MORPHO-AGRONOMICAL AND MOLECULAR MARKER BASED GENETIC DIVERSITY ANALYSES AND QUALITY EVALUATION OF SORGHUM [*Sorghum bicolor* (L.) Moench] GENOTYPES

by

Nemera Geleta Shargie

A dissertation submitted in accordance with
the academic requirements for the degree

of

Philosophiae Doctor

in the

Department of Plant Sciences (Plant Breeding)
Faculty of Natural and Agricultural Sciences at the
University of the Free State
Bloemfontein, South Africa

Supervisor: Professor M.T. Labuschagne (Ph.D.)
Co-Supervisors: Dr. C.D. Viljoen (Ph.D.) and Professor G. Osthoff (Ph.D.)

May 2003

Dedicated to

My Father

DECLARATION

I declare that the dissertation submitted hereby for the degree of Philosophiae Doctor in Agriculture in the University of the Free State is an original work and has not been previously submitted by me to another University.

I further concede copy right of the dissertation in favour of the University of the Free State.

Nemera Geleta

Nemera Geleta Shargie

May 2003

ACKNOWLEDGEMENTS

I would like to thank the following individuals, organizations and/or institutions for their contribution towards the success of this dissertation:

- Professor Maryke T. Labuschagne for her efficient supervision and continued support, motivation and enthusiasm.
- Dr. Chris D. Viljoen for his excellent co-supervision and vital theoretical and practical input.
- Professor G. Osthoff for his interest, consistent encouragement and guidance in the food quality evaluation.
- The Agricultural Research & Training Project (ARTP) from the World Bank through the Alemaya University (AU), Ethiopia, for the financial support of my study and the research. The Alemaya University for giving me the study opportunity.
- The sorghum growers of the study area for their willingness and cooperation during research sample collection.
- The Institute of Biodiversity Conservation and Research of Ethiopia for giving me the permission to take the sorghum accessions, included in the research, to South Africa.
- Ms Elizma Koen of the Genetics and Molecular Biology Laboratory for her excellent assistance, knowledge and help when required.
- Mrs C. Bothma for her professional assistance in the sensory evaluation of sorghum *injera*.
- Dr. A. Hugo, Miss Maryna de Wit, Miss Eileen Roodt and Mr. W. Combrick for their technical assistance during the chemical composition determination.
- Mrs. Christina Matla of the Small Grains Institute, Bethlehem, for her help in milling the samples included in the sensory evaluation.

- Mr. Shimelis Woldehawariat, the coordinator of the ARTP/AU for all his support throughout my study and the rest of the staff for their assistance in facilitating the research materials and logistics.
- Personnel of the Plant Breeding division especially Mrs. Sadie Geldenhuys for providing excellent support, encouragement and the creation of a good working environment.
- All personnel and students of the Molecular and Genetics Laboratory at the previous Dept. of Botany and Genetics (the current Dept. of Plant Sciences) for their technical support and assistance.
- Personnel of the Sorghum Improvement Program of the Alemaya University, especially Miss Birtukan Yimam, for the technical support during the field experiment.
- Mrs. Heleni Girma and my wife Shashitu Barkessa for their help in preparation of the *injera* for the sensory evaluation. In addition, my wife, for her care, encouragement, support and patience during the study.
- All colleagues and friends that directly or indirectly have influenced my studies.
- God, who made all these possible.

