90	5.33	4.29	4.29	4.88	3.33	3.55	2.33	0																																
WO11	3.22	3.55	5.11	5.11	2.90	3.73	5.33	3.91	3.73	2.90	3.33	4.10	3.91	3.91	3.91	2.45	3.55	2.90	2.33	0																																
YAH	4.90	4.49	5.80	5.80	3.35	5.11	3.91	4.49	3.91	4.49	4.88	4.29	4.90	4.90	3.33	4.29	2.74	3.33	4.10	3.06	0																																
WO11	4.88	5.11	6.04	6.04	5.11	5.33	7.49	4.29	5.80	4.88	4.90	5.33	4.29	4.29	3.91	4.90	3.55	3.55	4.29	3.55	3.73	0																																
YAH	5.80	5.33	5.80	8.01	3.06	8.55	7.70	5.33	5.55	4.49	6.55	6.55	6.29	6.29	5.80	3.22	4.49	4.49	4.49	2.45	4.88	5.33	0																																
YM2	4.88	4.29	3.91	3.91	3.22	4.90	4.90	2.02	4.10	3.55	3.73	4.49	4.88	4.88	2.90	3.06	3.22	2.90	3.55	2.90	2.74	3.91	3.73	0																																
YM3	5.80	5.33	3.73	4.90	4.90	4.29	6.55	4.90	5.11	4.10	5.11	5.56	5.33	5.33	3.33	4.88	3.33	4.10	4.49	4.10	4.88	6.29	6.04	3.73	0																																
YM4	4.10	3.33	4.90	6.29	4.49	3.55	3.91	5.33	5.11	4.49	4.29	6.04	5.80	5.80	3.73	5.11	5.33	5.33	3.73	3.73	3.55	5.33	6.55	4.10	5.11	0																																
YF2	4.88	4.88	6.55	5.55	3.22	4.49	4.10	4.88	3.73	3.91	3.33	4.49	4.29	4.29	3.91	4.10	3.91	3.22	4.29	2.53	3.06	5.11	4.90	3.55	4.49	4.49	0																																
YF3	3.55	5.55	5.11	6.04	3.55	4.90	7.49	4.29	5.80	3.91	5.33	5.33	4.88	4.88	3.22	4.10	3.22	3.91	4.88	3.22	4.10	3.91	4.49	4.29	3.33	4.90	5.11	0																																
YF4	3.73	3.73	4.90	4.49	3.06	3.55	5.55	4.10	4.29	4.10	3.55	6.04	4.90	4.90	4.10	3.91	4.10	3.91	3.73	2.45	3.22	3.73	4.29	3.73	3.91	3.91	3.33	3.33	0																																
WP2	5.11	4.29	6.04	6.04	3.91	4.10	6.83	3.56	5.80	3.22	3.33	6.83	5.11	5.11	3.91	3.06	3.91	5.11	4.29	3.55	4.90	4.29	5.33	3.91	4.90	4.90	3.22	5.11	3.73	0																																
WG10	6.04	4.29	3.55	5.56	4.29	4.49	5.33	4.29	4.90	4.88	4.49	7.49	4.88	4.88	5.11	5.33	5.55	4.29	3.91	3.55	4.49	5.55	4.90	3.55	4.90	5.33	3.55	6.04	3.73	4.88	0																																
YF1	3.22	2.90	4.29	4.29	2.59	3.06	4.49	2.59	3.39	2.90	2.74	3.73	2.90	2.90	2.33	3.06	2.53	2.90	2.90	2.02	1.88	2.90	3.73	1.76	3.06	2.16	2.33	3.22	1.88	2.90	3.22	0																																
YI2	3.22	2.53	3.91	3.55	2.33	3.39	4.10	2.33	2.74	2.90	2.45	3.39	3.22	3.22	2.33	2.45	2.53	2.33	2.59	1.51	1.63	2.90	3.39	1.51	2.74	2.45	2.02	2.90	1.89	2.90	2.59	0.59	0																																
YR5	2.90	2.90	5.11	5.11	2.90	3.73	4.10	3.55	3.39	3.91	3.33	4.10	4.29	4.29	2.90	3.39	2.90	3.22	3.22	2.02	2.16	3.91	3.39	2.33	3.33	2.16	2.33	3.91	2.45	4.29	3.55	1.03	1.03	0																																
YR6	4.90	6.83	4.49	5.33	3.73	5.55	5.11	3.33	5.55	5.80	4.88	6.04	5.33	5.33	4.90	4.29	4.49	3.73	4.49	4.10	3.91	6.29	5.55	3.33	3.55	4.88	3.73	3.33	3.91	5.80	4.90	3.73	3.06	3.73	0																																
YL1	3.06	2.74	4.10	3.73	2.45	3.55	3.91	2.45	2.90	3.06	2.33	3.55	3.39	3.39	2.45	2.33	2.45	2.45	2.74	1.53	1.76	2.74	3.55	1.53	2.90	2.59	2.16	3.06	2.02	2.74	2.74	0.70	0.09	1.15	3.22	0																																
YL2	3.73	4.10	3.73	4.49	3.06	4.29	5.55	4.49	3.22	4.10	4.29	5.55	5.33	5.33	2.45	3.55	3.33	3.73	4.10	2.45	2.90	4.10	4.29	3.73	3.22	3.91	3.73	2.45	3.55	4.10	3.73	2.45	1.88	2.74	3.91	2.02	0																																
YL4	3.22	2.90	3.91	3.91	2.59	3.39	4.49	2.59	3.06	2.90	2.45	3.39	3.22	3.22	2.33	2.45	2.33	2.33	2.33	1.51	1.63	2.59	3.06	1.51	2.74	2.45	2.33	2.90	1.88	2.90	2.90	0.39	0.39	1.03	3.33	0.49	2.16	0																																
YL5	4.29	3.55	4.29	5.55	4.88	5.80	5.80	5.11	5.80	6.55	5.80	6.83	7.70	7.70	4.29	4.10	3.91	3.91	4.53	3.55	4.10	5.11	4.10	3.91	3.38	4.49	5.11	4.29	3.38	6.04	5.55	3.55	2.90	3.55	3.33	3.06	3.73	2.90	0																																
YS2	2.90	2.53	3.91	3.55	2.02	3.33	4.10	2.33	2.74	2.53	2.45	3.06	2.90	2.90	2.33	2.16	2.33	2.02	2.33	1.26	1.63	2.90	3.06	1.26	2.74	2.45	1.76	2.53	1.88	2.53	2.59	0.59	0.19	1.03	2.74	0.29	1.88	0.39	2.90	0																																
YS3	3.73	3.73	5.33	3.73	3.06	3.55	7.11	2.74	4.29	4.90	3.22	4.29	4.10	4.10	4.10	3.22	3.73	3.33	3.73	2.45	3.55	4.49	4.29	2.45	3.51	3.55	3.06	4.10	2.90	3.73	4.10	1.88	1.63	2.16	3.55	1.76	3.22	1.88	4.10	1.38	0																																
YO1	3.91	3.55	3.22	5.11	3.22	4.10	5.33	3.22	3.33	3.91	4.10	5.80	4.88	4.88	3.55	3.06	3.22	2.53	3.22	2.53	3.73	4.29	3.73	2.53	3.33	3.73	3.91	3.22	3.33	3.91	3.55	2.02	2.02	2.59	4.49	1.88	2.74	2.02	3.91	1.76	3.33	0																																
YO2	4.49	4.10	5.33	4.90	2.74	4.28	5.11	3.33	3.22	2.74	3.22	3.56	4.10	4.10	3.06	2.90	2.74	2.74	3.73	2.16	2.30	3.73	3.91	2.16	3.55	3.55	2.16	3.33	2.90	2.45	3.73	1.38	1.15	2.16	3.91	1.26	2.59	1.38	4.10	0.92	2.59	2.45	0																																
YO4	3.06	2.74	3.73	3.33	2.16	3.22	4.29	2.45	2.59	2.45	2.30	3.22	3.06	3.06	2.45	2.02	2.45	2.16	2.45	1.15	1.76	2.74	2.90	1.38	2.90	2.59	1.88	2.45	2.02	2.74	2.74	0.70	0.29	1.15	2.90	0.39	1.76	0.49	3.06	0.09	1.51	1.63	0.81	0																																
YO6	2.74	2.74	4.10	3.73	2.16	3.56	3.91	2.45	2.90	2.74	2.30	3.22	3.06	3.06	2.45	2.02	2.16	2.16	2.45	1.38	1.76	2.74	3.22	1.38	2.90	2.59	1.88	2.74	2.02	2.45	2.74	0.70	0.29	1.15	2.90	0.19	2.02	0.49	3.06	0.09	1.51	1.63	1.03	0.19	0																																
YO8	3.22	2.53	3.91	3.91	2.30	3.39	3.73	2.59	3.06	2.90	2.45	3.39	2.90	2.90	2.30	2.45	2.59	2.30	2.59	1.51	1.38	2.90	3.39	1.51	2.74	2.45	1.76	2.90	1.88	2.90	2.30	0.59	0.19	1.03	3.06	0.29	1.88	0.39	2.90	0.19	1.63	2.02	1.15	0.29	0.29	0																																
YO10	2.90	2.59	3.91	3.55	2.02	3.33	4.10	2.30	2.74	2.59	2.45	3.06	2.90	2.90	2.30	2.16	2.30	2.02	2.30	1.26	1.63	2.90	3.06	1.26	2.74	2.45	1.76	2.59	1.88	2.55	2.59	0.59	0.19	1.03	2.74	0.29	1.88	0.39	2.90	0.01	1.36	1.76	0.92	0.09	0.09	0.19	0																																
YO11	4.90	6.29	7.49	5.80	4.90	5.11	6.55	6.83	7.11	5.80	3.91	6.55	6.83	6.83	6.29	5.55	5.80	5.33	6.29	3.39	4.29	6.67	5.11	4.49	4.88	5.11	4.10	5.33	4.88	5.33	5.80	4.10	3.73	4.49	5.11	3.91	4.88	3.73	5.33	3.39	5.11	5.33	3.91	3.56	3.56	3.73	3.39	0																																
YO12	3.22	2.90	3.91	3.91	2.30	3.06	4.10	2.53	3.39	2.53	2.45	3.39	3.22	3.22	2.02	2.74	2.30	2.53	2.53	1.76	1.63	2.53	3.39	1.51	2.74	2.16	2.02	2.90	1.88	2.53	2.90	0.19	0.39	1.03	3.39	0.49	2.16	0.19	3.22	0.39	1.88	2.02	1.15	0.49	0.49	0.39	0.39	3.73	0																																
YO13	3.56	2.53	3.91	3.91	2.30	3.73	4.10	2.90	3.06	2.90	2.74	3.39	3.56	3.56	2.53	2.45	2.53	2.30	2.53	1.51	1.38	2.53	3.06	1.26	3.06	2.74	2.30	3.22	1.88	3.22	2.59	0.81	0.39	1.26	3.39	0.49	2.16	0.59	2.90	0.39	1.88	2.30	1.38	0.49	0.49	0.39	0.39	3.73	0.59	0																																

2.93% (ranging from 0.01 to 5.88). Within the subgroups, the maize inbred lines from subgroup A had an average distance between them of 2.83% (ranging from 0.1 to 5.56), in subgroup B the average distance between the lines were 3.79% (ranging from 2.02 to 5.56), in subgroup C the average distance between the lines were 3.47% (ranging from 2.45 to 4.49) and in subgroup D the average distance between the lines were 1.62% (ranging from 0.01 to 3.22). Looking at the minor groups the average distances were all lower that in their subgroup. Therefore, the trend of the clustering pattern was more variation in the major clusters with less variation between lines in subgroups and even less variation within the minor groups. It was also further illustrated that the local and foreign inbred lines were not grouped separately.

Grouping of inbred lines into clusters of genetically related lines

When the different dendograms and consensus trees were compared with each other (Fig. 4.2 to 4.7), the lines were grouped into two major groups (I and II). Group I consisted of two subgroups (A and B) and group II consisted of two subgroups (C and D). Subgroup A was divided in three minor groups: 1 (WO5 and WO7), 2 (WO11, WO1, WO2 and WK2) and 9 (WO4, WI1, YM1 and WO8), subgroup B was divided in five minor groups: 7 (YF3, YL2, WO3 and WM1), 4(YO1, WF1 and WO10), 5 (YR6, YL5 and YM3), 6 (WP3, YM2 and YS3) and 3 (WR1, WF2, YM4, WP1 and WM2), subgroup C was divided in two minor groups: 8 (YO11, WK1 and YF2) and 10 (WR2, WP2 and YO2), and subgroup D was divided in four minor groups 11 (YF1, YO12, YL4 and YR5), 12 (YO13 and WO9), 13 (YL1, YI2 and YO9) and 14 (YO4, YO6, YO10, YS2 and WO6). WI2 and YF4 were not contained in any specific cluster.

