.6138 202 8X

U.O.V.S. BIBLIOTEEK

University Free State
34300000349344
Universiteit Vrystaat

HIERDIE EKSEMPLAAR MAG ONDER GEEN OMSTANDIGHEDE UIT DIE BIBLIOTEEK VERWYDER WORD NIE

01

Genetic variation among land races of *Cucurbita moschata* and inheritance of its fruit yield and quality.

By

Cousins Gwanama

Submitted in fulfilment of the requirements for the degree of

Philosophiae Doctor

in the Faculty of Agriculture

Department of Plant Breeding

The University of the Orange Free State

Supervisor: Prof. M.T. Labuschagne

Co-supervisor: Prof. A.M. Botha

July, 2000

Declaration

I, Cousins Gwanama, declare that the thesis submitted by me for the Philosophiae Doctor degree at the University of the Orange Free State is my own independent work and has not been submitted by me to another university or faculty. I further cede copyright of the thesis in favour of the University of the Orange Free State.

Signature..... Date..01/07/2000

Acknowledgements

I wish to extend my sincere gratitude to the following people and institutions for their assistance in this study:

1. The Deutscher Akademischer Austauschdienst (DAAD), Bonn, Germany, and the National Research Foundation (NRF), Pretoria, South Africa, for financial support.
2. The University of the Orange Free State, the Forest and Agricultural Biotechnology Institute/University of Pretoria and the University of Zambia for facilities
3 Prof. M.T. Labuschagne and Prof. A.M. Botha for advice throughout the study.
4. Mr. P. van Rooyen of the Grain Crops Institute, Potchefstroom, South Africa, for some statistical analyses
5. My wife, Mary, and family for moral support.

Contents

List of figures

List of Tables

List of abbreviations

ANOVA	Analysis of variance
cm	Centimetre
dap	Days after planting
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
EDTA	Ethyldiaminetetra-acetic acid
ha	Hectare
kg	Kilogram
m	Meter
mg	Miligram
mm	Millimetre
mM	Milimolar
ng	Nanogram
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphic DNA
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
TBE	Electrophoresis gel running buffer containing Tris, boric acid and EDTA
t/ha	Tons per hectare

Tris	Tris[hydroxymethyl]aminomethane
µl	Microlitre
µM	Micromolar

CHAPTER 1

INTRODUCTION

Tropical pumpkin, *Cucurbita moschata* Duchsne, is one the most important vegetables in Zambia. Its role as a traditional crop for household food security in the rural areas is well recognised. It is second only to maize in frequency of cultivation (Andren *et al.*, 1991). In the rural areas, it is grown mainly for its fruit and secondarily for the leaves. Its popularity has been spreading to the urban areas, although leaf consumption is more important in the towns than the fruit. Consequently the trend has been of production of fruit for food security in the country and as a cash crop for leaf vegetable production in peri-urban areas. As a cash crop, mean yearly incomes of over US$ 60 have been recorded. This is significant in a society where the majority of farmers earn less than US$ 200 per annum (Gwanama & Nichterlein, 1995).

The crop is also important for its contribution to nutrition in terms of abundant supply of minerals, vitamin C and, especially, β-carotene (provitamin A). Pumpkins contain up to 5 00 μg β-carotene per 100 g. This makes them a cheap choice for the prevention of xerophthalmia (an eye disease) which is reported to afflict at least 1.9% of children under the age of six.

The problems to be found in pumpkin production are twofold: agronomic and demand related. The former relates to the quality of cultivars as regards yield levels, stability of yield, suitability for tillage machinery, etc. The latter category involves market perceived quality attributes to earn a good price and encourage consumption. At the moment the cultivars grown in Zambia are land races which are very variable in both sets of attributes. To be able to supply fruits of consistent and high quality the intervention of deliberate crop improvement efforts is required. In particular, it would be preferable to produce hybrids as they normally have consistent and predictable performance and quality. At the moment such cultivars are lacking.

Scientific investigation in this species is quite backward in comparison to the major agricultural crops. A limited number of germplasm collections have been made and these have seldom been adequately characterised and evaluated for the important characteristics or even for genetic relatedness. In other words, there are only few, obscure breeding materials of reasonable genetic superiority. In such a case there is little choice but to begin with the local material which is adapted to the climate and endemic pests and diseases.

The objectives of this study, therefore, were

- to analyse the amount of genetic variation in *C. moschata* in Zambia and Malawi with the hope of finding material that is adequately genetically different to be used for crosses.
- to estimate genetic effects for important yield and quality traits this population.
- to choose putative parents for use in establishing crosses from which selection could be made for the traits of interest.
- to estimate the heritability for important traits in this population so as to have insights in the possible gains that could be obtained from selection.

The collection from Malawi was included because it was the only other available at the moment from the countries surrounding Zambia. Both conventional and molecular techniques were employed in the study to make as much progress as possible during the limited period. Ordinarily, materials for use in heritability and genetic effects analysis should be well known at the beginning from prior evaluation which requires at least two seasons and two environments. This requirement was circumvented by using the random amplified polymorphic DNA (RAPD) technique to quickly organise the material into related groups. However,

this organisation could only give insights as to genetic similarity, but could not possibly be associated with particular morphological characteristics.

CHAPTER 2

ANALYSIS OF GENETIC VARIATION IN *CUCURBITA MOSCHATA*

Chapter abstract

Knowledge of genetic relationships among genotypes is essential for the effective utilisation of germplasm, especially for poorly characterised species. Random amplified polymorphic DNA (RAPD) analysis provides a quick and reliable method for resolving genetic relationships. Although *Cucurbita moschata* Duch, also known as tropical pumpkin, is one of the most important vegetable crops in Africa, being adapted to a wide range of climatic and soil conditions, it is a scientifically neglected species. The objectives of this study were to (1) analyse the amount of genetic diversity in *C. moschata* land races grown in south-central Africa and (2) classify the land races to assist in selection of parent genotypes for improvement of fruit characteristics. Cluster analysis, based on 39 polymorphic and 105 monomorphic DNA fragments amplified by 16 primers, was used to show relationships among 31 genotypes obtained from Zambia and Malawi. The analysis revealed four clusters, with genotypes from Malawi mainly clustering in three clusters while all genotypes from Zambia and three from Malawi clustered in one cluster. The pair-wise mean genetic distance was 0.32±0.04 for samples from Malawi and 0.26±0.04 for samples from Zambia. The possible application of the resulting classification in breeding of *C. moschata* is discussed.

Key words: Cucurbita moschata, cluster analysis, RAPD's.

2.1 **LITERATURE REVIEW**

Tropical pumpkin (*Cucurbita moschata*) has not yet received attention as an important crop species. The few small germplasm collections on record have not been characterised phenotypically or otherwise. It is important at the beginning of a breeding programme to discriminate among available genotypes to establish the level of genetic diversity and, thereby, identify the most suitable material for crosses.

Molecular markers provide a quick and reliable method for estimating genetic relationships among genotypes of any organism. They can facilitate rapid screening of large numbers of genotypes for polymorphic loci (Thormann *et al.*, 1994). An overwhelming number of such marker assays is available. The most appropriate markers for any question should be those that (1) are heritable, (2) discriminate between individuals or populations, (3) are easy to measure and evaluate and (4) provide results that can be compared with similar studies (Westman & Kresovich, 1997).

Molecular markers may be divided into two broad groups: those that rely on hybridisation and those that employ the polymerase chain reaction (PCR). Hybridisation techniques, although very reproducible, are very laborious and

require a prior knowledge of the DNA sequences of the target sites to be assessed. Some PCR techniques also require sequence knowledge for the design of primers (Staub *et al.,* 1996). For species in which sequence information is lacking, those PCR procedures are to be preferred which do not require such information.

Among the PCR methods which require no sequence information, random amplified polymorphic DNA (RAPD) analysis (Williams *et al.,* 1990) is one of the simplest and fastest for use in genetic similarity studies. The RAPD assay is based on amplification of unknown genomic regions using short (oligonucleotide) random primers followed by electrophoresis and visualisation of these fragments on agarose gels. Similar technologies that use short primers are the arbitrarily primed PCR (AP-PCR, Welsh & McClelland, 1990) and the DNA amplification fingerprinting (DAF, Caetano-Anolles *et al.*, 1991). These are more costly than the RAPD because they use polyacrylamide and denaturing gels, respectively.

Some of the disadvantages of RAPD markers are as follows:

(1) RAPD markers are usually dominant, rather than codominant. Hence only one allele per locus is sampled in each analysis (Staub *et al.*, 1996).

(2) The genomic region of amplified DNA (nuclear or cytoplasmic) and the

sequence homology of fragments with similar mobility are not clear. In one interspecific study (Thormann *et al.*, 1994) 15 bands amplified by 14 RAPD primers were tested by Southern blotting for homology with equal mobility bands in their replicate samples. Three of them failed to hybridise. Further testing of other fragments found seven of them to be homologous with mitochondrial sequences and one with a chloroplast sequence. The limitation of sequence homology is especially important if higher levels of taxa are to be compared (Westman & Kresovich, 1997).

(3) Reproducibility is a major concern in RAPD analysis. This depends on the quality of reaction ingredients, the design and GC content of primers, the cycling temperature and number of cycles. Variation in results between different makes of thermocyclers has been observed even for the same reaction conditions (Westman & Kresovich, 1997). In many cases each study must first optimise these variables.

Nonetheless, the advantages of RAPD analysis far outweigh the limitations. These include a lack of the requirement for sequence information of the species, ease and speed of the assay, little amount of DNA required, no use of radioactivity and ability to provide markers in genomic regions with repetitive DNA sequences (dos Santos *et al.*, 1994; Thormann *et al.*, 1994; Welsh *et al.*, 1991). Unlike RFLP's, RAPD analysis does not even require high quality DNA.

Any plant organ can be used for DNA extraction. In some instances reproducible results have been obtained by directly using ground plant tissue as template.

Additionally, RAPD's are also advantageous over isozyme analysis (the benchmark of genetic analysis since the late 1950's) as they detect more polymorphism at about the same cost of analysis owing to the larger number of primers available compared to only around twenty enzyme systems in the latter (McDonald *et al.*, 1994).

Most studies to test the reliability of RAPD markers in estimating intraspecific genetic relationships found them to be as reliable as RFLP's (the benchmark of genetic analysis of the late 1980's). dos Santos *et al.* (1994) used 62 decamer primers and 15 cDNA probes to discriminate among 45 *Brassica oleracea* cultivars and found the two methods to have similar resolving power. They concluded that RAPD markers were to be preferred for genetic similarity studies owing to the ease of the technique. Hallden *et al.* (1994) used 50 RFLP probes and 92 decamer primers to determine genetic relationships among three doubled haploids of *Brassica napus*. Nearly identical dendrograms were obtained from the two methods and bootstrap analyses showed that both needed about 30 probes or primers to deduce these relationships. Thormann *et al.* (1994) compared 41 RAPD primers and 125 RFLP probes for their ability to distinguish

among accessions within and between species. They used 18 accessions from six cultivated *Brassica* species and one from *Raphanus sativus*. RAPD results were very similar to RFLP results for relationships within species, but conflicted for interspecific relationships to RFLPs. They concluded that RAPD's were not to be recommended for interspecific genetic studies.

In some cases it has been shown that RAPD's are capable of detecting more polymorphism than RFLPs, especially in self fertilised crops of narrow genetic base (Staub *et al.*, 1996). Garcia-Rodriguez *et al.* (1996) discriminated among nine genotypes of *Cucumis. melo*, using 39 RAPD markers. Previous RFLP analysis had failed to detect sufficient polymorphism for establishment of genetic relationships. In *Brassica rapa ssp. oleifera*, 75% of 220 RAPD markers detected polymorphisms between two parental genotypes compared to 67% of 74 RFLP probes (Tanhuanpää *et al.*, 1995a). Similarly, in *B. napus* RAPD's detected more polymorphism (48%) than RFLPs (38%) between two parental lines (Tanhuanpää *et al.*, 1995b).

The objective of this study was to analyse the amount of genetic variation in *C. moschata* land races grown in the south-central African region and classify them to assist in selection of parent genotypes in a breeding programme.

2.2 MATERIALS AND METHODS

2.2.1 Plant material

Thirty-one accessions (land races) were included in the investigation. Thirteen were obtained from a collection made in Malawi in 1990 and the rest from a collection made in Zambia in 1993 and 1994. The areas from which the accessions were collected lie between latitude 8^{o} -- 17^{o} S and longitude 23^{o} -- 36^{o} E and cover a wide range of agroclimatic zones from hot, semiarid lowlands to warm, wet highlands receiving above 1,200 mm of rainfall.