TABLE OF CONTENTS

PAGE

DECLARATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES x

LIST OF FIGURES xiii

SYMBOLS AND ABBREVIATIONS xv

APPENDICES xvii

CHAPTER 1 GENERAL INTRODUCTION ---- 1

REFERENCES ---- 4

CHAPTER 2 LITERATURE REVIEW ---- 6

2.1 Introduction ---- 6

2.2 Morpho-agronomic traits as markers ---- 8

2.2.1 Qualitative traits ---- 8

2.2.2 Quantitative traits ---- 8

2.3 DNA-based molecular marker systems (DNA fingerprinting) ---- 9

2.3.1 Restriction fragment length polymorphisms (RFLPs) ---- 12

2.3.2 Polymerase chain reaction (PCR) – based techniques ---- 13

2.3.2.1 Random amplified polymorphic DNA (RAPD) markers ---- 14

2.3.2.2 The amplified fragment length polymorphisms (AFLP's) ---- 14

2.3.2.3 Microsatellites or simple sequence repeats (SSRs) ---- 17

2.4 Genetic distance analysis ---- 18

2.5 Comparison of major marker systems ---- 20

2.6 Food quality characteristics ---- 22

2.6.1 Physical properties and chemical composition ---- 23

2.6.1.1 Physical properties ---- 23

2.6.1.2 Chemical composition ---- 23
2.6.1.2.1 Protein content ---- 23
2.6.1.2.2 Lipid content ---- 24
2.6.1.2.3 Carbohydrate content ---- 25
2.6.1.2.4 Polyphenol / Tannins ---- 27
2.6.2 Food quality / Sensory evaluation ---- 28
REFERENCES ---- 31
CHAPTER 3 PHENOTYPIC DIVERSITY FOR MORPHO-AGRONOMICAL TRAITS IN SORGHUM ---- 50
Abstract ---- 50
3.1 Introduction ---- 50
3.2 Materials and methods ---- 52
3.2.1 Qualitative traits ---- 52
3.1.1.1 Plant material ---- 52
3.1.1.2 Methods ---- 53
3.1.1.3 Data analysis ---- 53
3.2.2 Quantitative traits ---- 54
3.2.2.1 Location of the study ---- 54
3.2.2.2 Plant materials ---- 54
3.2.2.3 Parameters measured ---- 54
3.2.2.4 Statistical analysis ---- 59
3.3 Results and discussion ---- 59
3.3.1 Qualitative traits ---- 59
3.3.2 Quantitative traits ---- 61
3.3.2.1 Clustering ---- 61
3.3.2.2 Principal component analysis ---- 64
3.4 Conclusions ---- 70
REFERENCES ---- 71
CHAPTER 4 ANALYSIS OF GENETIC DIVERSITY BASED ON DNA MARKERS IN SORGHUM ---- 74
Abstract ---- 74

4.1 Introduction ---- 74
4.2 Materials and methods ---- 76
4.2.1 Plant material ---- 76
4.2.2 DNA extraction ---- 76
4.2.3 DNA concentration determination ---- 76
4.2.4 AFLPs ---- 77
4.2.4.1 Restriction endonuclease digestion and ligation of adaptors ---- 77
4.2.4.2 PCR amplification reactions ---- 77
4.2.4.3 AFLP analysis ---- 79
4.2.4.4 Statistical analysis ---- 79
4.2.5 Microsatellites (SSRs) ---- 79
4.2.5.1 SSR primers ---- 79
4.2.5.2 SSR amplification ---- 80
4.2.5.3 SSR locus visualization / gel analysis ---- 80
4.2.5.4 Data collection and statistical analyses ---- 80
4.3 Results and discussion ---- 82
4.3.1 AFLP markers ---- 82
4.3.1.1 Level of AFLP polymorphism ---- 82
4.3.1.2 Genetic distance and cluster analysis ---- 84
4.3.1.3 Principal component analysis ---- 87
4.3.2 Microsatellite markers ---- 88
4.3.2.1 Polymorphism of SSRs in sorghum accessions - 88
4.3.2.2 Genetic diversity ---- 89
4.3.3 Comparison of AFLP and SSR markers ---- 93
4.4 Conclusions ---- 95
REFERENCES ---- 96
CHAPTER 5 COMPARISON OF AFLP, SSR AND MORPHO-AGRONOMICAL TRAITS MARKERS FOR ESTIMATING GENETIC DIVERSITY IN SORGHUM ---- 99
Abstract ---- 99