If the Jukes Cantor distance tree (Fig. 4.2) and – consensus tree are compared (Fig. 4.4) with each other, they are very similar. Group I in the dendogram is divided in subgroup A, C and B, with group II only consisting of subgroup D. In the consensus tree group I is divided in subgroup A and B, and group II is divided in subgroup C and D. In subgroup A of the dendogram the minor groups are grouped starting with minor group two, then one and then nine. In the consensus tree subgroup A is divided in minor group one, followed by two, followed by nine. In

subgroup B the minor groups in the dendogram grouped as follows: minor group three, followed by four and five, followed by seven and six. In the consensus tree, subgroup B grouped as follows: minor groups four and seven, followed by three, followed by five and six. Subgroup C and D had the same minor groups in both trees. Looking at the minor groups, WO11 was not included in minor group two of the dendogram, but it was included in the consensus tree. In minor group seven of the distance tree, WO3 clustered with WM1 and WO6, but last mentioned was not included in minor group seven in the consensus tree. In the distance tree YO9 was included in minor group 12, but in the consensus tree it was included in minor group 13. It was also clear that the dendogram had more resolution at the bottom of the tree, than the consensus tree. In general, however, the two trees were closely correlated.

If the Kimura distance tree (Fig. 4.3) and – consensus tree (Fig. 4.5) are compared with each other, they are also very similar to each other. Group I and II consisted of the same subgroups in both trees. In subgroup A of the dendogram the minor groups are grouped starting with minor group two, then one and then nine. In the consensus tree subgroup A is divided in minor group one, followed by two, followed by nine. In subgroup B the minor groups in the dendogram grouped as follows: minor group three, followed by four and five, followed by seven and six. In the consensus tree, subgroup B grouped as follows: minor groups six and five, followed by seven, followed by three and four. Subgroup C and D had the same minor groups in both trees. In minor group two of the consensus tree the line WO11 was included in the cluster, but not in the dendogram. In minor group seven in the dendogram WO3 clustered with WM1 and WO6, but they were not included in minor group seven in the consensus tree. In the distance tree YO9 was included in minor group 12, but in the consensus tree it was included in minor group 13. In these trees, it was also clear that the dendogram had more resolution at the bottom of the tree, than the consensus tree. In general however, the two trees were closely correlated.

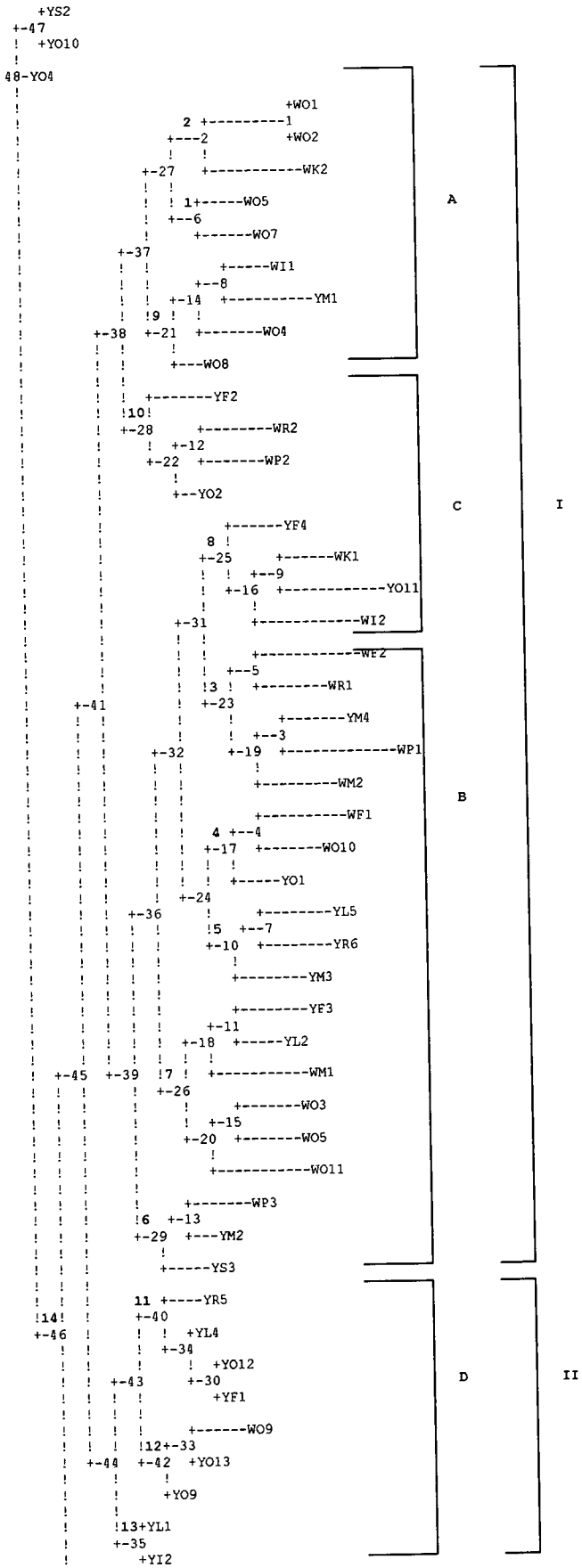


Fig. 4.2 Dendogram generated by Neighbour joining analysis from genetic distance matrix determined by Jukes and Cantor. The values on the dendogram represent the genetic distances among different inbred lines.

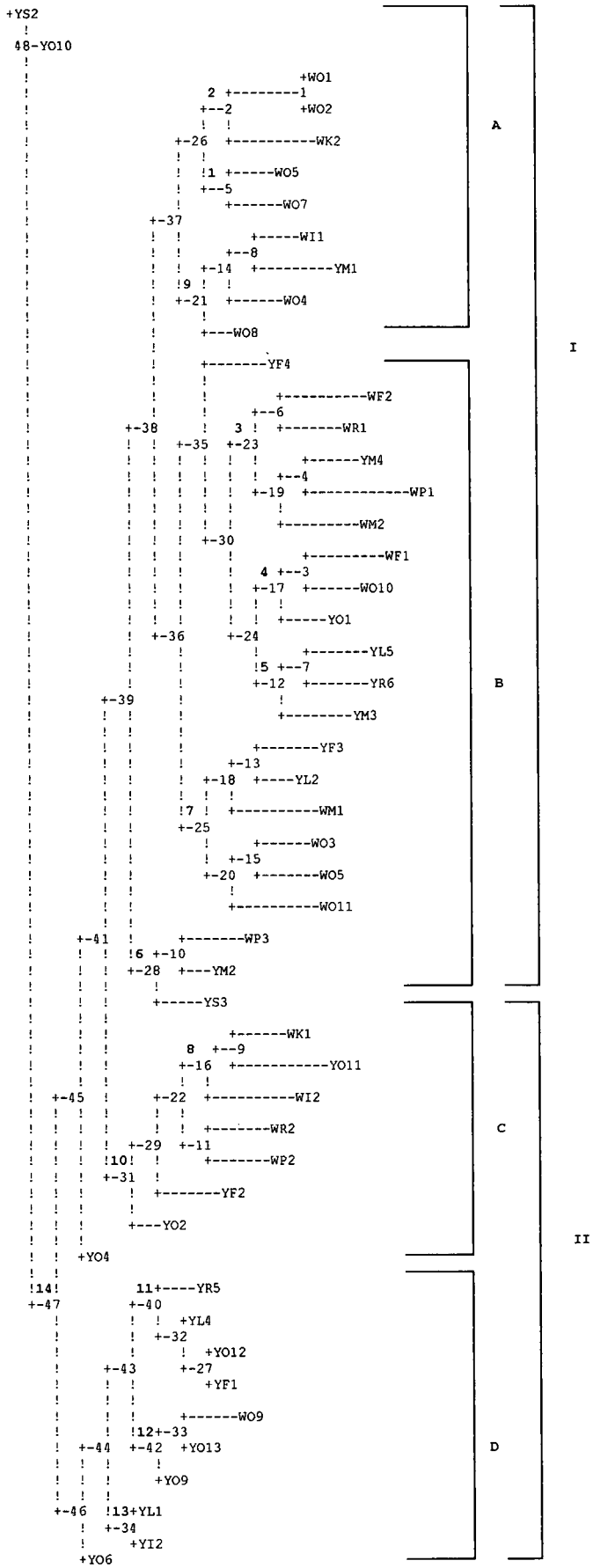


Fig. 4.3 Dendogram generated by Neighbour joining analysis from genetic distance matrix determined by Kimura.

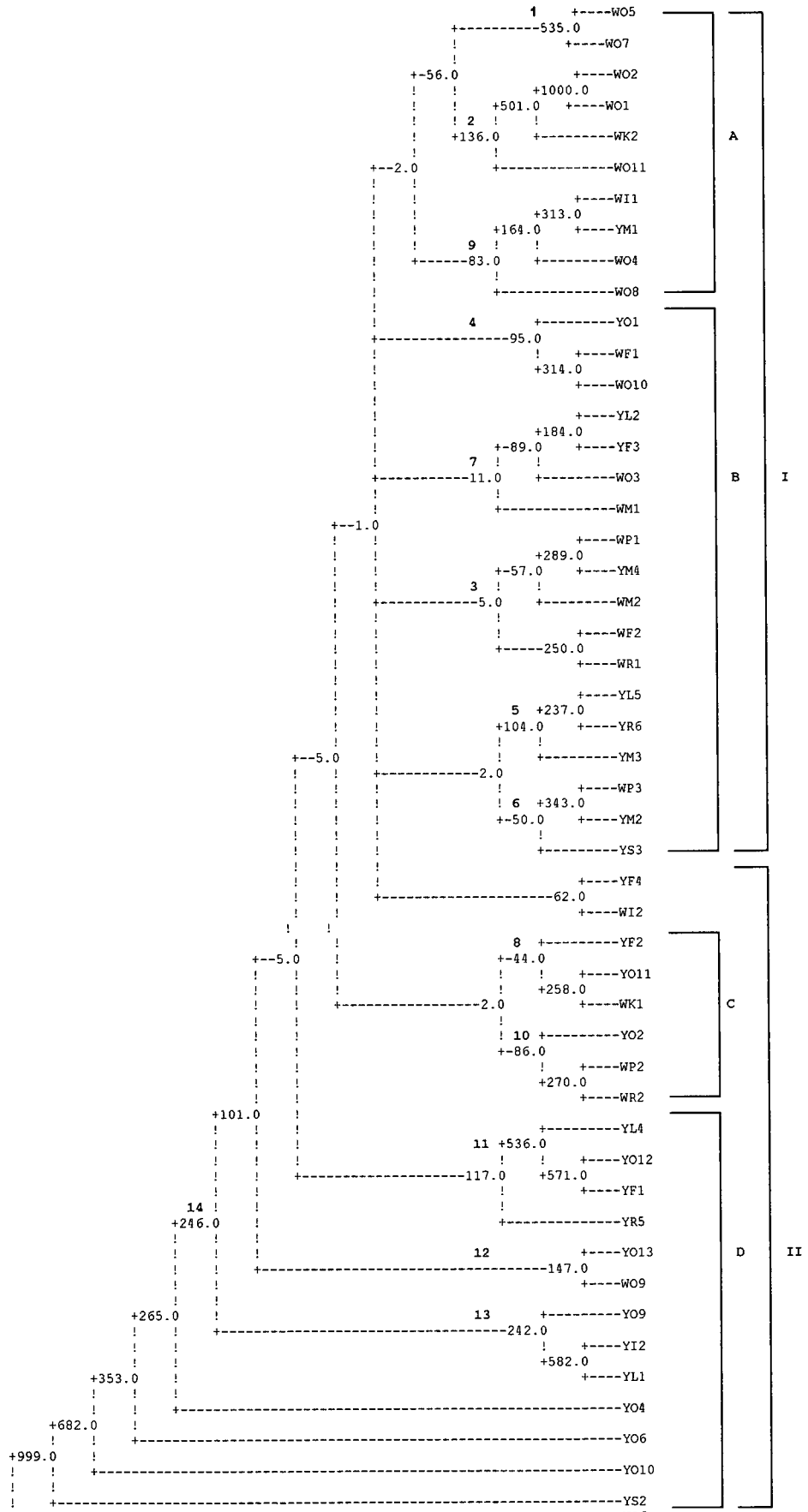


Fig. 4.4 Consensus tree generated by Neighbour joining analysis from genetic distance analysis determined by Jukes Cantor.

