GENETIC DIVERSITY IN UPLAND COTTON

(Gossynium hirsutum L.)

By

KLASSOU, Celestin

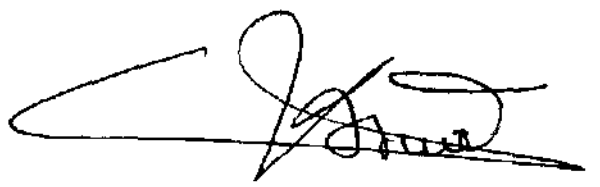
A thesis submitted to the Postgraduate School,
Ahmadu Bello University in partial fulfilment of
the requirements for the degree of Master of Science
in Crop Breeding

Department of Plant Science
Faculty of Agriculture, Ahmadu Bello University,
Zaria, NIGERIA

March, 1998.

Declaration

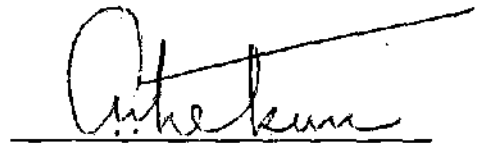
I hereby declare that this thesis has been written by me and that it is a record of my own research work. It has not been presented before in any previous application for higher degree.



C. Klassou
(Student)

28/5/98
Date

The above declaration is confirmed



Dr. C.A. Echekwu
(Chairman, Supervisory Committee)

29-07/98
Date

Certification

This thesis entitled "Genetic Diversity in Upland Cotton (*Gossypium hirsutum* L.) by Célestin KLASSOU meets the regulations governing the degree of Master of Science of Ahmadu Bello University, Zaria and is approved for its contribution to scientific knowledge and literary presentation.

..... Date 29/07/98

Dr C.A. ECHEKWU
(Chairman, Supervisory Committee)

..... Date 29/7/98

Dr (Mrs) P.E. OLORUNJU
(Member, Supervisory Committee)

..... Date 29/7/98

Dr S.O. ALABI
(Member, Supervisory Committee)

..... Date 29/7/98

Dr S.O. ALABI
(Head, Department of Plant Science)

..... Date 17/2/99

(Dean, Postgraduate School)

AHMADU BELLO UNIVERSITY
ZARIA

Dedication

This thesis is dedicated to the memories of my late father and mother, Mr NADJANSOU NGUIRANA TOBA and Mrs H. NADJANSOU née TCHOUTTA HARAINA.

Acknowledgement

I wish to express my sincere and profound appreciation to my major suprevisor, Dr. C.A. Echekwu. His keen interest in this study, right guidance coupled with thorough and prompt reading of the manuscripts left me with a great lesson to learn in his supervisory role to graduate students. His entire family is deeply acknowledged for their constant assistance. I am deeply indepted to the other Members of the Supervisory Committee; Dr.(Mrs.) P.E. Olorunju and Dr. S.O. Alabi. Their encouragement and advice to me to get this work done remain in my memory.

I wish to recognise and acknowledge the patience and understanding of all my lecturers and the entire staff of the Department of Plant Science, especially the late Professors C.C. Nwasike and L.B. Olugbemi who are here-in remembered.

I appreciate the patience and mutual support of all my classmates in the Department during this long academic session at the Ahmadu Bello University. I extend my special regards to C.C. Okonkwo and his entire family, Mrs Omonlara Adjadi, Jacques Beyo and S.T. Duala.

I also wish to thank the Cameroon Government through the Institute of Agronomic Research (IRA) for giving me a study leave for this Postgraduate study. I express my heartfelt appreciation and thanks to Dr. J.A. Ayuk-Takem, and Dr. J. Bindzi-Tsala, Director and Deputy Director of IRA respectively for their understanding and their contribution towards obtaining a sponsorship for me. My thanks and appreciation are extended to Garoua II Project (PGII) Sterring Committee and particularly S. Nzietchueng and Seiny Boukar Lamine, former and current PGII Coordinators for the financial and managerial support. The collaboration of my

colleagues J.Ekorong, B. Bachelier, and F.Bernard, research fellows at IRA, Cameroon is highly appreciated.

My appreciation goes to the "Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Département des Cultures Annuelles" (CIRAD-CA), France, for providing me with the Upland cotton germplasm and their assistance in the data processing. I extend my special regards to J. Gouthiere, E.Gozé, J. Chaume, B. Hau, A. Joly and P. Lethourmy.

I am very grateful to R. Okechukwu, IITA, Ibadan and the entire staff of the computer unit, Grain Legumes Improvement Programme (GLIP) for their assistance in processing and analysing the data.

Sincere gratitute is offered to Dr. F.B. Alisa for his continuous encouragement, and to all the other people too many to enumerate for the assistance rendered in data collection, fibre quality tests, preliminary data records, and the typing of this project report

I sincerely thank my wife, C.L. Klassou (née) Irana Clarisse Lamana, for enduring my long absence from home, taking care of our lovely children, Hervé Dahtoudi, Aurélie N'Yoldah and Maxime Goudongsou. I hope the experience she got during my absence will be benefical for our future. Her encouragement as well as tolerance and support of all the members of our family, nuclear and extended, are acknowledged.

I am profoundly indebted to all those friends and relatives who during my absence, have given some forms of help to my wife and family at home.

Finally and above all may Glory, Honour and Thanks go to the Almighty God.

ABSTRACT

Two hundred and fourty seven genotypes of upland cotton were grown at Maroua and Garoua (Cameroon), and Samaru (Nigeria), in 1994 to (i) investigate the nature and extent of their genetic diversity using Principal Component and Cluster Analyses, (ii) assess the relative contribution of different traits to their total variation, and (iii) determine the relation between the clustering patterns and the geographic distribution of these accessions.

A 16 x 16 partially balanced lattice experimental design, with four replicates, was used for this study. Plots consisted of single rows 5m long, spaced 1m apart at Maroua and Garoua, and 0.91m at Samaru with a plant to plant spacing of 0.40m. Data were collected on five competitive plants, wherever possible, for 25 agronomic and fibre technology traits. Analysis of variance for each trait was performed on plot mean basis.

In the combined analysis of variance across locations, all the upland cotton genotypes were significantly different for all the traits evaluated. The location x genotype effects were highly significant for all the traits measured and their patterns varied appreciably over environments, indicating differential response of different traits in varying environments.

At each location, significant variability for all the traits observed was also found among the genotypes of upland cotton selected. The seed cotton per plant on average was low at Samaru (43.5g), presumably due to late planting period, medium at Maroua (70.6g), and high at Garoua (I99g).

Simple linear correlations, principal component and cluster analyses were performed on adjusted treatment means. At Samaru, seed cotton per plant was positively associated with reproductive traits, boll number, number of flowers per plant, but negatively correlated with

(lays to first and 50% flowers opening (presumably due to an important shedding of young capsules), and fibre length. At Maroua coefficients of correlation were moderate.

At Garoua strong associations were observed between seed cotton yield, plant height, number of monopodia and sympodia, flower and boll number, percent fibre and 2.5% Span Length.

Two-dimensional scatter diagrammes displayed results of the principal component analyses. There was a poor resolution of individuals into groups due to the fact that the two principal components captured only about 30% of the total variation. Two large groups were formed. All genetic materials originating from African countries grouped together and showed good adaptation to the environments at the test sites.

The cluster arrangements delineated more precisely the individuals into 8, 7, and 6 groups at Samaru. Maroua, and Garoua, respectively. Despite some discrepancies, in general. cluster patterns were in agreement with pedigree information

The intercluster divergence measured by the Mahalanobis' D distance varied from 1.8 to .3.8, indicating a moderate to low degree of diversity within the genetic materials studied at each location.

TABLE OF CONTENTS

TITLE PAGE i

DECLARATION ii

CERTIFICATION iii

DEDICATION iv

ACKNOWLEDGEMENTS . . . . . v-vi

ABSTRACT . . . . . vii-viii

TABLE OF CONTENTS . . . . . ix-xi

LIST OF TABLES . . . . . xii-xiii

LIST OF FIGURES xiv

LIST OF APPENDIX xv

CHAPTER ONE

INTRODUCTION 1

CHAPTER TWO

LITERATURE REVIEW 4

2.1 Genetic Diversity in Crop Plants 4

2.2 Measurement of Genetic Diversity 5

2.3 Use of Multivariate Analysis 6

2.3.1 Use of Principal Component and Cluster Analyses in Plant Breeding 8

2.3.2 Classification 11

2.4 Traits Selected and Inter-relations 14

CHAPTER THREE

MATERIALS AND METHODS 16

3.1 Experimental Materials 16

3.2 Experimental Technique 16

3.3 Observations 17

3.4 Statistical Analysis 19

3.4.1 Analysis of Variance 19

3.4.2 Correlations Among Traits, and Principal Compnent Analysis 23

3.4.3 Classification 23

CHAPTER FOUR

RESULTS 24

4.1 Mean Performance 24

4.2 Simple Linear Correlations 30

4.3 Principal Component Analysis 35

4.3.1 Samaru Location 35

4.3.2 Maroua Location 38

4.3.3 Garoua Location 45

4.4 Hierarchical Clustering 49

4.4.1 Clustering at Samaru 50

4.4.2 Clustering at Maroua 52

4.4.3 Clustering at Garoua 54

4.5 Inter-Cluster distances 56

CHAPTER FIVE

DISCUSSION 60

CHAPTER SIX

SUMMARY AND CONCLUSION 68

REFERENCES 70

APPENDIX 80

LIST OF TABLES

Table 1: Traits observed for the evaluation of 247 Upland Cotton Genotypes and Measurements 18

Table 2: Form of analysis of variance with expected mean squares (EMS) for each trait in one location 21

Table 3: Form of analysis of variance with expected mean squares (EMS) for each trait accross the three locations 22

Table 4: Means, Standard errors of mean (S.E.$\bar{X}$), Ranges and coefficients of variaton (C.V%) for 247 genótypes of upland cotton grown at Samaru, Maroua, and Garoua in 1994 25

Table 5: Mean squares for the combined analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown in the three locations in 1994 27

Table 6: Estimates of genotypic variance (θ^2_G), phenotypic variance(θ^2_{ph}), and heritability (h^2) of the traits measured at each location 29

Table 7: Simple linear correlations among 22 traits of upland cotton grown at Samaru in 1994 31

Table 8: Simple linear correlations among 23 traits of upland cotton grown at Maroua in 1994 32

Table 9: Simple linear correlations among 23 traits of upland cotton grown at Garoua in 1994 33

Table 10: Principal component (PC) analysis of traits associated with 247 upland cotton genotypes grown in Samaru in 1994, showing eigen values and proportion of variation associated with the first five PC axes and eigen vectors of traits 36

Table 11: Principal component (PC) analysis of traits associated with 247 upland cotton genotypes grown in Maroua in 1994, showing eigen values and proportion of variation associated with the first five PC axes and eigen vectors of traits 42

Table 12: Principal component (PC) analysis of traits associated with 247 upland cotton genotypes grown in Garoua in 1994,

showing eigen values and proportion of variation associated with the first five PC axes and eigen vectors of traits 46

Table 13: Inter-cluster average values of D (Mahalanobis distance) for 247 genotypes of upland cotton grown at Samaru in 1994 57

Table 14: Inter-cluster average values of D (Mahalanobis distance) for 247 genotypes of upland cotton grown at Maroua in 1994 58

Table 15: Inter-cluster average values of D (Mahalanobis distance) for 247 genotypes of upland cotton grown at Garoua in 1994 59

LIST OF FIGURES

FIGURES

1. Scatter diagram for the major traits of the first two PC axes that described the variation of 23 traits measured in 247 genotypes grown at Samaru 39

2. Scatter diagram for 247 genotypes grown at Samaru in 1994 40

3. Scatter diagram for the major traits of the first two PC axes that described the variation of 23 traits measured in 247 genotypes grown at Samaru 43

4. Scatter diagram for 247 genotypes grown at Samaru in 1994 44

5. Scatter diagram for the major traits of the first two PC axes that described the variation of 23 traits measured in 247 genotypes grown at Samaru 47

6. Scatter diagram for 247 genotypes grown at Samaru in 1994 48

7. Dendrogram for 247 genotype of upland cotton grown at Samaru in 1994 51

8. Dendrogram for 247 genotype of upland cotton grown at Maroua in 1994 53

9. Dendrogram for 247 genotype of upland cotton grown at Garoua in 1994 55

LIST OF APPENDIX

APPENDIX

1. Upland cotton genotypes studied, abbreviations used in the figures, and source country 80

2. List of source countries, codes (FAO), and number of genotypes of upland cotton evaluated 85

3: Analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown at Samaru in 1994 86

4: Analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown at Maroua in 1994 87

5: Analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown at Garoua in 1994 88

6 - 13. Individual clusters at Samaru 89-96

14 - 19. Individual clusters at Garoua 97-102

20. Pedigree scheme of upland cotton materials created by or with the collaboration of the French Institute of cotton Research (IRCT) Montpellier, France 103

21. Pedigree scheme of some U.S. upland cotton materials 104

22. Pedigree scheme of some upland cotton materials from Paraguay, Argentina and Brazil 105

23. Pedigree scheme of some upland cotton varieties from Nigeria 106

CHAPTER ONE

INTRODUCTION

Cotton is harvested from about 35.2 million hectares, thus, occupying 25 per cent of the world's arable land, and 34 per cent of this area lay within the tropics. Of the 45 million tonnes of seed cotton produced (about 19-20 million tonnes of fibre) only 20 per cent come from the tropics (Lee, 1984; ICAC, 1996). Although most cotton is produced outside the tropics {>70% of the production by China (23.2%), U.S.A. (20.1%), India and Pakistan (21.8%), and former USSR (9.4%)}, cotton is of greater economic significance to tropical countries than the temperate ones, frequently being the major or sole source of foreign exchange (Hearn and Constable, 1984; Echekwu, 1991; Lançon, 1995). Although cotton is grown mostly for fibre, the seeds are also important. Cotton seed oil is used for culinary purposes, and oil cake residue is a protein-rich feed for ruminant livestock (Lee, 1984).

The cotton of commerce belongs to the order Malvales, family of *Malvaceae*, tribe of *Gossypieeae*, and the genus of *Gossypium*. The genus *Gossypium* consists of about 40 diploid species with 2n = 2x = 26 chromosomes, subdivided themselves into 7 genomes (A to G); and 4 allotetraploid species with 2n = 4x = 52 chromosomes (genome AD) (Fryxell, 1984; Fryxell *et al.*, 1992; Wendel and Albert, 1992; Stewart, 1994; Dessauw and Hau, 1997). Saunders (1961) speculated that the centre of origin of the genus Gossypium is in Central Africa because four of the seven diploid genome groups occur in the African continent.

There are four domesticated species of cotton. *Gossypium arboreum L.* and *G. herbaceum L.*, both diploids are native of the Old World. *Gossypium arboreum* remains an important crop in India, whereas *G. herbaceum*, important in earlier times, is today grown

mostly for local use in the drier areas of Africa and Asia. *Gossypium barbadense L.* and *G. hirsutum L.* evolved in the New World. Both are allotetraploids, or rather amphiploids in origin. *Gossypium barbadense*, commonly known as extra-long-staple, Egyptian, and Pima Cotton, and other names, supplies about 8% of the current world production of fibre. The fibre is used mostly for the production of luxury fabrics and sewing thread. *G. hirsutum* known most widely as Upland Cotton, contributes about 90 % of the current world production (Lee, 1984). Upland cotton fibres are used in manufacture of a variety of textile products, cordage, and other non-woven products (Lee, 1984).

Cotton is considered to be a normally self-pollinated crop, although natural crossing may range from zero to more than 50% (Niles *et al.*, 1984; Lançon, 1995). Cotton pollen is heavy and sticky and as a result, its transfer is effected by insects (Niles *et al.*, 1984). Control of pollination is relatively simple and hybridization is also relatively easy to accomplish. Intervarietal hybridizations, crosses between varieties or strains, have been used to develop many varieties (Demol *et al.*, 1992).

Selection of desirable forms from a wide array of available plant types and attributes is an important practice in plant breeding. For selection to be effective, useful germplasm with broad genetic base must be available (Hallauer, 1981; Mumm and Dudley, 1994; Bisht *et al.*, 1995). In self pollinated plant species such as cotton, germplasm is available in the form of homozygous lines which had been released as cultivars in specific ecological areas, and hybridization between these aimed at combining their best traits that are coded for by genes has been used to develop many of the improved varieties. The limitations resulting from normal pollination concerning biparental heredity necessitate a very careful selection of parent materials for the realization of success in hybridization programmes, especially when polygenic traits are involved. Selection of parents which are not closely related helps

in the creation of a population with a wide genetic base in which the increased amount of variability will make selection more efficient (Bhatt, 1970; Agarkova *et al.*, 1980; Arunachalam, 1980; Ariyo, 1987, 1990).

Assessment of genetic diversity or similarity among populations or species of plants may be based on analysis of pedigree and heterosis data (Cox *et al.*, 1985), biochemical and genetic markers like isozymes (Smith *et al.*, 1985a,b; Smith and Smith, 1987; Smith, 1984, 1988; Bernardo, 1983) and Restriction Fragment Length Polymorphisms (RLFPs) (Godshalk *et al.*, 1990) or quantitative traits (Gallais *et al.*, 1986; Onyilagha, 1986; Peeters and Martinelli, 1989; Brown, 1991; Souza and Sorrels, 1991a,b; Casler, 1995; Ouendeba *et al.*, 1995).

It is generally assumed by the plant breeders that cultivars originating from widely separated parts of the world are more likely to be genetically different. On this basis such cultivars are included in hybridization programmes in the hope that their presumed genetic diversity would provide a greater likelihood of promising genetic rearrangements (Bainiwal and Jatasra, 1980; Ziauddin *et al.*, 1980).

This study has been undertaken to:

1 - investigate the nature and extent of genetic diversity between Upland cotton genotypes from 42 countries across the world cotton producing regions, using Multivariate Analysis,

2 - assess the relative contribution of different characters to the total divergence,

3 - determine the relation between clustering patterns and geographic distribtuion of the genotypes.

CHAPTER TWÓ

LITERATURE REVIEW

2.1. Genetic Diversity in Crop Plants.

Genetic diversity can be considered as genetic difference or unlikeness, especially in related individuals or lines. It is a terminology utilized by the breeder to explain the extent of unrelatedness when evaluating crosses developed from different varieties; the progeny from such crosses are sometimes superior to progeny from crosses of inbred lines developed from a single variety.

N. I. Vavilov the Russian geneticist and plant breeder was the first, in 1926, to realise the essential need for a really broad genetic base for crop improvement although Smith *et al.* (1990), reported early inbreeding studies with maize (*Zea mays* L.) by Beal (1880) and Shull (1908) which had shown that there was a relationship existing between grain yield or grain yield heterosis and genetic diversity. Empirical studies by Moll *et al.* (1962, 1965) and Paterniani and Lonnquist (1963) showed a positive relationship between genetic distance, as measured by geographical distance, and both F_1 grain yield and grain yield heterosis in maize. Moll and Stuber (1974) reported that the comparison of inter- and intraspecific hybrids of alfalfa and cotton showed greater heterosis to be associated with greater diversity.

Maize breeders were the first to recognize the importance of genetic diversity in hybrid development programmes and routinely exploit this in the use of lines from different heterotic groups (Hallauer and Miranda, 1988).

The importance of genetic diversity existing in the parental material for enhancing the organization of breeding programmes has been emphasized by the works of several biologists on various farm crops (Murty and Arunachalam, 1966; Ziauddin *et al.*, 1980; Hawkes, 1981;

Gallais *et al.*, 1986; Melchinger *et al.*, 1991; Souza and Sorrels, 1991a,b). A great restriction in genetic base appeared to be accompanied by an increasing risk of genetic vulnerability (Duvick, 1984; Anonymous, 1997). Duvick (1984) recognized several sources of diversity available to today's farmers and plant breeders, (i) released cultivars and advanced elite lines, (ii) breeding pools, and (iii) gene banks. Thus, genetic diversity in crop plants is provided in other and less obvious ways.

2.2. Measurement of Genetic Diversity.

The genus *Gossypium* has a wide geographical distribution and shows large morphological and cytogenetical variations since it is made up of diploid and allotetraploid species. The events at the origin of the amphidiploid species generated a wealth of genetic diversity that exists primarily in *Gossypium hirsutum* (Duvick, 1984; Endrizzi *et al.*, 1984; Fryxell, 1984). However, the efficient use of this large genetic heterogeneity can only be optimized if this diversity is evaluated, quantified and classified in some way (Vanderborght and Depiereux, 1987; Amurrio *et al.*, 1995). Liu (1979), considered genetic distance as a measure of genetic divergence. The objective measurement of similarity (or distance) between populations has concerned geneticists, taxonomists, anthropologists, and plant and animal breeders for the last century (Goodman, 1972; 1973; Brown, 1991; Souza and Sorrels, 1991a,b).

Assessment of genetic diversity (or distance) between lines, populations, or races may be based on (i) analysis of pedigree and heterosis data, (ii) biochemical or genetic markers such as isozymes, and RFLPs, or (iii) quantitative traits based upon biometrical criteria (Poehlman and Borthakur, 1969; Hallauer and Miranda, 1988; Jensen, 1988). However, Melchinger *et al.*, 1991, recognized that although all three types of estimates may be used

to infer genetic relationships, the correspondence among them is sometimes low and never perfect, primarily because of sampling deviations and the failure of phenotypes to differentiate genotypes precisely.

Assessment of genetic diversity for the first type of estimates is based on the knowledge of the pedigree of the individuals considered (Lefort-Buson, 1985). This type stems from the concept of genes identity (Coefficient of Identity and Parentage Coefficient). The usefulness of this type of estimates lies upon the good knowledge of pedigree data, but is limited by the complexity of computation.

The second type of estimates is based on proteins and DNA analysis (Goodman, 1973; Vienne (De) and Damerval, 1985). This method of estimating genetic diversity might be of little practical value to the breeder because of the complex interaction and pleiotropic effects of genes expressed in quantitative characters he is greatly dealing with.

The third type of estimates is based on the collection of multiple measurements on one or several samples of individuals, followed by assessments of statistical distances between these samples.

2.3. Use of Multivariate Analysis.

According to Kendall and Buckland as cited by Dagnélie (1975), the expression "Multivariate analysis" is used rather losely to denote the analysis of data which are in the sense that each member bears the value of p variates.

Morrison (1967) considered multivariate statistics as consisting of statistical methods which have the objectve of studying the relationships between several dependent or independent variables. The methods make it possible to, (i) obtain a synthetic description of the overall variability, (ii) assess the relative contribution of each trait to the differences

among populations and, (iii) group similar populations together on the basis of information that can be directly used in a breeding programme (Goodman and McBird, 1977; Hussain *et al.*, 1977; Camussi, 1979; Camussi and Ottaviano, 1981). Edye (1976) advocated multivariate analyses of agronomic databases in a manner similar to those used commonly in numerical taxonomy. Such an analysis was shown to elucidate patterns of variation in those attributes which are of economic importance, irrespective of taxonomic or genetic distance measured by molecular, chemical, or other means. Such classifications were shown to be useful supplements to other classifications.

Multivariate statistics in the field of Genetics can be divided into two main components. The first one is concerned with techniques which analyse patterns of covariation found within a single sample (multiple correlation and regression, principal components, canonical correlation, factors analyses, etc) and the second is concerned with analysis of variation between samples (canonical variate, cluster, multidimensional scaling and network analyses, etc) (Bryant and Atchley, 1975).

Although there is paucity of information on the use of multivariate analyses of agronomic parameters in cotton, such analyses are also common in other disciplines and crop plants. A multivariate analysis scheme of quantitative characters has been used to measure genetic diversity within cereal species, and was found to be one approach in identifying parent potential in wheats (Bhatt, 1970, 1973), rice (Kanwal *et al.*, 1983), and oats (Souza and Sorrels, 1991a).

Using database consisting of qualitative or both quantitative and qualitative characters, Hamman (1972), Bhatt (1976), and Amurrio et al. (1995) illustrated the utility of multivariate methods for discriminating among tobacco (*Nicotiana tabacum* L.), wheat cultivars, and pea landraces, respectively. Subtle differences and relationships were shown

which would otherwise go undetected with standard univariate analyses. Hamman (1972) found that multivariate techniques resolved several phenotypic measurements into fewer, more interpretable and more easily visualized dimensions.

2.3.1. Use of Principal Component Analysis (PCA) and Cluster Analysis (CLA) in plant breeding.

PCA is a multivariate technique for examining relationships among several quantitative variables. It is used for summarizing data and detecting linear relationships. Plots of principal components are especially valuable tools in exploratory data analysis. Principal components can be used to reduce the number of variables in regression, clustering, and so on (Onyilagha, 1986; SAS, 1988). Given a data set with p numeric variables, p principal components can be computed. Each principal component is a linear combination of the original variables, with coefficients equal to the eigen vectors of the correlation or covariance matrix. The eigen vectors are customarily taken with unit-norm. The principal components are sorted in descending order of eigen values, which are equal to the variances of the components.

The purpose of cluster analysis is to place objects into groups or clusters suggested by the data, not defined *a priori*, such that objects in a given cluster tend to be similar to each other in some sense, and objects in different clusters tend to be dissimilar.

Principal component and cluster analyses techniques are usually applied together, and have been found to be complementary, each technique providing informations not evident from the other (Rhodes and Martin, 1972). Principal component analysis can be used for summarizing data and detecting linear relationships, and the components produced can be used to reduce the number of traits in clustering.

Brown (1991), used PC & CL analyses on agronomic trial data to give graphical presentation of relative cultivar performance, and to show interrelationships of cotton cultivars across the U.S. Cotton Belt, based on agronomic performance and fibre quality measurements. Cultivars in the Mississipi Delta, Central, and Texas High Plains regions did not cluster tightly, whereas cultivars in the Eastern, New Mexico, and San Joaquin Valley regions clustered comparatively tightly. This implied greater intercultivar differences in the first three regions than in the latter three, which showed the genetic base of the first three to be broader.

Except for Upland cotton, the use of the two techniques is seen increasingly in plant breeding literature. Goodman and Mck.Bird (1977), using principal component and cluster analyses of 20 ear characters for 219 Latin American races and subraces of maize, delineated 14 groups of races. The relationships among these groups were also indicated, and some subdivisions of the 14 racial groups were also described.

In order to identify the major characters which account for variation among *Dioscorea rotundata* in Nigeria, Onyilagha (1986), subjected 40 characters recorded on 14 high yielding accessions to Single Linkage Cluster Analysis (SLCA) and Principal Component Analysis (PCA) and found the results of SLCA to agree with that of PCA ordination on the first three axes. The PCA showed 5 main distinguishing characters between the yam accessions. The author found the discrimation between groups and description of differences between groups to be more strongly marked from the PCA than in the SLCA.

Peeters and Martinelli (1989), assessed the potential of using hierarchical cluster analysis to classify entries from a germplasm collection according to their degree of similarity. Results suggested that similarity was greatest among individual entries by country of origin and that hierarchical cluster analysis could be used as a tool to classify entries from

germplasm collections according to their respective pedigree, even when no passport data are available. Based on this technique, it was also shown that the segregative potential of entries could be estimated.

Illoh and Olarede (1991), studied the major characters which account for the variation among mango varieties in Nigeria, using PCA and Single Linkage Cluster Analysis (SLCA). The results from both methods divided the thirty one varieties collected in all ecological zones of the country into two or four groups.

Factor, principal component and canonical analyses were used by Ariyo (1993) to study the extent of genetic diversity among 30 accessions of West African okra (*Abelmoschus caillei*) (A. Chev.). Genetic variability among the accessions proved to be large. Factor analysis and principal component analysis produced similar results which were substantially different from those produced by canonical analysis.

Mumm and Dudley (1994), in a study aimed at classifying 148 U.S. maize inbreds with respect to heterotic grouping by RLFPs resulting from 46 probe-enzyme combinations, used cluster analysis to reveal associations among inbreds and assign them to groups by the average linkage method. In the cluster analysis, inbreds were ordered into two major groups generally coinciding with distinctions between breeding groups derived from "Lancaster Sure Crop" open pollinated variety and from Iowa Stiff Stalk Synthetic. The classification was found to generally agree with pedigree information and was supported by principal component and cluster analyses.

Casler (1995), using PC and CL analyses, studied patterns of variation in a collection of perennial ryegrass accessions aimed at (i) estimating and describing quantitative genetic variation in the USDA perennial ryegrass collection, (ii) using this information to develop recommendations for future germplasm exploration, and (iii) initiating development of a core

collection. Cluster analysis was useful in identifying groups of accessions with superior performance for four traits and in classifying accessions to identify a core collection which preserved most of the phenotypic variation found within the entire collection.

Ouendeba *et al.* (1995), used Ward's cluster and principal component analyses to investigate the nature and degree of divergence among African Pearl Millet [*Pennisetum glaucum* (L) R. Br.] Landrace populations. The cluster analyses revealed similarities between Nigerien and Senegalese and between Nigerien and Nigerian landrace populations. Four principal components were found to explain 92% of the total variation. It was advocated that these results could be useful in choosing potentially heterotic pearl millet populations for intercrossing to develop improved cultivars, synthetics, and hybrids for use in Africa.

2.3.2. **Classification**

Two early measures of similarity between taxa were Pearson's Coefficient of Racial Likeness (C.R.L.) and Mahalanobis' (1936) generalized distance (D^2 Statistic) (Goodman, 1968).

Generalized D^2 statistic has been commonly used for grouping diverse genetic stocks in various types of breeding materials. This statistical technique is recognized by many authors as a powerful tool in quantifying the degree of divergence among biological populations at the genetic level (Katiyar, 1978).

In Upland Cotton, Singh and Bains (1968) studied the genetic divergence for ginning outturn and its components using D^2 statistic. Significant differences were revealed between the varieties considering the 6 characters (ginning outturn, lint- index, seed- index, mean fibre weight, seed surface area, and fibre density) collectively. Seed surface area and lint-

index contributed maximaly towards the total divergence. The clustering pattern was found not to follow the pattern of the geographical distribution.

Lefort and Schwendiman (1974) used D^2 statistics analysis recapitulated by a dendrogram to set the distribution of cotton strains into three groups: one of a balanced type, the two others of extreme and opposed types. In another application of D^2 analysis to the study of genetic divergence in Upland Cotton under different environments, Singh and Gill (1984), also found genetic diversity not to be related to geographical diversity of the genotypes evaluated. No two strains were grouped together in all environments. Ekbote and Khorgade (1986) estimated the magnitude of divergence present in 33 genotypes of Upland cotton from different parts of India, using the D^2 statistic, which grouped the genotypes into clusters. Geographical origin also was found to have limited influence on the genetic diversity.

