CHARACTERISATION AND DIALLEL ANALYSIS OF COMMERCIALLY PLANTED COTTON (*GOSSYPIUM HIRSUTUM* L.) GERMPLASM IN TANZANIA

BY

EVERINA PAUL LUKONGE

Submitted in the fulfilment of the requirements for the degree of Philosophiae Doctor, in the Department of Plant Sciences (Plant Breeding), Faculty of Natural and Agricultural Sciences

UNIVERSITY OF THE FREE STATE
BLOEMFONTEIN, SOUTH AFRICA
NOVEMBER 2005

SUPERVISOR: PROF. M.T. LABUSCHAGNE
CO-SUPERVISOR: DR. L. HERSELMAN

DECLARATION

I hereby declare that this dissertation, prepared for the degree Philosophiae Doctor, which was submitted by me to the University of the Free State, is my original work and has not previously in its entirety or in part been submitted to any other University. All sources of materials and financial assistance used for study have been duly acknowledged. I also agree that the University of the Free State has the sole right to the publication of this dissertation.

Signed on _____ November 2005 at the University of Free State, Bloemfontein, South Africa.

Signature: _____

Name: Everina Paul Lukonge

i

DEDICATION

This piece of work is dedicated to my parents Christian Lukonge and Laurentina Christian, my husband Paul Machibya and my son Ebenezal Namala.

TABLE OF CONTENTS

LIST OF TABLES

ABREVIATIONS

ACP	Acyl carrier protein
AFLP	Amplified fragment length polymorphism
AMMI	Additive Main effects and Multiplicative Interaction
ANOVA	Analysis of variance
ARC	Agricultural Research Council
ASV	AMMI stability value
BC	Before Christ
bp	Base pairs
CEC	Cation exchange capacity
^0C	Degree Celcius
cm	Centimetre
cmol	Centimolar
CMS	Cytoplasmic male sterility
CTAB	Cetyltrimethylammonium bromide
CV	Coefficient of variation
Df	Degree of freedom
DR	Desaturation ratio
DNA	Deoxyribonucleic acid
dNTP	$2^{'}$-Deoxynucleoside $5^{'}$- triphosphate
dS/m	deci-Siemens/metre
DUS	Distinctiveness, Uniformity and Stability
EC	Electric conductivity
ECGA	Eastern cotton growing area
EDTA	Ethylenediaminetetraacetate
EM-AMMI	Expectation-maximisation - AMMI
ER	Elongation ratio
FAO	Food and Agricultural Organisation
fmol	Femtomoler
g	Gram
GC	Gas chromatography

G x E	Genotype x environment
GOT	Ginning outturn
GoT	Government of Tanzania
GRAS	Generally recognised as safe
g/tex	Gram per tex
HDL	High density lipoprotein
HVI	High volume instrument
IBPGR	International Board for Plant Genetic Resources
ICAC	International Cotton Advisory Committee
IPCA	Interaction principle component analysis
KAS	Ketoacyl-ACP synthase II
kg/ha	Kilogram per hectare
LDL	Low density lipoprotein
LDR	Linoleic desaturation ratio
l/ha	Litre per hectare
LSD	Least significant difference
m	Metre
m^2	Metre squared
masl	Metre above sea level
µg	Microgram
µl	Microlitre
µm	Micrometre
µM	Micromolar
meq	Milliequivalent
mg	Milligram
min	Minute
ml	Millilitre
mm	Millimetre
mM	Milimolar
MS	Mean squares
MUFA	Monounsaturated fatty acid
N	North

ng	Nanogram
nm	Nanometre
NCSS	Number cruncher statistical system
NSS	National soil service
NTSYS	Numerical taxanomy multivariate analysis system
ODR	Oleic desaturation ratio
OC	Organic carborn
PA	Phosphatidic acid
PAGE	Polyacrylamide gel electrophoresis
PBR	Plant breeders rights
PCA	Principle component analysis
PCR	Polymerase chain reaction
Pi	Cultivar performance measure
PIC	Polymorphic information content
pmol	Picomole
psi	Pound per square inch
PUFA	Polyunsaturated fatty acid
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RCBD	Randomized complete block design
RFLP	Restriction fragment length polymorphism
rpm	Revolution per minute
S	South
SCA	Specific combining ability
SDS	Sodium dodecyl sulphate
SE	Standard error
SED	Standard error deviation
SFA	Saturated fatty acid
SNP	Single nucleotide polymorphism
SS	Sums of squares
ssp	Subspecies
SSR	Simple sequence repeat

SV	Stability variance
TAG	Triacylglycerol
Taq	*Thermus aquaticus*
TCL and SB	Tanzania Cotton Lint and Seed Board
TE	Tris EDTA buffer
TN	Total nitrogen
ton/ha	Tons per hectare
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloric acid
TSP	Triple super phosphate
U	Unit
UFA	Unsaturated fatty acid
UFS	University of the Free State
USA	United States of America
UPGMA	Unweighted pair group method of arithmetic averages
UPOV	Union for the protection of new varieties
UV	Ultraviolet
v/v	Volume per volume
WCGA	Western Cotton Growing area
WCRC	World Cotton Research Conference
Wi	Wricke's ecovalence
w/v	Weight per volume
YS	Yield stability

CHAPTER 1

INTRODUCTION

Upland cotton (*Gossypium hirsutum* L.) is a very important textile fiber currently accounting for 90% of the commercially grown cotton worldwide. In the 1993/94 to 1997/98 seasons, cotton was the second most important oilseed crop in the world averaging one-fourth that of soybean (*Glycine max* L.) (Cherry and Leffler, 1984; Zhang, 2001; Jones and Kersey, 2002). Cotton is cultivated in the tropical and subtropical regions on a wide range of soil types as an annual crop, though it is basically a tropical perennial crop. Cotton is primarily used to produce lint which is the unicellular out-growth of the cottonseed. Cotton fibre is made up of a primary wall and secondary cellulose wall which develops after cell elongation ceased (Prentice, 1972; Poehlman, 1987; Kim and Triplett, 2001).

Cotton is harvested as seedcotton, which is then ginned to separate the seed and lint. The long lint fibres are processed by spinning, to produce yarn that is knitted into fabrics. The short fibres (fuzzy), covering the seeds are known as 'linters'. The first cut linters have a longer fibre length and are used in the production of belts, mattresses and mops. The second cut linters have a much shorter fibre length and are a major source of cellulose for both the chemical and food industry. These linters are used as a cellulose base in products such as high fibre dietary products as well as a viscosity enhancer (thickener) in ice cream, salad dressings and toothpaste. In the chemical industry, second cut linters are used in combination with other compounds to produce cellulose derivatives such as acetate, nitrocellulose and a wide range of other compounds (Gregory *et al.,* 1990; Pillay and Myers, 1999).

Delinted cottonseed can be processed to produce oil, meal and hulls. Cottonseed oil has been in common use since the middle of the nineteenth century and achieved GRAS (Generally Recognised As Safe) status under the United States Federal Food Drug and Cosmetics Act because of its common use prior to 1958 (ANZFA, 2002). Cottonseed oil is used in a variety of products including edible vegetable oils and margarine, soap and plastics. Cottonseed cake, meal flour or hulls derived from it is used in food products and for animal feed as carbohydrate

roughage, but is limited by the presence of natural toxicants in the seeds (gossypol and cyclopropenoid fatty acids) (Pillay and Myers, 1999).

Cultivation of cotton is of great importance for the national economy worldwide due to the increasing demand for cotton products. Cotton lint production increases season after season, for example in 1996/97, 19736 metric tons of lint were produced compared to 18714 metric tons during the 1994/95 season (ICAC, 2001). Meredith *et al*. (1997) stated that cotton yield has greatly increased since 1935 because of improved crop management and breeding. In South Africa, cotton is one of the five major crops produced commercially in the country and makes a significant contribution to the economy (Dippenaar-Schoeman, 1999).

In Tanzania, cotton is of great economical importance as it is the second most important cash crop after coffee, representing 15% of the country's total exports and almost 40% of agricultural exports (Bunyecha and Tamminga, 1995; Baffes, 2002). Following liberalisation of the cotton industry, strong competition from village to market level resulted in the deterioration of cotton quality. Furthermore, mixing of different types of cotton varieties led to poor cotton properties (TCL and SB, 2002). Available varieties have medium yields (1200 kg/ha at research level and 300-500 kg/ha at farmers level), medium ginning percentages (36.8-39.6%) and medium fibre strength (22-25g/tex). Based on improved spinning machines, fibre strength above 28 g/tex is recommended for international cotton fibre markets (Deussen, 1992; Hau, 1997).

Selection criteria used at present in conventional breeding programmes are usually based on phenotypic characteristics. Environmental conditions affect the phenotypic characteristics that are complicated by their polygenic nature (Antoni *et al*., 1991; Rivera *et al*., 1999). Although morphological markers are used to perform the above tasks it is difficult to characterise and it does not show high levels of variation.

In the past, repeated crossing and intensive selection between a few families with desirable traits, led to the narrowing of the cotton gene pool that resulted in low genetic variation between existing accessions of cultivated cotton genotypes. Therefore, development of molecular genetic analysis for diversity studies for the available germplasm is important for cotton improvement (Iqbal *et al*., 2001; Allen and Auld, 2002). Although crosses have been made between varieties

in the past in Tanzania, no diallel crosses were applied to determine compatibility, heterosis, heritability and correlation between characteristics. Stability of varieties has not been tested over different environments. Therefore, genetic diversity studies using molecular markers (amplified fragment length polymorphism or AFLP), morphological markers and oil and fatty acid content are important to determine genetic diversity. In combination with diallel crosses, heterotic groups and performances can be determined and stable varieties identified. Molecular marker knowledge will increase efficiency and effectiveness of marker-assisted breeding and in conservation of plant genetic resources in Tanzania.

Improved cotton varieties are urgently needed to improve the cotton market through cotton yield, high ginning percentages and good cotton quality as these factors affect lint price on the world market. The success of a breeding programme is mainly due to knowledge on the available germplasm especially genetic diversity (Meredith and Bridge, 1984; Pillay and Myers, 1999). The above knowledge is important to a plant breeder.

The objectives of this study were
1. To use the Gas Chromatography technique to study the fatty acid composition in 30 cotton varieties from Tanzania.
2. To use morphological characteristics to study the genetic diversity available in 30 different cotton varieties.
3. To use the AFLP technique to study the genetic diversity in 26 different cotton varieties and to build capacity on molecular marker-assisted breeding.
4. To compare genetic similarities and dendrograms from morphological and molecular markers and determine the relatedness between these varieties from this data.
5. To use seven parents in diallel crosses to study combining ability, heterosis, correlations and heritability of most important characteristics.
6. To study the genotype with environment interaction using 21 diallel F_1 progeny and the parents.

CHAPTER 2

LITERATURE REVIEW

2.1 History, origin and diffusion of cotton

2.1.1 History and origin

Cotton is harvested from almost 32.4 million hectares in more than 40 nations of the temperate and tropic regions of the world (Anonymous, 1981). The crop is grown as far as 47^0 degrees N latitude in the Ukraine and 37^0 N latitude in the USA. In the Southern hemisphere production extends to about 32^0 S latitude (Niles and Feaster, 1984). Cotton grows at an optimum temperature of 30^0C, where 15^0C is the minimum temperature for cottonseed germination and growth (Munro, 1987).

Various theories have been advanced to explain the selective value of lint in the evolution of the species, but there is no convincing evidence that it is of any use to a cotton plant growing wild in its natural habitat (Munro, 1987). However, the primary centres of diversity for the genus are west central and southern Mexico (18 species), northeast Africa and Arabic (14 species) and Australia (17 species) (Brubaker *et al*., 1999). Brown *et al*. (1999) reported that *Gossypium hirsutum* L. and *Gossypium barbadense* L. are natives of Mexico where they were domesticated originally.

In the light of increased knowledge of the distribution and relationships of primitive cottons, Santhanam and Hutchinson (1974) reported that the Asiatic species and races probably differentiated before domestication. Fryxell (1968) reported that cottonseeds can survive floating in seawater for at least a year with undiminished viability and can thus be distributed by ocean currents. Pursegloves (1968) agreed that the most likely explanation was that cottonseeds floated across the Atlantic from Africa to South America. The development of Old World cotton as a major raw material took place in Sind. This was found during excavation in Pakistan that was dated at approximately 3000 BC (Gulati and Turner, 1928). In Peru the New World tetraploid cottonseeds dated back to 2500 BC (Hutchinson, 1959). In Southern Mexico, cotton was dated around 3500 BC (Smith, 1968). Linted cotton species have been used for cotton

fabrics between 4000 and 3000 BC (Munro, 1987). The oldest archaeological remains of *G. hirsutum* are from the Tehuacan Valley of Mexico, 4000 to 5000 years ago.

It is assumed that *G. hirsutum* was probably first domesticated by pre-Columbian people of the Yucatan peninsula (Brubaker *et al.*, 1994). The wild *G. hirsutum* variety is 'Yucatanense', a sprawling perennial shrub with reproductive development controlled by photoperiod flowering under short day conditions. Variety 'Punctatum' arose from 'Yucatanense'. These early-domesticated varieties dispersed to the rest of Mesoamerica, northern South America and the Caribbean basin. Ethno botanical evidence suggested that landrace 'Latifolium' arose from this germplasm. Some accessions classified as 'Latifolium' show photoperiodic flowering while others are photoperiodic independent. In Guatemala, cotton was traditionally intercropped with pepper (*Capsicum spp*.). Cotton plants were removed as soon as first bolls began to open in order to prevent competition with the developing pepper. This practice would have eliminated late maturing genotypes. Selection for early maturity would have reduced seed dormancy and possibly photoperiod dependent flowering. The early maturing Latifolium genotypes diffused into the highlands of southern central Mexico (Brubaker *et al.,* 1999).

Mexican *G. hirsutum* types may have been grown in the Stephens Austin colony in Texas as early as 1821. Numerous introductions were probably made by soldiers returning from the Mexican-American war (1846-1848). These cultivars were subjected to strict selection to create varieties adapted to local conditions in various cotton growing regions of Northern America. Throughout these periods, outcrossing occurred between cultivars (Endrizzi *et al.,* 1985), collectively known as American Upland cotton. The resulting high yielding and adaptable varieties were dispersed to Europe, Asia and Africa. The limited genetic diversity of cultivated upland *G. hirsutum* has been observed by several researchers (Multani and Lyon 1995; Iqbal *et al.,* 1997; Iqbal *et al.,* 2001; Lu and Myers, 2002). A hypothesis to explain this is that genetic bottlenecks occurred upon importation of small quantities of seed from Mexico to America in the 19th century. For example, Burling's cotton in 1806 was smuggled out of Mexico in the stuffing of dolls. More bottlenecks may have occurred during the late stages of development of *G. hirsutum* Latifolium possibly as a result of rigorous selection (Lewis, 1962).

2.1.2 Diffusion of cotton in Africa

Seed multiplication of much longer staple in Egypt started in the early 19th century. The American civil war in 1861-1865 stimulated cotton growing of the American tetraploid species, especially upland varieties, which produced lint of markedly better length and fineness as well as better yields than Old World diploid cotton (Munro, 1987). In 1902, the Lancashire cotton manufacturers joined together to form the British Cotton Growing Association in the colonies. They conducted experiments and established cotton plantations to find out where cotton can grow successfully and its main effort was in Africa. Research stations were established and by 1945-1946 progress reports were published from experiment stations in Australia, South Africa, Zimbabwe, Sudan, Tanzania, Uganda, Malawi, Nigeria and West Indies (Cowley, 1966).

2.2 Evolution and genetics

Cotton is primarily a self-pollinated crop but there is about 1-32% natural outcrossing during field cultivation that depends mainly on location and pollinator availability (Poehlman, 1987; Abdalla *et al*., 2001).

Cotton belongs to the order *Malvales,* family *Malvaceae* and genus *Gossypium*. *Gossypium* includes about 45 diploid (2n=2x=26) species and five allotetraploid (2n=4x=52) species (cultivated and wild) (Brubaker *et al*., 1999). Diploid species comprise genomic groups A, B, C, D, E, F, G and K and allotetraploid species are made up of two subgenomic groups with affinity A and D genomes (Endrizzi *et al*., 1985; Stewart, 1995). There are four cultivated species, two Old World diploid species (*G. arboreum* L. and *G. herbaceum* L.) both A-genome (2n=26) that are native to southeast Asia and Africa and two New World allotetraploid species (*G. barbadense* L. and *G. hirsutum* L.) with the AD genome (2n=4x=52) from Central America and Northern South America (Endrizzi *et al*., 1985; Pillay and Myers, 1999; Iqbal *et al.,* 2001). The entire worldwide cotton production is from *G. barbadense* and *G. hirsutum* though *G. hirsutum* comprises 90-95% of the world cotton production (Iqbal *et al.,* 2001; Altaf Khan *et al*., 2002).

In allotetraploid species, the D-genome has 13 small chromosomes and the A-genome has 13 moderately large chromosomes in the haploid complement of 26. D and A genomes differ in the amounts of moderately repetitive DNA sequences (Geever *et al.,* 1989). Differences in the

quantities of repetitive DNA are thought to permit genome specificity during meiotic chromosome pairing (Mursal and Endrizzi, 1976). During the evolutionary process, diploid species with small chromosomes hybridised with a second diploid species with larger chromosomes. The spontaneous doubling created a 52 chromosome tetraploid species (2n=4x=52) with two groups of genomes A and D (AD) (Simmonds, 1984; Munro, 1987).

Gossypium arboreum (A2 genome) is still grown in Pakistan and India on marginal land for use in non-woven material and is helpful in breeding programmes as a donor of host-plant resistance genes. The A-genome cotton enhances genetic diversity of tetraploid cotton breeding programmes (Stanton *et al*., 1994), especially with the development of techniques for introgressing A-genome germplasm into AD-genome cultivars (Stewart, 1992). Hybrids between *G. hirsutum* and *G. arboreum* have led to the selection of genotypes with earlier maturity and an increased range of fibre traits (Wang *et al*., 1989; Stanton *et al*., 1994).

2.3 The importance of cotton

In developing countries cotton accounts for nearly 3% of the total crop area and is produced for various purposes (Fortucci, 2001).

(a) Contribution to agriculture and economy

In 2000, world cotton production amounted to 19 million tons. Cotton production contributes substantially to the national economy in some of the African developing countries. Even when the share of national income is small, the crop provides significant returns to areas specializing in production (Fortucci, 2001). In South Africa, since 1974, the area under cotton production increased by more than three fold (WCRC, 2003). Currently cotton is one of the five major crops produced commercially in the country. Cotton is Tanzania's largest export crop after coffee (Bunyecha and Tamminga, 1995).

(b) Contribution to agricultural export revenue

Cotton is one of the important commodities traded on the world market. On average, cotton exports accounts for nearly 20% of total agricultural export revenue for African countries. Such large revenues obviously have important multiplier effects on national economies and

household incomes. The relative importance of cotton export revenues has increased, particularly in sub-Saharan Africa (FAO, 2000).

(c) Indirect contributions

In addition to the direct impact of fibre exports, cotton is one of the basic materials for textiles. Cotton production contributes to employment, though it is difficult to obtain numbers of farmers or family members actually employed or involved, particularly in developing countries. Generally, small cotton farms use almost all of the global labour employed in cotton to produce 65% of the world's output on 72% of the planted area. In addition to direct farm employment, cotton production provides additional opportunities for rural employment in cotton ginning, transport and marketing in those countries with textile and clothing manufacturing industries (FAO, 2000). Agrimarket INFO (1998) and Dippenaar-Schoeman (1999) reported that the cotton industry in South Africa is one of the largest employment sectors and it significantly contributes towards social and economic uplifting in the country. In Tanzania labour is the major input and cotton provides employment to 500000 rural households (Baffes, 2002).

On average, households use about 35% of their total cash income obtained from cotton to buy food, 10% for clothing, 15% for production inputs and 40% for many other needs such as medical care, communication and education. Households with a school age child use about 40% of the cash income for the child's education (FAO, 2000).

2.4 Cotton development and advances

2.4.1 General cotton development

In nature, *G. hirsutum* is a perennial shrub that grows to about 1.5 meter in height. As the use of cotton increased, selection took place for more desirable field characteristics. Today modern upland cotton cultivars are high yielding, day length neutral and annual plants. Tall perennial cottons were replaced by the compact and heavy yielding annual crop (Munro, 1987). Unfortunately this was accompanied by reduction in genetic diversity (Anonymous, 1972; Endrizzi *et al*., 1985). Niles and Feaster (1984) stated that trends in cotton breeding formally were towards improvement of plant size, earliness, fibre quality, seed properties, environmental

stress tolerance, boll size, bolls per plant and pest resistance. Thus, cultivated cotton is a perennial plant with an indeterminate growth habit that has been adapted to annual crop culture (Kohel and Benedict, 1987).

Breeding to improve fibre quality traditionally focused on enhancing long fibre or fibre strength for ring yarn manufacturing systems. With the technological evolution of yarn manufacturing from solely ring based spinning to predominantly rotor and air-jet spinning, fibre profiles needs have been revised for these spinning systems. Successful rotor spinning requires high fibre strength for all yarn counts, along with fibre fineness for fine count yarns. Air-jet spinning requires minimum, but uniform fibre length, fibre fineness and to a lesser extent strong fibre. In contrast, ring spinning requires minimum fibre length, fibre strength and to a lesser extent minimum fibre fineness (May, 2002). Breeders have been successful in developing cultivars with stronger fibres that can withstand the forces associated with higher manufactured speed spinning machines (Deussen, 1992).

Cotton improvement has always been directed towards yield and yield components like locules, boll size, number of bolls per plant, seeds per boll, seed size, lint index, seed index and ginning outturn. Therefore, breeders applied different breeding methods for improvement like pedigree breeding (Munro, 1987), bulk population breeding (Allard, 1960), backcross breeding (Sikka and Joshi, 1960) and interspecific and intraspecific breeding for hybrid vigour or heterosis that is found in F_1 crosses within and between species (Hutchinson *et al.,* 1947).

A wide range of crosses has been tested for hybrid vigour and showed increase in yield ranging from 0-100% above the parental mean. The first cytoplasmic male sterility (CMS) line of commercial cotton was introduced by crossing *G. hirsutum* as male parent to *G. harknessii* L. (Meyer, 1975). Work on CMS and restorer genes is being carried out using the technique developed by Weaver and Weaver (Munro, 1987). The primary problem in production of hybrid cottonseeds involves the development of good combiners with dependable disease and pest resistance and the secondary problem is the cost of F_1 seed production (Tang *et al*., 1993b).

Zhang (2001) commented that though primarily *G. hirsutum* is cultivated in the world, morphological and cytological studies of cotton lagged behind due to its large genome and small chromosomes. Recently molecular biology studies have shown significant progress in genetic studies like:

1. Transgenics were produced with cotton as a major crop in which commercialisation of biotechnology initiatives by the private sector provided various forms of plant protection and the genes for resistance to insects, herbicides, disease, drought and cold resistance.

2. DNA marker systems were used in the construction of cotton linkage maps, screening for molecular markers linked to important agronomic traits genes, studying genetic diversity and heterosis mechanism.

3. Cotton regeneration and transformation: research continues to be focused on problems of regeneration and transformation. It is expected that conventional breeding and these new technologies will complement each other and improve the cotton industry.

In South Africa, the Agricultural Research Council (ARC)-Institute for Industrial Crops is responsible for agronomic and quality improvement of cotton. There are programmes responsible for developing new cultivars adaptable to the environment through gene manipulation, to produce cultivars tolerant to *Verticilium* wilt, nematodes and insects based on morphological characteristics such as hairiness, okra leaf, frego bract and red colour. About 1380 germplasm accessions from Central America (exotic), early released germplasm, registered cultivars and local cultivars from South Africa, Zimbabwe and Mozambique are maintained (Van Heerden *et al.*, 1987).

2.4.2 Cotton development and advances in Tanzania

Cotton was introduced to Tanzania around 1904 by German settlers as a plantation crop, but the attempt failed. During the 1920's new efforts focused on smallholder production, first in eastern and later in western Tanzania. Production of cotton on commercial scale started at Ukiriguru. In the Western Cotton Growing areas (WCGA's) (Mwanza, Shinyanga, Mara, Kagera, Kigoma, Tabora and Singida regions), breeding activities started at Ukiriguru Research Institute in 1939. Seed for sowing originally came from Uganda and consisted of mixtures of different US-cotton varieties. Selections from the mixtures cultivated, resulted in the release of the first variety

called *Mwanza local* in the 1940's. The first variety with good jassid resistance was released in 1946 (Lukonge and Ramadhani, 1999). In the following years, other varieties were developed with improved jassid resistance, high yield and high ginning percentage. In the early 1960's the focus moved to breeding varieties with bacterial blight resistance. Another important disease was fusarium wilt in areas surrounding the Lake Zone. In the 1960's, resistant material was released. Resistant varieties cultivated presently are UK77, UK82 and UK91 (Ramadhani and Lukonge, 1999).

In the Eastern Cotton Growing Areas (ECGA's) (Morogoro, Coast, Ruvuma, Arusha, Tanga, Kilimanjaro and Iringa regions), research started at Ilonga Research Institute in 1943. Cotton grown in the area consisted of heterogenous mixtures of varieties from Uganda, which was given the name of Coast Local. Early efforts to improve the genetic crop constitution had the primary objective of selecting material that was high yielding, had good jassid resistance and good lint quality. Further selection and variety testing led to the release of the multiline IL58. Various other commercial varieties were released (IL62, IL74 and IL85). Since 1985, bacterial blight disease in the ECGA's has become more noticeable. Resistant material was improved from crosses of Malawian and Nigerian material (Ramadhani and Lukonge, 1999).

Cotton research in Tanzania comprises of five sections (breeding, entomology, pathology, agronomy and fibre testing), all working together. However, the programme has been involved in exotic variety introductions for crossing purposes. There are about 200 accessions in the programme with important traits for breeding purposes and are mainly from outside the country (Lukonge and Ramadhani, 1999). In Tanzania, cotton is mainly produced in a subsistence agricultural system and is entirely rain fed (Jones and Kapingu, 1982; Baffes, 2002). The WCGA's produce about 90% of the total cotton in the country, while the rest (10%) comes from the ECGA's (TCL and SB, 2001).

2.5 Problems of the cotton sector in Tanzania

The major cotton production constraints in Tanzania include unfavourable weather conditions (mainly drought) in some regions, insect pests (American bollworm, jassids, lygus and aphids),

diseases (fusarium wilt, bacterial blight and verticillium wilt), weeds, competition with food crops, declining soil fertility and unsatisfactory marketing and seed distribution systems (Bunyecha and Tamminga, 1995).

In 1991, Ukiriguru Research station released a new cotton variety, UK91, which was superior (yield and resistant) to both UK77 and UK82. However, achieving higher yields at farmers' fields requires multiplication and release of enough UK91 seed to replace the older varieties so as to avoid mixing with existing varieties. Since the release in 1991, there was no enough seed produced to cover the WCGA's (Shepherd and Farolfi, 1999).

Tanzania's textile industry was started in the early 1970's as part of the government's efforts to industrialise the economy. More than 80% of mill capacity was under state ownership (Government of Tanzania, 1999b). Once government support came to an end, the industry was unable to survive international competition and some textile mills went out of business (Shepherd and Farolfi, 1999).

Infrastructure shortcomings severely impede the development of the cotton sector. Firstly, because most cotton must be transported by rail, the quality of rail services is vital to sectoral performance. Greater efficiency in rail transport will lower costs to growers. Secondly the road network in the Mwanza region, where most cotton is produced, requires considerable upgrading. As with rail transport, road improvements will increase efficiency and reduce costs, thereby leading to higher producer prices (Baffes, 2002).

Declining input, caused by removal of input price subsidies at farmer level, (mainly insecticides and fertilizers) led to poor quality cotton and low yields. Any quality decline due to reduced input use reflects relative prices and hence market forces (Bunyecha and Tamminga, 1995; Ramadhani et al., 1998).

Following reforms, as cotton prices rose in the late 1990's, price competition and overcapacity in ginning caused abandonment of zoning, leading to the mixing of infected and uninfected seed and ultimately reduction in cotton fibre quality. The northern and southern area varieties, which

were released for specific agroclimatic conditions of the area, were also mixed (Government of Tanzania, 1999a; TCL and SB, 2002).

2.6 Lipid and fatty acid composition

Lipids are a group of naturally occurring compounds or biological molecules which are readily soluble in organic solvents such as hydrocarbons, chloroform, benzene, ethers and alcohols but insoluble in aqueous solutions (Gurr and Harwood, 1991; Michael, 2001; Christie, 2003a).

The major roles of lipids can be described, although individual lipids may have several different roles: 1) Structural lipids: these lipids play an important part in biological structure/membranes which provide barriers that protect organisms against their environment like the surface of the skin, fur of animals, surface of leaves in plants and walls of micro-organisms (Harwood, 1996; Michael, 2001). They also occur within the cell, providing a structure in which many metabolic reactions take place (Salunkhe *et al*., 1992). 2) Storage lipids: fatty acids in the form of simple glycerides, constitute an important source of fuel in mammals and in many plants. Many seeds store triacylglycerols to provide energy for the germination process. In animals, storage fat may be delivered directly from fat in the diet or may be synthesised in the adipose tissue (Gurr and Harwood, 1991; Harwood, 1997). 3) Lipids in metabolic control: lipids participate in the transmission of chemical messages in living organisms, others are fat soluble vitamins, while others act as precursors for a range of molecules with diverse metabolic activities [lipophilic bile acids which are involved in lipid absorption (Gunstone, 1967; Gurr and Harwood, 1991)]. They contribute significantly as functional ingredients in improving the sensory characteristics of several processed products. However, the essential fatty acids characterised by polyunsaturated fatty acids with 6 and 9 carbons have to be supplied in the diet because animals cannot synthesise them endogenously. 4) Plant lipids are used by the industry for detergents, nylon and cosmetic manufacture, as highly stable lubricants and as a renewable source of fuel (Harwood, 1997).

Lipids are classified into two classes. (1) Simple or "neutral" lipids are those in which hydrolysis yield at most two types of primary products per mole. (2) Complex lipids or "polar" that yield three or more primary hydrolysis products per mole (Christie, 2003a). About 70% of

edible fats are derived from plant sources (Salunkhe *et al.*, 1992). Nearly all commercially important fats and oils of animal and plant origin consist almost exclusively of simple lipids, triacylglycerols (often termed 'triglycerides'). Triacyglycerols consist of a glycerol moiety with each hydroxyl group esterified to a fatty acid (Padley *et al.*, 1994). The remaining part is 2-monoacyl-*sn*-glycerols, diacylglicerols, tocophenols, waxes, free fatty acids and polar lipids including phospholipids and galactolipids (Christie, 2003a). A stereospecific numbering system has been recommended to describe these forms. The prefix '*sn*' is placed before the stem name of the compound, when the stereochemistry is defined (Figure 2.1).

```
          A                              B
          H                              H
          |                              |
 H  — C — OH           H — C — OOCR'   position sn-1
          |                              |
 HO ← C → H            R''COO←C → H       position sn-2
          |                              |
 H  — C — OH           H — C — OOCR''' position sn-3
          |                              |
          H                              H
```
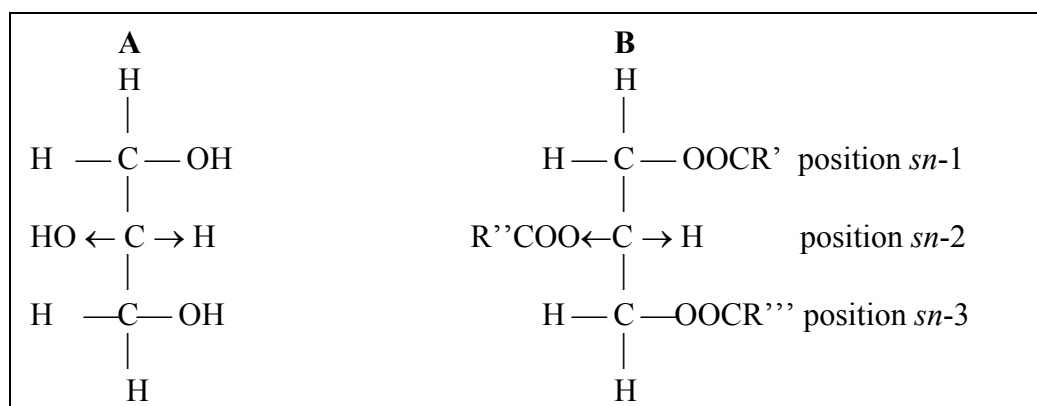
Figure 2.1 The stereochemical configuration of glycerol (A) and triacyl-*sn*-glycerol (B), (adapted from Christie 2003b)

Fatty acids are straight chain carbon acids usually with an even number of carbon atoms. Fatty acids are characterised by the number of carbon atoms (n) and number of double bonds (m) as (n:m) (Christie, 2003b). Fatty acids without double bonds are called saturated fatty acids (SFAs) like lauric (C12:0), myristic (C14:0), palmitic (C16:0) and stearic (C18:0) acids. Those with one bond are called monounsaturated fatty acids (MUFAs) like palmitoleic [(C16:1) (n-7)], oleic [(C18:1) (n-9)] and erucic [(C22:1) (n-9)] acid (Charley and Weaver, 1998) and those with more than one double bond are called polyunsaturated acids (PUFAs) like linoleic [(18:2) (n-6)], α-linolenic [(18:3) (n-3)] and δ-linolenic [(18:3) (2-6)] acid.

The relative amount of fatty acids present in oil and the distribution in triacylglycerol molecular species determine the physical, chemical, physiological and nutritional properties of vegetable oils (Murthi and Achaya, 1975; Padley *et al*., 1994). The composition of position *sn*-2 is of great importance when triacylglycerols are consumed and digested by animals, since 2-monoacyl-*sn*-glycerols are formed which can be absorbed by the intestine and utilised as such. Position *sn*-3 for example, is the last position to be acylated during triacylglycerol biosynthesis and this step is potentially important in the cellular control mechanism. Position *sn*-2 of the triacylglycerols of seed oils is greatly enriched in the polyunsaturated fatty acids (specifically linolenic and linoleic acids). Relatively little difference between the primary positions can be realised where less common fatty acids tend to be concentrated in position *sn*-3. Saturated fatty acids are concentrated in the primary positions and monoeonic acids are relatively evenly distributed (Figure 2.1) (Christie, 2003b).

Longer-chain fatty acids (C20-C24) are apparently concentrated in the primary positions with some preference for position *sn*-3. There are exceptions to these rules and in cacao butter for example, oleic acid is present largely in position *sn*-2. Minor differences only in the distribution of saturated and monoeonoic fatty acids between *sn*-1 and *sn*-3 have been observed but too few samples have been analysed for definitive comment. Some seed oils contain unusual fatty acids for example an allenic estolide was found entirely in position *sn*-3 in *Sapium sebiferum* (Christie, 2003b).

In the complex pathway of triacylglycerols biosynthesis, palmitate has different fates. One key enzyme is the β-ketoacyl- Acyl carrier protein (ACP) synthase II (KAS) (Harwood, 1996). In this pathway saturated fatty acids, palmitic and stearic acids are synthesised and stearic acid is subsequently desaturated to oleic, linoleic and linolenic fatty acids (Harwood, 1997). The majority of polyunsaturated fatty acids are synthesised through the 18:1 desaturase, in the endoplasmic reticulum (Browse, 1991). KAS is exclusively responsible for the condensation of C16:0-ACP with malonyl-ACP to stearoyl-ACP, thus determining the C16/C18 fatty acid ratio of seed oil. However, palmitate may be released from palmitoyl-ACP by an acyl-ACP-thioesterase and re-esterified on the chloroplast envelope to coenzyme A (C16:0-CoA). Alternatively, palmitoyl-ACP may be used within the chloroplast by an acyltransferase to form

phosphatidic acid (PA) that can subsequently be desaturated by plastidic enzymes. The palmitate content found in triacylglycerols (TAG) is determined by the competitive activity of a thioesterase, an acyltransferase and KAS II (Möllers and Schierholt, 2002).

The C18 polyunsaturated fatty acids, linoleic [(C18:2) (n-2)], α-linolenic [(C18:3) (n-3)] or *cis* 9, *cis* 12, *cis* 15-octadecatrienoic [(C18:2) (n-3)] and δ-linolenic or *cis* 9, *cis* 12-octadecadienoic acid [(C18:3) (n-6)] are major components of most plant lipids (Christie, 2003b). Harwood (1997) and Gurr and Harwood (1991) reported that common fatty acids of animal and plant tissues are C16 and C18 straight chain compounds with zero to three double bonds of *cis* (or Z) (the 2 hydrogen substituents are on the same side of the molecule) configuration.

Fats and oils account for a substantial portion of the calorific value of the human diet, being ingested in their natural form as components of whole foods or in their extracted form either as ingredients in processed foods or as cooking mediums, salad oils and spread (Krawezyk, 2001). Dietary intake of fatty acids significantly increases the levels of total cholesterol in the bloodstreams contributing to increased occurrence of arteriosclerosis and consequently a greater risk of cardiovascular disease (Stamler and Shekelle, 1988; Liu *et al.,* 2002). Fatty acids in fats and oils can themselves have significant effects on serum cholesterol levels (Figure 2.2).
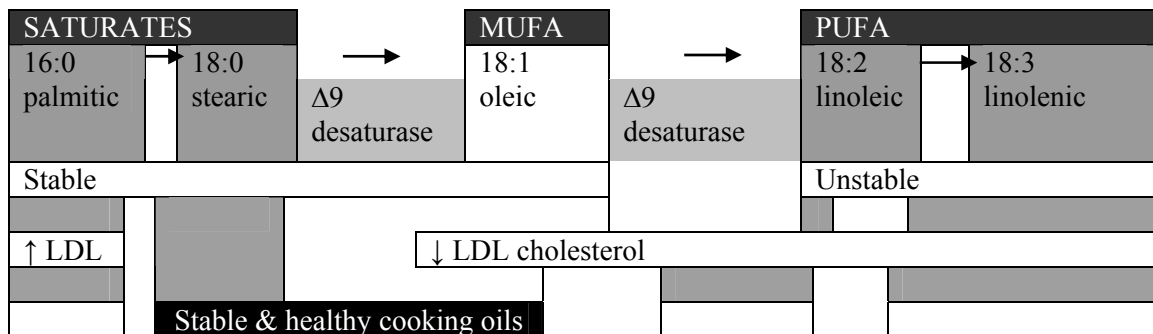
Figure 2.2 Schematic diagram of the biosynthetic pathway for the major saturated (SFA), monounsaturated (MUFA) and polyunsaturated (PUFA) fatty acids in oilseeds and their key nutritional and functional attributes (adapted from Liu *et al*. 2002)

Initially it was considered that all saturated fatty acids and in particular myristic acid (C14:0), stearic acid (C18:0) and palmitic acid (C16:0), the principle saturated fatty acids present in plant oils, had the undesirable property of raising serum low density lipoprotein (LDL) cholesterol levels (Mensink and Katan, 1992; Zock *et al.*, 1994). However, it was revealed that stearic acid (C18:0) does not raise LDL-cholesterol like other saturates and may lower the total cholesterol, thus considered to be neutral with respect to risk of cardiovascular disease (Dougherty *et al.*, 1995; Liu *et al.*, 2002). On the other hand, unsaturated fatty acids, such as monounsaturated oleic acid (C18:1) and polyunsaturated linoleic acid (C18:2) and α-linolenic acid (C18:3), have the beneficial property of lowering LDL-cholesterol, thus reducing the risk of cardiovascular disease (Mensink and Katan, 1992).

Highly unsaturated oils are unstable when exposed to high temperatures and oxidative conditions for long periods of time. This results in the development of short chain aldehyde, hydroperoxide and keto derivatives, imparting undesirable flavours and reducing the frying performance of the oil by raising the total level of polar compounds (Chang *et al.*, 1978). Polyunsaturated oils can, however, be converted into stable cooking oils by hydrogenation in which the carbon double bonds (unsaturated) are reduced to single bonds (saturated). However, partial hydrogenation results in the breakdown of naturally occurring *cis* carbon bonds and occasional reformation in *trans* configuration (Ray and Carr, 1985), forming *trans*-fatty acids (the two hydrogen constituents are on opposite sites) (Gurr and Harwood, 1991). In contrast to *cis*-unsaturated fatty acids, *trans*-fatty acids are known to be as potent as palmitic fatty acid in raising plasma LDL cholesterol levels (Noakes and Clifton, 1998) and lowering plasma high density lipoprotein (HDL) cholesterol (Zock *et al.*, 1994).

Although cotton is grown mostly for fibre, the seeds are an important source of oil. The estimated world production of cottonseed oil in 1985 was 3.57 million metric tons ranking fifth in vegetable oil production after soybean, palm, rapeseed and sunflower (Hatje, 1989). World production of cottonseed oil was about 4 million metric tons in both 1997 and 1998 (Jones and Kersey, 2002).

2.7 Molecular marker technology

Characterising genetic diversity and degree of association between and within varieties is the first step toward developing germplasm and crop cultivars. Successful crop improvement depends on genetic variability that arises from genetic diversity (Rana and Bhat, 2004). A lack of genetic diversity may limit breeding progress and gain from selection. A variety of molecular marker technologies have been used to study the genetic diversity and relationship within species and between their wild relatives (Cornelius and Sneller, 2002).

DNA fingerprinting involves the display of sets of fragments from specific DNA samples. It is an effective tool to increase the speed and quality of backcrossing conversion, thus reducing the time taken to produce crop varieties with desirable characteristics (Farooq and Azam, 2002; Murtaza *et al*., 2005). With the use of molecular techniques, it is now possible to hasten the transfer of desirable genes among varieties and to introgress novel genes from related species. Using DNA fingerprinting, polygenic characteristics can be easily tagged and genetic relationships between sexually incompatible crop plants can be established (Altaf Khan *et al*., 2002; Rana and Bhat, 2004).

A number of DNA fingerprinting techniques are presently available. These techniques have been developed over the past few years to provide genetic markers capable of detecting differences among DNA samples across a wide range of scales (Vos *et al.,* 1995; Blears *et al*., 1998). Molecular markers possess many advantages, which make them superior to morphological markers. Molecular markers offer a great scope for improving the efficiency of conventional plant breeding by carrying out selection not directly on the trait of interest but on molecular markers linked to that trait. Furthermore, these markers are used in mapping of specific genes, cultivar identification and biodiversity studies (Rana and Bhat, 2004). Molecular markers are not environmentally influenced and are detected in all plant growth stages (Kumar, 1999; Rungis *et al.,* 2000). DNA based markers are considered the most suitable markers for genetic distance estimates because of potentially large numbers of polymorphisms (Gepts, 1993).

Recently, with the advent of protein and molecular marker techniques, genetic diversity in crop germplasm has been assessed at protein and DNA level (Mueller and Wolfenbarger, 1999).

Protein based markers include isozymes (Market and Moller, 1959). DNA based markers include restriction fragment length polymorphism (RFLP) (Liu and Turner, 1993), random amplified polymorphic DNA (RAPD) (Williams *et al.,* 1990), amplified fragment length polymorphism (AFLP) (Zabeau and Vos, 1993), microsatellite or simple sequence repeat (SSR) (Akkaya *et al.,* 1992) and single nucleotide polymorphism (SNP) (Bojinov and Lacape, 2003).

2.7.1 Molecular markers and application in cotton
2.7.1.1 Restriction fragment length polymorphism (RFLP)

RFLP markers facilitate the selection of progeny with desirable genotypes in a short span of time, are co-dominant and can identify unique loci (Natalija, 2001). Polymorphisms detected by RFLP markers are reliable and can be used for accurate scoring of genotypes. Bands visible on an autoradiogram represent restriction fragments of the digested DNA that contain sequences homologous to the cloned sequences used as probe (Liu and Turner, 1993; Farooq and Azam, 2002). RFLP analysis is highly repeatable and produces one to five polymorphic fragments. RFLP requires relatively large amounts of pure DNA which is difficult to isolate in cotton due to abundance of phenolic compounds. It is labour intensive, time consuming and expensive compared to other newly developed polymerase chain reaction (PCR) based techniques (Tanksley *et al.*, 1989).

RFLP (Liu and Turner, 1993) analysis has been used in different studies including genetic diversity and determining genetic similarity among *Brassica oleracea* L. (Santos *et al.,* 1994) and *Zea mays* L. (Smith *et al.,* 1990). In a study of heterosis and combining ability of cotton, Meredith and Brown (1998) assayed 16 parents using RFLP analysis and observed that the correlation of genetic distance and midparent heterosis was small (r=0.08). Paterson *et al.* (1999) used RFLP analysis to determine genetic diversity in relation to evolution of diploid and allotetraploids in cotton and observed that allotetraploid A_t and D_t genomes and A and D diploid genomes were recombinationally equivalent despite a nearly two-fold difference in physical size. RFLP has been applied to several cotton species to study evolution, population genetics and phylogenetic relations but revealed low variation in cotton compared to other taxa (Brubaker *et al.,* 1994). In comparisons, levels of allozyme variation were higher than levels of RFLP variation (Wendel and Brubaker, 1993; Brubaker *et al.,* 1994). Since cotton is an allotetraploid with a large genome, it is desirable to have efficient DNA assay systems for

development of large numbers of polymorphic markers to cover the entire genome in a relatively short time frame (Brubaker *et al.,* 1994).

2.7.1.2. Random amplified polymorphic DNA (RAPD)

RAPD (Williams *et al.,* 1990) was another breakthrough to find a solution for breeding problems. RAPD analysis detects nucleotide sequence polymorphisms in DNA amplification based on assays using a single primer with an arbitrary nucleotide sequence (Altaf Khan *et al.,* 2002). RAPD analysis tends to provide only dominant markers. Despite this limitation, mapping using dominant markers linked in coupling is on a pre-gamete basis as efficient for mapping as co-dominant markers (Tingey and del Tufo, 1993). The RAPD technique is PCR based and requires low amounts of DNA and produces one to 10 polymorphic fragments per reaction. Generated DNA fragment patterns depend on the primer sequence and nature of template DNA (Williams *et al.*, 1990). Disadvantages of RAPD markers for phylogenetic studies include that the genomic origin (nuclear or cytoplasmic) of fragments and the sequence homology of fragments with similar mobility in a gel, are not known and the RAPD technique is not repeatable (Williams *et al.*, 1990; Karp *et al.*, 1997).

Lu and Myers (1999) studied the genetic relationships in 10 influential upland cotton varieties using RAPD markers and observed that the most important germplasm represented by highly influential cotton lines lacked variation at DNA level. Tatinen *et al.* (1996) studied genetic diversity of 16 near-homozygous elite cotton genotypes of *G. hirsutum* and *G. barbadense* using 135 RAPD markers as well as morphological characteristics. Both procedures generated dendrograms consisting of two clusters, one resembling *G. hirsutum* and the other *G. barbadense.* Classification of genotypes based on the two methods gave similar results with a correlation of 0.63 between genetic and taxonomic distances. Several genotypes were identified that were genetically and phenotypically distant from typical *G. hirsutum* and *G. barbadense.* RAPDs have been used to evaluate elite cotton commercial cultivars (Multani and Lyon, 1995; Iqbal *et al.*, 1997), tag the *cms-D8* restorer gene (Zhang and Zhang, 1997), tag genes influencing general combining ability effects for yield components (Lu and Myers, 2002) and construct bispecifc (Yu and Kohel, 1999) and trispecific (Altaf Khan *et al.,* 1998, 1999) genomic maps of cotton.

2.7.1.3 Amplified fragment length polymorphism (AFLP)

AFLP analysis was developed by Zabeau and Vos (1993). AFLP is a DNA fingerprinting procedure that takes advantage of RFLP and PCR to amplify a limited set of DNA fragments from a specific DNA sample (Vos *et al*., 1995). The basic difference between RFLP and AFLP analysis is that with RFLP only restriction sites determine polymorphism, but in AFLP restriction sites plus additional selective nucleotides determine polymorphism (Becker *et al*., 1995). AFLP is an efficient PCR based technique used to generate large numbers of polymorphic DNA fragments. This property is referred to as a high multiplex ratio (Rana and Bhat, 2004).

The AFLP technique requires no prior knowledge of nucleotide sequences because it uses adapters of known sequence ligated to restriction fragments and allows specific co-amplification of high numbers of restriction fragments. AFLP analysis provides a novel and powerful DNA fingerprinting technique for DNA of any origin and complexity. Depending on the resolution of the detecting system, typically 50-100 restriction fragments are amplified and detected. Different systems include denaturing polyacrylamide gel electrophoresis (PAGE) and automated capillary sequencers (Natalija, 2001; Altaf Khan *et al*., 2002).

The AFLP technique can be used to map chromosomes and fill gaps on chromosome segments on which no RFLP loci had previously been mapped (Becker *et al*., 1995). AFLP is a powerful, efficient, reliable, stable, reproducible and rapid assay with genome mapping applications. AFLP can be used for determining genetic relationships among populations, cultivar identification and germplasm evaluation (Thomas *et al*., 1995; Maughan *et al.,* 1996; Tohme *et al*., 1996). However, AFLP markers are dominant (Maughan *et al*., 1996; Sharma *et al*., 1996). AFLP as a tool for evaluating genetic relationships among populations and cultivar evaluation is reproducible even against the background of different combinations of *Taq* DNA polymerases and buffers (Tohme *et al.,* 1996; Altaf Khan *et al*., 2002). Its capacity to detect large numbers of independent genetic loci with minimal cost and time requirements makes it an ideal marker system for a wide array of genetic investigations (Maughan *et al*., 1996). AFLP is unique since common sets of primers can be established among different plant species for comparative studies. These unique characteristics make AFLP analysis an excellent method for detection and study of genetic polymorphism in a wide array of plant species (Altaf Khan *et al*., 2002).

The application of the AFLP technique in genetic diversity studies has shown great success among a wide range of crops like soybean (*G. max*) (Maughan *et al*., 1996) and sunflower (*Helianthus annuus)* (Liu *et al.,* 2003). In mapping studies, AFLP analysis has been used in different crops like rice (*Oryza sativa* L.). Studies showed that the AFLP technique was the most efficient way to generate large numbers of markers that are linked to target genes (Zhu *et al.,* 1998).

AFLP analysis has been applied in cotton to identify genes for resistance to fungal wilt diseases. It showed a greater potential compared to conventional breeding since it reduced the selection time and used small numbers of plants for detection of resistance genes (Bruce *et al.,* 2001). Liu *et al.* (2001) used AFLP analysis to determine whether rapid genomic changes associated with non-Mendelian genomic changes in early generations following polyploid synthesis also occurred in allopolyploid cotton *(Gossypium*) species. The extent of fragment additivity in newly combined genomes was ascertained for a total of approximately 22000 genomic loci and was observed in nearly all cases. This indicated that rapid and unexplained genomic changes did not occur in allopolyploid cotton. These data indicated that polyploid speciation in plants is accompanied by a diverse array of molecular evolutionary phenomena, which will vary among both genomic constituents and taxa (Liu *et al.,* 2001).

Altaf Khan *et al*. (1997) used AFLP analysis to study inheritance patterns of segregating loci and to establish linkage groups among trispecific cotton species in a segregating F_2 population. A total of 216 markers (194 AFLPs, 19 RAPDs and three morphological markers) were scored, of which 85 showed normal Mendelian inheritance. Preliminary evaluation results indicated that all measured quantitative traits showed a high degree of genetic variation. Significant deviation from the expected 3:1 dominant segregation ratio was observed. It was suggested that combined data from molecular, morphological and quantitative traits could be used to construct genetic linkage maps that would be useful for identifying alien introgressions and economically important traits in the trispecific F_2 population (Altaf Khan *et al*., 1997). Rana and Bhat (2004) found AFLP analysis to be more efficient for diversity study analysis and cultivar identification compared to RAPD analysis.

Pillay and Myers (1999) assessed the level of AFLP polymorphism in Old and New World cotton species. Four AFLP *Eco*RI-*Mse*I primer pair combinations produced a 10-fold increase in the number of DNA fragments per plant compared to RAPD analysis. AFLP data assigned the genotypes into groups corresponding with origin and/or pedigree relationships. Iqbal *et al*. (2001) used AFLP analysis to determine evolution in upland cotton and realised that at species level and above, genetic similarity based on AFLP analysis was in agreement with known taxonomic relationships. Abdalla *et al*. (2001) studied the genetic diversity and relationship of diploid (*G. herbaceum*, *G. raimondii* and *G. arboreum*) and tetraploid (*G. hirsutum* and *G. barbadence*) cotton. AFLP analysis was useful for estimating genetic relationships across a wide range of taxonomic levels and for analysing the evolutionary and historical development of cotton cultivars at genomic level.

In the process of breeding for low-gossypol seed and high-gossypol plants in upland cotton, Vroh Bi *et al*. (1999) used AFLP analysis to assess genetic similarity among germplasm and RFLP probes to tag the introgression of specific chromosome segments from parental species. Genetic similarity between upland cotton and wild species ranged from 29.5-43.2%, while similarity reached 80% between upland cotton and BC3 plants. Zhong *et al*. (2002) used AFLP analysis to assess day-neutrality in cotton backcross populations. Genetic distances among the recurrent parent and backcross populations ranged from 0.35-0.75. Genetic distances among the non-recurrent parent DPL 16 and backcross populations ranged from 0.16-0.38. Many AFLP markers tended to cluster together as linked blocks and were selected with the day-neutral flowering phenotype. This finding indicated that linkage drag was occurring during introgression of the day-neutral flowering trait.

2.7.1.4 Microsatellites or simple sequence repeats (SSR)

SSR (Akkaya *et al.,* 1992) analysis detects variation at individual loci and has been thought of as the "new allozyme". SSRs are found throughout genomes of many eukaryotes and constitute an abundant source of DNA markers (Natalija, 2001). SSRs consist of tandem repeated DNA sequences, like $(AT)_n$ surrounded with specific sequences and are mainly located in the non-coding part of the genome. Polymorphisms are due to variation in the number of detected repeats (Tautz, 1989). SSRs are amplified by PCR, using flanking primers. SSRs are highly polymorphic and provide co-dominant genetic markers following Mendelian inheritance, which

increases the efficiency and accuracy of population genetic measures compared to other markers like AFLPs and RAPDs (Russell *et al*., 1997). SSRs carry two-fold more information than AFLPs and RAPDs and 40% more information than RFLPs when the number of alleles per locus is the target (Pejic *et al.,* 1998). Unless useful primers have been designed in previous studies, it is necessary to screen the organism for microsatellites before primers can be developed. Screening is practically complex and expensive and may yield only a small number of potential microsatellite loci (Robinson and Harris, 1999).

SSRs have been used to study different crop species' structure for example in *Clusiaceae* (Aldrich *et al*., 1998) and to study variation within and between populations of *Myrtaceae* (Rossetto *et al*., 1999). A large number of SSR primer pairs were developed for cotton and are being utilised in various cotton genomic projects (Cantrell, 2000; Reddy *et al*., 2000). Liu *et al*. (2000) used 66 primer pairs to amplify 70 marker loci using 13 monosomic and 28 mono-telo-disomic cotton cytogenetic stocks. The aim was to develop anchor SSRs for cotton chromosomes to provide the basis for a framework genetic map. Forty two SSR loci were assigned to cotton chromosome arms. Twenty-six SSRs were not located on informative chromosomes. By screening *G. herbaceum* (2n=2x=26=2A1) and *G. raimondii* (2n=2x=26=2D5) accessions, 19 SSRs were clearly shown to occur on the A subgenome and 11 on the D subgenome. Sutirtha *et al*. (2001) used 11 cultivars to analyse genetic relationships among improved cotton genotypes and the association with F_2 hybrid performance. Ninety-six SSR primer pairs were screened that resulted in 102 polymorphic SSR markers. Similarities of 0.80 to 0.99 were observed indicating a narrow genetic base in cotton. Liu *et al*. (2003) studied inheritance and linkage mapping of fertility restoration for cytoplasmic male sterility in *G. hirsutum* by analysing monosomic and telosomic lines using SSRs. Results revealed that the *Rf1* locus could be located on the long arm of chromosome 4.

Molecular markers are becoming increasingly attractive markers in molecular breeding and diversity assessment (Powell *et al*., 1996; Rana and Bhat, 2004). The choice of which fingerprinting technique to use depends on the application (e.g. DNA genotyping, genetic mapping or population genetics), the organism under investigation (e.g. prokaryotes, plants, animals or humans) and resources (time and money) available. In most cases not one fingerprinting technique is ideal for all applications (Blears *et al.,* 1998).

2.8 Important characteristics in cotton improvement

Cotton cultivars may react differently to different production areas and have different characteristics but mainly are either hairy or non-hairy, normal or okra leaf shaped, frego bract or normal bract, reddish or green coloured varieties. Hairy plant types are used to resist jassids in Africa and Asia. Cotton varieties without hairs (glabrous) offer resistance to *Helioths* spp. and pink bollworm by incurring decreased egg laying that is associated with decreased trash content of fibre (Thaxton and El-Zik, 1994). Okra leaf shaped plants have a more open canopy, which permits 70% more light penetration. This reduces the numbers of boll weevil, pink bollworm and boll rot. Okra leaf shape is associated with accelerated fruiting rates, early maturity and production of fibre with less trash than normal leaf cultivars. Normal leaf cotton has taller plants than okra leaf cotton and lower fruit loss at 43% compared to 59% for okra leaf plants (Andries *et al.,* 1969; Reddy, 1974).

The frego bract trait is associated with a high level of resistance to boll weevil and can reduce boll weevil damaged squires up to 50% compared with normal bract (Jones, 1972; Jenkins, 1976). Frego bract is associated with delayed fruiting in maturity in addition to reduced yield (Jones, 1972; Thaxton *et al*., 1985). Red plant colour confers significant degrees of non-preference to the boll weevil and cotton aphid damage. Varieties with the smooth-leaf trait generally give higher fibre grades than those with normal or densely hairy leaves (Thaxton and El-Zik, 1994).

Yield: Yield refers to the total seedcotton and lint yield. Yield is a composite of many other traits, each influenced by many genes that have variable effects and are modified by environmental conditions and cultural methods (Christidis and Harrison, 1955; Meredith, 1984). Cotton varieties vary in yield potential, therefore varieties producing high seedcotton and high lint yield are important to the client. Andries *et al*. (1971) reported lower yields with okra leaf cotton compared to normal leaf cotton. Lint and seed yield are highly significantly positively related. As one tends to increase, so does the other. Selections should not be based on seedcotton yield, instead selections should be based on lint yield as lint yield depend on seedcotton and ginning outturn (Thaxton and El-Zik, 1994).

Number of bolls and boll size: The boll is the unit package of yield, thus high yield is achieved when the size and number of bolls per unit area is maximised. Kerr (1966) and Coyle and Smith (1997) suggested that prolificacy (boll number per plant or per unit area) is an important factor to consider during selection for yield improvement. Yield models described by Worley *et al.* (1976) supported this. Selection for boll size and seed size could positively influence lint yield, if a breeder selects for medium boll size, small seeds per boll and maintaining high ginning percentage (Coyle and Smith, 1997).

Plant height: Plant height is important as a contributor to yield and can determine vegetative and fruiting branches of the plant. Breeding for plant height variation is influenced by both yield potential and harvesting methods (Niles and Feaster, 1984). Kohel and Benedict (1987) observed plant heights between 0.95-1.07 m while Emeetai-Areke (1999) reported that the plant height of cotton ranged from 1.0-2.0 m. Generally plant height is highly affected by the environment.

Hairiness: Hairiness of the leaf and other parts of the plant is heritable and varies between varieties. Hairiness is important for insect pest (jassid) resistance. The hairs on leaves and stems interfere with the ovipostion and laying of eggs, thus reducing the rate of damage. Cotton leaf hairs are stellate and vary both in length and density from sparse to densely hairs (Munro (1987).

Ginning outturn (GOT): This is the percentage of lint obtained from a sample of seedcotton and varies between cotton varieties and for upland strains ranges between 30-40%. Christidis and Harrison (1955) and Munro (1987) reported that the range of varieties with regard to ginning percentages was shown to change little from year to year and from place to place. Singh and Singh (1980) and Carvalho and De-Carvalho (1995) studied genetic control of ginning outturn. Results indicated that ginning outturn is controlled by additive genetic effects. Singh *et al.* (1990) reported non additive effects controlling ginning outturn.

Seeds per boll: Munro (1987) reported that seeds per boll and the number of locules is a characteristic of the species or a variety. In *G. hirsutum* there is mainly eight seeds per locule and the locule per boll varies between three to five. Seeds are the units of production and fibres

grow from the outer cells of seed surfaces. The higher the number of seeds per boll, the more lint is produced because it increases the amount of surface area for lint production (Culp and Harrel, 1973). Therefore, breeding for increased bolls per unit land area, more seeds per boll, large seed surface area per unit seed weight and increased weight per unit seed surface are importanct (Smith and Coyle, 1999). Worley *et al.* (1976) reported that seeds per boll are the second largest contributor to yield.

Seed and lint index: Seed index refers to 100 seed weight. Seed index is important in determining the yield, especially in seed cotton. It varies between varieties and is highly affected by population density. Cottonseed measures about 10x6 mm and weighs about 80 mg (5-10 g per 100 seed) (Munro, 1987). Each pure line has its own particular mean seed weight to which it breeds true. It is a characteristic subjected to great influence of boll size and number of seeds per locules (Sikka and Joshi, 1960).

Lint index represents the absolute weight of lint borne by a single seed (or more often 100 seeds). Lint index has a direct relationship with the yield potential of a genotype but is affected by population density (Munro, 1987). It is a compound characteristic being a function of mean number of hairs per seed and mean hair weight. Sikka and Joshi (1960) reported that lint index is governed by two genetic systems, a single pair of factors having pleiotropic effects and a complex of modifiers, which have minor effects on lint production. Lint index is controlled by additive genetic effects (Singh *et al.*, 1990).

Fibre strength: This fibre quality trait is useful for spinners and processors. The inherent strength of individual cotton fibres is an important factor in the strength of the thread spun from them. High tensile strength of fibres is necessary for good spinning properties, especially with modern fast spinning machines (Niles and Feaster, 1984; Munro, 1987). Fibre strength is affected by environmental fluctuations (Christidis and Harrison, 1955).

Fibre length: This is the staple length that is universally recognised as the premier fibre property, because it is closely associated with the processing efficiency in manufacturing and determining the quality of the yarn produced. Fibre length variation can occur from boll to boll and plant to plant. Even on a single seed the hairs are not of the same length (Munro, 1987).

Fibre fineness (Micronaire): This is the measure of soft or silky feel. It is an important quality trait of cotton associated with long hairs and smaller cell diameter in combination with wall thickness. Fibre fineness determines the texture of cotton fibre into soft and silky or coarse and harsh and is affected by the environment (Sikka and Joshi, 1960). Christidis and Harrison (1955) reported that course lint is dominant over fine lint and is quantitatively inherited. Micronaire is acceptable anywhere within the base range of 3.5-4.9 units inclusive. The premium range is between 3.5-4.2, with values below 3.5 too fine and above 4.9 too course (Patil and Singh, 1994).

Uniformity of fibre length: It is an important fibre quality characteristic determining the maturity of the fibres. The value is important in determining the spinning performance and utility of the lint. Higher values are an indication that the yarn spun from such fibres will be uniform in size and strength, with less wastage of fibres. Uniformity varies between varieties and it is affected by environmental factors (Christidis and Harrison, 1955).

2.9 Genetic variance

Genetic variation is defined as the inherent characteristic of all living organisms that in the population provides the information necessary to choose the best selection strategy for that population and is partitioned into three components attributable to different causes (Meredith, 1984).

1. The additive variance effect is the average effect of genes. The resemblance between parents and offspring is largely due to additive genetic effects and is responsible for determining the response of the population to selection.

2. The dominance effect is the interaction of allelic genes. This represents the deviation of the heterozygote from the average of the parents.

3. Non-allelic interaction or epistatic effect is the interaction of non-allellic genes that influence a particular trait (the interaction deviation is the result of epistatic effect) (Meredith, 1984).

Genetic variation is described in statistical terms as:
The phenotypic variance of the population is a function of genotypic and environmental variance.

$V_P = V_G + V_E$

Where: V_P = Phenotypic variance, V_G = Genotypic variance, V_E = Environmental variance.

The genetic variance (V_G) is further divided into breeding value (additive variance), dominance variance and interaction variance.

$V_G = V_A + V_D + V_E + V_I$

Where: V_A = Additive variance, V_D = Dominance variance, V_E = Environmental variance, V_I = Interaction variance (Falconer, 1989).

The breeding value of a line is a function of the additive gene action. The additive genes are directly transported from the parents to the offspring and are responsible for the resemblance between relatives and can be used to calculate inheritance. The dominance additive ratio indicates the degree of dominance. Dominance:additive ratio of less than one refers to partial dominance, near one indicates complete dominance and greater than one indicates over-dominance (Falconer, 1989).

Kapoor (1994) and Turner *et al*. (1976) indicated that epistasis for seedcotton yield per plant, boll weight and ginning outturn was of duplicate type, thus additive and dominance gene effects have been found to be important in upland cotton. However, it varied from characteristic to characteristic. Gad *et al.* (1974) and Singh and Singh (1980) reported additive genetic variation for seedcotton yield, number of bolls, ginning outturn and lint index. Sayal and Sulemani (1996) reported over-dominance on lint percentage, seed index, lint index and staple length from a 8 x 8 diallel cross and additive effects for seedcotton yield. Carvalho and De-Carvalho (1995) studied fibre percentage and boll size in four varieties of *G. hirsutum* and 12 hybrids from a complete diallel set of crosses. Both traits showed incomplete dominance. Additive gene effects predominated in the control of both traits. Ahmad *et al*. (1997) observed additive gene action with partial dominance for bolls per plant, boll weight, seedcotton yield and seed index. Epistatic effects were involved in the expression of all the characteristics except for boll weight. Genetic markers can determine genetic variation, which makes it possible to determine the relationships between different genotypes and to forecast which pairings can produce new and superior gene combinations (Sharma *et al.,* 1996).

2.10. Diallel analysis

Hayman (1954) defined 'diallel cross' as the set of all possible matings between several genotypes. The genotypes may be individuals, clones, homozygous lines, etc. and if there are 'n' of them there are 'n^2' mating combinations ('n' inbred lines are crossed, 'n^2' progeny families are produced), counting the reciprocals separately. Diallel mating designs permit estimation of the magnitude of additive and non-additive components of heritable variance (Griffing, 1956; Mather and Jinks, 1977). Data obtained from such cross combinations can be analysed in several ways, but most commonly, analyses are based on the procedure proposed by Hayman (1954) and Griffing (1956). On the basis of these premises, a test for the validity of the additive-dominance model has been suggested. It is possible to obtain estimates of additive and dominance components of heritable components of variation from the mean squares of these mating designs (Hayman, 1954; Mather and Jinks, 1977).

2.10.1 Combining ability

Griffing (1956) proposed a more general procedure for diallel analysis, which makes provision for non-allelic interaction. According to this approach, mean measurement of a cross is partitioned into major components, apart from the general mean (μ) and environmental variance:

(1) General combining ability (GCA) is used to designate the average performance/contribution of the parents/line in hybrid combination (Sprague and Tatum, 1942). Falconer and Mackay (1996) defined it as the mean performance of the line in all crosses, when expressed as a deviation from the mean of all crosses. GCA consists of additive and additive epistatic variances (Matzinger, 1963).

(2) Specific combining ability (SCA) is used to designate those cases in which certain combinations do relatively better or worse than would be expected on the bases of the average performance of the line involved (Sprague and Tatum, 1942). It is the deviation to a greater or lesser extent from the sum of the GCA of its two parents. SCA consists of dominance and all types of epistatic variances are regarded as an estimate of effects on non-additive gene actions (Falconer and Mackay, 1996).

GCA and SCA effects help locate parents and crosses that are responsible for bringing about a particular type of gene action (Baker, 1978; Meredith, 1984). GCA and SCA effects and variances are effective genetic parameters of direct utility to decide the next phase of the

breeding programme (Arunachalam, 1976; Dabholkar, 1992). It helps selection of parents for construction of synthetics, selection of suitable F_1s for a multiple crossing or composite breeding programme and the possibility of employing an appropriate selection technique like modified mass selection, recurrent selection and reciprocal selection (Dabholkar, 1992). Differences in GCA have been attributed to additive, additive x additive and higher order interactions of additive genetic effects in the base population, while differences in SCA have been attributed to non additive genetic variance (Baker, 1978).

Abdalla *et al*. (1999) in a study on cotton combining ability and genetic variance for yield and earliness characteristics observed that additive gene action was greater than dominance gene action for both yield and earliness. Epistasis affects the estimates of GCA and SCA mean squares, variances and other effects in an unpredictable manner (Baker, 1978). El-Adl and Miller (1971) found that GCA of F_1s was more important than SCA for lint yield and the components of yield with the exception of lint percentage. Tang *et al.* (1993a), Lee *et al.* (1967) and Baloch *et al.* (1996) observed positive and negative GCA effects exerted by parents on boll weight, boll number, lint yield and lint percentage. SCA effects for lint percentage observed were significant and consistent across the environment. Theoretically the presence of significant GCA and SCA in the F_1 generation is a consequence of fluctuations in additive and dominance relationships among parents (Tang *et al.,* 1993a).

2.11 Heritability

Heritability is defined as the ratio of variance due to hereditary difference and genotypic variance to the total phenotypic variance (Meredith, 1984). The higher the ratio the more heritable the trait would be. If conversely, the ratio is smaller, the bigger the influence of the environment on the phenotypic expression of the trait. Thus, it expresses the proportion of the total variance that is attributable to the average effects of genes.

Heritability can be defined in two senses:
1. Broad sense heritability includes total genetic variance (Meredith, 1984). Dudley and Moll (1969) defined it as the ratio of total genetic variance to phenotypic variance and it expresses the extent to which individuals' phenotypes are determined by their genotype (Dabholkar, 1992).

$h^2 = V_G/V_P$

Where: h^2 = Heritability, V_G = Genetic value, V_P = Phenotypic value

2. Narrow sense heritability is the ratio of additive genetic variance to phenotypic variance (Dudley and Moll, 1969) and expresses the extent to which phenotypes are determined by the genes transmitted from parents. It is the breeding value (additive genetic variance) of the parents which determines the genetic properties of the progeny. Narrow sense heritability is used for determining selection progress estimates and selection indexes and determines the degree of resemblance between parents and offspring (Chaudhary, 1991; Meredith, 1984).

Narrow sense heritability measures the extent of correspondence between breeding values and phenotypic values and expresses the magnitude of genotypic variance in the population, which is mainly responsible for changing the genetic composition of the population via selection (Falconer, 1989). It provides a basis to predict accuracy with which selection for genotypes could be made based on phenotypic measurements of individuals or groups of individuals (Falconer, 1989; Dabholkar, 1992).

$h^2 = V_A/V_P$

Where: h^2 = Heritability, V_A = Breeding value, V_P = Phenotypic value

Also $h^2 = V_A/ (V_P = V_A/V_A + V_D + V_I + V_E)$

$V_p = V_A + V_D + V_I + V_E$

Where: V_A = Additive variance, V_D = Dominance variance, V_I = Interaction variance, V_E = Environmental variance (Falconer, 1989).

Heritability is a property not only of the characteristic being studied, but also of a population being sampled and the environmental conditions to which individuals have been subjected (Falconer, 1989; Dabholkar, 1992). Populations which are genetically more uniform are expected to show lower heritability than genetically diverse populations. Since environmental variance forms part of phenotypic variance, it affects the magnitude of heritability. Tang *et al.* (1992; 1996) observed a relatively high heritability for fibre length and strength of cotton and low heritability for elongation and micronaire. Dedaniya and Pethani (1994) reported that lint

yield per plant, seedcotton and number of bolls per plant had high to moderate heritability estimates. Siddiqui (1997) observed that heritability estimates were high for seedcotton per plant, 2.5% span length, plant height and weight of 25 burst bolls.

Luckett (1989) observed substantial additive effects and high heritability for boll size, span length, fibre strength and lint percentage on some of the plant materials. Lancon *et al*. (1993) observed a relatively high heritability for plant height, flowering earliness, number of bolls on vegetative branches, uniformity, fineness, maturity, strength, fibre percentage and seed weight. Carvalho *et al*. (1995) in a study of six *G. hirsutum* varieties and hybrids from a complete diallel set of crosses observed a low heritability estimate (0.19) for yield, which was controlled mainly by dominant genes.