2.2.2 DNA extraction

Eight seeds per accession were used for DNA extraction. Small pieces of about a quarter of a seed were cut from whole seeds previously soaked in water for 48 hours. The pieces were ground with a pestle and mortar in liquid nitrogen. The powder was thoroughly mixed and about 200 mg of it homogenised in 600 µl of extraction buffer (containing 200 mM Tris (pH 7.4), 250 mM NaCl, 25 mM EDTA (pH 8.0) and 17.3 mM SDS) in a 1.5 ml microcentrifuge tube. The homogenate was left at room temperature for 10 minutes to one hour after which it was centrifuged at 12,000 rpm for 15 minutes. The supernatant (300

µl) was transferred to a fresh microcentrifuge tube and digested with 2 µl of RNase A (10 mg/ml). The samples were extracted with one volume of chloroform/isoamyl alcohol (24:1). The supernatant was transferred to a fresh microcentrifuge tube and DNA precipitated by adding half a volume 7.5 M ammonium acetate and two volumes of absolute ethanol. After a further centrifugation, the alcohol was discarded and the DNA air-dried and re-dissolved in 50 -- 100 µl of sterile water. Quantification of the DNA was carried out electrophoretically on a 1 % agarose gel alongside cut and uncut λDNA

2.2.3 DNA amplification

Single arbitrary 10-base primers from the series OPB, OPK, OPL and OPM (Operon Technologies, Almeda, California, USA) were tested for their ability to amplify scorable and reproducible DNA fragments (Table 2.1). Primers resulting in faint or unreproducible bands were excluded from subsequent analysis. Twenty-one primers were accepted for the statistical analysis. Each 25 µl PCR reaction mixture consisted of 5 – 12 ng genomic DNA; 280 µM dNTPs; 3.5 mM $MgCl_2$; 0.001% gelatine; 0.3 µM primer; 0.6 units of *Taq* polymerase (Boehringer Mannheim); and 1 X amplification buffer [20 mM $(NH_4)_2SO_4$, 75 mM Tris pH 9.0, 1% Tween 20] supplied by the enzyme manufacturer. Samples

were subjected to the following thermal profile for amplification in an oven thermocycler (Hybaid Omnigene-TR3 CM 220): 3.5 min at 94 °C; 40 cycles of 30 s at 94 °C, 30 s at 35 °C and 1.5 min at 72 °C; and a final elongation step at 72 °C for five minutes. Visualisation of amplification fragments was accomplished on a 2 % agarose gel (Seakem LE, FMC, USA) in 0.5 X TBE buffer stained with ethidium bromide.

2.2.4 Statistical analysis

Data scored as presence (1) or absence (0) of amplification fragments were used to calculate Jaccard's coefficients of similarity for each pair of accessions as follows:

$$S_{ij} = X_{ij}/(X_{ij} + X_i + X_j)$$

where X_{ij} is the number of shared fragments between samples i and j, X_i and X_j are the numbers of fragments unique to each sample, respectively. Both polymorphic and monomorphic fragments from polymorphic primers were used in the analysis. Estimates of similarity between all pairs of accessions were converted to dissimilarity and expressed as Euclidean genetic distances (dissimilarities) by the formula $\sqrt{(1-S_{ij})}$ (Gower, 1966). Genetic distance estimates of pair-wise comparisons served as elements in the proximity matrix, D, in cluster analysis by the unweighted pair-group method using arithmetic averages (UPGMA; Sneath & Sokal, 1973).

2.3 RESULTS AND DISCUSSION

Sixteen primers yielding consistent and distinct amplification products were included in the analysis. Between five and 13 scorable fragments were amplified by each primer per genotype with an average of nine fragments per primer (Table 2.1). Between one and five polymorphic bands were scored for each primer. Figure 2.1 and Figure 2.2 show examples of the RAPD fragment variation that was obtained in the amplification. Bands amplified ranged in size between 300 and 3 000 bp. The polymorphic fragments were obtained in the range 500 -- 2 000 bp.

One hundred and forty-four fragments were generated out of which 39 (23 %) were polymorphic. Genotypes could be distinguished by a combination of fragments but no genotype-specific fragments were scored. Thus, differences between clusters reflected differences in frequencies rather than presence or absence of fragments.

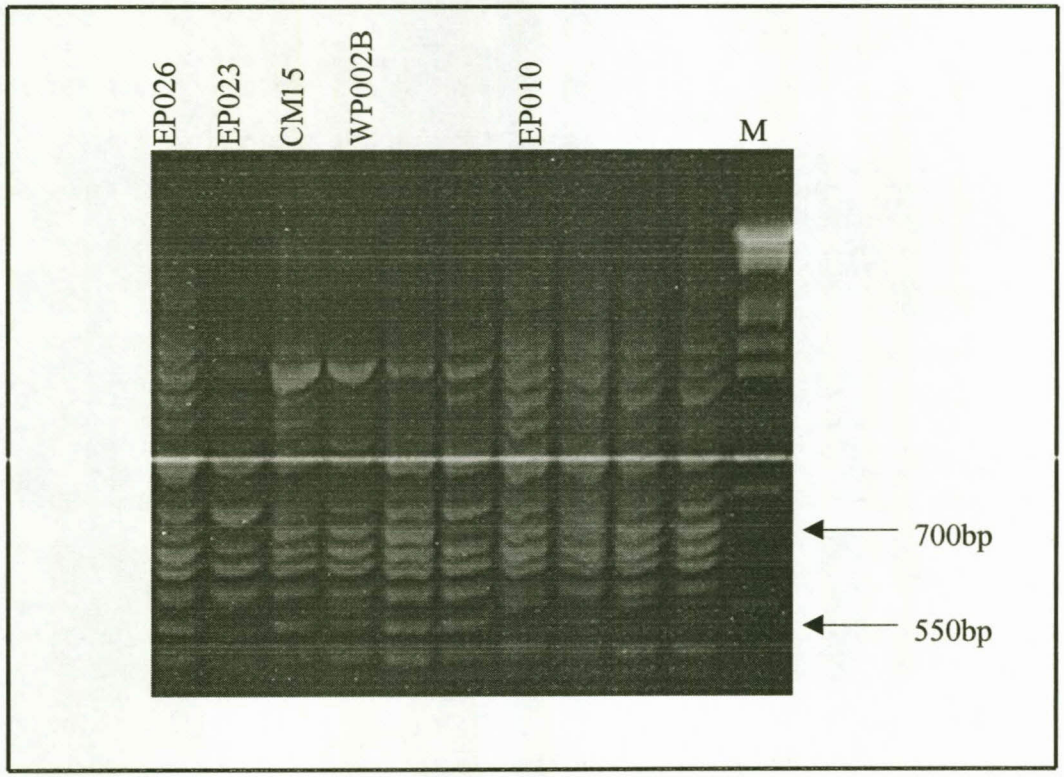

Figure 2.1 A portion of a RAPD profile resulting from amplification by primer OPL-15 of DNA extracted from pumpkin seed. A 700 bp fragment is absent in land race EP023. A 500 bp fragment is1 absent in land race EP010. The extreme right lane (M) contains bands from a λ Eco RI/ Hind III digest.

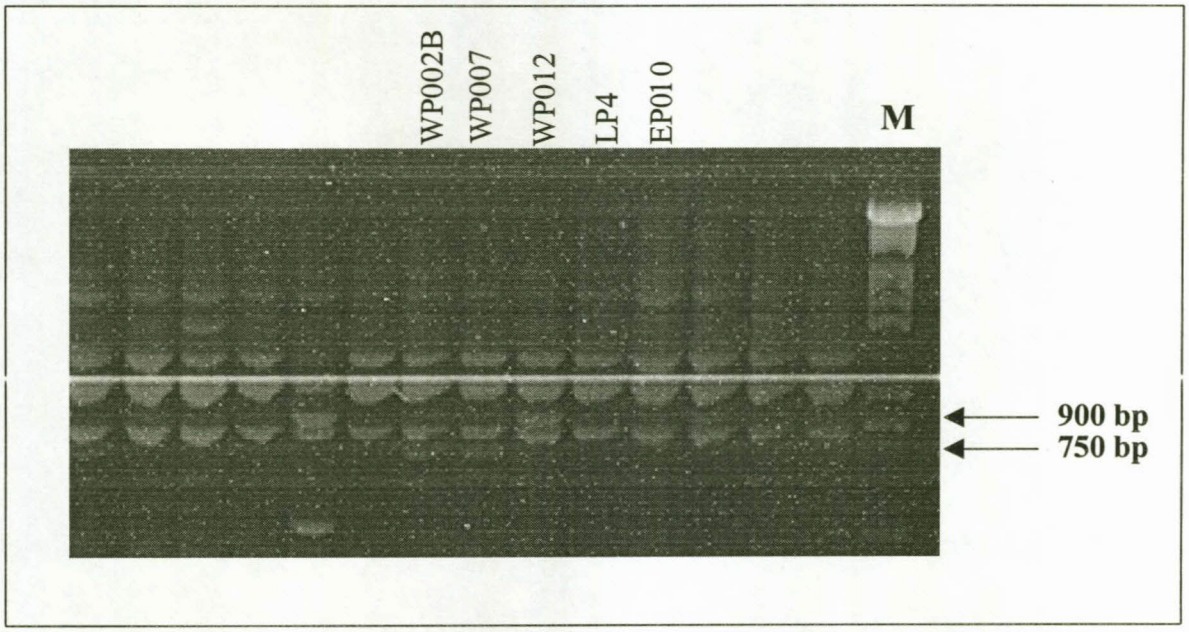

Figure 2.2 A portion of a RAPD profile resulting from amplification by primer OPK-12 of DNA extracted from pumpkin seed. A 900 bp fragment is present in land race WP012 and missing in WP002B, WP007, LP4 and EP010. A 750 bp fragment is present in land races WP002B and WP007 and absent in land raceWP012, LP4 and EP010. The extreme right lane (M) contains bands from a λ Eco RI/ Hind III digest.

Table 2.1. List of random primers used in the analysis of *Cucurbita moschata* and the yield of scorable DNA fragments.

Primer	Total number of fragments	Polymorphic fragments	
		Number	Percent
OPB-01	11	4	36
OPK-07	9	2	22
OPK-10	8	1	13
OPK-12	7	3	48
OPK-16	5	2	40
OPK-19	12	5	42
OPL-03	10	3	30
OPL-05	12	4	33
OPL-07	12	4	33
OPL-08	5	1	20
OPL-11	4	1	25
OPL-14	11	2	18
OPL-15	13	3	23
OPL-16	10	2	20
OPL-17	7	1	14
OPL-18	8	1	13
Total	144	39	
Average	9		23.2

The range of dissimilarities obtained was 0.13 (between EP002 and LP15) to 0.41 (between CPC74 and LP4) (Table 2.2). The mean genetic distance between all the samples was 0.31±0.04. The mean genetic distance for Malawian accessions was 0.32±0.04. Except for one outlier, LP4, genotypes from Zambia were generally more similar ($0.13 \leq D \leq 0.36$) with a mean genetic distance of 0.26±0.04. Accession LP4 was divergent from the rest of the samples, being characterised by absence of more fragments than the rest (Table 2.2), and had a mean genetic distance of 0.37±0.03.

The dendrogram produced by cluster analysis revealed four clusters (Figure 2.3). Although individual pair-wise dissimilarities were as big as 0.41 the dendrogram did not reveal any accessions separated by $D > 0.36$. This was due to the normalising of data before construction of the dendrogram. Clusters I, II and III contained only accessions from Malawi. Except for accession LP4, all accessions from Zambia grouped together in cluster IV, together with three samples from Malawi (Fig 2. 3).

The entire collection of Zambian accessions in cluster IV were joined by smaller dissimilarities in comparison to those in the other clusters. Preferred ecotypes in Zambia may probably have a similar genetic background. This cluster also includes the land race 'Munasangu' which is the predominant land race grown

for home and market consumption around the major urban areas in Zambia. Cluster IV represents the elite land races which could be used as testers for combining ability. Accessions CM15 and CM21 clustered close together in cluster IV. These accessions were collected within 100 kilometres of the Southern province of Zambia. Samples WP002B and WP012, collected from within 200 kilometres of the Western province of Zambia, also clustered together (Figure 2.3).

The accession LP4 clustered independently (Figure 2.3). Divergent genotypes may have good breeding values. Genotypes in the same cluster may represent members of one heterotic group. Maximum variability for selection in segregating populations may be achieved by utilising genotypes from different clusters as parents of crosses.

The dissimilarities obtained in this study show rather modest variation in land races of *C. moschata* cultivated in south-central Africa. The reason for this is high frequency of seed exchange by farmers over wide geographic-ethnic regions. At least 40% of farmers in Zambia obtain their cucurbit seed from other farmers annually (Gwanama & Nichterlein, 1995). It is probable that greater diversity in *C. moschata* would be seen if a wider collection from more distant geographical regions was included. However, inclusion of unadapted ecotypes in

a local breeding programme would necessitate a longer breeding cycle as it would be required to break the linkage drag of undesirable characteristics.

The mean genetic distance estimates, however, compare roughly with those in use in breeding programmes for other crops of the Cucurbitaceae. In *Cucumis melo,* the variation is much smaller, with dissimilarities from of 0.01 to 0.2 (Garcia-Rodriguez *et al.,* 1996). In *Citrullus lanatus*, Lee *et al.* (1995) found a mean genetic distance of 0.30 although individual distances among pairs of genotypes were as high as 0.91. Their study, which involved inbred lines, showed closer genetic relationships between individual genotypes within distantly related clusters.

In other outcrossing species greater variation is often observed, for example, in *Cymbidium* a mean genetic of 0.29 has been reported with a wider range of 0.08 to 0.5 (Obara-Okeyo & Kako, 1998). In *Zea mays,* dissimilarities of up to 1.0 (Mumm & Dudley, 1994) have been reported. However, comparison with maize is difficult as inbreds possessing maximum dissimilarity owing to homozygosity for all polymorphic loci were used.