Carvalho *et al.* (1995), on Brazilian cotton, *Gossypium hirsutum* var. *latifolium* Hutch., used D^2 statistic and canonical variables coupled with the grouping methods of Tocher and nearest neighbour to study divergence in seven cultivars grown at Viçosa in 1990. The seven cultivars were divided into four groups, and, in this case, there was an apparent association between genetic and geographic origin. Selection for disease resistance, especially against ramulose [causal agent *Colletotricum gossypii* var *calosporioides*], was an important factor in the divergence among the materials. The most divergent traits were plant height, fibre percentage and earliness.

Kalsy *et al.* (1995), evaluated 15 cotton (*Gossypium hirsutum*) parents along with their 105 F1 hybrids for seven related traits and subjected the data to D^2 statistic. The 120 genotypes were grouped into 10 clusters. Three of four crosses exhibited positive heterosis for seed cotton yield/plant. But unexpectedly, there was no correspondence between the

divergence of the parents as measured by the generalized distance between them and the extent of heterosis. The existence of ample genetic diversity among the genotypes, as seen from the inter-cluster distances, was found adequate for improvement by hybridization and selection for different combinations of characters.

In a particular study concerning the use of multivariate analysis to derive a system of classification for 15 races of maize from the region encompassed by southern Brazil, Bolivia, Paraguay, Uruguay, and northern Argentina, Goodman (1968) found the distances obtained to be relatively similar to the distances obtained from the commonly employed Mahalanobis' generalized distance technique.

With the objective to investigate the usefulness of various measures of statistical distance between races of maize, relative to their F_2 generations, and thereby providing an estimate of genetic divergence between the pairs of races involved in the crosses, Martinez *et al.* (1983) subjected five morphological characters to six statistical procedures: Euclidean distance, Mahalanobis' distance, generalized distance, modified generalized distance, approximate Dempster's distance, and Dempster's distance. The degrees of relationship estimated were in general agreement with previous studies based on classical taxonomic methods and multivariate analyses. The utility and use of generalized distance as a measure of genetic divergence are well illustrated in several other crop plants as reported by Murty and Arunachalam (1966), on sorghum; Rao *et al.* (1980), on safflower; Devi *et al.* (1991), on kenaf; Jagadev and Samal (1991), on niger (*Guizotia abyssinica* Cass.); Ziauddin *et al.* (1980), on Triticales; Gupta and Malik (1991), on wheat *Triticum sp.*; Bainiwal and Jatasra (1980), on pigeon pea, Dharmalingam and Kadambavanasundaram (1989), on cowpea (*Vigna unguiculata* (L.) Walp); Khan *et al.* (1990), and Katiyar (1978), on chikpea (*Cicer arietinum* L.)

2.4. **Traits selected and inter-relations.**

Selection progress for traits of interest can easily and rapidely be achieved if reliable heritability estimates are established. A measure of the the genetic associations among various traits is also important when planning breeding programmes (Dudley and Moll, 1969; Meredith, 1984).

Traits routinely studied by cotton breeders can be subdivided into highly heritable (number of vegetative branches, number of fruiting branch, internode length, lint percentage, linter percentage, boll weight, fibre length, fibre strength, fibre fineness, and maturity date), moderately heritable (node of the first fruiting branch, plant height, maturity ratio, micronaire reading, seed-index), and lowly heritable (yield per plant, number of bolls on fruiting branch) (Meredith, 1984; Lançon, 1987; 1995; Lançon *et al.*, 1993a).

Relationships among traits of economic importance is valuable in making decisions in breeding programmes, especially for selection purposes (Lukhele, 1981). Some of the principal correlations useful to a cotton breeder are listed below:

(i) positive favourable correlations between average boll weight and yield, seed-index and fibre elongation, fibre length and resistance of the yarn, fibre strength, maturity of the fibre and seed-index (Gutknecht and Fournier, 1982; Meredith, 1984; Dessauw and Hau, 1997);

(ii) positive unfavourable correlations between micronaire index and lint yield, ginning percent and fineness, maturity and micronaire index, elongation and fineness (Gutknecht and Fournier, 1982; Meredith, 1984; Lançon *et al.*, 1993b; Dessauw and Hau, 1997);

(iii) negative favourable correlations between node of the first fruiting branch and earliness, fineness and fibre strength, maturity and seed index (Lançon *et al.*, 1993b; Dessauw and Hau, 1997);

(iv) negative unfavourable correlations between lint yield and seed-index and strength, number of vegetative branches and earliness, fibre percentage and fibre length, seed-index and average boll weight, elongation and strength, maturity and fineness (Gutknecht and Fournier, 1982; Meredith, 1984; Lançon *et al.*, 1993b; Dessauw and Hau, 1997).

Though most of the phenotypic correlations follow the same trends as genetic correlations described above, Lançon *et al.* (1993b) found a small number to appreciably differ from them. These are flowering earliness and the production of vegetative branches, fibre percentage and elongation or standard fineness, 50% S.L. and strength or standard fineness, stelometre strength and elongation, standard fineness or seed index, elongation and standard fineness, and fineness and seed index.

Also, yield related traits and/or plant height are strongly influenced by environment and are much more complex to be improved than those related to fibre-quality (Lançon *et al.*, 1993b).

CHAPTER THREE

MATERIALS AND METHODS .

3.1. Experimental materials.

The materials used in this study are all, upland cotton (*Gossypium hirsutum* L.) genotypes with diverse genetic background originating from 42 different countries in five continents (Appendices 1 and 2). Seeds of the different genotypes were provided by the Montpellier (France) gene bank, in November 1993 and multiplied during the off-season at the Maroua research station of the Institute of Agronomic Research (IRA), Cameroon. Two hundred and seventy five genotypes randomly chosen from the "Cultivar Catalogue" (Banque de Gènes, Montpellier, July, 1993), were planted on the 29th of November 1993 under irrigation and strict selfing and were carried out on ridges of 5m long in order to get enough seeds for further evaluation of these materials during the planting season 1994/1995.

Only 200 genotypes produced sufficient quantities of cottonseed. Forty-seven genotypes were then added from elite breeding materials and cultivars of IRA Maroua Research Station and the Institute for Agricultural Research (IAR), Samaru, Nigeria. Nine genotypes from the latter group were repeated and randomized as different entities so as to meet the requirement of 256 entries for a 16x16 lattice square experimental design.

3.2. Experimental Technique.

In the 1994/95 growing season, the 256 entries were sown in a 16x16 partially balanced lattice, four replicates obtained by duplicating the simple lattice (Cochran and Cox, 1957; Steel and Torrie, 1960; Gomez and Gomez, 1984), in three locations, Samaru in Nigeria (latitude 11°11'N;longitude, 7°35'E), Maroua (latitude, 10°35'N; longitude,14°18'E) and Garoua (latitude, 9°20'N; longitude,13°23'E) in Cameroon.

Plots consisted of single rows which were 5m long, spaced 1m apart at Maroua and Garoua, and 0.91m at Samaru with a plant to plant spacing of 0.40m. Four to six cotton seeds were hand sown per hill on the 1st, 3rd, and 12th of July at Garoua, Maroua and Samaru, respectively, in 1994. After seedling emergence, the plants were thinned to two per hill, then to one at the beginning of square initiation. Missing hills were replaced for in most plots by replanting before thinning. At the three locations, multinutrient fertilizer (44kg N/ha, 20kg P_2O_5/ha, 30kg K_2O/ha, 10kg S/ha, 2kg B/ha at sowing in Maroua; 30Kg N/ha, 40kg P_2O_5/ha, 30kg K_2O/ha, 10kg S/ha, 2kg B/ha at sowing and 23kg N as urea on 40th day in Garoua; and 50kg N/ha, 25kg P_2O_5/ha and 0.7kg B/ha in Samaru) were applied. The plots were kept weed-free by hand-hoeing as necessary and the crops were protected against insect pests by insecticide treatments (eight weekly sprays of karate - 10 E.C.- starting at 40 days after sowing at the rate of 1.0 litre per hectare).

3.3. Observations

Data were taken from five randomly selected plants on morphological and agronomic traits *in situ*. At maturity, samples of 200g of seed cotton were taken from each plot, ginned, and the lint produced was subjected to fibre quality tests at the Fibre Technology Laboratory of the Institute for Agricultural Research, Samaru, Nigeria.

Observations were recorded on the traits listed in Table 1.

Table 1: Traits observed for the evaluation of 247 Upland cotton genotypes and measurement

	Code	Traits	Measurement
1	H	Plant height (cm)	The height in cm from the base of the plant to the base of the terminal bud
2	INPP	Number of internodes per plant	Total number of internodes counted in the main stem per plant
3	FB	Number of fruiting branches	Total number of fruit bearing branches regardless of whether or not the bolls in them mature and eventually open
4	VB	Number of vegetative branches	Total number of vegetative branches counted per plant
5	HFFB	Height of the first fruiting branch	Distance in cm from ground level to the insertion in main stem of the first fruiting branch
6	D50BO	Days to 50% boll opening	Number of days after planting when 50% bolls have open (derived from daily count of bolls open in the basic plot)
7	INFFB	Number of internodes of the first fruiting branch	Total number of internodes counted on the first fruiting branch
8	LFFB	Length of the first fruiting branch	Total length in cm of the first fruiting branch
9	BOPP	Number of bolls per plant	Total number of bolls produced by the plant counted at harvest
10	BOLT	Boll length	Half circumference in mm of the mature and still green capsule along the peduncle axis (mature captules per plant taken at the different stages)
11	BODIA	Boll diameter	Circumference in mm at the largest enlargement of the capsule (same as above for measurement)
12	DFBO	Days to first boll opening	Total number of days after planting when the cumulative number of open bolls is the same as the number of plants per plot
13	DFFO	Days to first flower opening	Total number of days after planting when the cumulative number of open flowers is the same as the number of boll
14	TPb	Number of teeth per bract	Total number of teeth counted on the bract covering the bolls
15	D50FO	Days to 50% flower opening	Number of days after planting when 50% of flowers have opened (derived from daily count of flowers)
16	FLPP	Number of flowers per plant	Total number of flowers produced per plant
17	SCFP	Earliness	Weight of seed cotton at first picking expressed as a percentage of the total seed cotton yield
18	YLPP	Yield per plant	Seed cotton in g harvested per plant
19	GIN%	Ginning percent	Weight of lint divided by the weight of seed cotton expressed as a percentage
20	S.I	Seed index	Weight in g of 100 seeds
21	2.5%S.L.	2.5% span length	The fibre length in mm measured by the Digital fibrograph in the test specimen spanned by 2.5% of the fibres scanned at the initial starting point.
22	50% S.L.	50% span length	The length in mm in the test specimen spanned by 50% of the fibres scanned at the initial starting point.
23	FBST	Fibre bundle strength	The strength of a bundle of cotton fibres measured on the stelometer with the low jaws holding the fibre bundle separated by 1/8inch spacer (1/8 gauge length T_1)
24	MIC	Micronare index	The resistance of a given weight of ginned cotton lint to air flow as an indication of the fineness of the sample
25	BBVB(%)	Percent bolls produced by vegetative branches	Proportion of bolls counted on vegetative branches out of the total bolls produced per plant.

3.4. Statistical analysis.

Plot means were used in all analyses for the 247 genotypes. The data were sorted out and entered under MSTATC (1993) software and the analysis were conducted at the "Centre de Coopération Internationale en Recherche Agronomique pour le Développement - Département des Cultures Annuelles" (CIRAD-CA), Montpellier, France.

3.4.1. Analysis of Variance

The data recorded were carefully examined in order to fulfil the requirements for the variance and multivariance analyses using appropriate statistical procedures. The frequency distribution of the observations were crosschecked with their expected values. Some suspected gross errors, from which no type of experimentation seems to be entirely free, were omitted and some data transformations performed (square root and logarithmic transformations of records on number of flowers and pods per plant, respectively) so as to minimize and stabilize the experimental errors wherever necessary. The observations with unit normal deviates exceeding 4.5 in absolute value were also considered outliers and excluded from the final analysis (Rodgers *et al.*, 1983).

First, an analysis of variance was conducted for all the traits using the General Linear Model (GLM) procedures of SAS (1988), which are appropriate in analyzing data arranged in non-orthogonal incomplete design with missing cells (omitted data) of our type, for each location and across locations. In the analyses, locations, which were actually experimental research stations were considered fixed effects. The genotypes selected for testing were either those which had the most promise as potential parental materials for hybridization schemes or the most readily available. The genotypes were also thus considered fixed effects in the analyses.

The statistical model used for the analysis in each location was:

$$y_{ijk} = \mu + \alpha_i + \gamma_j + \delta_{jk} + \epsilon_{ijk}$$

y_{ijk} = Observation for the i^{th} treatment in the k^{th} block within the j^{th} replication;

μ, α_i, γ_j and δ_{jk} : represent the effects of mean treatment, replication and the incomplete block, respectively.

ϵ_{ijk} : represents the experimental error.

The form of analysis of variance for each location is presented in Table 2.

The statistical analysis model used for the combined analysis was:

$$y_{ijkl} = \mu + \alpha_i + \beta_l + (\alpha\beta)_{il} + \gamma_{jl} + \delta_{jkl} + \epsilon_{ijkl}$$

Y_{ijkl} = Observed mean for the i^{th} treatment at the l^{th} place in the k^{th} block within the j^{th} replication;

$\mu, \alpha_i, \beta_l, (\alpha\beta)_{il}$, γ_{jl} and δ_{jkl} : represent the effects of mean, treatment, location, interaction, replication (nested in location), and block (nested in replication within location), respectively;

ϵ_{ijkl} represents the experimental error.

The form of combined analysis of variance is shown in Table 3. The block variance was used to test locations and replication sources of variation, while the error variance was the appropriate term for testing hypothesis about other sources of variation in the analysis of variance (Shultz, 1955; Cochran and Cox, 1957; McIntosh, 1983).

The treatment adjusted means obtained from the analysis of variance were subjected to further analyses:

- Correlation and Principal component analysis,
- Classification.

Table 2: Form of analysis of variance with expected mean squares (EMS) for each trait in one location.

Source of variation	d.f	MS	EMS
Replications	r - 1		
BLOCK(REP)	r(k -1)		
Genotypes	t* - 1	M2	$\sigma^2_\epsilon + r\theta^2_G$
Error	rk(k - 1) - (t - 1)	M1	σ^2_ϵ
Corrected Total	rk^2 - 1		

r = number of replication (r = 4)
k = block size (k = 16)
t = number of treatments (genotypes) (t = 247)
σ_ϵ = error variance
θ^2_g = Total genetic variance among entries
M subscript = The observed mean square of the subscripted effect
θ^2 = Denotes variance for fixed effect
t* = k^2 - a where a = 9.

Table 3: Form of analysis of variance with expected mean (EMS) for each trait across the three locations in 1994.

Source of variation	d.f.	MS	EMS
Location (LOC)	l - 1		
Replication (LOC)	l(r - 1)		
BLOCK (LOCxREP)	rl(k -1)		
Genotype (GENO)	t^* - 1	M3	$\sigma^2_\epsilon + rl\theta^2_G$
LOCxGENO	(l-1)(t-1)	M2	$\sigma^2_\epsilon + r\theta^2_{GL}$
Eror	l[rk(k-1)-(t-1)]	M1	σ^2_ϵ
Corrected Total	rlk^2-1		

l = number of location (fixed effect)
r = number of replications
k = bloc size
t = number of treatment (fixed effect)
M subscript = observed mean of the subscripted effect
$\sigma\epsilon^2$ = Error variance
ω^2_{GL} = Genotype x location interaction variance
θ^2_G = Total genetic variance among entries
θ^2 = Denotes variance for fixed effect

$t^* = k^2 - a$ where a = 9 (duplicated genotypes).

3.4.2. Correlation among traits and Principal Components Analysis (PCA).

Data were standardized to compute the correlation matrix of traits of each trial thereby removing the effects of scale (Souza and Sorrels, 1991). Principal components analyses were performed on the correlation matrix of traits of each trial using the "WINSTAT i.c. Procedure", a software developed by the "Centre de Coopération Internationale en Recherche Agronomique pour le Développement" (CIRAD). Montpellier, France.

3.4.3. Classification

A hierarchical clustering was carried out on each data set (i.e; for each location) using Ward's minimum variance method (SAS, 1988), which minimizes within clusters sums of squares. The sum of the first 10 principal component scores, representing more than 74% of the total variation, was used in each case for the clustering procedure, making the results equivalent to those from initially standardized data as the correlation matrix was used for principal component analysis (Brown, 1991). This method is a stepwise clustering process whereby each accession initially comprises its own cluster. Two clusters are combined in each step, minimizing the pooled within-cluster vaiance at each step. The last step creates a single all-inclusive cluster (Milligan, 1980; Casler, 1995). Hierarchical clusters are organized so that one cluster may be entirely contained within another cluster, but no other kind of overlap between clusters is allowed (SAS, 1988). Finally, one must take the decision of where to cut the tree so as to produce groups that are maximally related to each other for selected traits. It is less sensitive to error when one knows in some sense the biological material with which one is dealing (Amurrio *et al.*, 1995). Mumm and Dudley (1994) partitioned the dendrogram at the 0.90 level of dissimilarity to form sets of genotypes exhibiting ≤ 0.90 dissimilarity.

The distance between clusters was measured using the D^2 statistic. This was estimated using the STATICTCF, version 5, "Option Analyse Factorielle Discriminante" of the CIRAD, Montpellier, France.

CHAPTER IV

RESULTS

4.1. Mean performance

The means, standard errors, ranges and coefficients of variation for the 25 traits evaluated at each of the three locations, Samaru, Maroua, and Garoua is presented in Table 4. The means, ranges and coefficients of variation varied widely in values.

In Samaru, wide class ranges were observed on the percentage boll produced by vegetative branches (0 - 30), number of vegetative branches (0.0 - 2.5), earliness (10 - 70g), yield per plant (30 - 70g), and length of the first fruiting branch (15 - 40cm). High C.V.% of 57.2, 51.4, 37.3, 34.8, and 28.5, respectively, were thus observed. The lowest C.V.% (3.7) was observed on days to 50% flower opening.

In Maroua, wide class ranges were also observed on vegetative branches per plant, yield per plant (40 - 120g), and length of first fruiting branch (15 - 45cm), with C.V.% of 41.4, 31.6, and 23.3, respectively. Low values of coefficient of variation were observed on days to first boll (2.5%) and first flower (3.5%) opening, and days to 50% boll (2.4%) and flower (3.4%) opening, indicating an efficiency of the field design in the evaluation of the set of genotypes or an intra-genotype homogeneity for these traits.

In Garoua, wide ranges of records occured for vegetative branches per plant (1 - 6), yield per plant (100 - 300g), and length of first fruiting branch (30 - 70cm); the C.V.% were 39.3, 31.2, and 24.2, respectively. The lowest C.V.% (2.7%) was observed for days to 50% flower opening.

Table 4: Mean, Standard errors of mean (S.E. $\overline{X}$), Ranges and coefficient of variation (C.V.%) for 247 Genotypes of upland cotton grown at Samaru, Maroua and Garoua, 1994.

Traits	Mean ± S.E. $\overline{X}$	SAMARU Range	C.V%	Mean±S.E. $\overline{X}$	MAROUA Range	C.V.(%)	Mean±S.E. $\overline{X}$	GAROUA Range	C.V(%)
Plant height (cm)	102.963±13.750	75-125	11.6	118.826±12.807	80-140	12.9	123.941±17.777	80-160	12.4
No. of internodes per plant	13.790±1.246	11-17	10.3	-	-	-	-	-	-
No. of fruiting branches	15.358±1.141	13-18	8.7	16.728±1.333	14-19	10.8	14.506±1.571	12.5-17.5	15.4
No. of vegetative branches	1.166±0.527	0.0-2.5	51.4	1.523±0.480	1.0-3.5	41.4	2.292±0.859	1-6	39.3
Height of first fruiting branch (cm)	16.971±2.640	10.0-25.0	16.5	19.225±3.184	12.5-25.0	18.3	18.912±4.087	15-30	20.6
Days to 50% boll opening	-	-	-	117.288±4.326	110-130	2.4	122.102±5.587	110-140	2.7
No. of internodes of the first fruiting branch	3.196±0.467	2.5-4.5	22.1	3.736±0.503	2.0-5.0	21.4	5.684±0.718	3-7	19.6
Lenght of the first fruiting branch (cm)	22.406±5.451	15-40	28.5	31.406±5.045	15-45	23.3	44.051±8.546	30-70	24.2
No. of bolls per plant	2.606±0.246	2.00-3.25	15.2	2.707±0.212	2.2-3.2	13.3	3.390±0.242	2.75-4.00	9.2
Boll lenght (mm)	5.380±0.216	4.75-5.75	6.4	4.067±0.323	3.50-4.75	13.2	5.478±0.265	4.8-6.0	5.9
Boll diameter (mm)	3.427±0.121	3.2-3.8	5.5	2.705±0.171	2.25-3.00	11.3	3.143±0.149	2.8-3.6	6.2
Days to first boll opening	-	-	-	111.551±3.848	105-125	2.5	-	-	-
Days to first flower opening (DAP*)	81.883±3.101	75.0-87.5	4.0	66.785±2.826	60-75	3.5	69.876±5.395	60-90	4.5
No. of teeth per bract	10.567±0.774	7-12	9.8	11.190±0.659	10-13	10.1	12.257±0.681	11-14	6.9
Days to 50% flower opening (DAP)	101.690±2.343	97.5-107.5	3.7	87.686±4.080	75-100	3.4	92.415±7.641	70-120	3.5
No. of flowers per plant	3.990±0.463	3.0-5.5	14.5	6.901±0.449	6.0-8.0	10.8	5.979±0.650	4.5-7.5	14.1
Earliness (%)	36.758±11.700	10-70	37.3	-	-	-	-	-	-
Yield per plant (g)	43.516±9.631	30-70	34.8	70.571±16.487	40-120	31.6	198.893±44.915	100-300	31.2
Ginning percent	37.996±3.084	30.0-45.0	4.6	33.528±3.378	25-40	6.0	37.002±3.055	30-45	5.5
Seed index (g)	9.829±0.823	8-12	6.4	8.724±0.770	7-11	8.5	9.285±0.769	7-11	9.7
2.5% span lenght (mm)	26.101±1.486	22.5-30.0	4.1	25.381±1.627	20.0-30.0	4.4	26.186±1.598	22.5-30.0	4.3
50% span lenght (mm)	11.446±0.704	10-13	5.5	11.192±0.798	10-13	5.9	11.782±0.787	10-14	5.5
Fibre bundle strength (g/tex)	18.726±1.582	15.0-22.5	12.6	17.603±1.491	15-20	12.2	19.280±1.426	16-20	11.8
Micronaire index	3.934±0.368	3.5-5.0	10.1	3.212±0.308	3.0-4.0	14.0	3.729±0.368	3.0-4.5	11.6
Bolls (%) produced by monopodial branches	11.813±5.507	0-30	57.2	-	-	-	7.125±5.225	10-30	9.0

* Days after planting

- Traits discarded at that particular location showing anomalous heterogenous observations

The seed cotton per plant was highest in Garoua (199g), medium in Maroua (70.5g), and lowest in Samaru (43.5g). The ginning percentages were higher in Samaru and Garoua (38 and 37, respectively), but low at Maroua (33.5). The plant height did not vary much over the three locations, Samaru, Maroua, and Garoua and had values of 103cm, 119cm, and 124cm, respectively.

The mean squares from the individual locations analyses are presented in Appendices 3 to 5, while the results of the combined analysis are presented in Table 5. At each of the three locations highly significant differences were obtained between genotypes for all traits. Highly significant differences were also observed among the genotypes for the 20 traits measured in the combined analysis. The mean squares due to locations were highly significant. The Genotype x Location interaction mean square was not significant for fibre bundle strength. For all other traits apart from number of teeth per bract and days to 50% flowers opening which recorded significant mean squares, highly signicant values were obtained for the Genotypes x Location interaction effects.

Estimates of the genotypic and phenotypic variances and heritability for the 25 traits measured in the three locations are presented in Table 6.

At Samaru, the major portion of the phenotypic variance was attributed to Genotypic components for ginning percent (91.7%), 2.5 percent spar length (87.1%) and seed index (85.2%) and plant height (80.1%). The proportion of the phenotypic variance that was due to the genotypic variance was low for number of bolls per plant (33.3%) boll length (33.3%) and seed cotton yield (35.8%) at this site.

At Maroua, genotypic components that contributed most to the total variation observed were as at Samaru, ginning percent (91.2%), 2.5% span length, but also days to 50% bolls

Table 5: Mean squares for the combined analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown in the three locations in 1994.

	Mean squares of the traits measured											
Source of variation	DF	PH(cm)	FB	VB	HFFB(cm)	INFFB(cm)	LFFB(cm)	Log[BOPP]	BOLT(mm)	BODIA(mm)	DFFO	TPb
Location	3	125586.48**	1291.97**	345.40**	1582.63**	1759.46**	121951.05**	189.54**	635.10**	135.56**	64946.81**	747.66**
Replication(Location)	9	6978.86**	125.25**	21.51**	220.07**	35.16**	1258.14**	3.67**	4.07**	1.71**	49.99**	40.72**
Block(Location*Replication)	180	1415.27**	16.48**	1.11**	34.18**	2.07**	186.59**	0.30**	1.19**	0.25**	17.07**	3.33**
Genotypes	246	1796.29**	12.59**	3.09**	95.34**	1.67**	284.49**	0.27**	0.41**	0.11**	80.96**	3.16**
Location*Genotype	492	382.34**	4.44**	0.80**	16.96**	0.95**	101.31**	0.17**	0.21**	0.07**	44.36**	1.20*
Error	≤2142++	205.16	3.38	0.53	11.87	0.79	69.63	0.13	0.17	0.06	8.70	1.03

* and ** indicate significance at the 0.05 and 0.01 levels of pobability, respectively

\+ Trait abbreviated as defined in Table 1.

++ Error degree of freedom varied from one trait to another because of omitted data for these traits at each location

Table 5 (cont.)

Mean squares of the traits measured

Source of variation	DF	D50FO	[FLPP]"	YLPP(g)	GIN%	S.l(g)	2.5% S.L.	50% S.L.	FBST(g/tex)	MIC
Location	2	51694.27**	2273.02**	7149109.46**	5557.79**	310.50**	194.34**	88.58**	743.02**	141.80**
Replication(Location)	9	76.67**	6.39**	10642.66**	41.94**	21.63**	32.12**	39.41**	388 16**	4.06**
Block(i ocation*Replication)	180	20.39**	1.71**	3470.63**	14.25**	1.85**	3.99**	2.26**	20.81**	1.12**
Genotypes	246	112.67**	1.15**	3511.41**	108.50**	5.52**	25.55**	5.92**	16.48**	0.94**
Location*Genotype	492	85.65*	1.01**	2803.65**	5.54**	0.83**	1.48**	0.52**	4.69	0.23**
Error	$\leq$2142++	11.12	0.53	1541.89	3.76	0.58	1.23	0.42	5.13	0.18

* and ** indicate significance at the 0.05 and 0.01 levels of pobability, respectively

+ + Error degree of freedom varied from one trait to another because of omitted data for these traits at each location

Table 6: Estimates of genotypic variance (θ^2_G), phenotypic variance (θ^2ph) and heritatbility (h^2) for the traitsmeasured at each location.

	Samaru			Maroua			Garoua		
Traits	θ^2_G	θ^2_{ph}	h^2	θ^2_G	θ^2ph	h^2	θ^2_G	θ^2_G	h^2
H(cm)	144.5	180.46	0.80	92.5	151.38	0.61	249.4	308.40	0.80
INPP	1.03	1.54	0.67	-	-	-	-	-	-
FB	0.82	1.27	0.64	0.84	1.66	0.51	1.17	2.43	0.48
VB	0.16	0.25	0.64	0.12	0.22	0.54	0.49	0.70	0.71
HFFB(cm)	4.70	6.66	0.70	6.45	9.57	0.67	12.26	16.08	0.76
D50BO	-	-	-	13.88	15.80	0.88	27.09	29.76	0 91
INFFB	0.08	0.20	0.38	0.07	0.23	0.32	0.15	0.46	0.32
LFFB	17.97	28.17	0.64	9.71	23.14	0.42	41.87	70.48	0.59
BOPP	0.02	0.06	0.33	0.01	0.04	0.23	0.03	0.05	0.52
BOLT	0.01	0.04	0.33	0.02	0.10	0.24	0.04	0.07	0.63
BODIA	0.01	0.12	0.50	0.00	0.03	0.22	0.01	0.02	0.50
DFBO	-	-	-	10.76	12.71	0.85	-	-	-
DFFO	6.00	8.73	0.69	5.98	7.36	0.81	23.88	26.34	0.91
TPb	0.28	0.55	0.52	0.08	0.40	0.21	0.25	0.43	0.58
D50FO	1.80	5.30	0.341	10.81	13.08	0.83	50.05	52.62	0.95
FLPP[+]	0.14	0.22	0.062	0.06	0.20	0.31	0.19	0.37	0.52
SCFP	83.6	130.64	0.64	-	-	-	0.76	-	-
YLPP	31.9	89.10	0.36	125.8	251.82	0.50	963.9	1938.76	0.50
GIN%	8.50	9.28	0.92	10.44	11.46	0.91	8.12	9.16	0.89
S.I.	0.56	0.66	0.85	0.43	0.57	76.0	0.36	0.57	0.64
2.5% S.L.	1.8	2.13	0.87	2.25	2.57	0.87	2.11	2.43	0.87
50% S.L.	0.39	0.49	0.80	0.54	0.65	0.83	0.50	0.60	0.83
FBST	1.0	2.39	0.42	0.99	2.16	0.46	0.62	1.91	0.32
MIC	0.09	1.14	0.70	0.44	0.09	0.47	0.070	0.12	0.61
BBVB(%)	16.7	28.25	0.59	-	-	-	22.5	22.65	0.99

+ Trait abbreviations as defined in Table 2.
- Trait discarded at that particular location showing anomalous heterogenous observations

opening (87.8%), 50% span length (87.5%), days to first boll opening (84.8%), days to 50% (82.7%) and first (81.4%) flowers opening. The genotypic contribution to the phenotypic observed variance was moderate for number of internodes (31.9%) and length of the first fruiting branch (42.0%), boll length (23.7%) and diameter (22..2%), number of teeth per bract (20.7%), number of flowers per plant (31.2%), fibre bundle strength (46.0%) and micronaire (47.4%) and seed cotton yield (49.9%).