2.12 Heterosis

Hartl and Clark (1989) defined heterosis as the phenomena of enhanced hybrid performance. Falconer (1989) defined heterosis as the difference between the crossbred and inbred lines or the superiority over inbred lines. Heterosis is usually defined in two ways. For those interested primarily in the F_1 performance *per se*, the F_1 minus the highest performing parent, expressed as a percentage of that parent used and is referred to as "useful heterosis" (Meredith, 1984). Lamkey and Edwards (1999) and Meredith (1984) also defined heterosis as the F_1 minus mid parent expressed as a percentage of the mid parent, this is called "mid parent heterosis".

However, Flintham *et al*. (1997) stated that heterozygosity is an important prerequisite for heterosis, because heterosis can arise when over dominance at a given locus is a principal cause. Others, however, believe that dominance and epistasis are the underlying genetic basis of heterosis. Loci with no dominance do not cause heterosis. The amount of heterosis following a cross between two particular lines or populations depends on the square of the difference of gene frequency between the populations. If the populations crossed do not differ in gene frequency there will be no heterosis (Coors *et al.,* 1999).

Heterosis in the F_1 is $HF_1 = \Sigma dy^2$

Where: d = the deviation of the heterozygote from the homozygote mid parent, y = gene frequency.

Heterosis in cotton offers possibilities for increasing cotton production but the problem is to find proper cotton combiners (Tang *et al.*, 1993b). Heterosis is reduced by half in each subsequent generation. This is equally true when the joint effects of all loci are considered provided that epistatic interaction is absent. High levels of heterosis and specific combining ability for yield is ideal for development of hybrid cotton (Meredith, 1984). Subrahmanyam *et al.* (1989) observed marginal heterosis for fibre fineness where additive genetic and environmental components of variation were highly significant. Lee *et al.* (1967) and Miller and Marani (1963) obtained mid parent heterosis in F_1 hybrids of 26% for lint yield. Mid parent heterosis estimates for lint percentage and boll weight was smaller but significant. Xian *et al.* (1995) and Zhang and Zhang (1997) observed heterosis in all yield traits, which was greatest for seedcotton and lint yields.

Wang and Pan (1990) observed a significant parabolic regressional relationship between genetic distance values and heterosis where increased genetic distance was associated with increased heterosis. You *et al.* (1998) reported that path analysis showed that if mid-parent heterosis of boll weight and bolls per plant increased, mid-parent heterosis of seeds per boll and plant height were reduced and other traits were kept the same.

2.13 Correlations

Changes in one characteristic accompanied by a change in another variable is referred to as correlation. The coefficient of correlation is the measure of association between two characteristics. Correlations can be either positive, when an increase in one variable is accompanied by an increase in another one, or negative when an increase in one character is accompanied by a decrease in another (Falconer, 1989). Correlated characteristics are of interest for three reasons namely (1) genetic causes of correlation through the pleiotropic action of genes, (2) in connection to changes brought about by selection and (3) in connection with natural selection (Falconer and Mackay, 1996).

In plant breeding there are two types of correlations (Meredith, 1984):
1. Phenotypic correlation is the association between two characteristics that can be directly observed and determined from measurements of the two characteristics in a number of individuals of the population. Phenotypic values are determined by genotypic values

and environmental deviations. Phenotypic correlation comprises correlation due to environmental agencies and that due to non-additive genetic causes (Falconer, 1989; Dabholkar, 1992). If two characteristics have high heritabilties, correlation due to environmental agencies will be relatively less important (Falconer, 1989).

2. Genetic correlation is the correlation of breeding values which is a function of additive gene action (Falconer and Mackay, 1996). Genetic correlation between two or more characteristics may result from pleiotropic effects of genes or linkage of genes governing inheritance of two or more characteristics (Falconer, 1989). This expresses the extent to which two measurements are genetically associated. If the genetic correlation is high, selection for one trait will simultaneously results in changes of the other trait. This association may be either harmful or beneficial, depending upon the direction of genetic correlation and objectives of the breeder (Meredith, 1984).

Jixiang *et al.* (1996) and Hussain *et al.* (1998) observed that ginning outturn and lint index were positively and significantly correlated with each other. Fibre strength was also correlated with seedcotton yield. Tang *et al.* (1996) observed a high positive genetic correlation for boll weight with lint percentages, fibre strength and micronaire. Lint yield showed a genetic correlation with fibre strength and boll weight. Dedaniya and Pethani (1994) and Carvalho *et al.* (1995) observed that seedcotton yield per plant was positively correlated with number of bolls, plant height, boll weight, lint weight per plant and bundle strength tenacity. Negative correlation was observed among fibre strength and earliness, fibre length and fibre fineness, fibre length and fibre percentage as well as fibre fineness and fibre percentage (Carvalho *et al.*, 1995). Ibragimov (1989) observed close genetic correlation between relatively short fibre and high fibre outturn and high yield. Chen *et al.* (1991) observed that days from sowing to standard flowering date, days from sowing to practical flowering date, plant height and sympodia per plant were significantly positively correlated with each other and negatively correlated with first peak in cotton harvest.

2.14 Genotype x environment interaction

For the improvement of crop varieties, a better understanding of the potential of genotypes being used as parents and environments for which they are being developed, is necessary (Myers, 2004). The basic cause for differences between genotypes in yield stability is a wide

occurrence of genotype x environment (G x E) interactions. Genotype refers to the set of genes possessed by individuals that is important for the expression of traits under investigation. The environment is defined as all non-genetic factors that influence the expression of the trait and influence the growth and development of individuals. G x E interaction is a differential genotypic expression across environment (Basford and Cooper, 1998).

Accumulation of tolerance to a number of stresses is the key to wide adaptation and consequently selection in multiple environments is the best way to breed stable genotypes (Romagosa and Fox, 1993). When the effects of environmental differences are large, it may be expected that the interaction of G x E will also be large. As a result it is not only average performance that is important in genotype evaluation programmes, but also the magnitude of interactions (Gauch and Zobel, 1997). According to Romagosa and Fox (1993), G x E interaction reduces association between phenotypic and genotypic values of a genotype. This may cause promising selections from one environment to perform poorly in one and better in another environment, forcing plant breeders to examine genotypic adaptation.

An understanding of environmental and genotypic causes of G x E interaction is important at all stages of plant breeding, including ideotype design, parent selection based on traits and selection based on yield. This can be used to establish breeding objectives to identify ideal test conditions and to formulate recommendations for areas of optimal cultivar adaptation (Jackson *et al.,* 1998).

2.14.1 Crossover, non-crossover interaction and parametric and non parametric analysis

Some authors introduced G x E interactions as qualitative interactions (crossover) and quantitative interactions (non-crossover) (Baker, 1988). For non-crossover interactions, the true treatment differences vary in magnitude, but not in direction (the ranking of genotypes does not change from one environment to another), whereas for crossover interactions, the direction of true treatment differences varies. Crossover or qualitative interaction is important in agricultural production in contrast to non-crossover or quantitative interactions (Baker, 1988; Crossa, 1990).

For data sets with more than two genotypes and more than two environments, G x E interactions are commonly calculated by analysis of variance techniques leading to an estimated variance component. For a two-way table with n genotypes (rows) and m environments (columns), relationships between the amount of the variance component of G x E interactions and the rank changes of the genotype are of particular interest if a breeder is interested only in the existence of rank order differences over environments. Non-parametric analysis can be used for quantitative description of these relationships. Rank information (non-parametric) statistics for G x E interactions based on ranks provide a useful alternative to parametric approaches currently used, which are based on absolute data. Non-parametric statistics have advantages compared to parametric ones. (1) Reduction or even avoidance of the bias caused by outliers, (2) no assumptions are needed about the distribution of the analysed values and (3) homogeneity of variances and additivity (linearity) of effects are not necessary requirements. Statistics based on ranks and rank-orders are often easy to use and interpret (Crossa, 1990).

2.14.2 Statistical analysis of G x E interaction and stability concept

Different statistical methods have been proposed for estimation and partitioning of G x E interactions such as variance components, regression methods, multivariate analysis and cluster techniques (Hill, 1975; Cox, 1984; Freeman, 1985; Crossa, 1990). Analysis of G x E interactions is closely linked with the quantitative estimation of phenotypic stability of genotypes over environments (Kang, 1996). When significant G x E interactions are present, the effects of G x E are statistically non-additive, which means that differences between genotypes depend on the environment. Existing G x E interactions may, but must not necessarily, lead to different rank orders of genotypes in different environments (Crossa, 1990).

Numerous methods have been used in the search for an understanding of the cause of G x E interactions and can be categorised into two major categories (Van Eeuwijk, 1996). The first category involves factorial regression analysis of the G x E matrix (the yield matrix after the environment and genotype main effects are removed) against environmental factors, genotypic traits, or combinations thereof (Baril *et al.*, 1995). Fensham *et al.* (1998) using the first category when analysing 10 years of oat (*Avena sativa* L.) data, incorporated several genotypic covariates into a mixed model and indicated that plant type (plant height and kernel type) by environment interaction explained 50% of the observed variation. The second category involves correlation

or regression analysis, which relates the genotypic and environmental scores, derived from principal component analysis (PCA) of the G x E interaction matrix to genotypic and environmental covariates. This category is associated with the use of additive main effects and multiplicative interaction model (AMMI) (Gauch and Zobel, 1996).

2.14.3 Concepts of stability

A given crop species or cultivar can be successfully planted in an agro-climatic region depending on its adaptability and yield stability. Adaptability refers to good performance over a wide geographic region under variable climatic environmental conditions. Yield stability of yield is the ability of a genotype to avoid substantial fluctuations in yield over a range of environmental conditions (Heinrich *et al.,* 1983). Adaptability or stability of a cultivar often relates to physiological, morphological and phenological mechanisms. Grafius (1957) found that the tendency to stabilise yield depended on the temporal development of yield and yield components. According to Grafius (1957), yield is a product of several yield components and reductions in which one component may be compensated, to varying degrees, by an increase in other yield components.

Knowledge about the magnitude of G x E interactions is important in order to develop cultivars that combine high yield and stable performance over a wide range of environmental conditions. In the environment, individual genotypes may react to transient fluctuations in two different ways. Genotypes that are buffered against environmental variation and develop a similar phenotype over a range of environments possess a 'biological' or 'static' stability. This type is seldom a desired feature of crop cultivars, since no response to improved growing conditions would be expected. In contrast, 'agronomic' or 'dynamic' stability permits a predictable response to environments and stability according to the 'agronomic' concept, has no deviation from this response to environments (Becker and León, 1988). With quantitative traits, the majority of genotypes often react similarly to favourable or unfavourable environmental conditions. If a crossover (qualitative) type of G x E interaction (one that causes genotype rank changes) is present, the mean yield of genotypes selected via a method that combines yield and stability would usually be lower than that of genotypes selected on the basis of yield alone (Kang *et al*., 1991). Another way to clarify this is by examining the consequences to growers when researchers commit Type I error (rejecting the null hypothesis when it is true) and Type II

error (accepting the null hypothesis when it is false) relative to selection on the basis of yield alone and on the basis of yield and stability. Generally, Type II errors constitute the most serious risk for growers (Glaz and Dean, 1988). For the Kang *et al*. (1991) modified rank sum method, Type I and Type II error rates can be determined for stability components, but not for yield components.

Stability analysis techniques

Various methods of evaluating phenotypic stability have been suggested. Lin *et al*. (1986) investigated the statistical relationship between nine stability statistics and identified three concepts of stability:

Type 1: Stable genotype is characterised by a small variance across all environments. This type of stability is useful when the environments considered are not very diverse and is equivalent to the static concept of stability (Becker and León, 1988).

Type 2: A genotype is stable if its response to environments is parallel to the mean response of all genotypes in the trial. Type 2 stability is equivalent to the dynamic concept (Becker and León, 1988). Type 2 stability is primarily based on the interpretation of the regression coefficient in the joint linear regression model (Finlay and Wilkinson, 1963).

Type 3: A genotype is stable when residual mean squares from the regression model on the environmental index is small (smaller deviation from regression). Type 3 stability is also dynamic and the method of Eberhart and Russel (1966) can be used for its estimation.

Yield-stability statistic (YS) was proposed by Kang (1993), which is based on Shukla's (1972) stability variance statistic (SV). He partitioned G x E interaction into components, one corresponding to each genotype and termed each component as a stability variance. Lin *et al.* (1986) classified Shukla's stability variance as Type 2 stability, meaning that it was a relative measurement depending on genotypes included in a particular test. Kang *et al.* (1987) reported on the relationship between Shukla's stability variance and Wricke's ecovalence (Wi) (Wricke, 1962) and concluded that it was identical in ranking cultivars for stability (rank correlation coefficient = 1.00). This measure should be acceptable and useful to breeders and agronomists, as it provides contribution of each genotype in a test to total G x E interaction attributable to all

genotypes. Growers would prefer to use a high yielding cultivar that performs consistently from year to year. They may even be willing to sacrifice some yield if they are guaranteed, to some extent, that a cultivar would produce consistently from year to year (Kang *et al.,* 1991).

a. Cultivar performance

Cultivar performance measure *(P$_i$)* by Lin and Binns (1988) defined *P$_i$* of genotype *i* as the mean squares of distance between genotype *i* and the genotype with maximum response. The smaller the value of *P$_i$,* the smaller the distance to the genotype with maximum yield and the better the genotype. A pairwise G x E interaction mean square between the maximum and each genotype is also determined (Crossa, 1990).

b. Wricke's ecovalence (Wi)

Wricke (1962) proposed using the contribution of each genotype to the G x E interaction sum of squares as a stability measure and defined this concept or statistic as ecovalence *(Wi).* Ecovalence is simple to compute and is expressed as:

$$Wi = \Sigma_j \left[Y_{ij} - Y_i - Y_j + Y.... \right]^2$$

Where *Y$_{ij}$* is the mean performance of genotype *i* in the *j$_{th}$* environment and *Y$_i$* and *Y$_j$* are the genotype and environment mean deviations, respectively. *Y* is the overall mean. Genotypes with low *Wi* values (zero) have smaller deviations from the mean across environments and are more stable and possesses a high ecovalence (low value of *Wi* = high ecovalence). Becker and Léon (1988) illustrated ecovalence by using a numerical example by plotting yields of genotype *i* in various environments against the respective mean of environments.

c. Shukla's stability variance parameters ($\sigma^2{}_i$ and $d^2{}_i$)

Another statistic procedure is stability variance (SV) (Shukla, 1972). Based on the residuals from the additive model, the variance of cultivar *i* is defined as the variance of the cultivar across environments. For ranking purposes, stability variance (SV) is equivalent to ecovalence (Wi) (Wricke, 1962). Shukla's (1972) stability variance ($\sigma^2{}_i$) is considered as Type 2 stability (Lin *et al.,* 1986).

Shukla (1972) showed that genotypes could not be reliably described if the proportion of G x E interaction sums of squares due to heterogeneity among regression coefficients is small. Besides, there is lack of independence between performance and means of sites and between slopes and intercepts. Instead he proposed that G x E interaction sums of squares is partitioned into variance components (i^2) corresponding to each of the genotypes. On the basis of these variances, a genotype is stable if its stability variance (σ^2_i) is equal to environmental variance (σ^2_o), which means that $\sigma^2 = 0$. A relatively large value of σ^2_i indicate greater instability of genotype i. Since the stability variance is the difference between two sums of squares, it can be negative, but negative estimates of variances are not uncommon in variance components problems. Negative estimates of σ^2 may be taken as equal to zero as usual. Heterogeneity due to covariates such as environment and rainfall was removed from the G x E interaction sums of squares to derive another set of statistics, S^2i, for each genotype. Using his statistics, Shukla (1972) suggested that if a genotype becomes stable after applying the covariate, it can be suspected that the instability of the particular cultivar was introduced by the linear effects of that covariate. This approach is considered of practical importance because it identifies environmental factors that contribute to the heterogeneity in the G x E interaction.

d. Regression coefficient *(bi)* and deviation mean square *(S^2d_i)*.

Joint linear regression has been used as a method for analysing and interpreting the non-additive G x E interaction of two-way classification data. The G x E interaction is partitioned into a component due to the linear regression (b_i) of the ith genotype on environmental mean and deviation (d_{ij}):

$$(GE)_{ij} = b_iE_j + d_{ij}$$

The most widely used criteria for selecting high yield and stable performance are mean yield, regression response on site mean yield and deviations from regression (Eberhart and Russell, 1966; Freeman, 1973). The first measure is the slope *bi* from the regression of the yields of genotype *i* on an environmental index (Finlay and Wilkinson, 1963). Where *b* is equal to 1, it indicates that a cultivar reacts to a change in environment in the same way as the group mean. Finlay and Wilkinson (1963) proposed that regression coefficients approaching zero indicate stable performance. Regression coefficients approximating 1.0 indicate average stability. When this is associated with high mean yield, varieties have good general adaptability. When

associated with low mean yield, genotypes are poorly adapted to all environments. Regression values increasing above 1.0 describe genotypes with increasing sensitivity to environmental change (below average stability) and greater specificity of adaptability to high yielding environments. Regression coefficients decreasing below 1.0 provide a measure of greater resistance to environmental change (above average stability) and therefore, increasing specificity of adaptability to low yielding environments.

Stability analysis provides a method to characterise the response of a hybrid to varying environmental conditions. By far the most common technique in the commercial sector is based on the analysis developed by Eberhart and Russell (1966). In this analysis yields of a specific hybrid from many locations are regressed on the mean yield of all hybrids grown at the same set of locations. Maize breeders using this analysis tend to define a stable hybrid as one with high mean performance, a regression coefficient close to 1.0 and small deviations from regression (Eberhart and Russell, 1966).

Both Jensen and Cavalieri (1983) and Hallauer *et al*. (1988) noted that a large number of locations are necessary to obtain reliable estimates for the stability of a hybrid. Regression coefficients and cultivar mean yields over environments have been used to identify cultivars adapted to high or low environments and for general adaptability. Average phenotypic stability is shown by a regression coefficient of unity *(bi=1.0)*. A cultivar with *bi>1.0* reflects adaptability to high yielding environments and cultivars with *bi<1.0* imply adaptability to low yielding environments. Finlay and Wilkinson (1963) described the ideal cultivar as one possessing genetic potential in the highest yielding environment and with maximum phenotypic stability.

Eberhart and Russell (1966) proposed the use of two stability parameters to describe the performance of a variety over an array of environments. They proposed the regression of each cultivar on an environmental index as a function of the squared deviation. Breeders search for genotypes that show a stable high yield over years and locations. In general a genotype is considered stable when its performance across environments does not deviate from the average performance of a group of standard genotypes (Lin *et al*., 1986; Becker and León, 1988).

Eberhart and Russell (1966) proposed pooling the sum of squares for environments and G x E interactions and subdividing it into a linear effect between environments [with 1 degree of freedom (df)], a linear effect for G x E (with G-1 df) and a deviation from regression for each genotype (with E-2 df). The residual mean squares from the regression model across environments is used as an index of stability and a stable genotype is one in which the deviation from regression mean squares (S^2d_i) is small. Freeman (1973) and Hill (1975) reviewed the regression approach to study G x E interaction extensively. Pinthus (1973) proposed the use of the coefficient of determination (r^2_i), instead of deviation mean squares, to estimate stability of genotypes. However, (r^2_i) is strongly related to S^2d_i, but the application of (r^2_i) and b_i have the advantage that both statistics are independent of the units of measurement (Becker and Léon, 1988).

In discussing the most appropriate biometrical method, Becker and Léon (1988) noted that the regression approach is of little use if the regression coefficient (b_i) is included in the definition of "stability". For this reason, b_i is generally viewed by authors not as a measure of stability, but rather as additional information on the average response of a genotype to advantageous environmental conditions. Usually only a small part of G x E interactions can be explained by a heterogeneity of regression lines since in a normal series of trials most environments have near-average yield levels which lead to genotypes generally having b_i values close to 1. However, in situations where there is considerable variability from year to year or over environments, the regression approach may be useful in determining b_i as an indicator of response to variable environmental conditions. Becker and Léon (1988) cautioned that the choice of material will always influence an analysis of G x E interaction and has to be considered when discussing results.

Stability methods based on the G x E interaction sum of squares correspond to Type 2 stability, whereas the Eberhart and Russell (1966) method corresponds to Type 3 stability. Lin *et al.* (1986) pointed out that the parametric approach has the advantage of computational simplicity and only addresses certain aspects of stability, without giving an overall picture of the genotype's response. For example, a genotype may have Type 2 stability and simultaneously Type 3 instability (Crossa, 1990). Lin *et al.* (1986) recommended that scientists should firstly

define what type of stability they require, as well as understand the kind of environments that are to be used in the experiment, before considering what statistic to use.

Becker and Léon (1988) distinguished between two different concepts of stability, termed static stability and dynamic stability, respectively. Static stability is defined as a stable genotype possessing unchanged performance regardless of any variation of environments, implying that its variance among environments is zero. This is equivalent to the biological concept of stability and similar to Type 1 stability of Lin *et al.* (1986). Dynamic stability is defined as a genotype having a predictable response to environments and has no deviation from this response to environments. Becker (1981) termed this type of stability the agronomic concept to distinguish it from the biological or static concept. Becker and Léon (1988) stated that all stability procedures based on quantifying G x E interaction effects belong to the dynamic stability concept. Included are procedures partitioning G x E interaction, such as Wricke's (1962) ecovalence and Shukla's (1972), stability of variance, procedures using the regression approach such as proposed by Finlay and Wilkinson (1963) and Eberhart and Russell (1966) as well as non-parametric stability statistics.

Multivariate stability analysis techniques

Multivariate techniques are applied in stability analysis to provide further information on real multivariate response of genotypes to environments. According to Becker and Léon (1988) multivariate analysis has three main purposes: (1) to eliminate noise from the data pattern, (2) to summarise the data and (3) to reveal the structure in the data. Through multivariate analysis, genotypes with similar responses can be clustered, hypothesised and later tested and data can be easily summarised and analysed (Crossa, 1990). Multivariate analyses are appropriate for analysing two-way matrices of genotype and environments (Crossa, 1990). Becker and Léon (1988) defined the aim of the various multivariate classification methods as being to assign genotypes into qualitatively homogeneous stability subsets. Within subsets, no significant G x E interactions occurs, while differences among subsets are due to G x E interactions. However, Hohls (1995) reported drawbacks on multivariate analysis that, (1) numerous dissimilarity measures and clustering strategies exist and choosing between them can result in considerably different cluster groups and (2) non-existent structure could be forced onto the data.

Additive main effects and multiplicative interaction (AMMI)

The additive main effects and multiplicative interaction (AMMI) model combines analysis of variance (ANOVA) for genotype and environment main effects with principal component analysis (PCA) of the G x E interaction into a single model with additive and multiplicative parameters. It has proven useful for understanding complex G x E interactions (Kang, 1996). Results can be graphed in a very informative biplot that shows both main and interaction effects for both genotypes and environments. The AMMI model can partition data into a pattern rich model and discard noise-rich residual to gain accuracy. The AMMI model is used to separate estimated interaction components and adjust yield mean for the interaction. The advantage of using AMMI is that it accounts for a large proportion of variability in its first few components with subsequent dimensions accounting for diminishing percentage of pattern and increasing percentage of noise (Purchase, 1997). Gauch (1990) found AMMI useful for understanding complex interactions, gaining accuracy, improving selections and increasing experimental efficiency. The expectation-maximisation version, EM-AMMI, can impute missing data.

AMMI results can illuminate plant physiological processes that cause genotypes to interact with environments. It reveals the relative importance of various environmental factors or stresses. Most agricultural papers using AMMI provide a biological interpretation of AMMI genotype parameters. Analysis helped to identify morphological and physiological traits related to stress tolerance. A further advantage of these models is that they may be used for modelling and understanding interactions (Gauch, 1993).

The AMMI model is particularly useful in understanding G x E interactions and summarising patterns and relationships of genotypes and environments (Crossa, 1990). During the initial ANOVA the total variation is partitioned into three orthogonal sources, genotypes (G), environments (E) and G x E interactions. Romagosa and Fox (1993) observed that "in most yield trials, the proportion of sum of squares due to differences among sites ranged from 80-90% and variation due to G x E interactions was usually larger than genotypic variation". In AMMI analysis the interaction principle component analysis (IPCA) sum of squares alone is usually larger than for G. As genotypes and environments become more diverse, G x E tends to increase and may reach 40-60% of total variation. Normally the environmental main effect, which contributes up to 90% of the total variation, is fairly irrelevant, especially in selection

procedures. The AMMI model can produce graphs (biplots) that focus the data structure relevant to selection, in other words on the G and G x E sources (Romagosa and Fox, 1993).

In using the part of the AMMI analysis, PCA partitions G x E interaction into several orthogonal axes. Concern has been expressed by the number of axes the best AMMI model includes in its analysis, and how assessments and presentations of genetic stability can be made if too many axes are included. Gauch and Zobel (1996) stated that generally AMMI 1 and AMMI 2 models, with IPCA 1 and IPCA 2 respectively, are usually selected and that the graphical representation of axes, either as IPCA 1 or IPCA 2 against main effects, or IPCA 1 against IPCA 2, is not a problem and generally is informative. With AMMI 3 and higher models, IPCA 3 and higher axes are generally dominated by noise, have little or no predictive value and no biological interpretability and can thus be discarded.

Another primary use of the AMMI model is to improve the accuracy of yield estimates. Gains in accuracy of yield estimates are equivalent to increasing the number of replicates by a factor of two to five (Crossa and Cornelius, 1993). AMMI analysis offers a remarkably cost-effective means for improving research efficiency and increasing returns on investment (Gauch and Zobel, 1996). If, on the contrary, one is interested in genotypes that perform better, in one region, of which only a small sample of environments has been tested, one cannot predict interactions for each environment of that region. From a practical point of view, all interactions become unpredictable noise and it is reasonable to minimise G x E interactions, which is in accordance with the dynamic concept (Gauch and Zobel, 1996).

Like every other model, AMMI has its weakness. The nature of the residuals after fitting the additive main effects inevitably produces the appearance of multiplicative effects. Consequently the sum of squares for fitting the multiplicative term, which may be read directly from the latent root proportions of explained variation, will tend to be much larger than the expected value. Therefore, it is not possible to recommend a single model to be used at all times, because these models, depending on the type of data and research purposes, can be complimentary rather than being competitive.

Cotton genetic diversity characterisation, diallel and G x E studies

Based on literature, many studies have been done on cotton related to genetic diversity using morphological and molecular markers but few using cottonseed oil and fatty acids. Diallel studies have been done for different cotton genotypes for combining abilities, heritability and heterosis. Cotton genotypes G x E interaction analysis studies were done although further studies are still needed. Cotton in Tanzania is an important crop for farmers' income and country foreign currency. However, insufficient information is available on combining abilities for the available germplasm material and G x E studies are lacking. Cotton breeding programmes are currently based on morphological markers. With the limitations of morphological characterisation it is difficult to differentiate cotton varieties because of narrow genetic base. The need of molecular markers to study genetic diversity is necessary. This emphasizes the importance of this study for the Tanzanian cotton industry for improvement of cotton breeding programmes.

CHAPTER 3

EVALUATION OF OIL AND FATTY ACID COMPOSITION IN SEED OF VARIOUS COTTON VARIETIES*

3.1 INTRODUCTION

Cotton *(G. hirsutum)* in Tanzania has been important as a cash crop since 1935. The main product was lint for use in textiles for cloth manufacturing, but cottonseed oil, linters and husks were not extensively utilised. Cottonseed oil production in Tanzania has increased since the 1970's because of increased cotton production and increase in consumers' acceptance of cottonseed oil products. Cottonseed oil production in Tanzania has ranged from 213000 to 358000 metric tons for the last five years. Currently most vegetable oils are produced from cottonseed (60%), groundnut (20%), simsim (10%) sunflowers (6%) and a small amount from coconut and palm (COMESA, 2005).

Commercial cottonseed normally consists of about 10-15% linters, 35-40% hull and 50-55% kernels. Typical composition of cottonseed from various world sources indicates that seeds contain 5-12.8% moisture, 15.2-22.7% oil and 17.1-27.6% protein (Turner *et al.,* 1976; Lawhon *et al.,* 1977; Cherry *et al.,* 1981). A number of factors, including genetic and growing conditions, influence the composition of cottonseed (Wolf, 1988; Salunkhe *et al.,* 1992). Turner *et al*. (1976) and Lubbock and Raleigh (1996) reported the effect of environment on cottonseed oil content and some of the fatty acids.

Like other common vegetable oils, cottonseed oil consists predominantly of triglycerides. In addition, crude cottonseed oil contains a certain proportion of fatty acids in free form and minor amounts of a variety of other substances such as phospholipids, sterols and hydrocarbons. Some of these contribute considerably to the characteristic of the oil. It is furthermore a rich source of minerals, B vitamins and fat-soluble vitamins such as A, D and E (Murthi and Achaya, 1975).

Cottonseed oil is used in the manufacture of salad and cooking oils, shortenings and margarine. To a lesser extent, it is used in the packing of fish and cured meat. A small percentage of cottonseed oil is utilised to produce non-edible products like soap (Salunkhe *et al.,* 1992). Cottonseed oil contains 40-55% linoleic fatty acid, a principal and essential polyunsaturated fatty acid, 20-25% palmitic acid, a principle saturated fatty acid and 2-7% stearic acid. The main monosaturated fatty acid is oleic acid of 18-30%. Other fatty acids of small proportions are myristic and arachidic acids with about 0.5-2% cyclopropenoid acid (abbreviations and molecular formulas for individual fatty acids are presented in Table 3.1). Cottonseed oil has a 2:1 ratio of polyunsaturated to saturated fatty acid. Its fatty acids generally consists of about 70% unsaturated fatty acids [monounsaturated (oleic) and polyunsaturated (linoleic and linolenic)] and 27% saturated fatty acids (primarily palmitic, stearic, arachidic and myristic) (Scarth and McVetty, 1999; Smith and Rathore, 2001).

Lauric, myristic and palmitic acids are cholesterol raising saturated fatty acids but like other oils of plant origin, cottonseed oil is cholesterol free because of a high amount of polyunsaturated fatty acids. Pure cottonseed oil with a trans fatty acid level below 0.4% is virtually trans free (Jones and Kersey, 2002). However, cottonseed oil contains relatively high levels of palmitic acid (25%), which is considered to be a major contributor to increased levels of total blood cholesterol, especially LDL. Cottonseed oil contains more than 50% polyunsaturated fatty acids, mainly linoleic acid. Although linoleic acid is an essential fatty acid, it gives the oil a rapid drying quality due to the high percentage. This results in oxidative instability, which shortens the oil's shelf life and causes rancidity and off-flavours, though the high saturation of 27% contributes to stability. Stearic acid is being recognised increasingly as a neutral saturate in terms of disease because it is a saturated fatty acid but does not increase the LDL cholesterol level in blood. Oleic acid has the same LDL-lowering effect as linoleic acid, but it is not as susceptible to oxidation as linoleic acid (Liu *et al.,* 2002)

It is therefore desirable to develop new quality varieties with reduced contents of palmitic fatty acids at least to less than 5%. Correspondingly, the level of two other major fatty acids in cottonseed oil i.e. stearic acid and oleic acid should be raised. In assessing the prospects of achieving such reduction in palmitic acid and increase of stearic and oleic acids is by measuring the extent of natural variability for fatty acid composition in *G. hirsutum* varieties. The aim of

this study was, therefore, to assess the level and genetic basis of fatty acid variation between varieties in a Tanzanian collection of *G. hirsutum* germplasm with a view to identifying suitable parents for use in a hybridisation programme aimed at reducing the fatty acid content and increasing oil content.

Table 3.1   Common name, IUPAC/Systematic name, abbreviations and molecular formula for 14 fatty acids from cottonseed oil

Fatty acid	Common name	Abbre-viation	Complete formula	IUPACS/Systematic name	Molecular formular
Saturated fat acid	Myristic	C14:0	C14:0	Tetradecanoic	$CH_3(CH_2)_{12}COOH$
	Palmitic	C16:0	C16:0	Hexadecanoic	$CH_3(CH_2)_{14}COOH$
	Margaric	C17:0	C17:0	Heptadecanoic	$CH_3(CH_2)_{15}COOH$
	Stearic	C18:0	C18:0	Octadecanoic	$CH_3(CH_2)_{16}COOH$
	-	C19:0	C19:0	Nonadecanoic	$CH_3(CH_2)_{17}COOH$
	Arachidic	C20:0	C20:0	Eicosanoic	$CH_3(CH_2)_{18}COOH$
	Behenic	C22:0	C22:0	Docosanoic	$CH_3(CH_2)_{20}COOH$
	Lignoceric	C24:0	C24:0	Tetracosanoic	$CH_3(CH_2)_{22}COOH$
Monounsaturated	Palmitoleic	C16:1	C16:1c9 (n-7)	*cis*-9-Hexadecenoic	$CH_3(CH_2)_5CH= CH(CH_2)_7COOH$
	Oleic	C18:1	C18:1c9 (n-9)	*cis*-9-Octadecenoic	$CH_3(CH_2)_7CH= CH(CH_2)_7COOH$
	Eicosenoic	C20:1	C20:1c11 (n-11)	*cis*-11-Eicosenoic	$CH_3(CH_2)_9CH= CH(CH_2)_7COOH$
	Nervonic	C24:1	C24:1c15 (n-9)	*cis*-15-Tetracosanoic	$CH_3(CH_2)_7CH= CH(CH_2)_{13}COOH$
Polyunsaturated	Linoleic	C18:2	C18:2c9, 12 (n-6)	*cis*-9,12-Octadecadienoic	$CH_3(CH_2)_4(CH= CHCH_2)_2CH(CH_2)_6COOH$
	Linolenic	C18:3	C18:3c9,12,15 (n-3)	*cis*-9,12,15-Octadecatrienoic	$CH_3(CH_2)(CH= CHCH_2)_4CH(CH_2)_6COOH$

3.2 MATERIALS AND METHODS

3.2.1 Plant material

The 30 cottonseed varieties used in this study were obtained from Ukiriguru Research Institute in Tanzania. These were collections from different areas in the world like Zambia, Malawi, France, Mali, California, West Africa, USA and South Africa. Seeds for six varieties (UK91, Acala SJ$_2$, NTA 93-21, Okra leaf, Delcot 344 and MZ561) (numbers 25 to 30) (Table 3.2) were not enough and were multiplied again during the 2002/2003 season at the University of Free State, South Africa (UFS). Two duplicates were prepared for each variety.

Table 3.2 Names of 30 cotton varieties used in this study

CODE	Name	CODE	Name	CODE	Name
1	McNair 235	11	CIM 70	21	Dixie King
2	Frego bract	12	Cyto 12/74	22	SG 125
3	Reba W296	13	UK82	23	Coker 315
4	High gossypol	14	NTA 93-15	24	F-135
5	DP Acala90	15	Irma 1243	25	NTA 93-21
6	Aubarn 56	16	Guazuncho	26	MZ561
7	NTA 88-6	17	DP 4049	27	Okra leaf
8	BJA 592	18	HC-B4-75	28	Acala SJ$_2$
9	Stoneville 506	19	Nectariless	29	Delcot 344
10	Ezra	20	Des 119	30	UK91

3.2.2 Lipid extraction

The lipid extraction procedure described by Folch *et al.* (1957) was used. Cottonseeds were ground to a fine powder using a coffee grinder. A total of 0.5 g ground seed was added to 30 ml 2:1 (v/v) chloroform: methanol containing 10 mg/ml 0.001% Butylated hydroxy toluene and left in a refrigerator at 4^0C overnight. Samples were filtered through preweighed 18.5 cm Whatman No. 1 filter papers into a separation funnel. Sample flasks were washed three times with the 2:1 (v/v) chloroform: methanol mixture.

A total of 16.5 ml distilled water was added to each sample, shaken thoroughly and allowed to separate for one hour. The lower phase was drained into round bottom flask and 50 ml of the lower phase [Chloroform: methanol: water (86:14:1 v/v)] was added to each separation funnel and shaken thoroughly. After 15 min of separation, the lower phase was drained into the same round bottom flask. This step was repeated twice after which the upper phase was discarded. The contents of the flask were evaporated under vacuum in a rotary evaporator at 60^0C for 20 min at 15 psi. The remaining water was removed by addition of 30 ml methanol and evaporating each round bottom flask at 60^0C at 15 psi. The contents of each flask were washed six times with 5 ml portions of diethyl ether into a pre-weighed polytop. Diethyl ether was removed from each polytop by evaporation under a stream of nitrogen on a heating block at 60^0C for 20 min. Filter papers and polytops were dried at 50^0C overnight in a vacuum oven, removed and weighed the next morning. The oil solvent mixture was dried over anhydrous sodium sulphate,

the solvent evaporated under nitrogen and the oil percentage determined by weighing. The following formulas were used:

Fatty mass (g) = Polytop mass after extraction – Polytop mass before extraction

% Fat= <u>Fat mass (g) x 100</u>

Sample mass (g)

% FFDM = <u>FP mass after extraction (g) – FP before extraction (g) x100</u>

Sample mass (g)

% Moisture = 100 – (% Fat + % FFDM)

FFDM = fat free dry matter, FP = filter paper

3.2.3 Methylation

A modified procedure described by Slover and Lanza (1979) was used. A total of $\pm$ 10.0 mg extracted lipid was transferred to another polytop for methylation. This lipid was washed with 6x1 ml hexane into the test tube with a Teflon-lined cap. Hexane was removed by nitrogen evaporation without application of heat. Methanolic 0.5N NaOH (1 ml) was added, the tube capped and heated in a boiling water bath for 15 min. After the tube had cooled, 2 ml of BF_3/CH_3OH (14%) was added, the tube was recapped and heated in the boiling water bath for an additional 15 min. The tube was cooled and 1 ml hexane and 2 ml saturated aqueous NaCl were added. The tube was shaken vigorously for 1 min and allowed to stand for 10 min until the phases separated. The upper 70% hexane layer was transferred with a pasteur pipette to a 45x11 mm vial containing a 1 mm layer of anhydrous Na_2SO_4.

The vial was capped, shaken and allowed to stand for 20 min to remove traces of water. Then 100 µl hexane from each vial was transferred to a clean labelled auto-sampler vial and 900 µl hexane was added to each vial and stored below freezing point for gas chromatography analysis (GC).

3.2.4 Determination of fatty acid composition using gas chromatography

After methylation, fatty acid composition for all accessions were quantified using Varian GX 3400 flame ionisation gas chromatography, with a fused silica capillary, column and Chrompack CPSIL 88 (100 m length, 0.25 μm ID, 0.2 μm film thickness). Column temperature was 40-230^{0}C (hold 2 min; 4^{0}C/min; hold 10 min). Fatty acid methyl esters in hexane (1 μl) were injected into the column using a Varian 8200 CX Auto-sampler with a split ratio of 100:1. The injection port and detector were both maintained at 250^{0}C. Hydrogen was used as a carrier gas at 45 psi and nitrogen was used as a makeup gas. Chromatograms were recorded with Varian Star Chromatography Software version 4.0. Identification of sample fatty acids was made by comparing the relative retention times of fatty acid methyl ester peaks from samples with those of standards obtained from SIGMA (cat. No. 189-19).

3.2.5 Fatty acid ratio

In triacylglycerols biosynthesis, palmitate has different fates and beta-ketoacyl-ACP synthase II (KAS II) is the key enzyme in this pathway (Harwood, 1996). KAS II is responsible for the condensation of C16:0-ACP with malonyl-ACP to stearoyl-ACP. This process determines the C16/C18 fatty acid ratio of seed oil (Möllers and Schierholt, 2002). Therefore the C16/C18 was calculated as:

C16/C18 = (C16: 0 + C16: 1 + C16: 2)/(C18: 0 + C18: 1 + C18: 2 + C18: 3).

Other ratios determined were Elongation Ratio (ER). This is used to estimate the relative weight of the elongation pathway from oleic acid (C18:1) to eicosenoic (C20:1) and erucic acid (C22:1). Desaturation ratio (DR) is used to estimate the relative weight of the desaturation pathway from oleic acid (C18:1) to linoleic (C18:2) and linolenic (C18:3) within the overall fatty acid biosynthesis system (Pleines and Friedt, 1988; Velasco and Becker, 1998). Oleic desaturation ratio (ODR) and linoleic desaturation ratio (LDR) estimate the within desaturation pathway, the efficiency of the desaturation from oleic to linoleic (ODR) and from linoleic to linolenic (LDR).

$$ER \quad = \quad \frac{\%C20{:}1 + \%C22{:}1}{\%C20{:}1 + \%C22{:}1 + \%C18{:}1 + \%C18{:}2 + \%C18{:}3}$$

$$DR \quad = \quad \frac{\%C18{:}2 + \%C18{:}3}{\%C20{:}1 + \%C22{:}1 + \%C18{:}1 + \%C18{:}2 + \%C18{:}3}$$

$$ODR \quad = \quad \frac{\%C18{:}2 + \%C18{:}3}{\%C18{:}1 + \%C18{:}2 + \%C18{:}3}$$

$$LDR \quad = \quad \frac{\%C18{:}3}{\%C18{:}2 + \%C18{:}3}$$

3.2.6 Statistical analysis

The statistical package Agrobase (2000) was used to do simple descriptive analysis and correlation coefficients.

3.2.7 Genetic distances and dendrogram

Data for oil content, palmitic fatty acid, stearic fatty acid, oleic fatty acid and linoleic fatty acid were used for genetic distances and cluster analysis because they are important in health requirements and were present in each variety tested. Data were first grouped into three classes as low, medium and high using the following formula:

Highest-lowest/3 = x: Lowest $\le$ lowest + x; medium $\le$ lowest +2x and > lowest + x; highest > lowest + 2x.

For oil: the lowest value was $\le$ 17.38, medium >17.38 but $\le$ 20.77 and high was > 20.77; palmitic fatty acid: lowest value was $\le$ 21.19, medium >21.19 but $\le$ 23.15 and high was > 23.15, stearic fatty acid: lowest value was $\le$ 2.31%, medium >2.31 but $\le$ 2.55 and high was > 2.55%; oleic fatty acid: lowest value was $\le$ 15.84%, medium > 15.84 but $\le$ 17.17% and highest was > 17.17%; linoleic fatty acid: lowest value was $\le$ 55.62, medium > 55.62 but $\le$ 58.91% and highest was > 58.91%. Data were converted into a binary data matrix where presence was scored as 1 and absence as 0.

Genetic distances between pairs of varieties and cluster analysis were obtained using Number Cruncher Statistical system, NCSS 2000 (Hintze, 2000). Distance matrices for all pairs of varieties were constructed from oil and fatty acids data matrix using the Euclidean distance method (Kaufman and Rousseeuw, 1990).

$$GD = \sqrt{\sum[(X_i - Y_i)^2 / N]}$$

Where, GD is the genetic distance between individual X and individual Y; $i = 1$ to N; N is the total number of bands; Xi and Yi are the ith band scores (1 or 0) for individuals X and Y.

Cluster analysis was performed using the genetic distance matrices generated by the Eucledian distance method to reveal patterns of genetic relationships among varieties. Dendrograms were constructed using the unweighted pair group method of arithmetic averages (UPGMA) (Hintze, 2000), which minimises within cluster variance, to infer relationships among varieties. The co-phenetic correlation (Kaufman and Rousseeuw, 1990; Hintze, 2000) for each dendrogram was computed as a measure of 'goodness of fit' for each dendrogram. The two delta goodness of fit statistics, delta (0.5) and delta (0.1) were calculated to determine which clustering configuration fits the data best.

3.3 RESULTS

3.3.1 Oil and fatty acids content

Oil content and fatty acid profile results (example Figure 3.1) in *G. hirsutum* indicated variation between the 30 varieties of the cotton germplasm (Table 3.3). Fourteen fatty acids were identified in percentages of the total fatty acid of the seed oil, but not all varieties contained all fatty acids. The predominant fatty acids found in all 30 varieties were linoleic (essential) fatty acid (C18:2), palmitic (principle) fatty acid (C16:0), oleic (C18:1), stearic (C18:0), myristic (C14:0), palmitoleic (C16:1), arachidic (C20:0) and behenic fatty acid (C22:0) (Tables 3.1 and 3.3). Lignoceric (C24:0) was found in 24 varieties, margaric (C17:0) in four varieties, nervonic (C24:1) in nine varieties, nonadecanoic (C19:0) in one variety and eicosenoic (C20:1) in two varieties (Table 3.3). Fatty acid variations among varieties were observed mainly for linoleic fatty acid (51.97-62.19%), myristic fatty acid (0.36-0.93%), linolenic fatty acid (0.00-0.20%) and palmitic fatty acid (19.23-25.04%) (Tables 3.3 and 3.4).

Table 3.3 Means of fatty acids and oil content in 30 varieties from Tanzania

No	Name	20:0	20:1	22:0	24:0	24:1	18:0	18:1	18:2	18:3	14:0	16:0	16:1	17:0	19:0	Oil	SFA	UFA	PUFA	UN/SA	DR and ODR	LDR	C16/C18
1	McNair 235	0.28	0.00	0.16	0.12	0.13	2.09	15.25	54.41	0.12	0.93	25.04	0.41	0.00	0.00	20.71	28.63	70.33	54.53	2.46	0.78	0.002	0.35
2	Frego bract	0.33	0.00	0.18	0.09	0.16	2.67	16.90	52.43	0.10	0.74	24.90	0.59	0.08	0.00	23.25	28.98	70.17	52.52	2.42	0.76	0.002	0.35
3	Reba W926	0.27	0.00	0.15	0.08	0.12	2.63	16.96	53.15	0.17	0.65	25.10	0.48	0.00	0.00	22.29	28.89	70.89	53.32	2.45	0.76	0.003	0.35
4	High gossypol	0.27	0.00	0.12	0.13	0.00	2.45	17.17	52.33	0.12	0.75	24.93	0.55	0.08	0.00	20.73	28.74	70.17	52.45	2.44	0.75	0.002	0.35
5	DP Acala90	0.29	0.00	0.16	0.12	0.00	0.22	16.64	55.13	0.11	0.66	23.40	0.42	0.00	0.00	21.02	26.86	72.30	55.24	2.69	0.77	0.002	0.32
6	Aubarn 56	0.27	0.00	0.15	0.14	0.00	2.67	17.88	52.33	0.11	0.64	24.49	0.43	0.00	0.00	24.61	28.37	70.74	52.44	2.49	0.75	0.002	0.34
7	NTA 88-6	0.22	0.00	0.15	0.00	0.00	2.40	16.39	57.24	0.00	0.64	21.77	0.52	0.00	0.00	18.65	25.17	74.16	57.24	2.95	0.78	0.002	0.29
8	BJA 592	0.25	0.00	0.15	0.11	0.00	2.42	16.92	53.04	0.12	0.76	25.00	0.51	0.08	0.00	21.05	28.77	70.60	53.17	2.45	0.76	0.002	0.35
9	Stoneville 506	0.29	0.00	0.16	0.11	0.00	2.40	17.75	56.44	0.13	0.62	20.62	0.56	0.00	0.00	19.25	24.20	74.87	56.56	3.09	0.76	0.002	0.28
10	Ezra	0.27	0.00	0.18	0.23	0.00	2.69	17.32	53.71	0.13	0.68	23.22	0.49	0.00	0.47	19.19	27.73	71.64	53.84	2.58	0.76	0.002	0.32
11	CIM 70	0.31	0.00	0.15	0.11	0.11	2.79	18.28	54.41	0.11	0.69	23.60	0.45	0.10	0.00	19.04	27.46	72.13	54.52	2.63	0.76	0.002	0.32
12	Cyto 12/74	0.27	0.00	0.17	0.16	0.00	2.73	18.51	53.01	0.16	0.62	23.78	0.54	0.00	0.00	19.77	27.73	72.21	53.17	2.60	0.74	0.003	0.33
13	UK82	0.29	0.00	0.16	0.15	0.00	2.50	17.04	52.93	0.15	0.65	23.91	0.49	0.00	0.00	20.97	27.95	71.85	53.07	2.57	0.74	0.002	0.33
14	NTA 93-15	0.26	0.09	0.18	0.16	0.13	2.72	16.99	53.04	0.16	0.87	24.28	0.55	0.00	0.00	20.57	28.48	70.95	53.20	2.49	0.76	0.002	0.34
15	Irma 1243	0.25	0.00	0.16	0.11	0.17	2.67	18.26	51.97	0.10	0.74	24.12	0.51	0.00	0.00	18.88	28.06	71.01	52.07	2.53	0.74	0.001	0.34
16	Guazuncho	0.31	0.00	0.18	0.16	0.00	2.30	16.27	53.98	0.16	0.86	24.67	0.49	0.00	0.00	18.27	28.47	70.91	54.15	2.49	0.77	0.003	0.35
17	DP 4049	0.27	0.00	0.16	0.00	0.11	2.47	16.59	55.03	0.11	0.65	23.69	0.46	0.00	0.00	16.37	27.23	72.29	55.14	2.02	0.77	0.002	0.33
18	HC-B4-75	0.27	0.00	0.11	0.00	0.00	2.60	17.83	53.05	0.14	0.71	24.33	0.50	0.00	0.00	19.82	28.01	71.38	53.04	2.55	0.75	0.002	0.34
19	Nectarless	0.32	0.00	0.18	0.11	0.00	2.61	16.68	53.71	0.00	0.78	24.43	0.52	0.00	0.00	19.73	28.41	70.54	53.71	2.48	0.76	0.002	0.34
20	Des 119	0.25	0.00	0.14	0.11	0.11	2.44	16.88	53.65	0.17	0.82	24.20	0.48	0.00	0.00	13.98	27.97	71.29	53.82	2.55	0.75	0.003	0.34
21	Dixie King	0.28	0.00	0.14	0.10	0.00	2.52	16.38	54.49	0.11	0.78	23.93	0.51	0.00	0.00	16.68	27.76	71.48	54.60	2.57	0.77	0.002	0.33
22	SG 125	0.27	0.00	0.13	0.11	0.00	2.61	16.84	55.57	0.17	0.64	22.78	0.43	0.00	0.00	21.88	26.54	73.01	55.75	2.75	0.77	0.003	0.31
23	Coker 315	0.32	0.00	0.16	0.13	0.14	2.52	15.90	55.90	0.12	0.66	24.22	0.42	0.00	0.00	16.50	28.00	72.48	56.02	2.59	0.78	0.002	0.33
24	F 135	0.29	0.00	0.17	0.12	0.00	2.43	16.74	53.86	0.00	0.76	24.75	0.49	0.00	0.00	19.73	28.51	71.10	53.86	2.40	0.76	0.000	0.35
25	NTA 93-21	0.23	0.00	0.14	0.00	0.00	2.23	16.22	59.02	0.09	0.46	20.75	0.33	0.00	0.00	19.19	23.80	75.65	59.10	3.18	0.78	0.001	0.27
26	MZ561	0.19	0.00	0.11	0.00	0.00	2.28	17.19	58.88	0.14	0.40	20.06	0.28	0.00	0.00	20.31	20.30	77.23	59.02	3.31	0.77	0.002	0.26
27	Okra leaf	0.20	0.00	0.14	0.11	0.00	2.26	16.23	60.23	0.14	0.36	19.52	0.28	0.00	0.00	19.32	22.59	76.92	60.40	3.41	0.79	0.002	0.25
28	Acala SJ$_2$	0.21	0.00	0.11	0.00	0.00	2.07	17.08	58.82	0.16	0.46	19.96	0.29	0.00	0.00	23.25	22.80	76.35	58.98	3.35	0.78	0.002	0.26
29	Delcot 344	0.20	0.00	0.14	0.09	0.00	2.08	14.51	62.19	0.20	0.39	19.43	0.28	0.00	0.00	24.15	22.34	77.17	62.39	3.45	0.81	0.003	0.25
30	UK91	0.19	0.11	0.00	0.14	0.00	2.16	17.87	59.33	0.14	0.44	19.23	0.33	0.00	0.00	23.50	22.15	77.65	59.47	3.51	0.77	0.002	0.25

SFA=Saturated fatty acid, UFA = Unsaturated fatty acid, PUFA = polyunsaturated fatty acid, UN/SA = Unsaturated/saturated ratio, DR= Desaturation ratio, ODR = Oleic desaturation ratio, LDR = Linoleic

desaturation ratio, C16/18 = C16/C18 ratio

```
milliVolts    -1.73                         19.25                    +SR
                                                                <WI=2SR

                                                  <CS=0.50<ATT=16*

C14:0  ─┤─── 33.689

C16:0  ─┤              37.575
C18:1c9 ─┤═── 38.789

C18:0  ─┤──── 41.081
C18:1c9 ─┤═══ 42.241          42.101
C18:2c9,12* ─┤═ 43.543                      43.651
                44.299
:18:3c9,12,1 ─┤═ 45.376

C22:0  ─┤── 47.278


C24:1c15 ─┤═ 51.201                         <WI=4.0
```
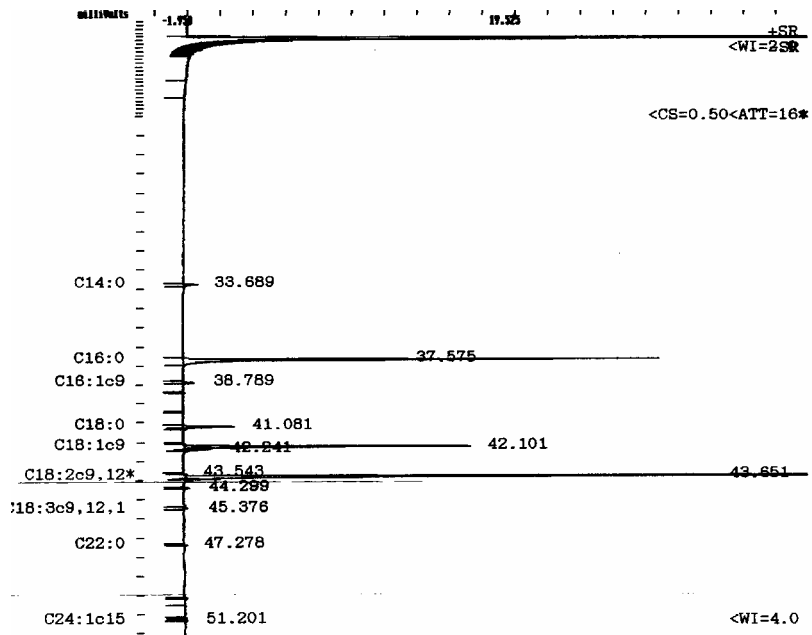
Figure 3.1   Graph developed by gas chromatography for fatty acid analysis (example of variety Reba W296)

Table 3.4 Descriptive statistical data for fatty acids and oil content showing the maximum, minimum, average and range

Name	20:0	22:0	24:0	24:1	18:0	18:1	18:2	18:3	16:0	16:1	17:0	Oil
Maximum	0.33	0.18	0.23	0.17	2.79	18.51	62.19	0.20	25.04	0.59	0.10	24.61
Minimum	0.19	0.00	0.00	0.00	2.07	14.51	51.97	0.00	19.23	0.28	0.00	13.98
Average	0.26	0.15	0.10	0.04	2.39	16.92	55.12	0.12	23.14	0.45	0.01	20.23
Range	0.14	0.29	0.23	0.17	0.75	4.30	11.43	0.21	6.35	0.55	0.10	11.13
STD Dev.	0.04	0.03	0.06	0.06	0.21	0.91	2.78	0.05	1.98	0.11	0.03	2.42
Variance	0.00	0.00	0.00	0.00	0.04	0.81	7.58	0.00	3.84	0.01	0.00	5.77
CV	14	22	57	161	8	5	5	5	8	24	257	12

Myristic acid, palmitic acid, stearic acid, arachidic acid and behenic fatty acids were the predominant saturated fatty acids observed. Other saturated fatty acids like margaric, nonadecanoic and lignoceric fatty acids were observed but not in all varieties. Monounsaturated fatty acids were palmitoleic, oleic and in few varieties eicosenoic and nervonic fatty acids were observed. Polyunsaturated fatty acids were linoleic and linolenic (Table 3.3). Total saturated fatty acid was about 26.32%, monounsaturated fatty acid was 17.39% and polyunsaturated was about 52.61% of the total fatty acid. This gives a ratio of polyunsaturated: saturated fatty acids of 1.99:1.00 equivalent to about 2:1. The total unsaturated fatty acid was about 70%, giving a ratio of unsaturated: saturated of about 3.1.

The environment had a large effect on some of the fatty acids as seen in the variation observed in some fatty acids between the 24 varieties (variety number 1 to 24) that were not multiplied at the UFS and six accessions (variety number 25 to 30) that were multiplied at the UFS (Table 3.3). Plants grown at the UFS were affected by winter.

The low temperature stopped growth at seed maturation stage and lead to premature seeds that increased the amount of some fatty acids like linoleic (for example in variety 1 to 24 the value was from 51.97-57.24% while for accessions 25 to 30 the value ranged from 58.82-62.19%). This was also observed for linolenic acid and oil content. Other fatty acids were reduced, for example palmitic acid, palmitoleic, stearic, myristic and arachidic (for example in the varieties 1 to 24 the value for myristic fatty acid ranged from 0.64-0.82% while for varieties 25 to 30 the value varied from 0.39-0.48%).

Delcot 344 had the highest linoleic acid (62.19%) and lowest oleic acid (14.51%) among varieties multiplied at the UFS, while NTA 88-6 had low oleic acid (16.39%) and high linoleic acid (57.24%) between varieties not multiplied at the UFS (Table 3.3). McNair 235 (25.04%), Reba W296 (25.10%), High gossypol (24.93%), BJA 592 (25.00%) and Nectarless (24.43%) had high palmitic fatty acid (principle fatty acid).

Margaric fatty acid was observed only in Frego bract (0.08%), High gossypol (0.08%), BJA 592 (0.08%) and CIM 70 (0.10%). Eicosenoic was observed in NTA 93-15 (0.09%) and UK91

(0.11%). Nonadecanoic (0.47%) was observed only in Ezra. Lignoceric fatty acid was not found in NTA 88-6, DP 4049, HC-B4-75, NTA 93-21, MZ561 and Acala SJ$_2$. Nervonic fatty acid was observed in McNair 235 (0.13%), Frego bract (0.16%), Reba W296 (0.12%), CIM 70 (0.11%), NTA 93-15 (0.13%), Irma 1243 (0.17%) DP 4049 (0.11), Des 119 (0.12%) and Coker 315 (0.14%). Alpha linolenic was not observed in NTA 88-6, Nectarless and F 135 (Table 3.3).

3.3.2 Fatty acid ratios

Ratios for fatty acids are shown in Table 3.3. Eicosenoic fatty acid (C20:1) was zero in almost all varieties except NTA 93-15 and UK91 and erucic fatty acid (C22:1) was not found in any of the varieties leading to zero ER (refer the ER formular section 3.2.5). The desaturation ratio (DR) showed that Delcot 344 had the highest value (0.81) and Cyto 12/74, UK82 and Irma 1243 had the lowest value (0.74) each. The average for all 30 varieties was 0.77. The oleic desaturation ratio (ODR) was similar to the DR ratio. The linoleic desaturation ratio (LDR) averaged 0.002 (Table 3.3).

3.3.3 Correlations

Correlations of oil content and different fatty acid composition as well as correlation between fatty acids are given in Table 3.5. Oil content was significantly negatively correlated with DR and ODR while correlation with C16/C18 ratio was significantly positive. A positive correlation between oil content and oleic, linoleic and linolenic fatty acids was observed.

Palmitic fatty acid was significantly positively correlated with arachidic, behenic, lignoceric, nervonic, stearic, linoleic, myristic fatty acids, palmitoleic, SFA and C16/C18 ratio. A significant negative correlation was observed between palmitic with unsaturated fatty acid (UFA), polyunsaturated fatty acid (PUFA), unsaturated:saturated ratio, DR, ODR and LDR. Linoleic showed a significant negative correlation with behenic, lignoceric, stearic, oleic, myristic, palmitoleic, SFA and C16/C18 ratio. A positive and significant correlation of linoleic was observed with palmitic, arachidic, UFA, PUFA, unsaturated: saturated ratio, DR and ODR (Table 3.5).

Table 3.5 Correlation results between oil, fatty acids and ratios for cotton varieties

	OIL	C20:0	C20:1	C22:0	C24:0	C24:1	C18:0	C18:1	C18:2	C18:3	C14:0	C16:0	C16:1	C17:0	SFA	UFA	PUFA	UN/SA	DR and ODR	LDR
C20:0	-0.23																			
C20:1	0.04	-0.01																		
C22:0	-0.31	0.65 *	0.17																	
C24:0	-0.06	0.52 **	0.20	0.60 *																
C24:1	-0.24	0.33	0.27	0.29	0.08															
C18:0	-0.13	0.52 **	0.24	0.42 *	0.45 *	0.25														
C18:1	0.08	0.06	0.02	-0.17	0.08	-0.09	0.58 **													
C18:2	0.16	0.71 **	0.14	-0.49 *	-0.50 *	-0.35	-0.75 **	-0.47												
C18:3	0.18	-0.18	0.14	0.24	0.11	-0.07	-0.15	-0.08	0.21											
C14:0	-0.34	0.71 **	0.27	0.55 *	0.46 **	0.41 *	0.41 *	0.01	-0.81 **	-0.21										
C16:0	-0.22	0.77 **	0.11	0.57 *	0.47 **	0.41 *	0.60 **	0.13	0.92 **	-0.19	0.87 **									
C16:1	-0.26	0.68 **	0.20	0.52 *	0.41 *	0.25	0.71 **	0.38 *	-0.83 **	-0.30	0.76 **	0.74 **								
C17:0	-0.13	0.27	-0.07	0.04	0.08	0.19	0.11	0.05	-0.29	0.03	0.20	0.29	0.29							
SFA	-0.23	0.79 **	0.14	0.60 *	0.53 **	0.41 **	0.66 **	0.18	-0.94 **	-0.19	0.88 **	0.99 **	0.78 **	0.28						
UFA	0.21	0.77 **	-0.13	0.61 *	-0.52 *	-0.39 *	-0.63 **	-0.17	0.94 **	0.23	-0.90 **	-0.99 **	-0.79 **	-0.30	0.99 **					
PUFA	0.17	-0.71 **	-0.14	-0.49 *	-0.49 *	-0.34	-0.75 **	0.47 **	0.10 **	0.22	-0.81 **	-0.92 *	-0.83 **	-0.29	0.94 **	0.95				
UN/SA	0.33	-0.75 **	-0.11	-0.60 *	-0.40 *	-0.43 *	-0.62 **	-0.15	0.89 **	0.23	-0.83 **	-0.95 *	-0.74**	-0.23	0.95 **	0.94	0.89 **			
DR=ODR	-0.78 **	-0.87 **	0.85 **	0.17	-0.46 *	-0.14	-0.78 **	-0.87 **	0.85 **	0.17	-0.46 **	-0.59 *	-0.69 **	-0.19	0.63	0.63	0.85 **	0.58 *		
LDR	-0.06	-0.02	0.09	0.80 **	-0.11	-0.17	-0.06	-0.02	0.09	0.80 **	-0.11	-0.11	-0.21	0.02	-0.10 *	0.12	0.11	0.13	0.07	
C16/18	0.60 **	0.14	-0.93*	0.19	0.89 **	0.42 *	0.60 **	0.14	-0.93 **	-0.19	0.89 **	0.10 **	0.76* *	0.31	1.00 **	-0.99	-0.93 **	-0.90 *	-0.61 **	-0.1

SFA=Saturated fatty acid, UFA = Unsaturated fatty acid, PUFA = polyunsaturated fatty acid, UN/SA = Unsaturated/saturated ratio, DR= Desaturation ratio, ODR = Oleic desaturation ratio, LDR = Linoleic desaturation ratio, C16/18 = C16/C18 ratio, * $p \leq 0.05$, ** $p \leq 0.01$

Significant positive correlations were observed between stearic acid with arachidic, behenic, lignoceric, oleic, myristic, palmitic, palmitoleic, SFA and C16/C18. A significant negative correlation was observed with linoleic, UFA, PUFA, DR and ODR. It was observed that the increase of C18:2 and C18:3 resulted in a significant decrease of C14:0, C16:1, C18:0 and C18:1. Palmitic was positively correlated with stearic, oleic and linoleic while stearic was positively correlated with oleic and negatively correlated with linoleic. The ODR was significantly negatively correlated with oil content, arachidic, lignoceric, stearic, oleic, myristic, palmitic and palmitoleic fatty acids. A significant positive correlation was observed between ODR and eicosenoic, linoleic, PUFA and unsaturated: saturated ratio.

3.3.4 Genetic distances

The genetic distances matrix for 30 varieties are shown in Table 3.6 and distances ranged from 0.00-0.92. Genetic distance for some varieties was 0.00, indicating that there was no difference for these varieties, for example CIM 70 and Cyto 12/74 and Frego bract and Reba W296, in relation to oil and the analysed fatty acids (Tables 3.3 and 3.6). The highest genetic distance of 0.92 indicated that there was a high genetic diversity between the varieties for example Dixie King with MZ561; Dixie King with Delcot 344, Aubarn 56 with Coker 315 and Dixie King with UK91. Varieties multiplied at the UFS (Okra leaf, UK91, Delcot 344, MZ561, Acala SJ$_2$ and NTA 93-21) had higher values for some fatty acids and some fatty acids were reduced because of the environmental effect that led to increased genetic distances between them and other varieties which were not multiplied at the UFS (Tables 3.3 and 3.6). The genetic distances between varieties that were not multiplied at the UFS ranged from 0.00-0.80, while genetic distances observed between varieties multiplied at the UFS ranged from 0.72-0.92 (Table 3.6).

3.3.5 Cluster analysis

Cluster analysis (dendrogram) for oil and four fatty acids (palmitic, oleic, stearic and linoleic) is shown in Figure 3.2. Based on these characteristics, the 30 varieties were grouped into different clusters. Two main groups (*I* and *II*) were revealed. The 24 varieties which were not multiplied at the UFS clustered together in group *I*, indicating that they were more closely related as compared to main group *II*. Main group *I* was further divided into two clusters *A* and *B*.

Table 3.6. Genetic distances for 30 varieties for oil content and fatty acids (oleic, palmitic, stearic and linoleic)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
2	0.72																												
3	0.72	0.00																											
4	0.60	0.60	0.60																										
5	0.60	0.44	0.44	0.60																									
6	0.72	0.44	0.44	0.72	0.60																								
7	0.82	0.82	0.82	0.60	0.82	0.89																							
8	0.72	0.44	0.44	0.44	0.44	0.60	0.72																						
9	0.82	0.89	0.89	0.72	0.89	0.82	0.60	0.82																					
10	0.60	0.60	0.60	0.60	0.72	0.44	0.82	0.72	0.72																				
11	0.60	0.60	0.60	0.60	0.72	0.44	0.82	0.72	0.72	0.00																			
12	0.60	0.60	0.60	0.60	0.72	0.44	0.82	0.72	0.72	0.00	0.00																		
13	0.60	0.72	0.72	0.44	0.72	0.60	0.72	0.60	0.60	0.44	0.44	0.44																	
14	0.60	0.44	0.44	0.44	0.60	0.60	0.72	0.60	0.82	0.44	0.44	0.44	0.60																
15	0.60	0.60	0.60	0.60	0.72	0.44	0.82	0.72	0.72	0.00	0.00	0.00	0.44	0.44															
16	0.44	0.72	0.72	0.60	0.60	0.60	0.82	0.72	0.72	0.44	0.44	0.44	0.44	0.60	0.44														
17	0.66	0.66	0.66	0.31	0.66	0.76	0.66	0.52	0.76	0.66	0.66	0.66	0.52	0.52	0.66	0.66													
18	0.60	0.60	0.60	0.60	0.72	0.44	0.82	0.72	0.72	0.00	0.00	0.00	0.44	0.44	0.00	0.44	0.66												
19	0.60	0.44	0.44	0.44	0.60	0.60	0.72	0.60	0.82	0.44	0.44	0.44	0.60	0.00	0.44	0.60	0.52	0.44											
20	0.72	0.60	0.60	0.44	0.60	0.72	0.72	0.44	0.82	0.72	0.72	0.72	0.60	0.60	0.72	0.72	0.52	0.72	0.60										
21	0.76	0.66	0.66	0.52	0.66	0.76	0.66	0.52	0.85	0.76	0.76	0.76	0.66	0.66	0.76	0.76	0.59	0.76	0.66	0.81									
22	0.82	0.44	0.44	0.72	0.60	0.60	0.72	0.60	0.82	0.72	0.72	0.72	0.82	0.60	0.72	0.82	0.76	0.72	0.60	0.72	0.76								
23	0.82	0.72	0.72	0.60	0.72	0.92	0.60	0.60	0.72	0.82	0.82	0.82	0.72	0.72	0.82	0.82	0.66	0.82	0.72	0.44	0.52	0.60							
24	0.72	0.72	0.72	0.44	072	0.82	0.60	0.60	0.60	0.72	0.72	0.72	0.60	0.60	0.72	0.72	0.52	0.72	0.60	0.60	0.66	0.82	0.72						
25	0.72	0.82	0.82	0.72	0.72	0.89	0.72	0.82	0.72	0.82	0.82	0.82	0.82	0.72	0.82	0.72	0.66	0.82	0.72	0.82	0.85	0.82	0.82	0.60					
26	0.72	0.89	0.89	0.82	0.82	0.82	0.82	0.89	0.60	0.72	0.72	0.72	0.72	0.82	0.72	0.60	0.76	0.72	0.82	0.89	0.91	0.89	0.89	0.72	0.44				
27	0.72	0.82	0.82	0.72	0.72	0.89	0.72	0.82	0.72	0.82	0.82	0.82	0.82	0.72	0.82	0.72	0.66	0.82	0.72	0.82	0.85	0.82	0.82	0.60	0.00	0.44			
28	0.82	0.72	0.72	0.82	0.60	0.82	0.82	0.72	0.82	0.89	0.89	0.89	0.89	0.82	0.89	0.82	0.76	0.89	0.82	0.82	0.85	0.72	0.82	0.72	0.44	0.60	0.44		
29	0.72	0.82	0.82	0.89	0.72	0.82	0.89	0.82	0.82	0.89	0.89	0.89	0.89	0.89	0.89	0.82	0.85	0.89	0.89	0.89	0.92	0.82	0.89	0.82	0.60	0.60	0.60	0.44	
30	0.82	0.82	0.82	0.89	0.72	0.72	0.89	0.82	0.72	0.82	0.82	0.82	0.82	0.89	0.82	0.72	0.85	0.82	0.89	0.89	0.92	0.82	0.89	0.82	0.60	0.44	0.60	0.44	0.44

1=McNair 235, 2=Frego bract, 3=Reba W926, 4=High gossypol, 5=DP Acala90, 6=Aubarn 56, 7=NTA 88-6, 8=BJA 592, 9=Stoneville 506, 10=Ezra, 11=CIM 70, 12=Cyto 12/74, 13=UK82, 14=NTA 93-15, 15=Irma 1243, 16=Guazuncho, 17=DP 4049, 18=HC-B4-75, 19=Nectarless, 20=Des 119, 21=Dixie King, 22=SG 125, 23=Coker 315, 24=F 135, 25=NTA 93-21, 26=MZ561, 27=Okra leaf, 28=Acala SJ$_2$, 29=Delcot 344, 30=UK91
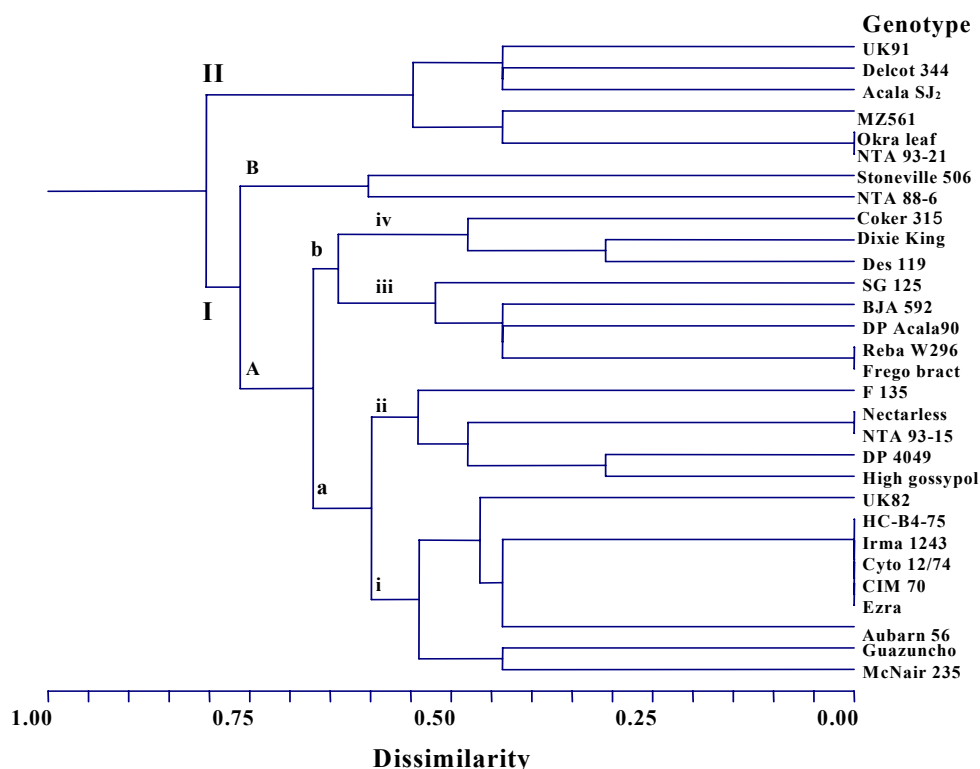
Figure 3.2 Dendrogram based on Euclidean genetic distance and UPGMA clustering for 30 cotton varieties using oil and four fatty acids

Cluster *A* was divided into two other clusters (*a* and *b*). Cluster *a* was subdivided into two subclusters *i* and *ii*. Subcluster *i* had two more subclusters. One subcluster included Guazuncho and McNair 235 with a genetic distance of 0.43 between them (Table 3.6). McNair 235 is an Eastern type variety (USA) (Niles and Feaster, 1984) and Guazuncho is from Argentina. The other subcluster had two groups where UK82 (from Tanzania) joined the group of HC-B4-75, Irma 1243, Cyto 12/74, CIM 70 and Ezra, all these varieties had a genetic distance of 0.00 between them. The genetic distance between this group and UK82 was 0.47. The grouping of CIM 70 and Cyto 12/74 might be because they originated from the same region (Pakistan) and could have shared the same genetic background and the other grouping might be due to similar oil and fatty acids contents.