5.1 Introduction ---------- 100
5.2 Materials and methods ---------- 101
5.2.1 Plant material ---------- 101
5.2.2 Methods ---------- 101
5.2.2.1 Morpho-agronomical traits ---------- 101
5.2.2.2 DNA markers ---------- 102
5.2.2.2.1 AFLP's ---------- 102
5.2.2.2.2 Microsatellites (SSR's) ---------- 102
5.2.3 Data analysis ---------- 102
5.3 Results and discussion ---------- 103
5.3.1 Level of polymorphism detection ---------- 103
5.3.2 Genetic diversity estimation ---------- 103
5.3.3 Clustering based on AFLP, SSR and morpho-agronomical markers ---------- 105
5.4 Conclusions ---------- 111
REFERENCES ---------- 112
CHAPTER 6 PHYSICO-CHEMICAL ANALYSIS OF SORGHUM AND SENSORY EVALUATION OF *INJERA* ---------- 115
Abstract ---------- 115
6.1 Introduction ---------- 116
6.2 Materials and methods ---------- 118
6.2.1 Selection of sorghum accessions ---------- 118
6.2.2 Methods ---------- 119
6.2.3 Chemical composition ---------- 119
6.2.3.1 Moisture content ---------- 119
6.2.3.2 Protein content ---------- 119
6.2.3.3 Lipid extraction and methylation ---------- 119
6.2.3.4 Fatty acid analysis ---------- 120
6.2.3.5 Starch content ---------- 120
6.2.3.6 Amylose content ---------- 122
6.2.3.7 Polyphenol / Tannin content ---------- 122

6.2.4 Sensory evaluation of *injera* ---------- 123
6.2.4.1 Sorghum samples ---------- 123
6.2.4.2 Methods of sensory evaluation---------- 124
6.2.5 Statistical analyses ---------- 125
6.3 Results and discussion ---------- 126
6.3.1 Chemical composition ---------- 126
6.3.2 Sensory evaluation of *injera* ---------- 132
6.3.3 Relationships between physiochemical properties and *injera* quality ---------- 137
6.4 Conclusions ---------- 139
REFERENCES ---------- 140
CHAPTER 7 GENERAL CONCLUSIONS ---------- 144
CHAPTER 8 SUMMARY ---------- 146
OPSOMMING---------- 148

LIST OF TABLES

TABLE PAGE

3.1 Local / cultivar name, collection site, altitude and status of sorghum samples used in the study ---------- 55

3.2 Character, descriptor and codes used for characterisation of sorghum accessions ---------- 57

3.3 List, code and descriptions of the quantitative characters recorded in the study ---------- 58

3.4 Estimates of H', partitioning into within and between collection sites for 10 qualitative characters in sorghum accessions ---------- 60

3.5 Estimates of the Shannon-Weaver diversity index, H', for 10 qualitative characters in sorghum accessions by location of collection ---------- 61

3.6 Distribution of the 45 sorghum accessions into seven clusters by location of origin using average values of quantitative characters ------- 64

3.7 Cluster means for 16 quantitative characters in 45 sorghum accessions ---------- 66

3.8 Correlation coefficients (n = 45) between quantitative traits and grain yield per panicle, head weight and grain number per panicle in sorghum ---------- 67

3.9 Principal component (PC) analysis of 16 quantitative traits in 45 sorghum accessions showing eigenvectors, eigenvalues and proportion of variations explained with the first eight PC axes ---------- 68

4.1 Adaptors and primers used for pre-selective and selective AFLP amplification reactions ---------- 78

4.2 Summary of the SSR-primer pairs used in this study ---------- 81

4.3 The number of AFLP fragments and degree of polymorphism determined for 45 sorghum accessions using eight AFLP primer combinations---------- 83

4.4 Number of alleles, size range (in base pairs), and PIC values for SSR loci found in 45 accessions of sorghum ---------- 90

4.5 Level of polymorphism and comparison of the amount of information obtained with AFLP and SSR markers in 45 sorghum accessions---------- 95