Analysis of a highly heterozygous species by the RAPD technique is expected to underestimate genetic distances, as RAPD markers are dominant markers (Staub

et al., 1996). Detection of heterozygous loci using codominant marker techniques would give more realistic genetic relationships. However, considering time and cost savings, RAPD's have shown adequate resolution of the land races to be used for breeding purposes. RAPD markers have also been used elsewhere in outcrossing species to provide quick insights into organisation of germplasm. For instance, in *Cymbidium* they have been shown to produce dendrograms that were consistent with known pedigree (Obara-Okeyo & Kako, 1998). Phenotypic characterisation in *Cucurbita* has traditionally been based on seed and fruit characteristics. These have proved useful in distinguishing samples of related species but exhibit inadequate variation for intraspecific discrimination of cultivars. In addition phenotypic markers in *Cucurbita* have been found to be unreliable and with no molecular basis (Wilson, 1989).

The modest genetic variation in the two collections studied, points to the need to widen the variability by further collections afield and to possibly introduce material from the primary centre of diversity - Central and North America. However, knowledge of the genetic relationships generated in this investigation will be of value in directing the exploitation of available germplasm. Moreover, differences in the performance of crosses may possibly be attributed to RAPD markers. This would later form a basis for a marker assisted breeding programme.

Table 2.2 Pair-wise distance matrix of the 31 *C. moschata* land races form Zambia and Malawi based on the D statistic, I.e. $\%D = (1-S_{ij})^{1/2}$

Accession	Accession number																													
Number	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2
1	.35	.23	.36	.36	.37	.35	.33	.31	.37	.38	.34	.35	.34	.36	.35	.34	.33	.32	.30	.33	.33	.30	.38	.33	.38	.35	.34	.27	.31	.26
2	.35	.36	.36	.36	.41	.36	.30	.35	.38	.40	.38	.32	.35	.34	.36	.34	.34	.34	.33	.29	.35	.37	.34	.32	.30	.33	.30	.30	.27	
3	.30	.28	.34	.27	.34	.31	.34	.36	.34	.37	.32	.27	.32	.31	.34	.34	.27	.31	.30	.31	.33	.31	.29	.31	.36	.31	.38	.26		
4	.32	.30	.29	.30	.36	.30	.29	.25	.31	.36	.30	.33	.32	.32	.29	.35	.31	.31	.28	.30	.31	.27	.31	.31	.32	.31	.35			
5	.35	.32	.32	.34	.39	.36	.32	.32	.37	.37	.37	34	.32	.34	.35	.38	.33	.37	.38	.34	.38	.39	.38	.32	.41	.36				
6	.28	.34	.36	.35	.34	.35	.32	.33	.32	.37	.35	.31	.36	.33	.32	.32	.33	.29	.26	.26	.35	.29	.28	.34	.26					
7	.31	.36	.39	.36	.35	.36	.34	.35	.33	.34	.36	.34	.35	.36	.35	.32	.32	.33	.34	.29	.34	.25	.32	.34						
8	.30	.29	.32	.31	.34	.31	.31	.33	.27	.38	.32	.26	.30	.29	.28	.29	.24	.26	.35	.28	.27	.30	.28							
9	.32	.30	.33	.32	.32	.25	.33	.33	.28	.33	.33	.28	.30	.30	.32	.30	.25	.31	.32	.22	.32	.32								
10	.26	.29	.33	.38	.38	.32	.32	.34	.28	.33	.32	.30	.31	.33	.31	.34	.30	.27	.32	.30	.30									
11	.34	.34	.33	.35	.32	.31	.36	.35	.34	.33	.31	.32	.34	.34	.31	.33	.32	.33	.35	.30										
12	.31	.27	.32	.28	.31	.28	.27	.31	.25	.32	.28	.26	.27	.29	.32	.25	.20	.31	.28											
13	.30	.27	.27	.26	.36	.33	.26	.30	.33	.31	.34	.28	.29	.30	.34	.35	.27	.29												
14	.24	.27	.29	.26	.34	.32	.32	.33	.24	.29	.32	.24	.27	.28	.27	.28	.24													
15	.30	.21	.24	.18	.32	.26	.24	.28	.22	.28	.27	.13	.16	.19	.25	.21														
16	.27	.29	.29	.26	.32	.31	.29	.30	.23	.28	.30	.23	.23	.23	.28															
17	.26	.32	.31	.27	.38	.30	.33	.29	.25	.33	.32	.30	.28	.27																
18	.25	.21	.28	.22	.36	.30	.24	.22	.21	.34	.28	.19	.25																	
19	.31	.23	.18	.20	.31	.27	.26	.30	.26	.32	.30	.24																		
20	.26	.21	.31	.22	.32	.28	.24	.29	.20	.33	.26																			
21	.31	.33	.33	.36	.36	.30	.28	.26	.26	.34																				
22	.31	.33	.35	.32	.36	.37	.36	.36	.31																					
23	.22	.23	.28	.24	.34	.30	.30	.27																						
24	.28	.24	.32	.29	.34	.28	.27																							
25	.32	.24	.30	.31	.34	.31																								
26	.35	.30	.30	.31	.35																									
27	.35	.33	.33	.32																										
28	.27	.21	.24																											
29	.33	.28																												
30	.29																													

*: 1 = CPC2, 2 = CPC74, 3 = CPC 77, 4 = CPC91, 5 = CPC114, 6= CPC119, 7= CPC139, 8 = CPC102, 9 = CPC89, 10 = CPC122, 11 = CPC55, 12 = CPC 128, 13 = CPC30, 14 = Lame pumpkin, 15 = EP002, 16 = EP021, 17 = EP014, 18 = Munasangu, 19 = MK5, 20 = LP15, 21 = EP026, 22 = EP023, 23 = CM15, 24 = WP002B, 25 = WP007, 26 = WP012, 27 = LP4, 28 = EP010, 29 = EP001, 30 = CM21, 31 = MK7

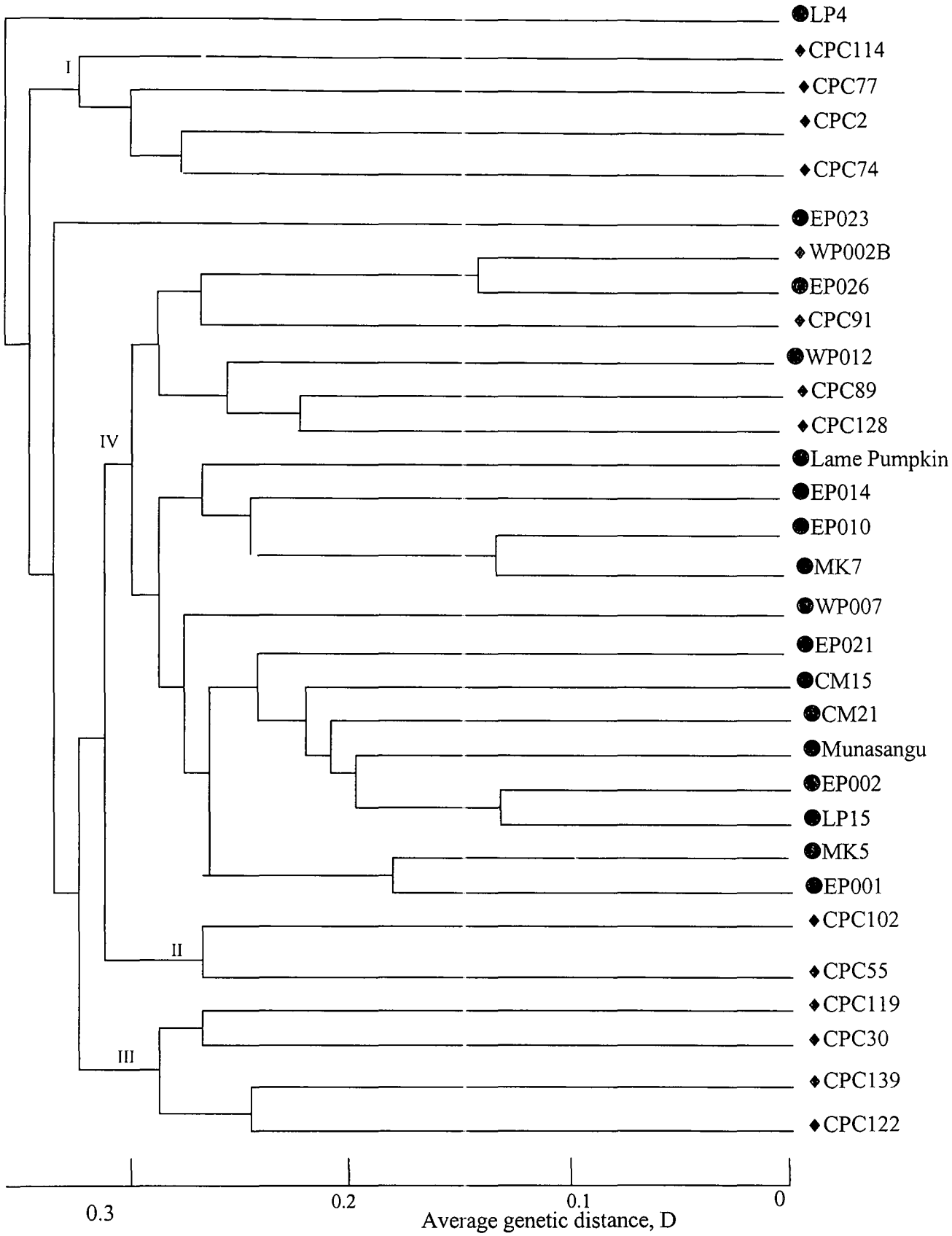

Figure 2.3. Dendrogram showing genetic relationships among 31 *C. moschata* genotypes based on average genetic distance. Symbols indicate country of origin: ● = Zambia, ♦ = Malawi

CHAPTER 3

GENETIC EFFECTS OF FRUIT YIELD AND QUALITY TRAITS OF *CUCURBITA MOSCHATA*

Chapter abstract

Estimation of genetic effects for a fixed set of genotypes may yield information on the presence of different types of gene action for important traits. Six land races of *Cucurbita moschata* (tropical pumpkin) and one set of their crosses were evaluated for fruit yield and quality traits. Variety effects and heterosis mean squares were significant for the days to female anthesis, the weight of the first fruit mature, the mean fruit weight and the soluble solids, indicating importance of both additive and non-additive gene action. Only variety effects were significant for the days to the first mature fruit. These characters could be improved by recurrent selection. Selection for heterosis would be effective except for the days to the first fruit to mature. Further analysis by offspring-parent regression indicated significant heritability for the days to the first mature fruit and the soluble solids. The non-significance of the other characters, despite significant variety effects, probably indicates the presence of inbreeding. Inclusion of selfs of the varieties is recommended in further tests to completely resolve the genetic effects. Significant negative correlation between the days to the first mature fruit and fruit yield makes it possible to increase yield by selecting for early maturity.

Key words: genetic effects, combining ability, heritability, heterosis, selection

3.1 LITERATURE REVIEW

3.1.1 Components of variance and heritability

The plant breeding process can be divided into three stages. Stage one involves the assembly of a pool of variable germplasm, stage two the selection of superior individuals from the pool and stage three the utilisation of superior individuals to create a better variety.

Estimation of genetic parameters of a population is useful in all these stages in deciding the appropriate breeding strategy that will utilise the genetic variance present. Dudley and Moll (1969) delineated six questions for which estimation of genetic variance components may be useful: (1) whether there is sufficient genetic variation in a population to allow further improvement of characters to take place, (2) the extent of testing required to identify superior genetic populations, (3) identification of the most promising genetic populations, (4) selection of the most rapid and efficient breeding procedure in improving important characters, (5) the type of variety to be produced and (6) whether the same breeding procedure would be effective in improving a number of important traits.

The phenotype of an individual can be divided into effects due to its genotype and the environment. Classically, a general linear model involving *r* genotypes $g_1, g_2, \ldots g_r$ in *s* environments $e_1, e_2, \ldots e_s$ is defined for the phenotypic mean, x_{ik}, of a character in terms of the population mean, μ, the genotypic effect, α_i, the effect of the environment, β_k and the genotype by environment interactions, γ_{ik} as follows (Jacquard, 1983):

$$x_{ik} = \mu + \alpha_i + \beta_k + \gamma_{ik} \quad \text{for } i = 1, 2, \ldots r \text{ and } k = 1,2, \ldots s$$

From this relationship the phenotypic variance, σ^2_P, is obtained as

$$\sigma^2_P = \sigma^2_G + \sigma^2_E + \sigma^2_{(GE)} + 2\sigma_{GE}$$

where σ^2_G is the total genetic variance, $\sigma^2_{(GE)}$, the genotype by environment variance, σ^2_E, the environmental deviations and σ_{GE} is the covariance between the genotypic and environmental values.