Subcluster *ii* contained five varieties, F 135, Nectarless, NTA 93-15, DP4049 and High gossypol. NTA 93-15 and DP 4049 have common traits from Deltapine varieties. Nectarless and

High gossypol are insect resistant varieties. The grouping of these varieties could be because some of them are sharing a common genetic background but also because of the similar oil and fatty acid contents. Cluster *b* of the main cluster *A* was subdivided into two subclusters *iii* and *iv*. Subcluster *iii* contained five varieties from which four varieties had related genetic backgrounds [BJA 596 and Reba W296 from West Africa and SG 125 and DP Acala90 from USA (Deltapine type varieties)]. The grouping with the rest was because of similar oil and fatty acid content. Subcluster *iv* contained three varieties (Coker 315, Dixie King and Des 119) and all originated from the USA. Cluster *B* of main group *I* contained two varieties (Stoneville 506 and NTA 88-6) with a genetic distance of 0.6. Stoneville 506 was from the USA (Deltapine type varieties) and NTA 88-6 from Mali. NTA varieties from Mali had traits from Deltapine varieties. Therefore there might be some related genetic background.

Main group *II,* contained the six varieties (Okra leaf, Delcot 344, UK91, MZ561, NTA 93-21 and Acala SJ$_2$), which were multiplied at the UFS. Their growth was influenced by the environment leadings to premature seeds, that increases the amount of linoleic and stearic content while the palmitic fatty acid was decreased (Table 5.3 and Figure 3.2).

3.4 DISCUSSION

The oil content results revealed a large variation from 13.98-24.61% though most of them were within the normal range (15.20-25.70%), with an average of 20.23%. This was in agreement with other reports (Lawhon *et al.,* 1977; Cherry *et al.,* 1981; Salunkhe *et al*., 1992). Total saturated fatty acid, monounsaturated fatty acid and polyunsaturated fatty acid percentages, gives a ratio of polyunsaturated: saturated fatty acids of 2:1. The total unsaturated fatty acid was about 70%, giving a ratio of unsaturated: saturated of about 3.1. This was in agreement with findings by Jones and Kersey (2002).

Variation in stearic acid was observed on McNair 235, NTA 93-15 and Guazuncho. These varieties had low stearic acid values of 2.07%, 2.22% and 2.29% respectively. It should be noted that these varieties were not multiplied at the UFS, it is therefore probably that the environment in which they were originally grown affected the results. Oleic fatty acid content did not reflect the difference between the two growing conditions, suggesting that it is stable as

was reported by Lubbock and Raleigh (1996). However, all cotton varieties had high palmitic fatty acid values (19.23-25.04%), as reported by Scarth and McVetty (1999) and Smith and Rathore (2001). Jones and Kersey (2002) reported palmitic fatty acid as a major contributor to increased levels of total blood cholesterol, especially LDL that leads to a greater risk to cardiovascular disease. Varieties High gossypol, Nectaless, McNair 235, Reba W 296 and BJA 592 had high palmitic fatty acid and all are pest and disease resistant, indicating that these varieties might have common traits.

A positive correlation between oil content and oleic, linoleic and linolenic fatty acids was observed. This was contradictory to results of Green and Marshall (1981) who observed negative correlation between these fractions in flax varieties. Möllers and Schierholt (2002) reported positive correlation between oil and oleic acid but oil and linoleic acid were negatively correlated in oilseed rape. Although linoleic and linolenic are positively correlated, it is still possible to select for genotypes with high linoleic content because linolenic content was low in all 30 varieties used in this study. Furthermore, the presence of a positive correlation of linoleic with oil and oleic acid will make simultaneous selection for all these characteristics possible. ODR mainly influences the desaturation of oleic to linoleic acids, thus, with the positive correlation between ODR and linoleic, it means as ODR increases linoleic increases too. The low value of LDR revealed that the efficiency of desaturation in cottonseed oil from linoleic to linolenic was low and leading to a significant decrease of linolenic content to about 0.12% in these 30 cotton varieties. The high average value of DR, which was similar to ODR indicated the efficiency of desaturation and increase of linoleic fatty acid (C18:2).

The increase of C18:2 and C18:3 resulted in a significant decrease of C14:0, C16:1, C18:0 and C18:1. Similar results were reported by Möllers and Schierholt (2002). Therefore, simultaneous selection for parents with increased stearic, oleic, linoleic fatty acids and reduced palmitic fatty acid might be difficult through conventional breeding. However, this might be possible through metabolic engineering in cotton, on two enzymatic steps in the fatty acid biosynthesis pathway. The technique involves targeting the expression of *Fat-B*, which likely regulates the proportion of monounsaturated fatty acid (mostly C18:1) to polyunsaturated fatty acid (mostly C18:2) (Chapman *et al*., 2001; Liu *et al*., 2002).

Genetic distances between varieties that were not multiplied at the UFS ranged from 0.00-0.80, while genetic distances observed between varieties multiplied at the UFS ranged from 0.72-0.92. Clustering of these six varieties was influenced by growing conditions, which affected oil and fatty acid content. These characteristics are quantitative traits, except for oleic fatty acid. Results confirmed the effect of environment on oil and fatty acid concentration as reported by Lubbock and Raleigh (1996). Apart from the varieties which were influenced by the environment, Stoneville 506 had the highest genetic distance (0.60-0.81) with the rest. NTA 88-6 had the highest genetic distance (0.85) with Aubarn 56. Apart from environmental effects, some varieties were grouped according to origin and pedigree relationships. For example, CIM 70 and Cyto 12/74, as well as most varieties from the USA and Mali that share Deltapine varieties in their pedigree, clustered together. These varieties were probably sharing common ancestors (same genetic background) (Niles and Feaster, 1984; Meredith and Brown, 1998).

3.5 CONCLUSIONS AND RECOMMENDATIONS

This study demonstrated that it is possible to differentiate closely related varieties by using fatty acid and oil values. It was revealed that linoleic, stearic and palmitic fatty acids are influenced by the environment and are probably quantitatively inherited. Therefore in studying genetic diversity using these characteristics, the variation of environment must be considered. The presence of high genetic distances between varieties, which were not multiplied at the UFS, indicated that although all varieties were of *G. hirsutum,* some had higher values than others. Therefore hybridisation for high oleic, linoleic and stearic fatty acid as well as reduced palmitic fatty acid is possible by selecting parents like CIM 70 and Cyto 12/74 for oleic and stearic fatty acid. Stoneville 506 can be selected as a parent for low palmitic fatty acid, increased oleic, stearic and linolenic fatty acids, though the oil content was at medium level (19.25%). Genetic distances showed a possibility of the varieties to form heterotic groups for improving these characteristics.

CHAPTER 4

STUDIES OF GENETIC DIVERSITY IN COTTON (*GOSSYPIUM HIRSUTUM* L.) VARIETIES IN TANZANIA USING AGRONOMICAL AND MORPHOLOGICAL CHARACTERS

4.1 INTRODUCTION

The study of the genetics and evaluation of cotton (*G. hirsutum*) diversity is important for improvement, efficient management and utilisation of the existing gene pool (Altaf Khan *et al*., 2002). Cotton is the most important textile fibre crop (Cherry and Leffler, 1984). In Tanzania, cotton is grown by small-scale farmers in two growing areas known as the Western Cotton Growing Areas (WCGA's) and Eastern Cotton Growing Areas (ECGA's). Cotton varieties grown in these areas are adapted to the environment and tolerant to some diseases and insect pests (Jones and Kapingu, 1982; TCL and SB, 2002).

Diversity analysis and maintenance of crop genotypes are essential processes in identification of genetic relatedness of available genetic resources. It facilitates the selection of potential parents for subsequent crossing and selection of progenies up to the final utilisation of cultivars in production schemes (UPOV, 1991; Van Esbroeck *et al*., 1999; Murtaza *et al*., 2005). This knowledge is important for germplasm collection and conservation (Pillay and Myers, 1999). Accurate morphological characterisation of varieties is an important process in breeding as variety characteristics like resistance to insect pests and diseases can be determined. For example, hairy leaf or stem varieties are resistant to insects like jassids while smooth leaf varieties reduce the trash content in harvested cotton (Meredith *et al.,* 1997). Utilising these morphological characteristics in breeding programmes help cotton growers to obtain high yields and good fibre quality with reduced dependence on pesticides. Meredith *et al*. (1997) reported a higher photosynthesis rate and improved fibre quality characteristics in varieties with subokra leaves compared to varieties with normal leaves.

Morphological markers can be monitored visually without specialised biochemical or molecular techniques. Although agronomical characterisation provides useful information to users, these characteristics are normally subjected to environmental influences and must be assessed during a fixed vegetative phase of the crop (Swanepoel, 1999). Morphological traits that are controlled by a single locus can be used as genetic markers, provided their expression is reproducible over a range of environments (Kumar, 1999). Besides environment, expression of morphological markers is altered by epistatic and pleiotropic interactions (Farooq and Azam, 2002; Rana and Bhat, 2004). Morphological features are mostly polygenic and in cotton display a large amount of variation in different environments. Examples include lint yield, fibre fineness (micronaire) and fibre maturity. Andries *et al*. (1971) reported that row spacing affects plant height. For such characteristics, conclusive results are obtained by repeats over years and/or locations. However, other traits like leaf colour, leaf shape and boll shape are consistent over environments and data from one or two tests normally give a good indication of relative performance (Verhalen *et al*., 2002).

Variation determination for cotton varieties in Tanzania has been done morphologically using phenotypic descriptors. This knowledge has been important mainly for selecting parental material for crossing to improve available varieties (Lukonge *et al*., 1999). Morphological properties mainly used were disease resistance, pest resistance, fibre quality, adaptability and yield components (boll size, boll number, branch number). Environment and cultivation practices had an effect on most of these morphological characteristics. Variation observed in farmers' fields indicated a seed-mixing problem, but since the characters were unknown, it was difficult to identify individual varieties (Hau, 1997). Therefore, the objectives of this study were to characterise and quantify genetic diversity in 30 varieties collected from different areas using agronomical and morphological markers. Information obtained was compared with molecular markers in Chapter 6 to assess the relatedness between these two methods of characterisation.

4.2 MATERIALS AND METHODS

4.2.1 Plant material and field management

Thirty cotton varieties, of which 25 were introduced from other countries and five were locally bred varieties (MZ561, UK82, UK91, IL74 and IL85) were evaluated at Ukiriguru Research Institute, Tanzania (Table 4.1). Field preparation started in October 2003. Farmyard manure at 7 ton/ha and triple super phosphate (TSP) fertilizer at 50 kg/ha were applied before field harrowing. A randomised complete block design (RCBD), replicated four times, was used. Planting in the field was done at the end of November 2003 and five seeds per hill were planted 2.5-5.0 cm deep. Plot sizes were three ridges of 4 m by 0.9 m and agronomic practices were performed (weeding for three times). Urea fertilizer (50 kg/ha) was applied six and 12 weeks (same amount) after planting. Insecticide (*Cypercal D*) was sprayed six weeks after planting (2.5 l/ha). During the first three applications, spraying was done weekly due to high insect pressure, followed by application once every two weeks as recommended. Three harvests were done (1st pick, 2nd pick and 3rd pick). Rainfall started late October 2003 at a low rate and this affected seed germination, but rain increased later. Therefore no drought related problems were experienced during the growing season. Rainfall and temperature data are shown in Appendix 4. 1.

4.2.2 Data collection and characterisation

Data on hairiness, leaf colour, leaf shape and stem colour were collected at 50% boll formation. Number of bolls, size of bolls, plant height and plant shape were characterised at harvesting. Ten bolls (boll sample) were collected randomly from each plot to determine ginning percentages, 100 seed weight, number of seeds per boll, seed fuzz, seed weight per boll, lint weight per boll and boll weight. Seedcotton yield (ton/ha) was determined after three harvests and lint yield per hectare was calculated using seedcotton per hectare and ginning percentages.

Table 4.1 Cotton varieties, origin and characteristics

No	Variety	Source	Characteristics	No	Variety	Source	Characteristics
1	HC-B4-75	USA/France	Drought tolerant	16	UK82	Tanzania	Medium yield, RBB
2	Coker 315	USA	Big bolls	17	Frego bract	USA	Resistant to insects
3	IL85	Tanzania	RBB	18	Delcot 344	USA	High GOT, no hairs
4	Guazuncho	Argentina	Drought resistant	19	DP 4049	USA	Few hairs, small boll, high GOT
5	SG 125	USA	High yield	20	BJA 592	Nigeria	Short staple, high GOT, RBB
6	CIM 70	Pakistan	Good FQ	21	Irma 1243	Cameroon	Good FQ
7	Aubarn 56	USA	RFW, RRN	22	IL74	Tanzania	RBB
8	Des 119	France	RFW	23	MZ561	Tanzania	Low yield, locally adapted
9	Cyto 12/74	Pakistan	Tall variety	24	Okra leaf	Israel	Resistant to insects, drought tolerant
10	NTA 93-15	Mali	High yield, SFW, med. FQ	25	McNair 235	USA	RFW
11	Acala SJ$_2$	USA	Large bolls, high. GOT, SBB	26	UK91	Tanzania	Medium yield, RBB, RFW
12	Stoneville 506	USA	RBB	27	NTA 93-21	Mali	High yield, SFW, high FQ, med. hairs
13	Dixie King	USA	High yield	28	High gossypol	Chadi	Insect resistant, high GOT
14	NTA 88-6	Mali	High yield, SFW, high FQ, few hairs	29	Reba B50	Central Africa	RBB
15	Reba W296	Central Africa	High yield, RBB, low FQ, SFW	30	DP Acala90	USA	Low yield, high GOT

GOT=Ginning outturn, RBB=Resistant to bacterial blight, RFW=Resistant to fusarium wilt, FQ=Fibre quality, SFW=Susceptible to fusarium wilt, RRN=
Resistant to root knot nematode, SBB= Susceptible to bacterial blight

Modified International Board for Plant Genetic Resources (IBPGR) descriptors for cotton were applied to measure 21 characteristics (Table 4.2 and Figure 4.1). Coding of morphological data into a binary matrix was applied in different ways. Characteristics that were similar in all varieties were not scored, for example petal colour, petal spot, flower branches and boll size. Characteristics with only two categories for description were scored as present (1) and absent (0) and entered into the data matrix. For example, if the genotype had cream pollen it was coded as 1 while yellow pollen was coded as 0. Normal leaf shape was coded as present (1) and okra leaf shape was coded as absent (0) (Appendix 4.2).

Characteristics with more than two categories, like stigma position, leaf hairs, leaf colour and fuzz density were coded by considering the whole range of diversity. For example, stigma position ranged between 1 (below anthers), 2 (same level) and 3 (above anthers). On coding, below anthers were coded as 1 while same level and above anthers for a specific variety was coded as 0. The same was applied for all characteristics with more than two categories.

Plant height was coded differently from other morphological characteristics using three classes: 3 (short), 5 (medium) and 7 (tall). If plant height was 80-95 cm it was classed as short, 95-110 cm as medium and above 110 cm as tall. Four plants were randomly selected from each plot, heights were measured and the mean for each plot was determined. Varieties with similar heights as the medium class were scored as 1 and while varieties with lower or higher heights than the medium class were scored as 0.

a. Boll shape

b. Pollen colour

c. Stigma position
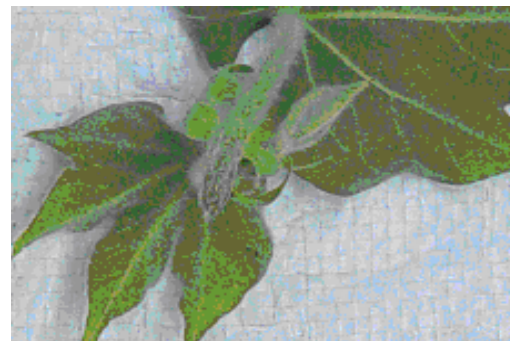
d. Few hairs

e. Many hairs

f. Okra leaf

g. Broad leaf

Figure 4.1 **Examples of morphological evaluated characteristics: a) =boll shapes, b) =pollen colour, c) =stigma position, d) =few hairs, e) =many hairs, f) =okra leaf, g) =broad leaf**

Table 4.2 List of morphological characteristics measured

Characteristic	Description
Pollen colour	Colour of pollen at 50% flowering
Stigma position	Position of stigma relative to anthers at 50% flowering
Flower type	Type of flowering at 50% flowering
Leaf shape	Leaf shape at 50% flowering
Leaf colour	Intensity of green colour
Leaf hairs	Leaf pubescence (lower side)
Stem hairs	Stem pubescence in upper part
Stem colour	Stem colour at 50% flowering
Bract dentition	Bract dentition at green maturity
Boll peduncle	Length of peduncle
Boll shape	Shape in longitudinal section
Boll prominence	Prominence of tip
Plant shape	Shape of plant at harvesting
Plant height	Height above ground measured in cm at harvest
Seed fuzz	Density of fuzzy
Hundred seed weight	Weight of 100 seeds (g)
Number of bolls per plant	Number of bolls per plant counted at harvest
Boll weight	Weight of the boll (g)
Ginning percentage	Content of lint (%)
Boll lint weight	Content of lint (g)
Boll seed weight	Content of seed (g)

4.2.3 Statistical analysis

Analysis of variance (ANOVA) and correlation coefficients on yield, yield components, ginning outturn (GOT) and fibre quality were calculated using Agrobase (2000).

Data were converted into binary data where presence was scored as 1 and absence as 0. Coded data for morphological characteristics were subjected to analysis using NCSS (Hintze, 2000). The Euclidean distance method (Kaufman and Rousseeuw, 1990) was used to construct distance

matrices. Cluster analysis was performed using the genetic matrices generated by the Euclidean distance method to reveal the patterns of genetic relationships among genotypes. Dendrograms were constructed using UPGMA clustering (Hintze, 2000). The co-phenetic correlation (Kaufman and Rousseeuw, 1990; Hintze, 2000) for each dendrogram was computed as a measure of 'goodness of fit' for each dendrogram. The two delta goodness of fit statistics, delta (0.5) and delta (0.1) were calculated to determine which clustering configuration fits the data best.

4.3 RESULTS

4.3.1 Agronomical characteristics
a.Yield and yield components
Plant height

Mean squares for plant height (Table 4.3) were significant, showing differences among varieties. NTA 93-15 (122.88 cm), Frego bract (122.13 cm), Cyto 12/74 (120.00 cm), Irma 1243 (118.56 cm), IL74 (118.00), MZ561 (116.56 cm), Acala SJ$_2$ (113.94 cm) and NTA 88-6 (112.06 cm) were significantly taller while Delcot 344 (89.88 cm), Stoneville 506 (89.63 cm), McNair 235 (89.00 cm), DP Acala90 (88.50 cm), Guazuncho (87.38 cm) and Des 119 (80.25 cm) were significantly shorter than the rest. For local varieties, as selection advanced, varieties became shorter. For example, for IL74 (118.00 cm), after selection the released variety IL85 was 105.69 cm. MZ561 was 116.56 cm but as selection continued, UK82 was 108.75 cm and the recent variety UK91 was 94.00 cm. The same was observed for the Reba variety material (Reba B50 and Reba W296), where Reba B50 was 107.38 cm and Reba W296 was 106.13 cm. Results indicated that selection for this trait concentrated on medium plant height (Table 4.3).

Boll weight

ANOVA results (Table 4.3) for boll weight showed significant differences among varieties. Dixie King (7.65 g), Coker 315 (7.35 g), Reba W296 (7.25 g), BJA 592 (7.15 g) Acala SJ$_2$ (6.85 g), HC-B4-75 (6.70 g), Stoneville 506 (6.70 g), DP Acala90 (6.58 g) and McNair 235 (6.53g) were significantly heavier than other varieties. Varieties Des 119 (5.40 g), Frego bract (5.18 g),

Table 4. 3 Variety means for different agronomical characteristics

No	Variety	Plant height (cm)	Boll weight (g)	GOT (%)	Bolls/ plant (no)	Lint/ boll (g)	Seed wt/ boll (g)	100 seed wt (g)	Fuzz grade	Lint/ha (ton/ha)	Seedcotton yield (ton/ha)	Seed/ boll (no)	Fibre length (mm)	Micronaire (units)	Maturity (%)	Fibre strength (g/tex)	Elongation (%)
1	High gossypol	97.50	5.00	35.95	21.40	1.80	3.78	11.8	6.50	0.77	2.15	30.00	32.3	4.25	92.00	40.70	7.50
2	HC-B4-75	103.44	6.70	41.48	25.78	2.78	3.80	12.13	7.25	1.08	2.61	27.00	30.5	4.59	88.00	26.70	8.00
3	NTA 93-15	122.88	6.10	43.43	33.45	2.65	3.33	12.05	6.75	1.15	2.64	34.00	27.8	4.41	89.00	31.10	10.20
4	BJA 592	105.56	7.15	38.53	27.65	2.60	4.13	12.45	6.25	1.20	3.10	37.00	29.3	3.86	86.00	29.80	7.30
5	RebaW296	106.13	7.25	37.23	28.33	2.70	4.38	12.68	6.00	1.02	2.73	32.00	28.8	5.42	85.00	31.30	8.00
6	Des 119	80.25	5.40	40.78	19.13	2.20	3.15	10.90	7.00	0.79	1.94	27.00	27.0	4.72	90.00	29.20	7.40
7	MZ561	116.56	6.18	38.65	40.23	2.43	4.00	13.15	6.25	0.88	2.15	25.00	28.8	4.91	96.00	32.20	8.33
8	CIM 70	105.44	5.75	37.15	31.95	2.18	4.18	10.78	6.25	0.86	2.33	31.00	30.0	4.31	92.00	31.10	8.60
9	Frego bract	122.13	5.18	40.95	34.65	1.85	3.23	10.93	7.00	0.94	2.63	37.00	32.5	3.92	85.00	29.60	7.50
10	McNair 235	89.00	6.53	35.80	24.08	2.68	3.73	11.25	6.75	0.96	2.35	34.00	32.0	4.81	91.00	31.40	7.80
11	UK91	94.00	5.93	37.80	21.80	2.24	3.76	12.08	5.75	1.06	2.76	30.00	31.0	4.33	93.00	33.60	8.60
12	Dixie King	101.19	7.65	40.45	21.03	2.98	4.30	12.75	6.75	1.08	2.66	35.00	28.5	4.38	86.00	24.90	8.00
13	IL85	105.69	6.00	35.33	36.58	2.13	3.73	11.80	6.50	1.07	3.03	28.00	28.3	4.66	95.00	30.53	8.38
14	Irma 1243	118.56	5.98	44.90	25.30	2.68	3.20	11.70	6.75	1.07	2.40	26.00	29.3	4.54	89.00	31.05	7.80
15	DP Acala90	88.50	6.58	40.08	20.35	2.65	3.78	10.83	6.50	0.91	2.24	32.00	30.0	4.47	95.00	31.60	8.00
16	UK82	108.75	6.05	38.93	44.28	2.28	3.73	12.38	6.25	0.98	2.57	27.00	31.0	4.45	94.00	31.80	8.10
17	DP 4049	97.75	6.15	42.25	21.15	2.60	3.50	10.78	7.25	1.13	2.66	33.00	29.3	4.83	90.00	29.10	8.00
18	Okra leaf	99.88	4.98	33.58	19.03	1.68	3.18	10.63	6.00	0.62	1.71	22.00	28.8	5.04	97.00	30.60	8.08
19	IL74	118.00	5.80	38.78	25.65	2.25	3.98	13.33	6.00	0.93	2.39	27.00	27.8	4.94	97.00	30.33	8.00
20	Guazuncho	87.38	5.80	41.40	28.83	2.40	3.23	10.43	7.00	1.28	3.09	30.00	32.3	4.34	89.00	30.90	8.00
21	SG 125	102.75	6.15	40.80	29.30	2.60	3.75	11.25	7.25	1.16	2.85	30.00	30.0	5.13	94.00	35.00	9.60
22	Delcot 344	89.88	5.93	41.18	22.33	2.48	3.38	11.35	7.50	0.95	2.24	28.00	28.0	4.00	85.00	27.10	7.60
23	Cyto 12/74	120.00	5.78	34.55	33.08	2.00	3.58	12.43	6.25	1.06	3.07	32.00	31.3	4.22	94.00	35.50	8.10
24	Acala SJ$_2$	113.94	6.85	37.93	26.28	2.60	4.18	13.35	6.50	1.06	2.68	27.00	27.8	3.68	84.00	28.20	8.00
25	Stoneville 506	89.63	6.70	39.55	22.53	2.65	4.00	12.78	7.00	0.85	2.13	27.00	27.3	4.44	90.00	34.00	8.00
26	Coker 315	100.88	7.35	41.43	26.48	2.85	4.00	11.15	7.25	1.08	2.59	30.00	29.5	3.62	91.00	32.00	8.10
27	Reba B50	107.38	5.98	36.83	32.95	2.20	3.73	12.33	6.25	0.89	2.42	24.00	32.3	4.55	89.00	29.60	7.50
28	Aubarn 56	105.25	6.45	39.93	28.95	2.58	3.88	12.15	6.00	1.06	2.61	33.00	29.0	4.43	93.00	34.20	8.90
29	NTA 88-6	112.06	6.43	43.98	27.48	2.83	3.48	12.75	6.25	1.02	2.32	28.00	30.5	3.95	88.00	33.80	7.70
30	NTA 93-21	100.25	5.75	44.43	28.75	2.55	3.05	10.93	6.25	1.31	2.96	28.00	31.3	4.86	95.00	34.18	8.90
	Mean	**103.69**	**6.18**	**39.35**	**27.62**	**2.43**	**3.68**	**11.84**	**6.58**	**1.19**	**2.53**	**29.65**	**29.80**	**4.40**	**91.00**	**31.37**	**8.13**
	LSD for V	13.26	0.72	2.13	8.28	0.30	0.51	0.64	0.56	0.00	0.51	0.85	0.00	0.09	0.01	0.98	0.16
	SED for V	7.97	0.43	1.28	4.98	0.18	0.31	0.48	0.34	0.00	0.301	0.51	0.00	0.05	0.01	0.59	0.09
	Repeat.(%)	62.00	67	79	53	73	57	78	60	99	55	97	99	97	99	95	97
	CV (%)	10.87	9.88	4.60	25.45	10.65	11.85	4.59	7.22	0.08	17.12	2.44	0.08	1.73	0.54	2.66	1.66
	Sign. For V	***	***	***	***	***	***	*	***	**	*	***	***	***	***	***	***

*p ≤ 0.05, ** p ≤ 0.01; *** p ≤ 0.001 level of significance; CV=coefficient of variation; LSD= least significance difference, SED=standard deviation; Sign. for V=significance for variety; Repeat.=repeatability.

GOT=Ginning Outturn, wt =weight

High gossypol (5.00 g) and Okra leaf (4.98 g) had significantly smaller and lighter bolls compared to other varieties. Boll weight for Mali varieties were reduced as selection continued [NTA 88-6 (6.43 g), NTA 93-15 (6.10 g) and NTA 93-21 (5.75g)]. The same was observed for Tanzanian material [MZ 561 (6.18 g), UK82 (6.05g) and UK91 (5.93 g)]. Dixie King (7.65 g), Coker 315 (7.35 g), Acala SJ$_2$ (6.85 g), DP Acala90 (6.58 g) and McNair 235 (6.53 g) had heavy bolls and originated from the USA (Eastern type varieties), though some were brought from France to the institute. This indicated that, during selection of these varieties, bigger bolls were one of the objectives as one of the yield components (Table 4.3).

Number of bolls

Number of bolls per plant was significantly higher for UK82 (44.28), MZ561 (40.23) and IL85 (36.56) compared to the rest of the varieties. These varieties originated from Tanzania, indicating the adaptation of these lines for this trait at this environment. Other varieties with high number of bolls included Frego bract (34.65), NTA 93-15 (33.45), Cyto 12/74 (33.08), Reba B50 (32.98) and CIM 70 (31.95). Des 119 (19.13) and Okra leaf (19.03) had the lowest number of bolls per plant (less than 20) compared to other varieties. In some varieties plant height was positively related to the number of bolls per plant because of increased plant branches and decreased boll weight. For example, UK82 (44.28), MZ561 (40.23), Frego bract (34.65), NTA 93-15 (33.45), Cyto 12/74 (33.08) and Reba B50 (32.98) had many bolls and were taller plants (Table 4.3). McNair 235, Dixie King and DP Acala90 had the lowest number of bolls per plant but with larger bolls, while Frego bract, Cyto 12/74 and CIM 70 (from Pakistan) had a larger number of bolls per plant but with the lowest boll weight. Therefore, as one trait increased, the other decreased (Table 4.3).

Ginning percentage (ginning outturn)

ANOVA results (Table 4.3) for GOT indicated highly significant differences among varieties. The highest GOT was obtained from Irma 1243 (44.90%), varieties from Mali [NTA 93-21 (44.43%), NTA 88-6 (43.98%), NTA 93-15 (43.43%)], DP 4049 (42.25%), HC-B4-75 (41.48%), Coker 315 (41.43%), Guazuncho (41.40%), Delcot 344 (41.18%), McNair 235 (40.98%), SG 125 (40.80%), Des 119 (40.78%), Dixie King (40.45%) and DP Acala90 (40.08%). Reba B50 (36.83%), High gossypol (35.95%), Frego bract (35.80%), IL85 (35.33%), Cyto 12/74 (34.55%) and Okra leaf (33.58%) had the lowest GOT values compared to others.

This indicated that varieties from Mali and Eastern types from the USA are important for crossing with Tanzanian material to improve GOT value (Table 4.3).

Fuzz grade

Analysis results (Table 4.3) for fuzz grade indicated that there were statistically significant differences among varieties. Varieties with higher fuzz grade were Delcot 344 (7.50), HC-B4-75 (7.25), DP 4049 (7.25), Coker 315 (7.25) and SG 125 (7.25). These were followed by Guazuncho (7.00), Stoneville 506 (7.00), Des 119 (7.00) and Frego bract (7.00). Varieties IL74 (6.00), Okra leaf (6.00), Aubarn 56 (6.00), Reba W296 (6.00) and UK91 (5.75) had low seed fuzz grade and were not significantly different between themselves.

Lint weight per boll

Statistically significant differences among varieties for lint weight per boll were observed. Dixie King (2.98 g) produced more lint compared to all other evaluated varieties. This was followed by Coker 315 (2.85 g), NTA 88-6 (2.83 g) and HC-B4-75 (2.78 g), which had significant higher values than the rest. Cyto 12/74 (2.00 g), Frego bract (1.85 g), High gossypol (1.80 g) and Okra leaf (1.68 g) had significantly lower lint per boll compared to other varieties (Table 4.3).

Seed weight per boll

The ANOVA (Table 4.3) indicated that variation in seed weight per boll was highly significant among varieties. Varieties that produced significantly heavier seeds compared to others were Reba W296 (4.38 g) and Dixie King (4.33 g), followed by CIM 70 (4.18 g), Acala SJ$_2$ (4.18 g), and BJA 592 (4.13 g), that had heavy bolls but with few bolls per plant. Irma 1243 (3.2 g), Okra leaf (3.18 g), Des 119 (3.15 g), High gossypol (3.1 g) and NTA 93-21 (3.08 g) had the lowest seed weight per boll.

Hundred seed weight

Results on 100 seed weight showed significant differences among varieties (Table 4.3). Acala SJ$_2$ (13.35 g), IL74 (13.33 g), MZ561 (13.15 g), Stoneville 506 (12.78 g), Dixie King (12.75 g) and NTA 88-6 (12.75 g) had significantly heavier seed compared to other varieties, followed by Reba W296 (12.75 g) and BJA 592 (12.68 g). Frego bract (10.93 g), NTA 93-21 (10.93 g), Des 119 (10.9 g), DP Acala90 (10.83 g), CIM 70 (10.78 g), DP 4049 (10.78 g), Okra leaf (10.63 g)

and Guazuncho (10.43 g) had the lightest seeds. The weight for all varieties was in the acceptable range of not less than 10 g per 100 seeds.

Seedcotton and lint yield

Seedcotton yield results shown in Table 4.3 indicated significant differences among varieties. BJA 592 (3.1 ton/ha), Guazuncho (3.09 ton/ha), Cyto 12/74 (3.07 ton/ha), IL85 (3.03 ton/ha) and NTA 93-21 (2.96 ton/ha) had significantly higher yields compared to other varieties. These were followed by SG 125 (2.85 ton/ha), UK91 (2.76 ton/ha), Reba W296 (2.73 ton/ha), Acala SJ$_2$ (2.68 ton/ha), Dixie King (2.66 ton/ha), DP 4049 (2.66 ton/ha), NTA 93-15 (2.64 ton/ha), Frego bract (2.63 ton/ha), Aubarn 56 (2.61 ton/ha), HC-B4-75 (2.61 ton/ha), Coker 315 (2.59 ton/ha) and UK82 (2.57 ton/ha). Des 119 (1.94 ton/ha) and Okra leaf (1.71 ton/ha) had the lowest yield. On lint yield per hectare, varieties NTA 93-21 (1.31 ton/ha), Guazuncho (1.28 ton/ha), BJA 592 (1.20 ton/ha), SG 125 (1.16t /tonha) and NTA 93-15 (1.15 ton/ha) had the highest lint yield. Okra leaf (0.62 ton/ha), High gossypol (0.77 ton/ha), Stoneville 506 (0.85 ton/ha), CIM 70 (0.86 ton/ha), MZ561 (0.88 ton/ha) and Des 119 (0.79 ton/ha) had the lowest lint yield per hectare. From this study it can be concluded that the ginning outturn and lint per boll determined the amount of lint per hectare.

Number of seeds per boll

Results of seed number per boll indicated that BJA 592 and Frego bract had the highest seed number (37), followed by Dixie King (35), McNair 235 (34), NTA 93-15 (33), Aubarn 56 (33) DP 4049 (32) Cyto 12/74 (32), DP Acala90 (32), Reba W296 (32) and CIM 70 (31). Okra leaf, Reba B50, MZ561, Irma 1243, UK82, IL74, Acala SJ$_2$, HC-B4-75, Des 119 and Stoneville 506 had the lowest number of seeds per boll (22 to 27) (Table 4.3). This indicated that varieties with a bigger number of seeds per boll had bigger boll weight and seed weight per boll.

b. Fibre quality

Fibre length

Fibre length of Frego bract (32.5 mm), Guazuncho (32.3 mm), Reba B50 (32.3 mm), High gossypol (32.3 mm), NTA 93-21 (32.0 mm), Cyto 12/74 (31.3 mm), UK82 (31.0 mm) and UK91 (31.0 mm) was significantly longer than the fibre of NTA 93-15 (28.0 mm), IL74 (28.0 mm), Acala SJ$_2$ (27.8 mm), Stoneville 506 (27.3 mm) and Des 119 (27.0 mm) while others were

in between, though they were still above the acceptable value ($\geq$ 25.15 mm) (Appendix 4.3 and Table 4.3).

Fibre strength

Results of fibre strength indicated significant differences among varieties ranging between 24.90-40.70 g/tex, though all were of acceptable strength (22.00 g/tex and above). High gossypol (40.70 g/tex), Cyto 12/74 (35.50 g/tex), SG 125 (35.00 g/tex), Aubarn 56 (34.20 g/tex), NTA 93-21 (34.19 g/tex), Stoneville 506 (34.00 g/tex), NTA 88-6 (33.87 g/tex), UK91 (33.67 g/tex) and MZ561 (32.20 g/tex) had the highest fibre strength while Delcot 344 (27.10 g/tex), HC-B4-75 (26.70 g/tex) and Dixie King (24.90 g/tex) had the lowest fibre strength. Results indicated that some varieties with low yield had high fibre strength for example High gossypol (Table 4.3).

Fibre maturity

High maturity values (84%-97%) were observed. IL74 (97%), Okra leaf (97%), MZ561 (96%), NTA 93-21 (95%), DP Acala90 (95%), IL85 (95%), Cyto 12/74 (95%), UK82 (94%), SG 125 (94%), UK91 (93%) and Aubarn 56 (93%) had the highest maturity values while Frego bract (85%), Delcot 344 (85%), Reba W296 (85%) and Acala SJ$_2$ (84%) had the lowest maturity values, though values were all in the acceptable range (80% and above) (Table 4.3).

Micronaire value (fibre fineness)

Micronaire values that measure fibre fineness, indicated significant differences among varieties. SG 125 (5.13 units), Okra leaf (5.04 units), IL74 (4.94 units), MZ561 (4.91 units), NTA 93-21 (4.86 units), DP 4049 (4.83 units), McNair 235 (4.81 units), Des 119 (4.72 units) and IL85 (4.66 units) had the highest micronaire values. Delcot 344 (4.00), NTA 88-6 (3.95), Frego bract (3.92), BJA 592 (3.86), Acala SJ$_2$ (3.68), Coker 315 (3.62) and Reba W296 (3.42) had the lowest micronaire values. The acceptable level of medium fibre fineness is between 3.5-4.9 units (Appendix 4.4). Fibres with micronaire values of 5.0-5.9 are coarse fibres while six and above are very course fibres. This study detected no varieties with very course fibres (Table 4.3).

Elongation

Elongation values were significantly different among varieties. NTA 93-15 (10.2 %), SG 125 (9.6%), NTA 93-21 (8.9%), Aubarn 56 (8.9%), CIM 70 (8.6%), UK91 (8.6%), IL85 (8.4%) and MZ561 (8.3%) had the highest elongation percentages. The lowest elongation percentages were observed in Frego bract (7.5%), Reba B50 (7.5%), High gossypol (7.5%), Des 119 (7.4%) and BJA 592 (7.3%). The acceptable percentage for elongation is 7% and above (Appendix 4.4). All varieties had acceptable elongation percentages (Table 4.3).

4.3.2 Correlation

The correlation analysis was conducted for 16 agronomical characteristics which were significantly different among varieties (Table 4.4). Significantly positive correlations were observed between plant height and seed cotton yield, lint yield, boll number and 100 seed weight. Boll weight was highly significantly positively correlated with seedcotton yield, lint yield, 100 seed weight, boll lint weight and boll seed weight. Ginning percentage was highly significantly positively correlated with boll lint weight and lint yield. Seedcotton yield was significantly positively correlated with boll number, boll lint weight, boll weight and boll seed weight. Hundred seed weight was significantly positively correlated with boll weight, lint weight and seed fuzz.

Seed number per boll was significantly positively correlated with seedcotton per hectare and lint yield per hectare. A highly significant negative correlation was observed between seed number per boll and fibre fineness (micronaire) and fibre maturity. A negative correlation was observed between fibre maturity and boll weight, lint weight per boll and seed fuzz, while a significant positive correlation was observed between fibre maturity and fibre fineness (micronaire), uniformity, strength and elongation. The fibre strength was significantly positively correlated with fibre maturity and fibre uniformity. A significant negative correlation was observed between fibre fineness (micronaire) and boll weight and seeds per boll. A positive and significant correlation of elongation was observed with lint per hectare, micronaire, maturity and uniformity. Fibre length was significantly and positively correlated with fibre uniformity and fibre strength, while a negative and significant value was observed with lint weight per boll.

Table 4.4 Correlation between 16 agronomical characteristics for cotton varieties

	PLH	BWT	GOT	L/HA	SC/HA	BNO	LWB	SWB	SFZ	HSW	MIC	MAT	UNF	STR	ELON	SPB
BWT	0.155															
GOT	-0.129	0.223														
L/HA	0.254 *	0.361 ***	0.345 ***													
SC/HA	0.327 **	0.308 **	-0.013	0.929 ***												
BNO	0.533 ***	0.119	-0.091	0.270 *	0.329 **											
LWB	0.070	0.848 ***	0.629 ***	0.430 ***	0.224 *	0.047										
SWB	0.111	0.314 **	-0.234 *	-0.034	0.046	-0.004	0.151									
SFZ	-0.225	0.145	0.274 *	0.046	-0.054	-0.249 *	0.273 *	-0.079								
HSW	0.339 **	0.447 ***	-0.150	0.088	0.169	0.216	0.288*	0.144	-0.253 *							
MIC	-0.107	-0.287 *	0.080	0.094	-0.149	-0.033	-0.143	0.004	0.013	-0.144						
MAT	-0.009	-0.291 *	-0.171	-0.136	-0.092	0.108	-0.297 *	0.067	-0.312 **	-0.105	0.6674 ***					
UNF	-0.169	-0.204	0.114	0.102	0.053	-0.002	-0.125	-0.133	-0.013	-0.243 *	0.4165 ***	0.321 **				
STR	-0.001	-0.259 *	-0.101	-0.072	-0.041	0.025	-0.249	-0.132	-0.249 *	-0.013	0.0639	0.463 ***	0.337 **			
ELON	0.157	-0.010	0.204	0.249 *	0.173	0.185	0.104	-0.009	-0.092	0.081	0.2919 *	0.339 **	0.289 **	0.211		
SPB	0.021	0.228	0.041	0.259 *	0.271 *	-0.018	0.172	-0.086	0.143	-0.108	-0.3544 **	-0.348 ***	-0.137	-0.023	0.069	
FLENG	0.021	-0.211	-0.101	0.097	0.140	0.178	-0.232 *	-0.151	-0.055	-0.197	-0.0172	0.059	0.505 ***	0.334 **	-0.144	0.188

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$; PLH=Plant height, BWT=Boll weight, GOT=Ginning outturn, L/HA=Lint per hectare, SC/HA=Seed cotton per hectare, BNO=Boll number per plant, LWB=Lint weight per boll, SWB=Seed weight per boll, SFZ=Seed fuzz, HSW=100 seed weight, MIC=Micronare, MAT=Maturity, UNF=Uniformity, STR=Strength, ELON=Elongation, SPB=Seed number per boll, FLENG= Fibre length

4.3.3 Morphological characterisation

Some morphological characteristics were common for all 30 varieties. For example, all varieties had cream petal colour, nectars and lacked petal spot. The lack of petal spot is associated with *G. hirsutum* and distinguishes it from *G. barbadence*. However, clear variation was observed for petal colour, stigma position, leaf colour, leaf shape, leaf size, leaf hair, stem hair, stem colour, boll dentition, boll size, boll shape, boll prominence, boll peduncle and plant shape (example in Figure 4.1). These characteristics were used for characterisation and the observed differences among varieties indicated the possibility of using morphological markers to differentiate varieties for germplasm collection and maintenance and for selection of suitable parents from the population (Appendix 4.2).

a. Genetic distances for 15 morphological characteristics

The genetic distance matrix for 30 varieties from a combination of 15 characteristics (Appendix 4.2) is given in Table 4.5 and distances ranged from 0.18-0.80. Cyto 12/74 and MZ561 were the most similar (0.18) while DP 4049 and Aubarn 56 were the most dissimilar (0.80) varieties. Two varieties; (Okra leaf and NTA 93-21) morphological characteristics were quite different from other varieties. For example, Okra leaf had okra type leaves compared to the normal leaves of all the other varieties. NTA 93-21 had smaller but normal leaves compared to other varieties. Okra leaf had the highest genetic distances with all varieties ranging from 0.55-0.77 (Table 4.5).

b. Cluster analysis

The dendrogram (Figure 4.2) consisted of two main groups, *A* and *B*. Each of the main groups divided into two clusters. Main group *A* consisted of cluster *1*, containing one variety (Okra leaf) and cluster *2* that divided into subclusters *a* and *b*. Subcluster *a* divided into two more subclusters; one contained BJA 596, Reba W296 (both from Central and West Africa) and High gossypol, Aubarn 56, Cyto 12/74, UK82, IL74, MZ561 and UK91. Four of these varieties (UK82, IL74, MZ561 and UK91) were local material developed in Tanzania. The second subcluster contained Frego bract and CIM 70.

Table 4.5 Genetic distances for 30 cotton varieties using 15 morphological characteristics

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
2	0.58																												
3	0.71	0.66																											
4	0.68	0.52	0.55																										
5	0.66	0.41	0.63	0.55																									
6	0.61	0.55	0.58	0.55	0.65																								
7	0.66	0.66	0.52	0.66	0.65	0.54																							
8	0.61	0.61	0.68	0.61	0.56	0.77	0.77																						
9	0.71	0.61	0.45	0.55	0.65	0.52	0.45	0.77																					
10	0.55	0.48	0.58	0.55	0.56	0.68	0.77	0.52	0.63																				
11	0.58	0.68	0.55	0.63	0.65	0.66	0.61	0.66	0.61	0.61																			
12	0.45	0.45	0.66	0.52	0.54	0.66	0.66	0.55	0.66	0.41	0.52																		
13	0.66	0.55	0.63	0.41	0.41	0.63	0.73	0.63	0.63	0.52	0.66	0.48																	
14	0.52	0.52	0.61	0.06	0.56	0.61	0.71	0.55	0.71	0.32	0.68	0.52	0.61																
15	0.68	0.63	0.61	0.68	0.65	0.48	0.55	0.71	0.55	0.66	0.58	0.68	0.71	0.63															
16	0.71	0.61	0.45	0.61	0.65	0.37	0.26	0.77	0.37	0.73	0.66	0.71	0.68	0.66	0.48														
17	0.71	0.66	0.58	0.66	0.65	0.36	0.58	0.77	0.63	0.68	0.71	0.75	0.73	0.61	0.55	0.52													
18	0.58	0.45	0.61	0.45	0.41	0.61	0.71	0.61	0.61	0.41	0.58	0.37	0.32	0.52	0.68	0.66	0.71												
19	0.63	0.52	0.66	0.58	0.56	0.71	0.8	0.55	0.66	0.32	0.68	0.52	0.55	0.45	0.68	0.75	0.71	0.52											
20	0.75	0.66	0.52	0.55	0.65	0.52	0.58	0.73	0.45	0.68	0.48	0.71	0.63	0.75	0.41	0.52	0.58	0.66	0.66										
21	0.58	0.58	0.41	0.58	0.65	0.55	0.66	0.61	0.61	0.48	0.45	0.58	0.61	0.52	0.63	0.61	0.55	0.52	0.58	0.55									
22	0.68	0.52	0.48	0.58	0.65	0.55	0.48	0.71	0.32	0.61	0.63	0.58	0.61	0.68	0.45	0.41	0.66	0.58	0.63	0.48	0.63								
23	0.73	0.58	0.41	0.52	0.65	0.48	0.41	0.75	0.18	0.66	0.58	0.63	0.61	0.73	0.52	0.32	0.61	0.58	0.68	0.41	0.58	0.23							
24	0.77	0.68	0.66	0.58	0.65	0.61	0.55	0.66	0.61	0.75	0.68	0.68	0.71	0.73	0.63	0.55	0.71	0.68	0.73	0.66	0.73	0.63	0.62						
25	0.55	0.48	0.63	0.55	0.41	0.63	0.68	0.58	0.68	0.52	0.61	0.41	0.45	0.55	0.75	0.68	0.73	0.32	0.61	0.73	0.55	0.66	0.65	0.66					
26	0.75	0.71	0.52	0.58	0.65	0.52	0.45	0.73	0.41	0.73	0.66	0.71	0.63	0.71	0.55	0.37	0.63	0.66	0.66	0.52	0.66	0.48	0.43	0.48	0.73				
27	0.61	0.55	0.63	0.55	0.57	0.68	0.73	0.58	0.68	0.52	0.55	0.55	0.63	0.55	0.75	0.73	0.73	0.55	0.61	0.68	0.55	0.75	0.65	0.66	0.45	0.77			
28	0.61	0.61	0.52	0.61	0.65	0.45	0.37	0.77	0.45	0.73	0.48	0.61	0.68	0.75	0.55	0.45	0.58	0.66	0.75	0.45	0.55	0.48	0.45	0.61	0.63	0.58	0.68		
29	0.66	0.48	0.45	0.71	0.65	0.63	0.58	0.63	0.52	0.52	0.61	0.61	0.68	0.54	0.55	0.52	0.63	0.61	0.55	0.58	0.48	0.41	0.57	0.75	0.63	0.63	0.68	0.58	
30	0.68	0.45	0.61	0.58	0.56	0.66	0.71	0.48	0.61	0.41	0.63	0.52	0.55	0.52	0.58	0.66	0.66	0.52	0.37	0.55	0.52	0.52	0.65	0.73	0.61	0.66	0.66	0.66	0.41

1= HC-B4-75, 2= Coker 315, 3= IL85, 4= Guazuncho, 5= SG 125, 6= CIM 70, 7= Aubarn 56, 8= Des 119, 9= Cyto 12/74, 10= NTA 93-15, 11= Acala SJ$_2$, 12= Stoneville 506, 13= Dixie King, 14= NTA 88-6, 15= Reba W296, 16= UK82,

17= Frego bract, 18= Delcot 344, 19= DP 4049, 20= BJA 592, 21= Irma 1243, 22= IL74, 23= MZ561, 24= Okra leaf, 25= McNair 235, 26= UK91, 27= NTA 93-21, 28= High gossypol, 29=Reba B50, 30= DP Acala90.

Varieties Reba W296, BJA 592 and CIM 70 were morphologically similar, especially for leaf size, leaf hairs and pollen colour. Reba W296 (Allen 51 x Coker 100) and BJA 592 were used as parental pedigree material for Tanzanian varieties. Cluster *b* contained varieties Acala SJ$_2$, Reba B50 (Stoneville 2B x Allen 50T), Irma 1243 and IL85. These varieties were similar for stem colour, plant shape and boll shape. Reba B50 and Irma 1243 both were from West Africa. IL85 had traits from Acala SJ$_2$ in its ancestral material. Most of varieties in main group *A* were from Africa, indicating that they were probably sharing the gene pool in their pedigree.
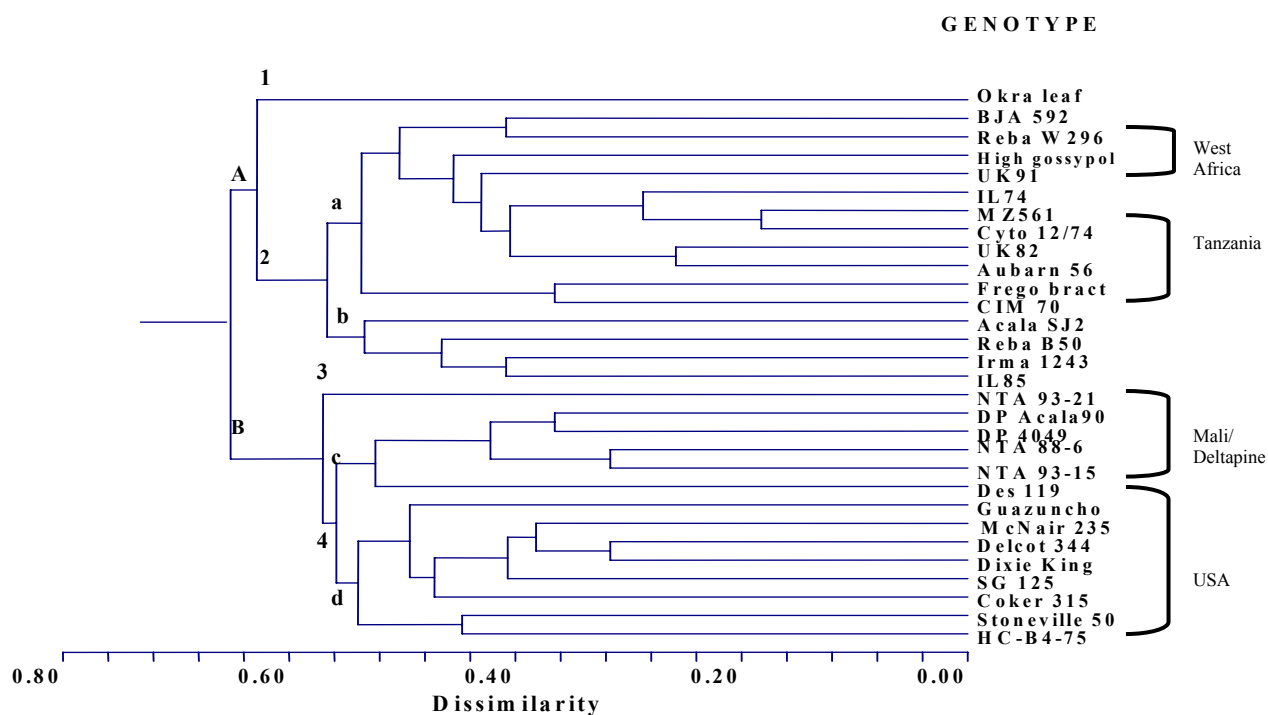
Figure 4.2 Dendrogram based on Euclidean distance and UPGMA clustering method for 30 cotton varieties using morphological data

Main group *B* was divided into two clusters, *3* and *4*. Cluster *3* contained a single variety NTA 93-21. Cluster *4* was divided into two subclusters *c* and *d*. Subcluster *c* contained varieties DP Acala90, DP 4049, Des 119 [(Stoneville 603 x Delcot 277) x Des 2134], NTA 88-6 [(Allen x Half and Half x DPMA) x ISA 205] and NTA 93-15. This group contained varieties from Mali (NTA material developed from some introductions from Deltapine varieties). This indicated the expression of pedigree relationship using qualitative morphological characteristics. Subcluster *d* was divided into two subclusters. One subcluster contained

Guazuncho, McNair 235 (with Coker as genetic background), Delcot 344, Dixie King, SG 125 and Coker 315. The other subcluster contained Stoneville 506 and HC-B4-75 (HAR x Coker). Subcluster *d* contained varieties developed as Eastern and Western types in the USA (Niles and Feaster, 1984) and their pedigree relatives. Generally, main group *B* contained varieties from the USA and their relations from Mali and Argentina. Qualitative morphological characteristics for some varieties showed a relationship with variety background, while others did not express any relationship with their ancestral parents. For example, Reba W296 grouped separately from Reba B50 while both were sharing Allen in their pedigree (Figure 4.2). Grouping of varieties related to origin revealed four groups, namely West and Central Africa, Tanzania, Mali (grouped with Deltapine material) and the USA. These four groups represented two areas of origin; group *A* contained most varieties from Africa except for the Acala SJ$_2$ and Okra leaf and group *B* contained the USA and Mali varieties that have Deltapine traits from the USA in their pedigree (Figure 4.2).

4.4 DISCUSSION

Results indicated that most of the varieties with high boll weight had low number of bolls per plant and visa versa. Therefore as number of bolls increased, boll weight decreased, except in two varieties (Okra leaf and Des 119). A significant positive correlation of GOT with boll lint weight and lint yield confirmed results of Jixiang *et al*. (1996) and Hussain *et al.* (1998). Hundred seed weight decreased as the number of seeds per boll increased. Differences were significant except for Okra leaf that had the lowest 100 seed weight and a low number of seeds per boll. Plant height increased number of bolls per plant due to increased plant branches, increasing seedcotton yield. Since seedcotton yield is negatively correlated to some fibre characters like micronare value, medium plant height is important for improvement of other traits. Higher values for number of seeds per boll were linked to heavier bolls, higher seedcotton yield per hectare and lint per boll and confirmed results of Tang *et al*. (1996). A relationship was observed between seed weight and lint weight per boll. Varieties with heavy seeds had heavy lint weight, probably because the lint develops on the seed surface. Therefore, as seed surface increases, lint development per seed increases. This is in agreement with results of Kittock and Pinkas (1975). Therefore, boll weight, boll number, plant height and 100 seed weight contributed significantly to seedcotton yield per hectare. Results were

supported by Dedaniya and Pethani (1994). Fibre length and fibre strength were significantly reduced by a higher number of seeds per boll. A negative correlation was observed between plant height and GOT which supported previous reports by Christidis and Harrison (1955). Micronaire values increased as seeds per boll decreased. Results indicated the positive and negative correlations present in cotton characteristics.

Yield components and fibre quality are important characteristics, besides to the fact that most of them have negative correlations. According to Ibragmov (1989), short fibres have close genetic correlation with high fibre outturn and high cotton yield. In conventional breeding programmes, selection for high seedcotton yield leads to low GOT causing breeders to select varieties with medium boll size, small seeds per boll (for medium seedcotton yield) and maintaining high ginning percentage and good fibre quality. A meaningful cotton improvement programme focuses on lint yield and quality, without sacrificing one for the other. Gannaway (1982) cautioned breeders to have balance in varietal development as improvement of one property results in rapid development of deficiencies in other properties. The genetics of lint yield and quality are complex and many components contribute to their total expression and inheritance. When using conventional breeding firstly there exists a need for gene pyramiding through backcrossing and selection to a certain point that the variety cannot loose one of the important traits. The second option of improving varieties is to use advanced methods like genetic engineering to transfer specific genes of interest. Based on these quantitative traits it is difficult to understand the actual performance of the particular variety because influences of environmental factors limits parental selection (Swanepoel, 1999).

Fibre results obtained from this study indicated that all varieties had acceptable ranges for all fibre quality characteristics except for Reba W296 that had a fibre fineness of 3.4 which is lower than the acceptable value of 3.5. Varieties NTA 93-21, NTA 93-15, SG 125, High gossypol, Cyto 12/74, and Aubarn 56 had good fibre quality. Acala SJ$_2$, Des 119, Reba W296, and BJA 592 had low fibre quality. Most of these fibre quality characteristics are affected by the environment. This is reflected by the fact that Tanzanian varieties (UK91, UK82, MZ561, IL74, and IL85), which have good adaptability to the Tanzanian environment, had good fibre

maturity. Varieties with low seedcotton and lint yield, for example High gossypol and Okra leaf, had high fibre strength and fibre fineness values.

Qualitatively inherited morphological characteristics can be used to determine the relatedness between varieties. Transfer of genes from parents to the offspring can be revealed, as was observed in this study. For example the grouping of Tanzanian varieties (MZ 561, IL74, UK82 and UK91) with their parents (BJA 592 and Reba W296) was observed. The same was observed where varieties from Mali grouped together with Deltapine varieties, their ancestral material. Clustering based on qualitative morphological characteristics depend on the origin of material and on the traits involved. Therefore, morphological characteristics controlled by qualitative traits can be used to characterise varieties for collection and maintenance of germplasm and for parental selection for improving local varieties.

4.5 CONCLUSIONS AND RECOMMENDATIONS

Best performers were identified in this study, therefore, improvement of cotton varieties in Tanzania can be achieved by introducing GOT, seedcotton yield per hectare, lint yield per hectare, fibre quality and 100 seed weight traits by crossing the present local varieties with other varieties like NTA 93-21 and NTA 93-15 which exhibit these favourable characteristics. Therefore, selection for boll size and seed size could positively influence lint yield if a breeder selected for medium boll size, small seeds per boll and maintaining high ginning percentage and good fibre quality. Okra leaf and High gossypol, although having low yield, can be used to improve pest control, fibre strength, fibre maturity and fibre uniformity traits. Another trait that can be incorporated from Okra leaf is early maturity.

A negative correlation was revealed between yield components and fibre quality. The negative association of these characteristics reduced the speed of cotton variety development and release in Tanzania. Transferring of important traits by conventional breeding was used where linkage and epistasis effects affected the breeding programme. Parental selection for hybridisation using morphological traits, the method currently being used by breeders, is influenced by the environment.

The study indicated that qualitatively inherited morphological characteristics can be used to characterise varieties for collection and maintenance of germplasm and for parental selection through heterotic groups (groups with large distances between them) for improving local varieties. However, since morphological traits are influenced by the environment this has to be confirmed on DNA level.

The current study concluded that conventional breeding which has been used for cotton breeding programmes needs to be supplemented with genetic mapping and transformation to identify and transfer specific genes in order to hasten progeny selection for increased yield and fibre quality.

CHAPTER 5

DETERMINATION OF VARIATION BETWEEN COTTON VARIETIES USING AMPLIFIED FRAGMENT LENGTH POLYMORPHISM

5.1 INTRODUCTION

Genetic diversity within lines and populations is fundamental for breeding and germplasm conservation (Pillay and Myers, 1999; Lübberstedt *et al*., 1998; Rana and Bhat, 2004; Murtaza *et al*., 2005). Knowledge of genetic diversity among breeding material could help to avoid the risk for increasing uniformity in elite germplasm and could ensure long term selection gain (Messmer *et al*., 1993). This is because crossing of a limited number of elite lines creates the danger of loosing genetic diversity (Lübberstedt *et al.,* 1998; Gutierrez *et al.,* 2002). Thus, classification and assignment of breeding lines to establish heterotic groups are major decisions in any breeding programme (Smith and Smith, 1992).

Genetic diversity can be measured using morphological and molecular markers (Gepts, 1993). Morphological markers are influenced by the environment and expression is altered by epistatic and pleiotropic interactions (Tanksley, 1983; Farooq and Azam, 2002). Most morphological markers are quantitatively inherited and in many cases, the heterozygous condition of morphological traits is not identifiable (Tatinen *et al*., 1996). Protein markers, for example isozymes, that reveal polymorphism at protein level are known as biochemical markers. However, their use is limited due to a limited number of protein and isozyme markers in any crop species and due to subjection to post-translational modifications (Kumar, 1999; Farooq and Azam, 2002). DNA-based molecular markers reveal polymorphism at DNA level. DNA markers are free of pleiotropic effects, thereby allowing any number of markers to be monitored in a single population (Kumar, 1999). Morphological, isozyme and nuclear DNA markers are inherited in a Mendelian manner while cytoplasmic markers are maternally inherited. The number of morphological and isozyme markers is limited compared to DNA markers which are numerous (Tanksley, 1983).

In cotton there is not a sufficient number of morphological markers to provide detailed coverage of the entire genome (Tatinen *et al*., 1996). Genetic markers represent genetic variation, which makes it possible to determine relationships between different genotypes and forecast which pairings can produce new and superior gene combinations (Kumar, 1999).

Different molecular markers were used to characterise the cotton genome for genetic diversity (Rana and Bhat, 2004; Murtaza *et al.,* 2005). These include RFLP (Brubaker *et al.,* 1994), RAPD (Iqbal *et al.,* 1997), AFLP (Abdalla *et al*., 2001) and microsatellites or SSR (Liu *et al.*, 2003). RFLP and isozyme revealed low levels of polymorphism in cotton (Wendel *et al.,* 1992; Brubaker *et al.,* 1994).

AFLPs provide high levels of resolution allowing delineation of complex genetic structures (Powell *et al*., 1996). AFLP analysis detects a large number of polymorphisms (Krauss, 1999) that are distributed across the genome and have a high multiplex ratio and each fragment is assumed to originate from a different area of the plant genome (Rafalski *et al*., 1996).

Agricultural research institutes in Tanzania based parental selection for hybridisation on morphological characteristics. Morphological markers are few, influenced by environment and have difficulty differentiating heterozygotes from homozygotes. Furthermore, specific growth stages restrict collection of data. Most morphological markers are quantitatively inherited. Molecular marker analysis of available germplasm material is important for cotton improvement, as they are abundant and not influenced by the environment. Therefore the aim of this study was to determine the genetic diversity between 26 cotton varieties from Tanzania (five locally bred in Tanzania and 21 exotics) usually used as parents in breeding programmes. AFLP analysis was used to analyse genetic diversity and to determine heterotic groups of varieties that can be used for improvement of available cotton varieties as well as improving the germplasm conservation programme.

5.2 MATERIALS AND METHODS

5.2.1 Plant material

Seeds of 26 cotton varieties (Table 4.1) were obtained from the Ukiriguru Research Institute in Tanzania. The material have different characteristics including leaf shape, leaf colour, stem colour, yield and yield components (boll size, boll shape, number of bolls and shape of bolls), fibre quality (fibre length, fibre strength, fibre uniformity and micronaire value), resistance to diseases and insects, level of hairiness, GOT and seed characteristics (Table 4.1).

Two plants were grown in two pots per variety in a glasshouse at the University of the Free State (UFS) in Bloemfontein, South Africa and at the Mikocheni Research Institute, Dar es salaam, Tanzania for DNA extraction. The soil was mixed with NPK 3:2:1+0.5 Zn fertilizer. Conducive germination and growing temperatures for day (28^0C) and night (16^0C) were maintained. Spraying of insecticide to control spider mites and aphids was done.

5.2.2 DNA extraction

Each variety was represented by two plants. DNA extraction was done using a modified monocot extraction procedure (Edwards *et al*., 1991). Young fresh leaves were collected on ice and frozen in liquid nitrogen and ground into a fine powder with a mortar and pestle. The ground powder was transferred to a clean 50 ml polypropylene centrifuge tube. A volume of 10 ml extraction buffer [0.1 M Tris(hydroxymethyl)aminomethine hydrochloric acid (Tris-HCl) pH 8.0, 0.05 M EthyleneDiamineTetraacetate (EDTA), 1.25% (w/v) Sodium dodecyl sulphate (SDS), 0.5 M NaCl, 0.2% (v/v) ß-mercaptho-ethanol and 0.4 M urea] preheated at 65^0C was added. A volume of 1 ml of Cetyltrimethylammonium bromide (CTAB) buffer (0.2 M Tris-HCl pH 8.0, 0.05 M EDTA and 10% (w/v) CTAB) and 2 ml of 5 M NaCl were added to the homogenate and incubated at 65^0C for 60 min with shaking every 10 to 20 min. Chloroform: isoamyl alcohol (24:1 v/v) (10 ml) was added, mixed and centrifuged at 10000 rpm for 15 min. The supernatant was transferred to a new tube and this step was repeated until the interface was clean. DNA was precipitated with two volumes (v/v) 100% ice cold ethanol, followed by overnight incubation at 4^0C. The precipitated DNA was spooled out with a sterile pasteur pipette and washed three times in 70% (v/v) ethanol. The DNA pellet was resuspended in 250 µl sterile water and stored at 20^0C.

Genomic DNA was treated with 0.4 mg/ml DNase free RNase through incubation for 2 h at 37^0C. DNA was treated with 0.75 M ammonium acetate and an equal volume of chloroform: isoamyl alcohol (24:1 v/v) and mixed gently, followed by centrifugation at 10000 rpm for 3 min. DNA was precipitated overnight from the aqueous phase with 500 µl ice cold 100% ethanol. After centrifugation at 10000 rpm for 15 min, DNA was washed twice with 70% (v/v) ethanol by centrifuging at 10000 rpm for 10 min, followed by air drying of the pellet. The pellet was resuspended in 50 µl TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) buffer and incubated overnight at 4^0C.

DNA concentration and purity was determined using a spectrophotometer (Hitachi U 2000) by measuring absorbances at 260 nm and 280 nm. The DNA concentration was calculated using the formula, [DNA] =Optimal density at 260 nm x dilution factor x constant (50 µg/ml) and DNA purity was determined using the 260/280 nm ratio. The quality, integrity and concentration of the DNA was confirmed by electrophoresis in a 0.8 % (w/v) agarose gel for 45 min at 60 volts with visualisation under UV light after staining with ethidium bromide. DNA samples were diluted to a working concentration of 200 ng/µl and stored at 4^0C.

5.2.3 AFLP analysis

AFLP reactions were done according to Herselman (2003). DNA was digested using *Eco*RI (rare 6-base cutter) and *Mse*I (frequent 4-base cutter) as described by Vos *et al.* (1995). Primer combinations of *Eco*RI and *Mse*I were represented as E- and M- respectively and are given in Table 5.1.

Restriction digestion and ligation of adapters

Genomic DNA (1 µl) was digested using 4 U *Mse*I and 1 x *Mse*I-buffer in a total volume of 50 µl for 5 h at 37^0C. Genomic DNA was further digested using 5 U *Eco*RI and NaCl to a final concentration of 100 mM and incubated overnight at 37^0C. Adapter ligation of the digested DNA was established by adding a solution containing 50 pmol *Mse*I-adapter, 5 pmol *Eco*RI-adapter, 1 U T4 DNA Ligase, 0.4 mM ATP and 1 x T4 DNA Ligase buffer followed by incubation at 16^0C overnight.

Preamplification reactions

Preamplification reactions were established in 50 µl reaction mixtures containing 5 µl restricted and ligated template DNA, 30 ng *Eco*RI primer + 1 and 30 ng of *Mse*I primer + 1 (Table 5.1), 1 x Promega polymerase buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% Triton x-100), 2 mM MgCl$_2$, 200 µM of each dNTP and 1 U *Taq* DNA polymerase (Promega, Madison, WI, USA). Amplification was performed using the following cycling profile: 30 cycles of 30 s at 94^{0}C, 60 s at 56^{0}C and 60 s at 72^{0}C. Amplification reaction quality and quantity were confirmed by electrophoresis in 1.5 % (w/v) agarose gels. Products were diluted accordingly using TE buffer (1:5, 1:10 or 1:20) before selective amplification.

Selective amplification

Selective amplification was carried out in a volume of 20 µl containing 5 µl diluted preamplification product, 30 ng *Mse*I-primer +3, 30 ng *Eco*R1 Fam or Ned - primer +3 (Table 5.1), 1 x Promega polymerase buffer, 2 mM MgCl$_2$, 200 µM of each dNTP, 100 µg/ml bovine serum albumin and 0.75 U Promega *Taq* DNA polymerase. Selective amplification was performed using the following cycling programme: one cycle of denaturation at 94^{0}C for 5 min followed by one cycle for 30 s at 94^{0}C, 30 s at 65^{0}C and 60 s at 72^{0}C. The annealing temperature was reduced by 1^{0}C per cycle during the next eight cycles, then 25 cycles were performed at 94^{0}C for 30 s, 56^{0}C for 30 s and 72^{0}C for 60 s followed by one last elongation for 5 min at 72^{0}C.

PCR products were prepared for capillary electrophoresis by adding 5 µl Fam labelled PCR product, 5 µl Ned labelled PCR product of each selective reaction, 99.5% deionised formamide and 6 fmol GENESCAN-1000 Rox^{TM} size standard. Samples were denatured at 94^{0}C for 10 min followed by quick cooling in ice slush. AFLP fragments were resolved using a Perkin Elmer Prism ABI 310 automated capillary sequencer (PE Biosystems, 2002). A total of eight primer combinations were screened. AFLP primer combinations were selected based on literature. Primer combinations and adapters used for screening 26 cotton varieties are presented in Table 5.1.