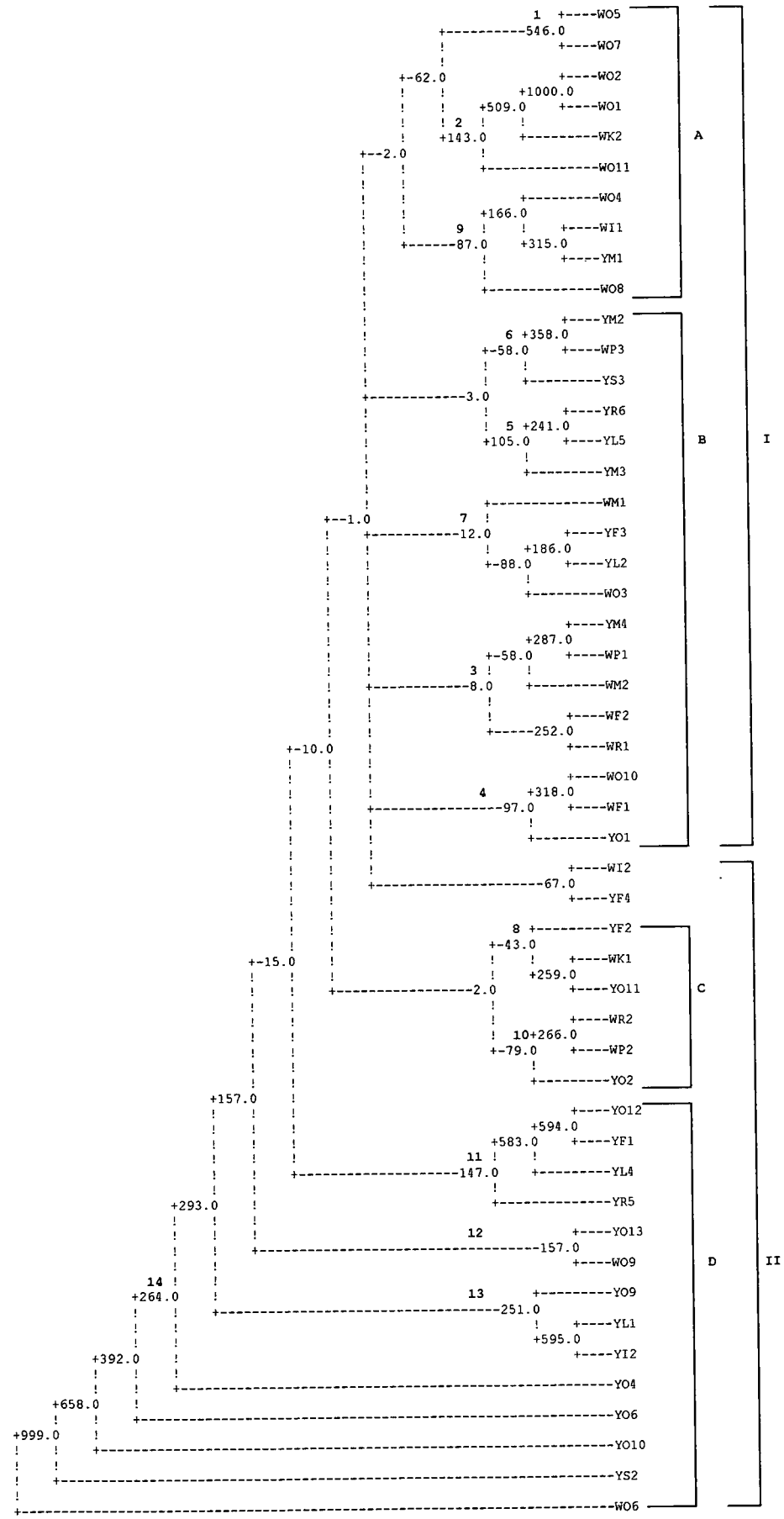


Fig. 4.5 Consensus tree generated by Neighbour joining analysis from genetic distance analysis determined by Kimura.

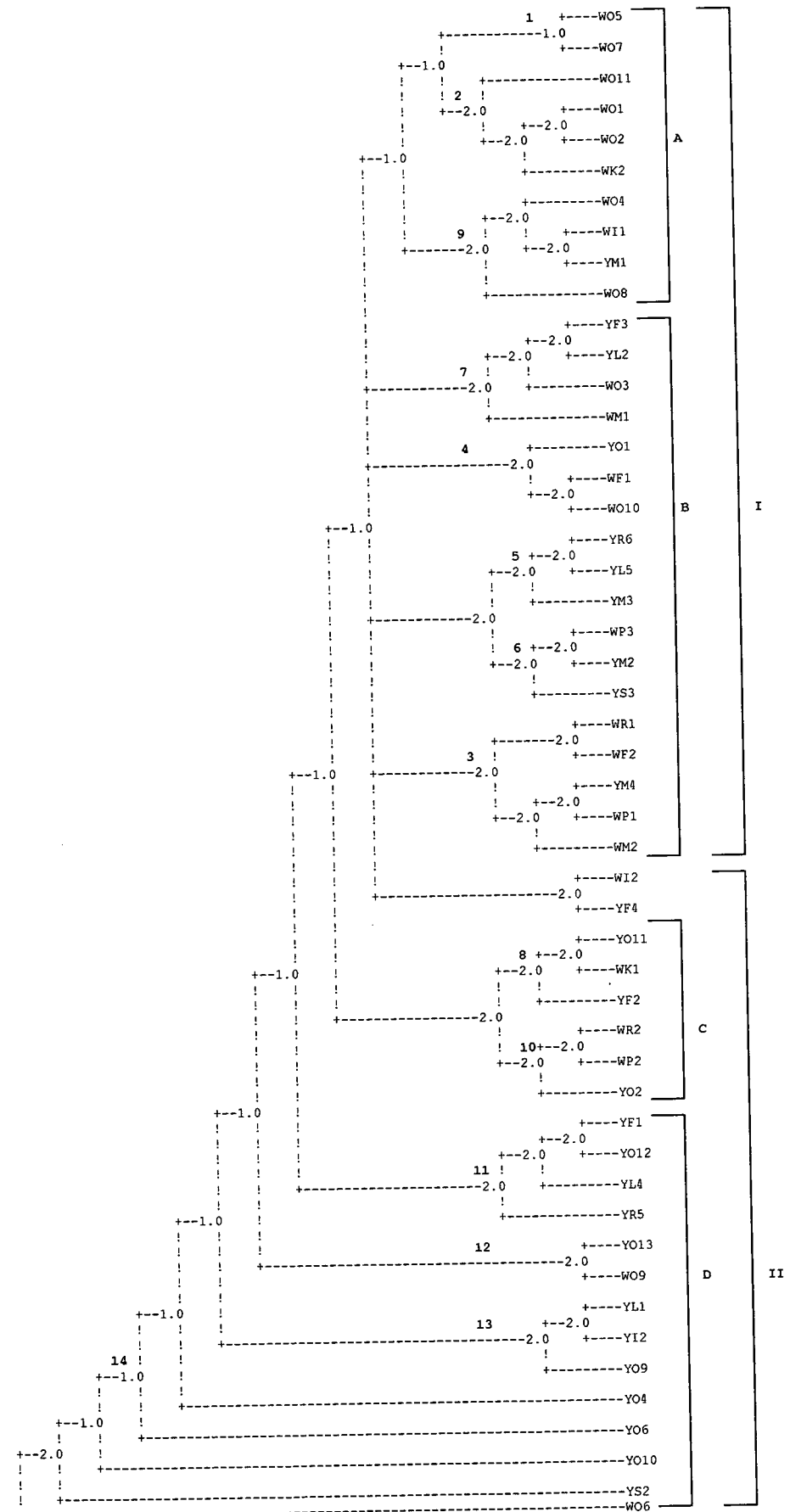


Fig. 4.6 Consensus tree generated by Neighbour joining analysis from the combination of genetic distance matrixes determined by Jukes and Cantor, and Kimura.

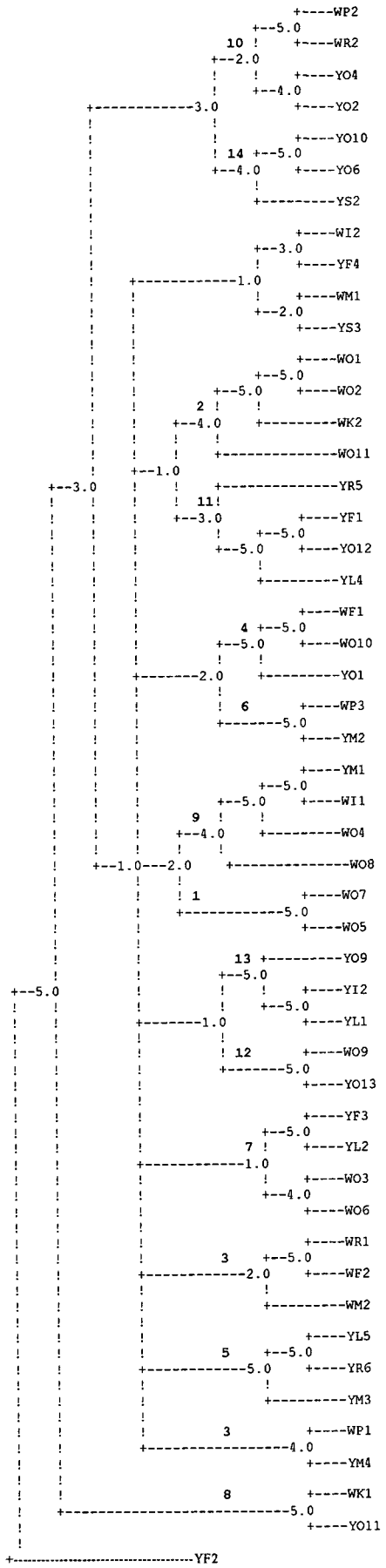


Fig. 4.7 Consensus tree generated by 1000 bootstrap replicates using DNA parsimony

When comparing the Jukes Cantor and Kimura (Fig. 4.2 and 4.4) generated trees it was evident that these trees were similar. Group I in the Jukes Cantor dendogram, was divided in subgroup A, C and B, with group II only consisting of subgroup D. In the Kimura dendogram, group I was divided in subgroup A and B, and group II was divided in subgroup C and D. The subgroups for both trees were the same. Line YF4 was grouped in minor group eight in the Jukes Cantor tree, but in the Kimura tree it was grouped separately.

If the two consensus trees (Fig. 4.4 and 4.5) are compared, one using the Jukes Cantor algorithm and the other one using the Kimura algorithm to determine the genetic distances, the trees look very similar. The only difference between the two trees is that in subgroup B, the order in which the minor groups clustered, differed. In the Jukes Cantor tree the groups clustered in the following order: four, seven, three, five and six; where as in the Kimura tree they clustered in the following order: six, five, seven, three and four. Although the order of the groups were different, all the groups had the same inbred lines clustered together and the rest of the trees were the same.

If the two consensus trees (Fig. 4.3 and 4.5) were compared individually with the combined consensus tree (Fig. 4.6), the only differences that occur once again, is the order in which the minor groups cluster together in subgroup B. In the combined consensus tree the minor groups clustered in the following order: seven, four, five, six and three.

If the combined consensus tree and the DNA parsimony consensus tree were compared, the groups and subgroups were clustered differently. Looking at the minor groups of the two trees, there were only a few differences. The parsimony tree had minor group three and five combined in one cluster, but in the combined tree they were clustered separately. Minor group six didn't include YS3 in the parsimony tree, but it was included in the combined tree. Minor group seven didn't include WM1 in the parsimony tree (like in the combined tree), but it included WO6 additionally. In the parsimony tree minor group ten had an additional line YO4. It was also clear that the resolution at the end of the parsimony tree was better than that of the combined consensus tree. The lines that formed part of

minor group 14 in the combined consensus tree (that had poor resolution), seemed to be the lines that grouped differently in the parsimony tree.

Discussion and conclusion

The results of this study indicated that, the lines were distributed into two major groups, four subgroups and 14 minor groups. From the distribution of lines in the dendogram the coherence of the AFLP technique and the analytical methods can clearly be seen, indicating its potential as a powerful tool for line identification. Although there were different groupings, the genetic distances were not that high, in other words the inbred lines are related to a large degree.

There were no definite groupings with hierarchical structure according to the genetic distances, the reason was that there was no hierarchical grouping in the group of maize inbred lines. The group of inbred lines were not developed in a natural selection system, and the relationship between the inbred lines were close. The clusters indicated the absence of any relation between geographic diversity and genetic diversity. This therefore, revealed a relatively low degree of genetic variability within the species. A possible reason can be that the inbred lines have been used all over the world to improve existing hybrids. This interbreeding has caused the variation in the group of material, in terms of place of origin, to be much less.