Decomposition of the σ^2_G into additive (σ^2_A) and non-additive (dominance) (σ^2_D) components yields a relationship that shows the contributions of the genetic causes of variation:

$$\sigma^2_P = \sigma^2_A + \sigma^2_D + \sigma^2_{(GE)} + \sigma^2_E + 2\sigma_{GE}$$

Additive genetic variance is the only component which is transmitted to offspring and, therefore, reflects the degree to which progenies are likely to

resemble their parents. The additive genetic variance of a single locus is determined by gene frequency and the average effect of substituting one allele by another. Additive genetic variance of a population is the sum of the additive genetic variances contributed by individual loci. Its proportion of the total variance, the (narrow sense) heritability (h_N^2), represents the breeding value of a population and is the primary parameter that determines gain from selection (Falconer, 1960)

Heritability is the resemblance, or covariance, among relatives. Knowledge of the heritability of a character is important since it indicates the possibility and extent to which improvement is possible through selection (Robinson *et al.*, 1949). From its definition it can be estimated either by the analysis of variance (ANOVA) or as regression of offspring on parents (Dudley & Moll, 1969; Falconer, 1960; Jacquard, 1983). If it is computed by the ANOVA, based on a single parent (half-sib family analysis):

$$h_N^2 = 2\sigma^2_A/\sigma^2_P = 2\sigma^2_A/(\sigma^2_A + \sigma^2_D + \sigma^2_{(GE)} + \sigma^2_E + 2\sigma_{GE})$$

Based on two parents (full-sib family analysis):

$$h_N^2 = \sigma^2_A/\sigma^2_P = \sigma^2_A/(\sigma^2_A + \sigma^2_D + \sigma^2_{(GE)} + \sigma^2_E + 2\sigma_{GE})$$

Estimation of heritability from regression of progeny on parents overcomes such limitations as small samples, high percentage of population required to be

selected, and slow rate of inbreeding (or cycle required to produce inbred populations) (Robinson *et al.*, 1949). If h_N^2 is measured from data consisting of parent-offspring pairs, it is equal to twice the regression of offspring on parents. If the data consists of mean-parent-offspring pairs h_N^2 is equal to the regression of progeny on parents (Jacquard, 1983). That is,

Parent-offspring data: $h_N^2 = 2\sigma_{OP}/\sigma^2_P$,

Mid-parent-offspring data: $h_N^2 = \sigma_{OP}/\sigma^2_P$

where the subscripts P and O denote parents and offspring, respectively.

However, if progenies are regressed on parents grown in the same environment, h_N^2 is overestimated due to confounding of genotype by environment interactions (Casler, 1981). Two approaches are available for removing the bias in h_N^2 estimated from parent-offspring regression. The first is to perform a covariance analysis and estimate the genotype-environment covariances and remove them from the h_N^2 equation. The second approach is to evaluate parents and progeny in different environments (at different locations in different years) and regress progeny means on parent performance from a different environment (Casler, 1981). Although the second approach produces genotype X environment covariances that are equal to zero, nevertheless, differential environmental expression of parents and progeny can have a drastic effect on the magnitude of the h_N^2 estimates (Frey & Horner, 1957).

3.1.2 Combining ability and heterosis

The concept of combining ability is used in connection with testing procedures to study and compare genotypes in hybrid combination. The term 'general combining ability' (*gca)* refers to the average performance of a genotype in hybrid combination. The performance of a particular cross may deviate from the general combining ability of its parents. This deviation is the 'specific combing ability' (*sca*) of the cross (Griffing, 1956). Heterosis refers to the deviation of progeny performance from the mean of its parents. In this respect, heterosis and sca are similar concepts, differing only in the way they are estimated. Heterosis is normally used with respect to a progeny's two parents, while sca refers to the deviation of the performance of a cross from the whole set of progeny resulting from crosses among several parents. *Gca* and *sca* represent the major types of gene action for quantitative traits. *Gca* is largely due to additive effects and higher order gene interactions, while *sca* is largely a function of non-additive genet action (Cukadar-Olmedo *et al.,* 1997). The importance of significant *gca* is that the best performing progeny may be produced by selecting from the cross of the parents having the highest *gca.*

The decision on what type of varieties to produce basically concerns whether F_1 hybrids should be produced. Only when the ratio of dominance to additive

genetic variance is in the overdominace range, that is, when better parent heterosis is indicated, is production of F_1 hybrids preferred (Dudley & Moll, 1969). In all other cases varieties other F_1 hybrids are to be favoured.

Estimation of combining ability and genetic effects requires a mating design. The term 'mating design' refers to the system of mating to develop progenies. Mating designs can be classified as one, two, three or four factor designs depending on the number of ancestors per progeny over which control is exercised. A set of half sib families or a polycross would constitute a one factor design. The diallel cross, North Carolina I, II and III are all two-factor designs, while the triallel and quadriallel crosses are three- and four-factor designs, respectively (Cockerham, 1963). It is always preferable to use the simplest design that will generate data for the variance components desired. A one-factor design is sufficient to detect presence of genetic variation. For separation into additive and dominance variance a two-factor design is necessary. More complex designs are required for estimation of epistatic variance (Dudley & Moll, 1969).

The diallel cross (Sprague & Tatum, 1942) is the most commonly used two-factor design for estimating genetic effects, choosing the type of breeding system to use and selecting materials that show the greatest promise. In this procedure, *p*

parents, usually inbreds, are chosen and crosses made among these. Griffing (1956) outlined four methods for computing genetic effects, *gca* and *sca:* (1) parents, one set of F_1's and reciprocal F_1's (all p^2 combinations) (2) parents, one set of F_1's, reciprocal F_1's excluded ($p[p+1]/2$ combinations) (3) all F_1's including reciprocal crosses, parents excluded ($p[p-1]$ combinations) (4) one set of F_1's only ($p[p-1]/2$ combinations). In these analyses a general model for variety effects (v_{ij}) was composed of *gca* (g_i and g_j) and *sca* (s_{ij}) effects of parents i and j by the relationship

$$v_{ij} = g_i + g_j + s_{ij}$$

The advantages of the diallel over the other two-factor designs for estimating combining abilities lies in its ability to evaluate more combinations for specific combining abilities. For this reason, the diallel cross is more popular for combining ability analysis, especially if F1 hybrids are desired. Another advantage is that in the analysis of variance components, the probability of failure to obtain genetic effects which can occur when these effects are present owing to the cancelling of opposite effects at different loci (pairs of loci), is less in a diallel when each variety is crossed with all other varieties (Eberhart & Gardner, 1966). The disadvantages of the diallel cross include the large number of crosses generated and the limitations of the type of inferences that may be made when used in genetic studies (Stuber, 1980).

For each of these methods two types of statistical models are distinguished: (1) situations where parents are deliberately chosen and constitute the population about which inferences are to be made (fixed effects model or Model I), (2) situations where parents can be considered to be a random sample of a larger population about which inferences are to be made (random effects model or Model II) (Griffing, 1956). In an analysis based on fixed effects model, one is concerned with the combining abilities of the actual parents and with the identification of superior combinations. In an analysis based on random effects model, inferences are to be made about the population from which the parents were sampled and these inferences are made from estimates of components of variance. Diallel analysis of *gca* and *sca* under fixed effects model requires no statistical or genetic assumptions (Baker, 1978).

When a random effects model is used, estimation of genetic variances is possible in addition to *gca* and *sca.* In diploid inheritance and with the parents completely inbred, the population σ^2_A and σ^2_D can be estimated from the *gca* (σ^2_g) and *sca* (σ^2_s) variances by $\sigma^2_A = 2\sigma^2_g$ and $\sigma^2_D = \sigma^2_s$ (Griffing, 1956). Significant σ^2_g implies additive genetic effects and high heritability (Quendeba *et al.,* 1993). From this relationship, the closer the ratio of $2\sigma^2_g/(2\sigma^2_g+\sigma^2_s)$ to unity the greater the predictability of F1 hybrids based on *gca* alone (Baker, 1978). Small ratios and significant σ^2_s implies non-additive genetic effects and low precision in

predicting hybrid crosses. Cultivar development is best achieved through use of specific, rather than general, good combiners (Prixely & Bjarnason, 1993). In the case where non-inbred parents are used, the additive and non-additive variances can be estimated using the relationships $\sigma^2_A = 4\sigma^2_g/(1+F)$ and $\sigma^2_D = \sigma^2_s/(1+F)^2$ where F is the inbreeding coefficient of the parents (Wright, 1985).

Interpretation of components of variance obtained from the random effects model requires limiting genetic assumptions. The most common of these requires that genes be independently distributed (i.e, linkage equilibrium) (Baker, 1978). It is common to include additional assumptions such as lack of epistasis, multiple allelism, genotype-environment interaction, reciprocal differences and differential gametic selection (Pooni *et al.,* 1984). These conditions are met only by a large population produced by repeated generations of random outcrossing without selection (Wright, 1985). The assumption of linkage equilibrium is, therefore, in most cases considered untenable (Baker, 1978).

It has long been theorised that failure of this assumption may result in underestimation or overestimation σ^2_A and σ^2_D (Hayman, 1954). Sughroue and Hallauer (1997) presented empirical data for maize inbreds comparing 12 random diallels created by five generations of intermating from an eight-parent

diallel and the original diallel itself. In six of the eight traits evaluated σ^2_A was underestimated and σ^2_D overestimated. That study showed that coupling phase linkages cause estimates of σ^2_A and σ^2_D to be overestimated, while repulsion phase linkages cause estimates of σ^2_A to be underestimated and σ^2_D to be overestimated.

Failure of the assumption may also be due to the effects of sample size when sampling the parents. Feyt (1976) (Cited in Baker, 1978) has pointed out that the genes at *n* loci cannot be expected to be independent unless a minimum of 2^n parents are used in the diallel cross. Hayman (1954) concluded that estimates of variance components could not be significant estimates of population parameters unless the number of parents exceeds 10. With fewer parents, a fixed model is recommended. For these reasons it is advised that if the primary purpose of an investigation is the estimation of genetic components of variation and to test the assumptions on which the estimates are based, the diallel design should not be the preferred design (Pooni *et al.*, 1984).

There are two possibilities implied by the reference population: ancestral or descendant reference. These two definitions have genetic implications over and above those imposed by the statistical assumptions of random and fixed genotype effects to which they correspond (Wright, 1985). 'Ancestral reference'

is a population from which the diallel parents can be considered a sample from which they were derived by inbreeding without selection. If diploid inheritance is assumed the diallel F1's are a sample of crosses from the ancestral reference population and this population is in linkage equilibrium, irrespective of the inbreeding status of their parents, whereas the S1's are inbred and so belong to a different population. Therefore, the assumption of linkage equilibrium need not be made for estimation of genetic variances (Wright, 1985).

The 'descendant reference' is the equilibrium population which could be generated by repeated cycles of random mating initiated by the diallel parents. The descendant population has an inbreeding coefficient of $F = p/2$ or $1/p$ for non-inbred or inbred parents, respectively. The S_1 progenies have an inbreeding coefficient of 1/2 or 1 and constitute a $1/p$th part of the full diallel set. Therefore, the full diallel is the descendant population (parents, F_1's, S_1's). This full diallel is in linkage equilibrium for individual loci but is not in equilibrium collectively for all loci as the S_1's have a higher probability of homozygosity at all loci than the F_1's. Although the individual *gca* estimates of these are meaningful, the variance of these, σ^2_g, has no simple interpretation in terms of σ^2_A (Wright, 1985).

The choice of crossing method depends on the experimental material and the objectives of an investigation. When information on *gca* and *sca* is of primary concern, methods three and four are more applicable with the distinction between them that method three is to be used if reciprocal effects are indicated. Method one and two are used if combining ability analysis is employed to determine suitable lines to combine into a synthetic variety or if considerable inbreeding occurs in the species (Griffing, 1956). However, the distinctions are not significant with large diallels as estimates of σ^2_g do not differ materially between the methods when p >10 (Pooni *et al.,* 1984). With small diallels it is expected that either

σ^2_g (method two) > σ^2_g (method one) >> σ^2_g (method three) ≈ σ^2_g (method four)

or

σ^2_g (method two) < σ^2_g (method one) << σ^2_g (method three) ≈ σ^2_g (method four)

The genetical expectations of σ^2_g and σ^2_s for method one, Model II are identical with the general definitions of additive and non-additive components of variation in a randomly mating population. This is not true for methods two, three and four for diallels with p < 20 (Pooni *et al.,* 1984).

3.1.3 Genetic effects for a fixed set of genotypes

Most genotypes of interest to breeders have been highly selected for traits of economic importance. In other cases, they have not been created by repeated cycles of inbreeding from some parent population. With such material, the assumption that varieties are a random sample from some equilibrium base population is completely invalid; and estimation of variance components does not provide useful information. The investigator is then dealing with genetic constants, and not variances (Eberhart & Gardner, 1966). Griffing's methods for estimating combining ability under the fixed effects model are somewhat limited in the amount of genetic information they yield. Gardner and Eberhart (1966) proposed a more general fixed effects model for open pollinated variety crosses which can also be used for inbreds and other populations as long as those varieties or inbreds are in Hardy-Weinberg equilibrium.

Their original model was restricted to two alleles per locus and only additive and dominance effects. They used the concept of heterosis to cover the same ideas as combining ability. In this model, it is possible to estimate the genetic contributions of homozygous (a_j) and heterozygous (d_j) loci, 'variety effect', v_j, 'average heterosis', $\bar{h}$, 'variety' or 'general heterosis', h_j, of the jth variety and the 'specific heterosis', $h_{jj'}$ which occurs when variety j is mated to variety j'. In

the Griffing methods *sca* is the only parameter to measure heterosis. Method three and four do not measure heterosis due to omission of parents while even their inclusion in method one and two fails to give a clear picture of heterosis and the genetic effects involved (Gardner & Eberhart, 1966). Even in cases where parents are not included Gardner and Eberhart's Analysis II is superior to Griffing's method four by providing two tests for non-additive genetic effects, namely d_j and $h_{j'}$.