Table 5.1. Adapter and primer sequences used for fingerprinting 26 cotton genotypes

*Mse*I-adapter	*Eco*RI-adapter
5'-GACGATGAGTCCTGAG-3'	5'-CTCGTAGACTGCGTACC-3'
3'-TACTCAGGACTCAT-5'	3'-CATCTGACGCATGGTTAA-5'
*Mse*I-primer + 1	*Eco*RI-primer + 1
(5'-GATGAGTCCTGAGTAAC-3')	5' GACTGCGTACCAATTCA-3'
*Mse*I-primers + 3	*Eco*RI-primers + 3
(5'-GATGAGTCCTGAGTAA-NNN -3')	(5'-GATGCGTACCAATTC-NNN -3')
*Mse*I+CAT	*Eco*RI+ACA/FAM
*Mse*I+CTG	*Eco*RI+AAC/NED
*Mse*I+CTA	*Eco*RI+ACT/FAM
*Mse*I+CAC	*Eco*RI+ACC/NED

5.2.4 Data scoring and statistical analysis

AFLP fragments data for selected primer combinations were coded using a binary unit character (1 for presence and 0 for absence). Data was summarised in a data matrix for all varieties based on both unique and shared fragments. Fragments smaller than 40 bp were excluded from the data matrix. The coefficient of genetic similarities between all pairwise comparisons were computed using Dice coefficient (Dice, 1945) using the SIM-QUAL (similarity of qualitative data) programme of Numerical Taxonomy Multivariate Analysis System (NTSYS-pc) version 2.02i software package (Rohlf, 1993). Dice coefficient $GS= 2a/(2a+b+c)$ where a is the number of DNA fragments present in both variety x and variety y; b and c represent the total number of DNA fragments in variety x and variety y with regard to all primer combinations. Cluster analyses were performed using UPGMA clustering (Sokal and Michener, 1958) and utilised to construct a dendrogram using the SAHN programme of NTSYS-pc. For each dendrogram, co-phenetic coefficients between the matrix of genetic similarities and the matrix of co-phenetic values were computed using appropriate routines of the COPH and MXCOMP programme of NTSYS-pc. The significance of the co-phenetic correlation observed was tested using the Mantel matrix correspondence test (Mantel, 1967).

Calculations for polymorphic information content (PIC) was done using the formula of the expected heterozygosity (Smith *et al.,* 2000) as: $PIC=1-\sum f^2 i$, where f is the percentage of

genotypes in which the fragment is present. This was used to identify primers that would distinguish varieties most efficiently.

5.3 RESULTS

5.3.1 Primer combinations and fragments

Eight selected AFLP primer combinations produced a total of 835 fragments varying in size from 40 to 538 bp, with an average of 104 bp per primer combination. Examples of electropherograms for one of the cotton varieties are given in Figure 5.1. A total of 309 fragments were polymorphic with an average of 39 polymorphic fragments per primer combination equivalent to 37% polymorphisms. Primer combinations M-CAT/E-AAC, M-CAT/E-ACA and M-CTA/E-ACT produced the highest numbers of amplified fragments compared to other combinations (132, 126 and 119 respectively) (Table 5.2).

M-CTG/E-ACC (76) amplified the lowest number of fragments compared to other primer combinations followed by M-CAC/E-ACC (95) and M-CAC/E-ACT (96). Even though some of the primer combinations amplified low numbers of fragments, they were able to distinguish some of the varieties. For example, M-CAC/E-ACT uniquely identified Delcot 344 and M-CTG/E-ACC uniquely identified High gossypol and Delcot 344. Primer combination M-CAT/E-AAC uniquely identified eight varieties followed by M-CTA/E-ACT (6), M-CAT/E-ACA (4), M-CTA/E-ACC (4) and M-CAC/E-ACC (4). Delcot 344 was uniquely identified from other varieties by almost all primer combinations. High levels of polymorphism were observed for primer combinations M-CAC/E-ACC (51.6%), M-CTG/E-ACT (45.5%) and M-CTA/E-ACC (39.4%) (Table 5.2).

PIC values, measuring the gene diversity for a specific locus, are presented in Table 5.2. PIC value is an indication of a high probability of obtaining polymorphism using that primer combination. PIC values ranged from 0.37-0.57 with an average of 0.47. M-CAC/E-ACC (0.57), M-CTG/E-ACT (0.54) and M-CTA/E-ACC (0.49) were identified for having high levels of PIC. M-CAC/E-ACT had the lowest PIC value (0.37).
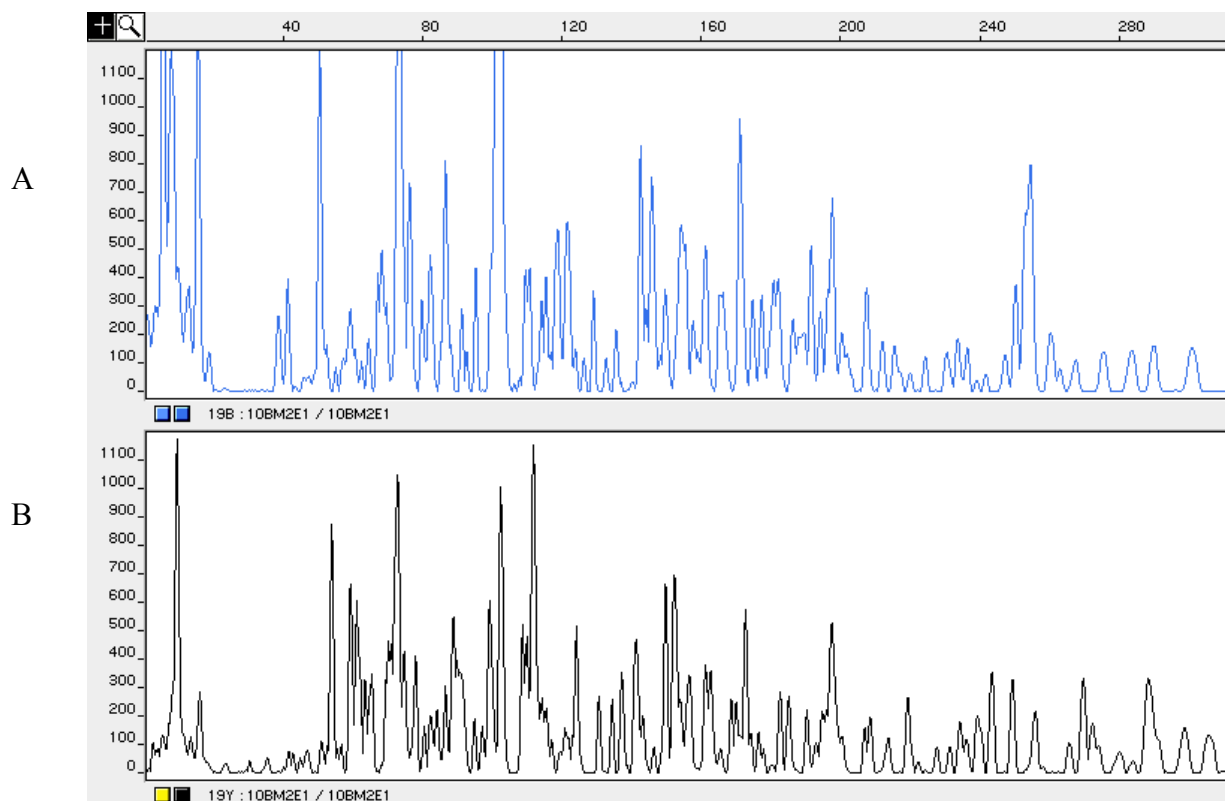
Figure 5.1 Examples of electropherogram AFLP fragments using Perkin Elmer Prism ABI 310: A) = Cotton variety McNair 235 DNA amplified using AFLP primer *EcoR*I + ACC Fam/*Mse*I + CAT and B) = Cotton genotype McNair 235 DNA amplified using AFLP primer *EcoR*I + AAC Ned/*Mse*I + CAT

Table 5.2 Information generated using 26 varieties and eight AFLP primer combinations

Primer combinations	Fragments range (bp)	Total fragments	Polymorphic fragments	% Polymophism	PIC	Unique Varieties
M-CAT/E-ACA	40-472	126	40	31.7	0.46	4
M-CAT/E-AAC	51-538	132	33	25.0	0.43	8
M-CTG/E-ACC	44-421	76	28	36.8	0.47	2
M-CTG/E-ACT	53-523	100	45	45.5	0.54	2
M-CTA/E-ACT	40-524	119	42	35.3	0.43	6
M-CTA/E-ACC	53-454	109	43	39.4	0.49	4
M-CAC/E-ACT	41-466	96	29	30.2	0.37	1
M-CAC/E-ACC	52-448	95	49	51.6	0.57	4
Total		835	309			
Average		104	39	37	0.47	

bp = base pair, PIC = polymorphic information content

5.3.2 Estimates of genetic distance

Dice genetic similarity estimates for each of the 26 cultivars are presented in Table 5.3. Pair-wise AFLP similarity estimates between varieties ranged from 0.894-0.979, with an average of 0.939. Genetic similarities estimates for 325 pair wise comparisons based on AFLP analysis displayed a normal distribution (Figure 5.2) with similarity values and frequency percentage values of 0% (0.800-0.840), 0% (0.841-0.880), 11.4% (0.881-0.920), 81% (0.921-0.960), 8.6% (0.961-1.000) and 0% (1.001-1.040).

Genetic similarities were high between some of the varieties. For example, 0.979 between McNair 235 and MZ561, 0.978 between Frego bract and Reba W296 and 0.977 between SG 125 and DP 4049. The lowest similarity value was observed between High gossypol and Cyto 12/74 (0.894), indicating that these two were the most diverse varieties. Generally, High gossypol, Cyto 12/74, Delcot 344, Okra leaf and Reba B50 had low similarities with the other varieties ranging from 0.894-0.940 (Table 5.3).

5.3.3 Cluster analysis

The dendrogram from UPGMA cluster analysis based on AFLP markers revealed two major groups *A* and *B* (Figure 5.3). Major group *A* contained Delcot 344 and group *B* contained the other 25 varieties. Delcot 344 has distinctive characteristics including reddish green coloured leaves with no leaf hairs and originated from the USA (Figure 5.3 and Table 5.3). The separate clustering indicated the high divergence of Delcot 344 from the rest.

The second major group (*B*) divided into two subgroups that ultimately contained four clusters. Cluster *I* contained 12 varieties with genetic similarity estimates ranging from 0.979 between McNair 235 and MZ561 to 0.956 between Aubarn 56 and HC-B4-75. This cluster was further divided into two subclusters. The upper most subcluster divided into two groups. The first contained HC-B4-75 (drought tolerant and susceptible to fusarium wilt), DP 4049, SG 125 and NTA 88-6. This group contained varieties from the USA except for NTA 88-6, which is from Mali, but has traits from Deltapine varieties (from the USA) in its pedigree.

Table 5.3. Genetic similarity estimates for 325 pair wise comparisons for 26 cotton varieties based on AFLP analysis

	HCB	RBW	FRB	DK	DP4	SG	CYT	RB5	N88	HGP	DES	MZ	McN	IL8	IL7	GUA	DEL	STN	AUB	N93	BJA	UK9	IRM	UK8	OKL
HCB																									
RBW	0.961																								
FRB	0.946	0.978																							
DK	0.954	0.960	0.953																						
DP4	0.970	0.953	0.946	0.953																					
SG	0.962	0.951	0.944	0.958	0.977																				
CYT	0.921	0.943	0.954	0.934	0.920	0.912																			
RB5	0.943	0.951	0.946	0.964	0.939	0.943	0.946																		
N88	0.957	0.944	0.937	0.949	0.964	0.967	0.910	0.943																	
HGP	0.936	0.924	0.918	0.931	0.937	0.935	0.894	0.925	0.945																
DES	0.944	0.941	0.947	0.937	0.948	0.939	0.923	0.927	0.950	0.936															
MZ	0.950	0.954	0.953	0.936	0.950	0.937	0.933	0.929	0.941	0.932	0.970														
McN	0.944	0.943	0.952	0.934	0.943	0.937	0.925	0.922	0.942	0.935	0.975	0.979													
IL8	0.953	0.939	0.927	0.946	0.957	0.952	0.910	0.939	0.964	0.942	0.956	0.953	0.953												
IL7	0.949	0.936	0.928	0.945	0.958	0.947	0.907	0.932	0.949	0.936	0.943	0.946	0.940	0.966											
GUA	0.945	0.927	0.919	0.944	0.947	0.947	0.900	0.924	0.945	0.950	0.939	0.937	0.940	0.957	0.953										
DEL	0.920	0.921	0.927	0.918	0.922	0.915	0.923	0.911	0.911	0.909	0.937	0.936	0.936	0.917	0.919	0.921									
STN	0.955	0.938	0.931	0.951	0.960	0.953	0.911	0.934	0.952	0.945	0.949	0.949	0.946	0.964	0.967	0.970	0.927								
AUB	0.939	0.938	0.940	0.934	0.946	0.938	0.918	0.920	0.935	0.923	0.947	0.955	0.949	0.943	0.952	0.940	0.940	0.958							
N93	0.947	0.929	0.917	0.936	0.949	0.940	0.901	0.920	0.940	0.930	0.926	0.934	0.930	0.946	0.949	0.944	0.920	0.951	0.938						
BJA	0.948	0.935	0.925	0.940	0.947	0.943	0.903	0.927	0.938	0.934	0.929	0.935	0.931	0.945	0.947	0.945	0.922	0.953	0.951	0.964					
UK9	0.947	0.927	0.926	0.934	0.952	0.940	0.900	0.920	0.941	0.938	0.932	0.935	0.934	0.941	0.950	0.943	0.918	0.946	0.937	0.954	0.954				
IRM	0.937	0.931	0.939	0.931	0.940	0.930	0.918	0.918	0.932	0.927	0.940	0.952	0.947	0.938	0.942	0.939	0.935	0.941	0.946	0.937	0.947	0.955			
UK8	0.950	0.932	0.925	0.939	0.953	0.947	0.903	0.929	0.947	0.940	0.930	0.938	0.931	0.949	0.953	0.950	0.915	0.958	0.948	0.952	0.965	0.957	0.955		
OKL	0.941	0.936	0.929	0.943	0.938	0.936	0.912	0.935	0.935	0.925	0.920	0.929	0.921	0.938	0.939	0.933	0.909	0.940	0.931	0.942	0.952	0.947	0.941	0.961	
ACL	0.943	0.934	0.923	0.938	0.944	0.944	0.907	0.925	0.937	0.928	0.927	0.936	0.927	0.945	0.952	0.937	0.916	0.945	0.940	0.947	0.953	0.952	0.946	0.958	0.945

HCB= HC-B4-75, RBW= Reba W296, FRB= Frego bract, DK= Dixie King, DP4= DP 4049, SG= SG 125, CYT= CYTO 12/74, RB5= Reba B50, N88= NTA 88-6, HGP= High gossypol, DES= DES 119, MZ= MZ561, McN= McNair 235, IL8= IL85, IL7= IL74, GUA= Guazuncho, DEL= Delcot 344, STN= Stoneville 506, AUB= Aubarn 56, N93= NTA 93-15, BJA= BJA 592, UK9=UK91, IRM= Irma 1243, UK8= UK82, OKL= Okra leaf, ACL= Acala SJ$_2$
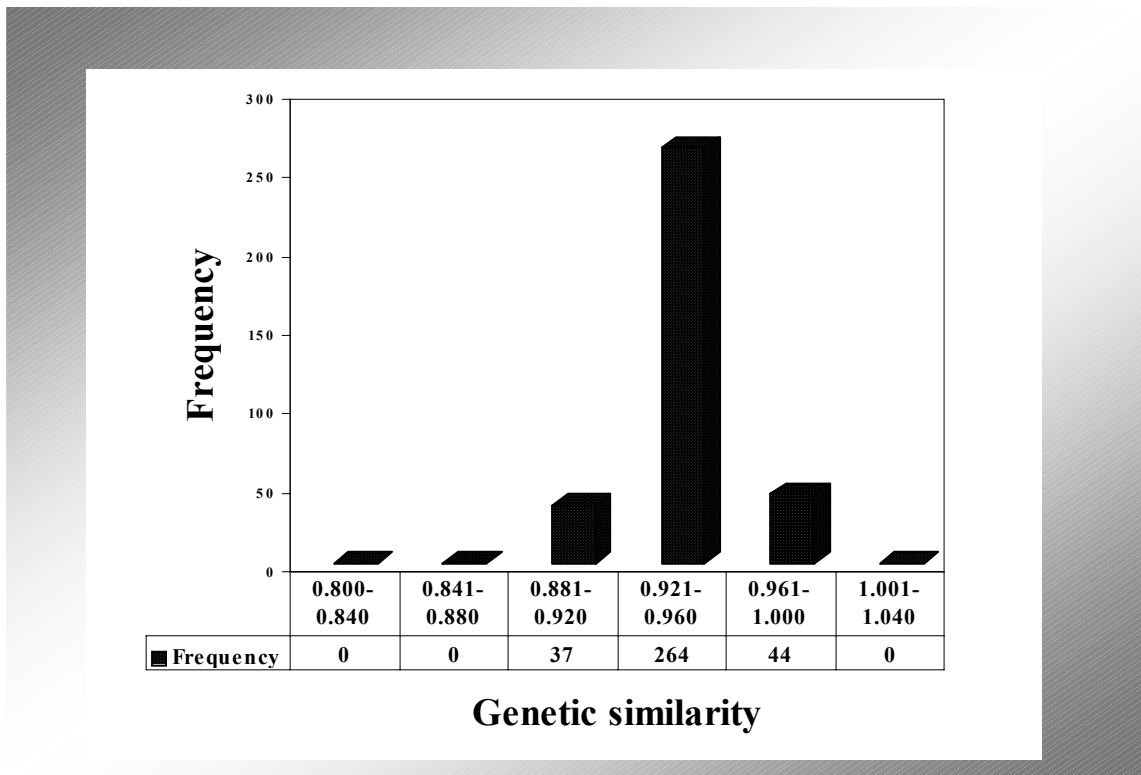
Figure 5.2 Frequency distribution of 325 pairwise genetic similarities obtained from AFLP data
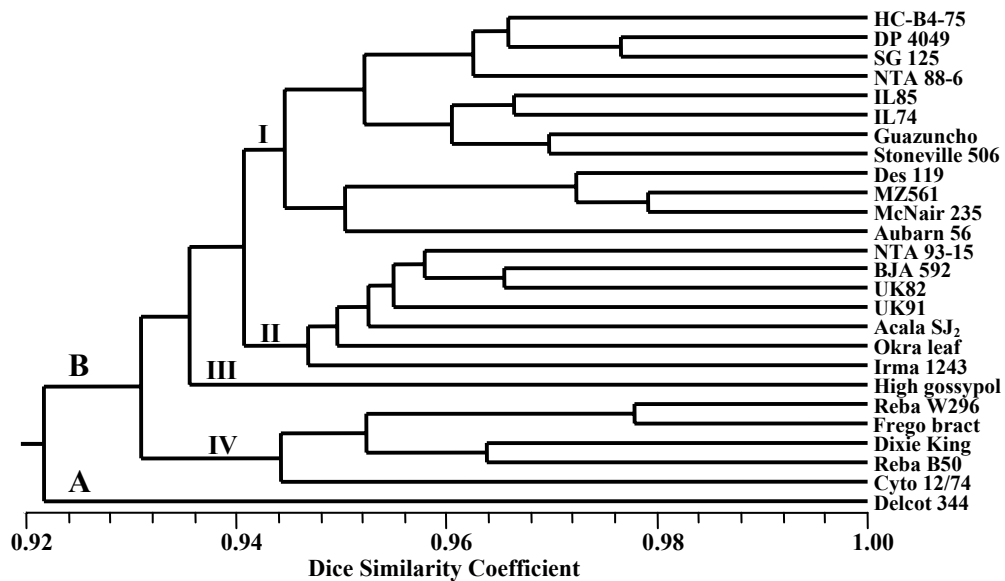
Figure 5.3 Dendrogram generated based on UPGMA clustering method and Dice coefficient using AFLP analysis among 26 cotton varieties

These varieties had high GOT values ranging from 40.5% to 43.9% (Table 4.3). The second group contained four varieties, Guazuncho (from Argentina, tolerant to drought), Stoneville 506 (resistant to bacterial blight from the USA), IL74 and IL85 (resistant to bacterial blight from Tanzania). The second subcluster contained four varieties, McNair 235, Des 119, Aubarn 56 (all from the USA and resistant to fusarium wilt) and MZ561 (from Tanzania).

Cluster *II* contained seven varieties; NTA 93-15, BJA 592 (short staple), UK82, UK91, Acala SJ$_2$ (large bolls), Okra leaf (okra leaf type and early maturing) and Irma 1243. NTA 93-15, BJA 592 and Irma 1243 originated from West/Central Africa (might have shared some genes). NTA 93-15 and Irma 1243 are susceptible to bacterial blight and fusarium wilt and have high GOT values. BJA 592, UK82 and UK91 are resistant to bacterial blight. UK82 and UK91 are Tanzanian varieties for the WCGA's clustered with BJA 592, their ancestor for bacterial blight resistance.

Cluster *III* contained High gossypol [(A333xFoster) x Allen MP-2 (a selection from Zaria Allen)] from Chadi and has resistance to insects due to high gossypol content. (Figure 5.3).

Cluster *IV* was composed of five varieties; Frego bract (insect resistant) and Reba W296 (Coker 100 x Allen 51-296) clustered together with a genetic similarity of 0.978. Dixie King (resistant to fusarium wilt) and Reba B50 (Stoneville B 1439 x A50T) clustered together. Cyto 12/74 (from Pakistan) joined them as a separate group with a genetic similarity of 0.944. RebaW296 and Reba B50 are bacterial blight and fusarium wilt resistant, have weak fibres and both originated from Central Africa.

Results from Table 5.3 and Figure 5.3 indicated that McNair 235 and MZ561, SG 125 and DP 4049, Frego bract and Reba W296, Des 119 and McNair 235 and Des 119 and MZ561 were the most similar varieties with genetic similarity values of 0.970 to 0.979. High gossypol and Delcot 344 were genetically the most distinctly related varieties. These varieties were morphologically different from the rest and were grouped in separate clusters. High gossypol has high gossypol content for insect resistance and has poor yield components compared to other varieties (Chapter 4). Delcot 344 has reddish green and smooth leaves, with high GOT

values and moderate yield components. Other distinctly related varieties included Okra leaf and Cyto 12/74.

5.4 DISCUSSION

The application of DNA-based technology in the assessment of genetic diversity of germplasm improves the precision and efficiency of parental selection in breeding programmes (Falconer and Mackay, 1996). The study revealed high similarity values among 26 selected cotton varieties, ranging from 0.894-0.978. These results were similar to results of Multani and Lyon (1995), Iqbal *et al.* (2001) and Lu and Myers (2002), confirming limited genetic diversity in cultivated cotton. Iqbal *et al*. (2001) hypothesised that the lack of diversity is due to a genetic bottleneck that occurred upon importation of small quantities of seed from Mexico in the 19th century. This genetic bottleneck was aggravated during the later stages of development of *G. hirsutum latifolium* through selection for early maturity (Lewis, 1962).

Based on genetic similarity values and the dendrogram, similar results were obtained by Lu and Myers (2002) on grouping McNair 235 with Stoneville 506 in the same cluster. These two varieties had a similarity value of 0.949 and clustered in main group *B* in subcluster *I*. Lu and Myers (2002) identified Delcot 344 to have least genetic similarity with all varieties, which reflected relative high genetic distances from other varieties. These results were similar to results obtained during this study as Delcot 344 grouped as an independent main group *A*. High gossypol clustered to a distinct cluster *III* in main group *B*. This might be due to its unique high gossypol content and probably other unique characteristics. Varieties like Cyto 12/74 and Okra leaf clustered with other varieties but with low similarity values. Genetic clustering of cotton varieties confirmed known morphological and agronomical traits data. Varieties with unique characteristics clustered separately. For example, Delcot 344 had high GOT values and reddish green leaves with few hairs, High gossypol had high gossypol content and poor yield components, while Cyto 12/74 were tall plants with good yield components but low GOT values and Okra leaf had an okra leaves type and poor yield components. These results confirmed the results reported by Roldan-Ruiz *et al*. (2000) that AFLP clustering was related to morphological traits of the genotypes evaluated.

Results for different primer combinations indicated that M-CAC/E-ACC and M-CTG/E-ACT had high polymorphic percentages (51.6% and 45.5% respectively) as well as high PIC values (0.57 and 0.54 respectively). This indicated that a primer combination can amplify a low number of fragments but reveal high polymorphism as indicated by M-CAC/E-ACC that had a high polymorphism percentage and PIC but lowest number of fragments. These primer combinations were reported by Abdalla *et al.* (2001) and Rana and Bhat (2004) as good primer combinations with a high resolution power that can be used for variety identification purposes in cotton. All primer combinations amplified 76 or more fragments and detected different levels of polymorphism.

All 26 cotton varieties were uniquely identified using eight AFLP primer combinations (Figure 5.3) despite of sharing of the same ancestors and low levels of genetic dissimilarity. AFLP analysis distinguished varieties according to traits contained in that variety. This led to the identification of different possible heterotic groups. HC-B4-75 and Guazuncho had different origins (Table 4.1) but are drought tolerant and grouped in cluster *I*. Varieties were furthermore distinguished according to known pedigree and origin data. Varieties with the same ancestor grouped together. For example, UK82 and UK91 grouped with their ancestor BJA 592 and Des 119, produced from crosses between Aubarn 257 x Stoneville 7 x Delcot 277 (Van Esbroeck and Bowman, 1998), grouped with Aubarn 56 and Stoneville 506 in cluster *I*. Poisson *et al.* (2003) reported the relationship of Guazuncho, Stoneville and Deltapine varieties, in their pedigree. These varieties grouped together in the current study. AFLP analysis was able to distinguish closely related Tanzanian varieties, for example UK82 and UK91 from IL74 and IL85. These varieties share the same ancestors, but have different environmental adaptabilities and resistances to fusarium wilt. Based on origin, Des 119, SG 125 and DP 4049 (all from the USA) clustered together in cluster *II* and contained Deltapine traits. Reba W296, Reba B50 and Frego bract clustering together in cluster *IV* originated from Central/West Africa with low GOT values (34.55% to 37.23%) (Table 4.3).

5.5 CONCLUSIONS AND RECOMMENDATIONS

Reduced genetic diversity of *G. hirsutum* as observed among the 26 studied varieties, emphasises the need to focus on introduction of more diverse cultivated cotton varieties from other countries into Tanzania. Introduction of germplasm resources should include other tetraploid species (*G. barbadense)* to enable improvement of the available material through hybridisation.

The heterotic groups identified could be used for improving cotton breeding programmes through hybridisation. Since Delcot 344 clustered separately, it could be crossed with any variety especially for improvement of high GOT values. Any variety in cluster *I* can be crossed with varieties in clusters *II, III* and *IV* as they are distantly related. Pejic *et al*. (1998) recommended the use of crosses between heterotic groups because these crosses are expected to perform better compared to crosses within heterotic groups. However, further molecular marker studies is important before marker-assisted selection breeding can be used for improvement.

Results indicated that AFLP analysis is a sensitive technique for detecting molecular markers for genetic studies in cotton. AFLP presents a better opportunity to detect polymorphisms among closely related genotypes (Pillay and Myers, 1999). Results obtained in this study on genetic analysis revealed a genetic similarity average of 0.939, indicating a narrow genetic base. Apart from the narrow genetic diversity present in cotton, because of its high multiplex ratio and robustness (Rana and Bhat, 2004), AFLP analysis managed to distinguish all varieties. AFLP is a promising marker system especially for cotton varieties which are closely related, with low levels of polymorphism and thus high levels of genetic uniformity.

AFLP is a useful molecular marker system to study genetic diversity and identify heterotic groups for cotton breeding programmes. However, based on reports from literature, SSR analysis is also effective as a marker system to reveal cotton genetic diversity. Since no SSR primers were screened during this study, future research should be conducted to compare the efficiency of AFLP analysis to detect genetic diversity in Tanzania cotton genotypes with other molecular marker systems, including SSR. The focus will be on improving cotton

germplasm through marker-assisted selection breeding in order to improve diversity for the cotton industry in Tanzania.

Future research, is needed to identify molecular markers linked to important traits and locating quantitative trait loci (QTL). Appropriate mapping populations must be developed for markers linked to traits, followed by screening of markers for usefulness in development of cotton varieties with improved yield, resistance (pests and diseases), fibre quality and adaptability to the environment.

CHAPTER 6

COMPARISON OF AGRONOMICAL, MORPHOLOGICAL AND AFLP MARKERS FOR COTTON GENETIC DIVERSITY STUDIES

6.1 INTRODUCTION

Characterisation and maintenance of crop germplasm are important for continuous supply of genetic variability for crop improvement (Ali *et al*., 2003). Diversity analysis of cotton (*G. hirsutum*) is an important component for efficient management and utilisation and identification of genetic relatedness of available genetic resources (Kumar, 1999). Future improvements of environmental resistance and cotton quality depend on diversity within genetic resources from which traits can be selected (Stewart, 1995). Cultivated cotton displays low levels of genetic diversity (narrow genetic base) (Pillay and Myers, 1999; Abdukarimov *et al*., 2003). Meredith (1998) stated that unless breeders broaden the genetic base of breeding programmes, yield stagnation will continue in the USA. This might apply to breeding programmes in Tanzania as well. The procedures for the International Union for the Protection of New Varieties of Plants (UPOV, 1991) and plant breeders rights (PBR) are based on criteria of distinctiveness, uniformity and stability (DUS) of genotypes as descriptors or genetic markers. There exists a need for cultivar specific DNA markers in cotton programmes for cultivar registration, plant patents and breeders' right protection as well as early detection of agronomical and economical traits (Pillay and Myers, 1999; Royo *et al*., 2003).

According to Arus and Gonzalez (1993) properties for good quality markers include (1) easy expression during the development of the plant, (2) easy recognition of all possible phenotypes (homozygotes and heterozygotes), (3) no effect on plant morphology of alternate alleles at marker loci and (4) low or null interaction among markers, allowing simultaneous use of multiple markers in segregating populations. Morphological (leaf shape, hairlines, boll size and boll shape) and agronomical (seed cotton yield, lint lint and fibre quality) characteristics for cotton have traditionally been used to distinguish varieties and provide useful information to users. However, the expression of the majority of these characteristics is significantly influenced by the environment causing problems for consistent identification (Lukonge and Ramadhani, 1999). Morphological markers cannot distinguish heterozygotes and

is time-consuming (Kumar, 1999; Swanepoel, 1999; Rungis *et al*., 2000). Depending on the insect population, 5-32% cross pollination is expected for cotton, which may lead to pollen contamination, adding to difficulties in genetic uniformity and stability assessment. These factors limit the use of morphological markers compared to molecular markers that are numerous and some molecular markers can identify heterozygotes (Meredith, 1995b). Molecular markers provide a number of practical applications including variety identification through DNA fingerprinting, development of genetic maps facilitating indirect selection of economical traits like disease resistance, cloning of important genes and in evolutionary and phylogenetic studies (Guthridge *et al.,* 2001; Altaf Khan *et al*., 2002).

Adugna (2002) used morphological and AFLP characterisation methods to study genetic diversity in *Linum usitatissimum* L. (linseed). Substantial differences based on clusters were observed that indicated different hierarchal patterns among accessions and a weak correlation (0.076) between the two genetic diversity matrices. Roldan-Ruiz *et al.* (2001) compared AFLP analysis and morphological characterisation among varieties of perennial ryegrass. Results indicated that these two methods were different, however some consistencies were found for closely related material. Ben-Har *et al*. (1995) and Bustin and Charcosset (1997) obtained similar results on maize.

The study of Van Esbroeck *et al.* (1999) on genetic similarity indices from agronomical and morphological traits for cotton revealed that several ancestral cultivars, developed from the earliest introductions from the USA, were similar to the most recent introductions from Mexico. This suggested that the original Mexican introductions were genetically similar. Bie *et al.* (2001) used RAPD analysis and stable and highly heritable agronomical traits in cotton to study genetic diversity and the relationship between the two methods. Classification of all genotypes based on the two methods gave similar results with a correlation of 0.76 between genetic and taxonomic distances. Federici *et al.* (2001) reported similar results on weedy rice based on AFLP analysis and morphological characteristics (colour and awns). AFLP analysis and morphological characterisation grouped cultivars in the same groups and cultivars were grouped according to geographic origin. However, Wu *et al.* (2001) reported differences between DNA markers and phenotypic characteristics (agronomical and fibre quality) for genetic diversity studies on cotton.

Tanzanian Agricultural Research Institutes use agro-morphological characteristics for parental selection for hybridisation. Therefore the main objectives of this study were to assess the genetic variation among 26 cotton varieties using agronomical, morphological and AFLP markers and compare the efficiency of these characterisation methods in cotton genetic diversity studies.

6.2 MATERIALS AND METHODS

6.2.1 Plant material
Treatments and experimental designs for agronomical, morphological characteristics and AFLP analysis studies are given in sections 4.2.1 and 5.2.1 respectively.

6.2.2 Agronomical and morphological data collection
Data collection procedures for agronomical and morphological characteristics are described in section 4.2.2.

6.2.3 DNA extraction
DNA extractions were done according to Edwards *et al.* (1991) as described in section 5.2.2.

6.2.4 AFLP analysis
AFLP analysis that included restriction, digestion and ligation of adapters, preamplification and selective amplification were performed according to Herselman (2003) as described in section 5.2.3.

6.2.5 Genetic similarities, clustering and spearman correlation analysis
Agronomical data on seed cotton yield, GOT, bolls per plant and fibre quality (fibre length, fibre strength and micronaire value) were grouped into three classes as low, medium and high using the following formula: Highest-lowest/3 = x: Lowest $\leq$ lowest + x; Medium > lowest and $\leq$ lowest +2x; Highest > lowest + 2x. For example, for yield the highest value was 3.1 ton/ha and the lowest was 1.71 ton/ha, therefore the lowest was $\leq$ 2.17 ton/ha, medium > 2.17 ton/ha and $\leq$ 2.62 ton/ha and highest > 2.62 ton/ha. Data were coded into a binary form and scored as present (1) and absent (0) and entered into the data matrix. AFLP fragments and

morphological data were coded as 1 for present and 0 for absent (sections 4.2.2 and 5.2.4). Coded data were subjected to analysis using the NTSYS-pc version 2.02i (Rohlf, 1993) computer programme. Similarity matrices were compiled for all pairs of varieties using Dice similarity coefficient. Cluster analysis was performed to reveal the pattern of genetic relationships among varieties using UPGMA clustering. Co-phenetic values for all methods (agronomical, morphological and AFLP) were performed as described in section 5.2.4. The NCSS computer package (Hintze, 2000) was used to determine the Spearman's rank correlation coefficient between agronomical, morphological and AFLP genetic similarities. Principle component analysis (PCA) biplots for agronomical, morphological and AFLP methods were performed using the NTSYS-pc computer package version 2.02i (Rohlf, 1993).

6.3 RESULTS

Morphological characterisation in the entire study was divided into two groups: qualitative characteristics (leaf shape, leaf colour, boll shape, leaf/stem hairs and pollen colour) that are not influenced by the environment (Appendix 4.2 and Table 4.2) and agronomical characteristics (seedcotton yield, GOT, bolls/plant, fibre length, fibre strength and micronaire value) that are influenced by the environment (Table 6.1).

The mean values (Table 6.1) for six agronomical characteristics namely seedcotton yield, GOT, boll/plant, fibre length, fibre strength and micronaire values, indicated a high variation among the 26 varieties. Variation of agronomical characteristics showed that some varieties out performed their respective means (Table 6.1). For example, NTA 93-15, SG 125, IL85, Cyto 12/74, HC-B4-75, Frego bract, UK82, Dixie King and Guazuncho, had higher values than the means for four and more characteristics, in contrast to High gossypol, Delcot 344, Okra leaf, Des 119 and IL74 that had less than three characteristics having values above average.

Table 6. 1 Variety means for different agronomical characteristics

No	Variety	GOT (%)	Boll/plant (no)	Seedcotton (ton/ha)	Fibre length (mm)	Micronaire (units)	Fibre strength (g/tex)
1	High gossypol	35.95	21.40	2.15	32.3	4.25	40.70
2	HC-B4-75	41.48	25.78	2.61	30.5	4.59	26.70
3	NTA 93-15	43.43	33.45	2.64	27.8	4.41	31.10
4	BJA 592	38.53	27.65	3.10	29.3	3.86	29.80
5	Reba W296	37.23	28.33	2.73	28.8	5.42	31.30
6	Des 119	40.78	19.13	1.94	27.0	4.72	29.20
7	MZ 561	38.65	40.23	2.15	28.8	4.91	32.20
8	Frego bract	40.95	34.65	2.63	32.5	3.92	29.60
9	McNair 235	35.80	24.08	2.35	32.0	4.81	31.40
10	UK91	37.80	21.80	2.76	31.0	4.33	33.60
11	Dixie King	40.45	21.03	2.66	28.5	4.38	24.90
12	IL85	35.33	36.58	3.03	28.3	4.66	30.53
13	Irma 1243	44.90	25.30	2.40	29.3	4.54	31.05
14	UK82	38.93	44.28	2.57	31.0	4.45	31.80
15	DP 4049	42.25	21.15	2.66	29.3	4.83	29.10
16	Okra leaf	33.58	19.03	1.71	28.8	5.04	30.60
17	IL74	38.78	25.65	2.39	27.8	4.94	30.33
18	Guazuncho	41.40	28.83	3.09	32.3	4.34	30.90
19	SG 125	40.80	29.30	2.85	30.0	5.13	35.00
20	Delcot 344	41.18	22.33	2.24	28.0	4.00	27.10
21	Cyto 12/74	34.55	33.08	3.07	31.3	4.22	35.50
22	Acala SJ$_2$	37.93	26.28	2.68	27.8	3.68	28.20
23	Stoneville 506	39.55	22.53	2.13	27.3	4.44	34.00
24	Reba B50	36.83	32.95	2.42	32.3	4.55	29.60
25	Aubarn 56	39.93	28.95	2.61	29.0	4.43	34.20
26	NTA 88-6	43.98	27.48	2.32	30.5	3.95	33.80
	Mean	39.35	27.62	2.53	29.8	4.40	31.37
	LSD for V	2.130	8.275	0.509	0.001	0.089	0.979
	SED for V	1.281	4.977	0.307	0.001	0.054	0.589
	Repeat.(%)	79	53	55	99	97	95
	CV (%)	4.60	25.45	17.12	0.08	1.73	2.66
	Sign. For V	***	***	*	***	***	***

Sign.=significant; * $p \leq 0.005$, ** $p \leq 0.01$; *** $p \leq 0.001$; CV=coefficient of variation; SED=standard deviation; LSD=least significance difference; V=variety; Repeat.=repeatability, GOT= ginning outturn, V= variety

6.3.1 Estimates of genetic similarities

Genetic similarities for agronomical data (16 data points) and AFLP analysis data (853 data points) are summarised in Table 6.2. Genetic similarities for all pairs (N = 325) ranged from 0.0-1.0 for agronomical and for AFLP analysis varied from 0.894-0.979, with means of 0.448

Table 6.2 Agronomical (above) and AFLP (below) genetic similarities

	HCB	RBW	FRB	DK	DP4	SG	CYT	RB5	N88	HGP	DES	MZ	McN	IL8	IL7	GUA	DEL	STN	AUB	N93	BJA	UK9	IRM	UK8	OKL	ACL
HCB		0.167	0.333	0.833	0.833	0.333	0.333	0.500	0.500	0.364	0.727	0.364	0.667	0.333	0.500	0.333	1.000	0.364	0.667	0.333	0.333	0.500	0.833	0.500	0.333	0.333
RBW	0.961		0.333	0.333	0.333	0.833	0.833	0.333	0.500	0.364	0.000	0.364	0.500	0.667	0.167	0.500	0.167	0.182	0.333	0.500	0.500	0.500	0.333	0.333	0.833	0.333
FRB	0.946	0.978		0.500	0.500	0.500	0.333	0.500	0.500	0.182	0.364	0.000	0.000	0.167	0.000	0.667	0.333	0.000	0.500	0.667	0.167	0.167	0.000	0.167	0.000	0.500
DK	0.954	0.960	0.953		1.000	0.500	0.500	0.333	0.333	0.364	0.727	0.364	0.500	0.500	0.333	0.500	0.833	0.364	0.500	0.500	0.500	0.667	0.667	0.333	0.500	0.500
DP4	0.970	0.953	0.946	0.953		0.500	0.500	0.333	0.333	0.364	0.727	0.364	0.500	0.500	0.333	0.500	0.833	0.364	0.500	0.500	0.500	0.667	0.667	0.333	0.500	0.500
SG	0.962	0.951	0.944	0.958	0.977		0.667	0.167	0.667	0.182	0.182	0.364	0.333	0.500	0.167	0.667	0.333	0.182	0.333	0.667	0.500	0.500	0.500	0.333	0.667	0.167
CYT	0.921	0.943	0.954	0.934	0.920	0.912		0.500	0.500	0.546	0.182	0.546	0.182	0.667	0.833	0.333	0.667	0.364	0.500	0.667	0.500	0.667	0.500	0.500	0.667	0.333
RB5	0.943	0.951	0.946	0.964	0.939	0.943	0.946		0.333	0.546	0.364	0.182	0.500	0.333	0.333	0.500	0.500	0.182	0.333	0.333	0.333	0.167	0.333	0.333	0.167	0.333
N88	0.957	0.944	0.937	0.949	0.964	0.967	0.910	0.943		0.182	0.182	0.364	0.500	0.333	0.333	0.500	0.500	0.182	0.500	0.500	0.500	0.333	0.667	0.500	0.333	0.167
HGP	0.936	0.924	0.918	0.931	0.937	0.935	0.894	0.925	0.945		0.400	0.400	0.727	0.546	0.546	0.546	0.364	0.600	0.546	0.364	0.000	0.546	0.546	0.364	0.546	0.364
DES	0.944	0.941	0.947	0.937	0.948	0.939	0.923	0.927	0.950	0.936		0.200	0.364	0.182	0.546	0.364	0.727	0.600	0.364	0.546	0.182	0.364	0.546	0.182	0.182	0.546
MZ	0.950	0.954	0.953	0.936	0.950	0.937	0.933	0.929	0.941	0.932	0.970		0.546	0.727	0.546	0.364	0.364	0.600	0.727	0.364	0.364	0.727	0.546	0.909	0.364	0.000
McN	0.944	0.943	0.952	0.934	0.943	0.937	0.925	0.922	0.942	0.935	0.975	0.979		0.667	0.667	0.333	0.667	0.546	0.833	0.333	0.167	0.667	0.833	0.667	0.667	0.333
IL8	0.953	0.939	0.927	0.946	0.957	0.952	0.910	0.939	0.964	0.942	0.956	0.953	0.953		0.333	0.500	0.333	0.364	0.500	0.500	0.364	0.333	0.500	0.667	0.667	0.333
IL7	0.949	0.936	0.928	0.945	0.958	0.947	0.907	0.932	0.949	0.936	0.943	0.946	0.940	0.966		0.333	0.500	0.909	0.833	0.500	0.167	0.667	0.667	0.667	0.333	0.333
GUA	0.945	0.927	0.919	0.944	0.947	0.947	0.900	0.924	0.945	0.950	0.939	0.937	0.940	0.957	0.953		0.333	0.364	0.333	0.833	0.333	0.500	0.500	0.333	0.333	0.167
DEL	0.920	0.921	0.927	0.918	0.922	0.915	0.923	0.911	0.911	0.909	0.937	0.936	0.936	0.917	0.919	0.921		0.364	0.667	0.333	0.333	0.500	0.833	0.500	0.333	0.333
STN	0.955	0.938	0.931	0.951	0.960	0.953	0.911	0.934	0.952	0.945	0.949	0.949	0.946	0.964	0.967	0.970	0.927		0.727	0.546	0.182	0.727	0.546	0.546	0.364	0.364
AUB	0.939	0.938	0.940	0.934	0.946	0.938	0.918	0.920	0.935	0.923	0.947	0.955	0.949	0.943	0.952	0.940	0.940	0.958		0.333	0.333	0.833	0.833	0.833	0.500	0.167
N93	0.947	0.929	0.917	0.936	0.949	0.940	0.901	0.920	0.940	0.930	0.926	0.934	0.930	0.946	0.949	0.944	0.920	0.951	0.938		0.333	0.500	0.500	0.333	0.333	0.333
BJA	0.948	0.935	0.925	0.940	0.947	0.943	0.903	0.927	0.938	0.934	0.929	0.935	0.931	0.945	0.947	0.945	0.922	0.953	0.951	0.964		0.500	0.167	0.333	0.333	0.500
UK9	0.947	0.927	0.926	0.934	0.952	0.940	0.900	0.920	0.941	0.938	0.932	0.935	0.934	0.941	0.950	0.943	0.918	0.946	0.937	0.954	0.954		0.667	0.667	0.667	0.333
IRM	0.937	0.931	0.939	0.931	0.940	0.930	0.918	0.918	0.932	0.927	0.940	0.952	0.947	0.938	0.942	0.939	0.935	0.941	0.946	0.937	0.947	0.955		0.667	0.500	0.167
UK8	0.950	0.932	0.925	0.939	0.953	0.947	0.903	0.929	0.947	0.940	0.930	0.938	0.931	0.949	0.953	0.950	0.915	0.958	0.948	0.952	0.965	0.957	0.955		0.333	0.000
OKL	0.941	0.936	0.929	0.943	0.938	0.936	0.912	0.935	0.935	0.925	0.920	0.929	0.921	0.938	0.939	0.933	0.909	0.940	0.931	0.942	0.952	0.947	0.941	0.961		0.500
ACL	0.943	0.934	0.923	0.938	0.944	0.944	0.907	0.925	0.937	0.928	0.927	0.936	0.927	0.945	0.952	0.937	0.916	0.945	0.940	0.947	0.953	0.952	0.946	0.958	0.945	

HCB= HC-B4-75, RBW= Reba W296, FRB= Frego bract, DK= Dixie King, DP4= DP 4049, SG= SG 125, CYT= Cyto 12/74, RB5= Reba B50, N88= NTA 88-6, HGP= High gossypol, DES= Des 119, MZ= MZ561, McN= McNair 235, IL8= IL85, IL7= IL74, GUA= Guazuncho, DEL= Delcot 344, STN= Stoneville, AUB= Aubarn 56, N93= NTA 93-15, BJA= BJA 592, UK9=UK91, IRM= Irma 1243, UK8= UK82, OKL= Okra leaf, ACL= Acala SJ$_2$

and 0.939 respectively. The first 50 variety pair wise similarities for AFLP analysis in descending order ranged from 0.979-0.953 while values for agronomical analysis ranged from 1.000-0.667. Absolute similarities and rankings (Table 6.3) were quite different between the two measurement systems (agronomical and AFLP analysis). From the set of 50 combinations, only five similar combinations were observed in both methods. High levels of genetic similarities were observed for AFLP analysis compared to agronomical analysis (Tables 6.2 and 6.3).

Agronomical similarity matrix data ranged from 0.00 to 1.00, indicating that some varieties were agronomically different while others were similar. The most similar varieties were HC-B4-75 and DP 4049 (1.00), DP 4049 and Dixie King (1.00), Stoneville 506 and IL74 (0.91) and UK82 and MZ561 (0.91). Low similarity values were observed for Acala SJ_2 and MZ561 (0.00), Acala SJ_2 and UK82 (0.00), Frego bract and IL74 (0.00), Frego bract and UK82 (0.00), Frego bract and McNair 235 (0.00), Frego bract and Stoneville 506 (0.00) and Frego bract and Aubarn 56 (0.00). Similarity values for other combinations ranged from 0.17-0.83. Based on low similarity values with most of the varieties, Acala SJ_2 and Frego bract were agronomicaly distinct from other varieties. Based on agronomical and AFLP analysis some varieties were consistently close, for example HC-B4-75 and DP 4049, DP 4049 and Delcot 344, HC-B4-75 and Dixie King, Dixie King and DP 4049 and to a lesser extent, DP 4049 and Irma 1243. Most variety pairs behaved irregularly from one method to the other (Tables 6.2 and 6.3). This irregularity was more clearly exhibited using PCA biplots (Figure 6.5). Pearson correlation coefficient between AFLP analysis and agronomical genetic similarities was $r = 0.0097$ at $p < 0.863$.

Based on morphological data (31 points), genetic similarities ranged from 0.17-0.92 with an average of 0.53, indicating that some varieties were morphologically similar while others were different. The most similar varieties were IL74 and Cyto 12/74, Delcot 344 and Dixie King, SG 125 and Dixie King, NTA 93-15 and NTA 88-6, NTA 93-15 and DP 4049, IL74 and MZ561, Delcot 344 and McNair 235, Stoneville 506 and Delcot 344, UK82 and Aubarn 56 and UK82 and UK91. Lowest similarity values were observed for Okra leaf and HC-B4-75 and NTA 93-15 and Aubarn 56. Based on morphological and AFLP analyses genetic similarity values, the correlation coefficient was $r = 0.05$ at $p < 0.33$ indicating low correlation though it was high

Table 6.3 Genetic similarity values for the first 50 pair wise comparisons for AFLP and agronomical data

AFLP Pair wise similarities			Rank	Agronomical pair wise similarities			Rank
Des 119	Acala SJ$_2$	0.979	1	HC-B4-75	Delcot 344	1.000	1
Reba W296	Frego bract	0.978	2	**Dixie King**	**DP 4049**	**1.000**	**2**
Dixie King	Acala SJ$_2$	0.975	3	MZ561	UK82	0.909	3
Des 119	MZ561	0.970	4	IL74	Stoneville 506	0.909	4
High gossypol	Acala SJ$_2$	0.970	5	**HC-B4-75**	**Dixie King**	**0.833**	**5**
Guazuncho	Delcot 344	0.970	6	**HC-B4-75**	**DP 4049**	**0.833**	**6**
HC-B4-75	**DP 4049**	**0.967**	**7**	HC-B4-75	Irma 1243	0.833	7
IL74	Delcot 344	0.967	8	Reba W296	SG 125	0.833	8
SG 125	Reba B50	0.966	9	Reba W296	Cyto 12/74	0.833	9
McNair 235	Acala SJ$_2$	0.966	10	Reba W296	Okra leaf	0.833	10
BJA 592	Irma 1243	0.965	11	Dixie King	Delcot 344	0.833	11
DP 4049	Reba B50	0.964	12	**DP 4049**	**Delcot 344**	**0.833**	**12**
NTA 88-6	McNair 235	0.964	13	Cyto 12/74	IL85	0.833	13
Aubarn 56	Acala SJ$_2$	0.964	14	McNair 235	Aubarn 56	0.833	14
Dixie King	Cyto 12/74	0.964	15	McNair 235	Irma 1243	0.833	15
IL85	Delcot 344	0.964	16	IL74	Aubarn 56	0.833	16
HC-B4-75	SG 125	0.962	17	Guazuncho	NTA 93-15	0.833	17
HC-B4-75	Reba W296	0.961	18	Delcot 344	Irma 1243	0.833	18
Irma 1243	Okra leaf	0.961	19	Aubarn 56	UK91	0.833	19
Reba W296	Dixie King	0.960	20	Aubarn 56	Irma 1243	0.833	20
DP 4049	**Delcot 344**	**0.960**	**21**	Aubarn 56	UK82	0.833	21
Dixie King	**DP 4049**	**0.958**	**22**	HC-B4-75	Des 119	0.727	22
Stoneville 506	Irma 1243	0.958	23	Dixie King	Des 119	0.727	23
DP 4049	IL85	0.958	24	DP 4049	Des 119	0.727	24
Delcot 344	Acala SJ$_2$	0.958	25	High gossypol	McNair 235	0.727	25
Irma 1243	Acala SJ$_2$	0.958	26	Des 119	Delcot 344	0.727	26
UK91	Irma 1243	0.957	27	MZ561	IL85	0.727	27
DP 4049	McNair 235	0.957	28	MZ561	Aubarn 56	0.727	28
IL85	IL74	0.957	29	MZ561	UK91	0.727	29
HC-B4-75	NTA 88-6	0.957	30	Stoneville 506	Aubarn 56	0.727	30
Des 119	McNair 235	0.956	31	Stoneville 506	UK91	0.727	31
MZ561	Stoneville 506	0.955	32	HC-B4-75	McNair 235	0.667	32
HC-B4-75	Stoneville 506	0.955	33	HC-B4-75	Aubarn 56	0.667	33
BJA 592	Acala SJ$_2$	0.955	34	Reba W296	IL85	0.667	34
UK91	Okra leaf	0.955	35	Frego bract	Guazuncho	0.667	35
Reba W296	MZ561	0.954	36	Frego bract	BJA 592	0.667	36
HC-B4-75	**Dixie King**	**0.954**	**37**	Dixie King	UK91	0.667	37
NTA 93-15	Acala SJ$_2$	0.954	38	Dixie King	Irma 1243	0.667	38
Frego bract	Cyto 12/74	0.954	39	DP 4049	UK91	0.667	39
NTA 93-15	BJA 592	0.954	40	**DP 4049**	**Irma 1243**	**0.667**	**40**
IL85	Acala SJ$_2$	0.953	41	SG 125	Cyto 12/74	0.667	41
IL74	Irma 1243	0.953	42	SG 125	NTA 88-6	0.667	42
BJA 592	Okra leaf	0.953	43	SG 125	Guazuncho	0.667	43
SG 125	Delcot 344	0.953	44	SG 125	NTA 93-15	0.667	44
MZ561	Acala SJ$_2$	0.953	45	SG 125	Okra leaf	0.667	45
Stoneville 506	NTA 93-15	0.953	46	Cyto 12/74	McNair 235	0.667	46
Frego bract	Acala SJ$_2$	0.953	47	Cyto 12/74	Guazuncho	0.667	47
Reba W296	DP 4049	0.953	48	Cyto 12/74	NTA 93-15	0.667	48
Frego bract	Dixie King	0.953	49	Cyto 12/74	UK91	0.667	49
DP 4049	**Irma 1243**	**0.953**	**50**	Cyto 12/74	Okra leaf	0.667	50

Note: Bolded pairs are ranked for both methods

compared to comparison between agronomical and AFLP analyses ($r = 0.0097$ at $p < 0.863$). AFLP analysis genetic similarities were described in section 5.3.2.

6.3.2 Cluster analysis

Dendrograms for agronomical, morphological, combined agronomical and morphological and AFLP analyses are presented in Figures 6.1, 6.2, 6.3 and 6.4, respectively. Based on agronomical clustering, the 26 varieties were grouped into two main groups (*A* and *B*). Group A contained three varieties (Acala SJ$_2$, BJA 592 and Frego bract). These varieties had low micronaire values (3.68-3.92) and more or less similar fibre strength and seedcotton yield values (Table 6.1). Main group *B* contained 23 varieties clustered into three clusters *I, II* and *III*. Cluster *I* comprised 13 varieties, which were further clustered into two subclusters. The first subcluster contained five varieties (HC-B4-75, Delcot 344, Dixie King, DP 4049 and Des 119) all from the USA. These varieties had high GOT values ranging from 40.5-42.3% (Table 6.1). The second subcluster contained eight varieties: MZ561 and UK82, both from Tanzania, clustered together and had many bolls/plant and low GOT values. Another four varieties, two from the USA (McNair 235 and Aubarn 58), one from West Africa (Irma 1243) and one from Tanzania (UK91) clustered together. These varieties had bolls/plant ranging from 21-25 and seedcotton ranging from 2.3-2.7 ton/ha. The last two of the eight varieties (IL74 and Stoneville 506) clustered together. These had more or less similar fibre length values (27.8 mm and 27.3 mm respectively).

Cluster *II* contained two varieties: Reba B50 and High gossypol that had similar seedcotton yield, GOT and fibre length.

Cluster *III* comprised eight varieties, which were clustered into two subclusters. The first subcluster contained seven varieties. Five varieties (Reba W296, SG 125, Okra leaf, Cyto 12/74 and IL85) in this subcluster had low GOT (< 38%) and more or less similar fibre length values except for SG 125 (Table 6.1). Other varieties (Guazuncho and NTA 93-15) had more or less similar micronaire values and fibre strength. The second subcluster comprised only NTA 88-6. Although NTA 88-6 grouped in cluster *III*, it was distinct from other varieties in the same cluster (Figure 6.1). Cluster *III* comprised varieties from different origins and most of these varieties had good fibre quality and high seedcotton yield (Table 6.1).
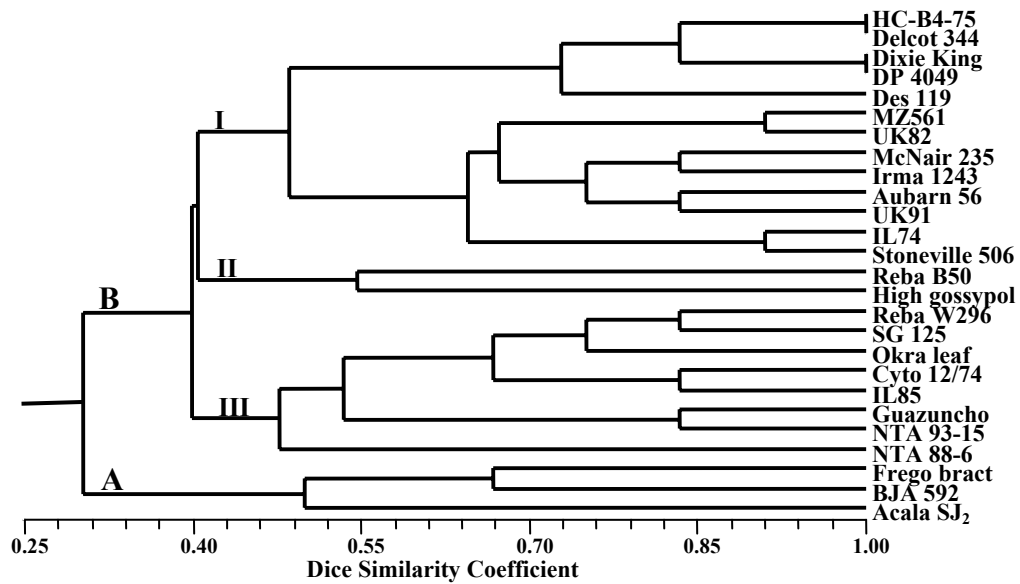
Figure 6.1 Agronomical characterisation of 26 varieties using UPGMA clustering method and Dice similarity coefficient

Based on morphological data, cluster analysis results (Figure 6.2) were similar to Figure 4.1 in section 4.3 except for few varieties. The difference occurred because in section 4.3.5, 30 varieties were analysed using NCSS computer package while in section 6.3.2 (Figure 6.2) only 26 varieties from which pure DNA for AFLP analysis could be obtained were used for comparison using NTSYS computer package. Genotypes (Figure 6.2) were grouped into two major groups *A* and *B*. Group *A* grouped into two clusters. Cluster *I* contained Okra leaf and cluster *II* was further grouped into two subclusters *i* and *ii*. Subcluster *i* contained Frego bract and subcluster *ii* was further grouped into two groups *a* and *b*. Group *a* contained Acala SJ$_2$ and group *b* was further divided into two groups. First group contained Irma 1243, IL85 and Reba B50 and the second group was further grouped into two more groups. First group contained seven varieties (UK91, UK82, Aubarn 56, High gossypol, IL74, MZ561 and Cyto 12/74). The other subgroup contained Reba W296 and BJA 592. All varieties in main group A were from Africa except for Acala SJ$_2$, Aubarn 56 (both from the USA) and Cyto 12/74 from Pakistan.

Major group *B* contained two clusters. Cluster *I* was further clustered into two subclusters *i* and *ii*. Subcluster *i* contained DP 4049, NTA 93-15, NTA 88-6 and Des 119 and subcluster *ii* contained Dixie King, SG 125, McNair 235, Delcot 344, Stoneville 506 and Guazuncho. Cluster *II* contained HC-B4-75. All varieties in major group *B* were from the USA and related to

Deltapine varieties. Therefore, varieties grouped according to origin and pedigree relationships. For example, group *b* in major group *A* contained most of varieties from Tanzania (UK82, UK91, IL74 and MZ561) and their ancestors (Reba W296 and BJA 592) while subcluster *i* of major group *B* contained Deltapine related varieties.
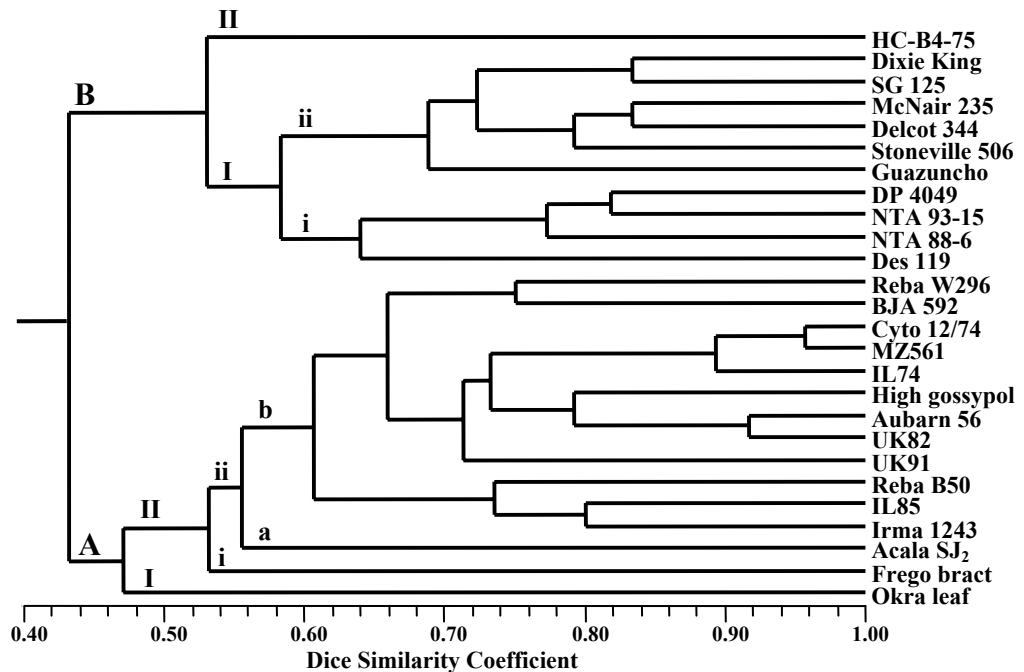
Figure 6.2 Morphological characterisation of 26 varieties using UPGMA clustering method and Dice similarity coefficient

The dendrogram for combined agronomical and morphological data is presented in Figure 6.3. Two major groups *A* and *B* were revealed. Group *A* comprised two clusters *I* and *II*. Cluster *I* contained one variety Acala SJ$_2$ and cluster *II* was divided into two subclusters *i* and *ii*. Subcluster *i* contained Okra leaf and subcluster *ii* was further grouped into two groups *a* and *b*. Group *a* contained two more groups. First group contained Reba B50 and second group had nine varieties (Cyto 12/74, MZ561, IL74, Aubarn 56, UK82, UK91, High gossypol, IL85 and Irma 1243). Group *b* contained three varieties (Reba W296, BJA 592 and Fregobract).

Major group *B* clustered into two clusters *I* and *II*. Cluster *I* comprised two varieties (SG 125 and Guazuncho). Cluster *II* was divided into two subclusters *i* and *ii*. Subcluster *i* contained four varieties (DP 4049, NTA 93-15, NTA 88-6 and Des 119) and second subcluster *ii* was further

divided into two groups *a* and *b*. Group *a* contained four varieties (Dixie King, Delcot 344, McNair 235 and Stoneville 506) and group *b* comprised HC-B4-75. Based on combined agronomical and morphological data clustering (Figure 6.3), few varieties grouped similar to the AFLP analysis dendrogram (Figure 6.4), for example UK82 and UK91. Another similar group contained McNair 235, Stoneville 506, SG 125, HC-B4-75 and DP 4049.
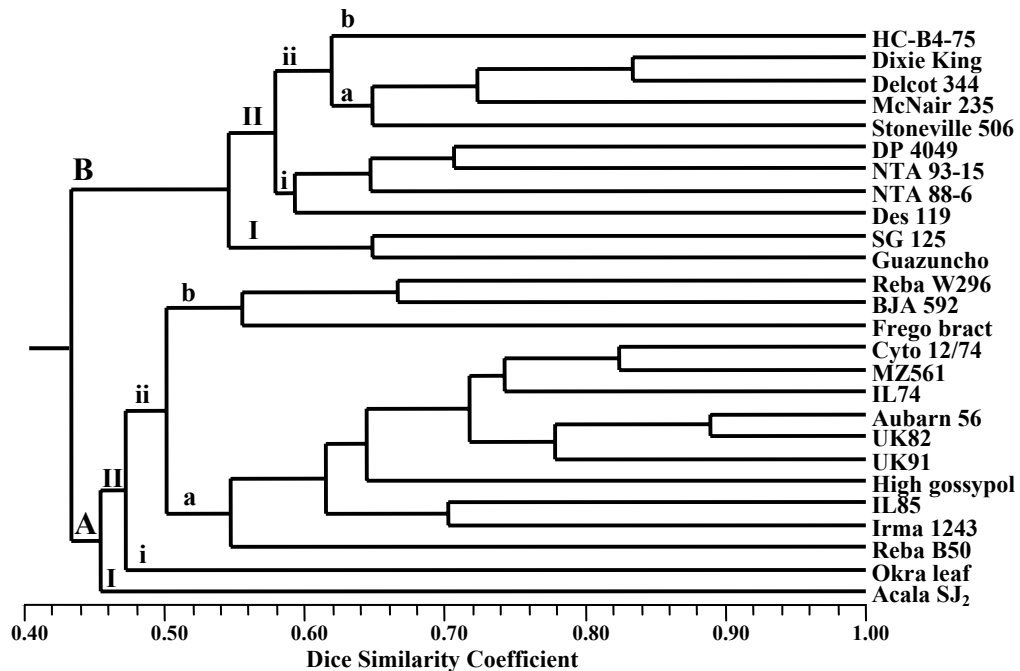
Figure 6.3 Agronomical and morphological characterisation of 26 varieties using UPGMA clustering method and Dice similarity coefficient

Detailed results for AFLP analysis dendrogram are presented in section 5.3.3 and Figures 5.3 and 6.4. The four dendrograms of agronomical, morphological, combined agronomical and morphological and AFLP analyses presented different grouping patterns although some varieties clustered similarly in all methods. For example, HC-B4-75, DP 4049, Stoneville 506, Des 119 and McNair 235 always clustered in the same main group as well as UK82 and UK91. This indicated some relationship among these characterisation methods of genetic diversity studies based on dendrogram analyses (Figures 6.1, 6.2. 6.3 and 6.4 respectively).
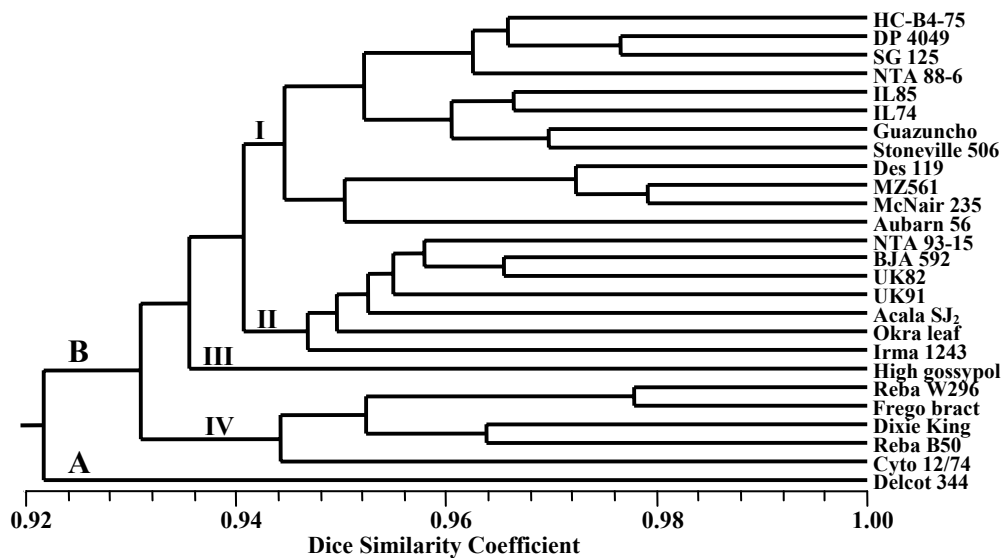
Figure 6.4 AFLP characterisation of 26 varieties using UPGMA clustering method and Dice similarity coefficient

Further comparisons using agronomical (boll/plant, GOT, seedcotton, fibre length, fibre strength and mironaire values), morphological (leaf shape, leaf colour, stem colour, leaf hairs, stem hairs, boll shape and boll prominence) and AFLP analyses were done using Spearman's correlation coefficient and PCA clustering for further examination of variety relationships. Spearman's correlation coefficient for the three methods indicated that AFLP and agronomical methods were not correlated (0.01) at $p < 0.863$, AFLP and morphological revealed weak correlation (0.1) at $p < 0.33$ and agronomical and morphological were correlated (0.2) at $P < 0.001$. Correlation between combined agronomical and morphological data with AFLP analysis indicated no correlation (0.03) at $p < 0.63$.

6.3.3 Principle component analysis of agronomical data

Principle component analysis based on agronomical characteristics (Figure 6.5a) was performed to study interrelationships between all varieties. PCA clustered varieties similarly to clustering of the agronomical dendrogram, as observed on agronomical main group *A* and subclustering of cluster *I* of main group *B* (Figure 6.1). Most of the varieties which were agronomically similar were clearly grouped for example UK82, MZ561, IL74 and Stoneville 506 as well as Dixie King and DP 4049 and HC-B4-75 and Delcot 344. Varieties that were agronomically different,

117

for example Frego bract, Reba W296 and High gossypol were clearly identified. NTA 88-6 which clustered separately in the dendrogram as a separate subcluster in cluster *III* of main group *B* (Figure 6.1) clustered within a group in the PCA biplot (Figure 6.5a). Based on cluster representation (Figure 6.1), some varieties were scattered in PCA indicating that though they were in the same clusters, they were not too close. In the dendrogram (Figure 6.1) cluster *II* contained Reba B50 and High gossypol, however these varieties were presented distantly in the PCA biplot (Figure 6.5a). Therefore, total agronomical data indicated high diversity among varieties. However, some varieties including UK91, UK82, DP 4049, HC-B4-75, Stoneville 506, Des 119, McNair 235, Reba W296 and Reba B50, based on agronomical data, clustered similarly to other clusters for morphological and AFLP analyses. The PCA biplot (Figure 6.5a) presentation was consistent with observations on characteristics expressed by variety means, genetic similarities (Table 6.1 and 6.2) and the dendrogram (Figure 6.1), except for a few varieties.
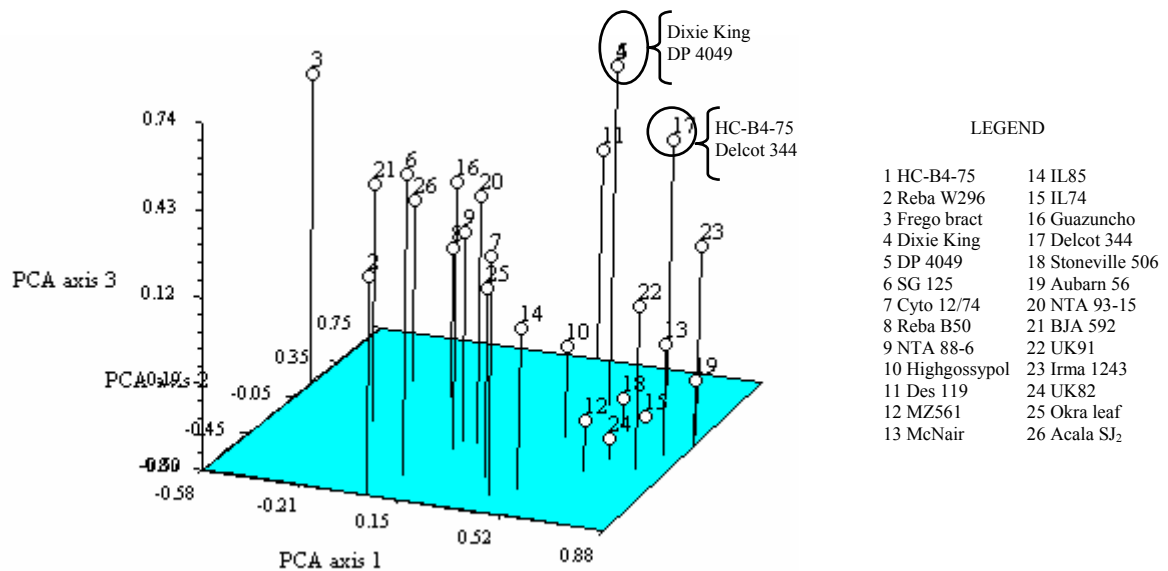
Figure 6.5a Agronomical data PCA biplot indicating relationships among 26 cotton varieties

6.3.4 Principle component analysis of morphological data

The PCA biplot for morphological analysis is presented in Figure 6.5b. Variety distribution was different from the agronomical biplot except for few varieties, but had some interesting similarities with the AFLP analysis PCA biplot (Figure 6.5b). For example, in both biplots

(morphological and AFLP analyses), Reba B50, Reba W296, Frego bract and Cyto 12/74 were distantly clustered from other varieties. Furthermore, HC-B4-75, DP 4049 and NTA 88-6 clustered as one group, Stoneville 506, BJA 592, UK82, UK91 and High gossypol representing another group were grouped similar to AFLP analysis PCA biplot. AFLP analysis and morphological (qualitative) biplots grouped most of the varieties according to origin and pedigree relationships.