5.1 Number of polymorphic bands, average and range of pair-wise genetic distance (GD) estimates among 45 sorghum accessions based on AFLP, SSR, and morpho-agronomical traits data ---------- 104

5.2 Correlation coefficients between genetic distance values estimated for the three techniques (AFLP, SSR and morpho-agronomical traits), with sample size of 990 ---------- 105

6.1 List of sorghum accessions included in the chemical composition determination ---------- 118

6.2 Average values and mean separation of the chemical compositions for 13 sorghum accessions---------- 127

6.3 Fatty acid compositions (%) of sorghum accessions ---------- 128

6.4 Average tannin content and mean separation for six sorghum accessions ---------- 129

6.5 Correlation coefficients (n = 39) between chemical and physical properties for 13 sorghum accessions---------- 130

6.6 Principal component (PC) analysis of eight physiochemical parameters in 13 sorghum accessions with eigenvectors, eigenvalues and proportion of variations explained by the first three PC axes ---------- 131

6.7 Rank sums and significance tests for *injera* colour ---------- 133

6.8 The rank sums and differences between products along with the significance levels for 'eye' quality of *injera* ---------- 134

6.9 Summarised results for under side appearance of *injera* from six sorghum samples ---------- 135

6.10 Summarised rank sums and significance tests for texture of *injera* -- 135

6.11 Rank sums and significance tests for *injera* taste results ---------- 136

6.12 Correlation coefficients between *injera* quality and some physico-chemical parameters in sorghum ---------- 138

LIST OF FIGURES

FIGURE PAGE

3.1 Geographic locations where the sorghum landrace accessions used in the study were collected. ---------- 56

3.2 Dendrogram showing cluster groups among the 45 sorghum accessions based on 16 quantitative traits data ---------- 63

3.3 Principal component plot of the 45 sorghum accessions, estimated using 16 quantitative traits data ---------- 69

4.1 Genescan electropherograms showing part of the AFLP runs of bulked DNA of two accessions (ETS 721 and ETS 2752) using *EcoRI*/ACA and *MseI*/CTA primers in the present study. ---------- 84

4.2 Frequency distribution of pair-wise genetic distance coefficients obtained for 45 sorghum accessions using AFLP data. ---------- 84

4.3 Dendrogram constructed based on AFLP data, showing genetic distance and cluster groups among 45 sorghum accessions ---------- 86

4.4 Plot of the 45 sorghum accessions against the first two principal components (PC1 and PC2) based on AFLP data ---------- 87

4.5 Agarose gel electrophoresis of SSR-PCR products amplified using primer sb6-36 $(AG)_{19}$ ---------- 89

4.6 Frequency distribution of pair-wise genetic distance coefficients among 45 sorghum accessions based on SSR data ---------- 91

4.7 Dendrogram constructed for 45 accessions of sorghum based on data from 43 ploymorphic SSR alleles ---------- 92

4.8 Plot of the 45 sorghum accessions against the first two principal components analysis (PC1 and PC2) computed using the SSR data ---------- 93

5.1 Dendrograms of 45 sorghum accessions constructed using dissimilarity matrix from (a) morpho-agronomic, (b) AFLP, (c) SSR, (d) AFLP + SSR, and (e) combined data. ---------- 110

6.1 Standard graph for glucose content determined by duplicate data points---------- 121

6.2 Standard amylose curve determined by duplicate data points ---------- 122

6.3 PC plot of 13 sorghum accessions analysed using eight physicochemical parameters. ---------- 132

6.4 *Injera / bidena* prepared from two sorghum accessions, *Ambajeette* (A) and ETS 1005 (B) showing the colour variation and the distribution of the 'eyes'---------- 134

6.5 The six sorghum *injera / bidena* samples evaluated by the panellists. ---------- 136

6.6 Bar graph showing the relationship between tannin content, *injera* quality, endosperm texture, grain colour and 1000-kernel weight in six sorghum samples ---------- 138