At Garoua, the phenotypic variance observed was most attributable to genotypic variance for boll percent on vegetative branches (99.6%), days to 50% flowers (93.1%) and bolls (91.0%) opening, days to first flower opening (90.7%), ginning percent (88.6%), and 2.5 and 50% span length (86.9% and 82.6%, respectively).

The portion of the phenotypic variance attributable to environmental components was sizeable for number of internodes of the first fruiting branch (67.8%), fibre bundle strength (67.8%), and number of fruiting branches (51.9%).

4.2. Simple Linear Correlations

All possible linear correlations among the traits considered were computed based upon adjusted mean values among the 247 genotypes at each location, and the results are presented in Tables 7 to 9.

In Samaru, correlation coefficients (r) were computed between 23 traits and the results are shown in Table 7. Many pairs of traits were correlated. Height of the first fruiting branch was significantly ($P \leq 0.01$) correlated with 16 out of the 22 other traits measured, followed by 50% span length, days to first flower opening, and number of vegetative branches that were associated with 14, 13, and 12 other traits, respectively. Number of teeth per bract was correlated ($P \leq 0.05$) with the fewest traits (three traits).

Table 7: Simple linear correlations between 23 traits of upland cotton grown at Samaru in 1994

	1	2	3	4	5	7	8	9	10	11	13	14	15	16	17	19	20	21	22	23	24	25	23
1	1.000																						
2	0.633**	1.000																					
3	0.647**	0.856**	1.000																				
4	0.315**	0.045	0.134*	1.000																			
5	0.411**	-0.046	0.022	0.570**	1.000																		
7	0.111	0.089	0.143*	0.200**	-0.081*	1.000																	
8	0.367**	0.038	0.152*	0.441**	0.261**	0.708**	1.000																
9	0.089	0.331**	0.312**	-0.005	-0.312**	0.364**	0.188**	1.000															
10	0.048	-0.020	-0.072	-0.012	-0.015	-0.062	0.038	-0.027	1.000														
11	-0.083	-0.211**	-0.239**	-0.082	-0.040	-0.046	0.000	-0.068	0.441**	1.000													
13	0.201**	-0.079	-0.041	0.301**	0.457**	-0.036	0.065	-0.214**	0.006	-0.021	1.000												
14	-0.094	-0.101	-0.062	-0.097	-0.062	-0.029	-0.016	0.073	0.131*	0.243**	-0.025	1.000											
15	0.158*	0.037	0.047	0.290**	0.220**	0.060	0.077	0.008	0.005	-0.031	0.576**	-0.016	1.000										
16	0.157*	0.403**	0.413**	0.185**	-0.203**	0.303**	0.214**	0.465**	-0.066	-0.093	-0.386**	-0.052	-0.085	1.000									
17	-0.263**	-0.064	-0.070	-0.343**	-0.367**	-0.003	-0.113	0.110	-0.073	0.012	-0.544**	0.039	-0.597**	0.184**	1.000								
18	0.158*	0.283**	0.261**	-0.071	-0.299**	0.296**	0.220**	0.655**	0.071	0.022	-0.287**	0.179**	-0.086	0.303**	0.123	1.000							
19	0.081	-0.053	-0.135*	0.143*	0.192**	-0.107	-0.022	-0.045	-0.093	-0.108	0.129*	0.070	0.108	-0.118	-0.062	-0.002	1.000						
20	-0.002	-0.073	-0.070	0.060	0.115	-0.131*	-0.102	-0.246**	0.424**	0.333**	0.021	0.072	-0.069	-0.147*	-0.084	-0.074	-0.308**	1.000					
21	0.261**	0.073	0.067	0.220**	0.323**	-0.174**	-0.051	-0.213**	0.146*	-0.024	0.173**	0.020	0.072	-0.166**	-0.199**	-0.180**	0.374**	0.221**	1.000				
22	0.320**	0.156*	0.136*	0.271**	0.355**	-0.176**	-0.023	-0.272**	0.117	-0.055	0.197**	-0.016	0.142*	-0.174**	-0.245*	-0.177**	0.308**	0.268**	0.825**	1.000			
23	0.250**	0.184**	0.177**	0.132*	0.276**	-0.125*	-0.022	-0.148*	-0.031	-0.080	0.173**	-0.003	0.051	-0.127*	-0.201**	-0.118	0.232**	0.103	0.386**	0.542**	1.000		
24	0.130*	0.082	0.060	0.168**	0.189**	-0.026	-0.038	-0.050	-0.163*	-0.177**	0.221**	-0.110	0.242**	-0.121	-0.384**	-0.069	0.206**	0.040	0.016	0.209**	0.259**	1.000	
25	0.229**	0.038	0.096	0.642**	0.378**	0.236**	0.432**	0.106	-0.065	-0.193**	0.163*	-0.077	0.181**	0.139*	-0.292**	0.052	0.083	-0.004	0.110	0.147*	0.100	0.155*	1.000

* and ** Significant at 0.05 and 0.01 levels of propability, respectively.

1 Plant height (cm)
2 No. of internodes per plant
3 No. of fruiting branches
4 No. of vegetative branches
5 Height of the first fruiting branch
6 Days to 50% boll opening
7 No. of internodes of the first fruiting branch
8 Length of the first fruiting branch
9 No. of bolls per plant
10 Boll length
11 Boll diameter
12 Days to first boll opening
13 Days to first flower opening
20 Seed index
22 50% span length
23 Fibre bundle strength
25 Percent bolls produced by vegetative branches
14 No. of teeth per bract
15 Days to 50% flower opening
16 No. of flowers per plant
17 Earliness
18 Yield per plant
19 Ginning %
21 2.5% span length
24 Micronare index

Table 8: Simple linear correlations coefficients between 22 traits measured on 247 genotypes of Upland cotton grown at Maroua in 1994.

	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	19	20	21	22	23	24
1	1.000																					
3	0.499**	1.000																				
4	0.343**	0.109	1.000																			
5	0.444**	0.087	0.458**	1.000																		
6	0.056	0.083	0.162*	0.157*	1.000																	
7	0.072	0.250**	0.080	-0.018	0.092	1.000																
8	0.169**	0.143*	0.235**	0.120	0.114	0.568**	1.000															
9	0.003	0.022	-0.030	-0.053	-0.041	0.142*	0.064	1.000														
10	0.076	0.134*	0.118	0.057	-0.035	-0.019	-0.089	-0.073	1.000													
11	0.031	0.017	0.087	0.081	0.018	-0.009	0.045	-0.080	0.450**	1.000												
12	0.062	0.055	0.163*	0.095	0.877**	0.045	0.099	-0.094	0.014	0.085	1.000											
13	0.139*	0.053	0.157*	0.190**	0.248**	0.041	0.033	-0.147*	-0.027	-0.013	0.307**	1.000										
14	0.056	-0.020	-0.074	0.089	0.120	-0.036	-0.008	-0.015	0.020	0.071	0.084	0.134*	1.000									
15	0.030	0.033	0.067	0.018	0.167**	0.018	0.024	-0.177**	0.040	-0.017	0.239**	0.199**	0.029	1.000								
16	-0.046	0.085	-0.045	-0.241**	-0.040	0.066	0.168**	0.301**	-0.038	-0.004	-0.016	-0.252**	-0.155*	-0.132*	1.000							
18	0.187**	0.079	-0.007	0.043	0.003	0.080	0.087	0.166**	0.127*	0.017	-0.009	-0.162*	-0.007	0.013	0.170**	1.000						
19	0.101	-0.041	0.045	0.092	-0.040	-0.192**	-0.107	0.099	0.055	0.056	-0.004	-0.021	0.022	-0.009	0.111	0.217**	1.000					
20	0.012	-0.045	-0.040	0.091	0.089	0.008	-0.036	-0.051	0.258**	0.116	0.066	-0.018	0.045	0.024	-0.082	0.170**	-0.097	1.000				
21	0.300**	0.121	0.121	0.346**	0.084	-0.116	-0.109	0.060	0.142*	0.025	0.067	0.077	0.065	0.045	-0.001	0.215**	0.453**	0.184**	1.000			
22	0.312**	0.153*	0.180**	0.347**	0.089	-0.107	-0.070	0.018	0.151*	-0.037	0.074	0.074	0.059	0.005	0.015	0.155*	0.465**	0.162*	0.878**	1.000		
23	0.336**	0.169**	0.159*	0.245**	0.102	0.033	0.035	-0.015	0.104	0.057	0.103	0.214**	0.050	-0.074	-0.119	0.123	0.239**	0.180**	0.431**	0.547**	1.000	
24	0.136*	0.068	0.060	0.125*	0.113	-0.026	0.009	0.019	0.079	0.044	0.117	0.192**	0.135*	-0.042	-0.031	0.058	0.311**	0.067	0.096	0.288**	0.315**	1.000

*, ** Significant at 5% and 1% propability, respective.

1 Plant height (cm)
2 No. of internodes per plant
3 No. of fruiting branches
4 No. of vegetative branches
5 Height of the first fruiting branch
6 Days to 50% boll opening
7 No. of internodes of the first fruiting branch
8 Length of the first fruiting branch
9 No. of bolls per plant
10 Boll length
11 Boll diameter
12 Days to first boll opening
13 Days to first flower opening
14 No. of teeth per bract
15 Days to 50% flower opening
16 No. of flowers per plant
17 Earliness
18 Yield per plant
19 Ginning %
20 Seed index
21 2.5% span length
22 50% span length
23 Fibre bundle strength
24 Micronare index
25 Percent bolls produced by vegetative branches

Table 9: Simple linear correlations between 22 traits of upland cotton grown at Garoua '- 1994

	1	3	4	5	6	7	8	9	10	11	13	14	15	16	18	19	20	21	22	23	24	25
1	1.000																					
3	0.629**	1.000																				
4	0.285**	0.011	1.000																			
5	0.521**	0.195**	0.646**	1.000																		
6	0.089	0.030	0.148*	0.118	1.000																	
7	0.078	0.053	0.138*	0.028	0.101	1.000																
8	0.223**	0.045	0.435**	0.434**	0.129*	0.686**	1.000															
9	0.278**	0.280**	0.140*	0.208**	0.073	0.105	0.141*	1.000														
10	0.029	-0.026	-0.046	0.043	-0.005	-0.130**	-0.005	0.096	1.000													
11	0.017	0.012	-0.117	-0.043	-0.012	-0.053	-0.010	-0.077	0.674**	1.000												
13	0.044	-0.075	0.213**	0.171**	0.159*	0.047	0.121	-0.011	-0.032	0.011	1.000											
14	-0.120	-0.073	-0.135	-0.181**	-0.019	-0.013	0.006	-0.093	0.196**	0.239**	-0.019	1.000										
15	0.017	-0.052	0.108	0.091	0.081	0.157*	0.174**	-0.012	-0.109	0.024	0.584**	0.011	1.000									
16	0.321**	0.308**	0.125*	0.188**	0.026	0.155*	0.178**	0.818**	0.118	-0.089	-0.052	-0.083	-0.103	1.000								
18	0.347**	0.297**	0.258**	0.291**	0.042	0.236**	0.279**	0.411**	0.107	0.081	-0.041	-0.013	-0.033	0.407	1.000							
19	0.086	0.097	0.078	0.159*	-0.040	0.013	0.119	0.008	0.112	0.199**	0.017	0.048	0.052	0.024	0.178*	1.000						
20	0.025	-0.017	-0.068	-0.015	0.044	-0.196**	-0.191**	-0.053	0.426**	0.353**	-0.063	0.122	-0.175**	-0.038	0.120	-0.169**	1.000					
21	0.355**	0.228**	0.280**	0.375	0.147*	-0.055	0.063	0.183**	0.266**	0.165**	0.105	-0.052	0.026	0.195**	0.231**	0.422**	0.122	1.000				
22	0.337**	0.250**	0.339**	0.402**	0.137*	0.010	0.130*	0.163*	0.295**	0.136*	0.053	-0.038	0.028	0.207**	0.256**	0.379**	0.220**	0.859**	1.000			
23	0.124	0.093	0.204**	0.271**	0.096	-0.064	-0.007	0.118	0.226**	0.121	-0.042	-0.026	-0.080	0.139*	0.147*	0.097	0.161*	0.395**	0.455**	1.000		
24	0.174**	0.174**	0.103	0.161*	0.043	0.038	0.096	0.088	-0.037	-0.095	-0.017	-0.039	-0.044	0.102	0.123	0.142*	0.207**	0.036	0.233**	0.05	1.000	
25	0.101	-0.016	0.396**	0.226**	0.244**	0.106	0.201**	-0.080	0.016	0.053	0.126*	0.023	0.057	-0.027	0.149*	0.086	-0.018	0.177**	0.205**	0.105	0.059	1.000

* and ** Significant at 0.05 and 0.01 levels of propability, respectively.

1 Plant height (cm)
2 No. of internodes per plant
3 No. of fruiting branches
4 No. of vegetative branches
5 Height of the first fruiting branch
6 Days to 50% boll opening
7 No. of internodes of the first fruiting branch
8 Length of the first fruiting branch
9 No. of bolls per plant
10 Boll length
11 Boll diameter
12 Days to first boll opening
13 Days to first flower opening
14 No. of teeth per bract
15 Days to 50% flower opening
16 No. of flowers per plant
17 Earliness
18 Yield per plant
19 Ginning %
20 Seed index
21 2.5% span length
22 50% span length
23 Fibre bundle strength
24 Micronare index
25 Percent bolls produced by vegetative branches

The magnitude of the correlation coefficients were, in general, low. The strongest positive correlations ($r > 0.5$) were recorded between plant height and either number of internodes per plant (0.63) or number of sympodia (O.65), number of internode per plant and number of sympodia (0.86), number of monopodia and either height of the first fruiting branch (0.57) or number of bolls produced by monopodia (0.64), length and number of internodes of the first fruiting branch (C.71), seed cotton yield and number of bolls per plant (0.65), 2.5% and 50% S.L. (0.82), and fibre bundle strength and 50% S.L.. The strongest negative strong correlations were recorded between earliness and either days to first or 50% flower opening (-0.54 and -0.60, respectively). Many pairs of traits were correlated in height of the first fruiting branch was significantly ($P \leq 0.01$) correlated with 15 out of the 23 traits measured, followed by 50% span length, days to first flower opening and number of vegetative branches that were associated with 14, 13 and 12 other traits, respecvtively. Number of teeth per bract was correlated ($P \leq 0.05$) to the fewest traits (three traits).

In Maroua, all possible linear correlation coefficients were computed between the 22 traits considered and the results are presented in Table 8. Fibre bundle strength, plant height, height of first fruiting branch, 2.5% and 50% span lengths, days to first flowering, and ginning% were significantly ($P \leq 0.01$) correlated with at least six other traits. 2.5% span length was correlated ($P \leq 0.05$) to the largest number of traits (11) while boll diameter was correlated ($P \leq 0.05$) with only one trait. The magnitude of correlation coefficients here were lower compared to those observed in Samaru. High positive coefficients were recorded between plant heigh and number of fruiting branches (0.50), length and number of internodes of the first fruiting branch (0.57), days to first and 50% boll opening (0.88), 2.5% and 50% S.L. (0.88), and fibre bundle strength and 2.5% S.L. (0.55).

In Garoua, all possible linear correlations were computed between the 22 traits considered, many of them displaying significant correlations between themselves (Table 9); sixteen traits were significantly ($P \leq 0.01$) correlated with at least five other traits. 2.5% span length was correlated ($P \leq 0.05$) with the highest number of traits (14), and number of teeth per bract and days to 50% flowers opening with the fewest (4 traits each). The strongest positive correlations were obtained between plant height and either number of fruiting branches (0.63) and node of the first fruiting branch (0.52), number of vegetative branches and node of the first fruiting branch (0.65), length and number of internodes of the first fruiting branch (0.68), number of flowers and bolls per plant (0.82), length and diameter of the boll (0.67), days to first and 50% flower opening (0.58), and 2.5 and 50% S.L. (0.86). Here again, 2.5% span length was correlated ($P \leq 0.01$) the highest number of traits (14), and number of teeth per bract and days to 50% flowers opening with the fewest (4 traits each).

4.3. **Principal Components Analysis**

Principal component analyses were performed on the correlation matrix of traits of each trial and the results for the three locations are presented in Tables 10, 11 and 12.

4.3.1. **Samaru location**

When the data for the 23 traits were subjected to principal component analysis, 20 components were produced with eigenvalues lying between 0.196 and 4.3 (Table 10). The first two principal components axes accounted for 18.7, and 15.8%, respectively, of the total variation and together accounted for 34.5% of this variation. The first five principal components axes accounted for 59.8% of the total variation.

Table 10: Principal component (PC) analysis of traits associated with 247 upland cotton genotypes grown in Samaru in 1994, showing eigen values and proportion of variation associated with the first five PC axes and eigen vectors of traits.

	PC1	PC2	PC3	PC4	PC5
Eigen values	4.301	3.631	2.264	2.001	1.564
Cummulative proportion of variation	18.701	34.488	44.329	53.029	59.827
Plant height(cm)	-0.273	0.266	0.194	0.075	-0.093
No. of internodes per plant	-0.095	0.349	0.366	-0.061	-0.225
No. of fruiting branches	-0.113	0.369	0.309	-0.040	-0.199
No. of vegetative branches	-0.310	0.134	-0.233	0.112	0.200
Height of the first fruiting branch	-0.353	-0.058	-0.116	0.051	0.126
No. of internods of the first branch (cm)	0.012	0.304	-0.292	0.145	0.131
Length of the first fruiting branch (cm)	-0.121	0.277	-0.307	0.231	0.236
No. of bolls per plant	0.126	0.364	-0.018	0.030	-0.066
Boll length (mm)	-0.007	-0.079	0.105	0.521	-0.121
Boll diameter (mm)	0.074	-0.139	-0.008	0.482	-0.111
Days to first flower opening (DAP*)	-0.296	-0.117	-0.211	-0.039	-0.313
No. of teeth per bract	0.057	-0.042	0.031	0.230	0.000
Days to 50% flower opening (DAP)	-0.240	0.007	-0.224	-0.046	-0.415
No. of flowers per plant	0.091	0.361	0.052	0.033	0.097
Earliness (%)	0.314	0.021	0.161	-0.022	0.370
Yield per plant (g)	0.126	0.314	0.053	0.149	-0.025
Ginning percent	-0.158	-0.058	0.044	-0.215	0.310
Seed index (g)	-0.064	-0.155	0.151	0.445	-0.095
2.5% span length (mm)	-0.282	-0.111	0.296	0.094	0.277
50% span length (mm)	-0.329	-0.094	0.313	0.068	0.208
Fibre bundle strength (g/tex)	-0.250	-0.039	0.263	-0.062	0.107
Micronaire index	-0.201	-0.007	-0.036	-0.222	-0.185
Bolls (%) produced by vegetative branches	-0.235	0.170	-0.258	0.059	0.230

* DAP : Days after planting

Fourteen principal components were necessary to explain more than 90% of the total variation.

The coefficients for the first five principal components are presented in Table 10. Height of the first fruiting branch was a primary source of variation with the largest coefficient (-0.353) in the first principal component axis. The 50% span length had the second largest coefficient (-0.329), followed by earliness (0.314), number of vegetative branches (-0.310), days to first flower opening (-0.296), and 2.5% span length (-0.282). The boll length (-0.007), and diameter (0.07), seed-index (-0.064), and number of internodes of the first fruiting branch (0.012) had the smallest absolute value coefficients of the first principal component axis. Number of fruiting branches (0.369), number of bolls per plant (0.364), number of flower per plant (0.361), number of internodes per plant (0.349), seed cotton yield per plant (0.314), and number of internodes of the first fruiting branch (0.304) had the largest coefficients in the second principal component axis. Traits that loaded the principal component three (9.8% of the total variation) were number of internodes per plant (0.366), 50% span length (0.313), number of fruiting branches (0.309), and length of first fruiting branch (-0.307). Principal component four explained 8.7% of the total variation and was loaded by boll length (0.521) and diameter (0.482) and seed-index (0.445). Principal component five accounted for 6.8% of the total variation and consisted of days to 50% (-0.015) and days to first flower opening (-0.313) and earliness (0.370) and ginning % (0.310).

The configuration of the traits that described the 247 genotypes of Upland cotton and the placement of these materials are presented in Figures 1 and 2. By viewing these figures simultaneously, one could observe that the traits were a bit far off the origin of the circle, implying a fairly high degree of correlation of these traits with the principal component axes. In general there was an association between genetic divergence and geographic origin. The

cotton genotypes from Samaru appeared well grouped at the top of the second axis associated with good seed cotton yield, number of flowers and bolls per plant, and the length of the first fruiting branch. This implied a good agronomic adaptation of the local cultivars. The genotypes from the other countries of Africa appeared at the left (Figure 2), associated with good fibre quality traits (Figure 1). The genotypes from Europe were well differenciated and appeared at the right (Figure 2), associated with earliness (Figure 1).

4.3.2. Maroua location

The principal component analysis grouped the 22 traits into 20 components axes with eigen values lying between 0.312 and 3.59 (Table 11). The first two components accounted for 16.31 and 10.64% of the total variation, respectively, representing a total proportion of variation of only 27%. The first five principal component axes eigen values accounted for 50% of the total variation, while the first fifteen accounting for 90%, indicating a relatively low degree of correlation among characters for the genotypes at Maroua, in 1994.

Also presented in Table 11 are the scores for the major traits describing the first five axes. The first principal component axis had its largest coefficients for 50% span length (-0.417), 2.5% span length (-0.390), fibre bundle strength (-0.351), plant height (-0.318), height of the first fruiting branch (-0.315), ginning % (-0.228), and number of vegetative branches (-0.228). Internodes of first fruiting branch (-0.004), number of bolls per plant (0.017), length of the first fruiting branch (-0.047), number of flowers per plant (-0.065) and days to 50% flowers opening (-0.058), and bract's teeth (-0.082) had the smallest absolute value of coefficient of the first principal component axis. Days to first boll opening (0.429), days to 50% bolls opening (0.424), ginning % (-0.311), length of first fruiting branch (0.258),2.5% span length (-0.255), and 50% span length (-0.249) had the largest

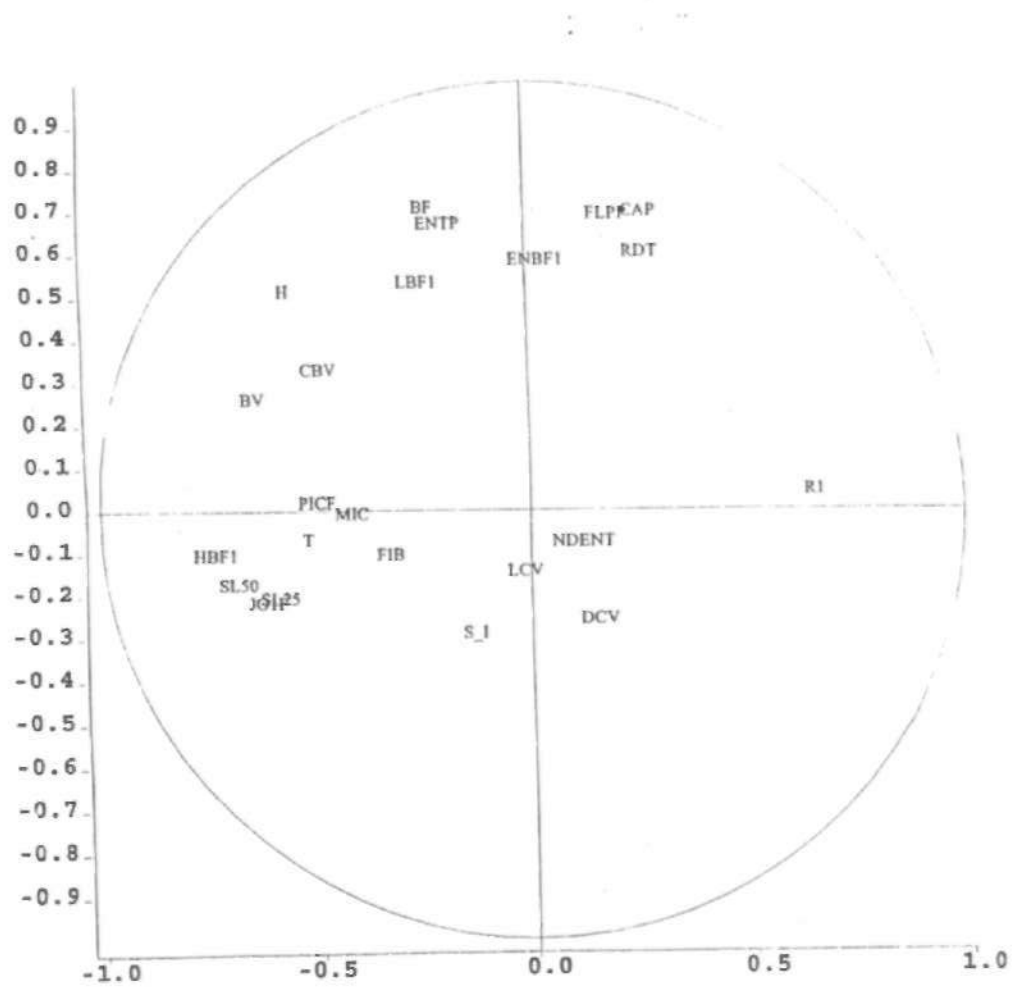
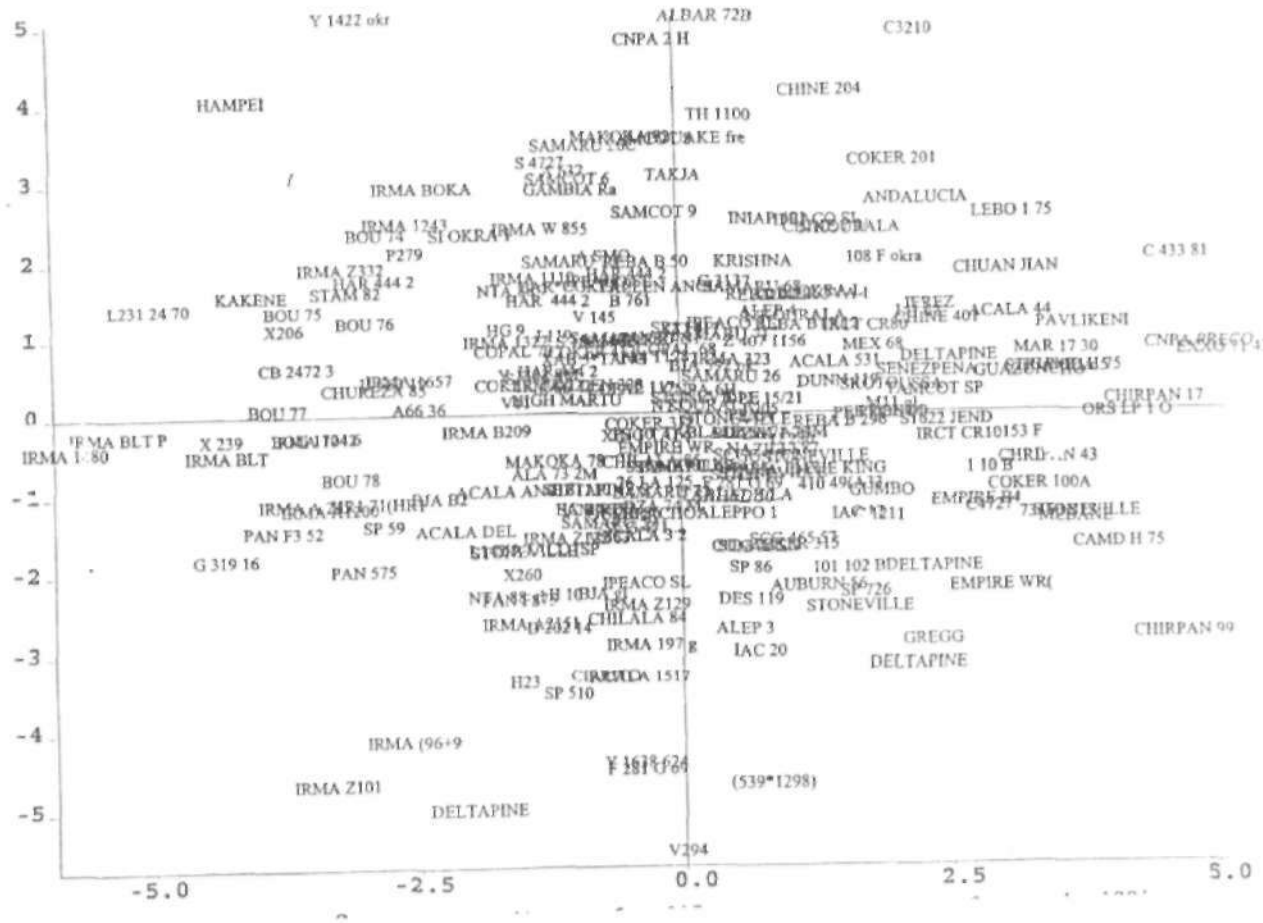


Figure 1: Satter diagram for the major traits of the first two PC axes that described the variation of 23 traits measured in 247 genotypes grown at Samaru in 1994.

H = Plant height (cm)
ENTP = Number of internodes per plant
BF = Numbers of : Fruiting branches
BV = Number of . Vegetative branches
HBF1 = Height of the first fruiting branch (cm)
ENBF1 = Number of internodes of the first fruiting branch
LBF1 = Length of the first fruiting branch (cm)
CAP = Number of of bolls per plant
LCV = Boll length
DCV = Boll diameter
JO1F = Days to first flower opening
NDENT = Number of teeth per bract
PICF = Days to 50% flowers opening
FLPP = Number of flowers per plant
R1 = Earliness (%)
RDT = Yield per plant (g)
FIB = Ginning percent
S.I. = Seed index (g)
SL25 = 2.5% span length (mm)
SL50 = 50% span length (mm)
T = Fibre bundle strength (g/tex)
MIC = Micronare index
CBV = Percent bolls produced by vegetative branches

Figure 2: Scatter diagram for 247 genotypes grown at Samaru in 1994.