Initial experiments with AFLP fingerprinting of a number of plant DNA's indicated that AFLP primers with at least three selective nucleotides at both the PstI and MseI primer were required to generate useful band patterns. Because primers with three selective bases tolerate a low level of mismatch amplification a two-step amplification strategy was developed for AFLP fingerprinting of complex DNAs. With the pre-amplification strategy the background smears in the fingerprint patterns were reduced, and fingerprints with particular primer combinations lacked one or more bands compared with fingerprints generated without preamplification. This is best explainded assuming that the direct preamplification with AFLP primers having three selective nucleotides resulted in a low level of mismatch amplification products, which caused the background

smears and gave discrete amplified fragments corresponding to repeated restriction fragments. An aditional advantage of the two-step amplification strategy is that it provided a virtually unlimited amount of template DNA for AFLP reactions (Vos *et al.*, 1995).

Only the reproducible amplification products that could be unambiguously scored in all the lines were included in the study. Bands that did not qualify were either too faint to score accurately or not reproducible. Some of the minor (low intensity) amplification products were not reproducible in all cases and were consequently discarded. Major (high intensity) products were generally reproducible across a range of near-optimum amplification conditions (Micheli *et al.*, 1994).

Lines WI2 and YF4 were very inconsistent, a possible reason for this could be because the lines are very inbred, and they had no predominant characters that could group them with a specific cluster, so they grouped together. This may also be the reason why their position in the trees was not consistent. A more consistent result may have been obtained if white and yellow types were treated separately.

The number of DNA markers required to distinguish any group of lines depends on the genetic variability among the lines and the size of the group (Myburg *et al.*, 1996). AFLP analysis was found to be highly effective in distinguishing genotypes from geographically separate areas (Pakniyat *et al.*, 1997). Good coverage has been demonstrated for AFLP's from PstI-MseI DNA digests in different lines (Powell *et al.*, 1997). Barrett and Kidwell (1998) have used 16 AFLP primers to detect genetic relationships among 54 wheat lines. The AFLP primer pairs detected 229 polymorphic bands. Bands excluded were those that in the amplification profiles appeared or disappeared when repeating the reaction across time and materials. A high level of polymorphism was detected among the AFLP markers and none of them were unique to distinguish the lines. Clustering of different lines have been done using AFLP's, for example wheat (Bohn *et al.*, 1999; Barrett *et al.*, 1998) and barley (Pakniyat *et al.*, 1997).

AFLP analysis has accelerated the use of DNA-based fingerprinting data for the identification and genetic analysis of lines, varieties and breeding lines (Vos *et al.*, 1995). This is certainly, in part, due to the fact that the technique does not require any previously obtained sequence data. This requirement has thus far precluded the use of most genome analysis techniques in all but the major crop species. We found that optimised reaction conditions were an absolute requirement for the generation of reproducible and informative AFLP profiles in maize. DNA samples were found to be stable and reproducible. These results indicated that the AFLP analysis could be used for the identification of closely related maize lines.

The results from this study will contribute to a large extent to the decisions that the breeder will make in the breeding programme. Since the inbred lines are grouped into clusters and the genetic distances between the lines are known, it is possible to make crosses between lines that will yield higher combining ability. The reason is that the shorter the genetic distance between two lines, the closer related the lines are. The closer related two lines are, that are crossed with each other, the lower the combining ability. The breeder will thus cross lines that have a larger genetic distance between them.

The implication of these results to the breeder and the breeding programme are significant. Knowing how closely related lines are to each other, is a good indication of how good the combining ability would be if two lines were to be crossed. This in return could make the whole breeding programme more labour effective, time effective, and cost effective, since less unnecessary crosses will be made, hybrids can possibly be developed quicker and because of the breeding programme being more structured in terms of quicker returns and possibly less input costs.

In conclusion, the results of this study indicate that AFLP analysis can be used for the identification and genetic analysis of maize lines. This study generated a combination of polymorphic AFLP loci that was used to construct a line fingerprinting profile. Genetic distance analysis of the maize inbred lines based on 107 AFLP loci indicated high genetic similarity among the groups and high genetic variation within the groups of lines included in the study. This data allows

the breeder, when making crosses between individuals that have been tested, to make crosses with higher combining ability, because of their genetic relatedness being known. That in return makes a breeding programme much more effective and saves a lot of time. This documentation will therefore allow further scientific studies on crosses that are made in the maize programme. It will also possibly contribute to more effective hybrid breeding.

CHAPTER 5

COMPARISON OF DIFFERENT METHODS OF GROUPING MAIZE INBRED LINES IN CLUSTERS.

Abstract

The study of genetic variation is a primary concern in population biology and is fundamental in the improvement of agricultural plants (Melchinger, 1993). Comparing results of different genetic diversity estimation methods may be indicative of their utility as a tool for plant breeders. Sixty-eight maize inbred lines from the Advanta Africa genebank were used to compare morphological characterisation, restriction fragment length polymorphism (RFLP) and amplified fragment length polymorphism (AFLP)-based genetic diversity estimates. The AFLP and RFLP analysis detected a similar clustering pattern of genetic diversity among the maize inbred lines, but it was difficult to compare the two methods since the available RFLP data were limited. The morphological clustering was different to those of the DNA based techniques. This emphasised the problems related to morphological characterisation, and indicated that DNA methods would be more reliable. The DNA based methods succeeded to identify groups that were genetically further apart, which will yield better hybrid progeny. Looking at the comparison of these methods to cluster the different inbred lines into clusters, the DNA techniques are definitely more reliable than the morphological characterisation.

Introduction

Obtaining accurate estimates of genetic diversity levels among germplasm sources may increase efficiency of breeding efforts to improve crop species (Barret *et al.*, 1998). Assignment of inbreds to heterotic groups and selection of donor lines for improving parents of an elite single cross are major decisions facing any maize breeder (Dudley *et al.*, 1991). Various methods, including morphological marker (Van Beuningen and Busch, 1997) and DNA marker (Karp *et al.*, 1996) analysis, have been used to quantify genetic diversity and relationships among genotypes.

Most, if not all, currently important heterotic groups in maize have been established empirically by relating the heterosis observed in crosses with the origin of their parents. Further classification criteria were based on morphological traits (Messmer *et al.*, 1992). Although morphology has proved useful for classifying maize races and populations (Goodman and Brown, 1988), it may not be appropriate for elite breeding germplasm. Furthermore, morphological characteristics can be affected by environmental conditions (Goodman and Paterniani, 1969). Smith and Smith (1989) concluded that morphological data may be a starting point in determining distinctiveness, but that RFLP (or AFLP) data offer better resolution of relatedness of maize inbreds (Gerdes and Tracy, 1994).

Restriction fragment length polymorphisms (RFLP) detect DNA polymorphisms through restriction endonuclease digestions followed by visualisation via DNA blot hybridisation (Mazur and Tingey, 1995). RFLP's offer several advantages as molecular markers. RFLP's are unaffected by environment, and show co-dominant inheritance. However, RFLP's are more polymorphic, more numerous and are better distinguished throughout the genome. Accordingly, RFLP's have been suggested as superior markers to assess diversity and relationships in maize breeding germplasm (Murray *et al.*, 1988; Walton and Helentjaris, 1987).

Amplified fragment length polymorphism (AFLP) is a recently developed DNA marker analysis system based on a novel, powerful combination of polymerase chain reaction (PCR) and restriction enzyme analysis (Vos *et al.*, 1995). AFLP's assay the presence/absence of restriction enzyme sites in combination with sequence polymorphisms adjacent to these sites (Mazur and Tingey, 1995). AFLP's are highly efficient compared with other DNA marker systems, they are reproducible, and exhibit intraspecific homology (Rouppe van der Voort *et al.*, 1997; Powell *et al.*, 1996; Tohme *et al.*, 1996).

Since this technology was recently developed, the relationship between AFLP-based, RFLP-based analysis and morphological characterisation has not been determined for maize. The objective of this study was to compare AFLP, RFLP and morphological genetic diversity estimates generated for 68 maize inbred lines from Advanta Africa Seeds.

Material and Methods

Plant Material

Sixty-eight maize inbred lines from the Advanta Africa genebank, originating from five different countries, were used in this study. These lines were all analysed using the morphological characterisation and AFLP analysis. However, only seven (YR5, YO11, YL1, YL2, YL4, YM4 and YL5) of the 68 lines were used in the RFLP analysis.

Agronomic traits

As described in chapter three.

Molecular marker assays

AFLP assays

As described in chapter four.

RFLP assays

GARST Laboratories (Advanta Africa's sister company) did the RFLP analysis on these seven inbred lines prior to the AFLP analysis. They presented us with the results of the research study. Advanta Africa Seeds in South Africa interpreted the data.

Data analysis

Associations among the 50 lines were revealed by cluster analysis using Kimura and Jukes and Cantor algorithms with Neighbour joining analysis. Neighbour joining analysis was also applied to the RFLP and morphological data.

Results

Phenotypic analysis

The dendograms were given in chapter three.

Genetic distance analysis

AFLP cluster analysis

The consensus trees were given in chapter four.

If the AFLP dendograms and the morphological dendograms are compared with each other, there were hardly any similarities between these two techniques. The only similarities were as follows: WO5 and WO7 clustered together in both trees, but in the morphological tree YS1 was also included in the cluster. Cultivars WO11, WO2 and WK2 also clustered together in both trees, but in the AFLP tree WO1 was included in the cluster. In the morphological tree WO1 was replaced by WI1. Cultivars WO2 and WK1, and YO6 and YO10 also clustered together in both trees. The rest of the groups clustered together very differently.

RFLP cluster analysis

If the RFLP dendograms (Fig. 5.1 and 5.2) and the AFLP dendograms are compared with each other, the lines in the RFLP clusters were grouped similarly to that of the AFLP dendograms. It was difficult to compare these two techniques with each other since there were only seven lines included in the RFLP analysis in comparison with the 50 lines included in the AFLP analysis. The clusters could not really be compared with each other, since the lines included in the RFLP analysis were representatives from different clusters. In the RFLP analysis both the UPGMA and Neighbour joining analysis results were similar. It was significant that all the Lancaster lines (YL5, YL4, YL1 and YL2) grouped together. Only YL4 and YL2 exchanged places in the two different trees. There were no correlations between the RFLP results and the morphological characteristics.

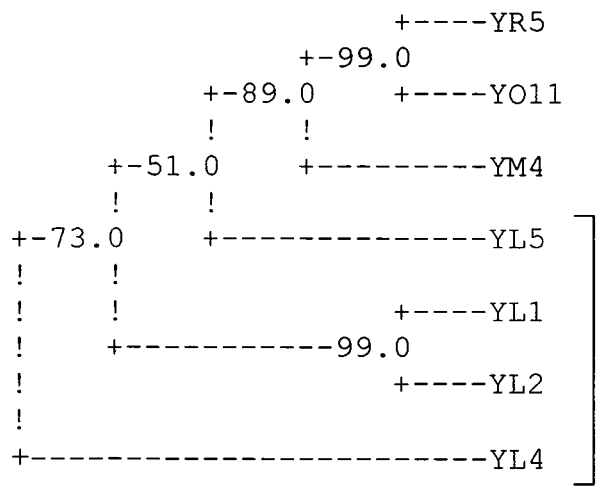


Fig. 5.1 Dendogram generated by Neighbour joining cluster analysis of pairwise distance data for the different maize inbred lines.

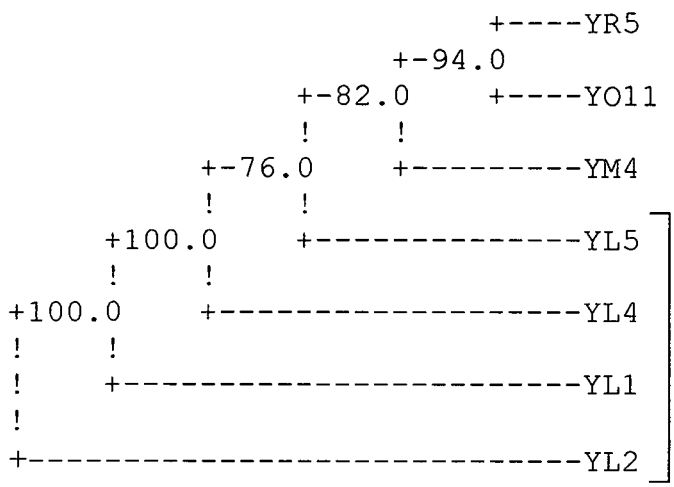


Fig. 5.2 Dendogram generated by UPGMA cluster analysis of pairwise distance data for the different maize inbred lines.