6.3.5 Principle component analysis of AFLP data

In the PCA biplot for AFLP analysis, varieties clustered closely in clear groups, except for few a varieties (Figure 6.5c). The most similar pairs were McNair 235, Des 119 and MZ561, another group contained HC-B4-75, DP 4049, NTA 88-6, Okra leaf and SG 125. These pairs correlated with subclusters of cluster *I* obtained during construction of the AFLP dendrogram (Figure 6.4). The main group contained UK82, Stoneville 506, IL74, Guazuncho, BJA 596, UK91, IL85 and Acala SJ$_2$. This main group correlates with varieties that clustered in clusters *I* and *II* in main group *B* of the AFLP dendrogram. Delcot 344 was the most dissimilar and most notably in the other direction. Delcot 344 clustered separately from all other varieties in the AFLP dendrogram (main group *A*, Figure 6.4). Cyto 12/74, Reba B50, Dixie King, Frego bract and Reba W296, although scattered, formed a group, which correlated to cluster *IV* of the dendrogram (Figure 6.4). Based on dendrogram (Figure 6.4) Aubarn 56 and Irma 1243 were clustered in different clusters *I* and *II* respectively. The PCA biplot (Figure 6.5c) indicated that these two varieties were closely related in one direction. AFLP analysis grouped varieties in relation to origin and pedigree relationships and indicated high similarity as many varieties were grouped together. PCA biplot results correlated with genetic similarities and the constructed dendrogram (Table 6.2, Figures 6.4 and 6.5c).
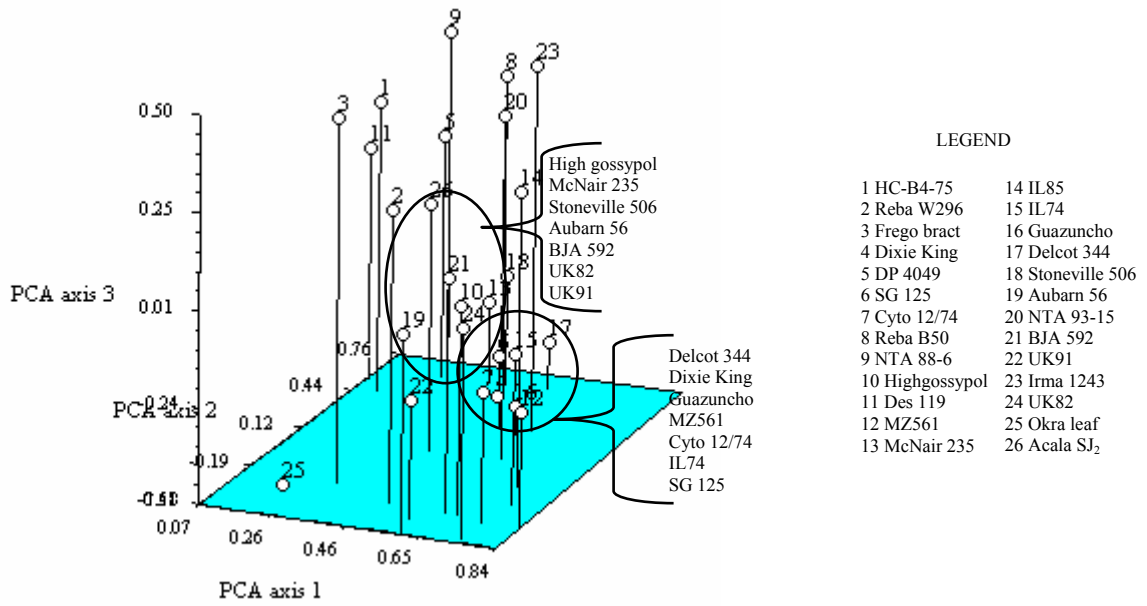
Figure 6.5b Morphological data PCA biplot indicating relationships among 26 cotton varieties
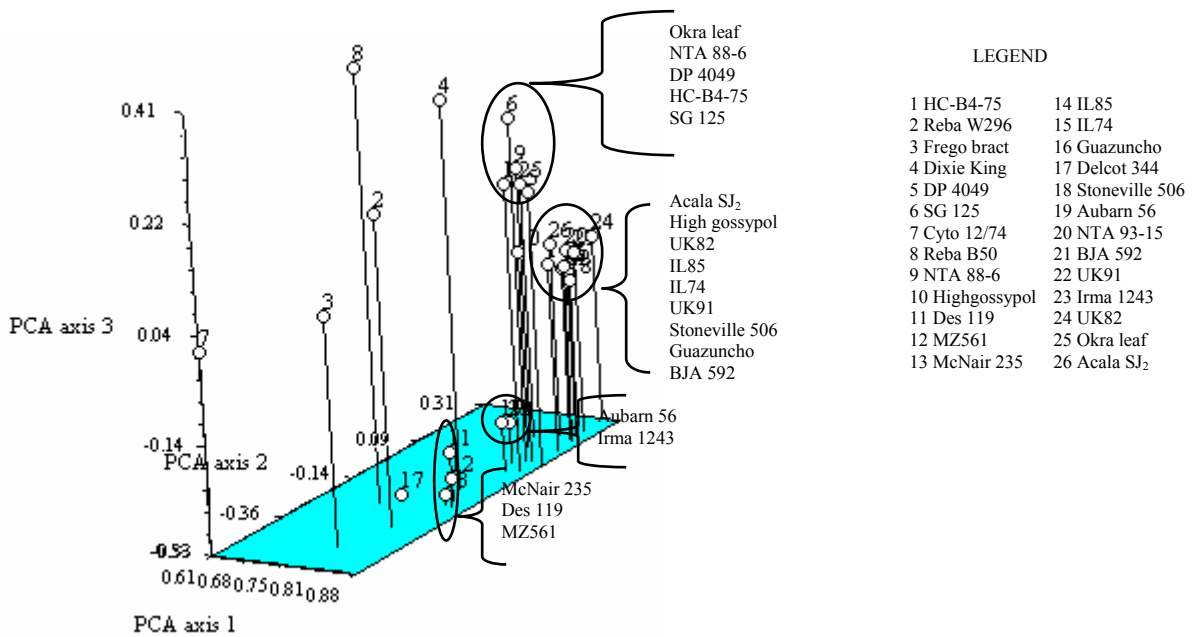
Figure 6.5c AFLP data PCA biplot indicating relationships among 26 cotton varieties

6.4 DISCUSSION

The main objectives for this study were to study genetic diversity among cotton varieties using agronomical, morphological and AFLP analyses and to compare these characterisation methods. Consideration of estimated genetic distance is important for comparative analysis of diversity levels (Roldan-Ruiz *et al.*, 2001). The overall mean of genetic similarities for agronomical data were 0.448, 0.530 for morphological and 0.500 for combined agronomical and morphological data and 0.939 for AFLP data. Therefore low genetic similarities were observed using agronomical and morphological compared to AFLP analysis. This is because agronomical characterisation deals with genetic similarities based on quantitative traits, which are influenced by the environment. According to Poehlman (1987), quantitative traits that are controlled by a number of genes with small effects are tremendously influenced by environment. Morphological data although based on qualitative traits, had few data points (31) compared to AFLP analysis (853).

Agronomical genetic similarities ranged from zero to one. The highest genetic similarities (1.0) indicated that varieties were closely related while the lowest genetic similarity (0.0) indicated that varieties were distantly related. Genetic similarities of AFLP did not support these wider range of agronomical genetic similarities. This indicated the presence of differences between these two methods. This was supported by weak Spearman correlation coefficient between AFLP and agronomical genetic similarities, indicating a complete absence of relationship between agronomical and AFLP analysis data as both methods are sampling different genetic information. Adugna (2002) obtained similar results.

The dendrogram constructed on the basis of agronomical traits was different from the AFLP (Figures 6.1 and 6.4). Few varieties clustered similarly but the majority grouped differently, probably due to the major effects of environment on the six quantitative traits used for characterisation. For example, out of 13 varieties in cluster *I*, seven (HC-B4-75, DP 4049, IL74, Stoneville 506, Des 119, MZ561 and McNair 235) clustered with cluster *I* of the AFLP dendrogam (Figure 6.1 and 6.4). This indicated partly a relationship between agronomical and AFLP grouping of varieties. However, based on agronomical data (Figure 6.1) varieties did not cluster according to collection areas or pedigree relationships.

Results from this study reflected the problem of quantitative trait characterisation which are highly influenced by the environment (Kumar, 1999). The AFLP method could be more important and reliable, as reported by Powell *et al*. (1996), Tohme *et al*. (1996), Barret and Kidwell (1998) and Swanepoel (1999) because it is not influenced by the environment. Furthermore, AFLP analysis is a powerful tool to discriminate and cluster closely related varieties as well as to trace origin and pedigree through genepool sharing.

Clustering based on morphological data, grouped most of the varieties according to origin and pedigree relationship. Two major groups related to origin were identified, one for Africa varieties and the other for USA varieties. Although an average similarity value of 0.53 was observed, morphological clustering was relatively similar to AFLP clustering. The correlation between morphological data and AFLP analysis was weak though high compared to that obtained for agronomical and AFLP data. According to Kumar (1999), morphological traits controlled by a single locus can be used as genetic markers, provided expression does not change over a range of environments. The weak correlation among morphological and AFLP was probably due to few data points (31) involved in morphological compared to AFLP analysis (853 data points).

Based on combined agronomical and morphological data, the average genetic similarity of 0.5 was observed. The grouping of varieties indicated some relationship with AFLP analysis. AFLP analysis provided more data points (853) compared to 16 data points for agronomical and 31 for morphological data. Efficiency increases with a higher amount of data points. AFLP analysis covers the entire genome, compared to agronomical and morphological analyses that focus on a few traits. AFLP markers are highly efficient compared to morphology and some other DNA markers since AFLP makers are reproducible and display intraspecific homology (Rana and Bhat, 2004). AFLP analysis is therefore a promising method for genetic studies in cotton.

AFLP analysis revealed high genetic similarities for MZ561 and McNair 235, Frego bract and Reba W296 and Des 119 and McNair 235, indicating close relationships, while a low similarity value (0.894) was observed between Cyto 12/74 and High gossypol. Delcot 344, High gossypol followed by Cyto 12/74, Irma 1243, Reba B50 and Okra leaf were the most distantly related varieties (Figure 6.4 and Table 6.2). Lu and Myers (2002) reported the high genetic distance of Delcot 344 with other varieties. Cross hybridising between distantly related varieties may

increase variation in the breeding population. Furthermore, clustering could be used to minimise germplasm collection and maintenance to avoid duplication.

The observed high genetic similarity average (0.936) in this study confirmed results reported by Pillay and Myers, (1999) and Abdukarimov *et al.* (2003) that cotton has low genetic diversity. Roldan-Ruiz *et al.* (2000) observed that when varieties with shared genepools were examined using AFLP markers, high similarity measures were produced which were linked to morphological similarities. Therefore, AFLP analysis can be used to conform variety pairs with shared genepools. In the current study, characterisation using agronomical, morphological and AFLP analyses independently provided the overall interrelationship of the studied varieties.

Generally genetic similarities, dendrograms and PCA biplots were related in discriminating and grouping varieties for each characterisation method involved in the study. Genetic similarities were able to group varieties in pairwise relationship and dendrograms grouped similar varieties in one cluster. The PCA biplots provided visual impression of how varieties were interrelated.

Many studies have been done to compare morphological and molecular based characterisation. Roldan-Ruiz *et al.* (2001) detected inconsistent relationships between morphology and AFLP using perennial ryegrass in a comparative study of molecular and morphological methods. Swanepoel (1999) reported a similar disparity between a phenotypic and molecular study on maize. Similar results were reported by Ben-Har *et al.* (1995), Bustin and Charcosset (1997), Wu *et al.* (2001) and Adugna (2002). Recently, an increasing number of studies demonstrated the capacity of molecular markers to efficiently discriminate varieties (Pejic *et al.,* 1998). In the present study, AFLP analysis exposed useful genetic relationships where varieties were dispersed more evenly compared to agronomical and morphological analyses. It provided more accurate and reliable relationships because it dealt with basic DNA sequences. Agronomical traits are expected to provide a general representation of variety relationships according to their growing environment. Therefore variety's performance (regarding quantitative traits) did not necessarily depend on the geographical origin or even pedigree relationship. Varieties that display high phenotypic similarities need not be genetically similar because the environment can manipulate phenotypic expression (Poehlman, 1987).

In this study the genetic diversity between varieties confirmed the usefulness of AFLP markers in studying genetic relationships as reported by Barret and Kidwell, (1998). However, Lübberstedt *et al.* (1998) and Swanepoel (1999) suggested that the combination of morphological and molecular markers could serve as a major source of information in separating closely related varieties. In the current study, AFLP separated closely related cotton varieties. For example varieties developed in Tanzania for different growing areas (IL74 and IL85 for the ECGA's) and (UK91 and UK82 for the WCGA's) were separated and Reba B50 and Reba W296 from West/Central Africa clustered together because of pedigree relationships. Guazuncho, Stoneville 506 and Des 119 (Figure 6.4) clustered together. Poisson *et al.* (2003) reported similarity between Guazuncho, Deltapine varieties and Stoneville varieties due to similar pedigree relationships.

AFLP analysis seems more efficient than morphology as it managed to arrange varieties according to origin and pedigree relationships (Chapter 5 and Figures 6.4 and 6.5c). However, one cannot undermine the role of morphological characterisation despite its limitations. Morphological characterisation has been used extensively for germplasm classification and identification, selection of the parents for cotton variety improvement as well as for germplasm collection, conservation and maintenance and is still useful in Tanzania. Its usefulness for characterisation is still valuable for farmers, breeders and germplasm maintainers and even for variety registration. Conventional breeding methods in combination with molecular markers have been successfully used to overcome breeding problems and in development of elite productive and highly adaptable cotton varieties (Abdukarimov *et al.,* 2003).

6.5 CONCLUSIONS AND RECOMMENDATIONS

In conclusion, low levels of correlation existed between agronomical, morphological, combined agronomical and morphological and AFLP based genetic similarities in the current study. AFLP analysis reflected the true expression of genotypes, while agronomical analysis encompassed the expression of genotype, environment and their interaction. Morphological (qualitative) characterisation is limited by few data points used for analysis. Agronomical and morphological characteristics are not consistent and few, whereas AFLP analysis appeared to provide more accurate estimates and utility of genetic diversity measurements. All methods have advantages and disadvantages for practical applications under different circumstances. Consequently, both

methods should continue rendering valuable services to farmers, breeders and genetic resource curators.

The overall findings from this study indicated that AFLP analysis and to a certain extent qualitative traits and quantitative traits, sufficiently detected genetic diversity to differentiate Tanzanian cotton varieties. Although all methods did not provide exactly the same description of relationships between varieties, there existed some consistency in discriminating varieties which were closely related and ones which were distantly related.

Although molecular markers like AFLPs analysis are more efficient and provide exciting insights (Kumar, 1999), they are limited due to initial costs, inadequate infrastructure and expensive chemicals. Application of DNA markers could accelerate the process of finding markers related to specific agronomical and morphological traits of interest, such as disease and pest tolerance (Spielmeyer *et al.,* 1998). *Gossypium hirsutum* has limited genetic diversity, therefore AFLP analysis may offer a powerful tool for analysing the inheritance and relationships of important traits in cotton breeding. Therefore, future research should focus on comparing the two methods in terms of feasibility, efficiency and accuracy by involving more tests over different environmental trials and years (for agronomic and morphological characterisation). Molecular analysis using more primer combinations and different molecular markers, along with costs and benefits, should be included.

CHAPTER 7

DIALLEL ANALYSIS ON VARIATION FOR YIELD AND FIBRE QUALITY OF TANZANIAN COTTON (*GOSSYPIUM HIRSUTUM* L.) GERMPLASM

7.1 INTRODUCTION

Cultivated cotton (*G. hirsutum* and *G. barbadense*) provide the world's main supply (> 99%) of raw cotton for factories and *G. hirsutum* produces a lint percentage of 35-45% (Ashraf and Ahmad, 2000). Efforts have been made in the past years to develop high yielding, good ginning outturn and good fibre quality cotton in Tanzania (Lukonge and Ramadhani, 1999). Most characteristics in cotton are quantitatively inherited enabling manipulation of genes (Meredith, 1984). Knowledge of genetic variation and heritability of yield and quality traits is essential to a cotton breeder, because the success of superior genotypes will depend on high heritability (Myers and Bordelon, 1995; Ashraf and Ahmad, 2000). The gene action and reproductive system of a genetic population provide information necessary to choose the best selection strategy for that population (Meredith, 1984).

Breeders of self-pollinated crops are primarily interested in combining desirable genes from different genotypes into a single genotype to create variability for a characteristic they wish to improve (Dabholkar, 1992). F_1s may exhibit superior performance due to dominance and/or non-allelic interaction. In advanced generations, however, linkage breaks and new combinations are formed. This leads to dissipation of superiority because the degree of dominance observed in the F_1 declines and combinations, which showed superiority due to non-allelic interaction, cease to exist. However, Tang *et al.* (1993a) and Reid (1995) reported F_2s being superior over parents, as they might have a broader range of adaptation than conventional cultivars due to high heterosis.

In order to improve productivity, one of the most important steps in a breeding programme is the detection of suitable parents. Combining ability analysis is used in breeding programmes to compare performances of lines in hybrid combinations (Griffing, 1956). Combining ability analysis is used to determine hybrid vigour in a population and thereby aid in selecting parents

for producing crosses and segregating populations (Meredith, 1984). In the current study combining ability analysis (general (GCA) and specific (SCA)) was used as a tool to differentiate good and poor combiners, followed by selection of appropriate crosses. One of the techniques widely used for this purpose in different crops, including cotton, is diallel analysis (Hayman, 1954; Dabholkar, 1992). Diallel cross analysis leads to identification of parents with additive and non-additive effects for specific characteristics. This in turn helps in choosing parents to be included in hybridisation or population breeding programmes (Murtaza *et al*., 2005). Apart from combining ability, diallel analysis can be used to study variation through heterosis and correlation coefficient analysis. In cotton, heterosis has a potential of increasing yield from 10-20% and improving fibre quality. One of the problems in using heterosis in cotton is defining a strategy for the selection of parents that will ultimately produce superior hybrids (Meredith and Brown, 1998). Prediction of genetic diversity and GCA of parents before crossing reduces the number of crosses and progeny to be screened and leads to a reduction in cost and time (Kumar, 1999).

Echekwu and Alaba (1995**)** and Coyle and Smith (1997) reported significant GCA for fibre strength, fibre length, fibre micronaire value, fibre uniformity, lint/boll, seed/boll, lint yield per unit land area and lint percentage. These results suggested a strong additive genetic variance for fibre quality parameters. It was observed that genotypes with good GCA estimates for fibre quality exhibited negative GCA for most of the within boll yield components (lint per boll, seed per boll). Tang *et al*. (1993b) observed that GCA accounted for more than 75% of sums of squares of hybrids and was significant for all fibre traits studied, indicating that additive genetic variation constituted the major portion of the genetic variation among hybrids. The F_1 hybrids showed significant positive correlations between GCA and total combining ability of parents (Xian *et al.,* 1995). Echekwu and Alaba (1995**)** reported that SCA mean squares were significant for lint percentage and boll size. Alam *et al.* (1992) observed a significant variance due to GCA and SCA for number of bolls/plant, plant height, seedcotton yield and ginning percentage. The magnitude of the GCA: SCA ratio indicated additive gene effects controlling all traits except ginning percentage, where the effect was non-additive.

Cotton breeders make use of heterosis but have the challenge of finding good combiners. The general approach is to cross genetically unrelated cultivars, but no information is apparent as

how to select these parents (Meredith and Brown, 1998). The idea of using more than two different parents in diallel crossing is to try and develop populations that are genetically more diverse since low levels of genetic diversity in cultivated cotton is well documented (Meredith, 1995a). Therefore, the improvement of these characteristics in the populations where selection will take place is important.

The aims of this study were 1) to study combining ability and identify suitable parents to be used in a cotton breeding programme for the improvement of yield and fibre quality, 2) to assess genetic and phenotypic correlation and heritability of yield and fibre quality using parents and 21 crosses of a half diallel cross and 3) to study the heterosis present in these characteristics.

7.2 MATERIALS AND METHODS

7.2.1 Parent material

Cotton material for this study was obtained from a 7x7 half diallel cross. Parents used in developing this material were obtained from the germplasm collection at the Tanzania Research Institute. Selection of parents was based on variability of traits contributing to yield (seedcotton and lint), ginning outturn (GOT), fibre quality, hairiness, yield components and insect and disease resistance (Table 7.1).

Table 7.1 Origin and characteristics of cotton parents used in 7x7 half diallel cross

	Variety	Source	Characteristics
1	Okra leaf	Israel	Resistant to insects, drought tolerant, early maturity
2	Acala SJ$_2$	USA	Large boll, high GOT, susceptible to bacterial blight
3	MZ561	Tanzania	Low yield, locally adapted
4	Delcot 344	USA	Good fibre quality, smooth leaves, reddish green leaves, early maturity
5	UK91	Tanzania	Medium yield, resistant (blight, fusarium), adapted to environment
6	Aubarn 56	USA	Resistant to fusarium wilt, resistant to root knot nematode
7	NTA 93-21	Mali	High yielding, susceptible to fusarium wilt, good fibre quality, small bolls, few hairs, early maturity

GOT= ginning outturn

7.2.2 Progeny

The 21 F_1 progeny used in this study were obtained by crossing the selected parents in a 7x7 half diallel mating design (Griffing, 1956). Seven varieties were grown and crossed in the greenhouse at the UFS for the season 2002/2003, but seeds were not enough since the crop was destroyed by low winter temperatures. Crosses were repeated in the field where seven ridges of 5 m long and 0.9 m wide were used per variety during the 2003/2004 season. Farmyard manure at 7 ton/ha and TSP feritilizer at 112 kg/ha was applied before field harrowing. Nitrogen fertilizer (Urea) was applied six and 10 weeks after planting at a rate of 50 kg/ha. American bollworm, aphids and red spider mites were controlled by a fortnightly spray of *Cypermethrine/dimethoate (Cypercal D)* mixture at 2.5 l/ha, following recommended cotton management practices. Other recommended field management practices like weeding four times for the entire season was applied. The trial was rain fed and no irrigation was applied.

Hand emasculation and pollination techniques were used for crossing. Flowers of the female parents were emasculated in the evening prior to opening and the pistil of the emasculated flower was pushed in a 3 cm long straw tube and folded over at the end. This helped to keep the stigma moist and avoid pollination from any foreign pollen. The next morning pollen from the paternal parent was put in the straw and the pistil pushed back into the straw with pollen. Pollination was done before 10 am to maximize seed setting and boll retention. Kausar *et al.* (1998) reported a decrease of boll retention from 30 % to 58.31 % when pollinations were delayed from 10 am to 11 am. Twenty-one combinations were raised from this mating design (Table 7.2). The seedcotton harvested from these crosses and their parents were ginned to get seeds that were used in this study.

7.2.3 Land preparation and field management

Twenty-eight entries (seven parents and 21 F_1 progeny) (Tables 7.1 and 7.2) were used in this study during the 2004/2005 season to evaluate yield (seedcotton and lint), yield components (boll/plant, seed/boll and boll weight), GOT and fibre quality (fibre length, fibre strength and micronaire). The experiment was conducted at four sites (Ukiriguru, Bwanga, Kanziga and Mwanhala) having different environmental conditions such as rainfall, temperature and soil types (Table 8.2). All these locations fall under the WCGA's (Figure 7.1).

Table 7.2 Description of cotton parents with half diallel crosses used in the study

	Okra leaf	Acala SJ$_2$	MZ561	Delcot 344	UK91	Aubarn 56
Okra leaf	s					
Acala SJ$_2$	Acala SJ$_2$ x Okra leaf	s				
MZ561	MZ561 x Okra leaf	MZ561 x Acala SJ$_2$	s			
Delcot 344	Delcot 344 x Okra leaf	Delcot 344 x Acala SJ$_2$	Delcot 344 x MZ561	s		
UK91	UK91 x Okra leaf	UK91 x Acala SJ$_2$	UK91 x MZ561	UK91 x Delcot 344	s	
Aubarn 56	Aubarn 56 x Okra leaf	Aubarn 56 x Acala SJ$_2$	Aubarn 56 x MZ561	Aubarn 56 x Delcot 344	Aubarn 56 x UK91	s
NTA 93-21	NTA 93-21 x Okra leaf	NTA 93-21 x Acala SJ$_2$	NTA 93-21 x MZ561	NTA 93-21 x Delcot 344	NTA 93-21 x UK91	NTA 93-21 x Aubarn 56

s = self pollinated
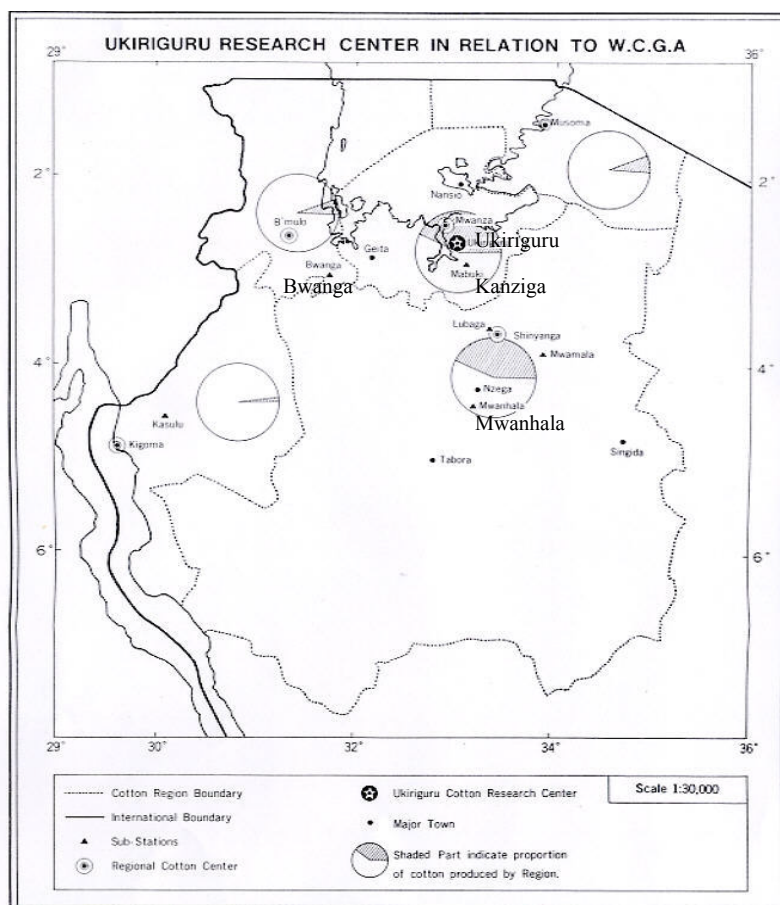
Figure 7.1 **Western Cotton Growing Areas showing the trial sites Ukiriguru, Kanziga, Bwanga and Mwanhala adapted from Nyambo (1982)**

Land preparation was done early October 2004 according to the standard procedures by ploughing and harrowing as described in section 7.2.2. Three ridges of 3 m long and 0.9 m wide, making a plot size of 8.1 m^2, were used per plot. A randomised complete block design (RCBD), replicated four times was used.

Planting dates varied from 30 November to 7 December 2004 for different sites. Seeds were placed at a depth of 3-4 cm at a spacing of 40 cm and four seeds were planted per hill. Seedlings were thinned to two plants per stand three weeks after planting. Field management was done as described in section 7.2.2. Plants were exposed to natural weather conditions where temperature and rainfall varied depending on the environmental conditions of the area (Table 8.1).

7.2.4 Data collection

Five plants in each plot were randomly selected for data recording for number of bolls per plant and seed per boll. Twenty full open bolls per plot were sampled, ginned and used to determine boll weight (g), 100 seed weight (g) and GOT. Matured bolls were hand picked. Two harvests were done at all sites except Kanziga, which had three harvests because of prolonged rains. Picking was done when the dew had evaporated. Lint samples (448) weighing 120 g each were sent to the Tanzania Cotton Board, in Dar es Salaam for analysis using a High Volume Instrument (HVI). Lint quality measurements included fibre strength, fibre length and fibre fineness (micronaire value).

7.2.5 Statistical analysis

Collected data were subjected to statistical analysis of variance using Agrobase (2000). Simple analysis of variance was done separately for each site to assess the performance of each genotype for each location. From the ANOVA the mean squares were used to evaluate levels of significant differences among replications and genotypes. Least significant differences (LSD) were used to separate mean differences and to rank genotypes. Coefficients of variation (CV) were used to determine the magnitude of experimental error.

7.2.6 Combining ability

Combining ability analysis was performed separately for each experimental site where the mean squares for combining abilities (GCA and SCA), GCA and SCA effects, variances and standard errors were provided. Analysis was based on the fixed effect (model 1) method II as recommended by Griffing (1956) when using few parents aiming at comparing the combining abilities of the actual parents involved in the experiment with identification of superior combinations.

The GCA:SCA ratio indicating whether the characteristic is controlled by additive or non additive genes (dominant), was calculated using ratios of the mean squares for GCA and SCA.

7.2.7 Heterosis

Loci with no dominance do not express heterosis (Falconer and Mackay, 1996). The amount of heterosis following a cross between two particular lines or populations depends on the square of the difference of gene frequency between the populations. Mid parent heterosis was computed as:

Heterosis (MP)(%)=(F_1-MP)/MP*100, where F_1= performance of F_1, MP= mid parent value.

7.2.8 Heritability

Heritability expresses the proportion of the total variance that is attributable to the average effects of genes. Broad sense heritability estimates were obtained from:

$h^2 = V_G/V_P$.

Where: h^2 = Heritability, V_G = Genetic value, V_P = Phenotypic value. Heritability values were obtained using GCA and SCA values.

$h^{2=} \sigma^2 g/\sigma^2 p$:

Where: h^2 = Heritability, $\sigma^2 g$ = Total genetic variance = $2\sigma^2 gca + \sigma^2 sca$, $\sigma^2 p$ = Phenotypic variance = $\sigma^2 g + \sigma^2 e$.

Standard error deviation of the heritability estimates were computed as SE (h^2)=[2/n_1+2+2/n_2+2](1-h^2), where n_1 and n_2 = are degree of freedom of the genotypes and error variances following Tekouano *et al*. (2002).

7.2.9 Correlation

Correlation refers to a change in one characteristic that is accompanied by a change in another. Phenotypic correlation estimates were estimated from the following formula:

rp = covp/$\sigma_{p(x)} \sigma_{p(y)}$

where: rp = phenotypic correlation between characteristics x and y

covp = phenotypic covariance

$\sigma_{p(x)} \sigma_{p(y)}$ = the root of the genetic variance of x and y respectively.

Phenotypic correlation was calculated from the combined analysis data based on genotype values over environment and replications.

Genetic correlation estimation:

$r_{g(x,y)} = \sigma_{g(x,y)} / \sigma_{g(x)} \sigma_{g(y)}$:

Where: $r_{g(x,y)}$ = the genetic correlation coefficient between any two characteristics, y and y, $\sigma_{g(x,y)}$ = is the genetic covariance of x and y, $\sigma_{g(x)}\sigma_{g(y)}$ = are the root of the genetic variance of x and y respectively.

Genetic correlation estimation was based on GCA effects values. Agrobase (2000) statistical package was applied for correlation analysis.

7.3. RESULTS

7.3.1 Yield and yield components

Analysis of variance for genotypes under individual sites indicated the presence of significant differences among genotypes (Table 7.3). Bolls/plant were significantly different among the genotypes at Kanziga where Aubarn 56 x Acala SJ$_2$ (24.5), NTA 93-21 x MZ561 (23.8) and NTA 93-21 x Aubarn 56 (23.3) had many bolls, while UK91 (13.8) and NTA 93-21 x Delcot 344 (14.0) had significantly lower numbers of bolls/plant. No significant differences among genotypes were observed for other sites like Ukiriguru. This site had many bolls/plant for MZ561 x Okra leaf (34.8) and MZ561 (31.3) and the lowest number for Aubarn 56 x Okra leaf (17.3) (Table 7.4).

Seed/boll were significantly different among genotypes at two sites (Ukiriguru and Bwanga) (Table 7.3). At Ukiriguru, UK91 x Delcot 344 (36.8) and at Bwanga Aubarn 56 x Delcot 344 (36.8) had significantly higher seed/boll while NTA 93-21 x Delcot 344 (28.0) at Bwanga and Okra leaf (28.8) at Ukiriguru had a significantly lower number of seed/boll (Table 7.4). At Kanziga, Aubarn 56 (36.3) and Aubarn 56 x Acala SJ$_2$ (35.5) were the best and Acala SJ$_2$ (26.8) was the poorest, while at Mwanhala NTA 93-21 x Delcot 344 (35.5) and UK91 x Delcot 344 (35.5) were the best and Okra leaf (27.0) had a low seed number. Parents Aubarn 56 and Delcot 344 showed high numbers of seed/boll at all sites (Tables 7.4).

Table 7.3 Mean squares for seedcotton yield, lint yield, yield components, GOT, fibre quality, GCA, SCA and GCA:SCA ratio for cotton genotypes grown at four sites

S.variation	Df	Bolls/plant (no)				Seed/boll (no)				Boll weight (g)			
		Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Replication	3	96.07	161.84 *	36.32	118.58	26.30	33.18	34.77 *	12.06	0.48	0.63	0.82 *	0.72 *
Genotype	27	59.33	29.82 *	36.45	63.96	10.73 *	14.38	20.23 *	13.07	1.18 **	1.05 **	1.31 **	1.33 **
GCA	6	24.48	8.84 *	19.88 **	43.79 **	6.06 *	5.73	5.66	7.82	0.80 **	0.79 **	0.72 **	0.98 **
SCA	21	12.21	7.06 *	6.04	8.05	1.72	2.98	4.89	1.97	0.15	0.11	0.22 **	0.15 **
Residual	81	45.78	15.95	24.60	41.06	6.17	17.61	12.48	11.47	0.29	0.26	0.29	0.23
GCA:SCA		2.6:1	1.2:1	3.3:1	5.4:1	3.5: 1	1.9:1	1.2: 1	3.9:1	5.4:1	7.1:1	3.3:1	6.7:1

Table 7.3 Continued

S. variation	Df	GOT (%)				Seedcotton yield (kg/ha)				Lint yield (%)			
		Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Replication	3	10.41 *	8.16 *	10.99 *	1.83	2376492 *	576330	62723	465322	91240	48741	12530	79088
Genotype	27	39.40 **	28.39 **	25.81 **	37.64 **	1011152 **	706910 **	396766 **	577521 **	421880 **	139818 **	67917 **	14.3759 **
GCA	6	41.64 **	28.16 **	25.33 **	37.97 **	674153 **	338402 **	278316 **	379004 **	231738 **	82964 **	752 **	114156 **
SCA	21	0.77	1.08	1.06	1.25	132398 **	58147 *	48013	89242	69393 **	21237 *	369	13592
Residual	81	2.96	3.00	2.73	2.46	234850	232588	113748	356989	94311	30670	14997	54266
GCA:SCA		54.4:1	26.6:1	23.9:1	30.4:1	5.1:1	5.8:1	5.8:1	4.2:1	3.3:1	3.9:1	2:1	8.4:1

Table 7.3 Continued

S. variation	Df	Fibre length (mm)				Fibre strength (g/tex)				Micronaire (unit)			
		Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Replication	3	0.01	0.00	0.00	0.00	0.15	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Genotype	27	9.72 **	2.93 **	9.95 **	9.10 **	16.91 **	19.19 **	10.94 **	53.48 **	0.26 **	0.22 **	0.37 **	0.21 **
GCA	6	6.52 **	1.02 **	6.61 **	7.45 **	8.76 **	7.56 **	4.79 **	10.43 **	0.13 **	0.08 **	0.18 **	0.16 **
SCA	21	1.28 **	0.65 **	1.31 **	0.80 **	2.94 **	4.01 **	2.15 **	14.21 **	0.05 **	0.05 **	0.07 **	0.02 **
Residual	81	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.05	0.00	0.00	0.00	0.00
GCA:SCA		5:1	1.6:1	5:1	9.4:1	2:1	1.9:1	2.2:1	0.73:1	2.9:1	1.6:1	2.6:1	6.7:1

S. variation = source of variation, GCA = General combining ability, SCA specific combining ability, GOT = ginning outturn, Df = degree of freedom, ≤ 0.05,** p $\leq$ 0.01

Table 7.4 Means for seedcotton, lint yield, yield components, GOT and fibre quality characteristics obtained at four different environments

Genotype	Bolls/plant (no)				Seed/boll (no)				Boll weight (g)				GOT (%)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Okra leaf	21.3	18.0	17.0	18.0	28.8	30.8	29.3	27.0	4.3	4.8	4.3	3.9	33.8	33.8	33.3	33.5
Acala SJ$_2$ x Okra leaf	20.8	19.5	16.5	16.8	31.8	32.5	31.0	30.0	5.4	6.0	5.8	5.3	35.8	33.5	35.9	35.9
MZ561 x Okra leaf	34.8	18.3	18.3	20.0	30.0	32.3	32.0	30.3	5.3	5.8	5.4	5.2	35.5	36.0	35.2	36.5
Delcot 344 x Okra leaf	20.5	15.5	18.3	18.5	32.3	33.8	35.3	30.3	4.8	5.8	6.2	56	39.3	37.0	38.1	38.8
UK91 x Okra leaf	19.5	16.0	18.8	19.5	31.0	32.8	31.8	33.8	5.5	5.8	5.2	5.2	35.5	32.3	34.2	35.2
Aubarn 56 x Okra leaf	17.3	20.3	16.5	15.0	30.8	32.0	33.0	31.3	5.4	6.0	5.7	5.5	36.5	34.8	37.0	39.0
NTA 93-21 x Okra leaf	23.3	18.3	26.5	18.5	31.8	32.5	34.0	31.0	5.8	6.0	6.0	5.4	41.8	38.5	39.2	42.7
Acala SJ$_2$	18.8	15.8	18.0	15.8	31.0	26.8	30.5	30.8	5.7	6.5	5.6	5.1	34.3	34.5	37.1	36.7
MZ561 x Acala SJ$_2$	24.5	22.5	24.3	16.5	32.8	33.3	35.0	34.0	5.97	6.3	6.1	5.8	36.3	35.0	35.8	38.3
Delcot 344 x Acala SJ$_2$	20.0	19.8	18.3	13.8	34.0	33.8	34.0	32.8	6.3	7.0	6.9	6.4	39.0	36.5	36.5	38.6
UK91 x Acala SJ$_2$	21.5	17.5	17.8	17.8	32.3	34.0	34.5	34.0	5.8	6.8	6.4	5.7	34.8	34.0	33.3	35.0
Aubarn 56 x AcalaS J$_2$	21.8	24.5	17.3	17.3	34.3	35.5	32.0	32.0	6.3	6.8	6.4	5.7	37.8	35.8	35.6	38.6
NTA 93-21 x Acala SJ$_2$	25.7	19.3	21.0	21.3	31.3	35.0	34.3	33.0	6.1	6.5	6.5	6.3	40.8	39.3	38.8	42.4
MZ561	31.3	17.3	27.5	32.8	29.8	30.0	30.5	31.5	5.1	5.5	5.3	5.2	34.8	34.5	34.7	36.1
Delcot 344 x MZ561	18.3	20.5	17.0	18.0	32.0	34.3	32.8	32.8	6.1	6.5	6.4	6.0	40.5	39.5	38.1	39.1
UK91 x MZ561	25.0	20.0	21.5	27.0	33.0	31.5	33.5	30.5	6.0	6.0	5.7	5.6	36.3	33.8	36.2	38.4
Aubarn 56 x MZ561	24.0	16.8	23.5	20.3	33.0	32.8	36.0	32.3	5.6	6.5	5.9	5.9	37.3	36.8	37.2	38.9
NTA 93-21 x MZ561	24.5	23.8	23.5	22.8	31.0	33.3	34.8	32.8	6.6	6.8	6.1	6.2	40.0	39.8	39.3	42.1
Delcot 344	25.0	15.8	16.5	15.5	32.5	35.3	33.5	34.0	5.9	6.5	6.1	5.6	43.8	40.0	41.3	42.3
UK91 x Delcot 344	23.8	18.8	20.8	19.5	36.8	33.8	35.3	35.5	6.9	7.3	6.4	6.3	39.8	36.3	36.9	38.2
Aubarn 56 x Delcot 344	22.8	16.5	20.8	21.3	33.5	33.3	36.8	32.5	6.3	6.8	6.8	6.2	40.5	40.0	38.4	39.9
NTA 93-21 x Delcot 344	22.8	14.0	18.0	23.3	32.5	31.3	28.0	35.5	6.4	6.0	5.9	7.1	44.5	42.5	42.9	44.8
UK91	24.3	13.8	18.3	25.3	31.0	34.5	32.0	31.8	5.6	6.0	5.2	5.7	34.8	35.3	34.1	35.6
Aubarn 56 x UK91	22.0	19.0	19.3	16.0	32.5	33.8	32.8	33.0	5.8	6.3	6.0	5.5	36.8	35.8	37.0	37.3
NTA 93-21 x UK91	28.3	20.0	17.5	21.3	34.0	34.3	34.5	31.0	6.4	6.5	6.5	5.9	40.5	37.3	39.4	41.6
Aubarn 56	19.5	19.0	22.3	19.0	33.8	36.3	36.0	31.5	5.7	6.5	5.6	5.3	38.5	36.0	38.1	38.4
NTA 93-21 x Aubarn 56	27.3	23.3	20.5	18.5	32.8	34.0	33.5	31.5	6.2	6.5	6.4	6.3	43.0	40.8	41.6	44.3
NTA 93-21	21.3	17.5	20.8	21.5	31.0	34.0	28.8	33.8	5.6	6.0	5.3	5.9	43.3	40.8	41.7	44.8
Grand mean	**23.2**	**18.6**	**19.8**	**19.7**	**32.2**	**33.1**	**33.0**	**32.1**	**5.8**	**6.3**	**5.9**	**5.7**	**38.4**	**36.8**	**37.4**	**30.0**
CV (%)	11.3	14.4	12.1	16.3	7.7	12.6	10.7	10.5	9.2	8.2	9.2	8.4	4.5	4.7	4.4	4.0
LSD	7.9	4.7	5.8	7.5	2.9	4.9	4.2	3.9	0.6	0.6	0.6	0.6	2.0	2.04	1.9	1.8
Significant	NS	*	NS	NS	*	NS	*	NS	**	**	**	**	**	**	**	**
R-squared (%)	33.8	50	35	39	42.4	25	39	30	37.7	59	61	67	82.1	76	77	84

Table 7.4 Continued

Genotype	Seedcotton yield (kg/ha)				Lint yield (kg/ha)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Okra leaf	1357.2	1522.3	750.5	1059.3	704.3	511.4	248.9	356.8
Acala SJ$_2$ x Okra leaf	2533.5	2194.6	1657.1	1931.3	1061.3	727.4	591.8	685.9
MZ561 x Okra leaf	3241.5	2219.6	1936.9	1682.0	1410.6	798.0	686.1	616.3
Delcot 344 x Okra leaf	2289.8	1777.1	1702.0	1741.8	1154.4	657.2	640.1	671.9
UK91 x Okra leaf	2774.2	1998.1	1762.2	1805.8	1119.7	643.8	601.2	640.9
Aubarn 56 x Okra leaf	2606.2	2062.9	1744.4	1571.5	1028.0	717.9	645.4	611.5
NTA 93-21 x Okra leaf	3088.2	2199.1	1454.4	1521.5	1849.1	843.3	568.4	651.5
Acala SJ$_2$	2673.9	2389.3	1833.5	1377.8	1070.2	811.0	680.4	508.7
MZ561 x Acala SJ$_2$	2768.3	2397.4	2076.1	2026.3	1635.6	840.3	743.7	773.1
Delcot 344 x AcalaSJ$_2$	2865.5	2521.6	2021.5	1898.5	1316.0	917.8	737.3	734.6
UK91 x Acala SJ$_2$	3149.1	2871.5	2205.5	1817.5	1208.4	979.6	733.9	639.5
Aubarn 56 x Acala SJ$_2$	3033.9	2270.4	2075.9	2250.3	1494.6	812.4	740.7	872.9
NTA 93-21 x AcalaSJ$_2$	3494.5	3113.7	2119.0	2309.0	1674.6	1216.1	821.3	981.3
MZ561	2962.6	2009.5	1717.6	2313.5	1225.1	690.3	596.6	833.0
Delcot 344 x MZ561	2543.3	2497.4	1836.1	1864.5	1280.6	979.1	699.3	729.7
UK91 x MZ561	3408.4	2424.4	1794.9	2131.8	1343.5	817.7	648.5	813.4
Aubarn 56 x MZ561	3400.9	2602.6	2326.3	2181.3	1417.2	954.4	863.5	850.3
NTA 93-21 x MZ561	3485.4	3328.2	2334.4	2152.0	1685.9	1316.7	917.5	907.7
Delcot 344	2371.5	2113.1	1468.9	1550.3	1452.8	844.1	605.1	642.7
UK91 x Delcot 344	3524.3	2614.6	2058.3	1944.5	1770.8	949.0	747.5	742.3
Aubarn 56 x Delcot 344	3279.9	2291.9	2018.2	1820.0	1531.2	914.1	775.5	725.2
NTA 93-21 x Delcot 344	3100.6	2149.3	1799.7	2912.3	1682.1	913.8	770.1	1303.0
UK91	3328.1	1769.7	1873.1	2282.3	1385.2	612.2	638.4	813.5
Aubarn 56 x UK91	3319.6	2427.4	1939.9	1690.5	1377.9	866.3	718.5	625.3
NTA 93-21 x UK91	3543.8	2904.8	2174.7	2347.3	2020.4	1086.3	856.6	973.2
Aubarn 56	2637.2	2206.0	2019.6	1822.0	1353.1	798.3	769.8	697.7
NTA 93-21 x Aubarn 56	3704.9	3067.3	117.8	2215.8	2255.3	1249.6	877.5	980.5
NTA 93-21	2803.4	2074.9	1972.2	2477.3	1325.7	847.5	821.2	1108.6
Grand mean	**2974.6**	**2357.8**	**1885.4**	**1953.4**	**1422.6**	**868.7**	**705.2**	**767.9**
CV (%)	12.2	13.4	15.8	16.1	13.7	11.3	14.4	16.9
LSD	570.2	567.4	396.8	702.9	361.3	206.0	144.1	274.1
Significant	**	**	**	**	**	**	**	**
R-squared (%)	64.4	52	54	37	61.0	61	61	48

Table 7.4 Continued

Genotype	Micronaire (unit)				Fibre length (mm)				Fibre strength (g/tex)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Okra leaf	5.2	5.2	4.3	5.1	27.4	29.6	29.2	27.8	30.6	33.7	34.41	31.1
Acala SJ$_2$ x Okra leaf	5.0	5.2	4.7	4.7	29.8	31.8	30.5	29.3	33.1	35.7	30.61	31.5
MZ561 x Okra leaf	4.8	5.0	4.3	4.9	27.6	32.0	32.0	28.2	32.6	35.0	34.80	29.0
Delcot 344 x Okra leaf	4.9	5.1	4.4	4.8	28.8	32.1	31.3	29.5	31.7	38.7	33.80	35.1
UK91 x Okra leaf	4.8	4.8	4.2	4.8	30.4	32.5	31.1	28.8	33.4	35.5	29.80	31.7
Aubarn 56 x Okra leaf	5.2	4.9	5.0	5.3	29.9	30.7	30.7	29.9	34.3	36.4	31.81	33.9
NTA 93-21 x Okra leaf	5.0	5.6	4.8	5.0	31.6	31.4	30.9	31.2	36.8	34.7	32.11	32.5
Acala SJ$_2$	4.5	5.1	4.4	4.5	32.6	32.0	30.6	29.2	35.6	34.9	31.11	30.6
MZ561 x Acala SJ$_2$	4.8	4.8	4.7	5.0	28.6	31.2	28.8	29.8	33.2	35.6	29.80	30.7
Delcot 344 x Acala SJ$_2$	5.0	5.1	4.3	4.7	29.1	32.0	32.5	29.8	35.4	36.1	34.01	35.0
UK91 x Acala SJ$_2$	4.7	4.5	3.7	4.5	33.2	33.1	32.4	31.6	37.3	36.0	33.40	34.5
Aubarn 56 x Acala SJ$_2$	5.2	4.9	4.5	4.9	31.1	30.3	28.7	28.3	32.7	35.2	31.21	31.8
NTA 93-21 x Acala SJ$_2$	4.4	4.8	4.5	4.9	33.2	33.1	2.1	32.4	39.1	39.5	32.70	34.6
MZ561	4.6	4.8	4.6	4.8	30.7	31.3	33.5	28.6	33.8	33.1	32.70	31.5
Delcot 344 x MZ561	4.7	5.2	4.3	4.8	29.6	32.0	31.8	30.1	32.4	38.5	32.00	30.3
UK91 x MZ561	4.6	5.1	4.2	4.9	29.7	31.3	3.0	27.5	31.7	34.8	34.31	38.5
Aubarn 56 x MZ561	5.0	4.9	4.4	5.0	31.1	31.4	30.1	28.6	37.3	28.9	34.31	32.0
NTA 93-21 x MZ561	5.1	5.2	4.6	5.0	32.3	32.7	32.4	31.4	35.0	36.4	34.10	36.9
Delcot 344	5.3	5.1	4.3	4.8	31.5	32.0	31.2	30.9	33.2	38.2	31.10	35.5
UK91 x Delcot 344	4.7	4.8	4.5	4.5	30.8	33.5	33.5	32.0	36.3	36.6	31.61	35.9
Aubarn 56 x Delcot 344	5.3	4.8	4.4	4.9	29.8	32.2	31.4	31.0	33.7	34.5	32.02	31.8
NTA 93-21 x Delcot 344	5.1	5.0	4.9	5.2	31.8	31.3	33.7	32.2	37.5	36.5	34.01	34.2
UK91	4.5	4.6	3.7	4.6	30.7	30.8	31.0	31.2	35.5	34.0	32.11	29.7
Aubarn 56 x UK91	4.9	5.0	4.8	4.7	29.6	31.9	29.5	29.9	34.2	35.6	31.11	33.2
NTA 93-21 x UK91	5.0	5.0	4.3	4.8	32.2	32.6	34.4	32.5	34.7	34.1	34.11	35.1
Aubarn 56	4.6	4.8	4.8	4.9	28.9	31.2	30.9	28.4	36.3	36.9	29.51	33.4
NTA 93-21 x Aubarn 56	5.2	5.2	4.7	5.4	31.9	31.9	31.8	30.7	34.8	39.5	32.80	32.6
NTA 93-21	4.9	4.6	4.54	5.1	31.8	31.9	34.7	32.1	36.4	38.3	36.10	37.6
Grand mean	**4.9**	**4.9**	**4.4**	**4.9**	**30.6**	**31.8**	**31.6**	**30.1**	**34.6**	**35.8**	**32.5**	**33.6**
CV (%)	0.3	0.3	0.2	0.45	0.17	0.05	0.04	0.01	1.2	0.04	0.03	0.64
LSD	0.02	0.02	0.01	0.03	0.06	0.02	0.014	0.004	0.49	0.02	0.013	0.25
Significant	**	**	**	**	**	**	**	**	**	**	**	**
R-squared (%)	99	99	99	99	99	99	100	100	97	100	100	99

* ≤ 0.05,** p ≤ 0.01, CV= coefficient of variation, LSD= Least squared deviation, R-squared= Repeatability squared.

Boll weight varied largely at all sites (Table 7.3). The boll weight of UK91 x Delcot 344 (6.9 g) was the highest at Ukiriguru, Delcot 344 x Acala SJ$_2$ (6.9 g) at Bwanga, NTA 93-21 x Delcot 344 (7.1 g) at Mwanhala while UK91 x Delcot 344 (7.3 g) and Delcot 344 x Acala SJ$_2$ (7.0 g) had significantly higher boll weights at Kanziga compared to other crosses. Parent Okra leaf had a significantly lower boll weight at all sites [Ukiriguru (4.3 g), Kanziga (4.8 g), Bwanga (4.3 g) and Mwanhala (3.9 g)] (Table 7.4).

Significant differences were observed among genotypes for GOT (Table 7.3). Ginning outturn results for parent NTA 93-21 and Delcot 344 were very high at all four sites ranging between 40-44% and was transmitted to their progeny where increased GOT was observed. For example, at Ukiriguru NTA 93-21 x Delcot 344 (44.5%), Delcot 344 (43.75%) and NTA 93-21 (43.25%) were the best and at Kanziga NTA 93-21 x Delcot 344 (42.5%), NTA 93-21 x Aubarn 56 (40.8%) and NTA 93-21 (40.8%) were the best. At Bwanga NTA 93-21 x Delcot 344 (42.9%), NTA 93-21 (41.7%) and NTA 93-21 x Aubarn 56 (41.6%) were the best and at Mwanhala NTA 93-21 x Delcot 344 (44.8%), NTA 93-21 (44.8%) and NTA 93-21 x Aubarn 56 (44.3%) ranked first. Okra leaf ranked last at all sites except at Kanziga where UK91 x Okra leaf (32.3%) was the lowest (Table 7.4).

Seedcotton yield and lint yield were significantly different where crosses of NTA 93-21 yielded higher compared to other crosses and parents (Tables 7.3 and 7.4). At Ukiriguru, NTA 93-21 x Aubarn 56 (3704.9 kg/ha) followed by NTA 93-21 x UK91 (3543.8 kg/ha) and UK91 x Delcot 344 (3524.3 kg/ha) ranked first for seedcotton yield while on lint yield NTA 93-21 x Aubarn 56 (2255.3 kg/ha) followed by NTA 93-21 x UK91 (2020.4 kg/ha) and NTA 93-21 x Okra leaf (1849.18 kg/ha) were the highest. At Bwanga NTA 93-21 x MZ561 (2334.4 kg/ha) and Aubarn 56 x MZ561 (2326.3 kg/ha) were the highest for seedcotton yield while for lint yield NTA 93-21 x MZ561 (917.54 kg/ha) and NTA 93-21 x Aubarn 56 (877.5 kg/ha) ranked first. NTA 93-21 x MZ561 (3328.2 kg/ha), NTA 93-21 x Acala SJ$_2$ (3113.7 kg/ha) and NTA 93-21 x Aubarn 56 (3067.3 kg/ha) ranked first at Kanziga for seedcotton yield. For lint yield the same combinations were the best. At Mwanhala NTA 93-21 x Delcot 344 (2912.3 kg/ha) and NTA 93-21 (2477.3 kg/ha) ranked first for seedcotton yield and lint yield values were 1303kg/ha and 1108.6 kg/ha respectively. The parent Okra leaf had the lowest values for seedcotton yield and lint yield at all sites (Table 7.4).

7.3.2 Fibre quality

Fibre quality analysis indicated significant differences among genotypes evaluated at all sites (Table 7.3). Micronaire values were high at all sites for all genotypes (Tables 7.3 and 7.4) except at Bwanga where genotypes UK91 (3.7) and UK91 x Acala SJ$_2$ (3.7) had the lowest values and the rest were of medium micronaire values (4.0 - 4.9). Coarse fibre (5.0 - 5.9) was mainly observed at Ukiriguru on Delcot 344 (5.3), Aubarn 56 x Delcot 344 (5.3), Okra leaf (5.2), Aubarn 56 x Okra leaf (5.2), Aubarn 56 x Acala SJ$_2$ (5.2), NTA 93-21 x Aubarn 56 (5.2) and NTA 93-21 x MZ561 (5.1). High micronaire values were observed at Kanziga on NTA 93-21 x Okra leaf (5.6), NTA 93-21 x Aubarn 56 (5.2), NTA 93-21 x MZ561 (5.2), Delcot 344 x MZ561 (5.2), Acala SJ$_2$ x Okra leaf (5.2), Delcot 344 x Acala SJ$_2$ (5.2), Okra leaf (5.2) and UK91 x MZ561 (5.1) while at Mwanhala high micronaire values were observed for NTA 93-21 x Aubarn 56 (5.4), Aubarn 56 x Okra leaf (5.3) and NTA 93-21 x Delcot 344 (5.2). Okra leaf had high micronaire values as male parent and NTA 93-21 when used as female parent (Table 7.4).

Based on fibre length at Ukiriguru, NTA 93-21 x Acala SJ$_2$ (33.2 mm), UK91 x Acala SJ$_2$ (33.2 mm), Acala SJ$_2$ (32.6 mm), NTA 93-21 x MZ561 (32.3 mm) and NTA 93-21 x UK91 (32.2 mm) had very long fibres (above 32 mm) while others had long fibres except for MZ561 x Okra leaf (27.6 mm) and Okra leaf (27.4 mm) which had medium fibre lengths. At Mwanhala, NTA 93-21 x UK91 (32.5 mm), NTA 93-21 x Acala SJ$_2$ (32.4 mm), NTA 93-21 x Delcot 344 (32.2 mm) and NTA 93-21 (32.1 mm) ranked first with very long fibres, the rest were long except for UK91 x MZ561 (27.8 mm) and Okra leaf (27.8 mm) that were of medium length. At Kanziga, UK91 x Delcot 344 (33.5 mm), NTA 93-21 x Acala SJ$_2$ (33.1 mm) and UK91 x Acala SJ$_2$ (33.1 mm) had very long fibres and the rest had long fibres. At Bwanga all genotypes had long to very long fibres and no medium length was observed, however NTA 93-21 (34.7 mm) and NTA 93-21 x UK91 (34.4 mm) were the best (Table 7.4).

Fibre strength was significantly different among genotypes at all sites (Table 7.3). Values ranged from strong (26-29 g/tex) to very strong ($\geq$ 30 g/tex) fibre for all genotypes at all sites. However, the performance was different from site to site. For example, at Mwanhala UK91 x MZ561 (38.5 g/tex) and NTA 93-21 (37.6 g/tex) ranked first. At Bwanga NTA 93-21 (36.1 g/tex), MZ561 x Okra leaf (34.8 g/tex) and Okra leaf (34.4 g/tex) had the highest fibre strength.

At Kanziga, NTA 93-21 x Aubarn 56 (39.5g /tex) and NTA 93-21 x Acala SJ$_2$ (39.5 g/tex) had the highest values. For Ukiriguru, NTA 93-21 x Acala SJ$_2$ (39.1 g/tex) and NTA 93-21 x Delcot 344 (37.5 g/tex) were the highest (Table 7.4).

7.3.3 Combining ability

Combining ability mean squares for the characteristics are presented in Table 7.3. Significant GCA mean squares for yield components, seedcotton yield, lint yield and fibre quality (micronaire, length and strength) indicated that additive genes controlled most of the characteristics. Significant SCA mean squares for boll weight were observed at Bwanga and Mwanhala and for seedcotton yield and lint yield at Ukiriguru and Kanziga. For micronaire values, fibre length and fibre strength significant SCA mean squares were observed at all sites. GCA mean square values were higher compared to the mean squares for SCA except for fibre strength at Mwanhala (Table 7.3).

7.3.4 General combining ability effects

Results for GCA effects are given in Table 7.5. Okra leaf had negative GCA effects for all characteristics except for micronaire values at Kanziga, Bwanga and Mwanhala and fibre strength at Bwanga. Acala SJ$_2$, known to have big bolls, had positive and significant GCA effects for boll weight, seedcotton yield, fibre length and fibre strength at some sites. Significant negative GCA effects for Acala SJ$_2$ was recorded on GOT and micronaire values at all sites. GCA effects for MZ561 was positive and significant for bolls/plant at all sites, seedcotton yield at Ukiriguru, Bwanga and Mwanhala, lint yield at Bwanga, fibre length at Bwanga and fibre strength and micronaire values at Bwanga and Mwanhala. GCA effects for MZ561 on boll weight and GOT was negative at all sites. The GCA effect for Delcot 344 was positive for seed/boll at Ukiriguru and Mwanhala, boll weight and GOT at all sites, fibre length (except at Ukiriguru), fibre strength at Kanziga and Mwanhala and micronaire value at Ukiriguru and Kanziga. Negative GCA effects were observed for bolls/plant and seedcotton yield. Positive and significant GCA effects were shown by UK91 for fibre length while the values for GOT, lint yield and micronaire value were negative at all sites. Aubarn 56 had positive GCA effects on seed/boll, boll weight, GOT, micronaire value and seedcotton yield while negative GCA effect values were observed on fibre length and fibre strength.

Table 7.5   General combining ability (GCA) effects of yield components, seedcotton, lint, GOT and fibre quality for cotton genotypes grown on four different sites

Genotype	Bolls/plant (no)				Seed/boll (no)				Boll weight (g)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Okra leaf	-0.77	-0.56	-1.10	-1.46	-1.36	-0.84	-0.97	-1.83	-0.62	-0.59	-0.50	-0.63
Acala SJ₂	-1.55	0.64	-0.85	-2.51	0.12	-0.81	-0.27	0.00	0.07	0.24 *	0.20 *	-0.03
MZ561	3.12 *	0.83 *	2.71 *	3.68 *	0.66 *	-0.84	0.06	-0.17	-0.10	-0.15	-0.14	-0.05
Delcot 344	-0.83	-1.37	-1.46	-1.35	0.98 *	0.64	0.50	1.11 *	0.24 *	0.24 *	0.38 *	0.35 *
UK91	0.37	-1.12	-0.74	1.57 *	0.37	0.47	0.23	0.47	0.13 *	0.05	-0.11	-0.01
Aubarn 56	-1.27	1.05 *	0.35	-1.18	0.78 *	0.99 *	1.31 *	-0.17	0.05	0.18 *	0.10 *	-0.01
NTA 93-21	0.93	0.52	1.10 *	1.24 *	-0.22	0.36	-0.86	0.58 *	0.23	0.02	0.07	0.36 *
SE±Gi	1.04	0.62	0.76	0.99	0.38	0.65	0.56	0.52	0.08	0.08	0.08	0.07
SE±Gi-Gj	1.59	0.94	1.17	1.51	0.59	0.99	0.83	0.79	0.13	0.12	0.13	0.11

* P ≤ 0.05

Table 7.5 Continued

Genotype	GOT (%)				Seedcotton yield (kg/ha)				Lint yield (kg/ha)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Okra leaf	-1.70	-1.63	-1.44	-1.90	-505.47	-374.05	-369.48	-361.66	-261.03	-170.98	-17.82	-172.39
Acala SJ₂	-1.59	-1.24	-0.98	-1.12	-67.15	142.80*	82.18*	-71.08	-94.46	19.58	1.67	-48.73
MZ561	-1.31	-0.49	-0.88	-0.74	108.43*	69.56	72.95*	115.15*	-17.49	15.26	11.43*	23.70
Delcot 344	2.66*	1.96*	1.59*	1.28*	-161.20	-87.14	-78.85	-38.44	28.86	7.78	-1.01	7.66
UK91	-1.56	-1.60	-1.55	-1.71	286.49*	-9.15	66.49*	74.93	25.56	-42.24	-1.33	-9.04
Aubarn 56	0.19	0.18	0.44*	0.28	91.44*	30.23	130.95*	-28.21	47.70	17.99	0.21	-9.10
NTA 93-21	3.33*	2.82*	2.82*	3.91*	247.47*	227.74*	95.76	309.31*	270.86*	152.61*	6.85*	207.91*
SE±Gi	0.27	0.27	0.25	0.24	74.78	74.42	52.04	92.19	47.39	27.02	4.70	35.95
SE±Gi-Gj	0.40	0.41	0.39	0.37	114.22	113.67	79.49	140.82	72.38	41.28	7.18	54.91

* P ≤ 0.05

Table 7.5 Continued

Genotype	Fibre length (mm)				Fibre strength (g/tex)				Micronaire (unit)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Okra leaf	-1.29	-0.50	-0.84	-0.92	-1.51	-0.35	0.07 *	-1.41	-0.11	0.15 *	0.05 *	0.08 *
Acala SJ₂	0.61	0.14 *	-0.60	-0.15	0.58 *	0.15 *	-0.69	-1.04	-0.12	-0.03	-0.05	-0.15
MZ561	-0.47	-0.11	0.19 *	-0.91	-0.77	-1.24	0.40 *	0.19 *	-0.11	0.03	0.02 *	0.02 *
Delcot 344	-0.17	0.32 *	0.46 *	0.64 *	-0.37	1.19 *	-0.05	0.53 *	0.13 *	0.05 *	-0.01	-0.05
UK91	0.32	0.25 *	0.28 *	0.43 *	0.22 *	-0.68	-0.18	1.07 *	-0.16	-0.15	-0.27 *	-0.18
Aubarn 56	-0.37	-0.39	-0.94	-0.62	0.33 *	-0.29	-0.88	-0.72	0.10 *	-0.04	0.20 *	0.11 *
NTA 93-21	1.35 *	0.29 *	1.46 *	1.52 *	1.54 *	1.19 *	1.32 *	1.38 *	0.05 *	0.03 *	0.06 *	0.17 *
SE±Gi	0.01	0.00	0.00	0.00	0.06	0.00	0.00	0.03	0.00	0.00	0.00	0.03
SE±Gi-Gj	0.01	0.00	0.00	0.00	0.09	0.01	0.00	0.05	0.00	0.00	0.00	0.01

* P ≤ 0.05, SE± (Gi) = standard error for any GCA effect, SE± (Gi-Gj)= standard error of the difference between any two GCA effects

The observed GCA effect for NTA 93-21 was positive and significant for bolls/plant, GOT, seedcotton, lint yield, fibre length, fibre strength and micronaire at all sites (Table 7.5). NTA 93-21 showed good combining ability for GOT, seedcotton yield, fibre length and fibre strength and outperformed all other parents (Table 7.5). This was followed by Delcot 344, which was a good combiner for seeds/boll, boll weight, GOT, fibre strength and fibre length. Aubarn 56 was a good combiner for seeds/boll and boll weight and average for seedcotton and lint yield. UK91 was a good combiner for fibre length and an average combiner for seedcotton and fibre strength. MZ561 was a good combiner for bolls/plant and seedcotton yield. Acala SJ_2 was an average combiner for boll weight, seedcotton yield, fibre length and fibre strength except at Ukiriguru. Okra leaf was a poor combiner for all characteristics studied except micronaire value and moderate for fibre strength (Table 7.5). Generally NTA 93-21 had positive and significant GCA effects for almost all characteristics at all sites, however values differed according to the environment and parents involved in crossing.

7.3.5 Specific combining ability effects

SCA effect estimates for GOT, boll weight, seedcotton yield and lint yield are presented in Table 7.6. Significant SCA effects for some characteristics indicated variation among F_1 hybrids. For GOT, nine combinations had positive and significant SCA effects, boll weight four combinations and seedcotton yield and lint yield both eight combinations. Combinations having positive and significant SCA effects for seedcotton yield had significant and positive SCA effects for lint yield except Delcot 344 x Acala SJ_2 that had significant SCA effect for seedcotton yield but not for lint yield, as well as MZ561 x Acala SJ_2 that had high SCA effect only for lint yield (Table 7.6). MZ561 x Okra leaf, NTA 93-21 x Acala SJ_2, NTA 93-21 x MZ561, UK91 x Delcot 344 and NTA 93-21 x Aubarn 56 were positive and significant for three of the four characteristics. Acala SJ_2 x Okra leaf, MZ561 x Acala SJ_2 and NTA 93-21 x UK91 had two significant SCA effects among the four characteristics. Delcot 344 x Okra leaf, UK91 x Okra leaf, UK91 x Acala SJ_2, Aubarn 56 x Delcot 344 and Aubarn 56 x UK91 were poor combinations with low SCA effects for all characteristics. Delcot 344 x Acala SJ_2 had negative SCA for GOT. UK91 x Acala SJ_2 had negative SCA for GOT, boll weight and lint yield. Aubarn 56 x UK91, Delcot 344 x MZ561 and UK91 x MZ561 had negative and significant SCA effects for seedcotton yield and lint yield (Table 7.6).

Table 7.6 Specific combining ability (SCA) effects for GOT, boll weight, seedcotton yield and lint yield for cotton combinations

Hybrid	GOT	Boll weight	Seedcotton yield	Lint yield
Acala SJ$_2$ x Okra leaf	0.31	0.03	167.20 **	99.13 **
MZ561 x Okra leaf	0.50 **	0.23	288.30 **	111.86 **
Delcot 344 x Okra leaf	0.03	-0.02	78.80	14.09
UK91 x Okra leaf	-0.29	0.09	90.20	2.08
Aubarn 56 x Okra leaf	0.36 **	0.09	50.10	-32.49
NTA 93-21 x Okra leaf	1.10 **	0.34	-44.40	43.64
MZ561 x Acala SJ$_2$	0.49 **	-0.02	-89.00	73.23 **
Delcot 344 x Acala SJ$_2$	-0.90	0.17	103.60 **	0.55
UK91 x Acala SJ$_2$	-0.75	-0.02	91.70	-18.03
Aubarn 56 x Acala SJ$_2$	0.05	1.30 **	37.00	12.90
NTA 93-21 x Acala SJ$_2$	0.41 **	0.03	224.60 **	79.83 **
Delcot 344 x MZ561	0.47 **	0.17	-107.60	-36.02
UK91 x MZ561	0.62 **	0.03	-49.20	-34.95
Aubarn 56 x MZ561	0.12	-0.05	187.30 **	46.78
NTA 93-21 x MZ561	0.03	0.43 **	220.30 **	81.14 **
UK91 x Delcot 344	-0.40	0.51 **	229.30 **	110.77 **
Aubarn 56 x Delcot 344	-0.31	0.09	95.00	11.00
NTA 93-21 x Delcot 344	0.76 **	0.01	69.00	40.51
Aubarn 56 x UK91	0.12	-0.26	-109.20	-61.53
NTA 93-21 x UK91	0.26	0.24 **	125.10	125.35 **
NTA 93-21 x Aubarn 56	0.98 **	0.04	207.50 **	197.61 **
SE± (S$_i$)	0.35	0.24	100.24	55.01
SE± (S$_i$-S$_j$)	0.48	0.33	138.34	75.92

** $p \leq 0.01$, SE± (S$_i$) = standard error for any SCA effect, SE± (S$_i$-S$_j$)= standard error of the difference between any two SCA effects

The GCA:SCA ratio variation (Table 7.3) was larger than one for all characteristics except for fibre strength at Mwanhala where SCA was higher than GCA and the ratio was 0.73:1. A high GCA:SCA ratio value was observed for example on GOT, where the ratio ranged from 23.9 to 54.4.

7.3.6 Heritability

Heritability (h^2) and correlations are both presented in Table 7.7. The estimates for h^2 for the studied genotypes indicated high heritability values for almost all studied characteristics ranging from 0.47 to 0.97. Ginning outturn (0.97 $\pm$ 0.02) had the highest heritability followed by seedcotton yield and lint yield with a heritability of 0.85 $\pm$ 0.1 each and fibre length (0.83 $\pm$ 06).