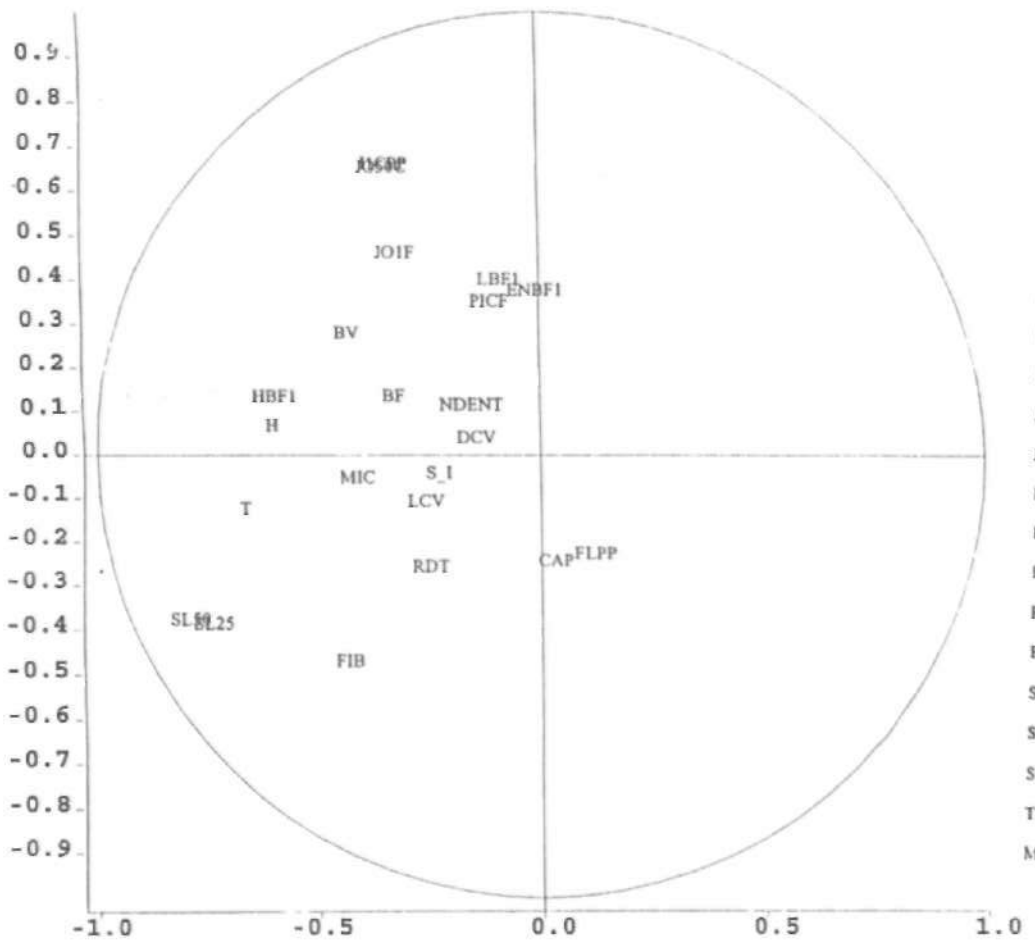
coefficient (in absolute values) in the second axis. The third principal component axis explained 9.1% of the total variation and was mainly associated with the length (-0.459) and internodes (-0.455) of the first fruiting branch, number of flowers per plant (-0.348), number of fruiting branches (-0.318), and number of bolls per plant (-0.310). Traits that loaded the fourth principal component axis (7.2% of the total variation) were length (0.421) and diameter of boll (0.374), days to 50% bolls opening (-0.348), days to first boll opening (-0.348) and number of flowers per plant (-0.302). The principal component axis five that accounted for 7.0% of the total variation had its highest loadings with boll length (0.453) and diameter (0.433), and seed-index (0.374).

The configuration of the 247 upland cotton genotypes and the traits used to describe them under the first two principal components are presented in Figures 3 and 4. In general, the genotypes from the African countries were found at the left (Figure 4) and were associated with fibre quality traits and days to first and 50% bolls opening and days to first flower opening (Figure 3). The genotypes from Cameroon (IRMA) were well differenciated and appeared to be associated with fibre length and earliness thus, indicating an agronomical superiority of local cultivars in their region of adaptation to cultivars from other regions. The genotypes from Côte D'Ivoire appeared to be associated with plant height and late maturity period. The genotypes from U.S. appeared more scattered compared to those from other regions, indicating the presence of more genetic diversity in that upland cotton population. The genotypes from Europe, unlike at Samaru appeared to be late maturing as genotypes from other regions.

Table 11: Principal component (PC) analysis of traits associated with 247 upland cotton genotypes grown in Maroua in 1994, showing eigen values and proportion of variation associated with the first five PC axes and eigen vectors of traits.

	PC1	PC2	PC3	PC4	PC5
Eigen value	3.588	2.340	2.009	1.591	1.540
Cummulative proportion of variation	16.307	26.945	36.075	43.308	50.307
Plant height (cm)	-0.318	0.041	-0.237	0.187	-0.220
No. of fruiting branches	-0.173	0.085	-0.318	0.1341	-0.056
No. of vegetative branches	-0.228	0.178	-0.141	0.182	-0.168
Height of the first fruiting branch (cm)	-0.315	0.084	-0.006	0.206	-0.246
Days to 50% boll opening (DAP*)	-0.185	0.424	0.099	-0.348	0.239
No. of internodes of the first fruiting branch	-0.004	0.242	-0.455	0.036	0.055
Lenght of the first fruiting branch (cm)	-0.047	0.258	-0.459	0.015	0.005
No. of bolld per plant	0.017	-0.163	-0.310	-0.280	0.083
Boll length (mm)	-0.137	-0.073	0.046	0.421	0.453
Boll diameter (mm)	-0.077	0.022	0.038	0.374	0.433
Days to first boll opening (DAP)	-0.182	0.429	0.134	-0.332	0.271
Days to first flower opening (DAP)	-0.170	0.298	0.195	-0.026	-0.192
No. of teeth per bract	-0.082	0.072	0.160	-0.008	0.033
Days to 50% flower opening (DAP)	-0.058	0.226	0.145	-0.018	0.045
No. of flowers per plant	0.065	-0.152	-0.348	-0.302	0.243
Yield per plant (g)	-0.131	-0.172	-0.226	-0.084	0.284
Ginning percent	0.228	-0.311	0.059	-0.257	0.021
Seed index (g)	-0.121	-0.030	0.093	0.182	0.374
2.5% span length (mm)	-0.390	-0.255	0.058	-0.109	-0.005
50% span length (mm)	-0.417	-0.249	0.048	-0.125	-0.048
Fibre bundle strength (g/tex)	-0.351	-0.084	0.003	0.004	-0.057
Micronaire reading	-0.217	-0.037	0.066	-0.131	0.024

* DAP : Days after planting



H = Plant height (cm)

BF = Numbers of · Fruiting branches

BV = Number of Fruiting branches

HBF1 = Height of the first fruiting branch (cm)

JO50C = Days to 50% boll opening

ENBF1 = Number of internodes of the first fruiting branch

LBF1 = Length of the fisrt fruiting branch (cm)

CAP = Number of of bolls per plant

LCV = Boll length

DCV = Boll diameter

JICPP = Days to first boll opening

JO1F = Days to first flower opening

NDENT = Number of teeth per bract

PICF = Days to 50% flowers opening

FLPP = Number of flowers per plant

RDT = Yield per plant (g)

FIB = Ginning percent

S.I. = Seed index (g)

SL25 = 2.5% span length (mm)

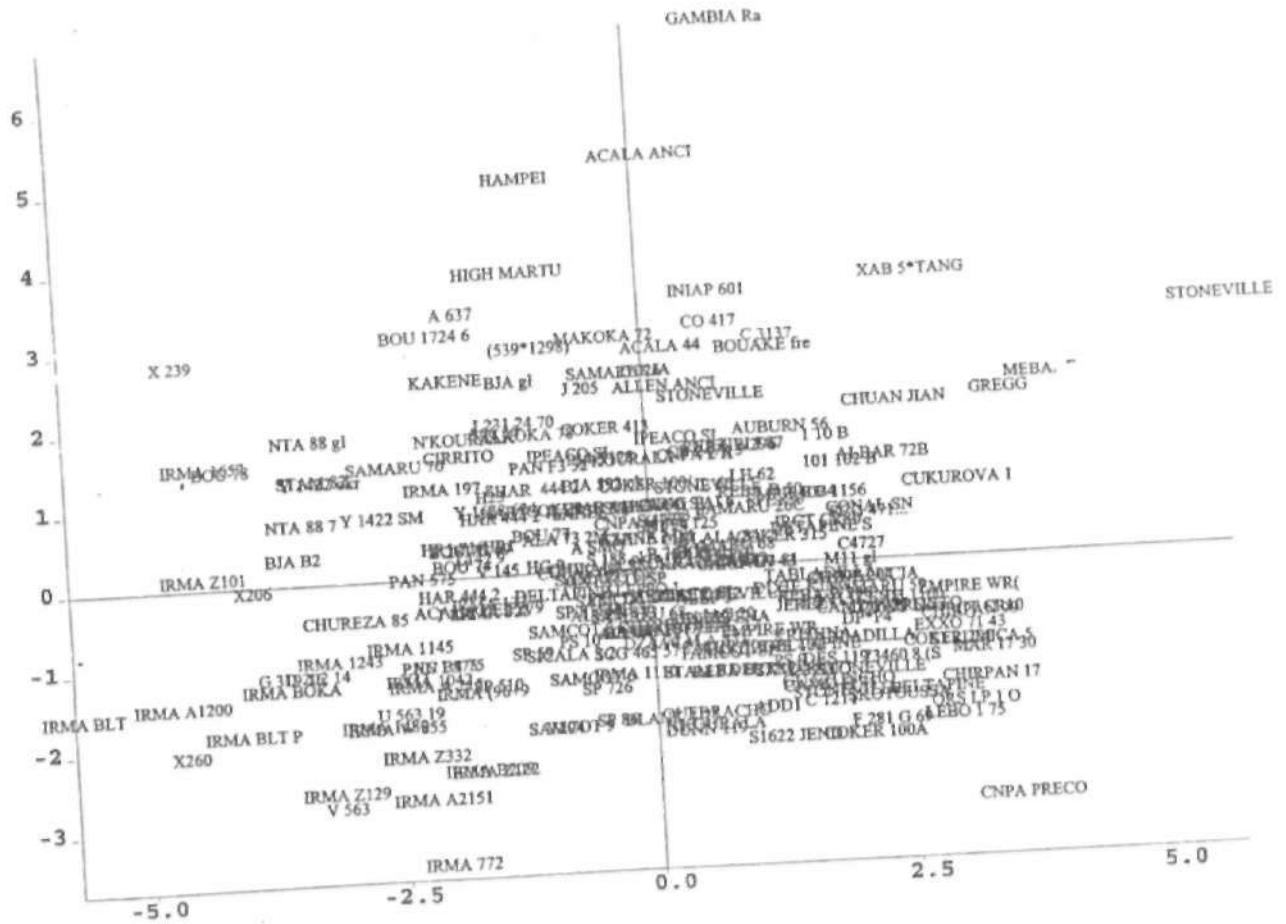
SL50 = 50% span length (mm)

T = Fibre bundle strength (g/tex)

MIC = Micronare index

Figure 3: Scatter diagram for the major traits of the first two PC axes described the variation of 22 triats of 247 Upland Cotton genotypes grown at Maroua in 1994.

Figures 4: Scatter diagram for the 247 genotypes of Upland Cotton grown at Maroua in 1994.



4.3.3. Garoua Location

The principal component analysis grouped the 22 traits measured on the 247 Upland cotton genotypes into 20 components with eigenvalues lying between 0.169 and 4.285 (Table 12). The first two principal components accounted for 19.5 and 11.8%, respectively, thus, accounting for 31.3% of the total variance. The first five components accounted for more than 54% of the total variation while the first fourteen for more than 90% of the total variation.

Also presented in Table 12 are the scores for the major traits describing the first five axes. The first principal component axis had the highest loading on 50% span length (-0.356), 2.5% span length (-0.334) and height of the first fruiting branch (-0.341). Plant height (-0.318), number of vegetative branches (-0.285), and seed cotton yield (-0.283), also accounted for the variation. The largest coefficients in the second axis were determined by boll length (0.438), boll diameter (0.393), seed index (0.387), number of internodes of the first fruiting branch (-0.309), and length of the first fruiting branch (-0.303). Traits that loaded principal component axis three (9.7% of the total variation) were number of flowers (0.411) and bolls (0.391) per plant, days to first flower opening (-0.359), and days to 50% flowers opening (-0.332). Principal component axis four (7.4% of the total variation) consisted mainly of morphological parameters [internodes of the first fruiting branch (0.440), boll diameter (0.380), length of the first fruiting branch (0.371), boll length (0.322), and bract's teeth (0.344)], while principal component axis five accounted for 6.2% of the total variance and consisted mostly of days to 50% (0.518) and days to first flower (0.476) opening and bolls (%) on vegetative branches (-0.305).

The configuration of the traits that described the 247 Upland cotton genotypes under the first two principal components are presented in Figures 5 and 6. Though there was a poor

resolution for individual groups, the genotypes could be sorted into two large groups, associated with the good fibre lengths and higher yield at the left side including genotypes from African countries: Cameroon, Chad, Côte d'Ivoire, Burkina Faso, and Togo. This indicates that the local cultivars from these African countries were well adapted with improved fibre quality traits.

4.4. **Hierarchical clustering.**

Hierarchical clustering carried out on each data set using Ward's minimum variance method delineated more clearly the groups of individuals and the results of the analyses are presented in the form of dendrograms of Figures 7, 8 and 9 (for purpose of good visualization, the values of semi-partial R-squared were multiplied by 1,000).

A fairly good resolution of individuals groups was expected with results from Samaru due to the fact that a maximal amount of variance was shown on the first two principal component axes. Thus cluster arrangements using Samaru data set served to determine *where to cut the "tree". Following the methodology described by Mumm and Dudley (1994)* and with the help of the knowledge gathered on the genetic material under study, the "tree" was cut at the semi-partial R-squared ≥ 100 level which corresponded to about 0.90 dissimilarity level. The semi-partial R-squared distance is a ratio of the variance between two clusters just joined divided by the total variance. It varies from 0 to 1. The same procedure was followed and applied to cluster arrangements of the two other locations. In general, the grouping agreed with pedigree information on the materials used.

The magnitude of intercluster distances is low, indicating a low degree of divergence between these clusters, thus a narrow genetic base of the upland cotton genotypes studies.

4.4.1. **Clustering at Samaru**

Following the methodology described hitherto, clustering genotypes based on dissimilarity of quantitative traits produced eight groups (Figure 7):

Cluster I (Appendix 6): It is a heterogeneous group comprised of 43 genotypes involving many stocks (Lone Star, Acala, Deltapine, Nkourala), and materials from Europe, and China. Samcot 9 from Nigeria was also assigned to this group (Allen type).

Cluster II (Appendix 7): Comprises 14 genotypes from Europe and America.

Cluster III (Appendix 8): Comprises 29 genotypes, designated as European material but actually they are not known precisely.

Cluster IV (Appendix 9): Comprises 41 genotypes all having in their genetic background an Allen type.

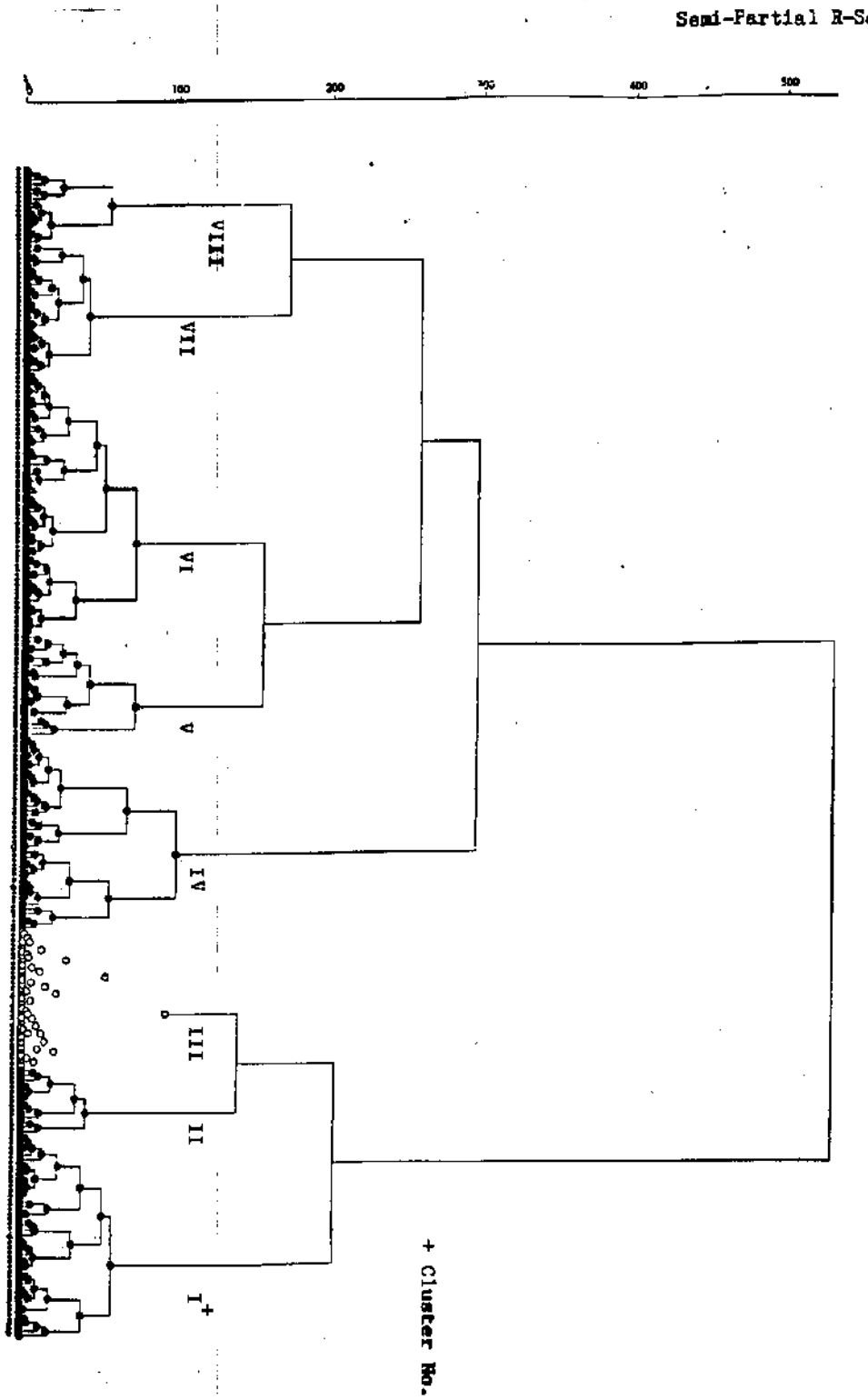
Cluster V (Appendix 10): Comprises 21 genotypes, half of which are of unknown genetic background. Here again we have material derived from crosses between Allen or Acala types and Lone Star types.

Cluster VI (Appendix 11): Comprises 56 genotypes derived from Lone Star stock (30%), a mixture of Allen*Nkourala with other US materials, some ATH (Arboreum*Thurberi*Hirsutum) derived types. Most of the materials are of complex genetic background; and 32% of them were of unknown genetic origin.

Cluster VII (Appendix 12): Comprises 27 materials from 6 different countries (French speaking African Countries) developed by, or with the collaboration of, the French Institute of Cotton Research (IRCT) and derived from HAR type.

Cluster VIII (Appendix 13): Comprises 16 genotypes from 5 different countries (3 continents), originating mainly from the old Allen*Nkourala*Triumph stock; sometimes one type is replaced by another US material. In this group, four discrepancies

Figure 7: Dendrogram for 247 genotypes of Upland Cotton grown at Samaru in 1994.



were noted with the inclusion of an Acala type, one Deltapine type and two HAR (*Hirsutum*Arboreum*Raimondi*).

4.4.2. **Clustering at Maroua.**

Figure 8 presents the results of cluster arrangements for Maroua. Compared to Samaru results, the resolution of individuals into groups was poor as shown by principal component analysis. Seven population clusters were formed:

Cluster I: Comprises 20 genotypes mainly developed in Cameroon, thus from Allen (IRCO 5028 type).

Cluster II: Comprises 48 genotypes involving Allen*Nkourala*Triumph old stock.

Cluster III: Comprises 28 genotypes of diverse genetic and geographic sources.

Cluster IV: Comprises 24 genotypes (Acala, Coker, Lone Star, Tanguis, Chinese material...).

Cluster V: Comprises 53 genotypes involving material from Europe, Texas programme, Lone Star stock.

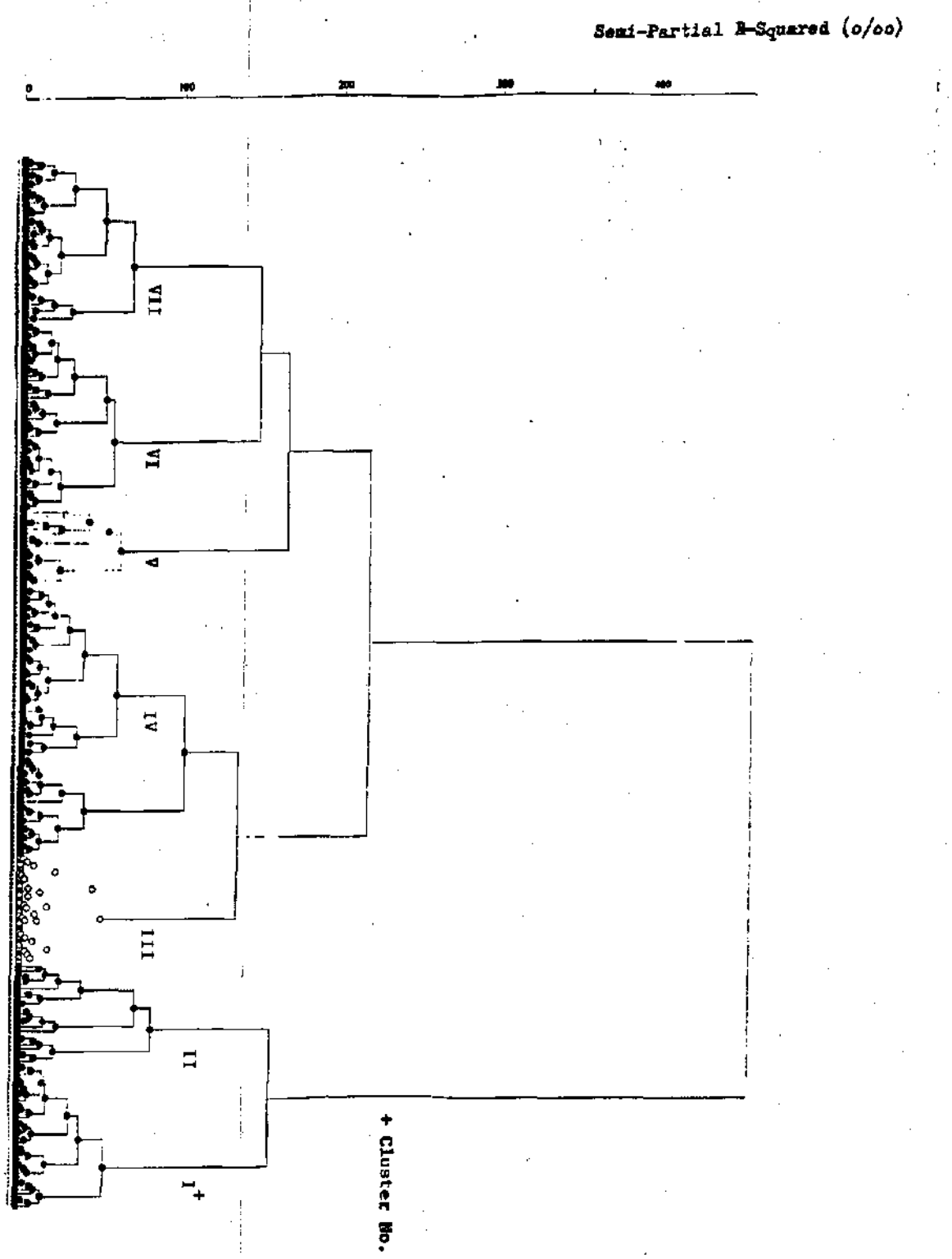
Cluster VI: Comprises 35 genotypes including Lone Star, HAR, Acala types...

Cluster VII: Comprises 39 genotypes belonging to Allen and HAR stocks mainly. Eleven genotypes are of unknown background, and two are of Deltapine and Acala types.

The last two groups are relatively uniform and place materials from the same country side by side.

However the compositions of the individual clusters is not here-in presented because the print-out do not include all the genotypes.

Figure 8: Dendogram derived from cluster analysis for 247 genotypes of Upland Cotton grown at Maroua in 1994.



4.4.3. **Clustering at Garoua.**

Figure 9 presents the results of cluster arrangements for Garoua. From the results of principal component analysis, the portion of variation explained by the first two axes was fairly good compared to Maroua even though it was still below average. The 247 genotypes under study were classified into 6 different groups (Appendixes 14 - 19):

Cluster I (Appendix 14): This group includes 18 genetic materials developed by, or with the collaboration of, the French Institute of Cotton Research (IRCT) in Chad, Cameroon, Côte d'Ivoire and Togo (Allen types predominantly).

Cluster II (Appendix 15): This group comprises 40 genetic materials derived from HAR and developed in Cameroon and Côte d'Ivoire.

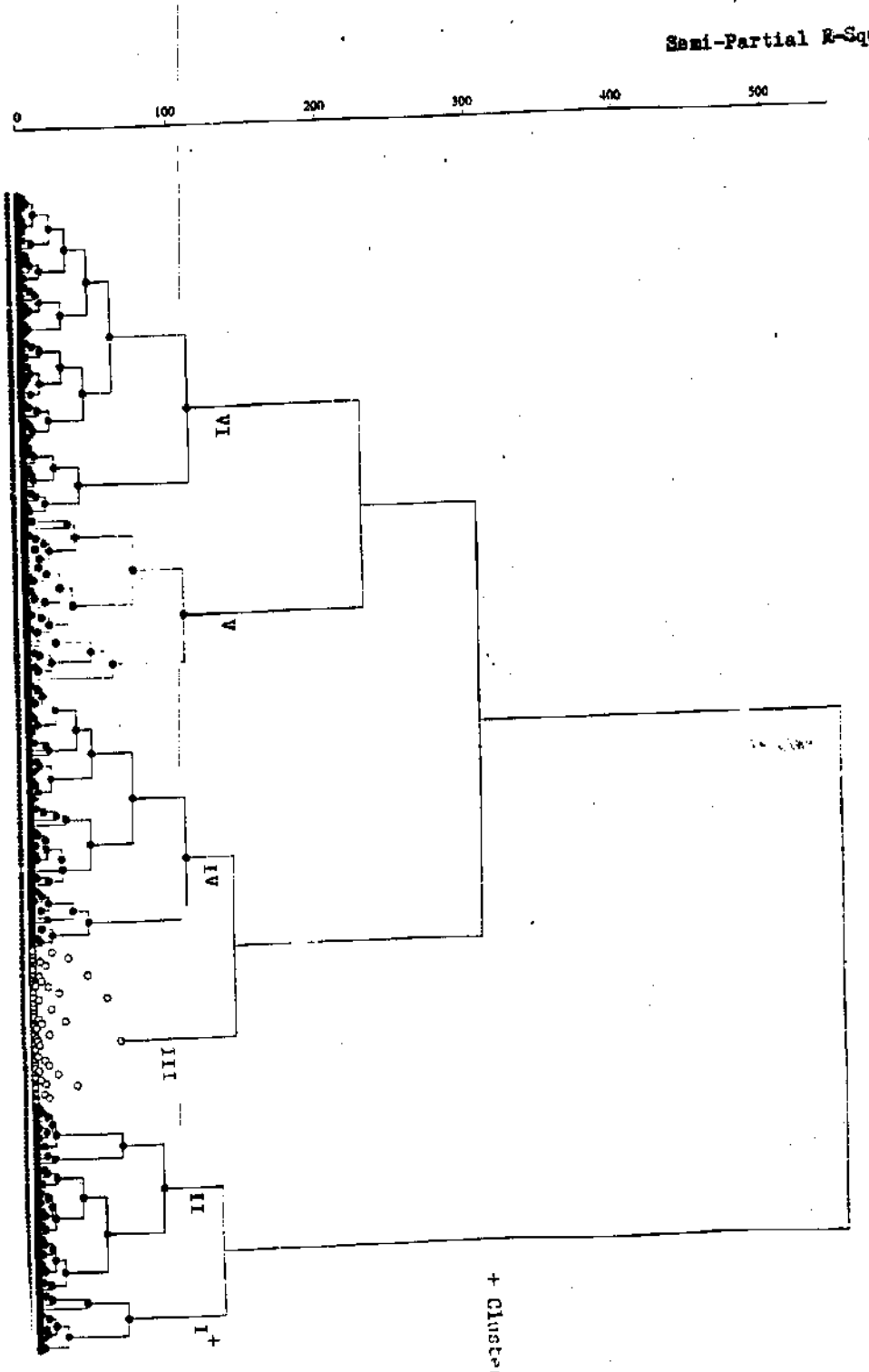
Cluster III (Appendix 16): This group comprises 33 genetic materials from Europe (former USSR and Central Europe). Some Lone Star (Empire and Coker types) and other materials of unknown origin are also present. This group is a little bit genetically heterogeneous.

Cluster IV (Appendix 17): In this group of 57 genotypes, we have the *Allen*Nkourala*Triumph* old stock, some Lone Star, Acala and Deltapine derived types.

Cluster V (Appendix 18): This group is comprised of 35 genotypes from diverse origins both genetic and geographic such as ATH types from Zaire, Lone star types from U.S.A., and HAR and Allen types from Côte d'Ivoire, Cameroon, and Nigeria. Many materials of unknown genetic background are included in this cluster.

Cluster VI (Appendix 19): This group is comprised of 69 individuals, predominantly Lone Star stock,some Allen and ATH material, and also European materials.

Figure 9: Dendrogram derived from cluster analysis for 247 genotypes of Upland Cotton grown at Garoua in 1994.



4.5. **Inter-cluster distances.**

Since the first ten principal component scores were used in each case for clustering procedures, and owing to the relationship between generalized distances and principal components, the same transformed sets of data were used to compute Mahalanobis D^2 distance between clusters.

The calculated D distances ($\sqrt{D^2}$) between clusters vary from 1.84 between groups 3 and 8 to 3.49 between clusters 1 and 4, at Samaru (Table 12). At Maroua location distances vary from 1.88 between groups 3 and 6 to 3.25 between groups 8 and 7 but were smaller compared to those separating different groups obtained at Maroua location (Table 13) while at Garoua the distances (Table 14) between population groups were medium compared to the other two locations. The smallest D(1.2) was found between clusters 1 and 2, and the highest (3.8) between clusters 2 and 6.