This model was later expanded to cover any number of alleles per locus and any number of loci. Parameters are defined for additive, dominance and additive by additive epistasis assuming diploid inheritance (Eberhart & Gardner, 1966). Since the definition permits an arbitrary number of alleles at each of an arbitrary number of loci for an unspecified amount of varieties, the necessary assumptions have been reduced to diploid inheritance, equilibrium within each variety, and no epistasis other than additive by additive. The main additional term in the expanded model is $aa_{jj'}$ which is the additive by additive epistasis. The relation between the two dominance effects, d_j and h_{ij} is evident from their mean expectations. The function of dominance effects defined as d_j is a measure of inbreeding depression in variety j, and the function defined as h_{ij} is a measure of heterosis in the cross of varieties j and j' if epistasis is lacking. When epistasis must be considered parameters exist for which it can be estimated, namely a_j and

1512583

$aa_{jj'}$. One advantage of this model is that if estimates of all parameters are obtained, and higher order epistasis and/or linkage are negligible, the performance of any entry can be predicted without actually growing all entries (Eberhart & Gardner, 1966).

3.1.4 Yield and quality traits of *C. moschata*

The most important utilisation of tropical pumpkin in Zambia and most of Africa is consumption of mature vegetable fruit, normally cooked and eaten as pudding or snack. Leaves are an important vegetable both in the country and in urban areas, while seed is roasted for snacks in the country (Gwanama & Nichterlein, 1995). The characters that have been found by correlation analysis to influence fruit yield are the length of the primary vine, the number of fruits per plant, the mean fruit weight and the weight of the first fruit to mature. Intermediate phenotypic correlations with yield ($0.5 < r < 0.7$) were obtained by Gopalakrishnan *et al.* (1980) and Chigwe (1992). In one study (Gwanama *et al.*, 1998), path analysis showed that the significant phenotypic correlations of these traits did not translate into significant direct genotypic effects except for the weight of the first fruit to mature and the number of fruits per plant. The mean fruit weight and length of the primary axis only indirectly influenced fruit yield through the weight of the first fruit to mature and the number of fruits per plant,

respectively. In a selection programme special attention should be given to the weight of the first fruit to mature and the number of fruits per plant. However the number of primary branches selected simultaneously with number of fruits per plant would also be helpful. Apart from these studies, literature on genetic components of these characters in the species is lacking.

Quality attributes differ from market to market. In general, not much consumer preference has been surveyed for this species. In one case study, Puerto Rican buyers scored fruit size, skin and internal flesh colour as the leading quality characteristics. The ideal pumpkin was of a size of 20-25 centimeter diameter, a piebald skin colour and a deep orange pulp colour (Carbonell *et al.*, 1990). Other markets will have different preferences. In the United States the leading characteristics are skin colour, flesh colour and soluble solids (Maynard, 1996). The soluble solids characteristic is important because it affects mature vegetable fruit palatability.

The objectives of this study were to:

- estimate genetic effects for important yield and quality traits in this population.
- to choose varieties to be used as parents for crosses from which selection could be made to improve economically important traits.

- to estimate the heritability for important traits in this population so as to have insights in the possible gains that could be had from selection.

3.2 MATERIALS AND METHODS

3.2.1 Choice of parents

On the basis of the cluster analysis (Chapter 2, Fig 2.1), accession LP4 and five land races taken randomly from cluster IV were chosen to be parents. Land race LP4 was chosen because it clustered independently (Fig 2.1). Seed of genotypes from Malawi proved to be inviable and so excluded from the trial. The parents used in the crossing strategy were EP021, EP026, LP4, Munasangu, MK7 and WP007.

3.2.2 Diallel crossing strategy

The six parental genotypes were sown at the University of Zambia on 1 September, 1998 in paired rows to effect hybridisation in a diallel cross. Each row consisted of four plants sown at a spacing of 2 m within and 2 m between rows. Basal dressing fertiliser was broadcast before planting at the rates of 123 kg/ha KCl and 309 kg/ha 'Compound D'(N:P:K = 10:20:10). Farmyard manure at the rate of 5 t/ha was drilled in shallow groves spaced 30 cm apart. Ammonium nitrate fertiliser was applied at the rate of 100 kg/ha six weeks after planting. Irrigation was applied once a week throughout the experiment.

Clean weeding was practiced manually. It was not necessary to control pests during the trial.

Staminate flowers began reaching anthesis 48 days after planting. Female flowers reached anthesis about 15 to 20 days later. Generally, the maturity of male and female flowers in pairs of rows for crossing did not coincide so that pollen had to be carried around from desired parents. Using this procedure, all plants were used both as donors and receptors. Both the male and female flowers of plants to be crossed were enclosed with glassine bags the day before anthesis. Limited studies of the Cucurbitaceae (for example, *Cucumis sativus*) have found their pollen viable for a few days after anthesis (Frankel & Galun, 1977). In this study pollen was employed for fertilisation for up to 48 hours and pollination effected before noon of each day for three weeks. Stamens were removed from the male flower of the paternal parent and rubbed unto the stigmas of the pistillate flower of the maternal parent. As only a few individuals were to be grown in the evaluation, representation of all the sampled plants for each set of parent varieties was essential. Seeds were packed in batches of 48 with 12 seeds coming from every female involved in a cross.

3.2.3 Field evaluation of F1 and parents

The experiment was sown at the University of Zambia Field Station on 4 October 1999. The pH of the experimental field at the beginning of the trial was 7.0 in the top 10 cm and 6.8 at 20 cm below the surface. The six open pollinated parental varieties and one set of crosses (15 crosses) were sown in a randomised complete block design replicated three times. Single row plots with four plants each at a spacing of 2 m between plants and 2 m between rows were used. The plots were initially sown with four seeds per station and thinned to one plant two weeks after emergence. A hive of bees was installed to effect pollination. Fertilisation and crop management were identical to those used during the crossing. Standard cultural practices were carried out, including irrigation when required as assessed by visual observation of the soil moisture at 10 cm below the soil surface.

Fruit yield-related characters measured were: the length of the primary vine (m) at 60 dap, the number of primary branches at 60 dap, the days to female anthesis (dap), the days to first fruit to mature(dap, when fruits were hard enough so that pricking with forefinger nail would not leave a scar), the weight of the first fruit to mature (kg) total mature fruit yield mature (kg, continuos harvest), the number of fruits per plant and the mean fruit weight (kg). Single

plant observations and measurements were taken from all four plants and plot means used for analysis.

For fruit quality two traits were measured. One representative fruit from each plant was ripped open and the thickness of the internal flesh was measured (cm). Pieces of about 25 g from the edible portion were obtained and the four pieces from each plot homogenised with a blender. This homogenate was taken for soluble solids analysis by the Brix method as described by Southgate (1986).

3.2.4 Statistical analysis

The data were subjected to analysis of variance using the MSTAT-C (1991) software. The soluble solids data were transformed ($\sqrt{[x+1]}$) before the analysis. Characters with significant genotype mean squares were subjected to further analysis.

3.2.4.1 Genetic effects

The genetic information of the varieties and the crosses were obtained using Analysis II of Gardner and Eberhart (1966):

$Y_{jj'} = \mu_v + (v_j + v_{j'})/2 + \gamma\bar{h} + \gamma(h_j + h_{j'}) + \gamma s_{jj'}$

Where

$Y_{jj'}$ = mean expectation of a cross between varieties *j* and *j'*

μ_v = mean of all varieties

v_j and $v_{j'}$ = the variety effects

$\bar{h}$ = the average heterosis contributed by the varieties

h_j and $h_{j'}$ = the contribution of each variety to the expression of heterosis

$s_{jj'}$ = the specific heterosis which occurs when varieties j and j' are mated

$\gamma = 0$ when *j* = j' and 1 when $j \neq j'$

3.2.4.*2* Correlation and heritability estimates

Simple phenotypic correlations between yield components were computed using plot means. The heritability for each trait with significant genotypic

mean square was calculated as the regression of offspring on mid-parent. The procedure suggested by Casler (1981) to break genotype-environment correlations was used. This involved regressing progeny values on mean-parent values from a different replicate within the same environment as follows:

$h^2_N = S_{xyj}/S_{xxj'}$, for $j \neq j'$

The standard error of heritability was calculated according to Falconer (1960) as:

$SE_h \approx \sqrt{[(S_{yy}/(N.S_{xx})]}$

Where S_{xy}, S_{xx}, and S_{yy} denote, respectively, progeny and mid-parent sum of products, the mid-parent sum of squares and the progeny sum of squares.

3.3 RESULTS AND DISCUSSION

3.3.1 Performance of varieties and their progeny

Equally vigorous growth was shown by all the genotypes with no differences in the primary vine length or the number of primary branches at 60 dap (Table 3.1) All the genotypes were of vining habit with mean length at 60 dap of 2.7 m and the average number of primary branches of 11.0. These data are comparable with those of Chigwe (1991) taken at the end of the season. In the current season stem development was assessed early in order not to disturb fruit set.

Significant differences for entries were detected for the days to female anthesis, days to the first fruit to mature, the weight of the first fruit to mature, the mean fruit weight and soluble solids (Table 3.1 and Table 3.2). The time to female anthesis ranged from 60 to 76 days with an average of 67 days (Table 3.1). These times were much longer than those reported by Maynard (1996) which varied between 29 and 48 days. The days to the first fruit to mature varied from 107 to 131 days with a mean of 114 days (Table 3.1). There were no reports in the literature for comparison of these data. The mean weight of the first mature fruit (2.6-5.2 kg with mean of 3.6 kg) and the mean fruit weight (2.4-4.7 kg with mean of 3.3 kg) (Table 3.1) were similar to another study using land races from the same collections in rain-fed conditions (Gwanama *et al.*, 1998).

They were also similar to the group classified as 'large fruited' inbreds and hybrids developed and evaluated in America (Maynard, 1996). The mean weight of the first to mature and the mean fruit weight were larger than results from Malawi (means 2.46 and 2.0 kg, respectively; Chigwe, 1991)

The number of harvested fruit per plant was low (1.3-3.5 with mean of 2.1, Table 3.1) due to abortions of immature fruits after the onset of rain at about the time of anthesis. When contrasted with other studies (Chigwe, 1991; Maynard, 1996; Gwanama *et al.*, 1998) this figure is acceptable. Maynard reported that the fruit number per plant for the larger fruited genotypes was always much lower than the smaller fruits ones. Smaller fruited genotypes (mean fruit weight less than 2 kg) yielded up to 6.9 fruit per plant on average (Maynard 1996). It is, therefore, not clear if the observed abortions were due to stress or whether it is a genetically controlled phenomenon to limit fruit load. Judging from the large standard errors for fruit number it would seem that more consistent fruit loads would be obtainable with improved control of the environment (possibly soil moisture and pollination). It seems probable that fruit set could be made consistent by planting the crop such that flowering and fruit set do not take place during predominantly humid periods. For the dry season crop the best time to plant would be September. Wet season plantings would be best in the middle of December.

The mean fruit yield per plant ranged from 4.3 to 9.3 kg with a mean of 6.7 kg (Table

3.1). This translates to 10.8-23.2 t/ha with an average of 16.7 t/ha. As yield per plant is the product of fruit number and mean fruit weight, significant differences were not detectable for yield despite significant differences in mean fruit weight (Table 3.1). The mean yield is slightly higher than in a previous study (Gwanama *et al.*, 1998) although top yields of up to 39 t/ha were recorded in that case. Yields of over 40 t/ha are possible from selections (Maynard, 1996). It should be noted, however, that the very high yields were reported for bush genotypes which are planted at much higher densities while still achieving similar numbers of fruit per plant

There was not much variation in the internal flesh thickness. The soluble solids ranged from 2.4 to 6.5 %. These results were consistent with those of Maynard (1996).

Table 3.1 Performance of *Cucurbita moschata* varieties (land races) and variety crosses for fruit yield and fruit quality characters

Land race/ cross	Trait									
	1	2	3	4	5	6	7	8	9	10
1. LP4	12.4	2.8	76 a-b[2]	129 a	3.2 c-e	3.1 b-e	1.8	5.5	27	4.4 a-b
2. WP007	13.8	3.1	66 e-g	109 b-c	3.2 c-e	3.1 b-e	1.8	5.7	33	4.5 a-b
3. Munasangu	10.8	2.5	62 g	115 b-c	3.6 b-e	2.9 c-e	2.5	7.3	32	3.8 b-c
4. EP026	11.2	2.5	65 e-g	108 c	5.2 a	4.7 a	1.8	8.2	35	4.1 a-c
5. EP021	11.7	2.5	62 g	109 b-c	2.7 d-e	2.4 e	3.5	8.2	30	4.4 a-b
6. MK7	9.3	2.5	74 a-c	115 b-c	4.0 a-d	4.2 a-c	1.5	6.8	36	3.8 b-c
7 = 1X2	12.8	2.9	75 a-b	131 a	2.8 d-e	2.5 d-e	1.8	5.2	31	4.4 a-c
8 = 1X3	11.0	2.2	71 b-e	113 b-c	3.7 b-e	3.4 b-e	1.7	6.0	37	3.9 a-c
9 = 1X4	9.0	2.9	77 a	112 a-b	3.1 c-e	3.2 b-e	1.3	4.8	34	3.9 a-c
10 = 1X5	10.4	2.7	72 a-d	116 b-c	3.6 b-e	3.3 b-e	1.8	5.9	36	4.1 a-c
11 = 1X6	12.3	2.8	61 g	112 b-c	2.6 e	2.4 e	2.8	6.7	34	2.6 d-e
12 = 2X3	12.5	2.6	64 f-g	112 b-c	3.5 b-e	3.0 c-e	2.7	8.1	31	3.1 c-e
13 = 2X4	8.1	2.9	66 d-g	112 b-c	3.7 b-e	3.3 b-e	2.2	8	33	2.4 e

Continued.