Discussion and conclusion

Comparison of different methods of grouping

Different cluster groups were formed using morphological, RFLP and AFLP analysis. Clusters drawn using morphological characteristics were different to those drawn with AFLP and RFLP data. Morphological traits had too much of an overlap to be able to correlate morphology with the genotype.

Although genotype is reflected in the morphology it is unfortunately inconsistent, because environment has an effect on how the morphology is expressed. From the comparison of methods it was clear that morphological characterisation is subjective, because the characters were measured visually. The variation within a certain character was also fairly small, thus making it difficult to cluster the inbred lines according to these parameters. The breeder assigned groupings based on non-selective characteristics and then tried to scientifically correlate the groupings. AFLP's and RFLP's are scientific methods and are not influenced by the environment. They are reliable and difficult to manipulate.

The development of inbred lines by self-pollination and then evaluation for hybrid performance is the basic procedure in most applied maize breeding programmes. Most breeders develop inbred lines with a certain idiotype in mind, but many correlation studies have shown that the values for agronomic traits of the parental lines have minimal predictive value for hybrid performance, particularly grain yield (Russell, 1974).

Controversial results have been reported in many crops with the use of morphological characterisation alone. Efisue (1993) has found that the combination of morphological and isozyme techniques has offered a better approach than morphological descriptors alone. Lübberstedt *et al.* (1998) concluded that RFLP analysis may serve as a major source of information for separation of closely related accessions, especially when integrated with phenotypic measures.

Some pairs of genotypes in the RFLP and AFLP consensus trees were observed together in both analyses. This suggests that the results of both analyses do have some correlation, but with the amount of RFLP data available it is difficult to really draw any conclusion. The similarity obtained in both dendograms produced by different analytical methods proves the clustering to be acceptable. This positive relationship between these different methods shows that AFLP's and RFLP's are both good techniques to cluster maize germplasm (Lübberstedt *et al.*, 1998; Gerdes and Tracy, 1994).

For the morphological analysis we used data collected in the 1998/1999 season on 25 characters to compile the dendograms. With the AFLP technique, data were collected over a period of three months. One hundred and seven AFLP fragment loci were used in our analysis to construct the dendogram. Using Neighbour joining analysis, specific distances between lines were obtained. It was possible to prove that by using the AFLP technique a large set of informative characters can be obtained in less time than with a morphological analysis. By using the AFLP technique, measures of genetic distance among the accessions were established, thus providing an identity for each line.

A major obstacle to increased use of molecular techniques in systematics is cost. Morphological data can be collected with minimal expenditures on supplies and equipment. The greatest barrier for morphological systematises probably is the initial set-up of the laboratory. Costs vary by specific discipline, but most molecular laboratories require tens of hundreds of thousands of dollars to establish and maintain. Because systematics is a relatively poorly funded subdiscipline of biology, these costs can be prohibitive. However, the value and need for molecular data in systematics is recognised, in spite of the expense (Hillis, 1987). Therefore, although morphological characterisation is influenced by the environment and is time consuming in general, among other disadvantages in relation to AFLP's and RFLP's, it can still be an important and practical means of making progress in germplasm evaluation for farmers and breeders (Lübberstedt *et al.*, 1998).

Application of these results in maize breeding programmes

In breeding programmes, parental lines should have a high degree of heterozygosity, as well as optimum genetic distance (Varghese *et al.*, 1997). Measurement of genetic distance facilitates identification of diverse parents in order to maximise the expression of heterosis (Smith *et al.*, 1990). A desired and high magnitude of heterosis is not always directly related to pronounced parental divergence (Arunachalam *et al.*, 1984) and intermediate divergent classes may also have a high probability of producing heterotic hybrids (Thakur and Zarger, 1989). When this procedure is applied for desired characters, it should result in better exploitation of heterosis in the F_1 hybids. Further investigations between parents and hybrids at

molecular level in maize should reveal useful information on the degree of heterosis and genetic distance between parents and its correlation with the performance of the progeny.

The improvements in marker techniques are likely to accelerate the process of finding markers associated with traits of interest. An important target of marker development work has been to provide tools for more efficient and sophisticated breeding. The close interaction between breeders and biotechnologists will be crucial to the effective implementation of molecular markers. However, it is equally clear that molecular markers offer a powerful set of tools for analysing the inheritance of a wide range of traits (Langridge and Chalmers, 1998).

In conclusion, to show congruency between morphological and molecular data is very difficult. The reason is that although molecular data are supposed to reflect the overall genotype of an organism, and the morphological data are supposed to be an expression of the genotype, the interaction of the environment on genotype is subjective. Environment affects the expression of the genotype. Therefore, morphology can be a valuable asset, but must not be seen as an independent entity to reflect the relatedness between lines or cultivars.

CHAPTER 6

CONCLUSION

In conclusion, using AFLP fingerprinting for the identification and genetic distance analysis of a collection of maize inbred lines, we found that there were no significant differences between the groupings according to the genetic distances. The reason was that there was no hierarchical grouping in the group of maize inbred lines. The group of inbred lines was not developed in a natural selection system, and the relationship between the inbred lines was close. The clusters indicated the absence of any relation between geographic diversity and genetic diversity. This therefore, revealed a relatively low degree of genetic variability within the species. A possible reason can be that the inbred lines have been used all over the world to improve existing hybrids. This interbreeding has caused the variation in the group of material, in terms of place of origin, to be much less.

The results from this study will however, contribute to a large extent to the decisions that the breeder will make in the breeding programme. Since the inbred lines are grouped into clusters and the genetic distances between the lines are known, it is possible to make crosses between lines that will yield higher combining ability. The reason is that the shorter the genetic distance between two lines, the closer related the lines are. The closer related two lines are, that are crossed with each other, the lower the combining ability. The breeder will thus cross lines that have a bigger genetic distance between them.

The implication of these results to the breeder and the breeding programme are significant. Knowing how closely related cultivars are to each other, is a good indication of how good the combining ability would be if two lines were to be crossed. The closer related the lines are, the poorer the combining ability. The results of this study will thus enable breeders to make better decisions about which lines to cross with each other. This in return could make the whole breeding

programme more labour effective, time effective, and cost effective. The results of this study thus indicate that AFLP analysis can be used for the identification and genetic analysis of maize inbred lines.

When looking at the morphological characterisation, the statistical analysis will allow the breeder to make objective choices for the selection of parental material to produce new hybrids. The dendograms in this study showed significant grouping, but more objective methods are needed to confirm the results since environment does influence the expression of the phenotype. These results will however, contribute to the breeding programme, since it will make the possibility of higher combining ability in crosses, as well as improved populations, higher.

Therefore, to show congruency between morphological and molecular data is very difficult. The reason being that although molecular data are supposed to reflect the overall genotype of an organism, and the morphological data are supposed to be an expression of the genotype, the interaction of the environment with genotype is significant, which affects the expression of the genotype. Therefore, morphology alone can not be used to reflect the relatedness between cultivars, scientifically based molecular marker systems must follow.

Morphological and molecular systematic techniques each have distinct advantages for phylogenetic reconstruction. Morphological techniques are applicable to an enormous range of plant material, and a large portion of organisms will continue to be studied primarily from morphological information. On the other hand, the potential molecular data set is incredibly extensive and, when fully utilised, should provide a detailed record of the history of life. Studies that combine the two approaches can thereby maximise both information content and usefulness (Hillis, 1987). In this study it was obvious that such combinations of molecular and morphological studies provide a truly comprehensive view of the genetic variation within a group of cultivars.

CHAPTER 7

SUMMARY

AFLP (Amplified Fragment Length Polymorphism) analysis has found widespread use in DNA fingerprinting because of its relative simplicity, low cost and high marker output. AFLP analysis also does not have the limitations associated with morphological characterisation and the use of other molecular markers such as RFLP's (Restriction Fragment Length Polymorphisms) and RAPD's (Random Amplified Polymorphic DNA markers). Furthermore, the AFLP technique has found widespread use in the identification of cultivars, varieties and breeding lines. The accurate identification of breeding material is extremely important, not only for the protection of breeder's rights, but also to accelerate plant-breeding programmes. The aim of this study was to use AFLP fingerprinting for the identification and genetic distance analysis of a collection of maize inbred lines from the Advanta Africa genebanks.

Fifty maize inbred lines from Africa (local), Argentina, Zimbabwe (CIMMYT), Garst (USA) and Thailand were fingerprinted using AFLP's. A total of 107 scored AFLP loci were used to calculate pairwise genetic distances. This revealed an average genetic distance of 4.38% between all lines studied. The trend of the clustering pattern was more variation in the major clusters with less variation between lines in subgroups and even less variation within the minor groups. It was also further illustrated that the local and foreign inbred lines were not grouped separately. Neighbour joining cluster analysis of the genetic distance data yielded a dendogram that indicated the absence of relation between geographic diversity and genetic diversity.

The resulting knowledge of genetic distance and identification of maize inbred lines in this study will contribute towards maize breeding programmes in Advanta Africa Seeds. It permits an organisation of germplasm resources and identification of parents for crossing blocks. This will enable the breeder to make more

scientific based choices, where both additive and non-additive sources of genetic variation contribute to the gain. Our results have shown that AFLP technology is a rapid, informative and precise technique for identification of maize inbred lines in this study.

Key terms: Amplified Fragment Length polymorphism (AFLP), AFLP markers, DNA fingerprinting, genetic distance, genetic variation, maize inbred lines, morphological characterisation, genetic maps.

OPSOMMING

AFLP (Amplified Fragment Length Polymorphism) analise het 'n wye toepassing gevind in DNA bandpatroon identifikasie a.g.v. die relatiewe eenvoudigheid daarvan, lae kostes en 'n hoë merker uitset. AFLP analise het ook nie die beperkings geassosieer met morfologiese karakterisering, en die gebruik van sommige molekulêre merkers soos RFLP's (Restriction Fragment Length Polymorphisms) en RAPD's (Random Amplified Polymorphic DNA markers) nie. Verder het die AFLP tegniek 'n wye toepassing gevind in identifikasie van kultivars, variëteite en teellyne. Die akkurate identifikasie van teelmateriaal is geweldig belangrik, nie net vir die beskerming van plantetelersregte nie, maar ook om teelprogramme te versnel. Die doel van hierdie studie was om die AFLP tegniek te gebruik vir die identifikasie en genetiese afstandsbepaling van 'n versameling van ingeteelde mielielyne wat teenwoordig is in Advanta Africa Seeds geenbanke.

Vyftig ingeteelde lyne van Afrika (lokaal), Argentinië, Zimbabwe (CIMMYT), Garst (Amerika) en Thailand se bandpatrone is bepaal met AFLP analise. 'n Totaal van 107 vasgestelde AFLP loci is gebruik om paargewyse genetiese afstande te bereken. Dit het 'n gemiddelde genetiese afstand van 4.38% aangetoon tussen al die ingeteelde lyne wat bestudeer is. Die algemene neiging van die groepering was meer variasie in die hoofgroepe, met minder variasie in die subgroepe en selfs minder variasie in die kleiner groepe. Dit was verder geïllustreer deurdat die lokale en oorsese lyne nie apart gegroepeer was nie. "Neighbour joining" analise van die genetiese afstand data het 'n dendogram gegee wat aangedui het dat daar geen verwantskap is tussen geografiese- en genetiese diversiteit nie.

Die kennis wat in hierdie studie gegenereer is van genetiese afstande en identifikasie van ingeteelde mielielyne sal bydra tot verbetering van mielie teelprogramme in Advanta Africa Seeds. Dit verskaf die kennis om kiemplasma bronne te organiseer, en om geskikte ouers te kies vir kruisingsblokke. Dit sal die teler in staat stel om meer wetenskaplik gebaseerde besluite te neem, waar beide additiewe en nie-additiewe bronne van genetiese variasie bydra tot gewas verbetering. Hierdie studie het aangetoon dat AFLP tegnologie vinnig, informatief en presies is vir die identifikasie van die ingeteelde mielielyne wat bestudeer is.

CHAPTER 8

List of References

AJMONE-MARSAN, P., LIVINI, C., MESSMER, M.M., MELCHINGER A.E. and MOTTO, M., 1992. Cluster analysis of RFLP data from related maize inbred lines of the BSSS and LSC heterotic groups and comparison with pedigree data. *Euphytica* 60(2): 139-148.

AKKAYA, M.S., BHAGWAT, A.A. and CREGAN, P.B., 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132: 1131-1139.