Table 13:Inter-cluster average values of D (Mahalanobis distance) for 247 genotypes of Upland cotton grown at Samaru in 1994

Cluster No	1	2	3	4	5	6	7	8
1	0.0000							
2	2.5022	0.0000						
3	2.1170	2.7748	0.0000					
4	2.2495	3.1033	2.2165	0.0000				
5	2.5332	3.0252	3.2658	2.6285	0.0000			
6	1.8374	2.6908	1.9100	2.2140	2.1207	0.0000		
7	2.5777	3.3827	3.1025	2.1427	2.7618	2.1084	0.0000	
8	3.1631	2.6606	2.6296	2.9096	3.4876	2.3297	2.4042	0.0000

Table 14: Inter-cluster average values of D (Mahalanobis distance) for 247 genotypes of Upland cotton grown at Maroua in 1994

Cluster No	1	2	3	4	5	6	7
1	0.0000						
2	2.4277	0.0000					
3	2.5491	2.2513	0.0000				
4	2.9179	2.1537	2.3332	0.0000			
5	2.4602	1.8779	1.9874	1.9420	0.0000		
6	3.2531	2.3002	2.7017	2.7158	2.1328	0.0000	
7	2.4451	2.0654	2.5648	2.9585	2.1234	2.2175	0.0000

Table 15: Inter-cluster average values of D (Mahalanobis distance) for 247 genotypes of Upland cotton grown at Garoua in 1994

Cluster No	1	2	3	4	5	6
1	0.0000					
2	2.5124	0.0000				
3	3.5014	2.4688	0.0000			
4	2.7390	2.3073	1.8499	0.0000		
5	3.7516	2.5384	2.3059	2.1749	0.0000	
6	2.3852	2.0011	2.0600	1.8651	1.8170	0.0000

CHAPTER V

DISCUSSION

A review by Meredith (1984) has shown that a considerable amount of useful variability is present in cotton. In this study, the presence of variability in the populations of upland cotton evaluated with respect to the 25 traits measured is confirmed by the values of ranges and means of these traits and also by the significant mean squares that were obtained among the cotton genotypes. Similar results have been reported in different collections of upland cotton by Lefort and Schwendiman (1974), Meredith (1984), Singh and Gill (1984), Brown (1991) and Lançon *et al.* (1992). The low values of coefficient of variation (2.5% to 22.1%) obtained for most of the traits is indicative of intra-genotypic homogeneity for these traits.

The relative magnitudes among different variances and components of variance are of interest to the geneticist in an academic way. Their practical significance, however, is in the estimation of heritability, i.e. the proportion of phenotypic variance that is transmissible from generation to generation. In this study, the high magnitudes of genotypic variances obtained as a component of the phenotypic variances even though inflated by genotypes x location effects that could be subtracted indicates that selection for these traits is likely to be effective.

The correlation coefficients determined for the combinations among the 25 traits in this study are estimates of phenotypic correlations. Although phenotypic correlations are admittedly only approximations of genotypic correlations, Miller *et al.* (1958) found them to be of comparable magnitude with a general tendency for the genotypic correlation to be slightly larger than its phenotypic counterpart. As a consequence, phenotypic correlations

give a fairly good indication about the direction of change in a given trait indirectly brought about when selection is actually being practiced on another trait.

The examination of the correlation coefficients between combinations of traits showed their interdependence and revealed the antagonism of expression between agronomic and fibre technology properties. In cotton both yield (seed cotton and lint) and fibre properties affect the economic net worth of the crop. Thus, many quantitative inheritance studies are directed mainly towards yield, the components of yield (lint per cent, bolls per plant and boll size, and seed-index) and fibre properties (fibre length, strength and fineness) (Meredith, 1984). The significant correlation obtained between seed cotton yield and number of bolls per plant, boll length and diameter, and seed-index compared fairly well with those reported elsewhere (Lefort and Schwandiman, 1974; Alabi, 1982; Meredith, 1984) and are interesting for future breeding programmes aimed at improving seed cotton yield. As stated by Robinson *et al.* (1952), the degree of phenotypic and genotypic correlations of traits is important in plant breeding particularly for quantitatively inherited traits because if genetic correlations are positively significant, selection for one trait will simultaneously result in changes (progress) of other traits (indirect selection). Lint yield was not measured but lint percentage (ginning percent) is one of its important component. The significant and positive associations of ginning % with fibre length and micronaire index over the three locations, and with seed cotton yield and fibre bundle strength over two locations are contrary to what is reported from other studies but could also be the result of genetic selection progress for these traits due to selection pressure that has been put on since the 1980s to improve lint yield, length, strength and fineness along with acceptable seed cotton yield. Similar results in different population of upland cotton were reported by Butany *et al.* (1966). Another evidence of

selection progress is brought about by the significant and positive association between fibre length and fibre bundle strength (Meredith and Bridge, 1971).

Earliness of crop maturity is also of interest to breeders. Days to first flower opening and days to 50% flower opening were negatively ($P \leq 0.01$) associated with earliness. This clearly suggests that in future breeding programme selection of strains early in days to first flower opening or days to 50% flower opening will help to improve earliness of upland cotton maturity.

The numerous phenotypic measurements could be resolved into fewer, more interpretable, and more easily visualized dimensions by the results of two multivariate techniques, namely, principal components and hierarchical cluster analyses.

The aim of principal component analysis is to resolve the total variation of a set of traits into linear independent composite traits which successively account for the maximal variability in the data. Sample traits are generally intercorrelated to varying degree, and hence not all principal components are needed to summarize the data adequately. Deciding upon the number of axes to retain in a reduced solution is subjective. In this study the first two principal component axes were retained. Less than 37% of the combined variance of all traits could be captured in them to summarize differences among samples of genotypes at each location. In this case the reduction in dimensionality was accompanied by a considerable loss of information. The first two components that explained 26.94, 31.29, and 34.49% of the total variation at Maroua, Garoua and Samaru, respectively, permit to equally ascribe the variability observed to agronomic and fibre technology traits. The first axis gives an account of an association between fibre technology traits (2.5% and 50% span lengths) on one hand and height of the first fruiting branch and/or plant height on the other hand, all opposed to earliness (in Samaru). The second one gives an account of the number and length

of fruiting branches, and size of bolls and seeds. In distinguishing among the genotypes therefore, fibre length, height of the first fruiting branch and/or plant height, and earliness would be most effective. Siddiqui *et. al.* (1976, 1977) found height and number of fruiting branches to be most important in contributing to divergence in different collections of upland cotton. Carvalho *et. al.* (1995), also found plant height, fibre percentage and earliness to be the most divergent traits.

The grouping of these accessions based upon principal component analysis showed that local cultivars tend to be agronomically superior in their region of adaptation to cultivars from other regions. African accessions distinguished themselves with yield and good fibre quality properties while upland cotton genotypes from Europe were early maturing crops. The differential response of the genotypes observed in the three locations indicated that environment plays a great role in determining the sample of traits for the trend of genotypes under study. Thus, there is a need for developing cultivars for each specific location.

The results of principal component analyses were not very discriminating, but one would have an overview of the relationships among the upland cotton genotypes and relate the placement of different groups or families to the traits accounting for that variability. Upland cotton genotypes from African countries showed superiority in the fibre quality traits, fibre length, bundle strength and micronaire index and could be potential sources of parental material in a breeding programme aimed at improving fibre quality properties. Specifically, genotypes from Nigeria generally had good yields and could therefore be used as parental materials to improve seed cotton yield in developing new varieties. Genotypes from European countries that were associated with earliness could be promising materials in developing early maturing varieties.

The traits measured were not very discriminating because they showed great interdependence among themselves. Selection of traits to calculate the genetic distance is the base of whatever classification (Arunachalam, 1981; Sanchez *et al.*, 1993; Amurrio *et al.*, 1995). The quantitative traits affected by environmental factors and interactions and showing high degrees of independence among themselves must be considered as criteria for classification (Goodman and Paterniani, 1969)New parameters that are now increasingly used to describe cotton crops such as "plant mapping" traits could be trully-descriptive and should be considered in further genetic diversity studies using multivariate analysis of biometrical data. Morphological traits determining yield can be readily described in such a study, and this would be very beneficial to cotton breeders.

Results from the cluster analysis also suggested that the hierarchical cluster method could be useful for determining the diversity among numerous genetic materials or the similarity of an entry relative to other entries (Souza and Sorrels, 1991a; Ouendeb *et al.*, 1995). It can be used to assign entries to their specific group of similarity, even when origin is unknown. Eight, seven and six distinct clusters were formed at Samaru, Maroua and Garoua, respectively.

The variation observed in the clustering in different environments could be due to differential response of different traits in the different environments. Considering the placements of the different genotypes on the scatter diagrams resulting from the principal component analyses, the first agglommerations (minimum variance) formed were among genotypes with more extreme scores. It is noteworthy that very divergent genotypes by geographic origin could fall into the same cluster: cluster I in Samaru and Cluster VI in Garoua contained upland cotton genotypes from the five continents. Some genotypes of the same genetic background but representing different selections could be distributed into

different clusters as observed in IRMA 1243 and IRMA (96 + 97) from Cameroon and the SAMCOT cultivars from Nigeria which have been placed into different clusters. Some genotypes of unknown pedigree but from the same ecogeographic origin could be distributed in different clusters. All these findings tend to show that hierarchical cluster analysis is a sensitive tool for measuring diversity among accessions. Conversely, genotypes of the same genetic origin could be distributed into different clusters as observed with the IRMA from Cameroon that are of IRCO type and SAMCOT 9 and SAMCOT 10 of Nigeria which fell into different clusters both at Samaru and Garoua.

Free clustering of the genotypes indicated dependante upon the directional selection pressure applied for realizing acceptable yield (lint and seed cotton) and fibre quality potentials in different countries and thus the well-evolved homeostatic devices will favour constancy of the associated traits and will thus show indiscriminate clustering. Substantial variability among the genotypes evolved in the same habitat could be due to genetically diverse nature of the parents used in developing genotypes, or genetic drift and selection forces under diverse environments (Katiyar, 1978).

With very few exceptions, the grouping patterns of the upland cotton genotypes as observed in the present study indicated that the genetic diversity was not directly related to geographic diversity. These results compare very well with those obtained in previous studies by Singh and Bains (1968), Ekbote and Khorgade (1986), Singh and Gill (1984), Gulin *et al.* (1990) and Brown (1991). Arunachalam and Ram (1967), on sorghum also reported similar results. Quite often, the quantitative traits like the ones measured in this study are economically important in the development of varieties. In order to bring together varieties with generally similar characteristics, quantitative traits are used for grouping. Nevertheless,

qualitative traits are often used for separating varieties when a limited range for quantitative traits is found within certain groups (Sneedon, 1970).

Despite the descrepancies observed in clustering, clusters broadly agreed with pedigree informations. Mumm and Dudley (1994) explained that discrepancies can result because of the methodology in the clustering process. First, since the clusters obtained with Ward's minimum variance method are non-overlapping, a genotype resulting from the cross of two inbreds could be grouped with only one of the parents. So, the grouping in the dendrogram is somewhat artificial in that group assignments indicate the group with which the line is similar. The dendrogram serves to portray the relative estimated relationships in the proximity matrix as closely as possible and in a manner that is easy to interpret, but it is not a perfect representation of those relationships.

The low estimates of inter-cluster distances (D) obtained in this study indicate that the 247 genotypes display a narrow range of diversity at each specific environment. An explanation for this low range of genetic diversity in the sampled population of upland cotton could be given. The development of the World upland cotton varieties are to a large extent dependent on the United States germplasm pool (Dessauw and Hau, 1997) which have been found to display a narrow genetic diversity (Richmond, 1951; Anonymous, 1997). These varieties are therefore expected to display only a limited amount of diversity despite their diverse geographic origin. But due to the interspecific origin of Upland cotton and search for a fairly broad genetic base in cotton in recent years by breeders, genetic diversty is still represented in cotton germplasm for breeders to work with.

The practical significance of grouping genotypes into different clusters and computing the statistical distance between them has been emphasized by Bhatt (1970) and Kanwal *et al.* (1983). Genotypes grouped into the same cluster presumably diverge very little from one

another as regards the aggregate of the traits measured. A crossing of genotypes belonging to the same cluster would not be expected to yield desirable segregates. Consequently crossing programme should be formulated such that the spectative parents belong to different clusters. The statistical distances (D), thus computed, represent the index of genetic diversity among clusters. It would appear logical to effect crosses genotypes belonging to the clusters seperated by high estimated statistical distances, and mean statistical distance at each location may be considered arbitrary as a guide line. But in choósing among the genotypes of a cluster, other practical considerations should be taken into account (Bhatt, 1970).

CHAPTER SIX

SUMMARY AND CONCLUSION

With the large increases of processing power of micro-computers and corresponding software developments, grounds are believed to exist at present for geneticians and plant breeders to evaluate high numbers of accessions when considering many traits simultenously (Peeters and Martinelli, 1989; Brown, 1991; Amurrio *et al.*, 1995). In this study, the treatment in multivariate of the data relative to twenty five traits, measured in two hundred and fourty seven genotypes of upland cotton (*Gossympium hirsutum* L.) grown in three different environments brought information about the genetic diversity present in this set of genotypes and gave the descriptive elements of the total variation observed as revealed by the univariate approach.

The examination of simple linear correlation between pairs of traits revealed their great interdependance, though in general they did not show high mutual correlations. Significant correlations were observed among yield components. Highest of these were the association between seed cotton yield and number of bolls per plant ($r=0.65^{**}$, and 0.411^{**}), and seed index and boll length ($r=0.424^{**}$, and 0.426^{**}) at Samaru and Garoua, respectively. An indirect selection for yield could be effective through the improvement of either boll number or seed-index.

The first two principal component axes that explained less than 36% of the total variation (great interdependence of traits) suggested equal contribution of agronomic and fibre technology traits to the total variation of the traits considéred in the sample of genotypes studied. Height of the first fruiting branch and/or plant height, fibre length and earliness in Samaru were imporatnt components of variability and contributed significantly to the total variation observed. The representation of the genotypes on these two axes provides elements of information in view of future breeding programmes aiming at improving seed cotton yield, fibre technology properties and earliness. Genotypes originating from African countries were

well adapted to their region and agronomically superior to other genotypes. Samaru materials (Nigeria) were well associated with seed cotton yield while materials from African French speaking countries (Cameroon, Chad, Côte d'Ivoire and Togo) were associated with good fibre technology properties. Materials developed in Europe distinguished themselves with earliness in maturity. These could be recommended as potential sources of parental material if one wishes to improve these traits in a breeding programme.

Hierarchical cluster analysis using ward's minimum variance grouped the two hundred and fourty seven genotypes of upland cotton into 8, 7 and 6 clusters at Samaru, Maroua and Garoua, respectively. The grouping pattern of the genotypes indicated that the genetic diversity was not directly related to the geographic diversity. There was also variation in the clustering in different environments presumably due to differential response of different traits in varying environments. It is obvious for the genetic relationship of genotypes to be obscurred by adaptation traits related to genotype x location variances.

Results from this study suggest that hierarchical cluster analysis could become a useful tool for the evaluation of a high number of genetic materials in several environments. Considering some agronomic and fibre technology traits, different genetic groups could be more or less clearly delineated, and clusters formed could provide a structure for sampling cultivars for genetic, breeding, or agronomic studies. There was correspondance between the geographic origin of the genotypes and their inclusion in particular clusters. Genotypes grouped into the same cluster presumably diverge very little from one another as regards the aggregate of the traits measured. Crossing of genotypes belonging to the same cluster is not likely to yield desirable segregantes. Consequently a crossing programme should be formulated such that the putative parents belong to different clusters (Bhatt, 1969; Ouendeba *et al.*, 1995). The categorisation of the diversity of many "entries" into groups with similar characteritics can also be used in building "Core" collections or as a prescreening programme for plant breeders (Herbert and Vincourt, 1985; Lefort-Buson, 1985; Ariyo, 1993; Casler, 1995).

REFERENCES

Agarkova, S. N., S. N. Chekalin, M. D. Varlakhov, E. I. Makogonov, 1980. [Genetic divergence and its use in breeding.] *S. Kh. Za rubzhom* 6, 18-21.

Alabi, S.O., 1982. Genetic study of quantitative characters in cotton, *Gossypium hirsutum*, L. using Diallel cross analysis. M.Sc Dissertation, Department of Plant Science, Faculty of Agriculture, A.B.U., Zaria, Nigeria.

Amurrio, J.M., A.M. de Ron, and A.C. Zeven, 1995. Numerical taxonomy of Iberian landraces based on quantitative and qualitative characters. *Euphytica 82*: 195-205.

Anonymous, 1997. Varieties: Development and Selection. *Cotton Physiology Today*, 8, 13-23.

Ariyo, O. J., 1987. Multivariate analysis and the choice of parents for hybridization in okra (*Abelmoschus esculentus* L. Moench). *Theor. Appl. Genet. 74*: 361-363.

Ariyo, O. J., 1990. Measurement and classification of genetic diversity in okra (*Abelmoschus esculentus*). *Ann. Appl. Biol. 116*: 335-341.

Ariyo, O. J., 1993. Genetic diversity in West African okra (*Abelmoschus caillei*) (A. Chev.) Stevels-Multivariate analysis of morphologica' and agronomic characters. *Genet. Res. & Crop Evol. 40*: 25-32.

Arunachalam, V., 1981. Genetic Distance in Plant Breeding. *Indian J.Genet. Plant Breed.* 41: 226-236.

Arunachalam, V. and J. Ram, 1967. Geographic diversity in relation to genetic divergence in cultivated sorghum. *Indian J. Genet. Pl. Breed. 27*: 369-380.

Bainiwal, C.R. and D.S. Jatasra, 1980. Genetic divergence in Pigeon pea. *Indian J. Genet. Pl. Breed. 40*(1): 153-156.

Beal, W.J. 1880. Report of the Michigan Board of Agriculture. pp. 287-288.

Bernardo, R. 1993. Estimation of coefficient of coancestry using molecular markers in maize. *Theor. Appl.Genet. 85*:1055-1062.

Bisht, I.S., R.K. Mahajan and R.S. Rana, 1995. Genetic diversity in South Asian okra (*Abelmoschus esculentus*) germplasm collection. *Ann. Appl. Biol. 126*: 539-550.

Bhatt, M.M.,1970. Multivariate anlysis approach to selection of parents for hybridization aimed at yield improvement in self-pollinated crops. *Aust. J. Agric. Res. 21*: 1-7.

Bhatt, G.M., 1973. Comparison of various methods of selecting parents for hybridization in common bread wheat (*Triticum aestivum* L.). *Aust. J. Agric. Res. 24*: 457-467.

Bhatt, G.M., 1976. An application of multivariate analysis to selection for quality characters in wheat. *Aust. J. Agric. Res. 27*: 11-18.

Brown, J.S., 1991. Principal Component and Cluster Analysis of Cotton Cultivar Variability across the U.S. Cotton Belt. *Crop Sci. 31*: 915-922.

Bryant, E.H., and R.A. Atchley, 1975. (Eds.). *Multivariate Statistical Methods*:
I: *Among-Groups Covariation.*
II: *Within-Groups Covariation.*
Dowden, Hutchinson & Ross, Inc.

Butany, W.T., M. Singh, and R.B. Mehra, 1966. Interrelationships between some characters in Hirsutum cotton. *Indian J. Genet. Pl. Breed. 26*:262-268.

Camussi, A. 1979. Numerical taxonomy of Italian populations of maize based on quantitative traits. *Maydica 24*: 161-174.

Camussi, A. and E. Ottaviano, 1981. Use of genetic effects of quantitative traits for the multivariate analysis of differences between populations. *In*: *Quantitative Genetics and Breeding* . Proceedings of the fourth Meeting of the Section Biometrics and Plant Breeding, September 2-4, Poitiers, France. pp. 295-304.

Carvalho L.P. (de), C.D. Cruz and C.F. de Moraes, 1995. Genetic divergence in Brazilian cotton, Gossypium hirsutum var. latifolium Hutch. *Brazilian J. Genet. 18*: 439-443.

Casler, M.D., 1995. Patterns of variation in a collection of Perennial Ryegrass Accessions. *Crop Sci. 35:* 1169-1177.

Cochran, W.G. and G.M. Cox, 1957. Experimental designs. John Wiley and Sons, N.Y.

Cox, T.S., Y.T. Kiag, M.B. Gorman, and D. Rodgers, 1985. Relationships between coefficient of parentage and genetic similarity indices in the soybean. *Crop Sci. 25*: 529-532.

Dagnélie, P. 1975. "Analyse statistique á plusieurs variables" *Gembloux-Presses Agron.*, 362p.

Demol, J.; Mergeai, G.; Hofs, J.L.; NDungo, V., 1992. (eds). L'amélioration du cotonnier. *In* Le cotonnier au Zaïre. Publication Agricole 29, Bruxelles, pp. 93-123.

Dessauw, D., et B. Hau, 1997. (Eds). *Manuel pour l'Amélioration du Cotonnier.* Session de formation du 17 février au 7 mars 1997. Programme COTON, CIRAD-CA, Montpellier, France.

Devi, B.S.; D. Subramanyam; P.V.R. Kumar; G. Saty-Anarayana; B. Krishnamurthy, 1991. Genetic divergence in Kenaf. *Indian J. Genet. Pl. Breed.*: *51*(2), 260-263.

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

Dharmalingam, V. and M. Kadambavanasundaram, 1989. Genetic divergence in cowpea (Vigna unguiculata (L) Walp). *Madras Agric. J., 76*(7): 394-399.

Duvick, D.N., 1984. Genetic Diversity in Major Farm Crops on the Farm and in Reserve. *Economic Botany, 38*(2), pp. 161-178.

Echekwu, C.A., 1991,. Status of cotton breeding in Nigeria. *Samaru Miscellaneous paper 128.* Institute for Agricultural Research, Ahmadu Bello University, Zaria. 21pp

Edye, L.A., 1976. Analysis of agronomic data from *Stylosanthes* introductions. *In* W.T. Williams (ed.) *Pattern analysis in agricultural science*. Elsevier Sci. Publ. Co., Amsterdam, the Netherlands.

Ekbote, A.P. and P.W. Khorgade, 1984, reed. 1986. Genetic divergence in Upland cotton (*Gossypium hirsutum* L.). *Punjabrao Krishi Vidyapeeth Res. Journal*, *8*(2), 1-4.

Endrizzi, J.E.; E.L. Turcotte, and R.J. Kohel, 1984. Qualitative Genetics, Cytology, and Cytogenetics. *In* R.J. Kohel and C.F. Lewis (eds) *Cotton*. Agronomy n° 24, ASA, CSSA, SSSA, Inc. Publ. Madison, Wisconsin, USA.

Fryxell, P.A., 1984. Taxonomy and Germplasm Ressources. *In* R.J. Kohel and C.F. Lewis (eds) *Cotton*. Agronomy n° 24, ASA, CSSA, SSSA, Inc. Publ. Madison, Wisconsin, USA.

Fryxell, P.A., L.A. Gaven and J. McD. Stewart, 1992. A review of *Gossypium* sect. *Grandicalyx* (Malvacea), including the description of six new species *Systematic Botany, 17*(1):91-114

Gallais, A., M. Lefort-Buson, and A. Charcobsset, 1986. Prediction of single cross values in forage Maize using biometrical distances evaluated from a two tester top cross design. *In*: Kearsey, N.J., and C.P. Werner (Eds): *Biometrics in Plant Breedding*, p. 21-30. Proceedings of the sixth Meeting of the EUCARPIA section, July 28th - August 11th, University of Birmingham, U.K. pp: 406.

Godshalk, E.B., M. Lee, and K.R. Lamkey, 1990. Analysis of the relationship of restriction fragment length polymorphisms to maize single-cross hybrid performance. *Theor. Appl. Genet. 80*: 273-280.

Gomez, K.A. and A.A. Gomez, 1984. Statistical Procedures for Agricultural Research. 2nd ed. John Wiley & Sons, Inc., N.Y. 684p.

Goodman, M. G., 1968. Genetic Distances: Measuring Dissimilarity Among Populations. *Yearbook of Physical Anthropology, 17*: 1-38.

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

Goodman, M.M., 1973. Genetic distances: Measuring dissimilarity among populations. *Yearbook of Physical Anthropology 17*: 1-38.

Goodman, M.M. and E. Paterniani, 1969. The races of Maize: III. Choice of appropriate characters for racial classification. *Economic botany 23*:265-273.

Goodman, M.M. and R.Mck.Bird, 1977. The race of maize: IV: Tentative grouping of 219 Latin American races. *Economic botany 31*: 204-211.

Gulin, V.V., Yu E. Dubrova, V.A. Kal'chenko, V.I. Muzyka, A.V. Atlanov, and D.S. Kazantsev, 1990. [Genetic differenciation of some varieties of cotton (*Gossypium hirsutum* L.).] *Genetika* (Moscova) *26*: 1515-1518.

Gupta, S.K., and S.K. Malik, 1991. Genetic divergence in wheat *Triticum* sp. at different environments. *Environment and Ecology 9*(1): 212-214.

Gutknecht, J., and J. Fournier, 1982. Contribution à l'étude de la maturité des cotons par la méthode optique. *Cot. Fib. Trop., 37*(3): 249-256.

Hallauer, A.R., 1981. Selection and Breeding Methods. *In*: Frey, K.J.: 1981 (ed.). Plant Breeding II. The Iowa State Univ. Press/Ames. 497pp.

Hallauer, A.R., and J.B. Miranda, Fo. 1988. *Quantitative genetics in maize breeding*. 2nd ed. Iowa State Univ. Press, Ames, IA.

Hamman, H.K., 1972. Utilization of multivariate statistics for discriminating among flue-cured tobacco varieties. *Tech. Bull. 212*. North Carolina Agric. Exp. St. (Raleigh, NC).

Hawkes, J.G., 1981. Germplasm Collection, Preservation, and Use.In Frey, K.J., 1981, (ed). pp 57-83. Plant Breedind II.The Iowa State Univ. Press/Ames. 497 pp.

Hearn, A.B. and G.A. Constable, 1984. Cotton. *In* P.R. Goldsworthy and N.M. Fisher (eds.) *The physiology of tropical food crops*. John Wiley & Sons, Inc. 664p.

Herbert, Y. et P. Vincourt, 1985. Distances calculées sur les critères biométriques; *In*: Lefort-Buson, M. et D. de Vienne (Eds), *Les distances génétiques: Estimations et applications*. INRA, Paris, 23-37.

Hussaini, S.H., M.M. Goodman, and D.H. Timothy, 1977. Multivariate analysis of geographical distribution of the world collecton of finger millets. *Crop Sci. 17*: 257-263.

ICAC: (Internatinal Cotton Advisory Committee), 1996. Cotton review of the World situation 1995/96, may_june. 16pp.

Illoh, H.C., and O. Olorade, 1991. Numerical taxonomic studies of mango (*Mangifera indica* L.) varieties in Nigeria. *Euphytica 51*:197-205.

Jagadev, P.N., and K.M. Samal. 1991. Multivariate analysis in niger (*Guizotia abyssinica* Cass.). Indian J. Genet. Pl. Breed. 51: 392-394.

Jensen, N.F., 1988. *Plant Breeding Methodology*. John Wiley & Sons, Inc. pp676.

Kanwal, K.S., R.M. Singh, J. Singh, and R.B. Singh, 1983. Divergent gene pools in rice improvement. *Theor. Appl. Genet. 65:* 263-267.

Kalsy, H.S., H.R Garg, P. Rathore, and J.S. Gill, 1995. Genetic divergence and heterosis in American cotton. *Crop improvement 22*:232-236.

Katyar, R.P., 1978. Genetic divergence for morphophysiological and quality determinants of yield in chickpea. *Indian J. agric. sci. 48*(8): 451-454.

Khan, I. A.; S. Imtiaz, and B.A. Maliik; 1990. Selection of diverse parents of chickpea (*Cicer arietinum* L.) by multivariate analysis and the degree of heterosis of their F_1 hybrids. *Euphytica 51*(3): 227-233.

Lançon, J., 1987. Behaviour of sixteen agronomic traits and fibre properties in two diallel crosses involving African and American varieties of Upland cotton (*Gossypium hirsutum* L.). *Cot. Fib. Trop. 42*: 255-262.

Lançon, J.: 1995. Effet de la densité de semis en sélection sur l'amélioration génétique du cotonnier: interactions, structures de correlations, hétérosis et valeur en lignée. Thèse de doctorat, Université Paris XI-Orsay, 125p.

Lançon, J., B. Bachelier, J.L. Chanselme, D. Dessauw, C. Klassou, E. N'Guessan, T.B. NGuyen, and E. Ousmane, 1992. Etude multilocal d'un diallèle à quatre géniteurs d'élite sélectionnés au sein du réseau coton africain.
I. Variabilité des localités. *Cot. Fib.Trop.* 47(4): 251-263.
II. Variabilité des parents. *Cot. Fib.Trop.* 47(4): 265-276.

Lançon, J., E. Gozé, G. Gawrysiak, B. Hau, B. Bachelier, J.L. Chanselme, D. Dessauw, C. Klassou, E. N'Guessan, T.B. NGuyen, and E. Ousmane, 1993a. Etude multilocal d'un diallèle à quatre géniteurs d'élite sélectionnés au sein du réseau coton africain. V. Effets génétiques. *Cot. Fib. Trop.* 48(4):253-282.

Lançon, J., E. Gozé, B. Hau, B. Bachelier, J.L. Chanselme, D. Dessauw, C. Klassou, E. N'Guessan, T.B. NGuyen, and E. Ousmane, 1993b. Etude multilocal d'un diallèle à quatre géniteurs d'élite sélectionnés au sein du réseau coton africain.
III. Corrélations entre variables. *Cot. Fib. Trop. 48*(1): 5-15.
IV. Hétérosis et homéostase. *Cot. Fib. Trop. 48*(1): 17-39.

Lee, J.A., 1984. Cotton as a World Crop. In Cotton, Agronomy Monograph No 24, ASA-CSSA-SSSA, Madison, WI, USA. 605 pp.

Lefort, P.L., et J. Schwendiman, 1974. Etude d'un matériel d'origine triple hybride *Gossypium hirsutum* x *G. arborium* x *G. raimondii*. I.- Application de l'analyse multivariable à la description des lignées de base. *Cot. Fib. Trop. 29*: 307-318.

Lefort-Buson, 1985.Utilisation des critéres biométriques. InLefort-Buson, M. and D. de Vienne, 1985. (ed.) *Les distances génétiques: Estimations et applications*. INRA, Paris. 181p.

Liu, L.F., 1979. [Genetic distance of quantitative characters and its estimation.] *Acta Genet. Sinica 6*: 348-355.