Traits : 1. Number of primary branches, 2. Primary vine length at 60 dap[1] (m), 3. Days to female anthesis, 4. Days to first mature fruit, 5. Weight of first mature fruit (Kg), 6. Mean fruit weight (kg), 7. Number of fruits per plant, 8. Fruit yield per plant (Kg), 9. Internal flesh thickness (mm), 10. Soluble solids (%)

Table 3.1 continued

Land race/ cross	Trait									
	1	2	3	4	5	6	7	8	9	10
14 = 2X5	11.1	2.5	63.0 g	111 b-c	3.9 b-e	3.1 b-e	2.2	6.9	29	5.1 a
15 = 2X6	10.5	2.4	64 f-g	114 b-c	3.5 b-e	3.3 b-e	2.5	8	28	3.8 b-c
16 = 3X4	8.4	2.7	69 c-f	113 b-c	3.4 c-e	2.7 d-e	2.7	7.4	29	3.6 b-d
17 = 3X5	11.4	3.0	64 f-g	113 b-c	4.8 a-b	4.3 a-b	1.6	6.3	31	4.6 a-b
18 = 3X6	12.0	2.4	62 g	117 b-c	3.1 c-e	3.1 b-e	1.5	4.3	35	4.0 a-c
19 = 4X5	10.8	2.9	63 f-g	111 b-c	2.9 d-e	3.1 b-e	1.7	5.2	31	4.6 a-b
20 = 4X6	10.9	3.0	66 e-g	110 b-c	4.4 a-c	4.1 a-c	2.3	9.4	30	3.8 b-c
21 = 5X6	10.2	2.5	63 fg	112 b-c	3.9 b-e	3.7a-d	2.0	7.3	30	3.5 b-d
Mean	11.0	2.7	67	114	3.6	3.3	2.1	6.7	32	4.0
Minimum	8.1	2.2	60	107	2.6	2.4	1.3	4.3	27	2.4
Maximum	13.8	3.1	76	131	5.2	4.7	3.5	9.3	37	6.5

Traits: 1. Number of primary branches, 2. Primary vine length at 60 dap[1] (m), 3. Days to female anthesis, 4. Days to first mature fruit, 5. Weight of first mature fruit (Kg), 6. Mean fruit weight (kg), 7. Number of fruits per plant, 8. Fruit yield per plant (Kg), 9. Internal flesh thickness (mm), 10. Soluble solids (%)

1. dap = days after planting

2. Means in the same column denoted by different letters are significantly different by Duncan's Multiple Range Test at $p \leq 0.05$.

3.3.2 Variety effects

In the Gardner and Eberhart model, 'variety effects' denotes the average performance of the parent varieties in the offspring. This is somewhat similar to (the traditional) Griffing's general combining abilities from Model I, method two where the parents are used as testers. The heterosis is partitioned in average, variety and specific components. The average heterosis is a measure of the mean difference of the progeny from the parents. The variety heterosis is the average heterotic contribution of each variety. This model is completely different from Griffing's models in the sense that Griffing's specific combining abilities confounds all heterotic effects and still contains some additive gene action (Eberhart & Gardner, 1966). The current model could also be used to estimate *gca* and *sca* if the parents are excluded from the analysis. Such an analysis would yield less genetic information than that obtained here.

The variety effects were significant for all the five characters with significant entry mean squares (Table 3.2). Heterosis was significant for all traits except the days to the first fruit to mature.

The significant variety effects for all the five traits imply presence of additive gene action in the population. These characters are likely to yield some gains to recurrent selection procedures. However, as pointed out by Gardner and Eberhart

(1966), estimation of genetic effects using varieties and their crosses does not separate additive gene action from the inbreeding portion of non-additive gene action in the variety effects. Separation of these would require the additional inclusion of selfs of the varieties. In a selection programme land races LP4 (1) (v_j = 8.11**) and MK 7 (6) (v_j = 6.78**) would be good sources of delayed flowering and late fruit maturity genes while Munasangu (3) (v_j = -5.22**) and EP021 (5) (v_j = -5.22**) would contribute to early maturity (Table 3.3).

Although early maturity is the more preferable phenomenon, late maturity genotypes would still perform well in the higher rainfall areas of Zambia. The country has three main agro-climatic regions: Region I in the South and East (rain season around three months, 600 mm of rain), Region II in the middle of the country (four months, 800 mm) and Region III in the north (five months, 800-1 200 mm). Incidentally, the LP4 was collected from Region I, MK7 from Region II, Munasangu and EP021 from the drier parts of Region II. The respective maturity genes were concentrated by farmer selection.

The time to female anthesis (beginning of reproductive phase) and the time to first fruit maturity are significantly correlated (r = 0.58*, Table 3.6). In most cases fruit maturity is the more important character. However, in situations where fruits are eaten immature, the time to beginning of reproductive phase would be a more relevant trait for which to select. The performance of the two

traits were not identical: Both variety and heterosis mean squares were significant for the days to anthesis while only variety effects were significant for the days to the first fruit to mature. The days to the first mature fruit could be expected to contain a larger portion of additive gene action, be more amenable to recurrent selection and produce more predictable progeny.

Table 3.2 Mean squares from Gardner and Eberhart Analysis II for genetic effects of fruit yield and quality traits

Source of variation	df	Days to female anthesis	Days to first fruit	Weight of first fruit	Mean fruit weight	Soluble solids
Entries	20	27.33**	37.4*	0.43**	0.39**	0.028**
Varieties (vj)	5	54.62**	87.50**	0.47*	0.47*	0.033*
Heterosis (h)	15	18.10**	20.69	0.41*	0.36*	0.026*
Average ($\bar{h}$)	1	4.68	1.60	0.08	0.09	0.053*
Variety (h_j)	5	29.87**	16.2	0.59**	0.54**	0.023
Specific ($h_{jj'}$)	9	13.05**	25.29	0.35*	0.30	0.024*
Error	40	4.03	16.16	0.16	0.14	0.011

* = significant at $p \leq 0.05$; ** = significant at $p \leq 0.01$.

The most desirable parent variety would be one with significant negative variety effects for days to female anthesis and days to first maturity and significant positive variety effects for weight of the first mature fruit, mean fruit weight and soluble solids. None of the land races used had all these qualifications. In addition to those already mentioned in respect of time to anthesis and maturity, land race EP026 had favourable significant variety effects for the weight of the first fruit (V_j = 1.53**) and the mean fruit weight (V_j = 1.35**). In a selection programme for early maturity and larger fruit, simultaneously, land races Munasangu, EP026 would be ideal candidates.

Improvement of days to female anthesis, days to first fruit to mature, the weight of the first fruit to mature and the mean fruit weight would be fairly easy as at least two genotypes are favourable in the first two categories and one genotype has desirable qualities in the third and fourth categories. The breeding cycle would be accomplished by simply hybridising either of the first two land races to the third and follow this by generations of conventional selection methods. If these traits were distributed among three or more land races the simultaneous improvement of the traits would require long cycle involving hybridisation of two genotypes followed by selection and hybridisation to a third genotype, followed by selection and so on. Alternatively, a diallel selective mating system could be used (Jensen, 1970). In this type of scheme three or more genotypes are mated in diallel fashion to simultaneously contribute genes to a gene pool in the F1 series. The F1 series can be directly subjected to mass selection as well as forming the basis for subsequent selective mating series which are themselves subjected

to mass selection. The diallel selective mating system shortens the breeding cycle when compared to the conventional selection.

Table 3.3 Estimates of variety effects (v_j) and variety heterosis (h_j) associated with the six parent land races of the diallel for days to female anthesis, days to first mature fruit, first fruit weight, mean fruit weight and soluble solids in pumpkin

Parent[1]	Days to anthesis		Days to first fruit		Weight of first fruit		Mean fruit weight		Soluble solids	
	v_j	h_j	v_j	h_j	v_j	h_j	v_j	h_j	v_j	h_j
1	8.11**	1.63	14.94**	-2.06	-0.44	-0.23	-0.32	-0.16	0.04	-0.015
	(76)	(76)[2]	(129)	(129)	3.2)	(3.2)	(3.1)	(3.1)	(4.4)	(4.4)
2	-1.88	0.22	-5.39**	4.69	-0.44	0.15	-0.32	-0.08	0.08	-0.132*
	(66)	(66)	(109)	(109)	(3.2)	(3.2)	(3.1)	(3.1)	(4.5)	(4.5)
3	-5.22**	1.97	0.94	-1.89	-0.04	0.24	-0.48	0.32	-0.09	0.079
	(62)	(62)	(115)	(115)	(3.6)	(3.6)	(2.9)	(2.9)	(3.8)	(3.8)
4	-2.56	3.56**	-6.39**	1.94	1.53**	-0.78**	1.35**	-0.62**	-0.02	-0.002
	(65)	65)	(108)	(108)	(5.2)	(5.2)	(4.7)	(4.7)	(4.1)	(4.1)
5	-5.22**	0.06	-5.39**	-0.06	-0.97*	0.82**	-1.02**	0.86**	0.06	0.160*
	(62)	(62)	(109)	(109)	(2.7)	(2.7)	(2.4)	(2.4)	(4.4)	(4.4)
6	6.78**	-7.44**	1.27	-2.64	0.36	-0.20	0.78	-0.32	-0.08	-0.090
	(74)	(74)	(115)	(115)	(4.0)	(4.0)	(4.2)	(4.2)	(3.8)	(3.8)
SE	1.83	0.97	3.67	2.59	0.37	0.26	0.35	0.24	0.096	0.068

* = significant at $p \leq 0.05$, ** = significant at $p \leq 0.01$; 1. Parents: 1 = LP4, 2 = WP007, 3 = Munasangu, 4 = EP026, 5 = EP025, 6 = MK7; 2 = numbers in brackets are parent means;

3.3.3 Variety and specific heterosis

Heterosis was significant for all the five traits except the days to the first fruit to mature. For these four traits both additive and non-additive gene action are important. The average heterosis was not significant for all traits except soluble solids (Table 3.2). Thus the performance of crosses as a set did not differ from their parents. In the case of soluble solids the crosses performed generally worse (mean 3.8 %) than the parents (mean 4.2%). The specific heterosis was significant for the days to female anthesis, the weight of the first fruit and the soluble solids (Table 3.2) Variety heterosis was significant for all traits except the days to the first fruit to mature and soluble solids (Table 3.2). Significant variety heterosis points to putative parents with generally desirable heterotic response in the crosses.

In common with the desirability of variety effects, the most desirable parent variety would be one with significant negative variety heterosis for days to female anthesis and days to first maturity and significant positive variety heterosis for weight of the first mature fruit, mean fruit weight and soluble solids. None of the land races used had all these qualifications. Land race MK7 had favourable variety heterosis for days to female anthesis (early flowering, h_j = -7.44**). Land race EP021 gave the most favourable variety heterosis for the fruit characteristics (weight of first fruit, h_j = 0.82*; mean fruit weight h_j = 0.86*; soluble solids h_j = 0.16*, Table 3.2). The fruit weight performance of

EP021 was not favourable (mean 2.4 kg, Table 3.3). All these requirements are met between MK7 and EP021, making them the best candidates for creation of a population for improvement of the five characters simultaneously by selection for heterosis. These two land races would be best for inclusion into a programme to produce synthetics with early maturity, bigger and more palatable fruit.

Specific heterosis was significant for the days to female anthesis, the weight of the first mature fruit and soluble solids. Desirable significant heterotic effects were obtained for days to anthesis in the combinations LP4 X MK7, WP007 X EP021 and EP026 X EP021 (h_{jj}' = -7.15**, -3.23**, -2.90*, respectively). For the weight of the first mature fruit favourable specific heterotic responses were obtained in the cross EP026 X MK7 (h_{jj}' = 0.94) and WP007 X EP021 (h_{jj}' = 0.21) in the case of soluble solids (Table 3.4). The ranking of all the above favourable specific heterotic response was in agreement with the *per se* performance of the crosses. According to Patil and Chopde (1983), agreement of *per se* performance and sca is an indication that selection of crosses on the basis of heterosis would be effective.

Table 3.4. Estimates of specific heterosis ($h_{jj'}$) for days to female anthesis, weight of the first fruit, mean fruit weight and soluble solids.

Cross	Days to female anthesis	Days to first fruit	Weight of first fruit	Mean fruit weight	Soluble solids
1X2	3.18* (75)[1]	9.25 (131)	-0.21 (2.8)	-0.15 (2.5)	0.16 (4.4)
1X3	-0.90 (70)	-5.67 (113)	0.40 (3.7)	0.44 (3.4)	0.01 (3.9)
1X4	2.85** (77)	3.17 (112)	0.08 (3.1)	0.26 (3.2)	0.05 (3.9)
1X5	2.02* (71)	-1.00 (116)	0.15 (3.6)	0.07 (3.3)	-0.08 (4.1)
1X6	-7.15** (61)[2]	-5.75 (112)	-0.42 (2.6)	-0.62 (2.3)	-0.14 (2.6)
2X3	-0.82 (64)	-2.12 (112)	-0.15 (3.5)	-0.11 (2.9)	-0.08 (3.1)
2X4	-1.73 (66)	-3.08 (112)	0.23 (3.7)	0.28 (3.3)	-0.20* (2.4)
2X5	-3.23** (63)	-2.92 (111)	0.07 (3.9)	-0.21 (3.1)	0.21* (5.1)
2X6	2.60* (64)	-0.33 (114)	0.06 (3.5)	0.20 (3.2)	-0.09 (3.8)
3X4	0.85 (69)	1.00 (113)	-0.33 (3.4)	-0.67 (2.7)	-0.04 (3.6)
3X5	0.68 (64)	2.17 (113)	0.68 (3.8)	0.67 (4.3)	-0.05 (4.6)
3X6	0.18 (62)	5.42 (117)	-0.60 (3.1)	-0.32 (3.1)	0.17 (4.0)
4X5	-2.90* (63)	0.00 (111)	-0.92* (2.9)	-0.57 (3.1)	0.03 (4.6)
4X6	0.93 (66)	-1.08 (110)	0.94* (4.4)	0.70 (4.1)	0.16 (3.8)
5X6	3.43** (63)	1.75 (112)	0.02 (3.8)	0.04 (3.7)	-0.10 (3.5)
SE	2.01	4.62	0.40	0.38	0.09

1 = numbers in brackets are cross means; 2 = numbers underlined are favourable specific heterotic effects; * = significant at $p \leq 0.05$; ** = significant at $p \leq 0.01$.