SYMBOLS AND ABBREVIATIONS

A = absorbance

AFLP = amplified fragment length polymorphism

AU = Alemaya University

bp = base pair

°C = degree Celsius

% = percent

CTAB = cetyltrimethyl ammonium bromide

cm = centimeter

DMSO = dimethyl sulphoxide

DNA = deoxyribonueclic acid

DNS = dinitrosalicylic acid

DNTP = deoxynucleoside triphosphate

EDTA = ethylenediamin tetra acetic acid

et al = 'et alii / alia' (and others)

g = gram

GD = genetic distance

h = hour

min = minute

H_2O = water

HCl = hydrochloric acid

KCl = potassium chloride

kg = kilograms

LSD = least significant difference

m = meter

M = molar

mol = mole

mg = milligram

$MgCl_2$ = magnesium chloride

ml = milliliter

mM = millimolar
NaCl = sodium chloride
NaOH = sodium hydroxide
NCSS = number cruncher statistical system
ng = nanogram
nm = nanometre
OD = optical density
PCA = principal component analysis
PCR = polymerase chain reaction
PIC = polymorphism information content
RAPD = random amplified polymorphic DNA
RFLP = restriction fragment length polymorphism
RFU = reflective fluorescent unit
rpm = revolutions per minute
SDS = sodium dodecyl sulphate
sec = second
SIP = Sorghum Improvement Program, Alemaya University
SSR = simple sequence repeat
TAE = Tris, acetic acid and EDTA
Taq = *Thermus aquaticus*
TE = Tris EDTA
Tris-HCl = (Tris [hydroxymethyl] aminomethane) hydrochloric acid
µg = microgram
µl = microlitre
µM = micromolar
UPGMA = unweighted pair group method using arithmetic averages
UV = ultraviolet

APPENDICES

Appendix I AFLP fragment scores for primer combination *EcoI* + ACA / *MseI* + CAA ---------- 150

Appendix II Amplified fragments score for 10 microsatellite loci: sb1-10, sb4-15, sb4-22, sb4-32, sb5-85, sb5-236, sb6-36, sb6-57, sb6-84 and sb6-342. ---------- 156

Appendix III Average values calculated for 16 quantitative traits in 45 sorghum accessions ---------- 160

Appendix IV Binary score of qualitative traits for 45 sorghum accessions ---------- 161

Appendix V Binary scores of quantitative traits for 45 sorghum accessions ---------- 164

Appendix VI An example of the pair-wise genetic distances estimated between some of the accessions based on morpho-agronomical data ---------- 169

Appendix VII An example of the pair-wise genetic distances estimated between some of the accessions based on AFLP data ---------- 170

Appendix VIII An example of the Pair-wise genetic distances estimated between some of the accessions based on microsatellites data ---------- 171

Appendix IX Partial view of the sensory evaluation of the *injera* --------- 172

CHAPTER 1

GENERAL INTRODUCTION

Analyses of the extent and distribution of genetic variation in a crop are essential for understanding the evolutionary relationships between accessions and to sample genetic resources in a more systematic fashion for breeding and conservation purposes. Sorghum [*Sorghum bicolor* (L.) Moench, 2n = 20] is fifth in importance among the world's cereals (Doggett, 1988). It is the major crop in warm, low-rainfall areas of the world. It is a crop with extreme genetic diversity (Subudhi *et al.*, 2002) and predominantly a self-pollinating crop, with various levels of outcrossing. The greatest variability is found in the northeast quadrant of Africa, which includes Ethiopia, Eritrea and Sudan, and most evidence points to this area as the likely principal area of its domestication (Vavilov, 1951; Doggett, 1970, 1988; House, 1985). According to Gebrekidan (1973, 1981), in Ethiopia, sorghum exists in tremendous diversity throughout the growing areas, which contain pockets of isolation with an extremely broad and valuable genetic base for potential breeding and improvement in the country and the world at large.