Lukhele, P.E., 1981. Estimation of Genetic Variability in a Sorghum Population (*Sorghum bicolor* L. Moench). M.Sc. Dissertation. Ahmadu Bello University, Zaria, Nigeria. 70 pp.

Mahalanobis, P.S. 1936. On the generalized distance in Statistics. *Proc. Nat. Inst.* (India) 2: 49.

Martinez, O.J.; M.M. Goodman and D.H. Timothy, 1983. Measuring racial differentiation in maize using multivariate distance, measures standardized by variation in F_2 populations. *Crop Sci. 23*: 775-781.

McIntosh, M.S., 1983. Analysis of combined experiments. *Agron. J. 75*:153-155.

Melchinger, A.E., M.M. Messmer, M. Lee, W.L. Woodman, and K.R. Lamkey, 1991. Diversity and relationships among U.S. maize inbreds revealed by RFLPs. *Crop Sci. 31*:669-678.

Meredith (JR.), W.R., 1984. Quantitative Genetics. *In* R.J. Kohel and C.F. Lewis (eds) *Cotton*. Agronomy n° 24, ASA, CSSA, SSSA, Inc. Publ. Madison, Wisconsin, USA.

Miller, P.A.; J.C. Williams; H.F. Robinson and R.E. Comstock, 1958. Estimates of genotypic and environmental variances and covariances in upland cotton and their implication in selection. *Agron. J. 50*: 126-131.

Milligan, G.W., 1980. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrica 45*: 325- 342.

Moll, R.H.; W.S. Salhuana, and H.F. Robinson, 1962. Heterosis and genetic diversity in variety crosses of maize. *Crop Sci. 2*: 197-199.

Moll, R.H.; J.H. Lonnquist; J. Vélez Fortuna, and C.E. Johnson, 1965. The relationship of heterosis and genetic divergence in maize. *Genetics 52*: 139-144.

Moll, R.H. and C.W. Stuber, 1974. Quantitative Genetics-Empirical results relevant to plant breeding. *Advances in Agronomy 24*:277-313.

Morrison, D.F., 1967. *Multivariate statistical methods*. McGraw-Hill Book Co., N.Y. 338 pp.

MSTATC, 1993. A software programme for the design, management and analysis of Agronomic Research Experiments. Michigan State University. Institute of International Agriculture.

Mumm, R.H. and J.W. Dudley, 1994. Classification of 148 U.S. Maize inbreds: I. Cluster analysis based on RLFPs. *Crop Sci. 34*:842-851.

Murty B.R., and V. Arunachalam, 1966. The Nature of Divergence in Relation to Breeding System in some Crop Plants. *Indian J. Genet. Plant Breed. 26A*: 188-197.

Niles, G.A., and C.V. Feaster, 1984. Breeding. p. 201-231. *In* R.J. Kohel and C.F. Lewis (ed.) Cotton. Agron. Monogr. 24. ASA, CSSA, and SSSA, Madison, WI.

Onyilagha, J.C., 1986. Numerical analysis of variation among Nigerian *Dioscorea rotundata* accessions. *Euphytica 35*: 413-419.

Ouendeba, B., G. Ejeta, W.W. Hanna, and A.N. Kumar, 1995. Diversity among African Pearl Millet Landrace Populations. *Crop Sci. 35*:919-924.

Paterniani, E., and J.H. Lonnquist, 1963. Heterosis in interracial crosses of corn (*Zea mays* L.). *Crop Sci. 3*:504-507.

Peeters, J.P., and J.A. Martinelli, 1989. Hierarchical cluster as a tool to manage variation in germplasm collections. *Theor. Appl. Genet. 78*:42-48.

Poehlman, J.M. and D. Borthakur, 1969. *Breeding Asian Field Crops*. Oxford & IBH Publishing Co. PVT.LTD.

Rao, V.R., M. Ramachandram, and J.R. Sharma, 1980. Multivariate analysis of genetic divergence in safflower. *Indian J. Genet. Pl. Breed. 40*: 73-85.

Richmond, T.R., 1951. Procedures and methods of breeding with special reference to American cultivated species. *Adv. Agron., 4*: 213-245.

Rhodes, A.M., and F.W. Martin, 1972. Multivariate studies of variations in Yams (*Dioscorea alata* L.). *J. Amer. Soc. Hort. Sci. 97*(5): 685-688.

Rodgers, D.M.; J.P. Murphy and K.F. Frey, 1983. Impact of plant breeding on the grain yield and genetic diversity of spring oats. *Crop Sci. 23*: 737-740.

Sanchez, J.J., M.M. Goodman, and J.O. Rawlings., 1993. Appropriate characters for racial classification in maize. *Econ. Bot. 47:* 44-59

SAS Institute Inc., 1988. SAS/STAT users guide: Release 6.03ed.
SAS Inst. Inc., Cary, NC.

Saunders, J.H., 1961. The wild species of *Gossypium* and their evolutionary history. Oxford Univ. Press, London.

Shull, G.H. 1908. The composition of a field of maize.
Am. Breeders' Assoc. Rep. 4: 296.

Shultz, E.F., Jr. 1955. Rules of thumb for determining expectations of mean squares in analysis of variance. *Biometrics 11*:123-135.

Siddiqui, J.A., R.C. Prasad, and R.B. Mehra, 1976. Hybrid performance in relation to genetic divergence in some varieties of *Gossypium hirsutum*. *Zeitschrift für Pflanzenzüchtung*, *77*(3), 215-221.

Siddiqui, J.A., R.B. Mehra, and V.P. Singh, 1977. *Divergence in some varieties of Upland cotton for yield and quality breeding*. Div. Genet., IARI, New Delhi, India, 858-862.

Singh, R.B. and S.S. Bains, 1968. Genetic divergence for ginning outturn and its components in Upland cotton. *Indian J. Genet. Pl. Breed. 28*:262-268.

Singh, T.H. and S.S. Gill, 1984. Genetic divergence in Upland cotton under different environments. *Indian J. Genet. Pl. Breed. 44*: 506-513.

Skiebel, K., C. Arnelang, I. Basse, H. Beelitz, 1991. Intensifying measurements in the breeding process with the help of statistical procedures. *Archiv. für Züchtungsforschury 21*(2): 113-120.

Smith, J.S.C., 1988. Diversity of United States hybrid maize germplasm: Isozymic and chromatographic evidence. *Crop Sci. 28*:63-69.

Smith, J.S.C., and O.S. Smith, 1987. Associations among inbred lines of maize using electrophoretic, chromatographic and pedigree data: I. Multivariate and cluster analysis of data from 'Lancaster Sure Crop' derived lines. *Theor. Appl. Genet. 73*:654-664.

Smith, J.S.C., and O.S. Smith, 1988. Associations among inbred lines of maize using electrophoretic, chromatographic and pedigree data: II. Multivariate and cluster analysis of data from 'Iowa Stiff Stalk' derived lines. *Theor. Appl. Genet. 76*:39-44.

Smith, J.S.C., M.M. Goodman, and C.W. Stuber, 1985a. Genetic variability within U.S. maize germplasm: I. Historically important lines. *Crop Sci. 25*:550-555.

Smith, J.S.C., M.M. Goodman, and C.V. Stuber, 1985b. Genetic variability within U.S. germplasm: II. Widely-used inbred lines 1970 to 1979. *Crop Sci. 25*:681-685.

Smith, O.S.; J.S.S. Smith; S.L. Bowen; R.A. Tenborg; S.J. Wall, 1990. Similarities among a group of elite maize inbreds as measured by pedigree F_1 grain yield, grain yield, heterosis, and RLFPs. *Theor. Appl. Genet. 80*(6): 833-840.

Sneedon, J.L., 1970. Identification of garden pea varieties. (I) Grouping, arrangement, and use of continous characters. *J. Natn. Inst. Agric. Bot., 12*: 1-6

Souza, E. and M.E. Sorrels, 1991a. Relationships among 70 North American oat germplasms: I. Cluster analysis using quantitative characters. *Crop Sci. 31*: 599-605.

Souza, E. and M.E. Sorrels, 1991b. Relationships among 70 North American oat germplasms: II. Cluster analysis using qualitative characters. *Crop Sci. 31*: 605-612.

Steel, R.G.D., and J.H. Torrie, 1960. *Principles and procedures of Statistics*. McGraw-Hill Company, Inc.; 481pp.

Stewart, J., McD., 1994. Potential for crop improvement with exotic germplasm and genetic engineering. *In* Challenging the future. Proc. World Cotton Research Conf. Brisbane, pp. 313-327.

Vanderborght, T., and E. Depiereux, 1987. A procedure of data analysis for the study of infraspecific diversity of plants with reference to *Phaseolus vulgaris* L. *Biom. Praxim. 27:* 113-123.

Vavilov, N.I., 1926. Studies on the origins of cultivated plants. *Bull. Appl. Bot. Plant Breed. 16*: 1-245.

Vienne (de), D. and C. Damerval, 1985. Distances calculées à partir de marqueurs moléculaires. *In*: Lefort-Buson, M. et D. de Vienne (Eds). *Les distances génétiques*: *Estimations et applications*. INRA, Paris, 39- 57

Wendel, J.F. and V.A. Albert, 1992. Phylogentics of the cotton genus (*Gossypium*): Character - State Weighted Parsimony Analysis of chloroplast. DNA Restriction Site Data and its Systematic and Biogeographic Implications. *Syst. Bot., 17* (1):115-143.

Ziauddin, A., R.P. Katiyar, and Radley Shyam, 1980.Genetic divergence in Triticales. *Indian J. Genet. Pl. Breed. 40:* 35-38.

APPENDIX

Appendix 1:Upland cotton genotypes studied, abbreviations used in the figures, and sources country.

S/No	GENOTYPES	Abbreviation	Country
1	(539*1298)*gl 14 1069	(539*1298)*g	IRA
2	1 10 B	110B	USA
3	101 102 B	101 102B	USA
4	108 F okra	108F okra	USR
5	153 F	153 F	USR
6	26 J A 125	26 J A 125	NIR
7	410 49(A333 57*HR 21	-(A 333 57*HR)	BDI
8	73460 8 (SINDOS 80)	-8 (SINDOS 80)	GRE
9	A 637	A 637	UGA
10	A SMO	A SMO	UGA
11	A66 36	A66 36	UGA
12	ACALA 1517 BR1	ACA 1517 BR1	USA
13	ACALA 44	ACALA 44	USA
14	ACALA 531 71046	ACA 531 7104	GRE
15	ACALA ANCIEN	ACA ANCIEN	USA
16	ACALA DEL CERRO	ACA DL CRO	USA
17	ADD1	ADD1	ELS
18	ALA 73 2M	ALA 73-M	ZAM
19	ALBAR 72B	ALBAR 72B	ZIM
20	ALEP 3	ALEP 3	SYR
21	ALEP 4	ALEP 4	SYR
22	ALEPPO 1	ALEPPO 1	SYR
23	ALLEN 333 57	ALN 333 57	CMR
24	ALLEN ANCIEN	ALN ANCIEN	CHD
25	ANDALUCIA	ANDALUCIA	SPA
26	AUBURN 56	AUBURN 56	USA
27	B 761	B 761	CAF
28	BJA 592 okra	JA 592 okra	CHD
29	BJA B2	BJA B2	CAF
30	BJA gl	BJA gl	CHD
31	BLANK 1 75	BLANK 1 75	USA
32	BOU 1724 6	BOU 1724 6	TOG
33	BOU 74	BOU 74	TOG
34	BOU 75	BOU 75	TOG
35	BOU 76	BOU 76	TOG
36	BOU 77	BOU 77	TOG
37	BOU 78	BOU 78	TOG
38	BOUAKE frego	BUAKE frego	CIV
39	C 1211	C 1211	USR
40	C 3137	C 3137	USR
41	C 433 81	C 433 81	FRA
42	C3210	C3210	USR
43	C4727	C4727	USR
44	CAMD H 75	CAMD H 75	USA
45	CAMD H 75 C	CAMD H 75 C	USA
46	CB 2472 3	CB 2472 3	BGD
47	CHILALA 84	CHILALA 84	ZAM
48	CHILALA 85	CHILALA 85	ZAM
49	CHINE 204	CHN 204	CPR
50	CHINE 401	CHN 401	CPR
51	CHINE 7 0/1	CHN 7 0/1	CPR
52	CHIRPAN 173	CHIRPA 173	BUL
53	CHIRPAN 996	CHIRPA 996	BUL
54	CHRIPAN 433	CHRIPA 433	BUL
55	CHUAN JIAN 3 sup	CHN J3 sprok	CPR
56	CHUREZA 85	CHUREZA 85	ZAM
57	CIRRITO	CERRITO	ECU
58	CNPA 2 H	CNPA 2 H	BRA
59	CNPA 6H	CNPA 6H	BRA
60	CNPA PRECOCE 1	CNPA PREC 1	BRA
61	CO 417	CO 417	IND
62	COKER 100/2	COKER 100/2	USA
63	COKER 100A gl	CKER 100A gl	USA
64	COKER 201	COKER 201	SPA

65	. COKER 312	COKER 312	USA
66	. COKER 315	COKER 315	USA
67	. COKER 413 E	COKER 413 E	USA
68	. CONAL SN	CONAL SN	NIC
69	. COPAL 117	COPAL 117	ELS
70	. COPAL 68	COPAL 68	ESL
71	. COPAL 70	COPAL 70	ESL
72	. CUKUROVA 1518	CKURVA 1518	TUR
73	. D 202 14	D 202 14	CIV
74	. DELTAPINE L61	DLTPNE L61	USA
75	. DELTAPINE SL GL	DLTPNE slGL	USA
76	. DELTAPINE SR1	DLTPNE SR1	USA
77	. DELTAPINE SR2	DLTPNE SR2	USA
78	. DES 119	DES 119	USA
79	. DETAPINE SL FREG	DTPNE slfrg.	USA
80	. *DIXIE KING III*	*DIXIE KING II*	*USA*
81	. DP*P4	DP*P4	USA
82	. DPL 15/21	DPL 15/21	TUR
83	. DUNN 119	DUNN 119	USA
84	. DZA 74 M	DZA 74 M	ZAM
85	. EMPIRE B4	EMPIRE B4	BKF
86	. EMPIRE WR (B2 B6	EM.WR (B26)	USA
87	. EMPIRE WR(B2 B3	EM.WR(B236)	USA
88	. EXD 71 438	EXD-71 438	GRE
89	. EZA 7A MM	EZA-7A-M	ZAM
90	. F 281 G 694 gl	F281-G694gl	CHD
91	. F 281 G 696 gl	F281-G696gl	CHD
92	. G 319 16	G319-16	CIV
93	. GAMBIA Ra 8 2	GMBIA Ra-8-2	PER
94	. GREGG	GREGG	USA
95	. GUAZUNCHO	GUAZUNCHO	ARG
96	. GUMBO	GUMBO okra	USA
97	. H 10	H 10	ISR
98	. H23	H23	ISR
99	. HAMPEI	HAMPEI	IND
100	. HAR 444 2 66	HAR-444-2-66	CIV
101	. HAR 444 2 65	HAR-444-2-65	CIV
102	. HAR 444 2 69	HAR-444-2-69	CIV
103	. HAR 444 2 70	HAR-444-2-70	CIV
104	. HAR*COKER 4 75	HAR*COK 4-75	MLI
105	. HC B4 75	HC-B4-75	BKF
106	. HG 9	HG 9	CHD
107	. HIGH MARTURITY	HI.MARTURITY	UGA
108	. HR1 71(HR1 F71)	HR.(HR1 F71)	CHD
109	. IAC 17	IAC 17	BRA
110	. IAC 20	IAC 20	BRA
111	. IL 66	IL 66	URT
112	. INIAP 601	INIAP 601	ECU
113	. IPEACO SL 23 6 680	IPsl 23-6-68	BRA
114	. IPEACO SL 2361	IPCOsl 2361	BRA
115	. IPEACO SL 7 1	IPCOsl 7-1	BRA
116	. IPEACO SL 7 SM6	IPCO sm6	BRA
117	. IRCT CR10	IRCT CR10	COS
118	. IRCT CR80	IRCT CR80	COS
119	. IRMA (96+97)	IRMA (96+97)	CMR
120	. IRMA 1042	IRMA A1042	CMR
121	. IRMA 1110 okra	IRMA 1110 ok	CMR
122	. IRMA 1145 gl	IRMA 1145gl	CMR
123	. IRMA 1243	IRMA 1243	CMR
124	. IRMA 1327	IRMA 1327	CMR
125	. IRMA 1480	IRMA 1480	CMR
126	. IRMA 1657	IRMA 1657	CMR
127	. IRMA 197 gl	IRMA 197gl	CMR
128	. IRMA 323	IRMA 323	CMR
129	. IRMA 772	IRMA 772	CMR
130	. IRMA A 215 GL	IRMA A215gl	CMR
131	. IRMA A1200	IRMA A1200	CMR
132	. IRMA A151 gl	IRMA A151gl	CMR
133	. IRMA B209 okra	IRMA B209ok	CMR

134 . IRMA BLT	IRMA BLT	CMR
135 . IRMA BLT PF	IRMA BLT PF	CMR
136 . IRMA BOKA	IRMA BOKA	CMR
137 . IRMA W 855	IRMA W855	CMR
138 . IRMA Z101 gl	IRMA Z101	CMR
139 . IRMA Z122 gl	IRMA Z122gl	CMR
140 . IRMA Z129 gl	IRMA Z129gl	CMR
141 . IRMA Z332 okra	IRMA Z332ok	CMR
142 . J 119	J 119	CIV
143 . J 205	J 205	IND
144 . JEREZ	JEREZ	SPA
145 . JIANSU N1	JIANSU N1	CPR
146 . KAKENE	KAKENE	BGD
147 . KRISHNA	KRISHNA	IND
148 . L142 9	L142-9	CIV
149 . L231 24 70	L231-24-70	CIV
150 . LEBO 1 75	LEBO 1-75	USA
151 . LH 62	LH 62	PAK
152 . M11 gl	M11 gl	USA
153 . MAKOKA 72	MAKOKA 72	MWI
154 . MAKOKA 78	MAKOKA 78	MWI
155 . MAR 17 30 76	MAR 17-30-76	USA
156 . MATACO (SP 347)	MTCO (SP347)	ARG
157 . MEBANE	MEBANE	USA
158 . MEX 68	MEX 68	MEX
159 . N'KOURALA 14 E 4	N'KRL14-E-4-3	MLI
160 . N'KOURALA 44 10	N'KRL 44-10	CHD
161 . N'KOURALA K1 C2	N'KRL K1-C2	MLI
162 . NAZILLI 87	NAZILLI 87	TUR
163 . NTA 88 7	NTA 88-7	MLI
164 . NTA 88-11gl	NTA 88-11gl	MLI
165 . ORS LP 1 Okra	ORS LPok	USA
166 . P14 T128	P14-T128	CHD
167 . P279	P279	PAR
168 . PAN 575	PAN 575	CHD
169 . PAN F3 52	PAN F3-52	MLI
170 . PAN F575	PAN F3-75	CHD
171 . PAVLIKENI 73	PAVLIKENI 73	BUL
172 . PEE DEE 0111	PEEDEE 0111	USA

173 .	PEXA	COS
174 .	PRONTO	USA
175 .	PS 10	IND
176 .	PS 9	IND
177 .	QUEBRACHO	ARG
178 .	REBA B 296	CAF
179 .	REBA B.50	CAF
180 .	REBA BTK12	CAF
181 .	REBA W296	CAF
182 .	S 188 okra	NIC
183 .	S 4727	BGD
184 .	S1622 JENDOUBA	USR
185 .	SA 66 51	UGA
186 .	SAHEL 50	IRA
187 .	SAMARU 26 C 54	CHD
188 .	SAMARU 26C	NIR
189 .	SAMARU 59 11	NIR
190 .	SAMARU 68	NIR
191 .	SAMARU 68 PAO NG	NIR
192 .	SAMARU 69	NIR
193 .	SAMARU 70	NIR
194 .	SAMARU 71	NIR
195 .	SAMARU 72	NIR
196 .	SAMCOT 10	NIR
197 .	SAMCOT 6	NIR
198 .	SAMCOT 8	NIR
199 .	SAMCOT 9	MAG
200 .	SAMIR 730	ZAI
201 .	SCG 465...57	ZAI
202 .	SCG 465...51	ZAI
203 .	SCG 471...57	ZAI
204 .	SCG 1572..528	BGD
205 .	SENEZPENA TABA	AUL
206 .	SI OKRA 1 4	AUL
207 .	SI OKRA L 22	AUL
208 .	SICALA 3 2	GRE
209 .	SKOTOUSSA	NIC
210 .	SMB 86	ARG
211 .	SP 510	ARG
212 .	SP 59	ARG
213 .	SP 726	ARG
214 .	SP 81110	CHD
215 .	SP 86	TOG
216 .	SR1 F4	TOG
217 .	STAM 82	USA
218 .	STAM F	USA
219 .	STONEVILLE 2B 59	USA
220 .	STONEVILLE 506	
221 .	STONEVILLE 731 ne	

222 .	STONEVILLE 7A	STONE 7A	USA
223 .	STONEVILLE 7A OK	STONE 7A OKR	USA
224 .	STONEVILLE 7A gl	STONE 7A gl	USA
225 .	STONEVILLE OKRA LEAF	STONE OKRA L	USA
226 .	STONEVILLE frego	STONE frego	USA
227 .	STRUMICA 5086	STRMCA 5086	YUG
228 .	TABLADILLA 100	TBLDILLA 100	SPA
229 .	TABLADILLA 16	TBLDILLA 16	SPA
230 .	TAKJA	TAKJA	THA
231 .	TAMCOT SP 21	TMCOT SP21	USA
232 .	TH 1100	TH 1100	PAK
233 .	U 280	U 280	NIC
234 .	U 563 19	U 563-19	CIV
235 .	V 145	V 145	NIC
236 .	U 563	V 563	CMR
237 .	V11	V11	CHD
238 .	V294 gl	V294	CHD
239 .	X 239	X 239	CHD
240 .	X 206	X206	CHD
241 .	X 260	X 260	CHD
242 .	XAB 5*TANGUIS 126 5	XB5*TNGS 126	BGD
243 .	XING TAI 68 71	XING TAI 68	CPR
244 .	Y 1422 SM 71	Y 1422SM71	CHD
245 .	Y 1422 okra	Y 1422 okra	CHD
246 .	Y 1638 624 GL	Y 1638 624GL	CHD
247 .	Z 407 1156	Z 407 1156	BDI

Appendix 2: List of source countries, codes (FAO), and number of genotypes of upland cotton evaluated

S/No	Country	Code	Genotype No
1	Argentina	ARG	8
2	Australia	AUL	3
3	Bangladesh	BGD	5
4	Brazil	BRA	9
5	Bulgaria	BUL	4
6	Burundi	BDI	2
7	Burkina Faso	BKF	3
8	Cameroon	CMR	25
9	Central African Republic	CAF	6
10	Chad	CHD	21
11	China People's Republic	CPR	6
12	Costa Rica	COS	3
13	Ecuador	ECU	2
14	El Salvador	ELS	4
15	France	FRA	1
16	Greece	GRE	4
17	India	IND	7
18	Iran	IRA	2
19	Israel	ISR	2
20	Ivory Coast	IVC	11
21	Madagascar	MAG	1
22	Malawi	MLW	2
23	Mali	MLI	6
24	Mexico	MEX	1
25	Nicaragua	NIC	5
26	Nigeria	NIR	13
27	Pakistan	PAK	2
28	Paraguay	PAR	1
29	South Africa	ZAF	1
30	Spain	SPA	5
31	Syria	SYR	3
32	Tanzania	URT	1
33	Thailand	THA	1
34	Togo	TOG	7
35	Turkey	TUR	4
36	Uganda	UGA	5
37	USA	USA	42
38	Union SSR	USR	7
39	Yugoslavia	YUG	1
40	Zaire	ZAI	4
41	Zambia	ZAM	6
42	Zimbabwe	ZIM	1

Appendix 3: Analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown at Samaru in 1994.

	Mean squares of the traits measured												
Source of variation	DF	'H(cm)	INPP	FB	VB	HFFB(cm)	INFFB	LFFB	log[BOPP]	BOLT(cm)	BODIA(mm)	DFBO	TPb	D50FO
Replication	3	9304.87**	169.09**	329.34**	47.08**	65.17**	58.29**	1983.79**	7.56**	2.46**	0.23**	121.69**	15.65**	42.10*
Block(Replication)	60	527.61**	5.51**	6.56**	0.92**	10.72*	0.90**	88.15**	0.20	1.73**	0.11**	23.82***	3.86**	21.68**
Genotype	246	721.84**	6.15**	5.10**	1.01**	26.67**	0.81**	112.67**	0.24**	0.18**	0.06**	34.90**	2.21**	21.20**
Error	≤714 ++	143.76	2.03	1.81	0.36	7.88	0.50	40.80	0.16	0.12	0.03	10.83	1.07	14.01

* and ** indicate significance at the 0.05 and 0.01 levels of pobability, respectively

Appendix 3 (cont.)

Source of variation	DF	[FLPP]+	SCPP(g)	YLPP(g)	GIN%	S.I.(g)	2.5% S.L.	50% S.L.	FBST	MIC	BBVB%
Replication	3	10.39**	3520.43**	4444.09**	16.88**	8.18**	61.56**	38.01**	342.52**	2.81**	177.40**
Block(Replication)	60	0.50**	424.11**	379.42*	7.88**	0.69**	4.45**	2.79**	23.01**	0.23*	70.63**
Genotype	246	0.88**	522.54**	356.41**	37.11**	2.63**	8.51**	1.94**	9.57**	0.54**	112.98**
Error	≤714 ++	0.33	187.92	228.69	3.08	0.39	1.1	0.39	5.56	0.16	46.21

* and ** indicate significance at the 0.05 and 0.01 levels of probability, respectively

\+ Trait abbreviations as defined in Table 1.

++ Error degree of freedom varied from one trait to another because of omitted data for these traits prior to the analysis of variance

Appendix 4: Analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown at Maroua in 1994.

	Mean squares of the traits measured											
Source of variation	DF	TH(cm)	FB	VB	HFFB(cm)	D50BO	INFFB	LFFB(cm)	Log|BOPP|	BOLT(mm)	BODIA(mm)	DFBOP	DFFO
Replication	3	9240.94**	28.26**	2.38**	71.27**	478.43**	27.17**	1512.79**	3.11**	7.03**	4.53**	183.01**	7.86
Block(Replication)	60	3289.14**	21.01**	0.95**	42.11**	24.79**	2.98**	333.57**	0.52**	1.40**	0.45**	27.68**	16.46**
Genotype	246	605.53**	6.66**	0.88**	38.26**	63.19**	0.94**	92.54**	0.17**	0.38**	0.11**	50.87**	29.42**
Error	714	235.51	3.28	0.40	12.44	7.68	0.64	53.69	0.13	0.29	0.09	7.81	5.48

Appendix 4 (cont.)

Source of variation	DF	TPh	D50FO	|FLPP|+	YLPP(g)	GIN%	S.I.(g)	2.5% S.L.	50% S.L.	FBST(g/tex)	MIC
Replication	3	99.09**	107.42**	6.42**	5149.44**	45.46**	53.69**	16.30**	37.96**	126.24**	7.67**
Block(Replication)	60	4.65**	22.80**	3.55**	2969.92**	29.29**	3.44**	4.33**	2.86**	17.27**	2.65**
Genotype	246	1.61*	52.30	0.80**	1007.28**	45.82**	2.29**	10.28**	2.60**	8.65**	0.38**
Error	714	1.28	9.06	0.55	504.17	4.04	0.55	1.28	0.44	4.67	0.20

* and ** indicate significance at the 0.05 and 0.01 levels of pobability, respectively
+ Trait abbreviations as defined in Table 1.

Appendix 5: Analysis of variance for agronomic and fibre technology traits from 247 genotypes of upland cotton grown at Garoua in 1994.

		Mean squares of the traits measured											
Source of variation	DF	*H(cm)	HB	VB	HFFB(cm)	D50BO	INTFB	LFFB(cm)	Log[BOPP]	BOLT(mm)	BODIA(mm)	DFFO	TPb
Replication	3	2399.33**	18.24**	14.77**	528.73**	111.12**	19.96**	279.60**	0.35*	2.76**	0.35**	22.26	7.38**
Block(Replication)	60	725.98**	32.59**	2.87**	83.68**	52.18**	3.49**	206.74**	0.27**	0.52**	0.28**	54.69**	1.98**
Genotypes	246	1233.61**	9.71**	2.79**	64.34**	119.03**	1.83**	281.90**	0.21**	0.27**	0.08**	105.35**	1.73**
Error	≤714++	236.10	5.04	0.82	15.29	10.68	1.24	114.40	0.10	0.10	0.04	9.81	0.72

++ Error degree of freedom varied from one trait to another because of omitted data for these traits prior to the analysis of variance

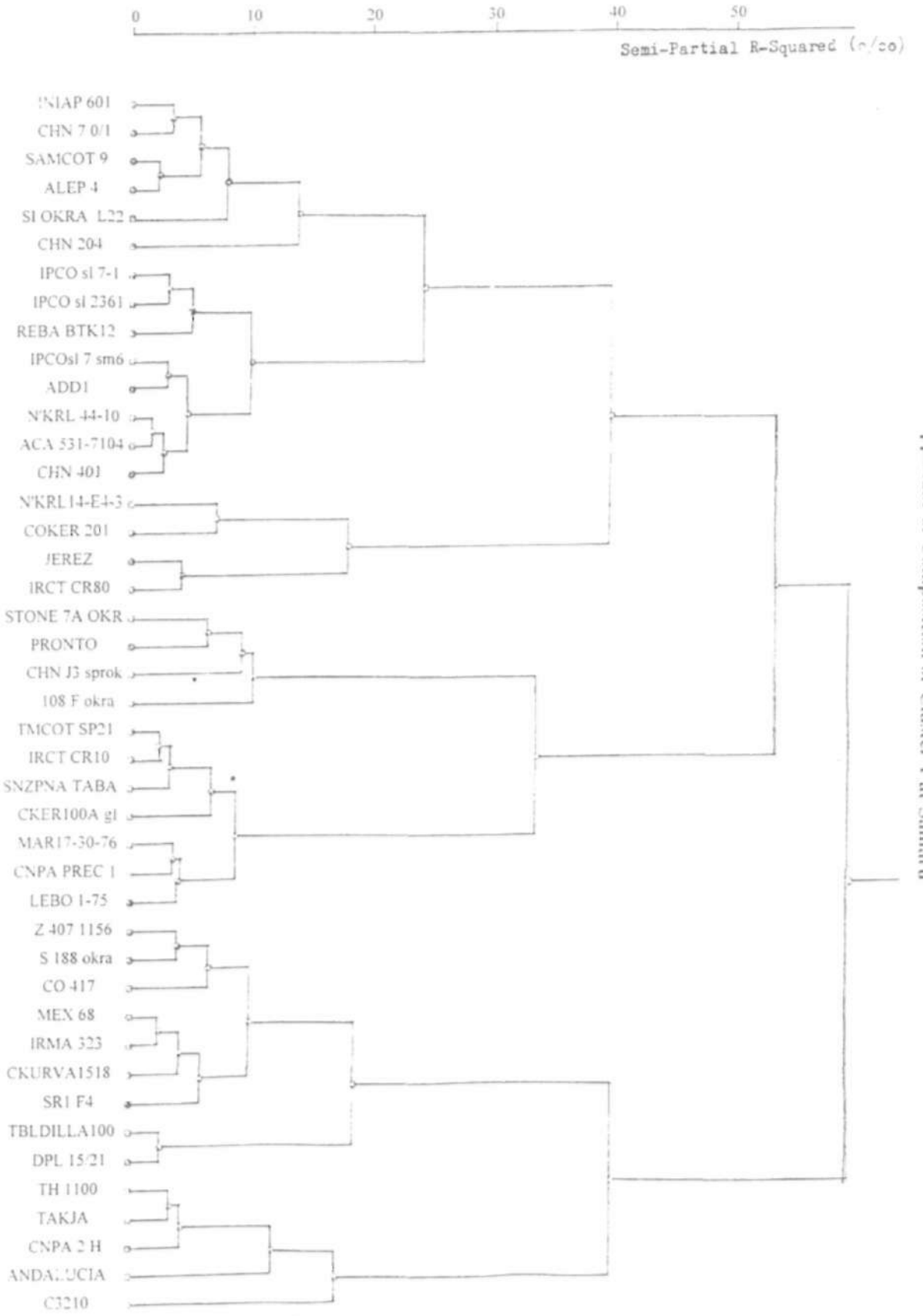
Appendix 5 (cont.)