3.3.4 Correlation of fruit yield components

Correlation analysis helps in identification of selection aids in an improvement programme. Additionally attention to correlation between traits is important to avoid unfavourable correlated response. A correlation coefficient between two characters that is very high suggests a strong heritable association and possibly a narrow sense base (Gaines, 1991).

Significant positive correlations were obtained between the length of the primary vine and the number of primary branches (r_p = 0.45**), between the days to female anthesis and the days to the first fruit to mature (r_p = 0.58**), between the weight of the first fruit to mature and the mean fruit weight (r_p = 0.89**), between the mean fruit weight and the fruit yield per plant (r_p = 0.39*), and the between fruit number per plant and the fruit yield per plant (r_p = 0.85**) (Table 3.5). The most important characters in selection for improved fruit yield are mean fruit weight and fruit number per plant. This was consistent with previous reports (Gopalakrishnan *et al.*, 1980; Chigwe 1991; Gwanama *et al.*, 1998). Lack of correlation between yield and plant habit traits (vine length and branch number) makes it possible to select for more compact plants without affecting the yield. However, due to the low genotypic variance for these traits little gain from selection would be realised.

There were significant negative correlations between the days to female anthesis and the number of fruits per plant (r_p= -0.45**), between the days to female anthesis and the fruit yield per plant (r_p = -0.44*), between the days to the first mature fruit and mean fruit weight

(r_p = -0.25*), between the days to the first mature fruit and the number of fruits per plant (r_p = -0.45**) and the between the days to the first mature fruit and fruit yield per plant (r_p= -0.51**) (Table 3.5). These negative correlations to fruit yield are favourable: selection for early maturity would result in higher yields higher yields. Nonetheless the negative correlation of early maturity with mean fruit weight is not favourable and would eventually reduce fruit yield.

Table 3.5 Phenotypic correlations (r_p)between fruit yield components of *C. moschata*

Trait	Primary vine length	Days to female anthesis	Days to first fruit	First fruit weight	Mean fruit weight	Fruit number	Fruit yield
Primary branches (No.)	0.45**	-0.11	0.13	-0.08	-0.05	0.00	-0.05
Primary vine length		-0.18	0.00	-0.14	-0.09	0.11	0.08
Days to female anthesis			0.58**	-0.06	0.03	-0.45**	-0.44**
Days to first fruit				-0.29*	-0.25*	-0.45	-0.51**
First fruit weight					0.89**	0.11	0.39**
Mean fruit weight						-0.08	0.26*
Fruit number							0.85**

3.3.5 Heritability of fruit yield components

The objective of the heritability estimations was to give insights as to the level of additive gene action involved (since this and inbreeding could not be isolated without the selfs of the varieties). Heritability values for the characters with significant genotypic mean squares ranged from 0.13 to 0.45. Significant heritability estimates were obtained for days to the first fruit to mature (h^2_N =0.42*) and the soluble solids (h^2_N =0.45*) (Table 3.6). This finding was consistent with the genetic effects analysis, especially for the days to the first mature fruit where additive gene action was the only component of relevance. These two traits are likely to be the most amenable to recurrent selection. The low heritabilities of the other traits may be attributable to insufficient testing. When estimated by regression methods, low or negative heritability values may result from inadequate testing procedures or insufficient samples (Hallaeur & Miranda, 1981). The likelihood of insufficient samples in this study is reflected by high standard errors. Falconer (1960) presents mathematical derivation to show that a large number of families or individuals within families is needed to obtain realistic standard errors. Standard errors would also be reduced by testing in a large number of environments.

Table 3.6. Heritability (h_N^2) and standard errors of heritability for fruit yield components and fruit quality characters of *C. moschata*.

Trait	h_N^2	SE	h^2/SE
Days to female anthesis	0.25	0.34	0.74
Days to first fruit to mature	0.42*	0.18	2.33
Weight of first fruit to mature	0.00	-	-
Mean fruit weight	0.13	0.27	0.48
Soluble solids	0.45*	0.12	3.75

* = h_N^2 significant if $h^2 \geq 2$ SE

3.4 **CONCLUSION**

The results of this study show that both additive and non-additive gene action are involved in the inheritance of the time to female anthesis, the weight of the first fruit to mature, the mean fruit weight and the soluble solids. Only additive gene action is important for the time to the first mature fruit. There was inadequate genetic variance for the other characters related to yield or quality. Varieties showing superior performance have been identified. Since the trial was carried out in one season and at a single site, all the mean squares and covariances obtained are confounded with genotype-environment interactions. Evaluations involving more environments will be necessary. The lack of published data for this species makes comparison impossible but leaves an opportunity for validation of these findings.

CHAPTER 4

GENERAL CONCLUSION

The first part of this study (Chapter 2) established the existence of adequate variation for breeding among the land races. The genetic distances were comparable with those found in germplasm in current use in other species of the Cucurbitaceae, such as watermelon (*Citrullus lanatus*) and melon (*Cucumis melo*). The use of RAPD markers proved a quick aid in organising the germplasm. In the absence of phenotypic characterisation, nonetheless, direct use of this germplasm, is still hampered by the lack of correlation of the classification with phenotypic characters. The classification would still be useful in that a reduced number of land races would be chosen for evaluation. Divergent genotypes would be ideal for inclusion in the evaluation; those land races in each cluster that are widely used would be given preference.

The land races from Zambia were found to be closely related, but sufficiently diverse for use in breeding. This was confirmed in the analysis of crosses and progeny (Chapter 3) where it was seen that variation was high in some traits and insignificant for most traits. Land race LP4, which was the only land race not from the same cluster, generally performed differently from the others, especially for crop maturity. This underscores the value of divergent genotypes. The collection from

Malawi was more diverse and would be expected to show greater variation and heterosis. Crosses between genotypes from Malawi and those from Zambia would be the most promising since there was correlation between geographical source of accessions and position in the dendrogram.

Further characterisation with DNA based markers is recommended. Progeny of divergent genotypes, as well as their parents could be characterised with a large number of primers to detect genotype-specific markers. These genotype-specific markers would be a basis of marker assisted selection. Traits that have been seen to vary among genotypes could be correlated with presence of markers to test for linkage.

This study has shown that some yield traits and soluble solids could be improved by selection. The understanding of genetic effects could be expanded by inclusion of other types of populations (selfs of parents, crosses selfed, crosses open pollinated, etc.) to resolve all types of genetic effects. The identification of parents with significant variety effects could also be exploited by developing inbreds from those parents. These could be the basis for genetic variance analysis by random effects models. Availability of inbreds would also facilitate linkage analysis. These various options would help in crop improvement as well as genetic analysis of tropical pumpkin.

CHAPTER 5

GENERAL SUMMARY

Tropical pumpkin (*Cucurbita moschata* Duchsne) is one of the most important vegetable crops in Africa. Apart from its phylogenetic relationships with other members of the genus, little information is available for this species. This study explored the amount of genetic variation in Zambia and Malawi for possible use in breeding. UPGMA cluster analysis from 39 polymorphic and 105 monomorphic DNA fragments amplified by 16 RAPD markers showed four clusters, with genotypes from Malawi mainly clustering in three clusters while all genotypes from Zambia and three from Malawi clustered in one cluster. Six land races from this classification were mated in a daillel fashion and parents and progeny evaluated for fruit yield and quality. Variety effects were significant for the time to female anthesis and first fruit maturity, the weight of the first fruit, mean fruit weight and soluble solids. These characters could be improved by selection. Only for soluble solids were the progeny as a set different from the set of parents. Both variety and specific heterosis were significant FOR some of these traits, offering opportunities for heterosis breeding.

ALGEMENE OPSOMMING

Tropiese pampoen (*Cucurbita moschata* Duchsne) is een van die mees belangrike groente gewasse in Afrika. Afgesien van die filogenetiese verwantskappe met ander lede van die genus, is daar min inligting beskikbaar van hierdie spesie. Hierdie studie het 'n ondersoek gedoen na die hoeveelheid genetiese variasie in Zambië en Malawi vir moontlike gebruik in teling. UPGMA groep analise van 39 polimorfiese en 105 monomorfiese DNA fragments gearnplifiseer deur 16 RAPD merkers het vier groepe aangetoon, met alle genotipes van Zambië en drie van Malawi in een groep. Ses landrasse van hierdie klassifikasie is gekruis in 'n dialleel analise en die ours en nageslag is geëvalueer vir vrug opbrengs en kwaliteit. Variëteits effekte was betekenisvol vir die tyd tot vroulike antese en eerste vrug rypwording, die gewig van die eerste vrug, die gemiddelde vrug gewig en die oplosbare vastestowwe. Hierdie eienskappe kan deur seleksie verbeter word. Net vir oplosbare vastestowwe was die nageslag verskillend van die stel ouers. Beide variëteit- en spesifieke heterose was betekenisvol vir sommige van hierdie eienskappe, wat moonlikhede skep vir baster teling.

REFERENCES

Andren, U., A. Nkomesha, L.P. Singongo & A. Sutherland. 1991. National seed availability study. Adaptive Research Planning Team, Ministry of Agriculture, Lusaka, Zambia.

Baker, R.J. 1978. Issues in diallel analysis. Crop Sci. 18: 533-536.

Carbonell, M.E., B..L Wessel, F Verela. & B. Luciano. 1990. Pumpkin *(Cucurbita moschata)* breeding priorities based on a survey of Puerto Rican consumers. J. Agric. Univ. Puerto Rico 74: 229-236.

Casler, M.D. 1981. Genotype X environment interaction bias to parent offspring regression heritability estimates. Crop Sci. 22: 540-542.

Caetano-Anolles, G. B.J. Bassam & P.M. Gressholf. 1991. DNA amplification using very short arbitrary oligonuceoltide primers. Biotechnology 9: 553-557.

Chigwe, C.F.B. 1991. Evaluation of Malawi pumpkin (*Cucurbita moschata* Duch ex Poir) germplasm collection. Progress Report, Project No. SRG/39/5. Bunda College of Agriculture, Lilongwe, Malawi. Pp 1-32.

Cockerham, C.G. 1963. Estimation of genetic variances. In: W.D Hanson and H.F. Robinson (ed.) Statistical genetics and plant breeding. Natl. Acad. Sci.- Natl. Res. Council. Washington, DC. Pp53-93.

Cukadar-Olmedo, B., J.F. Miller & J.J. Hammond. 1997. Combining ability of the stay green trait and seed moisture content in sunflower. Crop Sci. 37: 378-382.

dos Santos, J.B., J. Nienhuis, P. Skroch, J. Tivang & M.K. Slocum. 1994. Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. Theor. Appl. Genet. 87: 909-915.

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

Eberhart S.A. & C.O. Gardner. 1966. A general model for genetic effects. Biometrics 22: 864-881.

Falconer, D.S. 1960. Introduction to quantitative genetics. Longman, London.

Frankel, R. & E. Galun. 1977. Pollination mechanisms, reproduction and plant breeding. Springer-Verlag, Berlin/Hielderberg/New York.

Frey, K.J. & T. Horner. 1957. Heritability in standard units. Agronomy J. 49: 59-62.

Gaines, C.S. 1991. Associations among quality attributes of red and white soft wheat cultivars across locations and crop years. Cereal Chem. 68: 56-59.

Garcia-Rodriguez, E., I. Alvarez & R. Lozano. 1996. Random amplified polymorphic DNAs (RAPDs) as markers to determining genetic relationships

among *Cucumis melo* L genotypes. Proceedings of the VIth Eucarpia meeting on cucurbit genetics and breeding, 26 - 29 May, 1996, Malaga, Spain. Pp 186-193.

Gardner, C.O. & S.A. Eberhart. 1966. Analysis and interpretation of the variety cross diallel and related populations. Biometrics 22: 439-451.

Gopalakrishnan, T.R., P.K. Gopalakrishnan & K.V. Peter. 1980. Variability, heritability and correlation among some polygenic characters of pumpkin. Indian J. Agric. Sc. 50: 925-30.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrica 53: 325-338

Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Austr. J. Biol. Sc. 9: 463-493.

Gwanama, C., M.S. Mwala & K. Nichterlein. 1998. Path analysis of fruit yield components of *Cucurbita moschata* Duch. Trop. Agric. Res. Ext. 1: 19-22.

Gwanama, C. & K. Nichterlein. 1995. Importance of cucurbits to small-scale farmers in Zambia. Zambian J. Agric. Sc. 5: 5-9

Hallaeur, A.R & J.B. Miranda. 1981. Quantitative genetics in maize breeding. Iowa State Univ. Press, Ames, USA. Pp 45-113.