ANDERSON, E. and BROWN, W.L., 1952. Origin of Corn Belt maize and its genetic significance. In: Heterosis. J.W. Gowen (ed.). pp 124-148. Iowa State College Press, Ames, IA.

ARUNACHALAN, V., BANDOYPADHYAY, A., NIGAN, S.N., GIBBONS, R.W., 1984. Heterosis in relation to genetic divergence and specific combining ability in groundnut (*Arachis hypogea* L.). *Euphytica* 33: 33-35.

ARUS, P., TANKSLEY, S.D., ORTON, T.J. and JONES, R.A., 1982. Electrophoretic variation as tool for determining seed purity and for breeding hybrid varieties of *Brassica oleracea. Euphytica* 31: 417-428.

BARRETT, B.A. and KIDWELL, K.K., 1998. AFLP-Based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop Sci.* 38: 1261- 1271.

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

BASSIRI, A., 1976. Barley cultivar identification by use of isozyme electrophoretic patterns. *Can. J. Plant Sci.* 56: 1-6.

BEECHING, J.R., MARMEY, P., HUGLES, M.A. and CHARRIER, A., 1995. Evolution of molecular markers and approaches for determining genetic diversity in cassava germplasm. In: Proceedings of the 2nd Scientific Meeting of Cassava Biotechnology Network, Bogor, Indonesia. W. Roca & A.M. Thro (eds.). pp. 80-89. Working Document 150. CIAT, Cali, Colombia.

BEECHING, J.R., MARMEY, P., GAVALDA, M-C., NOIROT, M., HAYSON, H.R., HUGHES, M.A. and CHARRIER, A., 1993. An assessment of genetic diversity within a collection of cassava *(Manihot esculenta* Crantz) germplasm using molecular markers. *Annuals of Botany* 72(6): 515-520.

BIRD, R.M., GONZALEZ DE LEON, D. and HOISINGTON, D., 1997. Designation of SSR morphs. Maize Genetics Cooperation Newsletter, No. 71(63): 77-79.

BIRKENMEIER, E.H., SCHNEIDER, U. and THURSTON, S.J., 1992. Fingerprinting genomes by use of PCR with primers that encode protein motifs or contain sequences that regulate gene expression. *Mammalian Genome* 3(10): 537-545.

BOHN, M., UTZ, H.F. and MELCHINGER, A.E., 1999. Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. *Crop Sci.* 39: 228-237.

BONNETT, O.T., 1960. The inflorescences of maize. *Crop Sci.* 120(3107): 77-87.

BONNETT, O.T., 1954. The developmental anatomy of the corn plant. Proc. 15th Corn Res. Conf., pp. 40-47. Amer. Seed Trade Assoc.

BOPPENMAIER, J., MELCHINGER, A.E., BRUNKLAUS-JUNG, E., GEIGER, H.H. and HERRMANN, R.G., 1992. Genetic Diversity for RFLPs in European Maize Inbreds: I. Relation to performance of Flint x Dent Crosses for Forage Traits. *Crop Sci.* 32: 895-902.

BOTSTEIN, D., WHITE, R.L., SKOLNICK, M.H. and DAVIES, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

BROWN, A.H.D., GRACE, J.P. and SPEER, S.S., 1987. Designation of a "core" collection of perennial *Glycine*. *Soybean Genet. Newsl.* 14: 59-70.

BURR, B., EVOLA, S.V., BURR, F.A. and BECKMANN, J.S., 1983. The application of restriction fragment length polymorphism to plant breeding. In: Genetic engineering principle and methods. J.K. Setlow & A. Hollaender (eds.). pp. 45-49. Plenum Press, London.

BURSTIN, J., VIENNE D., DE DUBREUIL, P., DAMERVAL, C. and DE VIENNE, D., 1994. Molecular markers and protein quantities as genetic descriptors in maize. I. Genetic diversity among 21 inbred lines. *Theor. Appl. Genet.* 89 (7-8): 943-950.

CAETAO-ANOLLÉS, G.C., BASSAM, B.J. and GRESSHOFF, P.M., 1991. DNA amplification fingerprinting: a strategy for genome analysis. *Plant Mol. Biol. Rep.* 9: 553-557.

CAO, D. AND OARD, J.H., 1997. Pedigree and RAPD-based DNA analysis of commercial U.S. rice cultivars. *Crop Sci.* 37: 1 630-1 635.

CASTLEBERRY, R.M., CRUM, C.W. and KRULL, C.F., 1984. Genetic yield improvement of U.S. maize cultivars under varying fertility and climatic environments. *Crop Sci.* 24: 33-36.

CIAT, 1993. Biotechnology research Unit. Annual Report. Cali. Colombia.

COBB, D.B. and CLARKSON, J.M., 1994. A simple procedure for optimizing the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucl. Acid. Res.* 22(18): 3801-3805.

COWEN, N.M. and FREY, K.J., 1987. Relationship between genealogical distance and breeding behavior in oats (*Avena sativa* L.). *Euphytica* 36: 413-424.

COWNIE, P.J., 1993. The maize giant in a changing Africa and its challenge for cereal science and technology. In: Cereal science and technology impact on a changing Africa. J.R.N. Taylor, P.G. Randall and J.H. Viljoen (eds.). pp.31-52. CSIR, Pretoria, South Africa.

COX, T.S. and MURPHY, J.P., 1990. The effect of parental divergence on F2 heterosis in winter wheat crosses. *Theor. Appl. Genet.* 79: 241-250.

CREGAN, P.B. and BUSCH, R.H., 1978. Heterosis, inbreeding, and line performance in crosses of adapted spring wheat. *Crop Sci.* 18: 247-251.

DELLAPORTA, S.L., WOOD, J. and HICKS, J.B., 1983. A plant molecular DNA minipreparation version: II. *Plant Mol. Biol. Rep.* 1: 19-21.

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

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

DOYLE, P., 1993. Biotechnology in action. Science and Public Affairs London, Winter.

DUDLEY, J.W., SAGHAI MAROOF, M.A., RUFENER, G.K. and MAROOF, M.A.S., 1991. Molecular markers and grouping of parents in maize breeding programs. *Crop Sci.* 31(3): 718-723

DUVICK, D.N., 1977. Genetic rates of gain in hybrid maize yield during the past 40 years. *Maydica* 22: 187-196.

DU PLESSIS, J.G., 1993. Maize yield and grain quality as affected by low input production practices. In: Cereal science and technology impact on a changing Africa. J.R.N. Taylor, P.G. Randall and J.H. Viljoen (eds.). pp.163-172. CSIR, Pretoria South Africa.

ECHEVERRIGARAY, S., DELAMARE, A.P.L., TOSELLO, G., GAL, O., HILLEL, J. and LAVI, U., 1996. Evaluation of the genetic diversity among maize lines (*Zea mays L.*) by DNA fingerprinting analysis. *Ciencia Rural* 26(3): 501-503.

EDWARDS, K., JOHNSTONE, C. and THOMPSON, C., 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucl. Acid. Res.* 19: 1349- 1458.

EFISUE, A.A., 1993. Characterisation of IITA elite clones of cassava, yam and sweet potato by morphological and isozyme technique. B.Sc. (Hons) thesis. University of Ibadan. Ibadan, Nigeria.

ELLIS, R.P., MCNICOL, J.W., BAIRD, E., BOOTH, A., LAWRENCE, P., THOMAS, B. and POWELL,W., 1997. The use of AFLPs to examine genetic relatedness in barley. *Molecular Breeding* 3: 359-369.

EHRLICH, H.A., GELFLAND, D.H. and SNINSKY, J.J., 1991. *Crop Sci.* 252: 1643-1650.

EUBANKS, M.W., 1997. Molecular analysis of crosses between *Tripsacum dactyloides* and *Zea diploperennis (Poaceae). Theor. Appl. Genet.* 94(6-7): 707-712.

FELSENSTEIN, J., 1993. PHYLIP: *Phylogeny Inference Package, version 3.5p.* Computer program distributed by the University of Washington, Washington.

GALLAIS, A., 1979. The concept of varietal ability in plant breeding. *Euphytica* 28: 811-823.

GARDINER, J.M., COE, E.H., MELIA-HANCONCK, S., HOISINGTON, D.A. and CHOA, S., 1993. Development of a core RFLP map in maize using an immortalized F2 population. *Genetics* 134: 917-930.

GELFAND, D.H., 1989. *Taq* DNA polymerase. In: PCR Technology Principles and Applications for DNA amplification. H.A. Erlich (ed). pp 17-22. Stockton Press New York.

GERDES, J.T. and TRACY, W.F., 1994. Diversity of Historically Important Sweet Corn Inbreds as Estimated by RFLPs, Morphology, Isozymes, and Pedigree. *Crop Sci.* 34: 26-33.

GIBCO BRL, 1996. AFLP analysis system I. Technical Bulletin 34.

GOODMAN, M.M. and BROWN, W.L., 1988. Races of Maize . In: Corn and corn improvement. 3rd ed. G.F. Sprague and J.W. Dudley (eds). pp 33-79. Agron. Monogr. 18. ASA, Madison, WI.

GOODMAN, M.M., 1972. Distance analysis in biology. *Syst. Zool.* 21: 174-186.

GOODMAN, M.M. and PATERNIANI, E., 1969. The races of maize: III. Choices of appropriate characters for racial classification. *Econ. Bot.* 23: 265-271.

HALLAUER, A.R. and MIRANDA, J.B., 1988. Quantitative genetics in maize breeding (Second Edition). Iowo State University Press, Ames Iowa

HALLAUER, A.R., RUSSELL, W.A. and LAMKEY, K.R., 1988. Corn breeding. In: Corn and corn improvement. 3rd ed. G.F. Sprague and J.W. Dudley (eds.). pp.463-564. Agron. Monogr. 18. ASA, CSSA, and SSSA, Madison, WI.

HALLAUER, A.R., 1987. Maize. In: Principles of cultivar development. W.R.Ferh (ed). Macmillan Publishing Co., New York.

HELLENTJARIS, T., SLOCUM, M., WRIGHT, S., SCHAEFER, A. and NIENHUIS, J., 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphism. *Theor. Appl. Genet.* 72: 761-769.

HILLE, J., ZABEL, P. and KOORNNEEF, M., 1991. Genetic transformation of tomato and prospects for gene transfer. In: Genetic improvement of tomatoes. G. Kalloo (ed.). Monographs on Theoretical and applied Genetics Vol. 14. pp. 283. Springer- Verlag, Berlin Heidelberg.

HILLIS, D.M., 1987. Molecular versus morphological approaches to systematics. *Ann. Rev. Ecol. Syst.* 18: 23-42.

HOSAKA, K. and HANNEMAN, E., 1994. Random amplified polymorphic DNA markers detected in a segregating hybrid population of *Solanum chacoense x S. phureja. Jpn. J. Genet.* 69: 53-66.

HU, J. and QUIROS, C.F., 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant Cell Rep.* 10: 505-511.

JOHN, M.E., 1992. An efficient method for isolation of RNA and DNA from plants containing polyphenolics. *Nucl. Acid. Res.* 20 (9): 55-59.

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

JONES, M.G.K. and LINDSEY, K., 1990. Plant biotechnology. In: Molecular biology & biotechnology (second edition). J.M. Walker and E.B. Gingold (eds.). pp. 117- 158. Whitstable Litho Ltd, Whitstable, Kent.

JUGENHEIMER, R.W., 1976. Corn improvement, seed production, and uses. John Wiley & Sons, Illinois, USA.

KARP, A., SEBERG, O. AND BUIATTI, M., 1996. Molecular techniques in the assessment of botanical diversity. *Ann. Bot.* 78:143-149.

KEIM, P., SCHUPP, J.M., TRAVIS, S.E., CLAYTON, K., ZHU, T., SHI, L., FERREIRA, A. and WEBB, D., 1997. A High-Density Soybean Genetic Map Based on AFLP Markers. *Crop Sci.* 37: 537-543.

KELLY, J.D., 1995. The use of random amplified polymorphic DNA markers in breeding for major gene resistance to plant pathogens. *Hort Science.* 30(3): 461-465.

KIESSELBACH, T.A., 1980. The Structure and Reproduction of corn. University of Nebraska Press, Nebraska USA.

KOCHERT, G., 1994. RFLP technology. In: DNA-based markers in plants. Ronald L. Phillips & Indra K. Vasil (eds.). pp. 9-35. Kluwer Academic Publishers, Netherlands.

KOEBNER, R.M.D., KIRSCH, F., THORPE, C. and PRINS, R., 1998. AFLP as a source of STS markers in alien introgression. In: Proceedings of the 9th International wheat genetics symposium. A.E. Slinkard (ed), Saskatoon Saskatchewan, Canada. pp.118-122.