Source of variation	DF	D30FO	[FLPP]+	YLPP(g)	GIN%	S.I(g)	2.5% S.L.(mm)	50% S.L.(mm)	FBST(g/tex)	MIC	BBVB%
Replication(Rep)	3	78.81**	2.43*	22573.92**	63.80**	3.05**	20.02**	42.90**	696.80**	1.66**	7.66**
Block(Replication)	60	149.31**	1.93**	7154.00**	11.81**	1.80**	5.53**	1.50**	32.85**	0.62*	29.59**
Genotypes	246(245 for MIC)	210.47**	1.49**	7755.02**	36.65**	2.26**	9.73**	2.42**	7.64**	0.49**	90.61**
Error	≤714++	10.28	0.71	3899.41	4.16	0.81	1.27	0.42	5.16	0.19	0.40

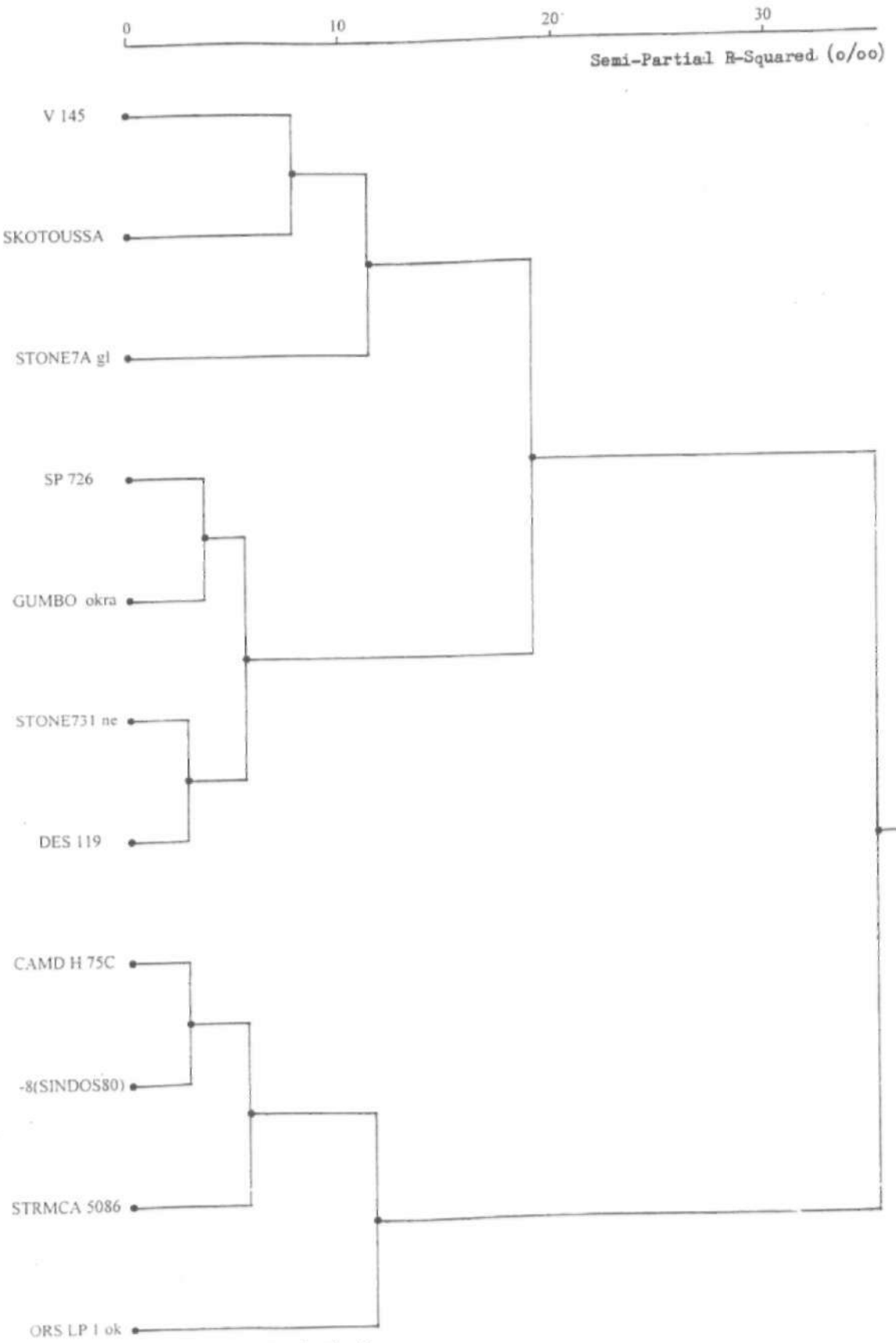
* and ** indicate significance at the 0.05 and 0.01 levels of pobability, respectively

+ Trait abbreviation as defined in Table 1.

++ Error degree of freedom varied from one trait to another because of omitted data for these traits prior to the analysis of variance



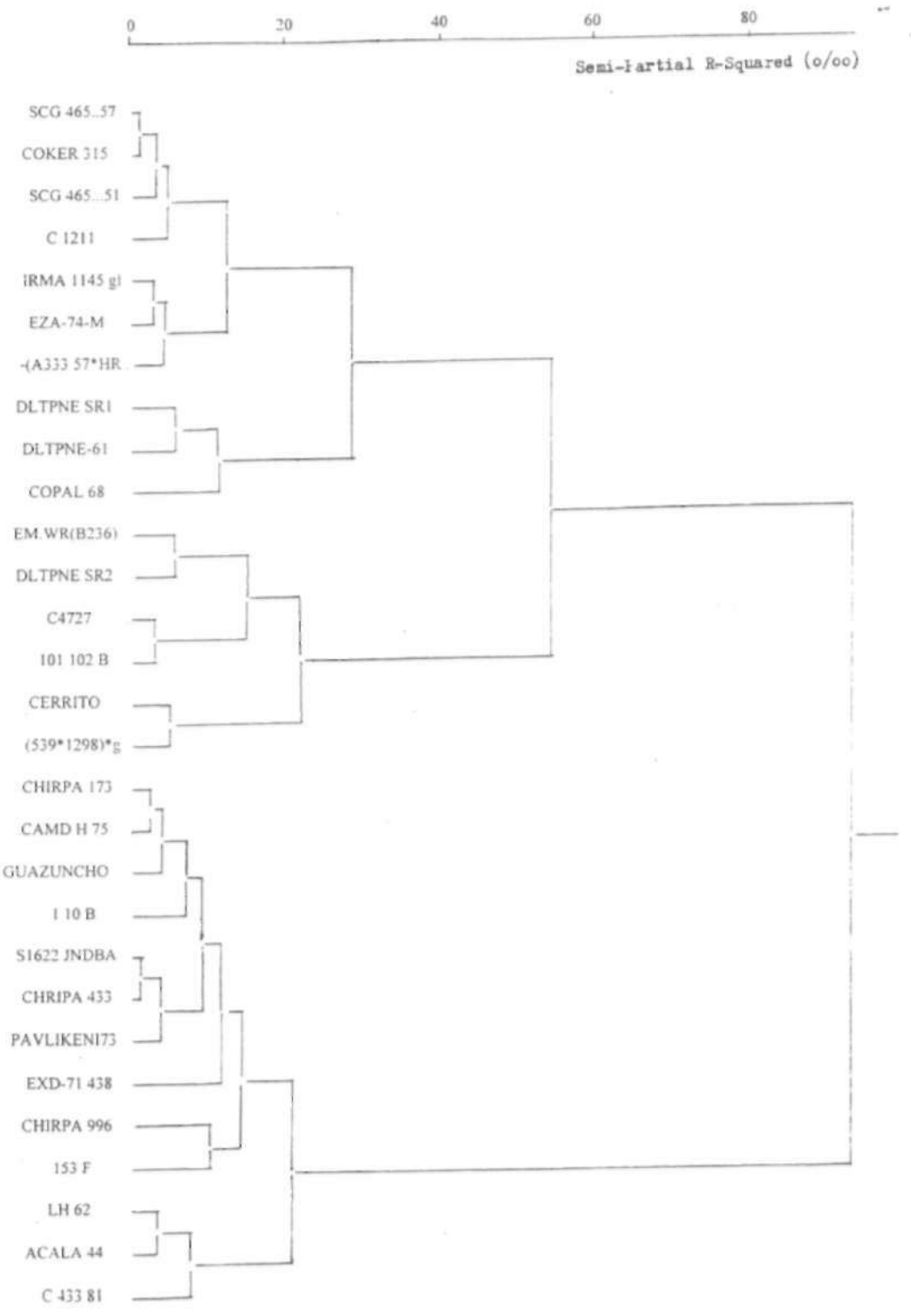
Appendix 6: Composition of Cluster I at Samaru



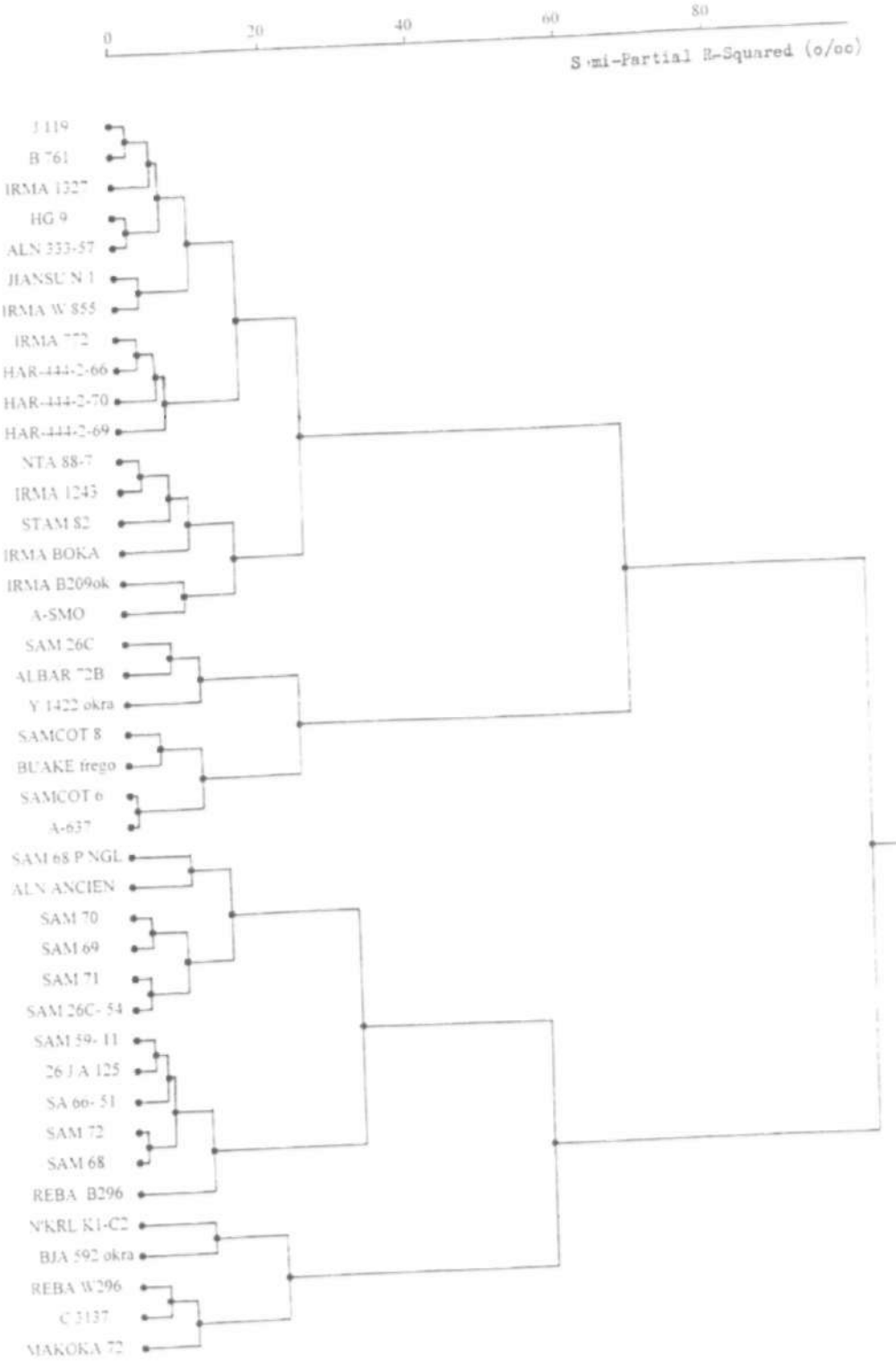
Appendix 7: Composition of Cluster II at Samaru

Note: Genotypes missing in the figure

1. STONE 2859 2. GREGG 3. MEBANE



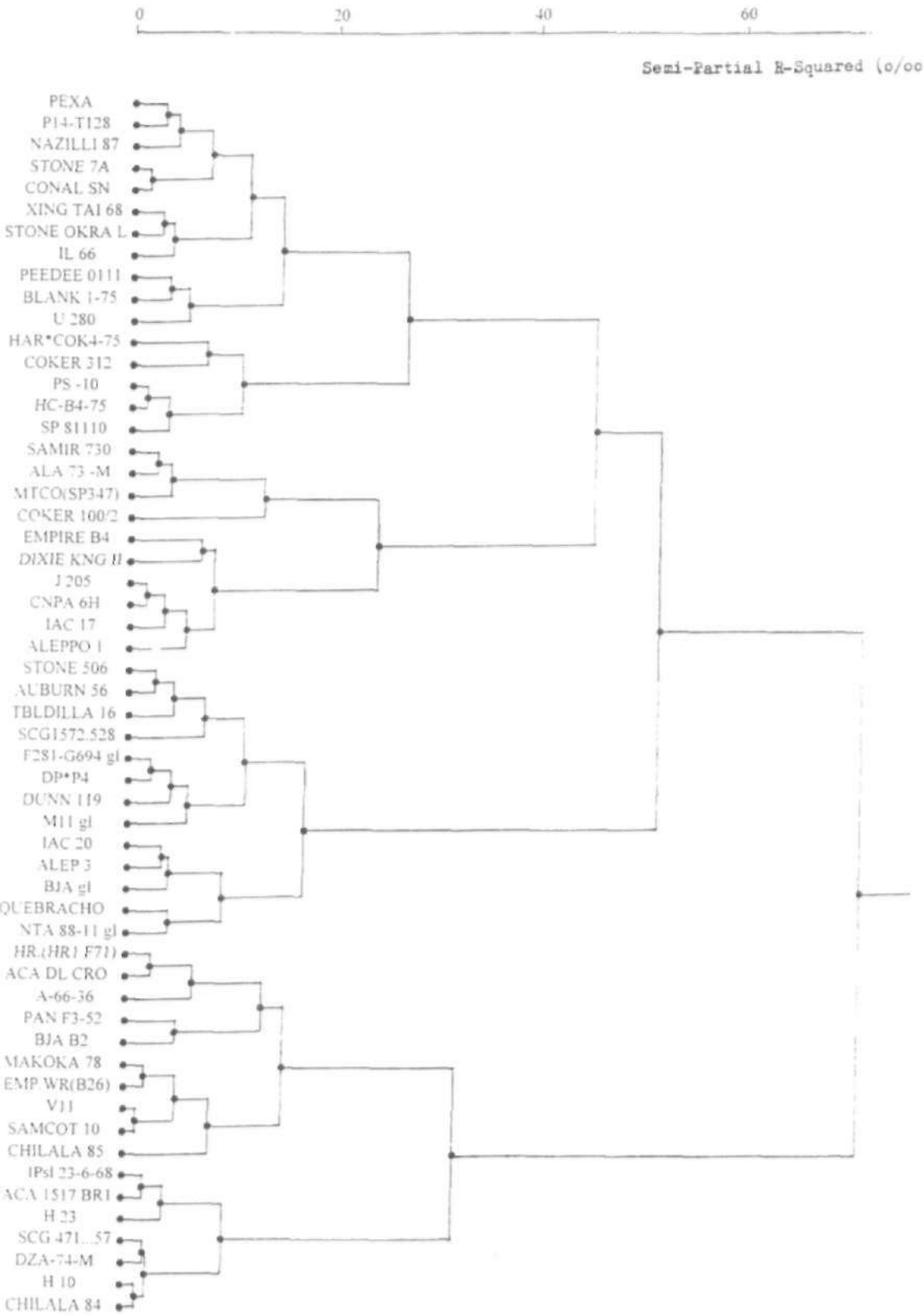
Appendix 8: Composition of Cluster III at Samaru.



Appendix 9: Composition of Cluster IV Samaru.

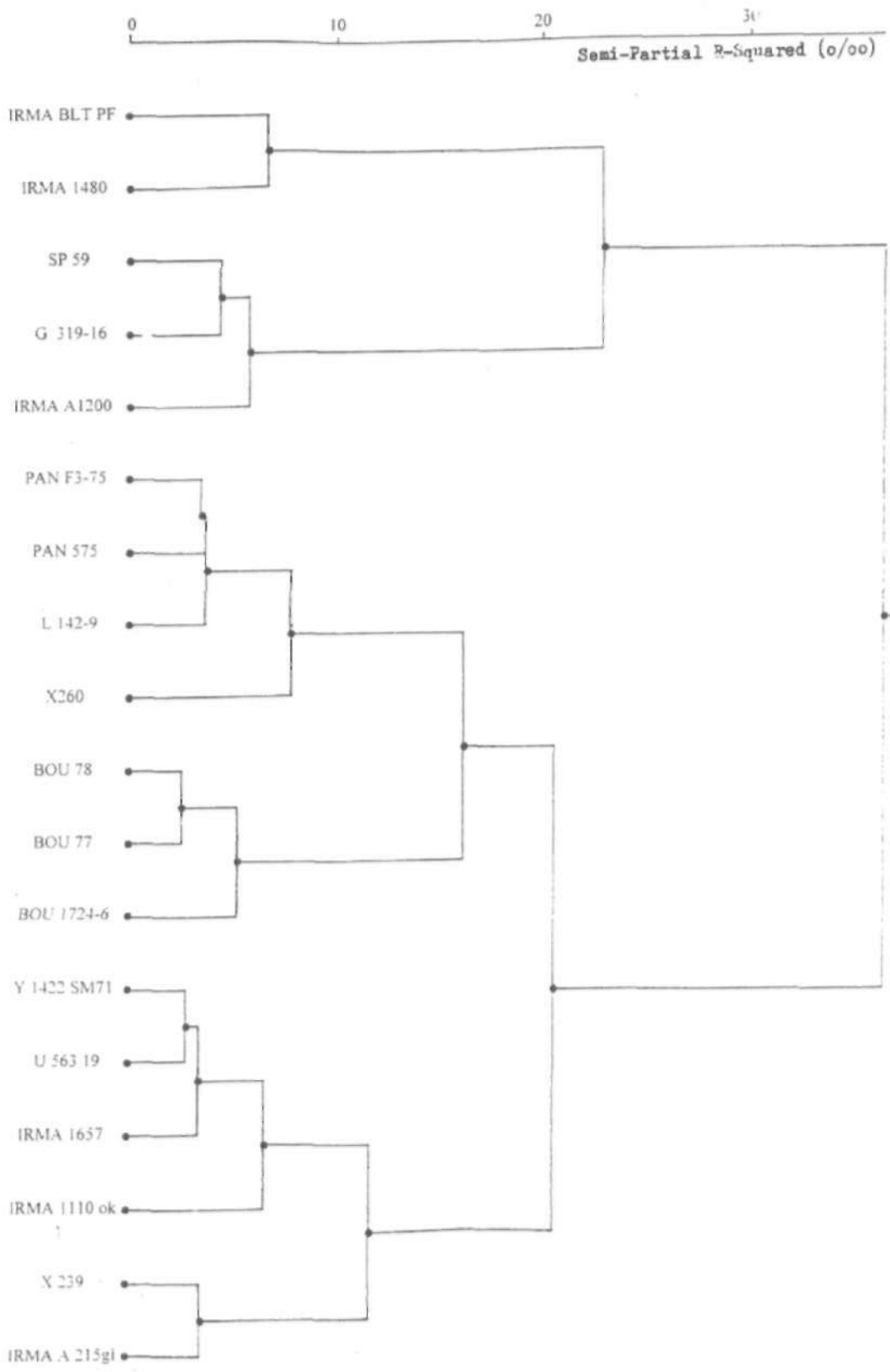


Appendix 10: Composition of Cluster V a Samaru.



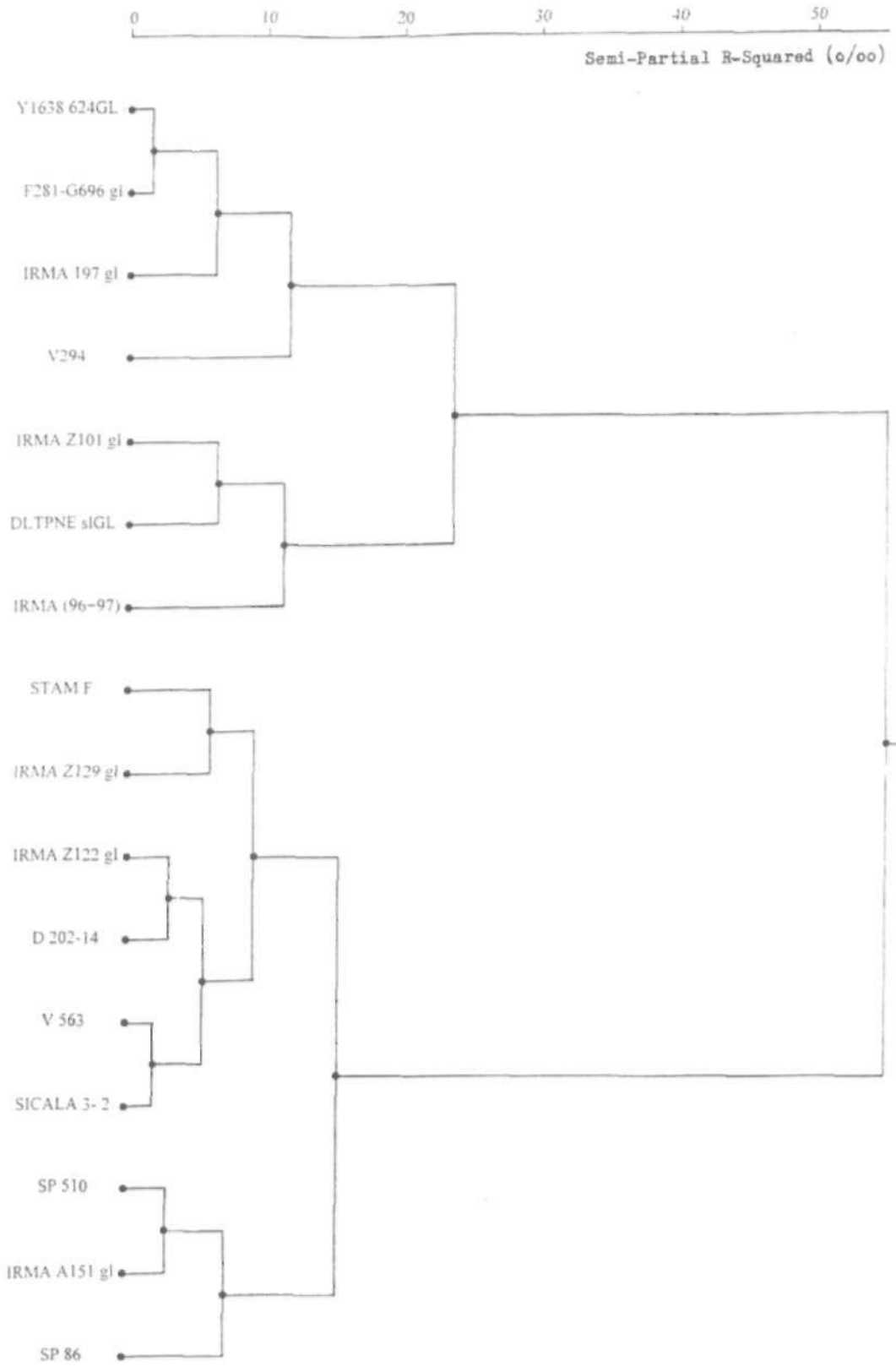
Appendix 11: Composition of Cluster VI Samaru.