Hallden, C., I.N.O. Nilson, I.M. Rading & T. Sall. 1994. Evaluation of RFLP and RAPD markers in a comparison of *Brassica napus* breeding lines. Theor. Appl. Genet. 88: 123-128.

Hayman, B.I.1954. The theory and analysis of the diallel cross. Genetics 39: 789-809.

Jacquard, A. 1983. Heritability: One word, three concepts. Biometrics 39: 465-477.

Jensen, N.F. 1970. A diallel selective mating system for cereal breeding. Crop Sci. 10: 629-635.

Lee, S.J., J.S. Shin, K.W. Park & Y.P Hong. 1995. Detection of genetic diversity using RAPD-PCR and sugar analysis in watermelon *[Citrullus lanatus* (Thumb.) Mansf.] germplasm. Theor. Appl. Genet. 92:719-725.

Maynard, D.N 1996. Performance of tropical pumpkin inbreds and hybrids. Proceedings of the VIth Eucarpia meeting on cucurbit genetics and breeding, 26 - 29 May, 1996, Malaga, Spain. Pp: 59-65.

McDonald, M.B., L.J. Elliot & P.M. Sweeney. 1994. DNA extraction from dry seeds for RAPD analysis in varietal identification studies. Seed Sci. Technol. 22: 171-176.

MSTAT-C. 1991. A software program for the design, management and analysis of agronomic research experiments. Michigan State University, USA.

Mumm, R.H. & J.W. Dudley. 1994. A classification of 148 maize inbreds: I. Cluster analysis based on RFLPs. Crop Sci. 34: 842-852.

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

Patil, V.D. & P.R. Chopde. 1983. Heterosis in relation to general and specific combining ability effects in linseed. Indian J. Genet. 43: 226-228

Pooni, H.S., J.L. Jinks & R.K. Singh. 1984. Methods of analysis and the estimation of the genetic parameters from a diallel set of crosses. Heredity 52: 243-253.

Prixely, K.V. & M.S. Bjarnason 1993. Combing ability for yield and protein quality among Modified Endosperm *opaque-2* maize inbreds. Crop Sc. 33: 1229-1234.

Quendeba, B., G. Ejeta., W.E. Nyquist., W.W. Hanna. & A. Kumar. 1993. Heterosis and combining ability of African pearl millet landraces. Crop Sci. 33: 735-739.

Robinison, H.F., R.E. Comstock & P.H. Harvey. 1949. Estimates of heritability and the degree of dominance in corn. Agronomy J. 41: 353-359.

Sneath, P.H.A & R.R. Sokal. 1973. Numerical taxonomy. W.H. Freeman and Co. San Francisco.

Southgate, D.A.T. 1986. Determination of food carbohydrates. Applied Science Publishers Limited, London.

Sprague, G.F. & L.A.. Tatum. 1942. General versus specific combining ability in single crosses of corn. J. Amer. Soc. Agron. 43: 923-932.

Staub, J., F. Serquen & M. Gupta. 1996. Genetic markers, map construction and their application in plant breeding. HortScience 31: 729-741.

Stuber, C.W. 1980. Mating designs, field layouts, and breeding records. *In*: W.R. Fehr & H.H. Hardley (ed). Hybridisation of crop plants. Am. Soc. Agronomy and Crop Sci. Soc. Am. Madison, Wisconsin, USA. Pg 80-101.

Sughroue, J. R. & A.R. Hallauer. 1997. Analysis of the diallel mating design for maize inbred lines. Crop Sci. 37: 400-405.

Tanhuanpää, P.K., J.P. Vilkki & H.J. Vilkki. 1995a. Association of a RAPD marker with linolenic acid concentration in the seed oil of rapeseed (*Brassica napus* L). Genome 38: 414-416.

Tanhuanpää, P.K., J.P. Vilkki & H.J. Vilkki. 1995b. Identification of a RAPD marker for palmitic acid concentration in the seed oil of spring turnip rape (*Brassica rapa* ssp *oleifera*). Theor. Appl. Genet. 91: 477-480.

Thormann, C.E., M.E. Ferreira, L.E.A. Camargo, J.G. Tivang, J.G. & T.C. Osborn. 1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88: 973-980.

Welsh, J. & M. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218.

Welsh, J., C. Petersen & M. McClelland. 1991. Polymorphisms generated by arbitrarily primed PCR in the mouse: application to strain identification and genetic mapping. Nucleic Acids Res. 19: 303-306.

Westman, A.L & S. Kresovich. 1997. Use of molecular marker techniques for description of plant genetic variation. *In*: Callow, J.A, B.V. Ford-Lloyd and H.J. Newbury (eds). Biotechnology and plant genetic resources. CAB International, UK. Pp 9-48.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski. & Tingey, S.S. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531-6535.

Wilson, H.D. 1989. Discordant patterns of allozyme and morphological variation in Mexican *Cucurbita*. Systematic Botany 14: 612 - 623.

Wright, A.J. 1985. Diallel designs, analyses, and reference populations. Heredity 54: 307-311.

APPENDICES

Appendix 1

List of *Cucurbita moschata* accessions employed in the study

Entry number	Accession number	Source
1	CPC2	Malawi
2	CPC30	Malawi
3	CPC55	Malawi
4	CPC74	Malawi
5	CPC77	Malawi
6	CPC89	Malawi
7	CPC91	Malawi
8	CPC102	Malawi
9	CPC114	Malawi
10	CPC119	Malawi
11	CPC122	Malawi
12	CPC128	Malawi
13	CPC139	Malawi
14	Lame pumpkin	Zambia
15	CM15	Zambia, Southern Province
16	CM21	Zambia, Southern Province
17	EP001	Zambia, Eastern Province
18	EP002	Zambia, Eastern Province

Continued

Appendix 1 continued

Entry number	Accession number	Source
19	EP010	Zambia, Eastern Province
20	EP014	Zambia, Eastern Province
21	EP021	Zambia, Eastern Province
22	EP023	Zambia, Eastern Province
23	EP026	Zambia, Eastern Province
24	LP4	Zambia, Luapula Province
25	LP15	Zambia, Luapula Province
26	MK5	Zambia, Central Province
27	MK7	Zambia, Central Province
28	Munansangu	Zambia, Lusaka Province
29	WP002B	Zambia, Western Province
30	WP007	Zambia, Luapula Province
31	WP012	Zambia, Western Province

APPENDIX-2

RAPD MARKER MATRIX USED TO GENERATE PROXIMITY MATRIX FOR UPGMA CLUSTER ANALYSIS

1 = Presence of DNA fragment, 0 = Absence of fragment, N = missing data

CPC2

01111111110111101111011111111111111101111111110110111111000111111101
11011111011011101111101011010111011111110101111111111111111111111111
111111111111

CPC30

01111111110011001111111111111111111111111111110110111111100111111111
10011111011111111111101011010111011111111111111111111111101111111111
111111111111

CPC55

11101111110011001111011111111111011011111111001101111110101111111111
10011111111011101011101011010111011111101NNNNNNNNN11111110111
1111111111101111

CPC74

11111111110111101111011111111111011111111111011111111100011111101
11011111011111111111101011011111011111111111111111111111111111111111
111111101111

CPC77

11111111110011101111111011111111011111111111101101111110001111111111
11011111011111101011101011011111011111111101111111111NNNNNN1111
111111111101111

CPC89

1110111111001110111111101111111111111011111110011011111111011111101
1001111101101111101110101101111101111111111111111111111111101111111
111111101111

CPC99

1110111111001110111101111111111111111111111111011011111100011111101
1001111101111110111110101101111111111111010111111111111111101111111
111111101111

CPC102

1110111111011100111101101111111101110111111100110111111000111111101
1001111111101110101110101101011101111111111111111111111111111111111
111111101111

CPC 114

0111111111011110111111111111111101110111101100111111111000111111101
1001111111111110111111111101011111111111011111111111111111111111111
111111101111

CPC119

1110111111001111111111111111111101111111011101101111111101111111101
1001111101111111111110101101011101111111111111111111111111101111111
111111111111

CPC122

1110111111001101111101101111111111111111011101101111110001111110101
1011111101101110101110101101011101111111011111111111111111101111111
111111111111

CPC125

1110111111001110111111111111111111110011111110011011111110011111101
1001111101111110101110101101011101111111111111111111111111101111111
111111101111

CPC139

1110111111001101111101111111111111111111101101110111111111111111101
1111111101111110101110101101011101111111111111111111111111101111111
111111101111

CM15

1110111111001100111111101111111111110111111100110111110100111111101
1001111101101110101110101101011111111111111111111111111111101111111
111111111111

CM21

0111111111001100111111111011N111NNNN1101111111001101111110001111 1
1011001111101111110101110101101011111111111101111111111111111101111
111111111111111

Lame pumpkin

0110111111001100111101101111111N1111111111110110111101001111111 10
1100111111111111010111010110101110111111111111111111111111110111111
111111111111

EP001

0111111111001100111111101111111111110111111100111111101000111111111
1001111111111111101110101101011101111111011111111111111111101111111
111111111111

EP002

1111111111001100111111101111111111110111111100110111101000111111101
1001111101111110101110101101011101111111111111111111111111101111111
111111101111

EP010

01111111110011001111111011111111111101111111101101111010001111111
10011111011111101011101011010111111111111111111111111111111101111111
111111101111

EP014

11101111110111001111111011111111111111111111111101111010001111110 1
10011111111011101011101011010111111111101111111111111111101111111
111111101111

EP021

11101111110011001111111011NNNNNNNN10011111110111111110100111 11
11011001111101111110101110101101011101111111111111111111111110111
111111111101111

EP023

01101111110011001111111011NNNN11NN10011110111011011111100111111
10110011111011111111011111011010111110111111011111111111111101001
111111111101111

EP026

11101111110011001111111011NNNNNNNN10011111110011011111100011111
11011101111101111110101110101101011111111111010111110111111110111
111111111111111

LP4

01101111111011101111111011111111011001111011001101111011111111111
10011111011111101011101011010111011111101011111111111111101111111
111111101111

LP5

11111111110011001111111011111111011101111111001101111010001111110 1
10011111011111101011101011010111011111111111111111111111101111111
111111111111

MK5

01111111110011001111111011111111111101111111001111111010001111110110011111011111101011101011010111011111011111111111111111011111111111111101111

MK7

NNNNN11111001100111111101111111101111111011111101111110011111110110011111011101011010111011111111111111111111111111111011111111111111

Munasangu

1111111111001100111111111011NNNNNNNNN11011111110111011110100011111101100111101111101011010111111111111111111111111111111101111111111111111

WP002B

11111111110111001111111111111111111101111111011101111110011111110100111110111101011010111111111111111111111111011111111111111

WP007

11111111111011001111111111111111111101111N11001111111110001111110110011111011111101011010111111111111111111111111010111111111111111

WP012

11101111110011001111111111111111111101111111001101111010001111110100111110111101011010111111111111111111111111011111111111111

Appendis 3

Estimators of heritability (σ_{OP}/σ^2P) from the regression of full-sib means on mid-parent means, where full-sibs and mid-parents are evaluated under varying circumstances (Adapted from Casler (1981)).

Estimation procedure	Parent and offspring evaluation for estimating heritability	Estimator of $\sigma_{OP}/\sigma^2{}_P$*
1	Same replicate, location and year	$(\hat{\sigma}^2{}_{PO} + \hat{\sigma}^2{}_{GL(PO)}/l + \hat{\sigma}^2{}_{wPO}/rl + \hat{\sigma}^2{}_{GY(PO)}/y + \hat{\sigma}^2{}_{GLY(PO)}/ly + \hat{\sigma}^2{}_{ePO}/rly)/\hat{\sigma}^2{}_P$
2	Different replicates; same location and year	$(\hat{\sigma}^2{}_{PO} + \hat{\sigma}^2{}_{GL(PO)}/l + \hat{\sigma}^2{}_{GY(PO)}/y + \hat{\sigma}^2{}_{GLY(PO)}/ly)/\hat{\sigma}^2{}_P$
3	Different locations, same year	$(\hat{\sigma}^2{}_{PO} + \hat{\sigma}^2{}_{GY(PO)}/y + \hat{\sigma}^2{}_{GLY(PO)}/ly)/\hat{\sigma}^2{}_P$
4	Different replicates and years, same location	$(\hat{\sigma}^2{}_{PO} + \hat{\sigma}^2{}_{GL(PO)}/l + \hat{\sigma}^2{}_{GLY(PO)}/ly)/\hat{\sigma}^2{}_P$
5	Different years; same replicate and location	$(\hat{\sigma}^2{}_{PO} + \hat{\sigma}^2{}_{GL(PO)}/l + \hat{\sigma}^2{}_{wPO}/rl + \hat{\sigma}^2{}_{GLY(PO)}/ly + \hat{\sigma}^2{}_{ePO}/rly)/\hat{\sigma}^2{}_P$
6	Different years and locations	$\hat{\sigma}^2{}_{OP}/\hat{\sigma}^2{}_P$

* $\sigma^2{}_P = \hat{\sigma}^2{}_p + \hat{\sigma}^2{}_{PL}/l + \hat{\sigma}^2{}_{wP}/rl + \hat{\sigma}^2{}_{PY}/y + \hat{\sigma}^2{}_{PLY}/ly + \hat{\sigma}^2{}_{ep}/rly$