Table 7.7   Genetic correlation (r_g) (below diagonal), phenotypic correlation (r_p) (above diagonal) and broad sense heritability (h^2) (diagonal and bold) for yield components, seedcotton yield, lint yield, GOT and fibre quality

	Boll weight	Bolls/plant	Seed/boll	GOT	100 seed weight	Seedcotton yield	Lint yield	Fibre length	Micronaire	Fibre strength
Boll weight	**0.74±0.3**	-0.167 **	0.420 **	0.220 **	0.415 **	0.429 **	0.344 **	0.275 **	0.041	0.214 **
Bolls/plant	-0.119 *	**0.80±0.4**	0.076	-0.014	0.107 *	0.449 **	0.467 **	-0.053	0.053	0.067
Seed/boll	0.504 **	-0.429 **	**0.73±0.3**	0.038	0.181 **	0.205 **	0.136 **	0.072	-0.069	0.060
GOT	0.517 **	0.073	0.384	**0.97±0.02**	-0.159 **	0.115 *	0.330 **	0.150 **	0.279 **	0.201 **
100 seed weight	0.664 **	-0.041	0.004	0.249	**0.47±0.9**	0.172 **	0.062	0.165 **	-0.161 **	0.043
Seedcotton yield	0.670 **	0.553 *	0.079	0.410	0.249	**0.85±0.1**	0.878 **	0.093	0.191 **	0.281 **
Lint yield	0.691 **	0.427 *	0.199	0.778 **	0.273	0.889 **	**0.85±0.1**	0.049	0.260 **	0.263 **
Fibre length	0.664 **	0.135	-0.002	0.531 *	0.268	0.795 **	0.825 **	**0.83±0.6**	-0.127 **	0.249 **
Micronaire	-0.062	-0.395	0.448	0.542 *	-0.073 *	-0.455 *	-0.047	-0.319 *	**0.74±0.02**	0.197 **
Fibre strength	0.729 **	0.082	0.157	0.570 **	0.201	0.809 **	0.853 **	0.942 **	-0.167 *	**0.74+0.02**

* $p \leq 0.05$, ** $p \leq 0.01$, GOT= ginning outturn

Seed/boll (0.73 ± 0.3) and boll weight (0.74 ± 0.3) had moderate heritability levels and hundred seed weight was the lowest heritability with 0.47 ± 0.9. Fibre quality heritability values revealed that fibre length had the highest (0.83 ± 0.6) heritability value followed by fibre strength and micronaire, with similar values (0.74 ± 0.02).

7.3.7 Correlations

Phenotypic correlations for most of the characteristics were positive and significant compared to genetic correlations (Table 7.7). Seedcotton yield had a highly significant positive correlation with bolls/plant (0.449), seed/boll (0.205), boll weight (0.429), 100 seed weight (0.172), lint yield (0.878), micronaire (0.191) and fibre strength (0.281) for phenotypic correlation. Seedcotton yield had a highly significant positive genetic correlation with bolls/plant (0.553), boll weight (0.670), lint yield (0.889), fibre length (0.795) and fibre strength (0.809), while micronaire value (-0.455) was significantly negatively correlated. Highly and significantly positive phenotypic and genetic correlations seen between yield (seedcotton and lint) and bolls/plant and boll weight indicated that these two characteristics were important to cotton yield. Bolls/plant had a significant negative phenotypic and genetic correlation with boll weight, indicating a negative association of these characteristics.

Micronaire value had a negative genetic correlation with boll weight and bolls/plant. Hundred seed weight and fibre length were significantly negatively correlated to micronaire value for both phenotypic and genetic correlations. The significant positive correlation between boll weight and seed/boll indicated the positive association for the traits governing these characteristics that might be common or linked. For both correlation studies GOT had high positive correlation with lint yield. Micronaire value was positively and significantly phenotypically correlated with seedcotton yield, lint yield, GOT and fibre strength. Fibre strength and seedcotton were negatively genetically correlated to micronaire. However, fibre length was positively phenotypically and genetically correlated with seedcotton yield, lint yield and fibre strength (Table 7.7).

7.3.8 Heterosis estimates

Heterosis values for the 21 combinations on four different sites varied from negative to positive (Table 7.8). Bolls/plant heterosis was negative for more than 50% of the combinations especially at Ukiriguru, Bwanga and Mwanhala. However, NTA 93-21 x Aubarn 56 (33.3) and MZ561 x Okra leaf (32.8) had high positive heterosis, while Delcot 344 x MZ561 (-35.1) had the highest negative heterosis for bolls/plant at Ukiriguru. Aubarn 56 x Acala SJ$_2$ (41.1) and NTA 93-21 x MZ561 (36.7) had high heterosis at Kanziga. At Bwanga NTA 93-21 x Okra leaf (40.4) had high heterosis, while NTA 93-21 x Delcot 344 (25.7) had the highest heterosis at Mwanhala.

Heterosis values for seed/boll were positive for most of the combinations. UK91 x Delcot 344 (15.8) had high heterosis at Ukiriguru, MZ561 x Acala SJ$_2$ (17.2) at Kanziga, NTA 93-21 x MZ561 (17.3) at Bwanga and UK91 x Okra leaf (14.9) at Mwanhala (Table 7.8). Boll weight heterosis values were positive and high at all sites except for Delcot 344 x Okra leaf (-5.9) at Ukiriguru, NTA 93-21 x Delcot (-4.0) at Kanziga and Aubarn 56 x UK91 (-0.6) at Mwanhala. At Ukiriguru the highest heterosis value was recorded for NTA 93-21 x MZ561 (21.4), at Kanziga for NTA 93-21 x MZ561 (17.4) and at Mwanhala for NTA 93-21 x Delcot 344 (23.6). At Bwanga NTA 93-21 x Okra leaf (26.3) had the highest heterosis and at this site no negative heterosis was observed among combinations for boll weight (Table 7.8).

Heterosis estimates recorded on combinations for GOT varied from site to site, for example at Ukiriguru NTA 93-21 x Okra leaf (8.4), NTA 93-21 x Aubarn 56 (5.2), Acala SJ$_2$ x Okra leaf (5.2) and NTA 93-21 x Acala SJ$_2$ (5.2) had high heterosis. At Kanziga NTA 93-21 x Aubarn 56 (6.2) and Delcot 344 x MZ561 (6.0) had the highest heterosis values. At Bwanga the highest heterosis was recorded on UK91 x MZ561 (5.3) and at Mwanhala NTA 93-21 x Okra leaf (9.1) had high heterosis value. Aubarn 56 x Delcot 344 (-1.5) had the lowest heterosis value at Ukiriguru, UK91 x Okra leaf (-6.5) at Kanziga, Delcot 344 x Acala SJ$_2$ (-6.8) at Bwanga and UK91 x Acala SJ$_2$ (-3.1) at Mwanhala. When Acala SJ$_2$ was used as male parent for GOT character development the combinations had low positive to negative mid parent heterosis, for example at Bwanga Delcot 344 x Acala SJ$_2$ (-6.8), UK91 x Acala SJ$_2$ (-6.5), Aubarn 56 x Acala SJ$_2$ (-5.3) and NTA 93-21 x Acala SJ$_2$ (-1.4) (Table 7.8).

Table 7.8 Mean mid-parent heterosis for yield components, yield and fibre quality characteristics of grown cotton on four sites

Hybrid	Bolls/plant				Seed/boll				Boll weight			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Acala SJ$_2$ x Okra leaf	5.0	15.6	-5.7	-0.7	6.3	13.0	3.8	3.9	8.2	6.7	17.4	16.0
MZ561 x Okra leaf	32.8	3.6	-17.9	-21.2	2.6	6.2	7.1	3.4	12.6	12.2	12.1	12.6
Delcot 344 x Okra leaf	-11.3	-8.2	8.9	10.5	5.3	2.3	12.4	-0.8	-5.9	2.2	19.2	16.5
UK91 x Okra leaf	-14.3	0.8	6.4	-9.8	3.8	0.4	3.7	14.9	10.0	6.9	9.7	7.4
Aubarn 56 x Okra leaf	-15.3	9.5	-15.9	-18.9	-1.6	-4.5	1.2	6.8	6.7	6.7	15.9	18.1
NTA 93-21 x Okra leaf	9.4	2.8	40.4	-6.3	6.3	0.4	17.2	2.1	17.1	11.6	26.3	9.9
MZ561 x Acala SJ$_2$	-1.0	36.4	6.6	-31.9	7.8	17.2	14.8	9.2	11.2	8.7	13.0	13.2
Delcot 344 x Acala SJ$_2$	-7.5	25.4	5.8	-12.0	7.1	8.8	6.3	1.2	7.9	12.0	18.9	20.4
UK91 x Acala SJ$_2$	2.4	18.6	-2.1	-13.4	4.0	11.0	10.4	8.8	2.5	12.5	18.9	4.6
Aubarn 56 x Acala SJ$_2$	15.2	41.0	-14.3	-0.7	5.8	12.7	-3.8	2.8	10.1	8.0	14.7	8.3
NTA 93-21 x Acala SJ$_2$	27.8	15.5	8.4	14.1	0.8	15.2	15.6	2.3	8.3	8.3	20.6	14.8
Delcot 344 x MZ561	-35.1	24.2	-22.7	-25.4	5.4	4.9	2.3	0.0	10.8	8.3	12.8	11.9
UK91 x MZ561	-10.1	29.0	-6.0	-6.9	8.6	-2.3	7.2	-3.6	12.9	4.4	8.8	3.3
Aubarn 56 x MZ561	-5.4	-7.6	-5.5	-21.7	3.9	-1.1	8.3	2.4	4.2	8.3	8.5	12.1
NTA 93-21 x MZ561	-6.7	36.7	-2.6	-16.1	2.1	3.9	17.3	0.4	21.4	17.4	15.3	11.0
UK91 x Delcot 344	-3.6	27.1	19.4	-4.3	15.8	-3.2	7.6	7.9	19.4	16.0	13.5	11.4
Aubarn 56 x Delcot 344	2.2	-5.0	4.5	23.2	1.1	-6.9	5.8	-0.8	7.5	3.9	15.6	14.2
NTA 93-21 x Delcot 344	-1.3	-15.8	-3.4	25.7	2.4	-9.8	-10.0	4.8	11.7	-4.0	5.1	23.6
Aubarn 56 x UK91	0.6	16.0	-4.9	-27.7	0.4	-4.6	-3.8	4.4	2.0	0.0	12.1	-0.6
NTA 93-21 x UK91	24.2	28.0	-10.3	-9.1	9.7	0.0	13.6	-5.3	13.3	8.3	23.3	1.0
NTA 93-21 x Aubarn 56	33.3	27.4	-4.7	-8.6	1.2	-3.2	3.5	-3.5	10.5	4.0	17.9	12.0

Table 7.8 continued

Hybrid	GOT				Seedcotton yield				Lint yield			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Acala SJ$_2$ x Okra leaf	5.2	-1.8	2.1	2.1	25.7	12.2	28.3	58.5	19.6	10.0	27.4	58.5
MZ561 x Okra leaf	3.7	5.5	3.7	4.9	50.1	25.7	56.9	-0.3	46.2	32.8	62.3	3.6
Delcot 344 x Okra leaf	1.3	0.3	2.3	2.3	22.8	-2.2	53.4	33.5	7.0	-3.0	49.9	33.1
UK91 x Okra leaf	3.7	-6.5	1.4	1.9	18.4	21.4	34.3	8.1	7.2	14.6	35.5	9.5
Aubarn 56 x Okra leaf	1.0	-0.4	3.7	8.5	30.5	10.7	25.9	9.1	-0.1	9.6	26.7	15.9
NTA 93-21 x Okra leaf	8.4	3.4	4.6	9.1	48.5	22.3	6.8	-13.9	82.2	24.1	6.2	-11.1
MZ561 x Acala SJ$_2$	5.1	1.5	-0.3	5.2	-1.8	9.0	16.9	9.8	42.5	11.9	-48.8	15.2
Delcot 344 x Acala SJ$_2$	0.0	-2.0	-6.8	-2.4	13.6	12.0	22.4	29.6	4.3	10.9	14.7	26.5
UK91 x Acala SJ$_2$	0.7	-2.5	-6.5	-3.1	4.9	38.1	19.0	-0.7	-1.6	37.7	11.3	-3.3
Aubarn 56 x Acala SJ$_2$	3.8	1.4	-5.3	2.6	14.3	-1.1	7.8	40.7	23.4	0.9	2.2	44.7
NTA 93-21 x Acala SJ$_2$	5.2	4.3	-1.4	4.1	27.6	39.5	11.4	19.8	39.8	46.6	9.4	21.4
Delcot 344 x MZ561	3.2	6.0	0.4	-0.3	-4.6	21.2	15.2	-3.5	-4.4	27.6	16.4	-1.8
UK91 x MZ561	4.3	-3.2	5.3	7.2	8.4	28.3	-0.0	-7.2	2.9	25.6	5.0	-1.2
Aubarn 56 x MZ561	1.7	4.3	2.2	4.3	21.5	23.5	24.5	5.5	9.9	28.2	26.4	-99.6
NTA 93-21 x MZ561	2.6	5.7	2.9	4.1	20.9	63.0	26.5	-10.2	32.2	71.3	29.4	-6.5
UK91 x Delcot 344	1.3	-3.7	-2.2	-1.8	23.7	34.7	23.2	1.5	24.8	30.3	20.2	1.3
Aubarn 56 x Delcot 344	-1.5	5.3	-3.3	-1.1	30.9	6.1	15.7	7.9	9.1	11.3	12.8	7.4
NTA 93-21 x Delcot 344	2.3	5.3	3.5	3.0	19.8	2.6	4.6	44.6	21.1	8.1	7.9	47.9
Aubarn 56 x UK91	0.3	0.4	2.6	0.9	11.3	22.1	-0.3	-17.6	0.6	22.8	2.1	-17.2
NTA 93-21 x UK91	3.9	-1.9	4.1	3.4	15.6	51.1	13.1	-1.4	49.1	48.8	17.4	1.3
NTA 93-21 x Aubarn 56	5.2	6.2	4.2	6.4	36.2	43.3	4.0	3.1	68.4	51.9	-38.0	8.6

Table 7.8 Continued

Hybrid	Fibre length				Fibre strength				Micronaire			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
Acala SJ$_2$ x Okra leaf	-0.6	3.3	2.0	3.1	-0.16	4.1	-6.6	2.1	2.9	0.9	7.5	-1.9
MZ561 x Okra leaf	-5.1	5.2	2.1	-0.1	1.4	4.9	3.7	-7.3	-1.9	0.0	-3.7	-1.7
Delcot 344 x Okra leaf	-2.2	4.2	3.6	0.3	-0.7	7.6	3.2	5.4	-6.7	-0.7	2.1	-2.9
UK91 x Okra leaf	4.6	7.6	3.3	-2.2	1.1	4.9	-10.4	4.3	-1.2	-1.8	3.1	-2.5
Aubarn 56 x Okra leaf	6.1	0.9	2.2	6.4	2.5	3.1	-0.5	5.1	6.0	-0.6	10.7	5.1
NTA 93-21 x Okra leaf	6.7	2.1	-3.2	4.2	9.9	-3.6	-8.9	-5.4	-0.8	14.1	9.8	-2.5
MZ561 x Acala SJ$_2$	-9.5	-1.4	-10.1	3.1	-4.2	4.7	-6.6	-1.1	5.6	-3.2	4.9	8.4
Delcot 344 x Acala SJ$_2$	-9.1	0.0	5.2	-0.8	2.9	-1.2	9.3	5.9	1.9	0.1	-1.6	1.7
UK91 x Acala SJ$_2$	4.9	5.4	5.2	4.6	4.8	4.6	5.7	14.4	4.4	-7.4	-9.9	-1.4
Aubarn 56 x Acala SJ$_2$	1.0	-4.1	-6.7	-1.8	-9.2	-1.9	2.9	-0.6	14.4	-0.8	-2.9	5.4
NTA 93-21 x Acala SJ$_2$	3.2	3.6	3.5	5.7	8.7	7.9	-2.7	1.3	-6.2	-1.0	-2.8	0.7
Delcot 344 x MZ561	-4.8	1.1	-1.7	1.1	-3.2	8.0	0.3	-9.5	-5.4	5.2	-3.0	1.1
UK91 x MZ561	-3.3	0.8	0.8	-8.1	-8.4	3.7	5.9	58.6	1.4	8.5	1.9	5.9
Aubarn 56 x MZ561	4.3	0.5	-6.5	0.5	6.5	-17.4	10.3	-1.2	8.5	1.9	-7.1	4.8
NTA 93-21 x MZ561	3.3	3.5	-4.9	3.6	-0.1	1.9	-0.9	6.7	7.7	10.4	2.0	1.5
UK91 x Delcot 344	-0.9	6.8	7.7	3.1	5.7	1.5	0.1	10.1	-4.2	-1.1	11.3	-4.9
Aubarn 56 x Delcot 344	-1.4	1.9	1.1	4.5	-3.0	-8.1	5.7	-7.6	6.4	-3.1	-2.5	1.9
NTA 93-21 x Delcot 344	0.5	-2.1	2.3	2.2	7.9	-4.6	1.2	-6.4	-0.1	3.2	11.9	3.9
Aubarn 56 x UK91	-0.9	2.9	-4.7	0.7	-4.7	0.4	0.9	5.1	7.6	6.4	13.9	0.1
NTA 93-21 x UK91	3.0	3.9	4.7	2.8	-3.3	-5.7	0.1	4.3	6.6	8.9	6.2	-0.6
NTA 93-21 x Aubarn 56	4.8	1.1	-3.0	1.7	-4.2	5.1	-0.1	-8.2	9.2	10.6	1.9	8.5

GOT = ginning outturn

High heterosis estimates were observed for seedcotton yield and lint yield. Hybrids with NTA 93-21 as female parent had high values for seedcotton yield at all sites, for example at Ukiriguru NTA 93-21 x Okra leaf (48.5) and NTA 93-21 x Aubarn 56 (36.2) while the lowest was for Delcot 344 x MZ561 (-4.6). At Kanziga NTA 93-21 x MZ561 (63.0), NTA 93-21 x UK91 (51.1) and NTA 93-21 x Aubarn 56 (43.3) had the highest values while the lowest was for Delcot 344 x Okra leaf (-2.2). At Bwanga, NTA 93-21 x MZ561 (26.5) had lower heterosis values compared to MZ561 x Okra leaf (56.9) and Delcot 344 x Okra leaf (53.0), while Aubarn 56 x UK91 had the lowest heterosis (-0.3). Acala SJ_2 x Okra leaf (58.5) had high heterosis on seedcotton yield at Mwanhala and the lowest heterosis value recorded for this site was Aubarn 56 x UK91 (-17.6). The same was observed on lint yield where NTA 93-21 as a female parent exhibited combinations with high heterosis values. Many combinations at Mwanhala had negative heterosis for seedcotton yield compared to other sites. MZ561 produced negative heterosis when used as male parent for seedcotton yield and lint yield especially at Mwanhala [Aubarn 56 x MZ561 (-99.6), NTA 93-21 x MZ561 (-6.5), Delcot 344 x MZ561 (-1.8) and UK91 x MZ561 (-1.2)] (Table 7.8).

Heterosis for fibre length behaved different from site to site except for NTA 93-21 x Acala SJ_2, Aubarn 56 x Okra leaf, UK91 x Acala SJ_2 and NTA 93-21 x UK91 that had positive heterosis at all sites (Table 7.8). NTA 93-21 x Okra leaf (6.7) and Aubarn 56 x Okra leaf (6.1) had high positive heterosis at Ukiriguru. At Kanziga UK91 x Okra leaf (7.6) and at Bwanga Delcot 344 x Acala SJ_2 (5.2) and UK91 x Acala SJ_2 (5.2) were the highest. Aubarn 56 x Okra leaf (6.4) had the highest heterosis at Mwanhala. For fibre strength the highest heterosis value was recorded for NTA 93-21 x Okra leaf (9.9) and NTA 93-21 x Acala SJ_2 (8.7) at Ukiriguru, Delcot 344 x MZ561 (8.0) at Kanziga, Aubarn 56 x MZ561 (10.3) at Bwanga and UK91 x MZ561 at Mwanhala (18.6) (Table 7.8).

Heterosis for micronaire value was high when NTA 93-21 and Aubarn 56 were used as female parents. However, high micronaire values cause coarse fibres, thus breeding has to balance the traits for medium micronaire values. The highest heterosis value for micronaire was observed on combinations of Aubarn 56 x Acala SJ_2 (14.4) at Ukiriguru and NTA 93-21 x Okra leaf (14.1) at Kanziga. At Bwanga Aubarn 56 x UK91 (13.9) and at Mwanhala NTA 93-21 x Aubarn 56 (8.5) had the highest heterosis values for micronaire. When Okra leaf and Acala SJ_2 were used as

male parents it produced negative heterosis for micronaire. This was observed for micronaire results for MZ561 x Okra leaf, Delcot 344 x Okra leaf and UK91 x Okra leaf and for Acala SJ_2 the negative heterosis was observed for combinations UK91 x Acala SJ_2, Aubarn 56 x Acala SJ_2 and NTA 93-21 x Acala SJ_2 (Table 7.8).

7.4 DISCUSSION

There was a high level of variation for characteristics studied among parents and the F_1 combinations. This suggested that selection could be done for improved yield and yield components, GOT and fibre quality. Parents varied significantly for GCA estimates. Significant values obtained from NTA 93-21, Delcot 344, Aubarn 56 and MZ561 indicated the possibility of good combining ability from these parents for the targeted characteristics. Genotypes with high positive GCA estimate for yield components for example MZ561 (bolls /plant) and Delcot 344 (boll weight) exhibited negative GCA estimates for fibre quality (fibre length and fibre strength) and vice versa, excluding NTA 93-21. This indicated that these characteristics are negatively correlated and could lead to difficulties in breeding programmes. These results agree with that reported by Thomson and Luckett (1988) and Coyle and Smith (1997). High GCA effects observed on lint yield and yield components were reported by El-Adl and Miller (1971) and Echekwu and Alaba (1995). These results indicated strong additive genetic variance.

High negative GCA effect estimates observed for Okra leaf for seedcotton yield were expected due to its low rate of boll retention. However, Okra leaf was included to exploit its drought tolerance, early maturity, high fibre strength and high micronaire traits in order to increase variability in cotton populations. Impact assessments of traits from Okra leaf will be done in further studies.

Positive and significant SCA effects were observed for seedcotton yield, boll weight, GOT and lint yield for some of the combinations. Significant SCA mean squares observed for boll weight was reported by Echekwu and Alaba (1995). The performance of these combinations indicated the possibility of improvement of these characteristics in the cotton programme. This was confirmed from the results that combinations with high SCA effects their parents had high GCA effects for the particular characteristics (Tables 7.5 and 7.6). Alam *et al*. (1992) reported that

152

significant mean squares for both GCA and SCA for specific characteristics (as observed on fibre quality) indicated that both additive and dominant gene effects were responsible for these characteristics. Singh (2002) suggested that SCA effects do not significantly contribute to the improvement of self-pollinated crops, except where commercial exploitation of heterosis is feasible. However, in self-pollinated crops like cotton, the additive x additive type of component is fixable in later stages.

Griffing (1956) and Machado *et al.* (2002) reported that crosses with high SCA values from parents with the highest GCA in a population should be efficient in choosing the more promising segregating population. Based on the obtained results, this depends on the type of genotype and characteristics selected for. According to Griffing (1956), GCA estimates close to zero indicate that genotype does not differ from the general mean of all crosses. Positive or negative values indicate that the respective parent is better or worse than the other parent included in a diallel as compared to the mean of the cross. Delcot 344 was reported as a good combiner for total yield and fibre strength (Tang *et al.,* 1993b; Meredith and Brown, 1998). In the current study Delcot 344 was a good combiner for yield components (seeds/boll and boll weight), GOT and fibre strength. Therefore, the high and significant positive GCA estimates generated by NTA 93-21, UK91 and Aubarn 56 on seedcotton yield, NTA 93-21 on lint yield, NTA 93-21 and Delcot 344 on GOT, fibre length and fibre strength, Aubarn 56 and Delcot 344 on seed/boll; MZ561, Aubarn 56 and NTA 93-21 on bolls/plant, and Aubarn 56, Acala SJ$_2$ and Delcot 344 on boll weight suggested that through breeding programmes these parents will generate enough variable material for improving cotton genotypes in Tanzania.

The GCA:SCA ratios for the studied characteristics were higher than one indicating the presence of additive genetic effects for most of the characteristics studied, except for fibre strength at Mwanhala. El-Dahan *et al.* (2003) obtained similar results. According to Ashraf and Ahmad (2000), high additive genetic variation for these characteristics suggested a possibility of improvement in these characteristics. Therefore normal breeding methods such as pedigree, backcrossing, or recurrent selection would be required to accumulate the additive genes in order to increase seedcotton yield and fibre quality.

High positive heterosis for boll weight was observed for more than 90% of the combinations. This agrees with Sambamurthy *et al.* (1995) who reported that in tetraploid cotton, boll weight and boll number for intraspecific hybrids are the major components of heterosis in yield and that this is usually observed in *G. hirsutum* crosses and not for *G. barbadense*. However, significant heterosis for boll weight does not signify high seedcotton yield. It can be assumed that increased seedcotton yield does not only depend on high heterosis for boll weight but either on boll weight in combination with other yield components like seed/boll, bolls/plant and seed weight. You *et al.* (1998) suggested that if mid-parent heterosis of boll weight and bolls per plant increased, mid-parent heterosis of seed/boll decreased, significantly improving mid-parent heterosis of lint yield. Results of this study are in agreement with Ashwathama *et al.* (2003). Hybrid combinations with high SCA values presented high values of heterosis (Tables 7.6 and 7.8). This association suggested that heterosis can be used for the choice of potential material where SCA cannot be estimated. Low heterosis values for micronaire were reported by Subrahmanyam *et al.* (1989). Xian *et al.* (1995) and Zhang and Zhang (1997) reported high heterosis for seedcotton and lint yield. These reports agree with the results shown in the current study.

The significant levels of phenotypic variation observed for the studied characteristics may involve additive components as suggested by Mather and Jinks (1977) and Khan *et al.* (2003). Genotypes with many seed/boll had large bolls and high seedcotton yield because as number of seed increased, boll weight increased which increased seedcotton yield. Lint yield can be increased, but depends on seedcotton yield and GOT of the genotype. This was reflected by the positive association present between yield components and seedcotton yield. High positive genetic and phenotypic correlations observed for seedcotton and lint yield agreed with findings by Gencer and Guvercin (2003). Azhar *et al.* (2004) reported significant phenotypic and genetic correlations for seedcotton yield with micronaire values and fibre strength. In the present population studied, seedcotton yield was genetically negatively correlated with micronaire. Fibre length and fibre strength were positively correlated phenotypically and genetically, indicating the association of these two fibre quality characteristics. This was supported by reports of Thomson and Luckett (1988) and Tang *et al.* (1996). However, micronaire was negatively phenotypically and genetically correlated to fibre length. A significant positive correlation observed between GOT and lint yield agrees with results of Jixiang *et al.* (1996) and Hussain *et al.* (1998).

Heritability values for the studied characteristics were moderate to high. This indicated that the phenotypic variance was additive in this material, although some non-additive action was observed. Dedaniya and Pethani (1994) reported high heritability values for seedcotton yield and lint yield as observed in the current study. Heritability estimates for fibre quality were high for fibre length while fibre strength and fibre fineness (micronaire) were similar. High heritability values for fibre length compared to other fibre quality characteristics agree with results reported by Tang *et al*. (1992; 1996) and Azhar *et al.* (2004). High heritability values indicated that in a future cotton breeding programmes for yield, yield components and fibre quality it is important to apply backcrossing in order to concentrate traits in the genotypes because many characteristics seems to be controlled by additive genes.

Improvement of cotton genotypes aims at combining high yield, high GOT values and improved fibre quality simultaneously. The observed negative correlation between bolls/plant and boll weight as well as seedcotton yield with fibre quality, will complicate the breeding process. However, the high variation level revealed in the population indicated that it was possible to identify good combiners and good combinations with positive and significant GCA and SCA effects respectively for those characteristics. Cotton breeding aims at increased seedcotton yield, resistance to insects and diseases and adaptability (characteristics important to producers). On the other hand, cotton breeding aims at increased lint yield, increased GOT and good fibre quality (characteristics for industrial ginners and textiles). Results indicated variation in the material evaluated especially for yield components, yield, GOT and fibre quality. Thus improvements for these characteristics in cotton breeding programmes are possible.

7.5 CONCLUSIONS AND RECOMMENDATIONS

Tanzanian cotton varieties have low to medium yield (seedcotton and lint) and fibre quality. Hence varieties with improved characteristics are required. Parents with high and positive GCA like NTA 93-21, Delcot 344, Aubarn 56 and MZ561 and combinations with high and positive SCA for the characteristics studied were identified during this study. This in turn, showed potential to increase variability in the cotton population in Tanzania and consequently improved cotton varieties. Estimated heterosis and heritability values indicated the presence of additive traits controlling the characteristics studied except for few characteristics that had both additive

and non-additive genes controlling the same characteristics. High GCA effects are mostly due to additive gene effects or additive x additive interactive effects (Griffing, 1956). Therefore breeders may utilise good general combiners in breeding programmes for improvement of cotton characteristics. It is recommended that breeders should breed for superior combining ability aimed at improving overall GCA for yield, yield components and fibre quality.

A correlation study among characteristics indicated positive and negative association among them. Therefore, since the study was done for one season, further studies on this population is necessary. However, preliminary results indicated that breeding techniques like recurrent selection, pedigree and backcrossing should be applied from the early stages because most of the characteristics seem to be controlled by additive traits. Backcrossing should be used to increase additive traits and selection for superior combinations. Therefore, through further studies and application of the above methods, this population will improve the cotton industry in Tanzania. It can furthermore be concluded that more exotic cotton material has to be imported from other areas to increase variability in the cotton population.

CHAPTER 8

GENOTYPE X ENVIRONMENT INTERACTION AND STABILITY ANALYSIS FOR TANZANIAN COTTON (*GOSSYPIUM HIRSUTUM* L.) GERMPLASM

8.1 INTRODUCTION

Genotype expression over environments is referred to as genotype by environment interaction (G x E) and is a main challenge facing plant breeders (Paterson *et al*., 2003). G x E interaction is of primary importance in many aspects of genomic research and is a special priority in the study of major crops grown in a wide range of environments (Saranga *et al.,* 2001). Detection of significant G x E interaction indicates that all phenotypic responses to agro-ecological conditions are not similar for all genotypes. According to Romagosa and Fox (1993) some genotypes may perform well in some environments but not so well in others. G x E interaction is a differential genotypic expression across environments and is important to geneticists and breeders because the magnitude of the interaction components provides information concerning the adaptation of a given variety (Myers, 2004). G x E interactions have assumed greater importance in plant breeding as they reduce the stability of genotypic values under diverse environments. The impact of the environment on phenology and growth varies depending on the crop species, crop variety and growth stages. G x E interaction may change the performance or development of a crop, therefore the extent of the environmental effect on a trait determines the importance of testing over locations and years. However, development of new crop varieties is expensive and time consuming (Romagosa and Fox, 1993).

Conventional analysis of variance assessment of G x E interaction becomes difficult as number of sites and years increases and due to multiple factors influencing the environment (Raun *et al.,* 1993). The term stability, thus, refers to the character of a crop that withstands fluctuations of environments (Myers, 2004). Yates and Cochran (1938) first proposed stability analysis for the interpretation of genotype by environment interaction. Their methodology held for the linear regression of variety yield on experimental "mean yield" in order to observe varietal stability across varying environments (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966). More recently, stability analysis has been adapted for use in comparing agronomic treatments across

different environments consisting of the linear regression of treatment mean yield on the environmental mean (Raun *et al.,* 1993; Unruh and Silvertooth, 1997).

The concept of stability has been defined in several ways and several biometrical methods including univariate and multivariate (Hill, 1975; Lin *et al.*, 1986; Becker and Léon, 1988; Crossa, 1990). Stability indices are usually univariate where as a genotype's response to varying environments is multivariate (Lin *et al.*, 1986). Through multivariate analysis, genotypes with similar responses cluster together and the data can be summarised and analysed more easily (Crossa, 1990).

Cotton originated from wild perennial plants adapted to semi-arid and subtropical environments that experience periodic drought and temperature extremes (Kohel, 1974). An understanding of G x E interaction for cotton genotypes based on their ability to meet defined targets for yield and/or fibre quality traits is important in the selection process by considering the ability of a specific line to meet one or more of these defined targets over a wide range of testing environments (Myers, 2004). Yield is a useful reference point from which to examine cotton G x E interactions. For example, in genetically equivalent cotton populations grown under well-watered and water-limited conditions results indicated that the water-limited condition was responsible for 50% yield reduction relative to the well-watered condition (Saranga *et al.,* 2001).

Tanzania presently has two main cotton growing areas, Western and Eastern. The WCGA is further divided into two zones, northern and southern because of environmental variation and specific adapted varieties for each zone (Jones and Kapingu, 1982; Ramadhani *et al.,* 1998; Lukonge and Ramadhani, 1999; TCL and SB, 2002). Cotton free market in the cotton industry, however, has resulted in seed mixing and disease (fusarium wilt) transmission (Hau, 1997; Go, 1999a; TCL and SB, 2002). This problem has resulted in impure varieties with poor fibre quality. The effort of having a single variety for the entire WCGA has been discussed and is still under discussion. The strategy to solve the problem, therefore is to breed for a variety that will suit both the northern and the southern zones.

Diallel crosses among random mating inbred lines, have long been used to understand the nature of genetic variation and the potential values of inbred lines (Mather and Jinks, 1977). The aims of this study were therefore to 1) evaluate the growth response of genotypes under four environments and to estimate the components of variance associated with interactions and 2) assess stability through five stability procedures for 21 diallel F_1 progeny and parents from Tanzania for the WCGA's.

8.2 MATERIALS AND METHODS

8.2.1 Plant material and experimental sites

Twenty eight cotton genotypes (including 21 F_1s and seven parents) (Tables 7.1 and 7.2) were studied. Material was selected according to differences in characteristics like maturity duration, presence of one or more non-preference traits for pests, resistance to disease, adaptability to environment, good fibre quality, good yield components and high yield (seedcotton and lint). Origin and description of the varieties are presented in Table 7.1 and section 7.2.1.

The study was conducted at four different cotton experimental sites of the WCGA's in Tanzania namely at Ukiriguru Research Institute, Kanziga and Bwanga (cotton experimental sites) (representing the northern cotton growing zone) and Mwanhala experimental site (representing the southern cotton growing zone) (Figure 7.1). The northern zone received more rain than the southern zone (Table 8.1), however the weather fluctuates year after year. The four different experimental sites, apart from different rainfall, had different maximum and minimum temperatures (Tables 8.1), soil types, elevation, latitude and longitude.

8.2.2 Soil sampling and analysis

Soil for the four sites were sampled randomly at six points at each site at two depths between 0-20 cm and between 30-50 cm. Soil from six points of the same depth was mixed thoroughly to form a composite sample for analysis. Samples were analysed at the Zonal Soils Laboratory, Ukiriguru Research Institute using standard methods as described in the National Soil Service (NSS, 1991). Data were obtained for pH (acidity/alkalinity), electrical conductivity

Table 8.1 Rainfall and temperature of the experimental sites for the cotton growing season August 2004 to May 2005

Month	Rainfall (mm)				Maximum/minimum Temperatures (C^0)			
	Ukiriguru	Kanziga	Bwanga	Mwanhala	Ukiriguru	Kanziga	Bwanga	Mwanhala
August	27.6	0.0	23.4	0.0	30.6/17.0	-	24.1/15.0	21.0/15.7
September	50.0	93.9	55.6	23.9	29.8/17.0	-	23.7/16.1	23.0/17.6
October	125.6	44.1	49.5	35.4	30.2/17.0	-	25.1/20.4	24.7/19.1
November	212.1	188.6	131.0	204.1	30.5/17.2	-	23.1/18.1	30.8/18.9
December	151.4	188.6	106.5	163.7	29.5/17.2	-	22.3/17.8	22.1/19.2
January	44.3	72.3	84.2	108.6	30.5/16.9	-	22.4/17.4	24.0/22.1
February	71.6	26.7	63.6	106.2	30.8/18.0	-	23.1/17.2	28.4/17.4
March	115.5	170.2	155.5	122.6	31.0/17.8	-	22.7/18.1	23.2/18.0
April	114.6	121.2	55.5	95.4	30.1/17.0	-	22.3/17.7	23.1/17.3
May	58.4	111.4	107.3	25.1	30.0/17.0	-	21.9/17.1	22.8/17.1
Total (10 months)	856.5	1017.0	832.1	885.1				
Average (10 months)	85.7	101.7	83.2	88.5	30.3/17.2	-	23.1/17.2	24.3/18.3
Weather (annual)	1000-1500	750-1000	1000-1500	750-1000	>26/15	19-26/13-15	>26/15	>26/15

(EC), organic carbon (OC), total nitrogen (TN), available phosphorus (P), cation exchange capacity (CEC), exchangeable cations/bases [calcium (Ca), magnesium (Mg), potassium (K) and sodium (Na)] and carbon/nitrogen (C/N) ratio.

8.2.3 Land preparation and field management
Land preparation was done as described in section 7.2.2 and 7.2.3.

8.2.4 Data collection
Five plants in each plot were randomly selected for data recording. Data collected before harvesting were plant height, number of bolls per plant and seed per boll. Sampling was done as described in section 7.2.4.

8.2.5 Data analysis
8.2.5.1 Analysis of variance
ANOVA was performed on yield, yield components, GOT and fibre quality for individual sites (section 7.2.5). Sites data were pooled and combined analysis for four different environments was performed using Agrobase (2000). Results were used to estimate the performance of genotypes across locations, G x E interaction effect and stability. Contribution percentages of genotypes and location to the total variation were estimated. After removing the replicate effect when combining data, G x E observations were partitioned into two sources of variation, (1) additive main effects for genotypes and environment and (2) non- additive main effect due to G x E interaction. ANOVA of combined data expresses the observed Y_{ij} mean yield of the ith genotype at the jth environment.

$$Y_{ij} = \mu + G_i + E_j + GE_{ij} + e_{ij.}$$

Where μ = general mean, G_i, E_j and GE_{ij} = the effect of genotype, environment and G x E interaction respectively and e_{ij} = the average of random errors associated with the rth plot that receives the ith genotype in the jth environment. The non-additive interaction *(GE $_{ij}$)* as defined in the above equation, implies that an expected value of the ith genotype in Y_{ij} depends not only on the level of *G* and *E* separately, but also on the particular combination of levels of *G* and *E* (Crossa, 1990).

8.2.5.2 Stability analysis

Five stability analysis procedures were used during this study to compare genotypes to understand which is the appropriate stability parameter for cotton. Agrobase (2000) was used for all analyses.

Lin and Binns cultivar performance measure (P_i)

The data set was analysed according to the procedure recommended by Lin and Binns (1988). The values estimated are the squares of the differences between genotype mean and the maximum genotype's mean at a location, summed and divided by twice the number of locations.

$$P_i = \frac{[n(Y_i - M..)^2 + (Y_{ij} - Y_i + M_j + M..)^2]}{2n}$$

Where Y_{ij} is the average response of the ith genotype in the jth environments, Y_i is the mean deviation of the genotype i, M_j is the genotype with maximum response among all genotypes in the jth locations and n is the number of locations. The smaller the value of P_i the less its distance to the genotype with maximum yield and the better the genotype.

Wricke's ecovalence (Wi)

According to Wricke (1962), ecovalence *(Wi)* is the contribution of each genotype to the genotype x environment interaction sum of squares. The ecovalence *(Wi)* or stability of the ith genotype is its interaction with environments, squared and summed across environments. Wricke stated that genotypes with low ecovalence have smaller fluctuations from the mean across different environments and are therefore more stable.

$$W_i = \sum_j [Y_{ij} - Yi - Y_j + Y..]^2$$

Where Y_{ij} is the mean performance of genotype i in j^{th} environment and Y_i and Y_j are the genotype and environment mean deviations respectively and $Y..$ is the overall mean.

Shukla procedure of stability variance

Shukla (1972) defined stability variance (SV) (σ^2_i) of genotype i as its variance across environments after the main effects of environmental means had been removed. Since the

genotype main effect is constant, the stability variance is based on the residual ($GE_{ij} + e_{ij}$) matrix.

$$Y_{ij} = \mu + a_i + E_j + g_{ij} + e_{ijk}$$

Where Y_{ij} is the yield of the ith genotype in the kth replicate of the jth environment, μ is the overall mean, a_i is the effect of ith genotypes, E_j is the effect of jth environment, g_{ij} is the interaction of ith genotype in the jth environment and e_{ijk} is the the random error

Eberhart and Russell's joint regression analysis

The important estimates for the procedure proposed by Eberhart and Russell (1966), are the regression coefficient *(bi)*, the deviation from regression for each genotype (S^2d_i) and the mean yield (kg/ha). The model $Y_{ij} = \mu_i + \beta_{ij} + \delta_{ij} + I_j$ defines stability parameters that may be used to describe the performance of a genotype over a series of environments. Y_{ij} is the genotype mean of the ith genotype at the jth environment, μ is the ith genotype mean over all environments, β_{ij} is the regression coefficient that measures the response of the ith genotype to varying environments, δ_{ij} is the deviation from regression of the ith genotype at the jth environment and I_j is the environmental index.

Additive Main Effects and Multiplicative Interaction Method (AMMI Model)

The additive main effect and multiplicative interaction (AMMI) method integrates analysis of variance and principal component analysis (PCA) into a unified approach (Gauch, 1988) and is especially useful in analysing multi-location trials (Gauch and Zobel, 1988). The AMMI analysis first fits the additive main effects of genotypes and environments by the usual analysis of variance and then describes the non-additive part and the G x E interaction by PCA. The AMMI model does not make provision for a specific stability measure to be determined and such a measure is essential in this study in order to rank genotypes in terms of stability. Purchase (1997) proposed the formula to calculate AMMI's stability value (ASV) as follows:

$$ASV = \sqrt{\left[\frac{\text{IPCA 1 SS}}{\text{IPCA 2 SS}} \ (\text{IPCA 1scores}) \right]^2 + \left[\text{IPCA 2scores} \right]^2}$$

ASV is the distance from zero in a two dimensional scattergram of IPCA 1 (Interaction Principal Component Analysis 1) scores against IPCA 2 scores.

Where SS = sums of squares

8.2.6 Combined comparison of stability analysis procedures

Spearman's coefficient of rank correlation (Steel and Torrie, 1980) was established statistically to compare the five stability analysis procedures, variety performance value *(Pi)*, ecovalence *(Wi)*, stability variance *(SV)*, deviation from regression *(S^2d_i)* and AMMI stability value *(ASV)* used in this study in the form of ranks. Spearman's rank correlation coefficient *(r_s)* can be described as:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Where $d_i = X_i - Y_i$ *(i = 1,2,……,n)* and d_i ranking numbers are whole numbers and when two or more equal ranking numbers occur, the average of the ranking numbers that they otherwise would have received, are ascribed to each genotype, *n* are genotypes arranged in the same order according to two stability parameters, X_i indicates the ranking number of the ith genotype for the first parameter and Y_i indicates the ranking number for the ith genotype of the second parameter. Correlation coefficients were estimated using Agrobase (2000). The significance of r_s can be tested by means of Student's *t* test, with *n-2* degrees of freedom where:

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1 - r_s^2}}$$

8.3 RESULTS

8.3.1 Soil

Soil types and soil nutrients results are presented in Table 8.2. Soil pH varied along the locations and ranged from 5.7-7.1 (Table 8.2). Bwanga had a pH range of 5.8-6.3 (slightly to medium acid), Kanziga 5.7-7.1 (medium acid to neutral), Mwanhala 5.8-6.0

Table 8.2 Description of soil analysis characteristics for the four sites

Site	Depth (cm)	pH H$_2$O	pH 1M KCl	Soil type	Soil local name	OC (%)	TN (%)	C/N ratio	Available P (mg/kg soil)	EC dS/m 25^{0}C	Na	K	Ca	Mg	TEB	CEC (me/10g)	Elevation (m asl)	Latitude (^{0}S)	Longitude (^{0}E)
Ukiriguru	0-20	6.7	6.7	Eutric Nitosol	Luseni	0.98	0.11	10	21.7	0.064	1.02	0.63	4.10	1.40	5.81	7.15	1236	2^{0}42'S	33^{0}1E
	30-50	7.0	7.0			0.43	0.04	10	5.6	0.047	1.58	0.84	5.46	2.40	10.28	12.40			
Kanziga	0-20	5.7	4.8	Eutric Nitosol	Luseni/Itogolo	1.29	0.13	10	60.2	0.141	0.76	0.74	4.81	1.30	7.61	9.42	1200	2^{0}58'S	33^{0}1E
	30-50	7.1	6.1			0.61	0.06	10	21.0	0.082	1.80	0.98	6.37	2.00	11.15	13.69			
Bwanga	0-20	5.8	4.8	Chronic Cambisol	Ibushi	0.96	0.10	10	7.0	0.200	0.97	0.54	3.51	0.98	5.95	7.14	1200	3^{0}02'S	31^{0}42'E
	30-50	6.3	5.5			0.61	0.06	10	4.9	0.041	1.36	0.84	5.46	1.84	9.50	11.40			
Mwanhala	0-20	5.8	5.0	Gleyic Luvisol	Itogolo plain	0.93	0.09	10	26.6	0.134	0.96	0.81	5.27	1.72	8.76	10.32	1250	4^{0}24'S	33^{0}09'E
	30-50	6.0	5.3			0.66	0.06	10	6.3	0.057	1.20	1.02	6.63	2.20	11.05	13.50			

OC= organic carbon, TN= total nitrogen, C/N= carbon nitrogeny ratio, Available P = available phosphorus, EC = electric conductivity, Exch. Cat. = exchangeable bases, Na = sodium, K = potassium, Ca = calcium, Mg = magnesium, CEC = cation exchange capacity, masl = metre above sea level

(medium acid) and Ukiriguru 6.7-7.0 (very slightly acid to neutral). Organic carbon at Kanziga was of medium range and other sites had low organic carbon. Total nitrogen ranged from 0.04-0.13 at all sites indicating low to very low nitrogen levels. Available phosphorus (P) ranged from 4.9-7.0 mg/kg at Bwanga, 21.0-60.2 mg/kg at Kanziga, 6.3-26.6 mg/kg at Mwanhala and 5.6-21.7 mg/kg at Ukiriguru (Table 8.2). High values were observed in the top soil while low values were observed in the subsoil. Mwanhala and Ukiriguru had high P in top soil and low P in subsoil while Kanziga soil had high P value for top and subsoil. Bwanga had medium P in the top soil and low P in subsoil.

The capacity of soils to retain and release nutrients especially N, P, and K for plants uptake (CEC) ranged from 7.14-13.69 cmol/kg (Table 8.2). All four sites had low CEC indicating low soil fertility. Sodium, potassium and phosphorus results indicated that sodium increased with depth. At Kanziga sodium levels in the soil was high. At Bwanga the sodium level was similar to other sites, however, potassium, calcium, phosphorus and magnesium levels were low (Table 8.2).

8.3.2 Yield and yield components

Analysis of variance (ANOVA) for yield components, yield (seedcotton and lint), GOT and fibre quality for individual sites showed significant differences among the mean squares for genotypes for all characteristics studied except for bolls/plant at Ukiriguru and seed/boll at Kanziga and Mwanhala (sections 7.3.1 and 7.3.2). Combined analysis for pooled data of the four locations showed significant differences among the mean squares for locations and genotypes for most of the characteristics and G x E interaction was significantly different for plant height, lint yield, fibre strength, fibre length and micronaire value (Table 8.3).

Plant height was significantly different for genotypes, locations and G x E (Table 8.3). MZ561 x Okra leaf (96.6 cm), MZ561 (95.6 cm) and Acala SJ_2 x Okra leaf (95.6 cm) were significantly taller across locations. Okra leaf (81.4 cm), NTA 93-21 x Delcot 344 (80.1 cm), NTA 93-21 x Aubarn 56 (79.9 cm) and Delcot 344 (77.7 cm) were significantly shorter (Table 8.4). The contribution of location to total variation was higher (28%) compared to genotype (12%) (Table 8.3).

Table 8.3 Combined analysis mean squares for cotton characteristics for four different environments

Source of Variation	Df	Plant height (cm) SS	MS	Cont. SS (%)	Bolls/plant (no) SS	MS	Cont SS (%)	Seed/boll (no) SS	MS	Cont. SS (%)	Boll weight (gm) SS	MS	Cont SS (%)	GOT (%) SS	MS	Cont. SS (%)
Total	447	86912			17997			5857			246			4881		
L.	3	23943	7981 **	28	1321	441 **	7	96	32 *	2	20	7.0 **	8	344	115 **	7
R(L)	12	5499	458 **	19	1238	103 **	64	319	27 *	71	8	0.7 **	38	94	8 *	20
G.	27	10386	385 **	12	2255	84 **	13	851	32 **	15	108	4.0 **	44	3294	122 **	68
GxL	81	11497	142 *	41	2862	35	16	726	9	12	24	0.3	10	249	3	5
Error	324	35585	110		10319	32		3866	12		87	0.3		902	3	
LSD (L)		2.3			1.2			0.8			0.1			0.4		
LSD (G)		6.1			3.3			2.0			0.3			1.0		
CV (%)		11.9			17.8			10.6			8.8			4.4		

Table 8.3 continue

Source of variation.	Df	Seedcotton yield (kg/ha) SS	MS	Cont SS (%)	Lint yield (kg/ha) SS	MS	Cont SS (%)	Fibre length (mm) SS	MS	Cont SS (%)	Fibre strength (g/tex) SS	MS	Cont SS (%)	Micronaire (units) SS	MS	Cont SS (%)
Total	447	243155001			73760125			1174			3405			48		
L	3	84028414	28009472 **	35	36150439	12050125 **	49	173	58.0 **	15	673	224.0 **	20	19	6.00 **	40
R(L)	12	10442602	870217 **	34	994798	82899	22	3	0.2	6	0.5	0.1	1	1.0	0.01 **	2
G	27	5038184	1865489 **	21	1410183	522292 **	19	452	17.0 **	39	896	33.0 **	26	16	0.60 **	33
GxL	81	22325270	275621	9	6779193	83694 **	9	472	6.0 **	40	1818	22.0 **	53	12	0.20 **	25
Error	324	75990531	234539		15733802	48561		75	0.2		18	0.1		0.1	0.00	
LSD (L)		106.8			48.6			0.1			0.1			0.0		
LSD (G)		282.4			128.5			0.3			0.2			0.0		
CV (%)		21.1			23.4			1.6			0.7			0.3		

* $p \leq 0.05$, ** $p \leq 0.01$, L= location, G= genotypes, Df= degrees of freedom, SS= sum of square, MS= mean square, LSD(L)= Least squared deviation for location, LSD(G)= Least squared deviation for genotype, CV= coefficient of variation, R(L)= reps in location, GxL= genotype x location, Cont SS = contribution sum of square

Table 8.4 Combined analysis means for cotton genotypes on yield, yield components, GOT and fibre quality for four different environments

No	Genotype	Height (cm)	B/plant (no)	Seed/boll (no)	B/weight (gm)	GOT (%)	Seedcotton (kg/ha)	Yield Lint (kg/ha)	Fibre length (mm)	Fibre strength (g/tex)	Micronaire (units)
1	Okra leaf	81.4c	18.4e	28.9d	4.3e	33.6e	1172.3d	455.3d	30.3bcd	32.5c	4.9ab
2	Acala SJ$_2$ x Okra leaf	95.2ab	18.4e	31.3c	5.6d	35.3c	2079.1c	766.6c	30.3bcd	32.7bc	4.9abc
3	MZ561 x Okra leaf	96.6a	22.8ab	31.1c	5.4d	35.8c	2270.0bc	877.8c	29.9cd	32.6bc	4.7c
4	Delcot 344 x Okra leaf	85.6c	18.2e	32.9bc	5.6d	38.3abc	1877.7c	780.9c	30.4bcd	34.8ab	4.8b
5	UK91 x Okra leaf	92.6abc	18.4e	32.3bc	5.4d	34.3e	2085.1c	751.4c	30.7b	32.6bc	4.6d
6	Aubarn 56 x Okra leaf	89.3bc	17.3e	31.8c	5.6d	36.8bc	1996.3c	750.7c	30.3bcd	34.1abc	5.1a
7	NTA 93-21 x Okra leaf	88.4bc	21.6b	32.3bc	5.8bc	40.5ab	2065.8c	978.1abc	31.3abc	34.0abc	5.1a
8	Acala SJ$_2$	87.8bc	17.1e	29.8d	5.7c	35.7c	2068.6c	768.8c	31.1abcd	33.1b	4.6d
9	MZ561 x Acala SJ$_2$	91.2b	21.9ab	33.8b	6.0abc	36.3c	2317.0b	998.1abc	29.6c	32.3c	4.8bcd
10	Delcot3 44 x Acala SJ$_2$	83.6c	17.9e	33.6b	6.7a	37.6b	2326.6b	926.4abc	30.9b	35.1a	4.8bcd
11	UK91 x Acala SJ$_2$	93.1abc	18.7e	33.7b	6.1ab	34.3e	2510.9ab	890.4bc	32.6a	35.3a	4.3d
12	Aubarn 56 x Acala SJ$_2$	90.4b	20.2e	33.4b	6.3ab	36.9bc	2407.6ab	980.1abc	29.6c	32.7bc	4.9abc
13	NTA 93-21 x Acala SJ$_2$	87.9bc	21.7abc	33.4b	6.4a	40.3ab	2759.3a	1173.3a	32.9a	36.5a	4.6d
14	MZ561	95.6a	27.2a	30.4d	5.3d	35.0d	2250.8bc	836.3c	31.0abcd	32.8bc	4.7c
15	Delcot3 44 x MZ561	82.5c	18.4e	32.9bc	6.3ab	39.3ab	2185.3bc	922.2abc	30.9b	33.3abc	4.8cd
16	UK91 x MZ561	91.2b	23.4ab	32.1bc	5.8bc	36.2c	2439.9ab	905.7b	30.2cd	37.3a	4.7c
17	Aubarn 56 x MZ561	90.9b	21.3bc	33.5b	5.9b	37.5b	2627.8a	1021.4ab	30.3bcd	33.1b	4.8b
18	NTA 93-21 x MZ561	91.3b	23.6ab	32.9bc	6.4a	40.3ab	2824.8a	1206.9a	32.2a	35.6a	4.9abc
19	Delcot 344	77.7c	18.2e	33.9abc	6.0abc	41.8a	1875.9c	888.7bc	31.4ab	34.5ab	4.9abc
20	UK91 x Delcot 344	89.9b	20.7c	35.3a	6.7a	37.8b	2535.4ab	1052.4ab	32.5a	35.1a	4.6d
21	Aubarn 56 x Delcot 344	87.6bc	20.2e	34.0abc	6.5a	39.7ab	2352.5abc	986.5abc	31.1abcd	33.0b	4.9abc
22	NTA 93-21 x Delcot 344	80.1c	19.5e	31.8bc	6.4a	43.7a	2490.5ab	1167.3a	32.3a	35.6a	5.0a
23	UK91	85.9c	20.4d	32.2bc	5.6d	34.9d	2313.3b	862.4c	30.9b	32.8b	4.4d
24	Aubarn 56 x UK91	85.1c	19.1e	33.0bc	5.9bc	36.7bc	2344.4b	897.0bc	30.2bcd	33.5abc	4.9abc
25	NTA 93-21 x UK91	87.8bc	21.8abc	33.4b	6.3ab	39.7ab	2742.6a	1234.6a	32.9a	34.5ab	4.8cd
26	Aubarn 56	89.2bc	19.9e	34.4ab	5.8bc	37.8b	2171.2bc	904.7b	29.9c	34.0abc	4.8cd
27	NTA 93-21 x Aubarn56	79.9c	22.4ab	32.9bc	6.4a	42.4a	2776.4a	1340.7a	31.6ab	34.9ab	5.1a
28	NTA 93-21	83.3c	20.3e	31.9bc	5.7d	42.6a	2331.9b	1025.7ab	32.6a	37.1a	4.7c
	Mean	87.9	20.3	32.6	5.9	37.9	2292.8	941.1	31.1	34.1	4.8
	LSD	6.1	3.3	2.0	0.3	1.0	282.4	128.5	0.3	0.1	0.0
	CV (%)	11.9	27.8	10.6	8.8	4.4	21.1	23.4	1.6	0.7	0.3

NB. Values in columns followed by similar letters are not significant different at 0.05 level of probability according to Duncan's multiple range test, LSD= Least squared deviation for genotype, CV= coefficient of variation, GOT= ginning outturn

ANOVA indicated a significant difference in bolls/plant among genotypes and locations while G x E interaction was not significantly different (Table 8.3). The difference between locations was significant as Ukiriguru gave a high number of bolls/plant and Kanziga was the lowest compared to other sites (Table 7.4). Parent MZ561 (27.2) with its combinations had many bolls/plant, for example NTA 93-21 x MZ561 (23.6), UK91 x MZ561 (23.4) and MZ561 x Okra leaf (22.8). Delcot 344 x Acala SJ$_2$ (17.9), Aubarn x Okra leaf (17.3) and Acala SJ$_2$ (17.1) had significant lower bolls/plant across sites (Table 8.4). The contribution of location to total sums of square (SS) was low (7%) compared to the contribution of genotype (13%) (Table 8.3).

Seed/boll differed significantly for location and genotypes while G x E interaction was not significant. The contribution of genotype to total SS was high (15%) compared to the contribution of location (2%) (Table 8.3). Significant differences were observed between UK91 x Delcot 344 (35.3), Aubarn 56 (34.4), Aubarn 56 x Delcot 344 (34.0) and Delcot 344 (33.9) for having many seed/boll compared to others. Three of the parents, MZ561 (30.4), Acala SJ$_2$ (29.8) and Okra leaf (28.9) had the lowest values (Table 8.4).

Boll weight was significantly different among locations and among genotypes but G x E interaction was not significant. The contribution of location to total SS was low (8%) compared to the contribution of genotype (44%) (Table 8.3). For example, boll weight at Kanziga was high as the soil had more organic matter (section 7.3.1). Crosses of Delcot 344 as male parent with UK91 (6.7 g), Aubarn 56 (6.5 g) and NTA 93-21 (6.4 g) as female parents exhibited large bolls. Okra leaf and its crosses as male parent resulted in small bolls (Table 8.4).

Significant GOT was observed among locations and genotypes while the G x E interaction was not significant. The contribution of location to total SS was low (7%) compared to the contribution of genotype (68%) (Table 8.3). Variation was observed at locations where Mwanhala had higher GOT and Kanziga lower values than other sites (section 7.3.1). Combinations NTA 93-21 x Delcot 344, NTA 93-21 x Aubarn 56, NTA 93-21 x Okra leaf, NTA 93-21 x Acala SJ$_2$ and NTA 93-21 x MZ561 showed increased GOT values ranging from 40.3-43.7%. NTA 93-21 (42.6%) and Delcot 344 (41.8%) were the parents with the

highest GOT values. MZ561, UK91 and Okra leaf (parents) and crosses UK91 x Okra leaf and UK91 x Acala SJ₂ GOT values ranged from 33.6-35.0% which is very low (Table 8.4).

Seedcotton and lint yield were significantly different among locations and genotypes but the G x E interaction was only significant for lint yield (Table 8.3). Female parent NTA 93-21 in crosses with MZ561 (2824.8 kg/ha), Aubarn 56 (2776.4 kg/ha), Acala SJ₂ (2759.3 kg/ha) and UK91 (2742.6 kg/ha) showed high seedcotton yield while Delcot 344 x Okra leaf (1877.7 kg/ha), Delcot 344 (1875.9 kg/ha) and Okra leaf (1172.3 kg/ha) were the lowest. Similar performances were observed for lint yield for the same combinations; NTA 93-21 x MZ561 (1206.9 kg/ha), NTA 93-21 x Aubarn 56 (1340.7 kg/ha), NTA 93-21 x Acala SJ₂ (1173.3 kg/ha) and NTA 93-21 x UK91 (1234.6 kg/ha). The poorest performers were UK91 x Okra leaf (751.4 kg/ha), Aubarn 56 x Okra leaf (750.7 kg/ha) and Okra leaf (455.3 kg/ha) (Table 8.4). The contribution of location to total SS was higher (35%) compared to the contribution of genotype (21%) to the total SS for seedcotton yield and for lint yield the contribution of location to total SS was higher (49%) compared to the contribution of genotype (19%).

8.3.3 Fibre quality

Fibre quality characteristics mean squares were all significant for locations, genotype and G x E interaction (Table 8.3). For fibre length, highly significant different values were observed for genotypes NTA 93-21 x UK91 (32.9 mm), NTA 93-21 x Acala SJ₂ (32.9 mm), NTA 93-21 (32.6 mm), UK91 x Acala SJ₂ (32.6 mm) and UK91 x Delcot 344 (32.5 mm). Lower fibre length values were observed for Aubarn 56 x Acala SJ₂ (29.6 mm) and MZ561 x Acala SJ₂ (29.6 g/tex) (Table 8.4). For fibre strength, UK91 x MZ561 (37.3 g/tex), NTA 93-21 (37.1 g/tex) and NTA 93-21 x Acala SJ₂ (36.5 g/tex) had high fibre strength values. The lowest values were observed for MZ 561 x Acala SJ₂ (32.3 g/tex) (Table 8.4). Fibre length and fibre strength values observed for all genotypes were above the medium value for fibre length (25.2-27.9 mm) and for fibre strength (22.0-25.0 g/tex). Micronaire values were high for NTA 93-21 x Aubarn 56 (5.1), Aubarn 56 x Okra leaf (5.1), NTA 93-21 x Okra leaf (5.1) and NTA 93-21 x Delcot 344 (5.0). Micronaire values above 5 are referred to as coarse fibre, which is not preferred for textiles. The medium value is 4.0-4.9, coarse fibre 5.0-5.9 and the very coarse fibre is ≥ 6.0. Apart from the highest four combinations, other genotypes had micronaire values in the acceptable range.

The contribution of genotype to the total variation was high (39%) for fibre length compared to location (15%) and fibre strength genotype contribution was 26% and location 20%. For fibre micronaire, the contribution to the total variation was high for location (41%) compared to genotype (33%) (Table 8.3).

8.3.4 Stability analysis

From the combined analysis lint yield (kg/ha) and fibre strength (g/tex) (sections 8.3.2, 8.3.3 and Table 8.3) had significant G x E interactions. As both characteristics are important to the world cotton industry and new spinning machines, further stability analyses of the 28 cotton genotypes were done using five statistical procedures.

Lin and Binns cultivar performance

According to Lin and Binns (1988), genotypes with low cultivar performance values (Pi) are stable. For lint yield, combinations NTA 93-21 x Aubarn 56, NTA 93-21 x UK91, NTA 93-21 x Acala SJ$_2$, NTA 93-21 x MZ561 and NTA 93-21 x Delcot 344 (all combinations with NTA 93-21), were stable across environments as they ranked in the first five positions. Parents Okra leaf, Acala SJ$_2$ and combinations Aubarn 56 x Okra leaf and UK91 x Okra leaf were unstable (Table 8.5). For both stability and the ANOVA the same five genotypes ranked as the best stable genotypes (Table 8.5), although there was a deviation for the ranking of the 3rd genotype.

Stability for fibre strength indicated that genotypes UK91 x MZ561, NTA 93-21, NTA 93-21 x MZ561, UK91 x Delcot 344 and NTA 93-21 x Acala SJ$_2$ with low Pi values, were stable across environments while MZ561 x Okra leaf and parents Okra leaf and UK91 were unstable as high Pi values were recorded according to Lin and Binns (1988) (Table 8.5). The overall mean revealed that UK91 x MZ561, NTA 93-21, NTA 93-21 x Acala SJ$_2$, NTA 93-21 x MZ561 and NTA 93-21 x Delcot 344 ranked first with low Pi values. NTA 93-21 x Acala SJ$_2$, ranked 3rd on overall mean (Table 8.5) but 5th on stability, UK91 x Delcot 344 which ranked 4th on Pi values changed to 8th on overall mean and NTA 93-21 x Delcot 344 moved from 8th for Pi values to 5th for overall mean, still there was not much difference for ranking with both procedures. The Lin and Binns (1988) stability concept was in harmony with the mean.

Table 8.5 Lin and Binns (1988) cultivar superiority measure, rank and means for lint yield and fibre strength for 28 cotton genotypes tested at four sites

No	Genotype	Lint yield (kg/ha)				Fibre strength (g/tex)			
		Pi	R	Mean	R	Pi	R	Mean	R
1	Okra leaf	549548.9	28	455.3	28	51.4	26	32.5	27
2	Acala SJ$_2$ x Okra leaf	282458.8	24	766.6	25	46.2	19	32.7	23
3	MZ561 x Okra leaf	188461.8	19	877.8	20	55.9	28	32.6	26
4	Delcot 344 x Okra leaf	265281.8	23	780.9	23	30.0	10	34.8	10
5	UK91 x Okra leaf	285086.6	25	751.4	26	46.3	20	32.6	26
6	Aubarn 56 x Okra leaf	302115.6	27	750.7	27	33.0	12	34.1	13
7	NTA 93-21 x Okra leaf	116925.9	8	978.1	12	37.5	16	34.0	14
8	Acala SJ$_2$	292785.1	26	768.8	24	47.3	22	33.1	19
9	MZ561 x Acala SJ$_2$	115262.9	7	998.1	9	50.8	25	32.3	28
10	Delcot 344 x Acala SJ$_2$	174623.9	15	926.4	13	26.5	6	35.1	7
11	UK91 x Acala SJ$_2$	210439.3	21	890.4	18	27.3	7	35.3	6
12	Aubarn 56 x AcalaS J$_2$	131159.5	11	980.1	11	45.4	18	32.7	24
13	NTA 93-21 x Acala SJ$_2$	57509.9	3	1173.0	4	25.8	5	36.5	3
14	MZ561	222194.6	22	836.3	22	46.5	21	32.8	22
15	Delcot 344 x MZ561	180024.8	16	922.2	14	49.2	25	33.3	17
16	UK91 x MZ561	174072.7	14	905.7	15	10.0	1	37.3	1
17	Aubarn 56 x MZ561	130186.8	10	1021.0	8	48.8	23	33.1	18
18	NTA 93-21 x MZ561	60054.7	4	1207.0	3	20.8	3	35.6	4
19	Delcot 344	173472.9	13	888.7	19	28.8	9	34.5	11
20	UK91 x Delcot 344	89147.3	6	1052.0	6	24.4	4	35.1	8
21	Aubarn 56 x Delcot 344	130041.9	9	986.5	10	43.6	17	33.0	20
22	NTA 93-21 x Delcot 344	64074.5	5	1167.0	5	27.6	8	35.6	5
23	UK91	196339.9	20	862.4	21	51.6	27	32.8	21
24	Aubarn 56 x UK91	183939.1	18	897.0	17	37.5	15	33.5	16
25	NTA 93-21 x UK91	27480.8	2	1235.0	2	28.9	10	34.5	12
26	Aubarn 56	183867.7	17	904.7	16	35.8	14	34.0	15
27	NTA 93-21 x Aubarn 56	13762.9	1	1341.0	1	35.3	13	34.9	9
28	NTA 93-21	141423.7	12	1026.0	8	15.9	2	37.1	2

Pi= cultivar superiority measure, R= rank

Wricke's ecovalence analysis

According to Wricke (1962), ecovalence (Wi) is defined as the contribution of each genotype to the G x E interaction. A genotype with a low Wi value has smaller fluctuations from the mean across different environments as environments contribute the least to G x E interaction. For lint yield, the Wi ranking procedure indicated that genotypes UK91 x MZ561, Aubarn 56 x Delcot 344, MZ561 x Okra leaf, Aubarn 56 x MZ561 and Aubarn 56 were the most stable while NTA 93-21 x Aubarn 56, NTA 93-21 x Okra leaf, NTA 93-21 x Delcot 344, NTA 93-21 and NTA

93-21 x UK91 were unstable across environments. Stability results for Wi were different from the overall means and Pi (Table 8.6). Combinations with MZ561 and Aubarn 56 as parents were stable and ones with NTA 93-21 as female parent were unstable.

Table 8.6 Wricke's ecovalence value, rank and mean for lint yield and fibre strength for 28 cotton genotypes tested at four sites

No	Genotype	Lint yield (kg/ha)				Fibre strength (g/tex)			
		Wi	R	Mean	R	Wi	R	Mean	R
1	Okra leaf	77017.1	22	455.3	28	18.9	25	32.5	27
2	Acala SJ$_2$ x Okra leaf	48310.8	16	766.6	25	2.4	6	32.7	23
3	MZ561 x Okra leaf	12441.2	3	877.8	20	17.9	23	32.6	26
4	Delcot 344 x Okra leaf	27471.0	11	780.9	23	18.7	24	34.8	10
5	UK91 x Okra leaf	25342.7	10	751.4	26	3.1	7	32.6	26
6	Aubarn 56 x Okra leaf	61500.2	19	750.7	27	1.0	4	34.1	13
7	NTA 93-21 x Okra leaf	209281.0	27	978.1	12	7.4	16	34.0	14
8	Acala SJ$_2$	76086.5	21	768.8	24	8.2	17	33.1	19
9	MZ561 x Acala SJ$_2$	34635.6	14	998.1	9	4.7	11	32.3	28
10	Delcot344xAcalaSJ$_2$	15059.5	9	926.4	13	0.9	3	35.1	7
11	UK91 x Acala SJ$_2$	65218.4	20	890.4	18	3.4	9	35.3	6
12	Aubarn 56 x Acala SJ$_2$	14538.8	7	980.1	11	1.1	5	32.7	24
13	NTA 93-21 x Acala SJ$_2$	27497.1	12	1173.0	4	12.9	21	36.5	3
14	MZ561	42901.1	15	836.3	22	5.1	12	32.8	22
15	Delcot 344 x MZ561	32426.5	13	922.2	14	20.2	26	33.3	17
16	UK91 x MZ561	9151.7	1	905.7	15	194.4	28	37.3	1
17	Aubarn 56 x MZ561	13471.8	4	1021.0	8	56.9	27	33.1	18
18	NTA 93-2 1x MZ561	51964.2	18	1207.0	3	5.2	13	35.6	4
19	Delcot 344	13809.6	6	888.7	19	12.7	20	34.5	11
20	UK91 x Delcot 344	80537.0	23	1052.0	6	5.9	14	35.1	8
21	Aubarn 56 x Delcot 344	12371.7	2	986.5	10	0.9	2	33.0	20
22	NTA 93-21 x Delcot 344	155309.0	26	1167.0	5	3.5	10	35.6	5
23	UK91	48895.5	17	862.4	21	12.6	19	32.8	21
24	Aubarn 56 x UK91	14732.0	8	897.0	17	0.9	1	33.5	16
25	NTA 93-21 x UK91	125985.0	25	1235.0	2	7.3	15	34.5	12
26	Aubarn 56	13610.7	5	904.7	16	13.1	22	34.0	15
27	NTA 93-21 x Aubarn 56	274523.0	28	1341.0	1	12.1	18	34.9	9
28	NTA 93-21	110709.0	24	1026.0	8	3.2	8	37.1	2

Wi= Wricke ecovalence value, R= rank

Results from Wi for fibre strength are summarised in Table 8.6. Aubarn 56 x UK91, Aubarn 56 x Delcot 344, Delcot 344 x Acala SJ$_2$, Aubarn 56 x Okra leaf and Aubarn 56 x Acala SJ$_2$ were stable, ranking in the first five positions. The lowest ranking entries for stability were UK91 x MZ561, Aubarn 56 x MZ561, Delcot 344 x MZ561, Okra leaf and Delcot 344 x Okra leaf. Stability results were different from overall means and Pi (Table 8.6). The most stable genotypes for fibre strength were combinations with Aubarn 56 as female parent and the least stable with MZ561 as male parent.