Appendix 12: Co. position of Cluster VII at Sam

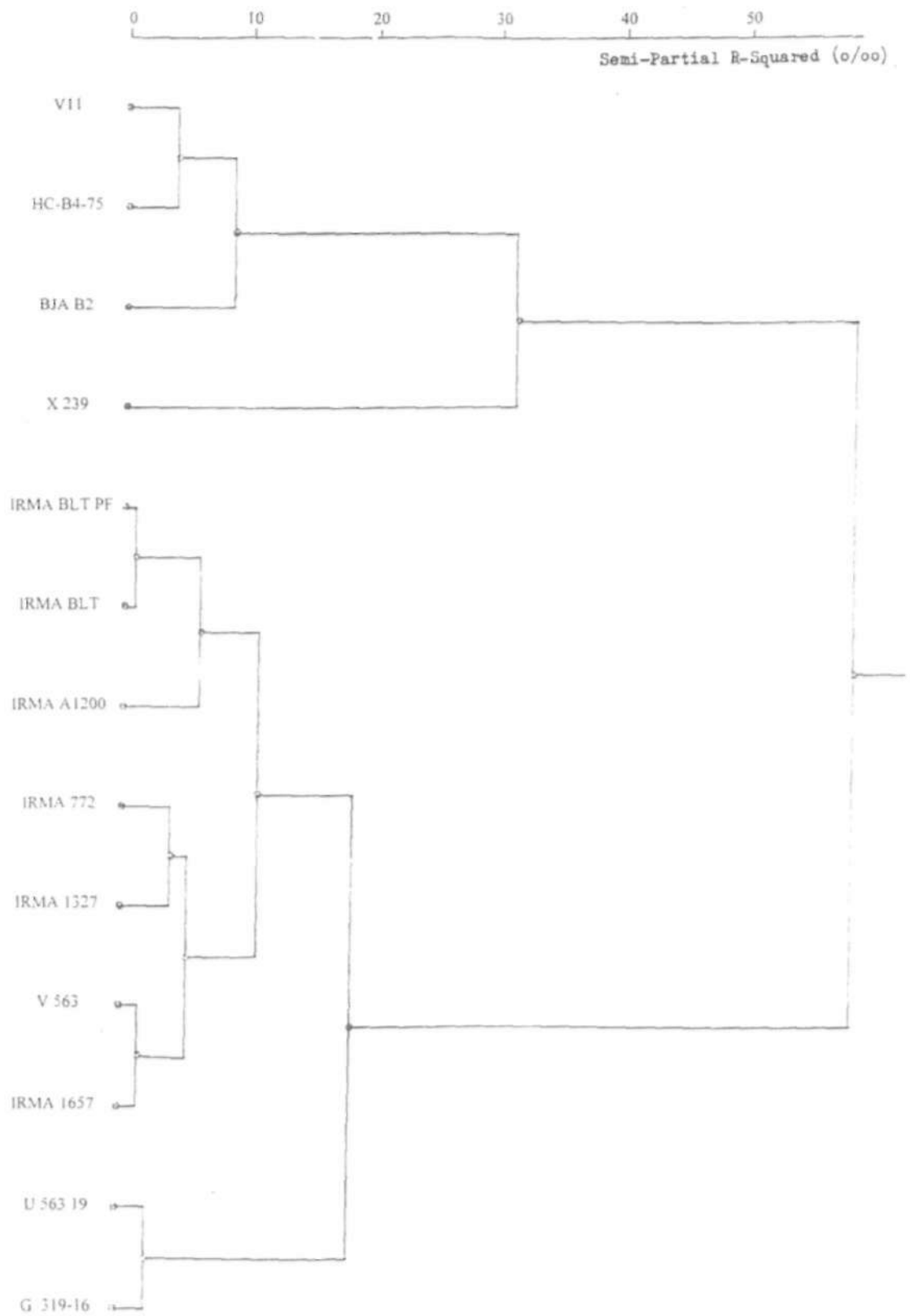


Note: Genotypes missing in the figure

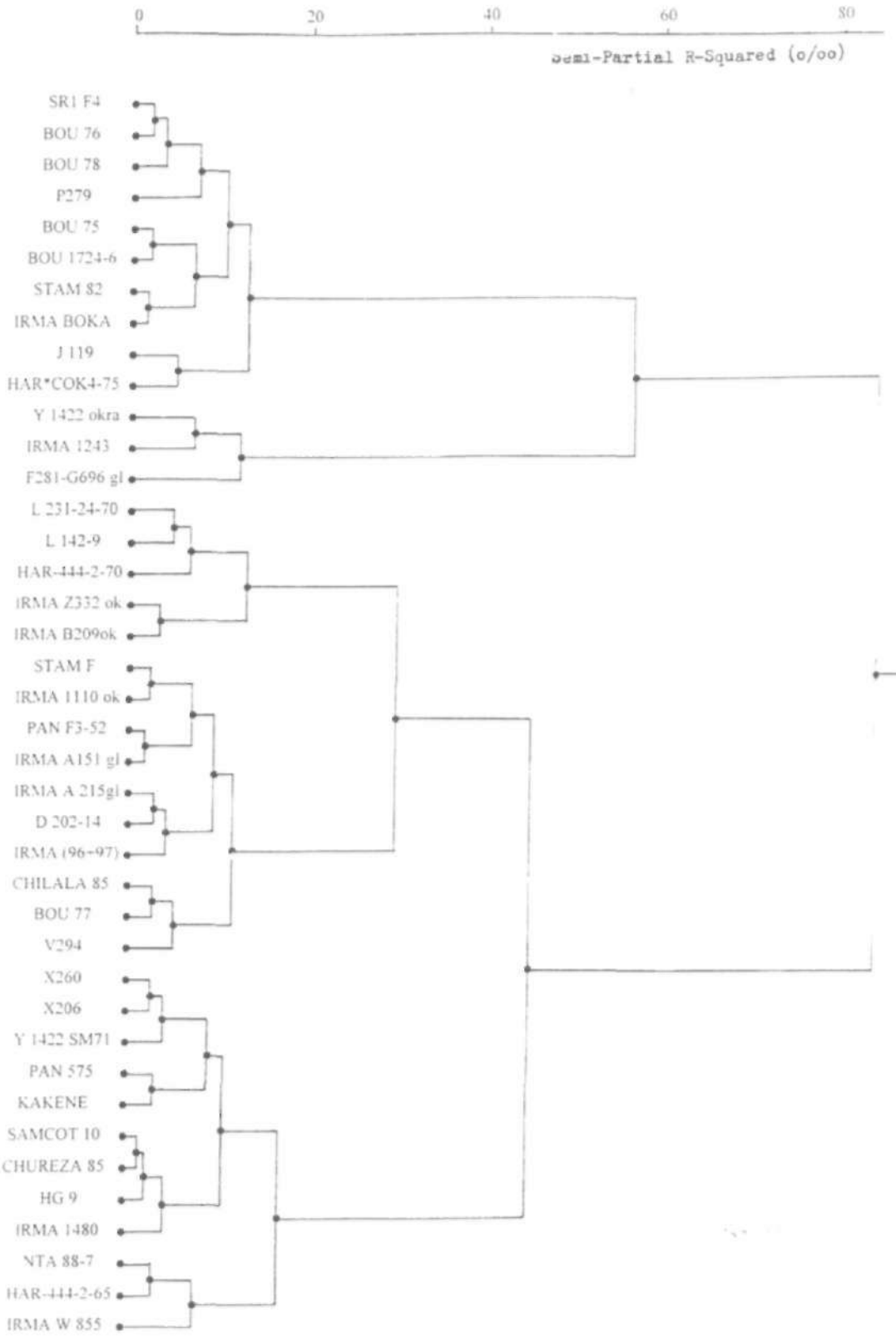
1. IRMA BLT
2. IRMA A 1042
3. IRMA Z 332 ok
4. HAR 444-2-65
5. L 231 24 70
6. BOU 76
7. BOU 75
8. BOU 74
9. CHUREZA 85



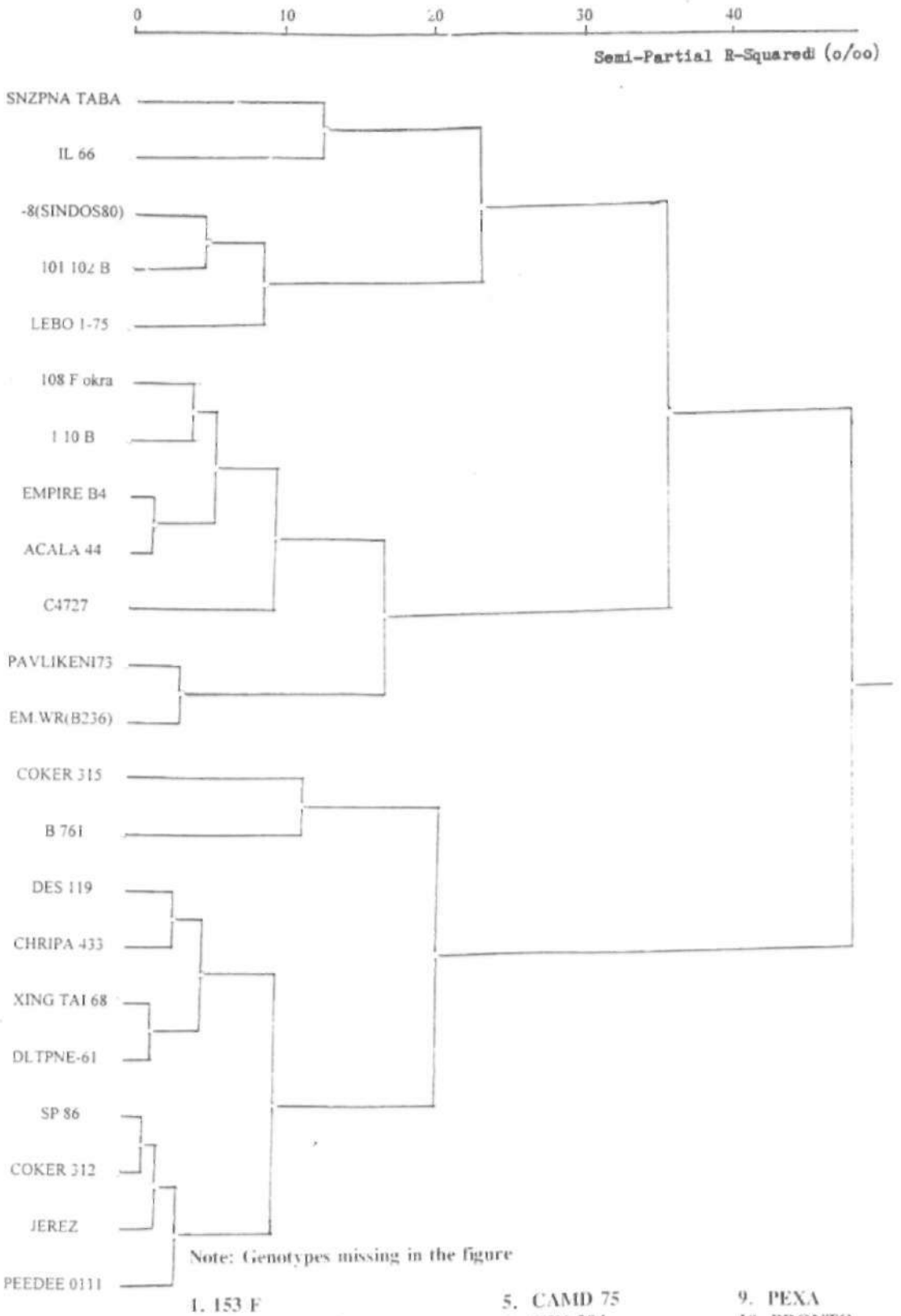
Appendix 13: Composition of Cluster VIII a



Appendix 14: Composition of Cluster at Garoua.



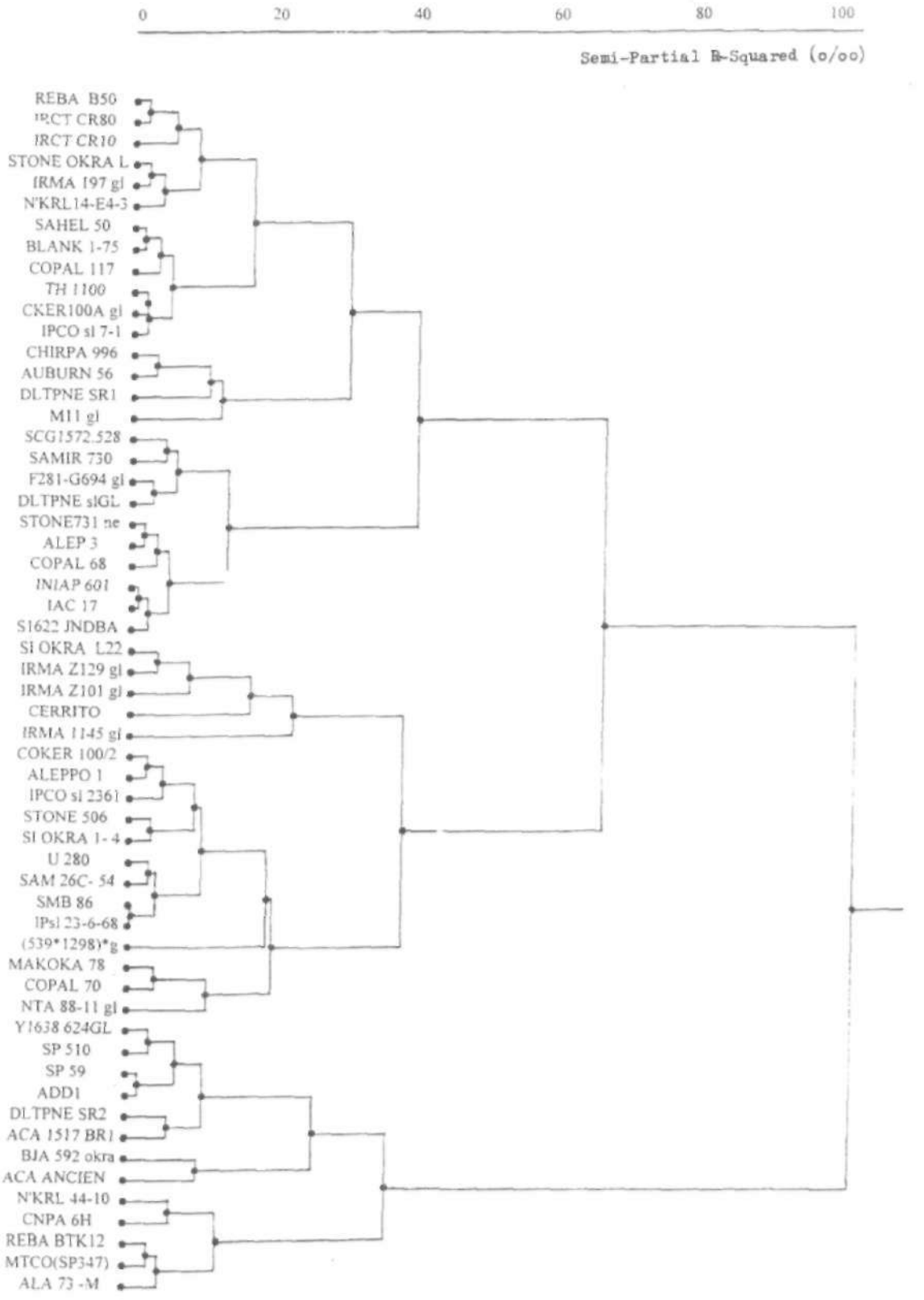
Appendix 15: Composition of Cluster at Garoua.



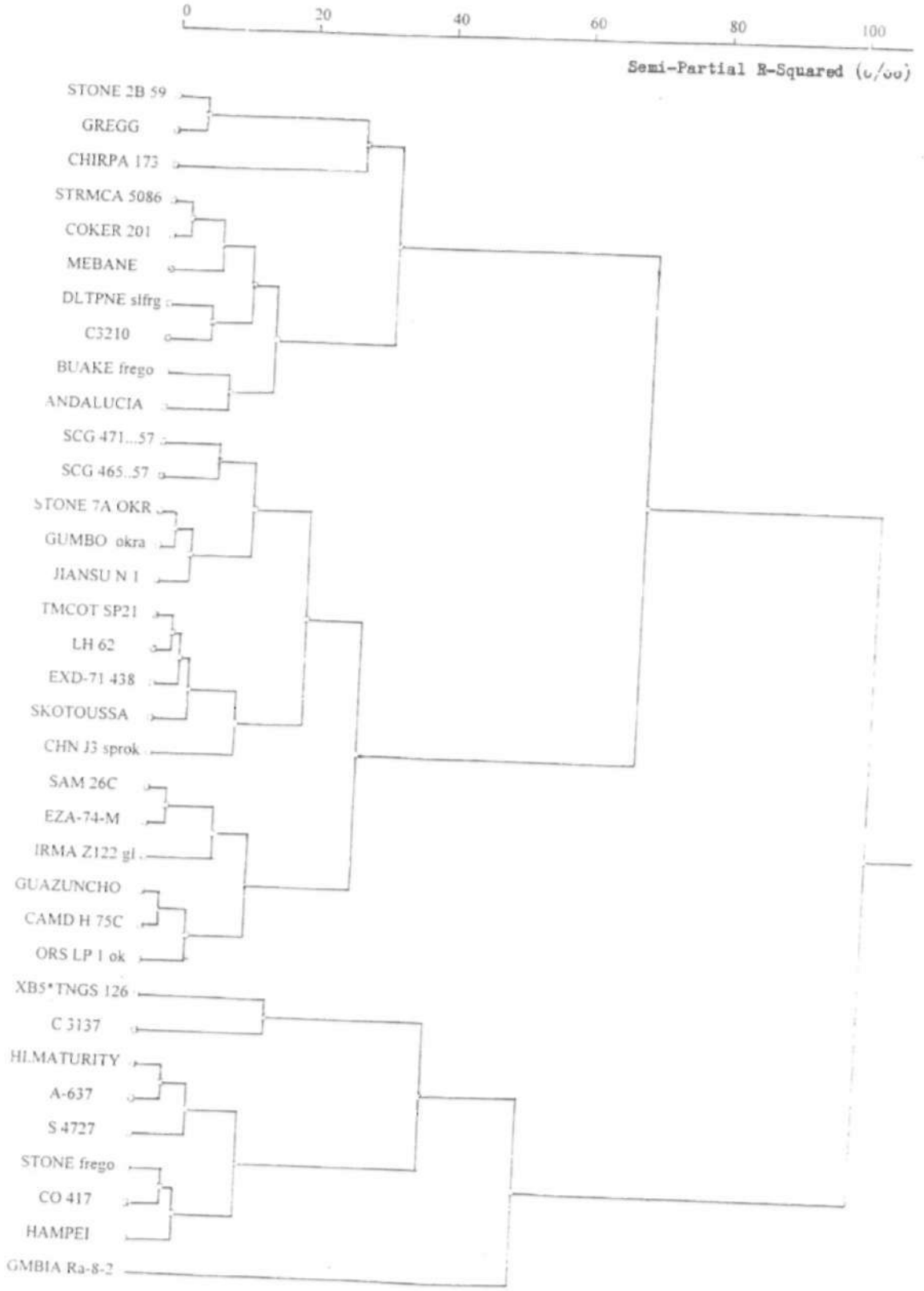
Appendix 16: Composition of Cluster II at Garoua.

Note: Genotypes missing in the figure

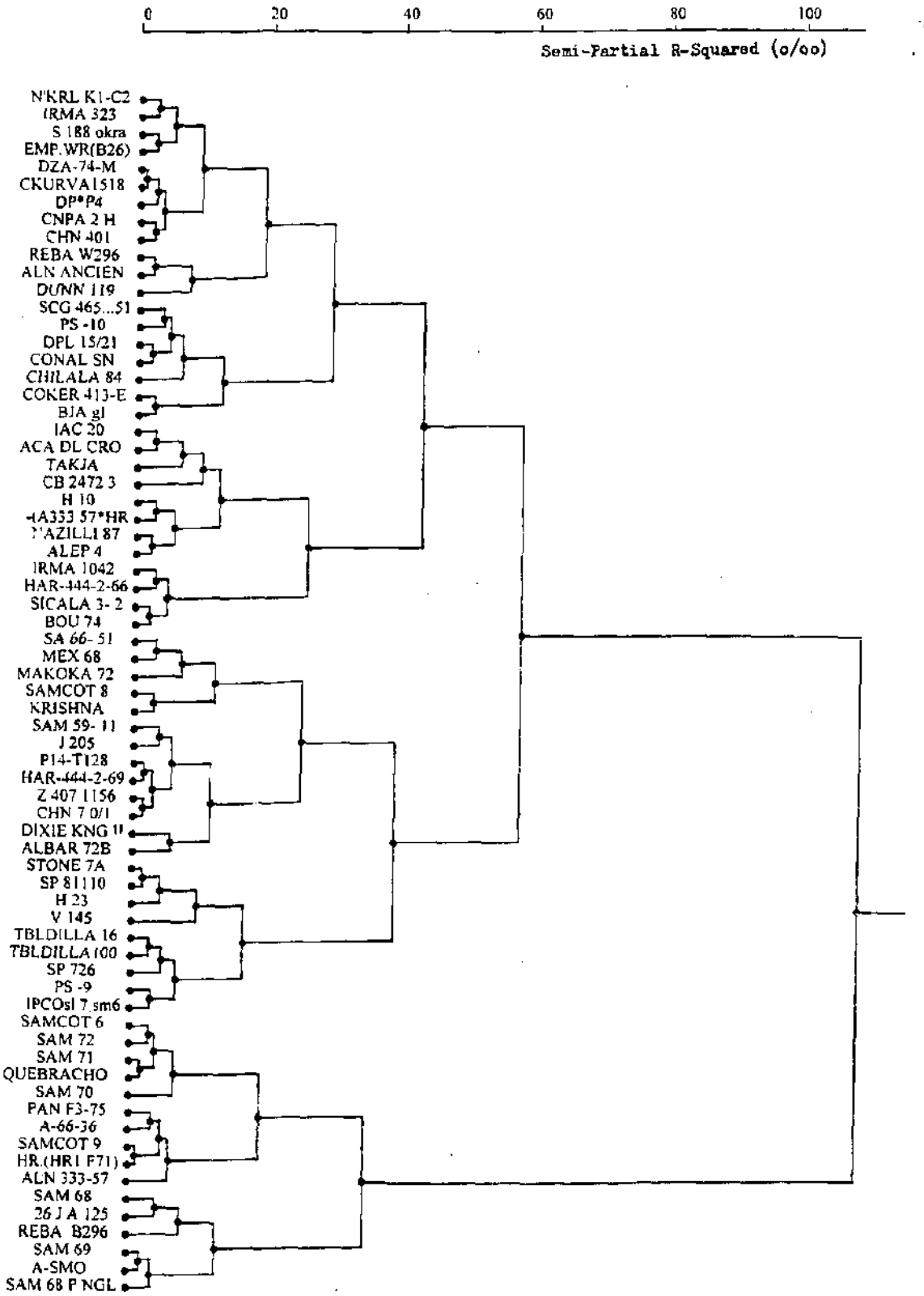
1. 153 F
2. ACA 531 7104
3. C 1211
4. C 433 81
5. CAMD 75
6. CHN 204
7. CNPA Prec 1
8. MAR 17-30-76
9. PEXA
10. PRONTO
11. STONE 7A gl.



Appendix 17: Composition of Cluster IV at Garoua.

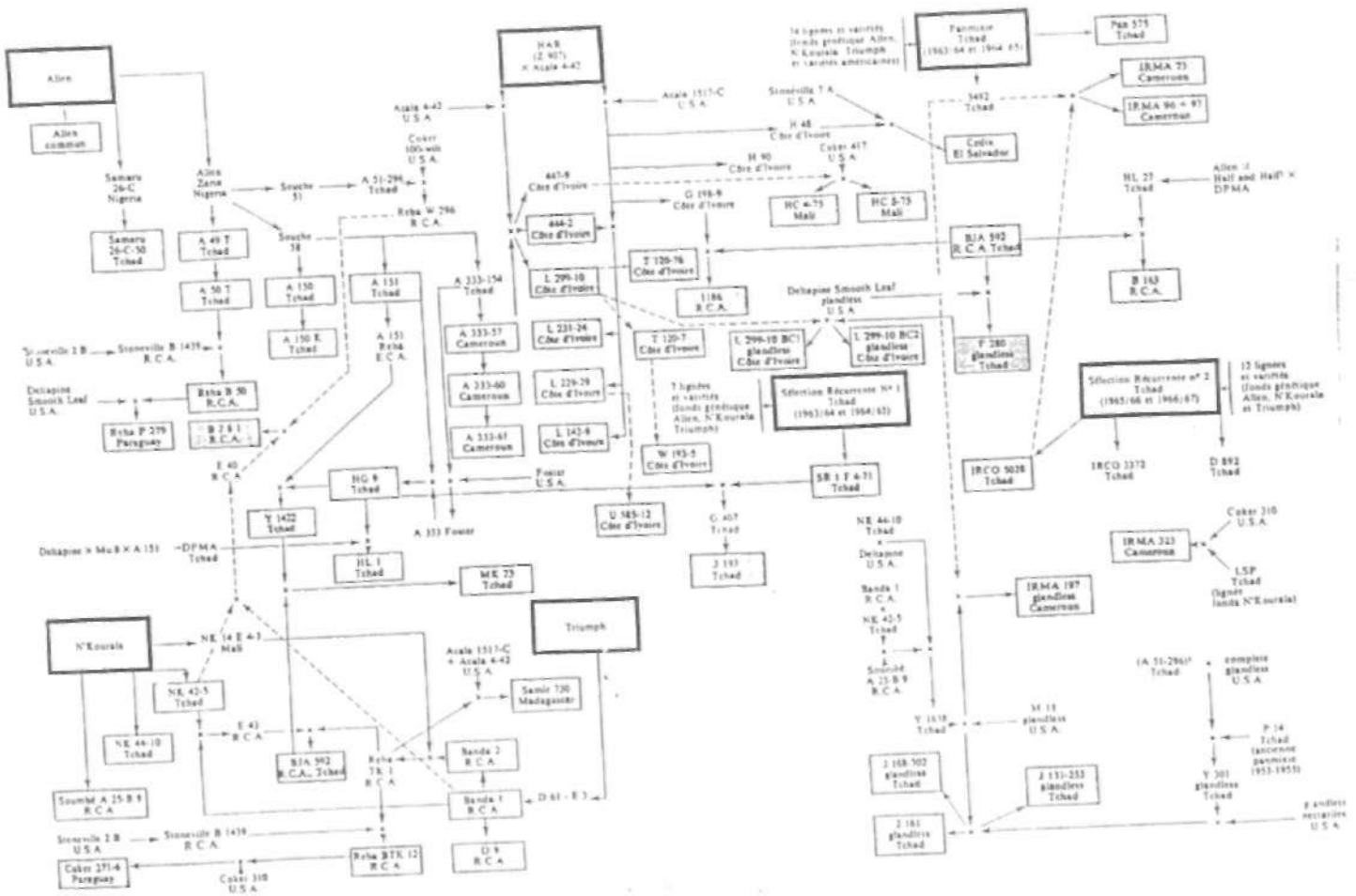


Appendix 18: Composition of Cluster V a
t Garoua.

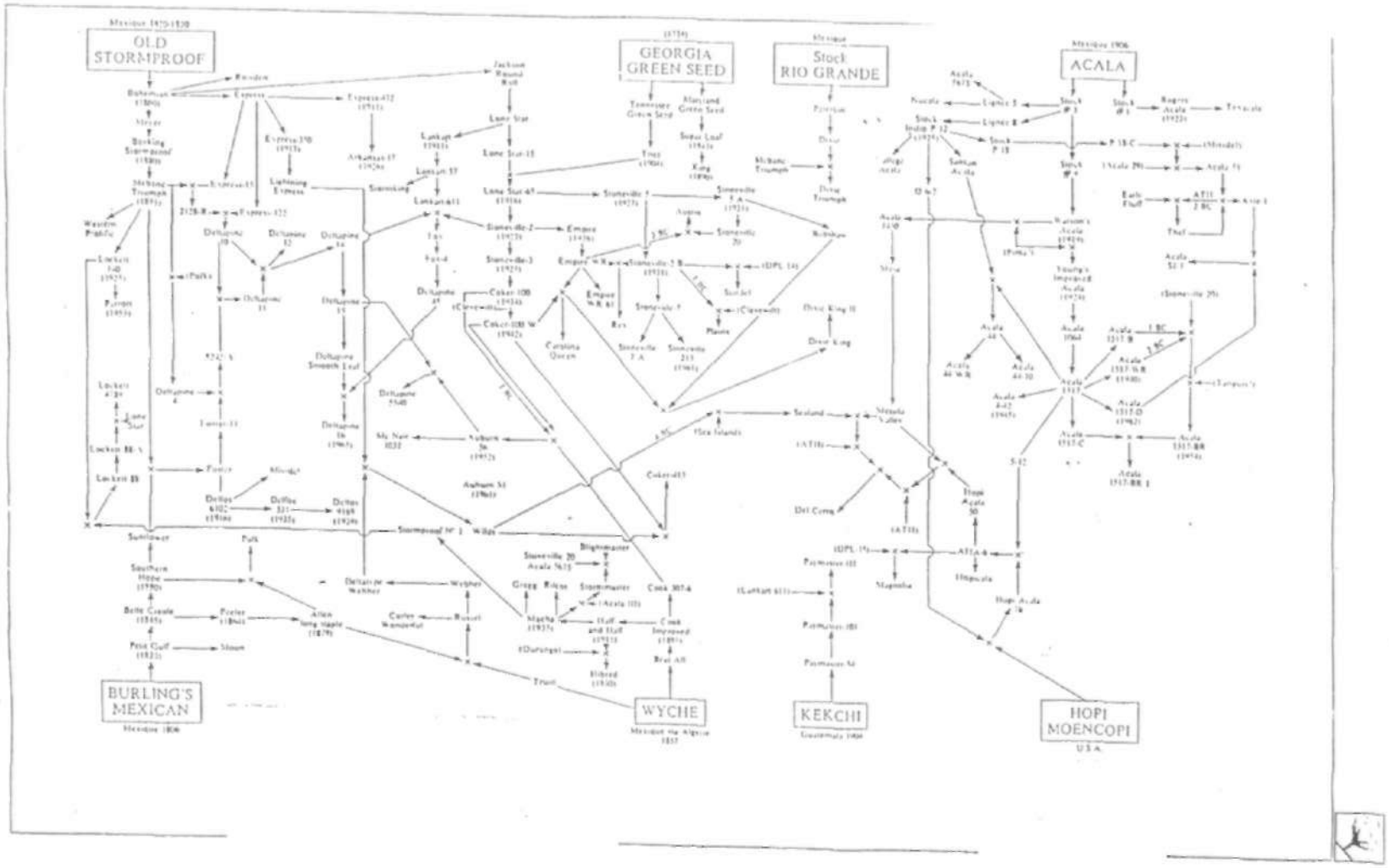


Appendix 19: Composition of Cluster VI at Garoua.

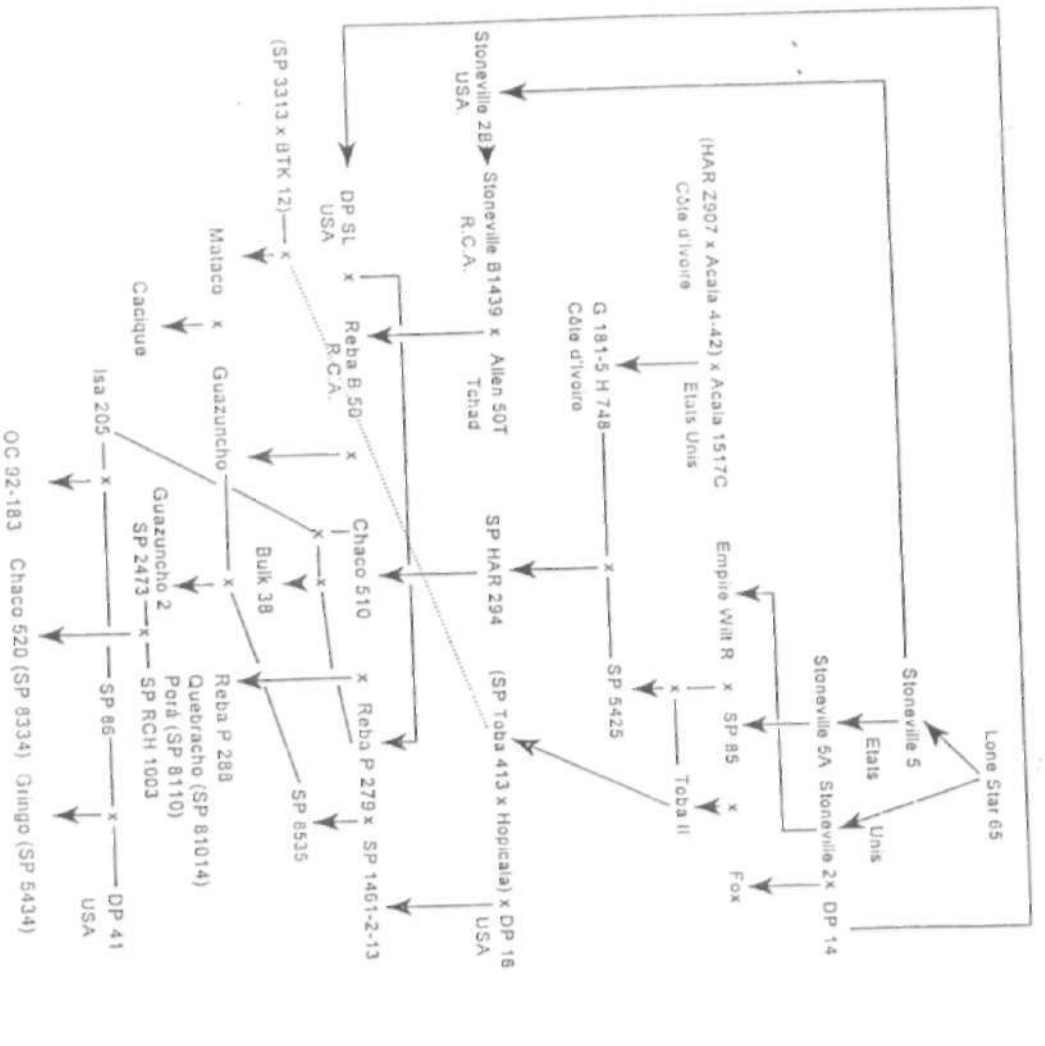
Appendix 20: Pedigree Scheme of Upland Cotton varieties Created by or with the collaboration of the French Institute of Cotton Research (IRCT) Montpellier, France.



Appendix 21: Pedigree Scheme of some U.S.A. Upland Cotton varieties.



Appendix 22: Pedigree Scheme of some Upland Cotton varieties from Paraguay, Argentina,



Appendix 23: Pedigree scheme of some upland cotton varieties developed in Nigeria.