KOEBNER, R.M.D., DEVOS, K.M. and GALE, M.D., 1994. Advances in the application of genetic markers in plant breeding. Asian Seed 1994, Chiang Mai, Thailand.

KULESHOV, N.N., 1933. World's diversity of phenotypes of maize. *Journal of the American Society of Agronomy* 25: 688-700.

LAMALAY, J.C., TEJWANI, R., RUFENER, G.K.I., 1990. Isolation and analysis of genomic DNA from single seeds. *Crop Sci.* 30: 1079-1084.

LANGRIDGE, P. and CHALMERS, K., 1998. Techniques for marker development. In: Proceedings of the 9th International wheat genetics symposium. A.E. Slinkard (ed). pp.107-117. Saskatoon Saskatchewan, Canada.

LEE, M., 1995. DNA markers and plant breeding programs. *Adv. Agron.* 55: 265-344.

LEE, M., GODSHALK, E.B., LAMKEY, K.R. and WOODMAN, W.W., 1989. Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. *Crop Sci.* 29(4): 1067-1071. Dep. Agron., Iowa State Univ., Ames, IA 50011, USA.

LIN, D., HUBBES, M. and ZSUFFA, L., 1993. Differentiation of poplar and willow clones using RAPD fingerprints. *Tree Physiology* 14: 1097- 1105.

LIU, X.C. and WU, J.L., 1998. SSR heterogenic patterns of parents for marking and predicting heterosis in rice breeding. *Molecular Breeding* 4: 263-268.

LIU, Z. and FURNIER, G. R., 1993. Comparison of allozyme, RFLP, and RAPD for revealing genetic variation within and between trembling aspen and bigtooth aspen. *Theor. Appl. Genet.* 87: 97- 105.

LOUIE, R., FINDLEY, W.R., KNOKE, J.K. and MCMULLEN, M.D., 1991. Genetic basis of resistance in maize to five maize dwarf mosaic virus strains. *Crop Sci.* 31(1): 14 –18.

LÜBBERSTEDT, T., MELCHINGER, A.E., FÄHR, S., KLEIN, D., DALLY, A. and WESTHOFF, P., 1998. QTL Mapping in Testcrosses of Flint Lines of Maize: III. Comparison across Populations for Forage Traits. *Crop Sci.* 38: 1278-1289.

MACKILL, D.J., ZHANG, Z., REDONA, E.D. and COLOWIT, P.M., 1996. Level of polymorphism and genetic mapping of AFLP markers in rice. *Genome* 39: 969-977.

MARSAN, P.A., EGIDY, G., MONFREDINI, G., DI SILVESTRO, S. and MOTTO, M., 1993. RAPD markers in maize genetic analysis. *Maydica* 38: 259-264.

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

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

MAZUR, B.J. and TINGEY, S.V., 1995. Genetic mapping and introgression of genes of agronomic importance. *Current opinion in Biotechnology* 6: 175-182.

McCLEAN, P.L., ERWING, J., LINCE, M. and GRAFTON, K., 1994. Development of a RAPD map of *Phaseolus vulgaris* L. *Ann. Rpt. Bean. Coop.* 37: 79-80.

MELCHINGER, A.E., 1993. Use of RFLP markers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. *Crop Sci.* 28: 621-628.

MELCHINGER, A.E., LEE, M., LAMKEY, K.R. and WOODMAN, W.L., 1990. Genetic diversity for restriction fragment length polymorphisms: Relation to estimated genetic effects in maize inbreds. *Crop Sci.* 30: 1033-1040.

MESSMER, M.M., MELCHINGER, A.E., BOPPENMAIER, J., BRUNKLAUS-JUNG, E. and HERRMANN, R.G., 1992. Relationship among early European maize inbreds: I. genetic diversity among flint and dent lines revealed by RFLPs. *Crop Sci.* 32: 1301-1309.

MESSMER, M.M., MELCHINGER, A.E., LEE, M., WOODMAN, W.L., LEE, E.A. and LAMKEY, K.R., 1991. Genetic diversity among progenitors and elite lines from the Iowa Stiff Stalk Synthetic (BSSS) maize population. Comparison of allozyme and RFLP data. *Theor. Appl. Genet.* 83: 97-107.

MICHELI, M.R., BOVA, R., PASCALE, E., D'AMBROSIO, E., 1994. Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucl. Acid. Res.* 22: 1921-1922.

MICHELMORE, R.W., KESSELI, R.V., FRANCIS, D.M., PARAN, I., FORTIN, M.G. and YANG, C.H., 1992. Strategies for cloning plant disease resistance genes. Molecular plant pathology 2. A practical approach. Oxford University Press, Oxford.

MICHELMORE, R.W., PARAN, I. and KESSELI, R.V., 1991a. Proceedings of the National Academy of Sciences of the USA. 88: 9828-9832.

MICHELMORE, R.W., PARAN, I. and KESSELI, R.V., 1991b. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. USA* 88: 9 828-9 832.

MILLAN, T.T. and CUBERO, A.M., 1995. Varietal identification in *Rosa* by isozymes and RAPD markers. In: Anonymous. 2nd Int Rosa Symp. ISHS/INRA Antibes. France (in press).

MOCK, J.J. AND FREY, K.J., 1970. Leaf number- An index of meioses in corn. *Crop Sci.* 10(6): 659-661.

MONEY, T., READER, S., QU, L.J., DUNFORD, R.P. and MOORE, G., 1996. AFLP-based mRNA fingerprinting. *Nucl. Acid. Res.* 24(13): 2616-2617.

MORELL, M.K., PEAKALL, R., APPELS, R., PRESTON, L.R. and LLOYD, H.L., 1995. DNA profiling techniques for plant variety identification. *Aus. J. Exp. Agric.* 35: 807-819.

MUEHLBAUER, G.J., SPECHT, J.E., THOMAS-COMPTON, M.A., STASWICK, P.E. and BERNARD, R.L., 1988. Near-isogenic lines- a potential resource in the integration of conventional and molecular marker linkage maps. *Crop Sci.* 28: 279-735.

MULLIS, K.B. and FALOONA, F.A., 1987. *Methods Enzymol.* 155: 335-350.

MURRAY, M.G., MA, Y., ROMERO-SEVERSON, J., WEST, D.P. and CRAMER, J.H., 1988. Restriction fragment length polymorphisms: What are they and how can breeders use them. In: Proceedings of the 43rd Annual Corn and Sorghum Research Conference. D. Wilkinson (ed.). pp. 72-87. ASTA, Washington, DC.

MURRAY, H.G. and THOMPSON, W.F., 1980. Rapid isolation of higher molecular weight DNA. *Nucl. Acid. Res.* 8: 4321.

MYBURG, A.A., BOTHA, A-M., WINGFIELD, B.D., 1996. Characterisation of RAPD markers linked to Russian wheat aphid resistance. *Plant Physiology* 111: 89.

NEI, M. and LI, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.* USA 76: 5269-5273.

NIENHUIS, J., TIVANG, J. and SKOCK, P., 1995. Genetic relationships among cultivars and landraces of lima bean (*P. lunatus* L.) as measured by RAPD markers. *J. Amer. Soc. Hort. Sci.* 120: 300- 306.

NICHOLSON, P., REZANOOR, H.N. and SU, H., 1993. Use of random amplified polymorphic DNA (RAPD) analysis and genetic fingerprinting to differentiate isolates of race O, C and T of *Bipolaris maydis*. Journal of Phytopathology 139(3): 261-267.

NYBOM, H., 1994. DNA fingerprinting-A useful tool in fruit breeding. *Euphytica* 77: 59-64.

NYBOM, HY., ROGSTAD, S.H. and SCHAAL, B.A., 1990. Genetic variation detected by use of the M13 'DNA fingerprint' probe in *Malus, Prunus* and *Rubus* (Rosaceae). *Theor. Appl. Genet.* 79: 153-156.

O' TOOLE, J.C., 1989. Breeding for drought resistance in cereals: emerging new technologies. In: Drought resistance in cereals. F.W.G. Baker (ed). pp.81-94. CAB International; Wallingford; UK.

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

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

PAUL, S., WACHIRA, F.N., POWELL, W. and WAUGH, R., 1996. Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis*(L.) O. Kuntze) revealed by AFLP markers. *Theor. Appl. Genet.* 94: 255-263.

PERKIN-ELMER, 1996. AFLP Plant Mapping Kit- the PCR marker of choice for plant mapping.

PHELPS, T.L., HALL, A.E. and BUCKNER, B., 1996. Microsatellite repeat variation within the yl gene of maize and teosinte. *Journal of Heredity* 87(5): 396-399.

POWELL, W., THOMAS, W.T.B., BAIRD, E., LAWRENCE, P., BOOTH, A., HARROWER, B., McNICOL, J.W. and WAUGH, R., 1997. Analysis of quantitative traits in barley by the use of amplified fragment length polymorphism. *Heredity*. In press.

POWELL, W.M., MORGANTE, C.A., HANAFREY, M., VOGEL, J., TINGEY, S. and RAFALSKI, A., 1996. Comparison of RFLP, RAPD, AFLP, and SSR (microsatelite) markers for germplasm analysis. *Molecular Breeding* 2: 225-228.

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

ROLFS, A., SCHULLER, I., FINCKH, U. and WEBER-ROLFS, I., 1992. PCR: Clinical Diagnostics and Research. Springer- Verlag, New York.

ROUPPE VAN DER VOORT, J.N.A.M., VAN ECK, H.J., DRAAISTRA, J., VAN ZANDVOORT, P.M., JACOBSON, E. and BAKKER, J., 1998. An online catalogue of AFLP markers covering the potato genome. *Molecular Breeding* 4: 73-77.

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

RUANO, G., BRASH, D.E. and KIDD, K.K., 1991. The first few cycles. Amplifications: A Forum for PCR users 4: 5-7.

RUSSELL, W.A., 1974. Comparative performance for maize hybrids representing different eras of maize breeding. In: Proc. Annu. Corn and Sorghum Ind. Res. Conf. 29th, Chicago, IL., 10-12 Dec. Am. Seed Trade Assoc. pp.81-101. Washington, DC.

SAGAI-MAROOF, M.A., BIYASHEV, R.M., YANG, G.P., ZHANG, Q. and ALLARD, R.W., 1994. Extraordinarily polymorphic microsatellite DNA in barley: Species diversity, chromosomal locations, and population dynamics. *Proc. Natl. Acad. Sci.* 91: 5 466-5 470.

SAIKI, R.K., GELFLAND, D.H. STOFFEL, S., SCHART, S.J., HIGUCHI, R., HORN, G.T., MULLIS, K.B. and EHRLICH, H.A., 1988. *Science* 239: 487-491.

SALHUANA, W., POLLAK, L.M., FERRER, M., PARATORI, O. AND VIVO, G., 1998. Breeding potential of maize accessions fromn Argentina, Chile, USA and Uruguay. *Crop Sci.* 38: 866-872.

SAMEC, P. and NASINEC, V., 1995. Detection of DNA polymorphism among pea cultivars using RAPD technique. *Biologia Plantarum* 37(3): 321-327.

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

SIMONS, G., VAN DER LEE, T., DIERGAARDE, P., VAN DAELEN, R., GROENENDIJK, J., FRIJTERS, A., BÜSCHGES, R., HOLLRICHER, K., TÖPSCH, S., SCHULZE-LEFERT, P., SALAMINI, F., ZABEAU, M. and VOS, P., 1997. AFLP-based fine mapping of the Mlo gene to a 30-kb DNA segment of the barley genome. *Genomics* 44: 598-604.

SMITH, J.S.C. and SMITH, O.S., 1991. Restriction fragment length polymorphisms can differentiate among U.S. maize hybrids. *Crop Sci.* 31: 893-899.

SMITH, O.S., SMITH, J.S.C., BOWE, S.L., TEMBORG, R.A., WALL, S.J., 1990. Similarities among a group of elite maize inbreds as measured by pedigree F1 grain yield, grain yield, heterosis and RFLP. *Theor. Appl. Genet.* 80: 833-840.

SMITH, J.S.C. and SMITH, O.S., 1989. The description and assessment of distances between inbred lines of maize: I. The use of morphological traits as descriptors. *Maydica* 34: 141-150.

SNEATH, P.H.A. and SOKAL, R.R., 1973. Numerical taxonomy. The principles and practice of numerical classification. W.H. Freeman and Co., San Francisco.

SOBRAL, B.W.S. and HONEYCUTT, R.J., 1994a. The polymerase chain reaction. In: Genetics, plants, and the polymerase chain reaction. K.B Mullis, F., Freré, Gibbs, R.A. (eds). Birkhauser, Boston, pp. 304-319.