Shukla's stability variance

According to Shukla's stability variance (1972), stable genotypes are those having minimum stability variance (SV). Results for stability variance and overall means are summarised in Table 8.7 for both lint yield and fibre strength with their ranking orders. Lint yield stability analysis results indicated that UK91 x MZ561, Aubarn 56 x Delcot 344, MZ 561 x Okra leaf, Aubarn 56 x MZ561 and Aubarn 56 were stable. NTA 93-21 x Aubarn 56, NTA 93-21 x Okra leaf, NTA 93-21 x Delcot 344, NTA 93-21 x UK91 and NTA 93-21 were unstable genotypes. Shukla's and Wricke's procedures gave similar ranking orders, which was different from the overall mean and Pi rankings.

Fibre strength stability values indicated that Aubarn 56 x UK91, Aubarn 56 x Delcot 344, Delcot 344 x Acala SJ$_2$, Aubarn 56 x Okra leaf and Aubarn 56 x Acala SJ$_2$ were stable, while the lowest ranking for stability analysis according to Shukla's were UK91 x MZ561, Aubarn 56 x MZ561, Delcot 344 x MZ561, Okra leaf and Delcot 344 x Okra leaf. Shukla's stability variance for fibre strength showed that combinations with Aubarn 56 and Delcot 344 were stable while combinations with MZ561 and Okra leaf as parents were unstable. Results of Shukla's agreed with the Wricke's procedure and stability results were different from overall means and Pi (Table 8.7).

Table 8.7 Shukla's stability variance, ranks and mean for lint yield and fibre strength for 28 cotton genotypes tested at four sites

No	Genotype	Lint yield (kg/ha)				Fibre strength (g/tex)			
		SV	R	Mean	R	SV	R	Mean	R
1	Okra leaf	107369.7	22	455.3	28	26.2	25	32.5	27
2	Acala SJ$_2$ x Okra leaf	66150.4	16	766.6	25	2.6	6	32.7	23
3	MZ561 x Okra leaf	14645.3	3	877.8	20	24.9	23	32.6	26
4	Delcot 344 x Okra leaf	36226.6	11	780.9	23	25.9	24	34.8	10
5	UK91 x Okra leaf	33170.5	10	751.4	26	3.6	7	32.6	26
6	Aubarn 56 x Okra leaf	85088.9	19	750.7	27	0.6	4	34.1	13
7	NTA 93-21 x Okral eaf	297287.7	27	978.1	12	9.8	16	34.0	14
8	Acala SJ$_2$	106033.5	21	768.8	24	10.9	17	33.1	19
9	MZ561 x Acala SJ$_2$	46514.2	14	998.1	9	5.8	11	32.3	28
10	Delcot 344 x Acala SJ$_2$	18404.9	9	926.4	13	0.5	3	35.1	7
11	UK91 x Acala SJ$_2$	90427.9	20	890.4	18	3.9	9	35.3	6
12	Aubarn 56 x Acala SJ$_2$	17657.3	7	980.1	11	0.7	5	32.7	24
13	NTA 93-21 x Acala SJ$_2$	36264.1	12	1173.0	4	17.8	21	36.5	3
14	MZ561	58382.6	15	836.3	22	6.4	12	32.8	22
15	Delcot 344 x MZ561	43342.2	13	922.2	14	28.2	26	33.3	17
16	UK91 x MZ561	9921.8	1	905.7	15	278.2	28	37.3	1
17	Aubarn 56 x MZ561	16125.1	4	1021.0	8	80.8	27	33.1	18
18	NTA 93-21 x MZ561	71396.2	18	1207.0	3	6.6	13	35.6	4
19	Delcot 344	16610.1	6	888.7	19	17.4	20	34.5	11
20	UK91 x Delcot 344	112423.9	23	1052.0	6	7.8	14	35.1	8
21	Aubarn 56 x Delcot 344	14545.5	2	986.5	10	0.4	2	33.0	20
22	NTA 93-21 x Delcot 344	219789.4	26	1167.0	5	4.1	10	35.6	5
23	UK91	66989.9	17	862.4	21	17.2	19	32.8	21
24	Aubarn 56 x UK91	17934.6	8	897.0	17	0.4	1	33.5	16
25	NTA 93-21 x UK91	177681.8	25	1235.0	2	9.6	15	34.5	12
26	Aubarn 56	16324.7	5	904.7	16	17.9	22	34.0	15
27	NTA 93-21 x Aubarn 56	390967.7	28	1341.0	1	16.5	18	34.9	9
28	NTA 93-21	155747.4	24	1026.0	8	15.9	8	37.1	2

SV= Shukla's stability variance, R= rank

Eberhart and Russell joint regression procedure

The Eberhart and Russell (1966) procedure involves the use of joint linear regression where the yield of each genotype is regressed on the environmental mean yield. The appropriate analyses of variance for the regression model for both lint yield and fibre strength are indicated in Table 8.8. The genotype's performance across environments is generally expressed in terms of three parameters, the mean yield, the regression coefficient *(bi)* and the deviation *($S^2 d_i$)* from the regression. The genotype should have a high mean yield (kg/ha), unit regression coefficient *(bi = 1.0)* and as small as possible deviations from the regression *($S^2 d_i$ = 0)*. It is, however,

specifically the deviation from the regression (S^2d_i) which is used as a measure of genotype stability across environments. Results (Table 8.8) indicated the significant differences for genotypes and G x E interaction [Var x Envi (linear)] for lint yield while for fibre strength, genotypes did not differ significantly.

Table 8.8 Analysis of variance for stability analysis according to the joint regression model (Eberhart and Russell, 1966)

Source of variation	Df	Lint yield (kg/ha)		Fibre strength (g/tex)	
		SS	MS	SS	MS
Total	447	14257874.6		846.8	
Varieties (genotypes)	27	3525473.6	130573.1 **	224.1	8.3
Env + in Var.x Env.	84	10732400.9	127766.7	622.7	7.4
Env. in linear	1	9037603.1		168.3	
Var. x Env. (linear)	27	1120956.7	41516.9 **	130.9	4.9
Pooled deviation	56	573841.1	10247.2	323.5	5.8
Residual	336	4182150.0	12446.9	4.6	0.1
Grand mean		941.1		34.1	
CV (%)		0.2		0.7	

Env= environment, Var= variety, CV= coefficient of variation, Df= degree of freedom, SS= sums of square, MS= mean squares, ** $p \leq 0.01$

For lint yield, according to Eberhart and Russell (1966), genotypes NTA 93-21 x Acala SJ$_2$ and Aubarn 56 x Acala SJ$_2$ were the most stable with low S^2d_i approaching zero and bi approaching one (Table 8.9). Stable genotypes for lint yield were combinations with Acala SJ$_2$ as male parent. According to the of Finlay and Wilkinson (1963), genotypes NTA 93-21 x Delcot 344 and NTA 93-21 x MZ561 that had high mean values and regression coefficients *(bi)* approaching 1 indicated general adaptability, however according to Eberhart and Russell's (1966) method were not stable because S^2d_i values were high. Combinations NTA 93-21 x Aubarn 56 (1.9), NTA 93-21 x Okra leaf (1.8) and NTA 93-21 x UK91 (1.6) had the highest regression values and were unstable according to Eberhart and Russell and according to Finlay and Wilkinson, (1963) based on high mean yield, were increasingly sensitive to environmental changes. Combinations were from the same mother parent (NTA 93-21) (Table 8.9).

Table 8.9 Joint regression stability parameters and rank for lint yield mean and fibre strength mean of 28 cotton genotypes

No	Genotype	Lint yield (kg/ha)					Fibre strength (g/tex)				
		Mean	R	(bi)	(S^2d_i)	R	Mean	R	(bi)	(S^2d_i)	R
1	Okra leaf	455.3	28	0.6	-4804.1	4	32.5	27	-0.2	5.2	22
2	Acala SJ$_2$ x Okra leaf	766.6	25	0.6	-11659.6	22	32.7	23	1.6	0.3	5
3	MZ561 x Okra leaf	877.8	20	1.1	-7626.8	10	32.6	26	0.7	8.7	25
4	Delcot 344 x Okra leaf	780.9	23	0.7	-9097.4	13	34.8	10	1.1	9.3	26
5	UK91 x Okra leaf	751.4	26	0.7	-10466.7	17	32.6	26	1.7	-0.0	1
6	Aubarn 56 x Okra leaf	750.7	27	0.6	-11164.8	20	34.1	13	1.3	0.2	4
7	NTA 93-21 x Okra leaf	978.1	12	1.8	-12208.5	25	34.0	14	1.1	3.7	20
8	Acala SJ$_2$	768.8	24	0.6	5054.5	5	33.1	19	1.5	3.4	19
9	MZ561 x Acala SJ$_2$	998.1	9	1.3	-8994.7	12	32.3	28	1.8	0.3	6
10	Delcot 344 x Acala SJ$_2$	926.4	13	0.8	-9895.7	15	35.1	7	0.6	0.0	2
11	UK91 x Acala SJ$_2$	890.4	18	0.7	5967.1	8	35.3	6	0.9	1.7	12
12	Aubarn 56 x Acala SJ$_2$	980.1	11	1.0	-5406.3	7	32.7	24	1.2	0.4	9
13	NTA 93-21 x Acala SJ$_2$	1173.0	4	1.1	-268.10	1	36.5	3	2.2	1.8	16
14	MZ561	836.3	22	0.8	1869.3	2	32.8	22	0.3	1.1	11
15	Delcot 344 x MZ561	922.2	14	0.8	-3475.9	3	33.3	17	2.0	6.9	24
16	UK91 x MZ561	905.7	15	0.9	-9167.3	14	37.3	1	-1.4	80.3	28
17	Aubarn 56 x MZ561	1021.0	8	0.8	-11370.9	21	33.1	18	-1.1	15.4	27
18	NTA 93-21 x MZ561	1207.0	3	1.1	12912.9	26	35.6	4	0.5	1.7	14
19	Delcot 344	888.7	19	1.2	-11096.7	19	34.5	11	1.8	4.6	21
20	UK91 x Delcot 344	1052.0	6	1.5	-10004.9	16	35.1	8	1.4	2.5	17
21	Aubarn 56 x Delcot 344	986.5	10	1.1	-8717.0	11	33.0	20	0.9	0.4	8
22	NTA 93-21 x Delcot 344	1167.0	5	1.0	64918.2	28	35.6	5	0.9	1.7	13
23	UK91	862.4	21	1.0	11891.6	24	32.8	21	1.0	6.3	23
24	Aubarn 56 x UK91	897.0	17	1.0	-5080.9	6	33.5	16	1.3	0.1	3
25	NTA 93-21 x UK91	1235.0	2	1.6	-11817.3	23	34.5	12	-0.0	0.3	7
26	Aubarn 56	904.7	16	0.9	-7179.1	9	34.0	15	2.3	1.8	15
27	NTA 93-21 x Aubarn 56	1341.0	1	1.9	-10815.0	18	34.9	9	2.1	2.6	18
28	NTA 93-21	1026.0	8	0.6	16112.4	27	37.1	2	0.5	0.8	10
	Standard error of beta			0.2					0.9		

bi= regression coefficient, S^2d_i= deviation from regression, R= rank

In stability analysis for fibre strength, the S^2d_i measure was approaching zero for most of the genotypes but the regression coefficient *(bi)* was high for most of the genotypes. However, Aubarn 56 x Okra leaf had a fibre strength mean of 34.1 g/tex, 1.3 bi approaching 1 and 0.2 S^2d_i approaching zero, theoretically indicating stability. UK91 x MZ561, Aubarn 56 x MZ561, Delcot 344 x Okra leaf and MZ561 x Okra leaf with high S^2d_i values (Table 8.9), were considered unstable. Genotypes UK91 x Okra leaf and Delcot 344 x Acala SJ$_2$ had low S^2d_i values, however the bi value for UK91 x Okra leaf was high and for Delcot 344 x Acala SJ$_2$ low.

This indicated them to be unstable genotypes according to Eberhart and Russell (1966) (Table 8.9).

Additive main effects and multiplicative interaction (AMMI) model

Results for AMMI analysis of variance for cotton genotypes on lint yield and fibre strength according to the best AMMI model fit are shown on Table 8.10. The ANOVA for lint yield indicated that G x E interaction was partitioned into three interaction principle component axis (IPCA). The IPCAs are ordered according to importance. The IPCA 1 axis was significant for lint yield and explained 66.3% of the total G x E interaction sums of squares percentage at 35.8% of the interaction degrees of freedom. The second IPCA (IPCA 2) explained 27.1% of the total G x E interaction sums of squares percentage however it was not significant. The IPCA 2 was included in the discussion especially for the IPCA scores biplot to determine if it had any relationship and impact. The remaining 6.4% being the residue or noise which are not interpretable were discarded as described by Purchase (1997).

Table 8.10   Analysis of variance of interaction principal components analysis in AMMI for lint yield and fibre strength of 28 cotton genotypes

Source	Df	Lint yield		Fibre strength	
		SS	MS	SS	MS
Total	447	73760099		3405.53	
Environments	3	36150413	12050138.0 **	673.08	224.4 **
Reps within Env.	12	994798.1	82899.8	0.51	0.1
Genotype	27	14101894	522292.4 **	896.39	33.2
Genotype x Env.	81	6779191	83693.7 **	1817.73	22.4 **
IPCA 1	29	4493143	154936.0 **	1041.48	35.9 **
IPCA 2	27	1831029	67815.9	494.96	18.3 **
IPCA 3	25	455019.2	18200.8	281.30	11.3 **
Residual	324	15733802	48561.1	17.82	0.1
Grand mean		941.085		34.13	
R-squared		0.7867		0.99	
CV (%)		23.42		0.69	

IPCA= Interaction principle component analysis, Env.= environment, CV= coefficient of variation, Df= degrees of freedom, SS= sum of squares, MS= mean squares, ** $p \leq 0.01$, R-squared= repeatability squared, Env= environment

IPCA scores of genotypes in the AMMI analysis was reported by Gauch and Zobel (1988) as indication of the stability of a genotype over environment. High IPCA scores, either negative or positive, indicate that a genotype is specifically adapted to a certain environment. IPCA scores around zero indicate that the genotype is stable over all environments sampled. In accordance with this concept from the AMMI biplot for lint yield (Figure 8.1a), genotypes Aubarn 56 x Acala SJ$_2$, Aubarn 56 x UK91, Aubarn 56 x Delcot 344, UK91, MZ561x Okra leaf, Aubarn 56, Delcot 344, Aubarn 56 x MZ561, UK91 x Acala SJ$_2$ and Delcot 344 x Acala SJ$_2$ were the most stable over environments when IPCA 1 scores and means were taken into account. In contrast NTA 93-21 x Aubarn 56, NTA 93-21 x UK91, UK91 x Delcot 344 and NTA 93-21 were unstable on the high yielding environment. Aubarn 56 x Okra leaf, Acala SJ$_2$ x Okra leaf, Acala SJ$_2$ and Okra leaf were unstable on the low yielding environment. Means versus IPCA 1 scores indicated that NTA 93-21 x Acala SJ$_2$, NTA 93-21 x MZ561 and NTA 93-21 x Delcot 344 were stable and adapted to the high yielding environment.

On the other hand, Delcot 344 x Okra leaf and UK91 x Okra leaf were low yielding and stable and adapted to low yield environments. Okra leaf as male parent and most of its crosses resulted in low lint yielding. Kanziga was the only stable low yielding environment. Ukiriguru was an unstable and high yielding environment, while Bwanga and Mwanhala were unstable and low yielding environments (Figure 8.1a). However, when IPCA 2 was considered, the stability biplot changed. For example, for IPCA 1 and IPCA 2 scores, Aubarn 56 was the most stable genotype followed by NTA 93-21 x Acala SJ$_2$, MZ561 x Okra leaf, Aubarn 56 x Delcot 344, Delcot 344, Delcot 344 x Okra leaf, UK91 x Okra leaf and UK91 x MZ561. NTA 93-21 and NTA 93-21 x Aubarn 56 were adapted to specific environments (Figure 8.1b).

Based on these results the other option was to calculate AMMI stability values (ASV). This was reported to produce a balanced measurement between the two IPCA scores (Purchase, 1997). According to ASV ranking (Table 8.11), Aubarn 56 was the most stable genotype followed by NTA 93-21 x Acala SJ$_2$, MZ561 x Okra leaf, UK91 x MZ561 and Aubarn 56 x UK91. ASV values agreed with the IPCA scores biplot (Figure 8.1b). On the other hand, NTA 93-21 x Aubarn 56, NTA 93-21 x Okra leaf, NTA 93-21 x UK91, NTA 93-21 x Delcot 344 and UK91 x Delcot 344 were unstable based on ASV ranking.
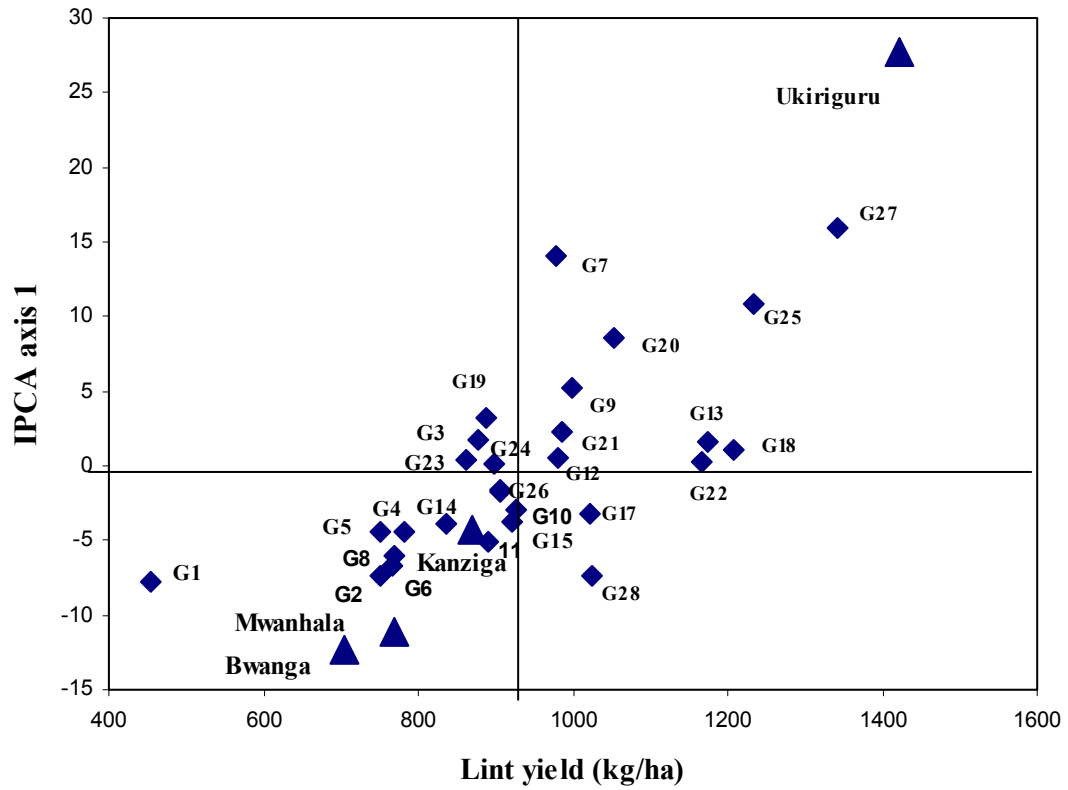
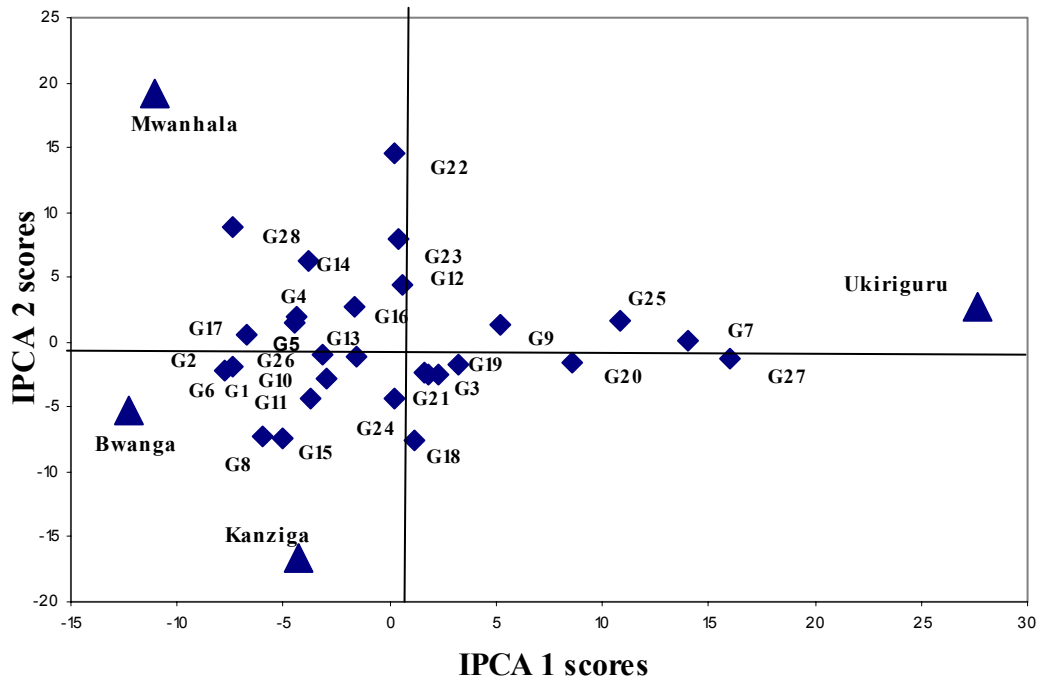**Figure 8.1a AMMI biplot for lint yield mean (kg/ha) and IPCA 1 scores**

Figure 8.1b AMMI biplot for lint yield IPCA 1 and IPCA 2 scores

G1=Okra leaf, G2=Acala SJ$_2$ x Okraleaf, G3= MZ561 x Okra leaf, G4= Delcot 344 x Okra leaf, G5= UK91 x Okra leaf, G6= Aubarn 56 x Okra leaf, G7= NTA 93-21 x Okra leaf, G8= Acala SJ$_2$, G9= MZ561 x Acala SJ$_2$, G10= Delcot 344 x Acala SJ$_2$, G11= UK91 x Acala SJ$_2$, G12= Aubarn 56 x Acala SJ$_2$, G13= NTA 93-21 x Acala SJ$_2$, G14= MZ561, G15= Delcot 344 x MZ561, G16= UK91 x MZ561, G17= Aubarn 56 x MZ561, G18= NTA 93-21 x MZ561, G19= Delcot 344, G20= UK91 x Delcot 344, G21= Aubarn 56 x Delcot 344, G22= NTA 93-21 x Delcot 344, G23= UK91, G24= Aubarn56 x UK91, G25= NTA 93-21 x UK91, G26= Aubarn 56, G27= NTA 93-21 x Aubarn 56, G28= NTA 93-21.

Table 8.11   Mean lint yield, mean fibre strength, IPCA 1 and IPCA 2 scores, AMMI stability values and rank of 28 cotton genotypes

		Lint yield (kg/ha)					Fibre strength (g/tex)					
No Genotype		IPCA1 scores	IPCA2 scores	Mean	R	ASV	R	IPCA1 scores	IPCA2 scores	Mean	R	ASV	R
1	Okra leaf	-7.7	-2.2	455.3	28	12.3	22	-0.1	-0.2	32.5	27	0.2	2
2	Acala SJ$_2$ x Okra leaf	-6.7	0.6	766.6	25	10.5	18	0.2	0.4	32.7	23	0.5	10
3	MZ561 x Okra leaf	1.8	-2.5	877.8	20	3.7	3	0.7	-0.0	32.6	26	0.9	20
4	Delcot 344 x Okra leaf	-4.4	1.9	780.9	23	7.1	11	-0.4	0.9	34.8	10	1.1	24
5	UK91 x Okra leaf	-4.5	1.5	751.4	26	7.2	12	0.1	0.4	32.6	26	0.5	8
6	Aubarn 56 x Okra leaf	-7.4	-1.8	750.7	27	11.7	20	-0.1	0.3	34.1	13	0.3	5
7	NTA 93-21 x Okra leaf	14.1	0.9	978.1	12	22.0	27	0.4	-0.5	34.0	14	0.8	16
8	Acala J$_2$	-6.0	-7.3	768.8	24	11.9	21	0.6	-0.2	33.1	19	0.9	19
9	MZ561 x Acala SJ$_2$	5.2	1.4	998.1	9	8.2	16	0.4	0.4	32.3	28	0.7	15
10	Delcot 344 x Acala SJ$_2$	-3.0	-2.8	926.4	13	5.5	10	-0.2	-0.2	35.1	7	0.3	6
11	UK91 x Acala SJ$_2$	-5.1	-7.4	890.4	18	10.9	19	0.2	-0.4	35.3	6	0.5	9
12	Aubarn 56 x Acala SJ$_2$	0.6	4.4	980.1	11	4.5	7	0.1	0.3	32.7	24	0.3	4
13	NTA 93-21 x Acala SJ$_2$	1.6	-2.4	1173.0	4	3.4	2	0.6	0.3	36.5	3	1.0	21
14	MZ561	-3.9	6.3	836.3	22	8.7	17	0.2	-0.6	32.8	22	0.6	14
15	Delcot 344 x MZ561	-3.7	-4.3	922.2	14	7.2	13	0.6	1.0	33.3	17	1.4	26
16	UK91 x MZ561	-1.7	2.7	905.7	15	3.7	4	-3.5	-0.1	37.3	1	5.0	28
17	Aubarn 56 x MZ561	-3.2	-1.0	1021.0	8	5.1	8	0.2	-2.3	33.1	18	2.3	27
18	NTA 93-21 x MZ561	1.1	-7.6	1207.0	3	7.8	14	-0.6	-0.1	35.6	4	0.8	17
19	Delcot 344	3.3	-1.8	888.70	19	5.4	9	-0.4	0.9	34.5	11	1.1	22
20	UK91 x Delcot 344	8.6	-1.6	1052.0	6	13.5	24	-0.2	0.1	35.1	8	0.4	7
21	Aubarn 56 x Delcot 344	2.3	-2.4	986.50	10	4.3	6	0.2	-0.1	33.0	20	0.3	3
22	NTA 93-21 x Delcot 344	0.3	14.6	1167.0	5	14.6	25	0.3	-0.4	35.6	5	0.6	13
23	UK91	0.4	7.9	862.4	21	7.9	15	0.8	-0.5	32.8	21	1.2	25
24	Aubarn 56 x UK91	0.2	-4.3	897.0	17	4.3	5	-0.8	0.6	33.5	16	0.2	1
25	NTA 93-21 x UK91	10.8	1.6	1235.0	2	17.0	26	-0.4	-0.6	34.5	12	0.8	18
26	Aubarn 56	-1.6	-1.1	904.7	16	2.6	1	0.3	0.4	34.0	15	0.5	11
27	NTA 93-21x Aubarn 56	15.9	-1.3	1341.0	1	25.1	28	0.5	0.9	34.9	9	1.1	23
28	NTA 93-21	-7.4	8.9	1026.0	8	12.7	23	-0.4	-0.0	37.1	2	0.6	12

IPCA= Interaction principle components axis, ASV= AMMI stability values, R= rank

AMMI analysis for fibre strength results revealed three significant G x E IPCAs (Table 8.11). IPCA 1 explained 57.3% of the total G x E sums of squares explanation at 35.8% of G x E degrees of freedom. IPCA 2 explained 27.2%. The contribution of IPCA 3 (15.5%) was relatively small and was discarded as residues or noise because it could be difficult to draw the right conclusions from this principle factor. With three significant IPCAs, two biplots were plotted as IPCA 1 scores against means and IPCA 2 scores against IPCA 1 scores (Figures 8.2a and 8.2b). For fibre strength all genotypes except UK91 x MZ561were stable though with different directions.
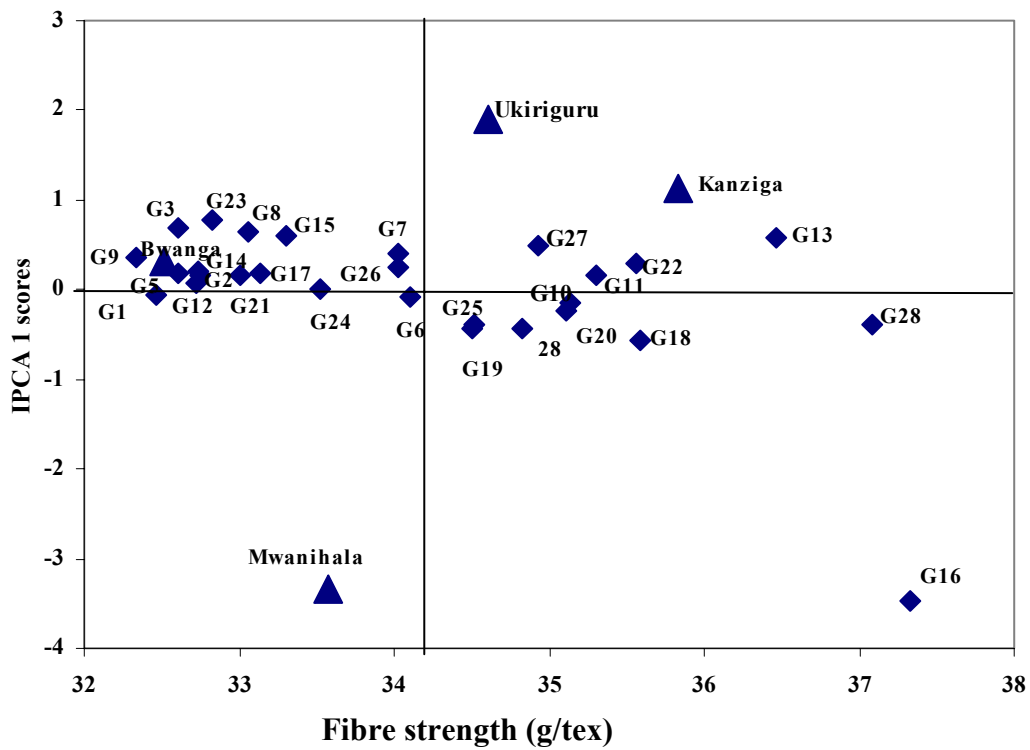
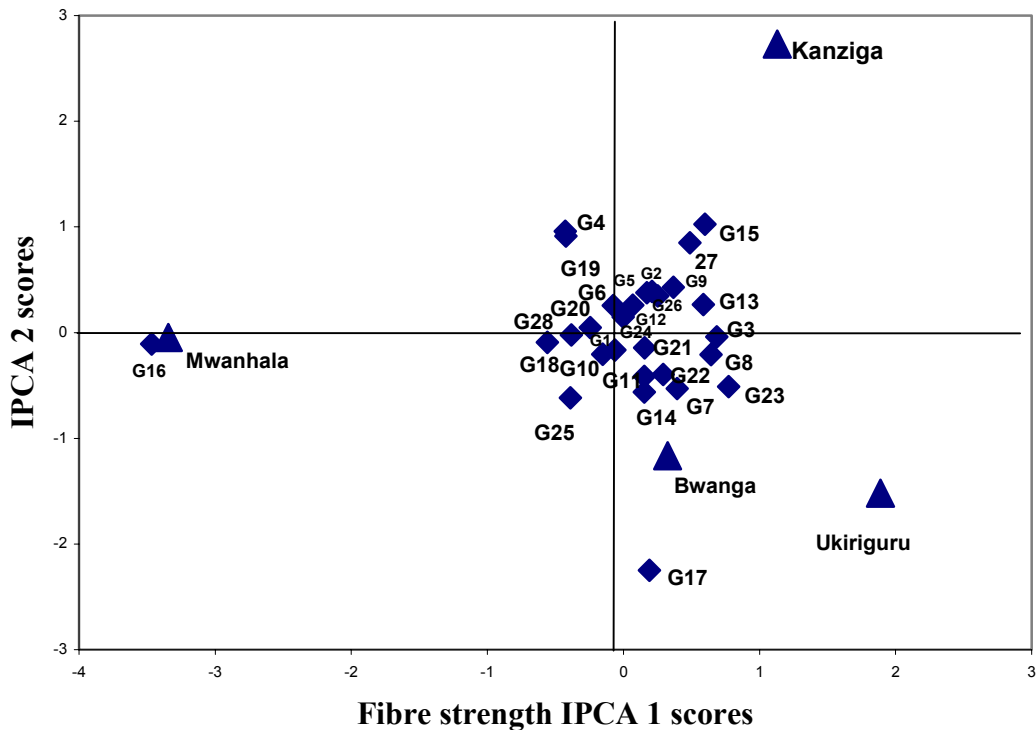**Figure 8.2a AMMI biplot for fibre strength means (g/tex) and IPCA 1 scores**

Figure 8.2b AMMI biplot for fibre strength IPCA 1 and IPCA 2 scores

G1=Okra leaf, G2=Acala SJ$_2$ x Okraleaf, G3= MZ561 x Okra leaf, G4= Delcot 344 x Okra leaf, G5= UK91 x Okra leaf, G6= Aubarn 56 x Okra leaf, G7= NTA 93-21 x Okra leaf, G8= Acala SJ$_2$, G9= MZ561 x Acala SJ$_2$, G10= Delcot 344 x Acala SJ$_2$, G11= UK91 x Acala SJ$_2$, G12= Aubarn 56 x Acala SJ$_2$, G13= NTA 93-21 x Acala SJ$_2$, G14= MZ561, G15= Delcot 344 x MZ561, G16= UK91 x MZ561, G17= Aubarn 56 x MZ561, G18= NTA 93-21 x MZ561, G19= Delcot 344, G20= UK91 x Delcot 344, G21= Aubarn 56 x Delcot 344, G22= NTA 93-21 x Delcot 344, G23= UK91, G24= Aubarn 56 x UK91, G25= NTA 93-21 x UK91, G26= Aubarn 56, G27= NTA 93-21 x Aubarn 56, G28= NTA 93-21.

Results indicated that most of the genotypes grouped in one quadrant, the other two quadrants had few while the fourth had none. Aubarn 56 x Okra leaf, followed by Aubarn 56, NTA 93-21 x Okra leaf, Delcot 344, NTA 93-21 x UK91 and Aubarn 56 x UK91 were the most stable. However, NTA 93-21, NTA 93-21 x Acala SJ$_2$ NTA 93-21 x Delcot 344, NTA 93-21 x MZ561, Delcot 344 x Acala SJ$_2$ UK91 x Delcot 344, NTA 93-21 x Aubarn 56 and MZ561 x Acala SJ$_2$ were stable on high yielding environments (Figure 8.2a). MZ561 x Okra leaf, Acala SJ$_2$ UK91, Aubarn 56 x MZ561, Aubarn 56 x Delcot 344, MZ561, Acala SJ$_2$ x Okra leaf, Aubarn 56 x Acala SJ$_2$ Okra leaf and MZ561 x Acala SJ$_2$ were stable for low yielding environments. UK91 x MZ561 was the only unstable genotype in relation to fibre strength. IPCA 1 scores and fibre strength means, indicated that Bwanga and Kanziga were stable environments (Bwanga was stable as low yielding environment and Kanziga as high yielding environment). Ukiriguru was unstable as high yielding environment and Mwanhala as low yielding environment.

The biplot for IPCA 1 and IPCA 2 scores (Figure 8.2b) indicated a different picture where most of the genotypes were grouped at the centre. The most stable genotype was Aubarn 56 x UK91 followed by Okra leaf, Delcot 344 x Acala SJ$_2$, UK91 x Delcot 344, Aubarn 56 x Delcot 344, Aubarn 56 x Acala SJ$_2$ and Aubarn 56 x Okra leaf (Figure 8.2b). This biplot confirmed the instability of UK91 x MZ561 as observed in Figure 8.2a. Other genotypes which were unstable apart from UK91 x MZ561 were Aubarn 56 x MZ561, Delcot 344 x Okra leaf and Delcot 344 x MZ561. All environments were unstable.

Based on this information and variation for stable genotypes, the *ASV* measure was used as a measurement for the AMMI procedure as proposed by Purchase (1997). The *ASV* ranking indicated that Aubarn 56 x UK91 followed by Okra leaf, Aubarn 56 x Delcot 344, Aubarn 56 x Acala SJ$_2$ and Aubarn 56 x Okra leaf were stable. Based on *ASV*, combinations of Aubarn 56 as female parent were more stable. In contrast to this, genotypes UK91 x MZ561, Aubarn 56 x MZ561, Delcot 344 x MZ561, UK91 and Delcot 344 x Okra leaf were the most unstable (Table 8.12). Results from IPCA scores biplots were in harmony with *ASV* rankings.

8.3.5 Comparison of stability procedures

Comparison of stability parameters for cotton genotypes was done for five stability measures applied using rank levels. Results are given in Table 8.12 (lint yield) and Table 8.13 (fibre strength). The stability procedures were: cultivar superiority measure *(Pi)* (Lin and Binns, 1988), ecovalence variance *(Wi)* (Wricke, 1962), stablity variance *(SV)* (Shukla, 1972), deviation from regression *(S^2d_i)* (Eberhart and Russell, 1966) and AMMI's stability value *(ASV)* (Gauch and Zobel, 1996; Purchase, 1997). The ANOVA overall mean for lint yield (kg/ha) and fibre strength (g/tex) were included in the comparison to support these stability parameters and for comparison purposes.

For lint yield (Table 8.12) the stability ranks of Shukla's stability variance *(SV),* Wricke's ecovalence *(Wi)* and AMMI stability value *(ASV),* were similar. MZ561 x Okra leaf, UK91 x MZ561, Aubarn 56, Aubarn 56 x Delcot 344, Aubarn 56 x Acala SJ$_2$, Delcot 344 x Acala SJ$_2$, Aubarn 56 x MZ561, Delcot 344 and Aubarn 56 x UK91 ranked within the first 10 stable genotypes for all three procedures. Eberhart and Russell deviation from the regression grouped these genotypes as intermediate stable but the difference was not too high. Lin and Binns cultivar superiority measure *(Pi)* deviated from the results obtained by the first four mentioned methods and grouped stable genotypes as unstable except for Aubarn 56 x Delcot 344 and Aubarn 56 x MZ561 which were grouped as stable. Lin and Binns ranking was similar to the ANOVA mean ranking. When mean ranking was included Aubarn 56 x MZ561 was accepted as a stable and high lint yield genotype though the ranks were 4th on Wricke and Shuklas, 8th on *ASV* and means and 10th on Lin and Binns procedures. Okra leaf was unstable almost for all procedures and ranked last (28th) for mean and *Pi,* 22nd for Wricke, Shukla and *ASV* while Eberhart and Russell ranked Okra leaf as 4th. When all five stability measures and mean ranks were averaged, the overall ranking identified genotypes Aubarn 56 x Delcot 344, NTA 93-21 x Acala SJ$_2$, UK91 x MZ561, Aubarn 56, MZ561 x Okra leaf, Aubarn 56 x Acala SJ$_2$, Aubarn 56 x UK91, Aubarn 56 x MZ561, Delcot 344 and Delcot 344 x MZ561 as stable genotypes. Unstable genotypes were Aubarn 56 x Okra leaf, NTA 93-21 x Okra leaf, NTA 93-21 and NTA 93-21 x Delcot 344.

Table 8.12 Ranking of different stability procedures for 28 cotton genotypes for lint yield

No	Genotype	Lin and Binns Pi	R	Wricke Wi	R	Shukla SV	R	Eberhart and Russell (S^2d_i)	R	AMMI ASV	R	ANOVA Mean	R	Overall R
1	Okra leaf	549548.9	28	77017.1	22	107369.7	22	-4804.1	4	12.3	22	455.3	28	22
2	Acala SJ$_2$ x Okra leaf	282458.8	24	48310.8	16	66150.4	16	-11659.6	22	10.5	18	766.6	25	21
3	MZ561 x Okra leaf	188461.8	19	12441.2	3	14645.3	3	-7626.8	10	3.7	3	877.8	20	5
4	Delcot 344 x Okra leaf	265281.8	23	27471.0	11	36226.6	11	-9097.4	13	7.1	11	780.9	23	13
5	UK91 x Okral eaf	285086.6	25	25342.7	10	33170.5	10	-10466.7	17	7.2	12	751.4	26	15
6	Aubarn 56 x Okra leaf	302115.7	27	61500.2	19	85088.9	19	-11164.8	20	11.7	20	750.7	27	25
7	NTA93-21 x Okra leaf	116925.9	8	209281.0	27	297287.7	27	-12208.5	25	22.0	27	978.1	12	28
8	Acala SJ$_2$	292785.1	26	76086.5	21	106033.5	21	5054.5	5	11.9	21	768.8	24	20
9	MZ561 x AcalaSJ$_2$	115262.9	7	34635.6	14	46514.2	14	-8994.7	12	8.2	16	998.1	9	12
10	Delcot 344 x Acala SJ$_2$	174623.9	15	15059.5	9	18404.9	9	-9895.7	15	5.5	10	926.4	13	11
11	UK91 x Acala SJ$_2$	210439.3	21	65218.4	20	90427.9	20	5967.1	8	10.9	19	890.4	18	17
12	Aubarn 56 x Acala SJ$_2$	131159.5	11	14538.8	7	17657.3	7	-5406.3	7	4.5	7	980.1	11	6
13	NTA 93-21 x Acala SJ$_2$	57509.9	3	27497.1	12	36264.1	12	-268.1	1	3.4	2	1173.0	4	2
14	MZ561	222194.6	22	42901.1	15	58382.7	15	1869.3	2	8.7	17	836.3	22	14
15	Delcot 344 x MZ561	180024.8	16	32426.5	13	43342.2	13	-3475.9	3	7.2	13	922.2	14	10
16	UK91 x MZ561	174072.7	14	9151.7	1	9921.9	1	-9167.3	14	3.7	4	905.7	15	3
17	Aubarn 56 x MZ561	130186.8	10	13471.8	4	16125.1	4	-11370.9	21	5.1	8	1021.0	8	8
18	NTA 93-21 x MZ561	60054.7	4	51964.2	18	71396.2	18	12912.9	26	7.8	14	1207.0	3	16
19	Delcot 344	173472.9	13	13809.6	6	16610.1	6	-11096.7	19	5.4	9	888.7	19	9
20	UK91 x Delcot 344	89147.3	6	80537.0	23	112423.9	23	-10004.9	16	13.5	24	1052.0	6	18
21	Aubarn 56 x Delcot 344	130041.9	9	12371.7	2	14545.5	2	-8717.0	11	4.3	6	986.5	10	1
22	NTA 93-21 x Delcot 344	64074.5	5	155309.0	26	219789.4	26	64918.2	28	14.6	25	1167.0	5	27
23	UK91	196339.9	20	48895.5	17	66989.9	17	11891.6	24	7.9	15	862.4	21	19
24	Aubarn 56 x UK91	183939.1	18	14732.0	8	17934.6	8	-5080.9	6	4.3	5	897.0	17	7
25	NTA 93-21 x UK91	27480.8	2	125985.0	25	177681.8	25	-11817.3	23	17.0	26	1235.0	2	23
26	Aubarn 56	183867.7	17	13610.7	5	16324.7	5	-7179.1	9	2.6	1	904.7	16	4
27	NTA 93-21 x Aubarn 56	13762.9	1	274523.0	28	390967.7	28	-10815.0	18	25.1	28	1341.0	1	24
28	NTA 93-21	141423.7	12	110709.0	24	155747.4	24	16112.1	27	12.7	23	1026.0	8	26

Pi= cultivar superiority measure, Wi= Wricke ecovalence, SV= stability variance, S^2d_i= deviation from regression, ASV= AMMI stability value, ANOVA= analysis of variance, R= rank

Comparison of stability parameters for fibre strength is summarised in Table 8.13. The ranking of Wi, SV, S^2d_i and ASV were more or less similar in identifying stable genotypes, although there was some divergence especially on S^2d_i values. For example Wi, SV and ASV identified Aubarn 56 x UK91 as 1st but it was ranked 3rd by S^2d_i. The same was observed on Aubarn 56 x Delcot 344 ranked 2nd by Wi and SV but 8th by S^2d_i and 3rd by ASV. Okra leaf ranked 2nd by ASV, but 25th by Wi and SV and 22nd by S^2d_i Cultivar superiority performance *(Pi)* diverged from other methods and was in harmony with means (Table 8.13).

The overall ranking of genotypes for all stability parameters showed that the most stable genotypes for fibre strength were Delcot 344 x Acala SJ$_2$, Aubarn 56 x UK91, Aubarn x Okra leaf, Aubarn 56 x Delcot 344, NTA 93-21, Aubarn 56 x Acala SJ$_2$, UK91 x Okra leaf, Acala SJ$_2$ x Okra leaf, UK91 x Acala SJ$_2$ and NTA 93-21 x Delcot 344 (Table 8.13). Unstable genotypes were Aubarn 56 x MZ561, UK91 x MZ561, MZ561 x Okra leaf, Delcot 344 x MZ 561, UK91, Okra leaf and Delcot 344 x Okra leaf (Table 8.13).

Correlation coefficient and stability parameters comparison
Stability comparison was done using Spearman's rank correlation coefficient to compare the ranks and relationships among parameters (Table 8.14). Correlation results revealed that for five stability measurement procedures and overall means, some were positively and others negatively correlated. Lin and Binns *(Pi)* and overall mean showed the greatest deviation from other procedures since negative correlations were observed.

For lint yield, positive and highly significant rank correlation coefficient values were observed to SV with Wi (1.000), SV with ASV (0.940) and Wi with ASV (0.940). Correlation for these parameters (SV, Wi and ASV) were positive but not significant to S^2d_i except for ASV. A highly positive correlation coefficient value was observed between Pi and mean. Negative correlations were observed for Pi and mean with Wi, SV, ASV and S^2d_i stability measures. For fibre strength, negative correlations were observed for mean with Wi, SV, S^2d_i and ASV. High positive and significant correlation values were observed between SV

Table 8.13 Ranking of different stability procedures for 28 cotton genotypes for fibre strength

No	Genotype	Lin and Binns Pi	R	Wricke Wi	R	SV	R	Shukla R	Eberhart and Russell (S^2d_i)	R	AMMI ASV	R	ANOVA Mean	R	Overall R
1	Okra leaf	51.4	26	18.9	25	26.2	25	25	5.2	22	0.2	2	32.5	27	22
2	Acala SJ$_2$ x Okra leaf	46.2	19	2.4	6	2.6	6	6	0.3	5	0.5	10	32.7	23	8
3	MZ561 x Okra leaf	55.9	28	17.9	23	24.9	23	23	8.7	25	0.9	20	32.6	26	26
4	Delcot 344 x Okra leaf	30.0	10	18.7	24	25.9	24	24	9.3	26	1.1	24	34.8	10	23
5	UK91 x Okral eaf	46.3	20	3.1	7	3.6	7	7	-0.0	1	0.5	8	32.6	26	7
6	Aubarn 56 x Okra leaf	33.0	12	1.0	4	0.6	4	4	0.2	4	0.3	5	34.1	13	3
7	NTA93-21 x Okra leaf	37.5	16	7.4	16	9.8	16	16	3.7	20	0.8	16	34.0	14	16
8	Acala SJ$_2$	47.3	22	8.2	17	10.9	17	17	3.4	19	0.9	19	33.1	19	21
9	MZ561 x AcalaSJ$_2$	50.8	25	4.7	11	5.8	11	11	0.3	6	0.7	15	32.3	28	14
10	Delcot 344 x Acala SJ$_2$	26.5	6	0.9	3	0.6	3	3	0.0	2	0.3	6	35.1	7	1
11	UK91 x Acala SJ$_2$	27.3	7	3.4	9	3.9	9	9	1.7	12	0.5	9	35.3	6	9
12	Aubarn 56 x Acala SJ$_2$	45.4	18	1.1	5	0.7	5	5	0.4	9	0.3	4	32.7	24	6
13	NTA 93-21 x Acala SJ$_2$	25.8	5	12.9	21	17.8	21	21	1.8	16	1.0	21	36.5	3	17
14	MZ561	46.5	21	5.1	12	6.4	12	12	1.1	11	0.6	14	32.8	22	15
15	Delcot 344 x MZ561	49.2	25	20.2	26	28.2	26	26	6.9	24	1.3	26	33.3	17	27
16	UK91 x MZ561	10.0	1	194.4	28	278.2	28	28	80.3	28	5.0	28	37.3	1	24
17	Aubarn 56 x MZ561	48.8	23	56.9	27	80.8	27	27	15.4	27	2.3	27	33.1	18	28
18	NTA 93-21 x MZ561	20.8	3	5.2	13	6.6	13	13	1.7	14	0.8	17	35.6	4	12
19	Delcot 344	28.8	9	12.7	20	17.4	20	20	4.6	21	1.1	22	34.5	11	20
20	UK91 x Delcot 344	24.4	4	5.9	14	7.7	14	14	2.5	17	0.4	7	35.1	8	11
21	Aubarn 56 x Delcot 344	43.6	17	0.9	2	0.4	2	2	0.4	8	0.3	3	33.0	20	4
22	NTA 93-21 x Delcot 344	27.6	8	3.5	10	4.1	10	10	1.7	13	0.6	13	35.6	5	10
23	UK91	51.6	27	12.6	19	17.2	19	19	6.3	23	1.2	25	32.8	21	25
24	Aubarn 56 x UK91	37.5	15	0.9	1	0.4	1	1	0.1	3	0.1	1	33.5	16	2
25	NTA 93-21 x UK91	28.9	10	7.3	15	9.6	15	15	0.4	7	0.8	18	34.5	12	13
26	Aubarn 56	35.8	14	13.1	22	17.9	22	22	1.8	15	0.5	11	34.0	15	18
27	NTA 93-21 x Aubarn 56	35.3	13	12.1	18	16.5	18	18	2.6	18	1.1	23	34.9	9	19
28	NTA 93-21	15.9	2	3.2	8	15.9	8	8	0.9	10	0.6	12	37.1	2	5

Pi= cultivar superiority measure, Wi= Wricke ecovalence, SV= stability variance, S^2d_i = deviation from regression, ASV= AMMI stability value, ANOVA= analysis of variance, R= rank

187

Table 8.14 Spearman's coefficient rank correlation of stability parameters and means for lint yield (below diagonal) and fibre strength (bold and above diagonal)

	Pi	Wi	SV	S^2d_i	ASV	Mean
Pi		**0.166**	**0.166**	**0.137**	**0.039**	**0.917 ****
Wi	-0.177		**1.000 ****	**0.905 ****	**0.760 ****	**-0.058**
SV	-0.177	1.000 **		**0.905 ****	**0.760 ****	**-0.058**
S^2d_i	-0.325	0.369	0.369		**0.727 ****	**-0.115**
ASV	-0.113	0.940 **	0.940 **	0.439 *		**-0.227**
Mean	0.971 **	-0.212	-0.212	-0.292	-0.132	

* p ≤ 0.05, ** p ≤ 0.01, Pi= cultivar superiority measure, Wi= Wricke ecovalence, SV= stability variance, S^2d_i = deviation from regression, ASV= AMMI stability value

and *Wi* (1.000), *SV* with *ASV* (0.760) and *Wi* with *ASV* (0.760). Eberhart and Russell stability measure *(S^2d_i)* was positive and significant to ASV (0.727), *Wi* (0.905) and *SV* (0.905). Cultivar superiority *(Pi)* was positive and significant to mean (0.917).

8.4 DISCUSSION

Different rainfall patterns resulted in low yield especially at Bwanga and Mwanhala where little rains were received at the time of bud formation, boll setting and maturity. At Kanziga low rain was experienced during January and February, then resumed in March and prolonged to April, resulting in rejuvenation and boll formation (Table 8.1). Soil pH increased with depth confirming results of Rochester and Constable (2003). Wright (1999) and McLeod (2001) reported that low phosphorus and potassium lead to premature senescence of fruits. The negative relationship of sodium with potassium and phosphorus might be one of the limiting factors that led to differences in yield and fibre quality results for different locations including Bwanga and Mwanhala.

General performance and genotypic stability using five stability parameters and mean for cotton genotypes were revealed. Combined analysis (ANOVA) across four environments indicated that NTA 93-21 x Aubarn 56, NTA 93-21 x Delcot 344, NTA 93-21 x UK91, NTA 93-21, NTA 93-

21 x MZ561 and NTA 93-21 x Acala SJ$_2$ were the best performers (Table 8.4). These genotypes were identified as good performers for yield components, yield (seedcotton and lint), GOT and fibre quality. The percentage contribution to the total variation was high for location compared to genotypes for plant height, seedcotton and lint yield. This suggested that these characteristics were controlled by many genes (polygenic). Characteristics controlled by many genes are influenced by environment. Gad *et al*. (1974) and Singh and Singh (1980) reported additive genetic variation for seedcotton and lint yield. Variation in location indicated more heterogeneous experimental fields, differences in soil fertility, moisture and even land levelling (Tables 8.1 and 8.2). Low seedcotton yield observed at Bwanga and Mwanhala (section 7.3.1 and Table 7.4) could be due to low total nitrogen and organic carbon (Table 8.2) observed at these sites. Silvertooth *et al.* (2001) reported that the crops demand for nitrogen is measured by fruit retention and vigour. Therefore in combination with other factors including rainfall, this could lead to low seedcotton yield.

For bolls/plant, seed/boll, GOT, fibre length and fibre strength, the contribution of genotypes to total SS variation was high compared to location. This indicated that these characteristics were controlled by a single or few major genes. This was reported by Sayal and Sulemani (1996). The significant differences observed between locations can be attributed by environment and soil types. Single or few major genes determined fibre quality, as indicated by the high percentage contribution of genotypes to the total variation. However, high interaction between fibre quality and environmental factors can influence the realisation of the potential of the genotype for fibre quality. This can suggest that though few major genes control fibre quality, the genotype's fibre quality can be modulated by environmental changes especially during floral anthesis. For micronaire value the contribution of location was high (40%) compared to genotype (33%) indicating that many genes are controlling this characteristic. Christidis and Harrison (1955) and Sikka and Joshi (1960) reported that micronaire value is quantitatively inherited and affected by the environment.

According to Myers (2004), in the past breeders were looking for higher yielding genotypes but lately, however, stable and sustainable yields under varying environmental conditions have gained importance over increased yield. Lint yield is an important character for cotton genotypes that needs improvement. It is controlled by many genes since it depends on the

seedcotton yield and ginning outturn of a variety. Fibre strength is an important fibre quality trait, especially due to improvements in textile processing, particularly advances in spinning technology that need strong fibres. This has led to increased emphases on breeding cotton for both improved yield and improved fibre properties. Results (Table 8.3) indicated that lint yield and fibre strength G x E interactions were highly significant. Although the ANOVA can partition the total variance into main effects, as an additive model it fails to exhaustively analyse non-additive G x E interaction. Therefore different stability measurements for lint yield and fibre strength were needed to analyse and identify stable cotton genotypes against the environment.

Based on stability, cultivar superiority measure *(Pi)* was in harmony with the mean results, as was reported by Purchase (1997), indicating that *Pi* is a performance rather than stability measurement. *Pi* indicated that among the top 10 genotypes ranked for lint yield and fibre strength, NTA 93-21 x MZ561, NTA 93-21 and NTA 93-21 x UK91 were stable for both measured characteristics. From the concept that a stable genotype relates to the performance of the best genotype in the trial might lead to biasing of the high yielding genotype where the low yielding genotype will always be unstable. This concept made cultivar superiority measure *(Pi)* to be in harmony with the mean. Cultivar superiority measure *(Pi)* showed the greatest divergence from other procedures as was noted by Purchase (1997).

Results obtained from *Wi* and *SV* stability methods were similar for measuring stability. This indicated that the two procedures are equivalent for ranking purposes as Wricke and Weber (1980) and Purchase (1997) noted. Shukla's stability measure and Wricke's ecovalence both have Type II stability according to Lin *et al*. (1986) and fall into the dynamic stability concept of Becker and Léon (1988). It was observed that genotypes Aubarn 56 x Delcot 344, Aubarn 56 x Acala SJ_2, Delcot 344 x Acala SJ_2 and UK91 x Okra leaf were the most stable for both characteristics (lint yield and fibre strength). These two stability measures efficiently ranked genotypes from most to least stable and can be used in cotton stability studies. However, apart from identifying stable and the unstable genotypes, it fails to locate intermediates and their low and high yielding environments.

The Eberhart and Russell method uses joint regression analysis by considering the deviation from the regression line to be as small as possible (S^2d_i=0) and regression coefficient approaching a unit (*bi*=1). Eberhart and Russell stability can be related to Becker and Léon (1988) dynamic concept of stability and to Type III stability as defined by Lin *et al*. (1986). Using this procedure it is difficult to identify stable and unstable genotypes because two different values (S^2d_i and *bi)* are used. Purchase (1997) suggested that, the use of this model in describing G x E interaction and stability of genotypes is recommended on condition that it is used in conjunction with other methods preferably multivariate methods of analysis. Therefore, the concept of this method is more of measuring the stability for genotypes sensitive to high yielding or low yielding environments to which genotypes may be specifically adapted. This was also noted by Becker and Léon (1988). This method identified genotypes for high yielding sites like Ukiriguru and Kanziga and low yielding sites like Bwanga and Mwanhala.

Results for three stability procedures, Wricke's, Shukla's and Eberhart and Russell's stability methods on lint yield identified genotypes Aubarn 56, Aubarn 56 x Acala SJ$_2$, MZ561 x Okra leaf and Aubarn 56 x UK91 as most stable. For fibre strength, genotypes Aubarn 56 x Delcot 344, Delcot 344 x Acala SJ$_2$, Aubarn 56 x Okra leaf, Aubarn 56 x Acala SJ$_2$ and UK91 x Okra leaf were identified as the most stable. Genotype Aubarn 56 x Acala SJ$_2$ was the most stable for both characteristics when both characteristics were considered with the three methods.

When IPCA biplots for lint yield were included, Aubarn 56 x Acala SJ$_2$ was still identified as the most stable followed by MZ561 x Okra leaf and Aubarn 56 x UK91. For fibre strength, genotypes Aubarn 56 x Delcot 344 and Aubarn 56 x Acala SJ$_2$ were identified as most stable while genotypes UK91 x MZ561 and Delcot 344 x MZ561 were unstable. Results for IPCA 1 and IPCA 2 scores biplots were in harmony with ASV ranking and these were in harmony with the other three stability methods ranking (*Wi, SV* and $S^2d_i)$. AMMI analysis is used to include ANOVA and PCA into a unified approach (Becker and Léon, 1988). Becker and Léon (1988) identified AMMI as a relatively simple method using IPCA 1 and IPCA 2 scores to determine *ASV* which can be used to identify superior genotypes through ranking. Generally, using these stability parameters, most results were related in terms of stability for the 28 cotton genotypes. However, *Pi* differed from other procedures.

The correlation rank coefficient for *Wi* and *SV* was unity (r=1): Wricke and Weber (1980), Kang *et al.* (1987) and Lin *et al.* (1986) reported a similar relationship between *SV* and *Wi* and concluded that these methods were identical in ranking cultivars for stability.

Based on the current study the combinations with NTA 93-21 were shown as good performers using ANOVA. However, on stability measures most of them were shown to be unstable or adapted to high yielding environments for both characteristics (lint yield and fibre strength). This indicated that most NTA 93-21 combinations can not be recommended for both environments, northern and southern zones of the WCGA's. Soil nutrients variation observed (Table 8.2) and rainfall patterns contributed to genotype's performance especially for polygenic characteristics.

Based on stability measures used in the current study, Lin and Binns should not be used for cotton genotype stability as it is based on the performance relative to the best performer in the trial, rather than stability measure. Shukla's and Wricke's methods apart from few limitations can be useful in cotton genotype stability studies. Eberhart and Russell's procedure of deviation from the regression was useful in the stability, however, the use of the two values, bi and S^2d_i can be difficult in interpretation. AMMI appears to be more accurate in describing both G x E stability and locating genotypes to specific environments. According to Lin *et al.* (1986), there is a need of a breeder to elaborate the type of stability required and the environment to be used before planning an experiment.

8.5 CONCLUSIONS AND RECOMMENDATIONS

From this study it can be recommended that breeders should understand the underlying concepts and practical limitations of genotype stability that can be the basis for breeding for fluctuating environmental conditions. Genetic improvement of yield and fibre quality will be fruitless if the responses of the the new genotype to the growing environment prevent full realisation of the enhanced genetic potential of the genotype. Using ANOVA the best genotypes for most of the characteristics can be recommended for future studies since analysis of variance identified them as best genotypes.

From the stability results through ranking by using five stability measures, the genotypes were identified as stable and unstable. Since the aim is to breed for stable varieties with improved yield and fibre strength, it can be recommended to continue doing stability analysis studies for these genotypes over more locations and years to understand the interaction between G x E x Y (year). This will be multipurpose taking into account variation of environment as well as that these genotypes are still at an early generation stage of crossing.

Since the best genotypes identified by ANOVA and stability procedures were different, future studies are needed to identify their specific environments. The procedures of Shukla, Wricke and Eberhart and Russell are useful in characterising cotton genotypes stability. However, AMMI analysis was able to identify stable genotypes over environments. Therefore, for cotton stability studies AMMI is recommended as the best method. It is able to give interaction results for genotype, environment and the relationship between genotype and environment. Results can be used to identify stable and unstable genotypes so that the unstable genotypes can be adapted to their specific environments. The highly significant correlation observed between *Wi, SV* and *ASV* indicated the close relationship among them in measuring stability for these cotton genotypes. However, *Wi* and *SV* are too general, making identification of specific adaptability of the genotype difficult. These two methods can work hand in hand with AMMI for more elaborative results for recommendation of genotypes and environments.

CHAPTER 9

GENERAL CONCLUSIONS AND RECOMMENDATIONS

The current study mainly focused on genetic diversity, diallel evaluation and G x E stability analysis. Genetic diversity is important for efficient management of germplasm and utilisation of material in breeding programmes. Based on genetic diversity using agronomical characteristics, oil content and fatty acids (quantitative traits), high levels of variation were seen indicating a possibility of improvement through selection and hybridisation. These quantitative characteristics are influenced by the environment indicating that environment is a factor for consideration when using agronomical characteristics and oil and fatty acids for genetic diversity studies.

Qualitatively inherited morphological characteristics can be used to characterise varieties for collection and maintenance of germplasm and for parental selection through heterotic groups to improve local varieties. Since morphological traits do not cover the entire genome, this has to be confirmed on DNA level. Furthermore, conventional breeding has to be supplemented with genetic mapping and transformation to identify and transfer specific genes in order to hasten progeny selection for increased yield and fibre quality.

AFLP results indicated a narrow genetic base as shown by high genetic similarities. The heterotic groups identified could be used for increasing diversity in cotton breeding programmes through hybridisation. Furthermore, cultivated cotton varieties and other tetraploid species (*G. barbadense)* from other countries should be introduced to increase variation in order to enable improvement of the available cotton material in Tanzania.

Low levels of correlation existed between agronomical, morphological, combined agronomical and morphological and AFLP analysis. AFLP analysis reflected the true expression of genotypes, while agronomical analyses encompassed the expression of genotype, environment and their interaction. Morphological (qualitative) characterisation was limited by few data points used for analysis. AFLP is a promising marker system especially for cotton varieties, which are closely related, with low levels of polymorphism and thus high levels of genetic uniformity.

Although all methods did not provide similar description of relationships between varieties, there existed some consistency in discriminating varieties which were closely related and ones which were distantly related. All methods have advantages and disadvantages for practical applications under different circumstances. It can be recommended that all methods should be used in plant breeding and germplasm conservation in Tanzania. Future research should focus on comparing these methods in terms of feasibility, efficiency and accuracy by involving more tests over different environmental trials and years (for agronomical and morphological characterisation). Furthermore, molecular analysis should concentrate on the use of other DNA marker systems, including SSRs, to generate more information and to develop markers linked to genes of important traits like disease and pest resistance.

Parents with high and positive GCA as well as combinations with high and positive SCA for the characteristics studied were identified. This in turn showed potential to increase variability in the cotton population in Tanzania and consequently improve cotton varieties. Estimated heterosis and heritability values indicated the presence of additive traits controlling the characteristics studied except for a few characteristics that had both additive and non-additive genes controlling the same characteristics. Therefore, breeders may utilise good general combiners in breeding programmes using breeding techniques like recurrent selection, pedigree and backcrossing for improvement of cotton characteristics.

Based on oil content, fatty acids, agronomical characteristics, morphological characteristics, AFLP analysis, diallel evaluation and G x E studies, the most promising genotypes were identified. For example, Delcot 344 and NTA 93-21 with their combinations were the best for most of the characteristics studied. These are recommended for cross hybridisation with local varieties for cotton breeding programme improvement in Tanzania. Other genotypes with moderate performance were NTA 93-15 and Aubarn 56. Based on the stability study, the combinations for the above genotypes were unstable but high yielding. Due to high positive GCA and SCA values these genotypes can still be improved through breeding techniques like pedigree and backcrossing.

A significant positive correlation between seedcotton yield and lint yield with boll weight and bolls/plant indicated that improvement of one characteristic could influence the other

characteristic. However, the significant negative correlation observed between yield components and fibre quality has to be considered in cotton improvement through selection for these characteristics.

Combined analysis of variance across environments indicated the effect of environmental factors in the expresion of yield and fibre quality. It is recommended that breeders should understand the underlying concepts and practical limitations of genotype stability as basis for breeding for fluctuating environmental conditions. Multilocational evaluation helps to express the genotypes and can offer more opportunities of selecting genotype for many or specific environments. Based on ranking it is recommended that further stability analysis studies for these genotypes over more locations and years should be done to understand the interaction between G x E x Y (year).

The procedures of Shukla, Wricke and Eberhart and Russell are useful in characterising genotype stability. For cotton stability studies, AMMI is recommended as the best method. It is able to give interaction results for genotype, environment and the relationship between genotype and environment. Results can be used to identify stable and unstable genotypes. Unstable genotypes should be adapted to specific environments. *Wi* and *SV* analyses are too general, making identification of specific adaptability of genotypes difficult. These two methods can work hand in hand with AMMI for more elaborative results for recommendation of genotypes and environments. For consistent results on genotype stability, it is necessary for a breeder to identify the type of stability method and the environment to work in before planning an experiment.

From this study, knowledge was obtained on genetic diversity using oil content, fatty acids and molecular markers which were not applied before in Tanzania. Information on combining abilities and G x E stability was obtained for the available germplasm, although further studies are still needed. Therefore, the study was of importance for gaining more knowledge for improving the cotton industry in Tanzania

CHAPTER 10

SUMMARY

Key words: cotton (*Gossypium hirsutum* L.), genetic diversity, oil, fatty acids, morphological, agronomical, AFLP, diallel, genotype x environment, stability.

1. Thirty cotton varieties were studied. Seven parents generated 21 F_1 diallel progenies. The main objective was to evaluate genotypes through oil content, fatty acids, morphological and molecular characteristics. Statistical procedures were performed with Agrobase, NCSS and NTYSS computer programmes. Correlation, heterosis, heritability, G x E interaction and stability values were obtained. Stability across four environments was analysed by five stability measures, that of Shukla, Wricke, Lin and Binns, Eberhart and Russell and the AMMI analysis.

2. Significant differences were observed for oil and fatty acids content. Oil content results revealed an average of 20.23%. The ratio for polyunsaturated:saturated fatty acid was 2:1 and for unsaturated:saturated 3:1. The predominant fatty acids were linoleic, palmitic, oleic, stearic, myristic, palmitoleic, arachidic and behenic. Palmitic fatty acid was high while oleic and stearic fatty acids were low for almost all varieties studied. Oil and fatty acid content were significantly influenced by the environment. CIM 70 and Cyto 12/74 can be selected as parents for increased oleic and stearic fatty acid and Stoneville 506 as a parent for reduced palmitic fatty acid, increased oleic, stearic and linolenic fatty acids in hybridisation.

3. ANOVA results indicated significant differences between varieties for all agronomical characteristics studied, where NTA 93-21 and NTA 93-15 were the best performers almost for all characteristics. Genetic distances for morphological (qualitative characteristics) analysis ranged from 0.18 to 0.80. Two major groups, one including many varieties from Africa and the other including varieties from the USA, were revealed.

4. AFLP analysis showed high genetic similarities (0.85 to 0.98). Varieties were distinguished according to pedigree and origin. The increased genetic similarity observed, indicates the need for the introduction of more diverse cotton material. Comparison of agronomical and AFLP markers for determining genetic diversity

indicated that genetic similarities between varieties were low for agronomical compared to AFLP markers. The correlation coefficient between AFLP analysis and agronomical genetic similarities was not significant,

5. Highly significant differences between genotypes over environments were observed and NTA 93-21 crosses outperformed the seven parents for almost all characteristics. The parent Okra leaf performed poorly for many characteristics. Trial sites Ukiriguru followed by Kanziga were high yielding environments.

6. Combining ability analysis results indicated significant variation in GCA estimates among parents for different characteristics. Genotypes having high positive GCA estimates for yield components exhibited negative GCA estimates for fibre quality, excluding NTA 93-21. Some combinations indicated positive and significant SCA effects for seedcotton, boll weight, GOT and lint yield. The GCA/SCA ratios for most of the studied characteristics were larger than one.

7. Correlation among characteristics indicated highly significant positive and negative correlations. The observed high and significant GCA and SCA effects for those characteristics are promising for cotton population improvement. Positive heterosis for boll weight was observed for more than 90% of the combinations. Hybrids with high SCA values presented high values of heterosis. Heritability values for the studied characteristics were moderate to high.

8. The combined ANOVA over four environments indicated significant differences between genotypes and locations for all characteristics. The percentage contribution of location to total variation was high compared to genotype variation except for a few characteristics. Highly significant differences of G x E for lint yield and fibre strength were observed.

9. Three stability procedures (Wricke, Shukla and ASV) identified similar stable genotypes. Eberhart and Russell's procedure, was in harmony with the above three methods but with some divergence. Cultivar superiority measure deviated from the other four measures. Therefore the procedures of Wricke, Shukla and AMMI can be applied for cotton stability studies.

OPSOMMING

Sleutelwoorde: katoen (*Gossypium hirsutum* L.), genetiese diversiteit, olie, vetsure, morfologies, agronomies, AFLP, dialleel, genotipe x omgewing, stabilitieit

1. Dertig katoen cultivars is gebruik in die studie. Sewe ouers het 21 F1 dialleel nageslagte gegenereer. Die hoof doelwitte was om die genotipes te evalueer deur bepaling van olie inhoud, vetsure, morfologiese en molekulêre eienskappe. Statistiese prosedures is uitgevoer met Agrobase, NCSS en NTYSS rekenaar programme. Korrelasie, heterose, oorerflikheid, GxE interaksie en stabiliteitswaardes is bereken. Stabiliteit is oor vier omgewings getoets met vyf stabiliteitsmetings, die van: Shukla, Lin en Binns, Eberhard en Russel en AMMI.

2. Betekenisvolle verskille is waargeneem vir olie en vetsuur inhoud. Olie inhoud was gemiddeld 20.23%. Die verhouding van polie-onversadigde:versadigde vetsure was 2:1 en onversadig: versadig was 3:1. Die oorwegende vetsure was linoleïen, palmitien, oleïen, stearien, miristien, palmitoleïen, eikosanoë, en dokosanoë suur. Palmitien suur was hoog, terwyl oleïen en stearien vetsure baie laag was vir meeste inskrywings. Daar was 'n groot omgewingseffek op olie en vetsure. CIM 70 en Cyto 12/74 kan gebruik word as ouers vir verhoogde oleïen en steariensuur inhoud en Stoneville 506 as ouer vir verlaagde palmitiensuur, verhoogde oleïen, stearien en linoleïensuur in kruisings.

3. ANOVA resultate het betekenisvolle verskille aangetoon tussen cultivars vir gemeette agronomiese eienskappe. NTA 93-21 en NTA 93-15 het die hoogste waardes gehad vir byna al die eienskappe. Genetiese afstande vir morfologiese (kwalitatiewe eienskappe) het gewissel van van 0.18 tot 0.80. Twee hoofgroepe, een met die meeste inskrywings van Afrika en die ander met inskrywings van die VSA is gekry.

4. AFLP het hoë genetiese ooreenkomste getoon (0.85 tot 0.98). Cultivars is onderskei volgens hulle stamboom en oorsprong. Hierdie hoë genetiese ooreenkoms het aangedui dat meer diverse katoen materiaal gebruik moet word. Vergelyking van agronomiese en AFLP merkers vir genetiese diversiteit het getoon dat genetiese ooreenkomste vir agronomiese eienskappe tussen cultivars laag was in vergelyking met die AFLP merkers. Die korrelasie koeffisiënt tussen die AFLP's en agronomiese merkers was nie betekenisvol nie.