SOBRAL, B.W.S. and HONEYCUTT, R.J., 1994b. High output genetic mapping in polyploid using PCR-generated markers. *Theor. Appl. Genet.* 86: 105-112.

SOLLER, M. and BECKMANN, J.S., 1983. Genetic polymorphism in varietal identification and genetic improvement. *Theor. Appl. Genet.* 67: 25-33.

SOUTHERN, E.M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.

SPRINGER, P.S., EDWARDS, K.J. and BENNETZEN, J.L., 1994. DNA class organization on maize Adh1 yeast artificial chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* 91(3): 863-867.

STAM, P. and ZEVEN, A.C., 1981. The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. *Euphytica* 30: 227-238.

STEGEMANN, H., 1984. Retrospect on 25 years of cultivar identification by protein patterns and prospects for the future. *Seed Sci. & Technol.* 15: 20-31.

STUBER, C.W., BROWN A.H.D., CLEGG M.T., KAHLER A.L. and WEIR B.S., 1990. Molecular markers in the manipulation of quantitative characters. Sinauer Associates Inc.; Sunderland, Massachusetts; USA.

SWOFFORD, D.L., 1991. PAUP: *Phylogenetic analysis using parsimony, version 3.0s.* Computer program distributed by the Illinois Natural History Survey, Champaign, IL.

TAI, T.H. and TANKSLEY, S.D., 1990. A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant. Mol. Biol. Rep.* 9: 297-303.

TANKSLEY, S.D., YOUNG, N.D., PATERSON, A.H. and BONIERBALE, M.W., 1989. *Bio/Technology* 7: 257-264.

TANKSLEY, S.D., 1983. Molecular markers in plant breeding. *Plant Mol. Biol. Rep.* 1: 3-8.

THAKUR, H.L. and ZARGER, M.A., 1989. Heterosis in relation to genetic divergence and specific combining ability in indian mustard *Brassica juncea* L. Czern. *Indian J. Genet. Plant Breed.* 49:223-226.

THOMAS, C.M., VOS, P., ZABEAU, M., JONES, D.A., NORCOTT, K.A., CHADWICK, B.P. and JONES, J.D.G., 1995. Identification of amplified fragment length polymorphism (AFLP) markers tightly linked to the tomato Cl 9 gene for resistance to *Cladosportum fulvum. Plant* 8: 785-794.

TING, Y.C., 1994. Molecular markers of anther culture-derived plants. : Maize-Genetics-Cooperation-Newsletter, No.68, 19-20. : Boston College, Chestnut Hill, Boston, MA 02167, USA.

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

TRAVIS, S.E., MASCHINSKI, J. and KEIM, P., 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax, a* critically endangered plant, using AFLP markers. *Mol. Ecol* 5: 735-745.

VAN BEUNINGEN, L.T. and BUSCH, R.H., 1997. Genetic diversity among North American Spring Wheat Cultivars: III. Cluster analysis based on quantitative morphological traits. *Crop Sci.* 37: 981-988.

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

VAILLANCOURT, L.J. and HANAU, R.M., 1992. Genetic and morphological comparisons of Glomerella (Colletotrichum) isolates from maize and from sorghum. *Experimental Mycology* 16(3): 219-229.

VARGHESE, Y.A., KNAAK, K., SETHURAJ, M.R., ECKE, J., 1997. Evaluation of random amplified polymorphic DNA (RAPD) markers in *Hevea braziliensis. Plant Breeding* 116: 47-52.

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

WALTON, M., and HELENTJARIS, T., 1987. Application of restriction fragment length polymorphism (RFLP) technology to maize breeding. In: Proc. 42nd Ann. Corn and Sorghum Res. Conf. American Seed Trade Association. ASTA. D. Wilkinson (ed.). pp. 48-75. Washington, DC.

WANG, Z., WANG, B. and ZENG, M., 1996. RAPD analysis of mtDNAs from multiplasmic CMS lines. Maize-Genetics-Cooperation-Newsletter, No. 70, 12. Academia Sinica, Beijing, China.

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

WAUGH, R., and POWELL, W., 1992. Using RAPD markers for crop improvement. *Tibtech* 10: 186-192.

WELLHAUSEN, E.J., ROBERTS, L.M., HERNANDEZ, E. and MANGELSDORF, P.C., 1952. Races of maize in Mexico. Bussey Inst. Harvard Univ. Press, Cambridge.

WELSCH, J., HONEYCUTT, R.J., MCCLELLAND, M. and SOBRAL, B.W.S., 1991. Parentage determination in maize hybrids using the arbitrary primed polymerase chain reaction (AP-PCR). *Theor. Appl.Genet.* 82: 473-476.

WILLIAMS, J.G., KUBELIK, A.R., LIVAK, K.J., RAFALSKI, J.A. and TINGEY, S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acid. Res.* 18: 6 531- 6 535.

WINTER, P. and KAHL, G., 1995. Molecular marker technologies for plant improvement. *World J. Microb. & Biotech.* 11: 438-448.

WHYTHE, I.V. and GEVERS, H.O., 1988. Diallel analysis of resistance of eight maize inbred lines to *Sphacelotheca reiliana. Phytopathology* 78(1): 65-68.

WU, K.S. and TANKSLEY, S.D., 1993. Abundance, polymorphism and genetic mapping of microsatelite in rice. *Molec. Gen. Genet.* 241: 225-235.

YAMAMOTO, T., NISHIKAWA, A. and OEDA, K., 1994. DNA polymorphisms in *Oryza sativa* L. amplified by arbitrary primer PCR. *Euphytica* 78: 143-148.

YOUNG, N. D., 1994. Constructing a plant genetic linkage map with DNA markers. Ronald L. Phillips & Indra K. Vasil (eds). pp. 39- 58. Kluwer Academic Publishers. Netherlands.

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

YU, K. and PAULS, K.P., 1994. Optimization of DNA-extraction and PCR procedures for random amplified polymorphic DNA (RAPD) analysis in plants. In: PCR Technology: Current innovations.Griffin, H.G., Griffen, A.M. (eds). CRC Press, Inc.

YU, K. and PAULS, K.P., 1992. Optimization of the PCR program for RAPD analysis. *Nucl. Acids Res.* 20: 2606.

YU, K., VAN DEYNZE, A. and PAULS, K.P., 1993. Random amplified polymorphic DNA (RAPD) analysis. In: Methods in plant molecular biology and biotechnology. pp. 287-301. CRC Press, Inc.

ZABEAU, M. and VOS, P., 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application number 92402629.7. Publication number 0 534 858 A1.

ZANONI, U. and DUDLEY, J.W., 1989. Comparison of different methods of identifying inbreds useful for improving elite maize hybrids. *Crop Sci.* 29: 577-582.

ZEHR, B.E., DUDLEY, J.W., CHOJECKI, J., SAGHAI MAROOF, M.A. and MOWERS, R.P., 1992. Use of RFLP markers to search for alleles in a maize population for improvement of an elite hybrid. *Theor. Appl. Genet.* 83: 903-911.

ZEVEN, A.C., KNOTT, D.R. and JOHNSTON, R., 1983. Phylogenetic distribution and genetic mapping of a (CCC)n microsatellite from rice (*Oryza sativa* L.). *Plant. Mol. Biol.* 21: 607-614.

APPENDICES

Data matrix of AFLP markers scored for genetic distance analysis

Table 9.1 Input data matrix used for Neighbour joining analysis of 107 AFLP markers.

WM1
11111111111111111111110000000001100000000000000000000000000000
00000000000000000000000000000000000000

WM2
00011100010000000000001111111111001111111000000000000000000000
00000000000000000000000000000000000000

WF1
10100010110101000000001000000000001000110111111111111111111000000
00000000000000000000000000000000000000

WF2
10100110000000100000001010010010010010110010001111100001100011111
11000000000000000000000000000000000000

WI1
00000000000100101000001000001000010111000000100000000100100100000
00111111000000000000000000000000000000

WI2
10110001000100011100000000100100000011000010000001100000110010010
00010000111100000000000000100000000000

WP1
01000100011011000010000110000000011100010001000001100100100000101
00101000100111110000000000000000000010

WP3
11100010000100000000001011000000000101001100000000000000000000100
00010010011000000100001001011100000000

WR1
10110000100100000000001001000001000000000000100100100110000101011
00000000001000100110000000000000000000

WR2
00010100100000010000101100000000010010111010000000100000100001100
00000100010001000001111000000000000000

WK1
0010000100000000000000010010000000000000001111000010110000110100000
00010000001000000011000111110000000000000000

WK2
0110010000000000000000001000000000000110101100011000000000001000010
0000000011100001000000000000110000100000000

WO1
0000100100001010100010110000000000000000000101000100000000000001000001
1000000000000101100000000000001111011111100

WO2
0011100010010000000000010000000000000000000010000100010000000101001000
0100000000000100000000100011100100001111101

WO3
0011100010010000000000010000000000000000000010000100010000000101001000
0100000000000100000000100011100100001111101

WO4
1010001000000101000000000100000010000000110001010000100010000000000000
000110000000000000110100000000000000000100000

WO5
0100000000010100000000010100010000000101001010010000000001000010000010
000000010000000000101000000000100000101 0000

WO6
1110100000000010000000001000001001000000000100100000000000100000001000
000010110000000100010010000000100110000000 0

WO7
0000000000010100100000100000000010000000100000000010010100011010 01
0000000000000101000101010000000000000000100

WO8
00100001000101000010001001000000010000000100000000110100000001001000
000000000000011000010000100010100000000010

WO9
0000000100010110001000100010100000000000000110000000000000000000000000
0000000001000000000010000100000000000000000

WO11
0000000000101000000000000000000100000000000000001000001000000000000100000
000001100000110000101100000000000100000000

YM1
0000000010010100000000000000111000000000100100000000010011100111000 1000
00000000001010000000000000000010100000000010

YM2
1110001001010000100000000000100110000000001000000000000000000001000
0100000000100100001000111000000100101100

YM3
0000000000010110000011100010100000010100101000001001010000000000000
0010110000010010010100001010000010000 1000

YM4
1000000000000000000000100001000000000000000000100010100110000000010
0000000000000100010000000000000110000000

YF2
1010000000000100010000000000100000000101001001000100011101000000 1
0010000101000000000000000010011100000000 0

YF3
0011010101000001001000001000010010010000000100000100100001000000000
0001010010000000100000000000000010000000

YF4
0011000000000000000000000000000000000000000000000000000000000000000
0001000000000000000000000000000100000001

WP2
0000000000000100000010100000000000101000010001000000000000000001001
0001011001101001000000000100000000000000

WO10
0010101000010111011010010000000010001000000110000001000000000000000
0010000110000000001010010010000 1000000000

YF1
0010010000010100100100000000101000001000000000000101000000000010 0
0010000001000000000000100000000000000001

YI2
0001010100000100000000000000001000010000000000000000000000000000000
0000110000000000000000000000000000100000000

YR5
0000000010110100000100100000000011000000000100110110011000000000011
0010010110100000000000000100000 1000000000

YR6
0100000000000000000000000000000000000000000000000000000000000010000
0000000000000000000000000000000000000000

YL1
0000001000010111000000000000000010100001010110000000000001010000000
0000001010000000000010000000000000000000

YL2
0010000000000000000000000000000000000000000000000000000000000000000
0000000000000000001000000000001000000000

YL4
1000010001100110010100000000100110010010100000100101011000000000001
0000000000010000000100000000001000000000

YL5
0000000000000000000001000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000

YS1
0001000000010010000011000000100000010000000000100000010000000010101
0100000000000000000000000000010000010000000

YS3
0111000000000100100000100010000000010010101010000010110000100000000
0000000000000000000001000000000000000000

YO1
0000000000000000000001000000000000000000001000000000000000000000000
0001001000000001010001100000010100000000

YO2
0000000000000000000001000000000000000000001000000000000000000000000
0000000000000000000000000000000000000000

YO4
0100000000000000000001000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000

YO6
0000000000000000000000000000000000000000000000000000000000000000000
0000000000001000000000000000000000000000

YO9
0000000000000000000001000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000

YO10
0010000000101111011000100010001000000100010000101001000100001000100
0000011111010001000000000000100010000000

YO11
00100000000000000000000000000000000000000000000000000000000000000000
00010000000000000000000000000100000000

YO12
00000000000000000000000000010000000000000000000010000000000000000000
00000000000001000000000000000000000000

YO13
11000010001001110000000000000000010100000000000000000011000000001
00001101000000000000000000000000100000001

Each test string shows the name of the cultivar or line and the presence (1) and absence (0) of AFLP markers.