5. Hoogs betekenisvolle verskille is gesien tussen genotipes oor omgewings, en NTA 93-21 kruise het die sewe ouers oortref vir byna alle eienskappe. Die ouer "Okra leaf" het swak presteer vir baie eienskappe. Ukiriguru gevolg deur Kanziga was die hoogste opbrengs omgewings.

6. Kombineervermoë analise resultate het betekenisvolle variasie vir GCA waardes aangetoon. Genotipes wat hoë positiewe GCA waardes gehad het vir opbrengs komponente het negatiewe GCA waardes gehad vir veselkwaliteit, behalwe NTA 93-21. Sommige kombinasies het positiewe betekenisvolle SCA effekte gehad vir saadkatoen, bolgewig, GOT en lintopbrengs. Die GCA/SCA verhoudings vir meeste gemeette eienskappe was meer as een.

7. Daar was hoogs betekenisvolle positiewe en negatiewe korrelasies tussen eienskappe. Die hoë en betekenisvolle GCA en SCA effekte vir hierdie eienskappe is belowend vir toekomstige verbetering van katoen populasies. Positiewe heterose is gesien vir bolgewig in meer as 90% van die kombinasies. Basters met hoë SCA waardes het hoë heterose waardes gehad. Oorerflikheidswaardes vir die gemeette eienskappe was gemiddeld tot hoog.

8. Die gekombineerde ANOVA oor vier omgewings het betekenisvolle verskille tussen genotipes en omgewings aangetoon vir alle eienskappe. Die persentasie bydrae van omgewing tot totale variasie was hoog in vergelyking met die bydrae van die genotipes behalwe vir 'n paar eienskappe. Hoogs betekenisvolle verskille vir GxE vir lint opbrengs en veselsterkte is gesien.

9. Drie stabilteitsprosedures (Wricke, Shukla en ASV) het dieselfde stabiele genotipes uitgewys. Eberhart en Russel se presedure was baie dieselfde as die ander metodes met enkele afwykings. Cultivar superioriteits meting het afgewyk van die ander vier metodes. Daarom kan Wricke, Shukla en die AMMI prosedures gebruik word vir katoen stabiliteitsstudies.

REFERENCES

Abdalla, A. M., A.A. Aboul-El-Zahab, S.R.H. Radwan and P. Dugger, 1999. Combining ability for yield and earliness of Pima x Egyptian cotton cultivars crosses. In: D. Richter (Ed.), Proceedings Beltwide Cotton Conferences pp 473-477. National Cotton Council Memphis, TN.

Abdalla, A.M., O.U.K. Reddy, K.M. El-Zik and A.E. Pepper, 2001. Genetic diversity and relationship of diploid and tetraploid cottons revealed using AFLP. *Theoretical and Applied Genetics* **102**: 222-229.

Abdukarimov, A.S., S. Djataev and I. Abdukarimov, 2003. Cotton research in Uzbekistan: Elite varieties and future cotton breeding. Proceedings World Cotton Research Conference pp 5-15. Cape Town, South Africa.

Adugna, W.G., 2002. Genetic diversity of linseed (*Linum usitatissimum* L.) in different environments. PhD Thesis pp 183-201. Department of Plant Sciences, UOVS, Bloemfontein, South Africa.

Agrimarket INFO, 1998. Cotton in South Africa. Agricultural News. Newsletter of the National Department of Agriculture (R.S.A.). Directorate communications (Pub).

Agrobase, 2000. Agronomix Software Inc, 71 Waterloo st, Winnipeg, Manitoba, R3NOS4, Canada.

Ahmad, R.T., I.A. Khan and M. Zubair, 1997. Diallel analysis for seed-cotton yield and its contributing traits in upland cotton *(Gossypium hirsutum). India Journal of Agricultural Sciences* **67:** 583-585.

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan, 1992. Length polymorphism of simple sequence repeat DNA in soybean. *Genetics* **132**: 1131-1139.

Alam, A.K.M.R., N.C. Roy and H. Islam, 1992. Heterosis and combining ability in upland cotton *(Gossypium hirsutum). Annals of Bangladesh Agriculture* **2:** 31-39.

Aldrich, P.R., J.L. Hamrick, P. Chavarriaga and G. Kochert, 1998. Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera. Molecular Ecology* **7**: 933-944.

Ali, A.M., O.M. Ahmed, B.C. Misaka, A.H. Latif, K. Elsiddig and E.A. Babiker, 2003. Characterisation of cotton germplasm and its utilization in breeding for major production constraints in Sudan. Proceedings World Cotton Research Conference pp 16-21. Cape Town, South Africa.

Allard, R.W., 1960. Principles of Plant breeding pp 485. John Wiley and Sons, Inc. New York.

Allen, R.D. and D.L. Auld, 2002. Molecular Genetic Tools for the use of exotic germplasm resources in cotton improvement. Departments of Biological Sciences and Plant and Soil Sciences. *http://genetics.biol.ttu.edu.pdf/proposal.pdf.*

Altaf Khan, M., G.O. Myers and J.McD. Stewart, 2002. Molecular markers, genomics and cotton Improvement. In: M.S. Kang (Ed.), Crop Improvement pp 253-284. Food products Press, New York.

Altaf Khan, M., G.O. Myers, J.McD. Stewart, J. Zhang and R.G. Cantrell, 1999. Additional of new markers to the trispecific cotton map. Proceedings Beltwide Cotton Conference pp 439. National Cotton Council, Memphis, TN.

Altaf Khan, M., J.McD Stewart, R.G. Cantrell, M. K. Wajahatullah and J. Zhang, 1997. Molecular and morphological genetics of trispecific F_2 population of cotton. Proceedings Beltwide Cotton Conference pp 448-452. National Cotton Council, Memphis, TN.

Altaf Khan, M., J. Zhang, J.McD. Stewart and R.G. Cantrell, 1998. Integrated molecular map based on a trispecific F_2 population cotton. Proceedings Beltwide Cotton Conference pp 491. National Cotton Council, Memphis, TN.

Andries, J.A., A.G. Douglas and R.C. Albritton, 1971. Performance of normal, okra and superokra leaf types in three-row width. In: J.M. Brown (Ed.), Proceedings Beltwide Cotton Conferences pp 59-60. National Cotton Council Memphis, TN.

Andries, J.A., J.E. Jones, L.W. Sloan and J.G. Marshall, 1969. Effects of okra leaf shape on boll rot, and other important characters of upland cotton (*G. hirsutum* L.). *Crop Science* **9:** 705-710.

Anonymous, 1972. Genetic vulnerability of major crops. *Cotton* pp 269-281. National Academy of Sciences. Washington, D.C.

Anonymous, 1981. Cotton Council International and National Cotton Council of America, Memphis, TN.

Antoni J., V. Scott and G.K. John, 1991. RAPD markers, a new technology for genetic mapping and plant breeding. *AgBiotech News and Information* **3**: 646-648.

ANZFA, 2002. Draft assessment report: Oil and linters derived from insect protected cotton containing event 15985. Full assessment S15, Application A436. Australia New Zealand Food Authority.

Arunachalam, V., 1976. Evaluation of diallel cross by graphical and combining ability methods. *Indian Journal Genetics* **36**:358-366.

Arus, P. and M. Gonzalez, 1993. Marker assisted selection. In: M.D. Hayman, N.O. Bosemark and I. Romagosa (Eds.), Plant Breeding: Principles and prospects pp 314-331. Chapman and Hall, London.

Ashraf, M. and S. Ahmad, 2000. Genetic effects for yield components and fibre characteristics in upland cotton (*Gossypium. hirsutum* L.) cultivated under salinized (NaCl) conditions. *Agronomie* **20**: 917-926.

Ashwathama, V.H., B.C. Patil, S.R. Kareekatti and T.S. Adarsha, 2003. Studies of heterosis for biophysical traits and yield attributes in cotton hybrids. Proceedings World cotton research conference pp 239-247. Cape Town, South Africa.

Azhar, F.M., M. Naveed and A. Ali, 2004. Correlation analysis of seed cotton yield with fibre characteristics in *Gossypium hirsutum. International Journal of Agriculture and Biology* **4**:656-658.

Baffes, J., 2002. Tanzania's Cotton Sector. Constraints and Challenges in a Global Environment. Africa Region Working Paper Series No. 42.

Baker, R.J., 1978. Issues in Diallel Analysis. *Crop Science* **18**: 533-536.

Baker, R.J., 1988. Tests for crossover genotype x environment interactions. *Canadian Journal of Plant Science* **68:** 405-410**.**

Baloch, M.J., H. Bhutto, R. Rind and G.H. Tunio, 1996. Combining ability estimates in 5x5 diallel *intra-hirsutum* crosses. *Pakistani Journal of Botany* **27**: 121-126.

Baril, C.P., J.B. Denis, R. Wustman and F.A. Van Eeuwijk, 1995. Analyzing genotype-by-environment interaction in Dutch potato variety trials using factorial regression. *Euphytica* **82**: 149-155.

Barret, B.A. and K.K. Kidwell, 1998. AFLP based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop Science* **38:** 1261-1271.

Basford, K.E. and M. Cooper, 1998. Genotype x environmental interactions and some considerations of their implications for wheat breeding in Australia. *Australian Journal Agricultural Research* **49**: 154-174.

Becker, H.C., 1981. Correlations among some statistical measures of phenotypic stability. *Euphytica* **30**: 835-840.

Becker, H.C. and J. Léon, 1988. Stability analysis in plant breeding. *Plant Breeding* **101**: 1-23.

Becker, J., P. Vos, M. Kuiper, F. Salamini and M. Heun, 1995. Combined mapping of AFLP and RFLP markers in barley. *Molecular Genetics* **249**: 67-73.

Ben-Har, A., A. Charcosset, M. Bourgoin and J. Guard, 1995. Relationships between genetic markers and morphological traits in maize inbred lines collection. *Euphytica* **84**: 145-154.

Bie, S., F.L. Kong, Y.Y. Zho, G.M. Zhang, Q.Y. Zhang and X.G. Wang, 2001. Genetic diversity analysis of representative elite cotton varieties in three main cotton regions in China by RAPD and its relation with agronomic characteristics. *Scientia Agricultura Sinica* **34:** 597-603.

Blears, M.J., S.A. De Grands, H. Lee and J.T. Trevors, 1998. Amplified Fragment Length Polymorphism (AFLP): A review of the procedure and its applications. *Journal of Industrial Microbiology and Biotechnology* **21**: 99-114.

Bojinov, B. and J.M. Lacape, 2003. Molecular markers for DNA fingerprinting in cotton. *http://Bojinovo_WCRC3_texte.doc.*

Brown, A.H.D., C.L. Brubaker, J.M. Stewart, M.J. Kilby and J.P. Grace, 1999. Production of fertile hybrid germplasm with diploid Australia *Gossypium* species for cotton improvement. *Euphytica* **108**: 199-213.

Browse, J.C.S., 1991. Glycerolipid metabolism: biochemistry and regulation. *Annual Review Plant Physiology Plant Molecular Biology* **42**: 467-506.

Brubaker, C.L., A.H. Paterson and J.F. Wendel, 1994. Reevaluating the origin of domesticated cotton (*Gossypium hirsutum*) using nuclear restriction fragment length polymorphisms (RFLPs). *American Journal of Botany* **81**: 1309-1326.

Brubaker, C.L., F.M. Bourland and J.E. Wendel, 1999. The origin and domestication of cotton pp 3-31. John Wiley and Sons, Inc., New York.

Bruce, R.L., L.A.B. Akbari, A.E. Cook, S. Kailasapillai and K.R. Lyon, 2001. DNA markers for resistance to fungal diseases in cotton. Publication. Australia Cotton Cooperative Research Centre. *http://www.mv.pi.csiro.au/Publicat/conf/coconf00/PLNTPATH/66/66.htm*

Bunyecha, K.F. and K.J. Tamminga, 1995. Analysis of factors determining the prices difference between producer and export prices for cotton. Field Note: **43**. FSRP. LZARI.

Bustin, J. and A. Charcosset, 1997. Relationships between phenotypic and genetic markers distance: Theoretical and experimental investigations. *Heredity* **78**: 447-483.

Cantrell, R.G., 2000. Germplasm engineering in cotton. Proceedings Beltiwide Cotton Conference pp 500. National Cotton Council, Memphis, TN.

Carvalho, L.P., C.D. Cruz, C.F. Moraes and L.P. De-Carvalho, 1995. Diallel analysis of yield and other traits in cotton. *Revista Brasiliera de Genetica* **18**: 93-97.

Carvalho, L.P. and L.P. De-Carvalho, 1995. Genetic control of fibre percentage and boll weight in cotton. *Revista Ceres* **42:** 626-636**.**

Chang, S.S., R.J. Paterson and C.T. Ho, 1978. Chemical reactions involved in the deep fat frying of foods. *Journal of the American Oil Chemists Society* **55**: 718-727.

Chapman, K.D. I., A. Brown, S. Sparece, S.A. Kinney, A.J. Kinney, A.J. Ripp and K.G. Pirtle, 2001. Transgenic cotton plants with increased seed oleic acid content. *Journal of the American Oil Chemists Society* **78**: 941-947.

Charley, H. and C. Weaver, 1998. Foods: A scientific approach (Third edition) pp 582. Prentice-Hall Inc. Merrill.

Chaudhary, R.C., 1991. Introduction to plant breeding pp 261. Oxford and IBH Publishing Co. PVT. LD. New Delhi-Bombay.

Chen, Z.F., Z.W. Zhang, Z.Y. Li and Y.F. Wang, 1991. Study on the correlation between the earliness of Upland cotton and its yield and fibre quality. *China Cottons* **5**:16-17.

Cherry, J.P. and H.R. Leffler, 1984. Seed. In: R.J. Kohel and C.F. Lewis (Eds.), Cotton pp 511-569. ASA, Madison, WI.

Cherry, J.P., L.A. Kohel and W.H Powell, 1981. Cottonseed quality factors affecting feed and food uses. In: J.M. Brown (Ed.), Proceeding Beltwide cotton Conference pp. 266-283. National Cotton Council Memphis, TN.

Christidis, B.G. and G.J. Harrison, 1955. Cotton growing problems. McGraw-Hill Book Company, Inc. New York.

Christie, W.W., 2003a. What is lipid? http:// *www.lipid.co.uk*

Christie, W.W., 2003b. Triacylglycerols: Structure, composition and analysis. The Scottish Research Institute, Invergowrie, Dundee, Scotland. http://www.lipid.co.uk/infores/Lipids/TAG/index.htm.

Common market for Eastern and Southern Africa (COMESA), 2005. Common investiment area-Tanzania. *webmaster@comesa.int*

Coors, J.G., S. Pandey, M. van Winkel, A.R. Hallaver, D.C. Hess, K.R. Lamkey, A.E Melchinger and C.W. Stuber, 1999. Preface pp 524. In: J.G. Coors and S. Pandey (Eds.), Genetics Exploitation of Heterosis in Crops. ASA, CSS, SSSA, Madison, WI.

Cornelius, B.K. and C.H. Sneller, 2002. Yield and molecular diversity of soybean lines derived from crosses of Northern and Southern Elite parents. *Crop Science* **42**: 642-647.

Cowley, E.J., 1966. Development of the cotton growing industry in Nigeria, with special reference to the work of the British Cotton growing Association. *Cotton Growing Review* **43:** 169-95.

Cox, D.R., 1984. Interaction. *Statistics Review* **52:** 1-31.

Coyle, G.G. and C.W. Smith, 1997. Combining ability for within-boll yield components in cotton (*Gossypium hirsutum* L.). *Crop Sci*ence **37**: 1118-1122.

Crossa J., 1990. Statistical analysis of multi-location trials. *Advances in Agronomy* **44:** 55-85.

Crossa, J. and P.L. Cornelius, 1993. Recent development multiplicative models for cultivar trials. In: D.R. Buxton, R. Shibles, R.A. Fortberg, B.L. Blad, K.H. Asay, G.M. Paulsen and R.F. Wilson (Eds.), International Crop Science pp 571-577. Crop Science society of America, Madson, WI.

Culp, T.W. and D.C. Harrel, 1973. Breeding methods for improving yield and fiber quality upland cotton (*Gossypium hirsutum* L.). *Crop Science* **13**: 686-689.

Dabholkar, A.R. 1992. Elements of Biometrical Genetics. Concept Publ. Camp., New Delhi, India.

Dedaniya, A.D. and K.V. Pethani, 1994. Genetic variability correlations and path analysis in deshi cotton (*Gossypium arboreum* L.). *Indian Journal of Genetics and Plant Breeding* **54:** 229-234.

Deussen, H., 1992. Improved cotton fiber properties-the textile industry's key to success in global competition. In: C.R. Benedict and G.M. Jividen (Eds.), Cotton fiber cellulose: Structure function and utilization conference pp 43-63. National Cotton Council, Memphis, TN.

Dice, L.R., 1945. Measures of amount of ecologic association between species. *Ecology* **26**: 297-302.

Dippenaar-Schoeman, A., 1999. Spiders as biochemical agents. ARC-news. In: Agricultural Newsletter pp 12-16. National department of Agriculture Directorate Communications (Pub).

Dougherty, R.M., M.A. Alman and J.M. Iacon. 1995. Effect of diets containing high or low amounts of stearic acid on plasma lipoprotein fractions and fecal fatty acid excretion of men. *American Journal of Clinical Nutrition* **61**: 1120-1128.

Dudley, J.W. and R.H. Moll, 1969. Interpretation and use of estimates of heritability and genetic variance in plant breeding. *Crop Science* **9:** 257-262.

Eberhart, S.A. and W.A. Russell, 1966. Stability parameters for comparing varieties. *Crop Science* **6**:36-40.

Echekwu, C.A. and S.O. Alaba, 1995. Genetic effects of yield and its components in interspecific crosses of cotton. *Discovery and innovation* **7:** 395-399.

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

El-Adl, A.M. and P.A. Miller, 1971. Transgressive segregation and the nature of the gene action for yield in an intervarietal cross of upland cotton. *Crop Science* **11**:381-384.

El-Dahan, M.A.A., E.O. Leid, M. Lopez and J.C. Gutierrez, 2003. Combining ability of yield and yield components in upland cotton (*Gossypium hirsutum* L.) under drought stress conditions. Proceedings World cotton research conference pp 43-48. Cape Town, South Africa.

Emeetai-Areke, T.E., 1999. Genetic variability for yield and quality characteristics in South Africa cotton (*Gossypium hirsutum*) germplasm. PhD Thesis, Department of Plant Breeding, UOVS, Bloemfontein, South Africa.

Endrizzi, J.E., L. Turcotte and R.J. Kohel, 1985. Genetics, cytology and evolution of *Gossypium*. *Advances in Genetics* **23**: 271-375.

Falconer, D.S. 1989. Introduction to quantitative genetics (second edition) pp 438. Longman, New York, USA.

Falconer, D.S. and T.F.C. Mackay, 1996. Introduction to quantitative genetics (fourth edition) pp 464. Longman Group Limited, England.

FAO, 2000. FAO statistical Database**.**

Farooq, S. and F. Azam, 2002. Molecular markers in plant breeding: Concepts and characterization. *Pakistan Journal of Biological Sciences* **5**: 1135-1140.

Federici, M.T., D. Vaughan, N. Tomooka, A. Kaga, X.W. Wang and K. Doi, 2001. Analysis of Uruguayan weedy rice genetic diversity using AFLP molecular markers pp 8. Research article. Molecular Biology and Genetics. Universidad Catolica de Valparaiso, Chile.

Fensham, A.B., A.R. Barr, B.R. Cullis and D. Elham, 1998. A mixed model analysis of 10 years of oat evaluation data: use of agronomic information to explain genotype x environment interaction. *Euphytica* **99**: 43-56.

Finlay, K.W. and G.N. Wilkinson, 1963. The analysis of adaptation in a breeding program. *Australian Journal of Agriculture Research* **14**:742-754.

Flintham, J.E., W.J. Angus and M.D. Gale, 1997. Heterosis, overdominance for grain yield, and alpha-amylase activity in F_1-hybrids between near isogenic *Rhl* dwarf and tall wheats. *Journal Agriculture Science* **129**: 371-378.

Folch, J., M. Lees and S.G.H. Sloane, 1957. A simple method for the isolation and purification of total lipids from animal tissue. *Journal of Biological Chemistry* **226**: 497-509.

Fortucci, P., 2001. The contributions of cotton to economy and food security in developing countries. *http://www.ICAC.org/icac/meetings***.**

Freeman, G.H., 1973. Statistical methods for the analysis of genotype x environment interactions. *Heredity* **31**: 339-354.

Freeman, G.H., 1985. The analysis and interpretation of interactions. *Journal of Applied Statistics* **12**: 3-10.

Fryxell, P.A., 1968. The typification and application of Linnaean binomials in *Gossypium*. *Brittonia* **20**: 378-386.

Gad, A.M., M.A. El-RawaL, M.A. Bisher and A.A. El-Kishen, 1974. Studies on gene action in the interspecific cross of cotton. 1. Manifestation of gene effects. *Egyptian Journal Genetics Cytology* **3:** 117-124.

Gannaway, J.R., 1982. Breeding for high strength cotton. Proceedings Beltwide Cotton Conference pp 80-92**.** National Cotton Council, Memphis, TN.

Gauch, H.G., 1988. Model selection and validation for yield trials with interaction. *Biometrics* **44:** 705-715.

Gauch, H.G., 1990. Using interaction to improve yield estimates. In: M.S. Kang (Ed.), Genotypes x environment interaction and plant breeding pp 141-150. Department of Agronomy Louisiana State University, Baton rouge, Louisiana.

Gauch, H.G., 1993. Prediction, parsimony and noise. *American Scientist* **81:** 468-478.

Gauch, H.G. and R.W. Zobel, 1988. Predictive and postdictive success of statistical analyses of yield trials. *Theoretical and Applied Genetics* **76**: 1-10.

Gauch, H.G. and R.W. Zobel, 1996. AMMI analysis of yield trials. In: M.S. Kang and H.G. Gauch (Eds.), Genotype by environment interaction pp 85-122. CRC Press, Boca Raton, New York.

Gauch, H.G. and R.W. Zobel, 1997. Identifying mega-environments and targeting genotypes. *Crop Science* **37**: 311-326.

Geever, R.F., F. Katterman and J.E. Endrizzi, 1989. DNA hybridization analysis of a *Gossypium* allotetraploid and two closely related diploid species. *Theoretical and Applied Genetics* **77**: 234-244.

Gencer, O. and R.S. Guvercin, 2003. Heritability of earliness criteria in relation to yield in cotton (*Gossypium hirsutum* L.) in the Harran Plain of the southeastern Anatolia project (GAP) area. Proceedings World cotton research conference pp 49-56. Cape Town, South Africa.

Gepts, P., 1993. The use of molecular and biochemical markers in crop evolution studies. In: M.K. Hecht (Ed.), Evolutionary biology pp 51-94. Plenum Press, New York.

Glaz, B. and J.L. Dean, 1988. Statistical error rates and their implications in sugar cane clone trials. *Agronomy Journal* **80:** 560-562.

Government of Tanzania (GoT), 1999a. Review of impact of taxes, levies and fees on the agricultural Sector: Final report. Ministry of Agriculture and cooperatives, Dar es salaam, Tanzania.

Government of Tanzania (GoT), 1999b. 1998/99 Marketing review of cotton. Ministry of Agriculture and cooperatives, Dar es salaam, Tanzania.

Grafius, J.E., 1957. Components of yield in oats. A geometric interpretation. *Agronomy Journal* **49:** 419-423.

Green, A.G. and D.R Marshall, 1981. Oil quantity and quality in lint seed *Australian Journal of Agriculture Research* **32**: 599-607.

Gregory, S.R., E. Hernandez and B.R. Savoy, 1990. Cottonseed processing. In: W. C. Smith and J. T. Cothren (Eds.), Cotton: Origin, History, Technology and Production pp 793-819. John Wiley & sons, New York.

Griffing, B., 1956. Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Science* **9:** 463-493.

Gulati, A.M. and A. J. Turner, 1928. A note on the early history of cotton. Technical Laboratory Bulletin No 17. Indian Central Cott. Cttee.

Gunstone, F.D., 1967. An introduction to the chemistry and biochemistry of fatty acids and their glycerides pp 1-209. Chapman and Hall, London.

Gurr, M.I. and J.L. Harwood, 1991. Lipid Biochemistry. An introduction (Fourth edition) pp 1-406. Chapman and Hall, London.

Guthridge, K.M., M.P. Dupal, R. Kolliker, E.S. Jones, K.F. Smith and J.W. Forster, 2001. AFLP analysis of genetic diversity within and between populations of perennial ryegrass (*Lolium perenne* L.). *Euphytica* **122**: 191-201.

Gutierrez, O.A., S. Basu, S. Saha, J.N. Jenkins, D.B. Shoemaker, C.L. Cheatham and J.C. McCarty, 2002. Genetic distance among selected cotton genotypes and its relationship with F_2 performance. *Crop Science* **42**: 1841-1847.

Hallauer, A.R., W.A. Russell and K.R. Lamkey, 1988. Corn breeding. In: G.F. Sprague and J.W Dudley (Eds.), Corn and Corn improvement pp. 463-564. ASA, Madison, Wisconsin.

Hartl, L.D. and A.G. Clark, 1989. Principles of population genetics (second edition) pp 1-682. Snauer Associates, Inc. Publishers, Sunderland.

Harwood, J.L., 1996. Recent advances in the biosynthesis of plant fatty acids. *Biochimica et Biophisca Acta* **1301**: 7-56

Harwood, J.L., 1997. Plant lipid metabolism. In: J.B. Harborne (Ed.), Plant biochemistry pp 237-272. Academic press.

Hatje, G., 1989. World importance of oil crops and their products. In: G. Robbelen, R.K. Downey and A. Ashir (Eds.), Oil Crops of the world pp 1-21. New York: McGraw-Hill.

Hau, B., 1997. The Tanzania Cotton Consultancy Mission report. A study commissioned by the FSR Lake Zone Programme pp 1-60. CIRAD Montpelier.

Hayman, B.I., 1954. The theory and analysis of diallel crosses. *Genetics* **45:** 155-172**.**

Heinrich, G.M., C.A. Francis and J.D. Eastin, 1983. Stability of grain sorghum yield components across diverse environments. *Crop Science* **23:** 209-212.

Herselman, L., 2003. Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. *Euphytica* **133**: 319-327.

Hill, J., 1975. Genotype x environment interactions a challenge for Plant Breeding. *Journal Agricultural Science* **85**: 477-493.

Hintze, J.L., 2000. Number Cruncher Statistical System (NCSS). Statistical System for Windows. Kaysville, Utah.

Hohls, T., 1995. Analysis of genotype x environment interactions. *South Africa Journal of Science* **91**: 121-124.

Hussain, S.S., F.M. Azhar and M. Sidiq, 1998. Genetic correlation, path coefficient and heritability estimates of some important plant traits of upland cotton. *Sarhad Journal of Agriculture* **14:** 57-59.

Hutchinson, J.B., 1959. The application of genetics to cotton improvement. Empire Cotton Growing Corporation pp 179-23**5***.* Cambridge University Press London.

Hutchinson, J.B., R.A. Silow and S.G. Stephens, 1947. The evolution of *Gossypium* pp 138. Oxford University Press, London.

Ibragimov, P., 1989. Genetic correlations between characters. *Khlopok* **4:** 45.

International Cotton Advisory Committee (ICAC), 2001. Production and Trade policies affecting the cotton Industry Washington D.C.

Iqbal, M.J., N. Aziz, N.A. Saeed, Y. Zafar and K.A. Malik, 1997. Genetic diversity evaluation of some elite cotton varieties by RAPD analysis *Theoretical and Applied Genetics* **94:** 139-144**.**

Iqbal, M.J., O.U.K. Reddy, K.M. Ez-Zak and A.E. Pepper, 2001. A genetic bottleneck in the evolution under domestication of upland cotton. (*Gossypium hirsutum* L.) examined using fingerprinting. *Theoretical and Applied Genetics* **103:** 547-554.

Jackson, P., M. Robertson, M. Cooper and G.L. Hammer, 1998. The role of physiological understanding in Plant Breeding: From a breeding perspective. *Field Crops Research* **49**: 11-37.

Jenkins, J.N., 1976. Boll weevil resistant cottons. Boll weevil suppression, management and elimination technology **71: 45-49.** USDA/ARS.

Jensen, S.D and A.J. Cavalieri, 1983. Drought tolerant in US maize. *Agriculture and Water Management* **7:** 223-236.

Jixiang, W., Z. Jun, J.I. Daofan and W.L. Partott, 1996. Analysis of the interactions between genetic effects and environment for yield traits in upland cotton. *Hereditan Beijing* **17**: 1-4.

Jones, J.E., 1972. Effect of morphological characters on cotton insect and pathogens. In: J.M. Brown (Ed.), Proceedings Beltwide Cotton Conference pp 88-92. National Cotton Council, Memphis, TN.

Jones, J.B. and P. Kapingu, 1982. Cotton Breeding report. In: B. Nyambo (Ed.), The Golden Jubilee Ukiriguru 50 years of Research 1932-1982 pp 17-20. Tanzania Agricultural Research Organization.

Jones, L.A. and J.H. Kersey, 2002. Cottonseed article. National cottonseed products Association (NCPA) (USA) *http://www. cottonseed.com/publication/csobro.Asp*.

Kang, M.S., 1993. Simultaneous selection for yield and stability in crop performance trials: Consequences for grower. *Agronomy Journal* **85:** 754-757.

Kang, M.S., 1996. Using genotype by environment interaction for crop cultivar development. *Advances in Agronomy* **62:** 199-252.

Kang, M.S., D.P. Gorman and H.N. Pham, 1991. Application of a stability statistic to international maize yield trials. *Theoretical and Applied Genetics* **81:** 162-165.

Kang, M.S., J.D. Miller and L.L. Darrah, 1987. A note on relationship between stability variance and ecovalence**. *Journal of Heredity* **78:** 107.

Kapoor, C.J., 1994. Studies of quantitative characters in upland cotton (*Gossypium hirsutum* L). Challenging the Future. Proceedings World Cotton Research Conference pp 297-298. Brisbane Australia.

Karp, A, K. Edwards, M. Bruford, B. Vosman, M. Morgante, O. Seberg, A. Kremer, P. Boursot, P. Arctander, D. Tautz and G. Hewitt, 1997. Opportunities and challenges. *Nature Biotechnology* **15**: 625-628.

Kaufman, L. and P.J. Rousseeuw, 1990. Finding groups in data: An introduction to cluster analysis pp 342. A Wiley-Interscience Publication, New York.

Kausar, M.N.S., R.A. Kainth, U.H. Rahman, I. Ullah and W.S. Khan, 1998. Estimation of hybrid vigour levels for economic hybrid cotton seed production. World Cotton Research Conference Abstracts pp 70. Athens, Greece.

Kerr, T., 1966. Yield components in cotton and their interrelations with fiber quality. Proceeding 18th Cotton Improvement Conference pp 276-283. Memphis, TN.

Khan, M., F.M. Azhar and I.A. Khan, 2003. Inheritance pattern of seed cotton yield and its yield components under Nitrogen stress condition. *International Journal of Agriculture and Biology* **4**:589-593.

Kim, H.J. and B.A. Triplett, 2001. Cotton fibre growth in planta and *in vitro*. Models for plant cell elongation and cell wall biogenesis. *Plant Physiology* **127**: 1361-1366.

Kittock, D.L. and L.l.H. Pinkas, 1975. Relationship of seed/boll to other yield components and fibre quality of Pima cotton. *Crop Science* **15**: 316-319.

Kohel, R.J. 1974. Influence of certain morphological characters on yield. *Cotton Growing Review* **51:** 281-292.

Kohel, R.J. and C.R. Benedict, 1987. Growth Analysis of Cottons with different Maturities. *Agronomy Journal* **79:** 31-34.

Krauss, S.L., 1999. Complete exclusion of nonsires in analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP). *Molecular Ecology* **8**: 217-226.

Krawezyk, T., 2001. Fat in dietary guidelines around the world. *INFORM* **12**: 132-138.

Kumar, L.S., 1999. DNA markers in plant improvement. An overview. *Biotechnology Advances* **17:** 143-182.

Lamkey, J.R. and J.W. Edwards, 1999. Quantitative genetics of heterosis. In: J.G. Coors and S. Pandey (Eds.), The Genetics and Exploitation of Heterosis pp 31-48. American Society of Agronomy and Crop Science, USA.

Lancon, J., E. Goze, G. Gawrysiak, B. Hau, B. Bachelier, J.L. Chanselme, D. Dessauw, C. Klassou, E.N. Guessan, T.B. Nguyen and E. Ousmane, 1993. Multi-site trial of a diallel with four elite parents bred within the cotton research African network. *Coton et Fibres Tropicales* **48**: 253-282.

Lawhon, J.T., C.M. Carter and K.F. Mattli, 1977. Evaluation of the food use potential of sixteen varieties of cottonseed. *Journal of the American Oil Chemists Society* **54**:75-80.

Lee, J.A., P.A. Miller and J.O. Rawlings, 1967. Interaction of combining ability effects with environments in Diallel crosses of upland cotton (*Gossypium hirsutum* L.) *Crop Science* **7:** 477-481.

Lewis, H., 1962. Catastrophic selection as a factor in speciation. *Evolution* **16**: 257-271.

Lin, C.S. and M.R. Binns, 1988. A superiority measure of cultivar performance for cultivar x location data. *Canadian Journal Plant Science* **68**: 193-198.

Lin, C.S., M.R. Binns and L.P. Lefkovitch, 1986. Stability analysis: Where do we stand? *Crop Scence* **26:** 894-900.

Liu, L., W. Guo, X. Zhu and T. Zhang, 2003. Inheritance and fine mapping of fertility restoration for cytoplasmic male sterility in *Gossypium hirsutum*. *Theoretical and Applied Genetics* **106**: 461-469.

Liu, Q., S. Singh and A. Green, 2002. High Oleic and High Stearic cottonseed Oils: Nutritionally Improved Cooking Oils Developed using gene silencing. *Journal of the American College of Nutrition* **21**: 205-211.

Liu, S., R.G. Cantrell and J.C. McCarty, 2000. Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. *Crop Science* **40**:1459-1469.

Liu, S., S. Saha and D. Stelly, 2001 Chromosomal assignment of microsatellite loci in cotton. *Journal of Heredity* **91**: 326-32.

Liu, Z. and G.R. Turner, 1993. Comparison of allozyme, RFLP and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. *Theoretical Applied Genetics* **87**: 97-105.

Lu, H. and G.O. Myers, 1999. DNA variation in ten influential upland cotton varieties by RAPDs. Proceedings Beltwide Cotton Conferences pp. 484. National Cotton Council, Memphis, TN.

Lu, H. and G.O. Myers, 2002. Genetic relationships and descrimination of the influential upland cotton varieties using RAPD markers. *Theoretical and Applied Genetics* **105**: 325-331.

Lübberstedt, T, A.E. Melchinger, C. Dußle, M. Vuylsteke and M. Kuiper, 1998. Relationship among early European maize inbreds: IV. Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD and pedigree data. *Crop Science* **40**: 783-791.

Lubbock, T.X. and N.C. Raleigh, 1996. Relationship between cotton seedling cold tolerance and physical and chemical properties. Proceedings Beltwide cotton conference pp 1170-1171. National cotton council, Memphis, TN.

Luckett, D.J., 1989. Diallel analysis of yield components, fibre quality and bacterial blight resistance using spaced plants of cotton. *Euphytica* **44:** 11-20.

Lukonge, E. and O.S. Ramadhani, 1999. Review of Plant Breeding and agrobiodiversity in the Lake Zone. Breeding and Agrobiodiversity workshop proceedings pp 6-11. Agricultural Research and Development Ukiriguru, Mwanza, Tanzania.

Lukonge, E., O.S. Ramadhani, T. Kibani, E. Temu and F. Joram, 1999. Evaluation of cotton characters. Field Note pp 1-43. ARI Ukiriguru.

Machado, C.F., J.B. Santos, G.H. Nunes and M.A.P. Ramalho, 2002. Choice of common bean parents based on combining ability estimates. *Genetics and Molecular Biology* **25**:179-183.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209-220.

Market, C.L. and F. Moller, 1959. Multiple forms of enzymes: tissue, ontogenetic and species specific pattern. *Proceedings of National Academy Science* **45:** 753-763.

Mather, K. and J.I. Jinks, 1977. Introduction to Biometrical Genetics pp 231. Chapman and Hall, London, UK.

Matzinger, D.F., 1963. Experimental estimates of genetic parameters and their applications in self fertilizing plants. In: W.D. Hanson and H.F. Robinson (Eds.), Statistical genetics and plant breeding no. 982. NAS-NRC.

Maughan, P.J., M.A. Saghai Maroof, G.R. Buss and G.M. Huests, 1996. Amplified Fragment Length Plymorphisms (AFLPs) in soybean; species diversity, inheritance and near isogenic lines analysis. *Theoretical and Applied Genetics* **93**: 392-401.

May, O.L., 2002. Quality improvement of upland cotton. *Journal of Crop Production* **5**: 6-9.

McLeod, I.G., 2001. The effect of water logging and ion interaction on the development of premature senescence in irrigated cotton. PhD Dissertation. University of New England, Armidale, Australia.

Mensink, R.P. and M.B. Katan, 1992. Effect of dietary trans fatty acids on high-density and low-density lipoprotein cholesterol levels in healthy subjects. *New England Journal of Medicine* **323**: 439-445.

Meredith W.R.Jr., 1984. Quantitative inheritance. In: R.J. Kohel and C.F. Lewis (Eds.), Cotton. Agronomic Monographs pp 131-150. ASA, CSSA, SSSA, South Segoe, Madison, WI.

Meredith, W.R.Jr., 1995a. Use of molecular markers in cotton breeding. In: G.A. Constable and N.W. Forrester (Eds.), Challenging the future. Proceedings World Cotton Research Conference pp. 303-308. Melbourne, Australia.

Meredith, W.R.Jr., 1995b. Strength and limitations of conventional and transgenic breeding. In: J.M. Brown (Ed.), Proceeding Beltwide Cotton Conference pp 166-168. National Cottton Council, Memphis, TN.

Meredith, W.R.Jr., 1998. Continued progress in breeding for yield in the USA. Proceedings World Cotton Research Conference pp 97-101. Athens, Greece.

Meredith, W.R.Jr. and J.S. Brown, 1998. Heterosis and combining ability of cottons originating from different regions of the United States. J*ournal of Cotton Science* **2**: 77-84.

Meredith, W.R.Jr. and R.R. Bridge, 1984. Genetic contributions to yield changes in upland cotton. In: W.R. Fehr (Ed.), Genetic contributions to yield changes in five major plants pp 75-87. CSSA Spec Publ 7. Madison, WI.

Meredith, W.R.Jr., J.J. Heitort, W.T. Pettigrew and S.T. Rayburn, 1997. Comparison of obsolete and modern cotton cultivars at two nitrogen levels. *Crop Science* **37:** 1453-1457.

Messmer, M.M., A.E. Melchinger, R.G. Herrmann and J. Boppenmaier, 1993. Relationship among early European maize inbreds: II. Comparison of pedigree and RFLP data. *Crop Science* **33**: 944-950.

Meyer, V.G., 1975. Male sterility from *Gossypium harkinessi. Journal of Heredity* **66:** 23-27.

Michael, W.K., 2001. Basic chemistry of biomolecules. *mking@medicine.indstate.edu*.

Miller, P.A and A. Marani, 1963. Heterosis and combining ability in diallel crosses of upland cotton (*Gossypium hirsutum* L.*). Crop Science* **3**: 441-444.

Möllers, C. and A. Schierholt, 2002. Genetic variation of palimitate and oil content in a winter oilseed rape doubled haploid population segregating for oleate content. *Crop Science* **42**: 379-384.

Mueller, U.G. and L. Wolfenbarger, 1999. AFLP genotyping and fingerprinting. *TREE* **14**: 389-394.

Multani, D.S. and B.R. Lyon, 1995. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. *Genome* **38**: 1005-1008.

Munro, J.M., 1987. Cotton (Second edition), Tropical Agriculture series pp 161. Longman Scientific and technical, John Willey 7 Sons, Inc, New York.

Mursal, I.E.J. and J.E. Endrizzi, 1976. A reexamination of the diploid like behavior of polyploid cotton. *Theoretical and Applied Genetics* **47:** 171-178.

Murtaza, N., M. Kitaoka and G.M. Ali, 2005. Genetic differentiation of cotton cultivars by polyacrylamide gel electrophoresis**.** *Journal Central Europe Agriculture* **6**: 69-76.

Murthi, K.S. and K.T. Achaya, 1975. Cottonseed: Chemistry and Technology pp 285-306. New Delhi. Publication and information Directorate.

Myers, G.O., 2004. Estimation of potential breeding value and genotype stability of cotton strains and varieties. *http/www.cottoninc.com*.

Myers, G.O. and F. Bordelon, 1995. Inheritance of yield components using variety trial data. Proceedings Beltwide Cotton Conference pp 510-513. San Antonio, TX, USA.

Natalija, B., 2001. Major plant genetic marker systems. Lithuanian University of Agriculture. Lithuania. *http://www.agsci.kvl.dk/breed/SBA/NordicBaltic/abstract2.htm*

National Soil Service (NSS), 1991. Soils of Mikindani Estate (Mtwara region) and their suitability for cashew, mango, lime, hybrid sisal and teak cultivation, pp 65. Ministry of Agriculture, Livestock Development and Cooperatives. ARI Mlingano, Tanga, Tanzania.

Niles, G.A. and C.V. Feaster, 1984. Breeding. In: R.J. Kohel and C.F. Lewis (Eds.), Cotton Agronomy Monograph pp 201-231. ASA, CSSA, SSSA, South Segoe, Madison, WI.

Noakes, M. and P.M. Clifton, 1998. Oil blends containing partial hydrogenation on interesterified fats: differential effects on plasma lipids. *American Journal of Clinical Nutrition* **98**: 242-247.

Nyambo, B., 1982. Ukiriguru Research Center in relation to WCGAs'. The golden Jubilee Ukiriguru 50 years of research 1932-1982. Tanzania Agricultural Research Organization.

Padley, F.B., F.D. Gunstone and J.L. Harwood, 1994. Occurrence and characteristics of oils and fats. In: F.D. Gunstone, J.L. Harwood and F.D. Padley (Eds.), The lipid handbook (second edition) pp 47-223. Chapman and Hall, London.

Paterson, A.H, C.L. Brubaker and J.F. Wendel, 1999. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. *Genome* **42**: 184-203.

Paterson, Y.S., M. Menz and C.X. Jiang, 2003. QTL analysis of genotype x environment interactions affecting cotton fiber. *Theoretical and Applied Genetics* **106**: 384-396.

Patil, N.B. and M. Singh, 1994. Development of medium-staple, high-strength cotton suitable for rotor spinning systems. In: G.A. Constable and N.W. Forrester (Eds.), Proceedings World Cotton Research Conference pp 264-267. Brisbane Australia.

Pejic, L., P. Ajmone-Marsan, M. Morgante, V. Kozumplic, P. Castiglion, G. Taramino and M. Motto, 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. *Theoretical and Applied Genetics* **97**: 1248-1255.

Perkin-Elmer (PE) Applied Biosystems, 2002. *http://www.perkin.elmer.com*

Pillay, A.A. and G.O. Myers, 1999. Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. *Crop Science* **39:** 1881-1886.

Pinthus, M.J., 1973. Estimate of genotypic value: A proposed method. *Euphytica* **22:** 121-123.

Pleines, S. and W. Friedt, 1988. Breeding for improved C18 fatty acid composition in rapeseed (*Brassica napus* L.). *Fat Science Technology* **90**: 167-171.

Poehlman, J.M., 1987. Breeding field crops (second edition) pp 724. Avi Publishing Company Inc Westport Connectient.

Poisson, J., O.M. Royo, L. Erazzu, S. Ibalo and A. Montenegro, 2003. First Okra leaf cotton cultivar bred and released in Argentina**.** Proceedings World Cotton Research Conference pp 259-262. Cape Town, South Africa.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski, 1996. The comparison of RFLP, RAPD, AFLP and SSR markers for germplasm analysis. *Molecular Breeding* **2:**225-238.

Prentice, A., 1972. Cotton: With special reference to Africa pp 282. Longman, London.

Purchase, J.L., 1997. Parametric analysis to describe G x E interaction and yield stability in winter wheat. PhD Thesis pp 4-83. Department of Agronomy, Faculty of Agriculture, University of Orange Free State, Bloemfontein, South Africa.

Pursegloves, J. W., 1968. Tropical Crops: Dicotyledons pp 73. Longman, London.

Rafalski, J.A., J.M. Vogel, M. Morgante, W. Powell, C. Andre and S.V. Tingey, 1996. Generating and using DNA markers in plants. In: B. Birren and E. Rai (Eds.), Non-Mammalian Genomic Analysis: A practical Guide pp 75-134. Academic Press, London.

Ramadhani, O.S. and E. Lukonge, 1999. Cotton perspectives in breeding and agrobiodiversity: *Breeding and Agrobiodiversity workshop proceedings* pp 36-46. Agricultural Research and Development Ukiriguru Mwanza Tanzania.

Ramadhani, O.S., R. Kileo, E. Temu and R. Msabaha, 1998. Survey on cotton production constraints in Meatu District. Meatu district council and ARID Ukiriguru.

Rana, M.K. and K.V. Bhat, 2004. A comparison of AFLP and RAPD markers for genetic diversity and cultivar identification in cotton. *Journal of Plant Biochemistry and Biotechnology* **13**: 19-24.

Raun, W.R., H.J. Barreto and R.L Westerman, 1993. Use of stability for long term soil fertility experiments. *Agronomy Journal* **85**: 159-167.

Ray, J.D. and B.T. Carr, 1985. Empirical modelling of soybean oil hydrogenation. *Journal of the American Oil Chemists Society* **62**: 1218-1222.

Reddy, O.U.K., T.D. Brooks, K.M. El-Zik and A.E. Pepper, 2000. Development and use of PCR-based technologies for cotton mapping. Proceedings Beltwide Cotton Conference pp. 483. National Cotton Council, Memphis, TN.

Reddy, R.S.C., 1974. Effects of three leaf shape genotypes *Gossypium hirsutum* L. and row types on plant microclimate, boll weevil survival, boll rot and important agronomic characters. PhD dissertation. Louisiana State University, Baton Rouge, LA.

Reid, P.E., 1995. Performance of F_1 and F_2 hybrids between Australia and USA commercial cotton cultivars. In: G.A. Constable and N.W. Forester (Eds.), Challenging the future. Proceeding World Cotton Research Conference pp 346-349. Brisbane, Australia.

Rivera, R, K.J. Edwards, J.H.A. Baker, G.M. Arnold, G. Ayad and T. Hodgkin, 1999. Isolation and characterization of polymorphic microsatellites in *Cocos nucifera* L. *Genome* **42**: 668-675.

Robinson J.P. and S.A Harris, 1999. Amplified fragment length polymorphism and microsatellites: a phenetic perspective. *http://webdoc.sub.gwdg.de/ebook/y/1999/whichmark/m12/Chap12.htm*.

Rochester, I.J and G.A. Constable, 2003. Leaf nutrient concentrations in cotton cultivars grown on slightly sodic soil. World Cotton Research Conference pp 681-688. Cape Town, South Africa.

Rohlf, F.J., 1993. NTSYS-pc. Numerical taxonomical and multivariate analysis system. Exeter Software, Setauket, New York, USA.

Roldan-Ruiz, I, F.A. van Eeuwijk, T.J. Gilliland, P. Dubreuil, C. Dillmann, J. Lallemand, M. De Loose and C.P. Baril, 2001. A comparative study of molecular and morphological methods of describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. *Theoretical and Applied Genetics* **103**: 1138-1150.

Roldan-Ruiz, I, T.J. Gilliland, E. Calsyn, T.J. Gillilad, R. Coll, M.J.T. Van Eijk and M. De Loose, 2000. Estimating genetic conformity between related ryegrass (*Lolium*) varities. II. AFLP characterization. *Molecular Breeding* **6**: 593-602.

Romagosa, I. and P.N. Fox, 1993. Genotype x environmental interaction and adaptation. In: M.D Hayward, N.O. Bosemark and I. Romagosa (Eds.), Plant breeding: Principles and Prospects pp 373-390. Chapman and Hall, London.

Rossetto, M., A. McLauchlan, A. Harris, F.C.L. Henry, R.J. Baverstock, P.R. Lee, L.S. Maguire and K.J. Edwards, 1999. Abundance and polymorphism of microsatellite markers in the tea tree (*Melaleuca attemifolia, Myrtaceae). Theoretical and Applied Genetics* **98:** 1091-1098.

Royo, O.M., J. Poison and A. Montenegro, 2003. Morphological traits characterisation and some yield and fibre quality components evaluation of the Argentinian cotton genetic resources**.** Proceedings World Cotton Research Conference pp 267-270. Cape Town, South Africa.

Rungis, D., D. Llewellyn, L. Dennis and B.R. Lyon, 2000. Characterization and mapping of the *XCM* resistance locus in Australian cotton cultivars by the use of molecular marker techniques. Proceedings 10th Australian Cotton Conference pp 5. Brisbane, Australia.

Russell, J.R., J.D. Fuller, M. Macaulay, B.G. Hatz, A. Jahoor, W. Powell and R. Waugh, 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical and Applied Genetics* **95**: 714-722.

Salunkhe, D.K., J.K. Chavan, R.N. Adsule and S.S. Kaman, 1992. World Oilseeds: Chemistry, Technology and Utilization pp 554. Van Nostrand Reinhold, New York.

Sambamurthy, J.S.V., D.M. Reddy and K.H.G. Reddy, 1995. Studies on the nature of genetic divergence in upland cotton (*Gossypium hirsutum* L.). *Annals of Agricultural Research* **16:** 307-310.

Santhanam, V. and J. B. Hutchinson, 1974. Cotton. In: J.B. Hutchinson (Ed.), Evolutionally studies in World Crops pp 89-100. Cambridge University Press, London.

Santos, J.B., J. Nienhuis, P. Scroch, J. Tivang and M.K. Slocum, 1994. Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. *Theoretical and Applied Genetics* **87**: 909-915.

Saranga, Y., M. Menz, C.X. Jiang, R.J. Wright, D. Yakir and A.H. Paterson, 2001. Genomic dissection of genotype x environment interactions conferring adaptation of cotton to arid conditions, *http://www.genome.org/cgi/doi/10.110/gr.157201*.

Sayal, O.U. and M.Z. Sulemani, 1996. Comparison of gene action controlling the qualitative traits in some early maturing cultivars of American cotton (*Gossypium hirsutum* L). *Sarhad Journal of Agriculture* **12:** 137-145.

Scarth, R. and P.B.E. McVetty, 1999. Designer oil canola. A review of new food grade Brassica oils with focus on high oleic, low linolenic types. *[Rachael_Scarth@umanitoba*] Department of Plant Sciences, University of Manitoba, Winnipeg, Canada.

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

Shepherd, A.W. and S. Farolfi, 1999. Export crop liberalization in Africa: A review pp 100. Rome: Food and Agriculture Organization.

http://www.fao.org/waicent/faoinfo/agricultul/ags/AGS/ecl.pdf

Shukla, G.K., 1972. Some statistical aspects of partitioning genotype-environment components of variability. *Heredity* **29**: 237-245.

Siddiqui, M.A., 1997. A study of variability and heritability of some quantitative character in *Hirsutum* cotton. *Journal of Maharashtra Agricultural Universities* **21** 256-258.

Sikka, S.M. and A.B. Joshi, 1960. Problems of cotton improvement. Cotton. *India Monograph* 137-335.

Silvertooth, J.C., E.J. Norton and E.R. Norton, 2001. Summary of nitrogen management experiments in irrigated cotton. *http://ag.arizon.edu/pubs/crops/az1224/* The University of Arizon College of Agriculture and life sciences.

Simmonds, N.W., 1984. Principles of crop improvement pp 339. Longman, London.

Singh, H., 2002. Genetic architecture of yield and its associated traits in bread wheat. PhD Thesis, Rajasthan Agriculture University. Bikaner, Rajasthan, India.

Singh, D.P. and R.B. Singh, 1980. Genetics of ginning characters in upland cotton Indian *Journal of Agriculture Science* **50**: 537-540.

Singh, M., T.H. Singh, G.S. Chahal and L.S. Randhawa, 1990. Genetic analysis of lint yield and its components in cotton. *Crop Improvement* **17:** 64-67.

Slover, H.T. and E. Lanza, 1979. Quantitative analysis of food fatty acid by capillary gas chromatography. *Journal American Oil Chemistry Society* **56**: 933-943.

Smith, C.E.Jr, 1968. Plant remains. In: D.S. Byers (Ed.), Prehistory of the Tehuacan Valley pp 220-255. Austin, Texas.

Smith, C.W. and K. S. Rathore, 2001. Molecular genetic improvement of cottonseed for feed and food. Texas food and fibres commission Annual report. Texas A & M University/ Soil and Crop Sciences Dept.

Smith, C.W. and S.G. Coyle, 1999. Association of fiber quality parameters and within boll yield components in upland cotton. *Crop Science* **37:** 1775-1779.

Smith, J.S.C. and O.S. Smith, 1992. Fingerprinting crop varieties. *Advances in Agronomy* **47**: 85-129.

Smith, J.S.C., S. Kresovich, M.S. Hopkins, S.E. Mitchell, R.E. Dean, W.L. Woodman, M. Lee and K. Porter, 2000. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Science* **40**: 226-232.

Smith, O.S., J.S.C. Smith, S.L. Brown, R.A. Tenborg and S.J. Wall, 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F_1 grain yield heterosis and RFLPs. *Theoretical and Applied Genetics* **80**: 833-840.

Sokal, R.R. and C.D. Michener, 1958. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin **28**: 1409-1438.

Spielmeyer, W., A.G. Green, D. Bittisnich, N. Mendham and E.S. Lagudah, 1998. Identification of quantitative trait loci contributing to Fusarium wilt resistance on AFLP linkage map of flax. *Theoretical and Applied Genetics* **97:** 633-641.

Sprague, G.F. and L.A. Tatum, 1942. General versus specific combining ability in a single cross of corn. *Journal American Society Agronomy* **34:** 923-932.

Stamler, J. and R. Shekelle, 1988. Dietary cholesterol and human coronary heart disease. *Archieves of Pathology and Laboratory Medicine* **112**: 1032-1040.

Stanton, M.A., J.M. Stewart, A.E. Percival and J.F. Wendel, 1994. Morphological diversity and relationships in the A-genome cottons, *G. arboreum* and *G. herbaceum. Crop Science* **34**: 519-527.

Steel, R.G. and J.H. Torrie, 1980. Principles and procedures of statistics pp 633. McGraw-Hill, New York.

Stewart, J.McD., 1992. Germplasm resources and enhancement strategies for disease resistance. Proceeding*s* Beltwide Cotton Conferences pp 1323-1325. National cotton Council, Memphis, TN.

Stewart, J.McD., 1995. Potential for crop improvement with exotic germplasm and genetic engineering. In: G.A. Constable and N.W. Forrester (Eds.), Challenging the future. Proceedings World Cotton Research Conference pp. 313-327. CSIRO, Melbourne, Australia.

Subrahmanyam, Y., N.P. Mehta, J.C. Patel, M.C. Bhalod, J.K. Iyer and V.G. Munshi, 1989. Genetic studies of fibre fineness in Asiatic cotton. *Journal of the Indian Society for Cotton Improvement* **14:** 153-158.

Sutirtha, B., J.N. Jenkins, D.B. Shoemaker, C. Cheatham, J. McCarty, A.G. Osman and S. Saha, 2001. Use of SSR markers for analysis of genetic relationship among improved cotton cultivars and its association with F_2 hybrid performance. Plant and Animal Genome 1X Conference. San Diego, CA.

Swanepoel, E., 1999. Identification of genetic distances and heterotic groups of inbred maize (*Zea mays*) lines using DNA fingerprinting. MSc Thesis, Department of Plant Breeding, University of Orange Free State, Bloemfontein, South Africa.

Tang, B., J.N. Jenkins and J.C. McCarty, 1992. Genotypic stability of cotton varieties, resistant germplasm and their F_2 hybrids. In: D.J. Herber (Ed.), Beltwide Cotton Proceedings Research Conference pp 583-587. National Cotton Council, Nashville, TN.

Tang, B., J.N. Jenkins, J.C. McCarty and C.E. Watson, 1993a. F_2 hybrids of host plant germplasm and cotton cultivars: Heterosis and combining ability for lint and yield components. *Crop Science* **33:**700-705**.**

Tang, B., J.N. Jenkins, J.C. McCarty and C.E. Watson, 1993b. F_2 hybrids of host plant germplasm and cotton cultivars: Heterosis and combining ability for fiber qualities. *Crop Science* **33**: 706-710.

Tang, B., J.N. Jenkins, J.C. McCarty, C.E. Watson and R.G. Creech, 1996. Evaluation of genetic variances, heritabilities, and correlations for yield and fiber traits among cotton F_2 hybrid populations. *Euphytica.* **91**: 315-322.

Tanksley, S.D, 1983. Molecular markers in plant breeding. *Plant Molecular Biology Reporter* **1:** 3-5.

Tanksley, S.J., N.D. Young, A.H. Paterson and M.W. Bonierbale, 1989. RFLP mapping in plant breeding: New tool for an old science. *Bio/Technology* **7**: 257-264.

Tanzania Cotton Lint and Seed Board (TCL and SB), 2001. Varieties and Characteristics: *htt://www.tancotton.co.tz/research.html*.

Tanzania Cotton Lint and Seed Board (TCL and SB), 2002. Cotton sector development strategy. *Final draft*. Dar es salaam, Tanzania.

Tatinen, V., R.G. Cantrell and D.D. Davis, 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. *Crop Science* **36**: 186-192.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acid Research* **17**: 6463-6471.

Tekouano, A., R. Ortiz and K.P. Baiyeri, 2002. Phenotypic and genotypic correlation in Musa populations in Nigeria. *Africa Crop Science Journal* **10**: 121-132.

Thaxton, P.M. and K.M. El-Zik, 1994. Development of MAR cotton germplasm with morphological mutant traits**.** In: G.A. Constable and N.W. Forrester (Eds.), Proceedings of the World Cotton Research Conference pp 244-250. Brisbane Australia*.*

Thaxton, P.M., K.M. El-Zik and L.S. Bird, 1985. Influence of glabrous and hairy plants, leaf and bract types near-isogenic cotton lines on lint yield, earliness and fiber quality. In: J.M. Brown (Ed.), Proceedings Beltwide Cotton Production Conference pp 81-84. National Cotton Council, Memphis, TN.

Thomas, C.M., P. Vos, M. Zabeau, D.A. Jones, K.A Norcott, B.P. Chadwick and J.D.G. Jones, 1995. Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato Cf-9 gene for resistance to *Cladosporium fulvum*. *Plant Journal* **8**: 785-794.

Thomson, N.J. and D.J. Luckett, 1988. Heterosis and combining ability effects on cotton. 1 Combining ability. *Australian Journal of Agricultural Research* **39**:973-90.

Tingey, S.V. and J.P. del Tufo, 1993. Genetic analysis with random amplified polymorphic DNA markers. *Plant Physiol*ogy **101:** 349-352.

Tohme, J., D.O. Gonzales, S. Beebe and M.C. Duque, 1996. AFLP analysis of gene pool of a wild bean collection. *Crop Science* **36**: 1375-1384.

Turner, H.R., H.H.Jr. Ramey and S.Jr. Worley, 1976. Relationship of yield seed quality and fiber properties in upland cotton. *Crop Science* **16**: 578-580.

Unruh, B.L. and J.C. Silvertooth, 1997. Planting and irrigation timing effects on the yield of upland and pima cotton. *Journal Production Agriculture* **10:** 74-79**.**

UPOV, 1991. International convention for the protection of new varieties of plant http://www.upov.int

Van Eeuwijk, F.A., 1996. Between and beyond additivity and non-additivity: The statistical modeling of genotype by environment interaction in Plant Breeding. PhD dissertation, Wageningen Agricultural University, Wageningen, the Netherlands.

Van Esbroeck, G. and D.T. Bowman, 1998. Cotton germplasm diversity and its importance to cultivar development. *Journal Cotton Science* **2**: 121-129.

Van Esbroeck, G.A., D.T. Bowman, D.S. Calhoun and O.L. May, 1999. Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars. *Crop Science* **39**: 323-328.

Van Heerden, H.G., W.H. Van Staden, L.J.B. Rossouw and G.G. Theron, 1987. Release of breeding lines (*Gossypium hirsutum* L.) adapted to South African conditions. *South Africa Journal Plant Soil* **4**: 203-205.

Velasco, L.D.F. and H.C. Becker, 1998. Variability for the fatty acid composition of the seed oil in a germplasm collection of the genus Brassica. *Genetic Research and Crop Evolution* **45:** 371-382.

Verhalen, L.M., B.E. Greenhagen and J.C. Banks, 2002. Choosing which cotton variety to grow pp 4. Current Report. Division of Agricultural Sciences and Natural Resources.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Pelema, M. Kuiper and M. Zabeau, 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acid Research* **23**: 4407-4414.

Vroh Bi, I., A. Maquet, J.P. Baudoin, P. du Jardin, J.M. Jacquemin and G. Mergeai, 1999. Breeding for "low gossypol seed and high-gossypol plants" in upland cotton. Analysis of tri-species hybrids and backcross progenies using AFLPs and mapped RFLPs. *Theoretical and Applied Genetics* **99**: 1233-1244.

Wang, D.J., G.P. Li and B.L. Li, 1989. Study on distance hybridisation of cotton and character transfer through backcrossing. *China Cottons* **3:** 6-7

Wang, X.D. and J.J. Pan, 1990. Studies on the relationship between genetic distance of parents and yield heterosis in hybrid cotton. *Acta Agronomica Sinica* **16**: 32-38.

Wendel, J.F. and C.L. Brubaker, 1993. The gene pool of *G. hirsutum.* In: D.J. Heber and D.A. Richter (Eds.), Proceedings Beltwide cotton conference pp 1556. National Cotton Council, Memphis, TN.

Wendel, J.F., C.L. Brubaker and A.E. Percival, 1992. Genetic diversity in *Gossypium hirsutum* and the origin of upland cotton. *American Journal of Botany* **79**: 1291-1310.

Williams, J.G.K., A.R. Kubelic, K.J. Livak, J.A. Rafalski and S.V. Tingey, 1990. DNA polymorphism amplified by arbitrary primers is useful as genetic markers. *Nucleic Acids Research* **18:** 6531-6535.

Wolf, W.J. 1988. Effects of agricultural practices, handling, processing and storage on legumes and oilseeds. In: E. Karmas and R. S. Harris (Eds.), Nutritional Evaluation of Food Processing pp 119-152. AVI, New York.

World Cotton Research Conference (WCRC), 2003. Importance of cotton in South African agriculture. *http://www.cottonconference3.co.za/importance.htm*

Worley, J., R. Smith, H.H. Ramey, D.C. Harrel and T.W. Culp, 1976. Ontogenetic model of cotton yield. *Crop Science* **16**: 30-34.

Wricke, G., 1962. Über eine methode zur erffassung der ökologischen Streubreite in feldversuchen. *Z. Pflanzenzüchtg*. 47: 92-96.

Wricke, G. and W.E. Weber, 1980. Erweiterte analyse von wechselwirkungen in versuchsserien. In: Köpcke and Überla. (Eds.), Biometrie heute und morgen. Springer-Verlag, Berlin

Wright, P.R., 1999. Premature senescence of cotton (*Gossypium hirsutum* L.) predominantly a potassium disorder caused by an imbalance of source and sink. *Plant and soil* **211**: 231-239.

Wu, Y.T, T.Z. Zhang and J.M. Yin, 2001. Genetic diversity detected by DNA markers and phenotypes in upland cotton. *Yi-Chuan-Xue-Bao* **28**: 1040-1050.

Xian, X.X., X. Wang, Z. Yin and L. Xie, 1995. Analysis of combining ability and heterosis for parental varieties in upland cotton. *Journal of Hebei Agricultural University* **18:** 34-40.

Yates, F. and W.G. Cochran, 1938. The analysis of groups of experiments. *Journal Agriculture Science* **28**: 556-580.

You, J., J. Liu and J. Sun, 1998. Analysis of heterosis and its components in intraspecific crosses between upland cotton and breeding lines developed from primitive race stocks *(Gossypium hirsutum* L.*). Acta Agronomica Sinica* **24:** 834-839.

Yu, J. and R.J. Kohel, 1999. Update of the cotton genome mapping. Proceedings Beltwide Cotton Research Conference pp 439. National Cotton Council, Memphis, TN.

Zabeau, M. and P. Vos, 1993. Selective restriction fragment amplification: a general method for DNA finger printing. *European Patent Application 92402629*. Publication no. EP 0534858.

Zhang, T.G.W., 2001. Present status and prospect on cotton genomic studies. Department of Agronomy, Nanjing Agricultural University, Nanjing.

Zhang, X.L. and X.L. Zhang, 1997. Prediction and utilization of hybrid heterosis between dominant glandless lines and conventional cotton varieties. *China Cottons* **24:** 20-21.

Zhong, M., J.C. McCarty, J.N. Jenkins and S. Saha, 2002. Assessment of day-neutral backcross populations of cotton using AFLP markers. *Journal Cotton Science* **6**: 97-103.

Zhu, J., M.D. Gale, S. Quarries, M.T. Jackson and G.J. Bryan, 1998. AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics* **96**: 602-913.

Zock, P.L., J.H.M. de Vries and M.B. Katan, 1994. Impact of myristic versus palmitic acid on serum lipids and lipoprotein levels in healthy women and men. *Arterioscler Thromb* **12**: 911-919.

APPENDICES

Appendix 4.1 Rainfall and temperatures from October 2003 to March 2004

Month	Rainfall		Temperature	
	Amount (mm)	Number of days	Minimum	Maximum
October	22.5	7	16.9	32.3
November	100.5	9	17.5	30.5
December	206.1	15	17.0	30.5
January	86.0	12	17.2	30.6
February	116.5	7	17.0	30.0
March	66.5	10	16.5	30.5

Appendix 4.2 Specific morphological characters for different varieties

Variety	Pollen colour	Stigma position	Leaf colour	Leaf shape	Leaf hair	Stem hairs	Stem colour	Bract dent.	Bract size	Boll shape	Boll prominence	Boll peduncle	Plant height	Plant shape
HC-B4-75	cream	above	medium	palmate	medium	medium	reddish	medium	medium	conical	strong	long	medium	conical
Coker 315	cream	same	dark	palmate	medium	medium	rgreen	medium	medium	ovate	strong	medium	medium	conical
IL 85	yellow	above	light	palmate	strong	strong	rgreen	medium	medium	ovate	medium	medium	short	conical
Guazuncho	cream	same	dark	palmate	strong	strong	rgreen	medium	medium	ovate	strong	medium	medium	cylindric
SG-125	cream	same	dark	palmate	weak	weak	rgreen	medium	medium	ovate	strong	medium	medium	conical
CIM-70	cream	above	medium	palmate	strong	strong	rgreen	medium	medium	rounded	strong	medium	medium	conical
Auburn 56	yellow	same	dark	palmate	strong	strong	rgreen	medium	medium	rounded	medium	long	medium	conical
Des 119	cream	above	dark	palmate	medium	medium	rgreen	medium	medium	conical	medium	short	short	cylindric
Cyto 12/74	yellow	same	dark	palmate	strong	strong	reddish	conical	medium	ovate	medium	medium	medium	globose
NTA 93-15	cream	above	dark	palmate	Medium	medium	reddish	medium	medium	ovate	strong	medium	short	globose
Acala SJ$_2$	yellow	above	medium	palmate	strong	medium	rgreen	medium	medium	ovate	medium	long	tall	globase
Stoneville 506	cream	above	dark	palmate	medium	medium	rgreen	medium	medium	ovate	strong	long	medium	globose
Dixie King	cream	above	dark	palmate	weak	weak	rgreen	medium	medium	ovate	strong	medium	medium	cylindric
NTA 88-6	cream	above	dark	palmate	medium	medium	reddish	medium	medium	rounded	strong	medium	short	conical
Reba W296	yellow	same	medium	palmate	medium	strong	rgreen	medium	medium	rounded	medium	medium	tall	globose
UK82	yellow	same	light	palmate	strong	strong	lgreed	medium	medium	rounded	medium	medium	medium	conical
Frego bract	cream	same	medium	palmate	strong	strong	rgreen	fine	fine	rounded	strong	medium	short	conical
Delcot 344	cream	above	dark	palmate	weak	medium	rgreen	medium	medium	ovate	strong	medium	tall	globose
DP 4049	cream	below	dark	palmate	medium	medium	reddish	medium	medium	ovate	strong	medium	medium	globose
BJA 592	yellow	same	medium	palmate	strong	strong	rgreen	medium	medium	ovate	medium	medium	short	cylindric
Irma 1243	cream	above	medium	palmate	strong	v.strong	rgreen	medium	medium	ovate	medium	medium	Tall	conical
IL74	yellow	same	light	palmate	medium	strong	rgreen	medium	medium	ovate	medium	medium	short	globose
MZ561	yellow	same	light	palmate	strong	medium	rgreen	medium	medium	ovate	medium	medium	medium	globose
Okra leaf	yellow	below	dark	digtate	strong	strong	rgreen	medium	medium	rounded	medium	short	medium	globose
McNair 235	cream	above	dark	palmate	weak	medium	rgreen	medium	medium	ovate	strong	short	medium	conical
UK91	yellow	below	dark	palmate	strong	strong	reddish	medium	medium	rounded	medium	medium	medium	cylindric
NTA 93-21	cream	above	dark	palmate	strong	medium	reddish	medium	medium	ovate	strong	short	tall	conical
High gossypol	yellow	same	medium	palmate	strong	strong	rgreen	medium	medium	ovate	medium	long	medium	conical
Reba B50	yellow	same	light	palmate	medium	medium	rgreen	medium	medium	ovate	medium	medium	short	conical
DP Acala90	cream	same	dark	palmate	medium	medium	rgreen	medium	medium	ovate	medium	medium	short	cylindric

Pollen colour=cream, yellow, dark yellow; Stigma position relative to anthers= below anthers, same level with anthers, above anthers; Intensity of green colour= light, medium, dark; Leaf shape=palmate to digitate, digitate, lanceolate; Leaf hairs (leaf pubescence lower side)= absent or very weak, weak, medium, strong, very strong; Stem hairs (stem pubescence in upper part)= absent or very weak, weak, strong, very strong; Stem colour=light green, dark green, reddish green, reddish; Bract dentition (at green maturity)=fine, medium, coarse; Bract size (at green maturity)= small, medium, large; Boll shape (boll shape in longitudinal section)= rounded, elliptical, ovate, conical; Boll prominence (boll prominence of tip)=weak, medium, strong; Boll peduncle (boll length of peduncle)= short, medium, long; Plant height at green maturity= short, medium, tall; Plant shape= cylindrical, conical, globules

Appendix 4.3 Fibre quality expression

Fibre Micronaire		Fibre Maturity		Fibre Uniformity		Fibre Strength		Fibre Length		Fibre Elongation	
Value	Explanation	(%)	Explanation	(%)	Explanation	(g/tex)	Explanation	(mm)	Explanation	(%)	Explanation
≤ 3.0	Very fine	0.70-0.80	Immature	≤ 77	Non-uniform	≤ 17	Very weak	≤ 25.15	Short	≤ 5.0	Very slight
3-3.9	Fine	0.80-0.85	Below average	77-79	Low	18-21	Weak	25.15-27.94	Medium	5-5.8	Slight
4-4.9	Medium	0.85-0.95	Mature	80-82	Medium	22-25	Medium	27.94-32.00	Long	5.9-6.7	Middling
5-5.9	Coarse	0.95-1.00	Above average	83-85	High	26-29	Strong	≥ 32.00	Very long	6.8-7.6	High
≥ 6.0	Very course	1.00	Very mature	≥ 85	Very uniform	≥ 30	Very strong			≥ 7.6	Very high

g/tex= gram per tex