Known under the generic name *bishinga* by Oromo people, various types of sorghum are widely cultivated on the highlands of eastern Ethiopia. Landraces are more preferred by the farmers due to their adaptation to specific environmental conditions and additional characters such as storability, food quality, and/or amount and quality of by-products. The diverse growing environments and the preference of farmers to grow landraces are ideal for maintenance of a wide range of sorghum types. But, according to Klingele (1998) the crop is facing a serious challenge from shrinking of individual land holdings due to the expansion of a cash crop chat (*Catha edulis*). Moreover, Maxted *et al.* (2002) indicated that though most genetic diversity of immediate and potential use to plant breeders is found among landraces there is evidence that it is being rapidly eroded.

Evaluation of genetic diversity can indicate which landraces carry the greatest genetic novelty, and are the most suitable for rescue, and possible future use in crop improvement. Furthermore, to improve and stabilize production and utilization of sorghum in the area, new lines of sorghum should also yield equal or better than existing landraces familiar to farmers. Evaluation of genetic diversity levels among adapted, elite germplasm can provide predictive estimates of genetic variation among segregating progeny for pureline cultivar development (Manjarrez-Sandoval et al., 1997). The use of germplasm developed within the same region targeted for cultivar improvement reduces the risk of losing essential adaptive characteristics through recombination (Allard, 1996). To improve yield and other consumer preferred traits through the use of landraces; therefore, complete information of the genetic diversity and the physical and chemical properties of sorghums available in the region is a priority.

The accurate, fast, reliable, and cost-effective identification of plant populations and varieties is essential in agriculture as well as in pure and applied plant research (Morell *et al.*, 1995). Traditionally, taxonomists classify genetic resources in sorghum based on morphological traits (Stemler *et al.*, 1977). In the first instance, this usually involves description of variation for morphological traits, particularly morpho-agronomical characteristics of direct interest to users. While these methods are very effective for many purposes, morphological comparisons may have limitations including subjectivity in the analysis of the character; the influence of environmental or management practices on the character; limited diversity among cultivars with highly similar pedigrees; and confining of expression of some diagnostic characters to a particular stage of development, such as flowering or seed maturity. Menkir et al. (1997) indicated that important traits, which are related to habitat adaptation and particular end use of the crop, exhibit enormous variability among sorghum germplasm. Hence, classifying germplasm accessions based solely on morphological characters may not provide an accurate indication of the genetic divergence among the cultivated genotypes of sorghum.

These considerations have led to the exploration or adoption of other techniques for genetic diversity estimation and cultivar identification, including cytogenetic analysis; isozyme analysis; and molecular techniques that analyse polymorphism at the DNA level directly. Molecular markers are nowadays widely used as tools to assess the soundness of morphological classification in crop plants. The amplified fragment length polymorphisms (AFLPs) and microsatellites or simple sequence repeats (SSRs) DNA markers have proved to be efficient and reliable in supporting conventional plant breeding programmes (Paterson *et al.*, 1991; Morell *et al.*, 1995; Kumar, 1999).

In this study, the level of genetic diversity was determined among 34 sorghum accessions that were sampled directly from farmers' fields and 11 elite breeding lines, using morpho-agronomic traits, DNA marker techniques (AFLP's and microsatellite's or SSR's), and evaluated for chemical composition and food quality characteristics.

General Objectives

1. To analyze the extent of genetic diversity in sorghum accessions from the eastern Ethiopian highlands using morpho-agronomical characters;
2. To examine the genetic diversity among accessions using DNA (AFLP and SSR) markers;
3. To verify how useful AFLP and SSR's markers are in determining distinctiveness of sorghum accessions;
4. To provide an example of the combined use of AFLP and SSR profiles, and highly reliable morpho-agronomical characters for diversity assessment;
5. To assess variability for chemical composition and quality characteristics, and see its integration with morpho-agronomical and DNA marker data;
6. To examine the distribution of genetic variation in different localities and give recommendations for future conservation and breeding strategies.

REFERENCES

Allard, R.W. 1996. Genetic basis of the evolution of adaptedness in plants. Euphytica 92:1-11.

Doggett, H. 1970. Sorghum. Longmans. Green & Co. Ltd. London.

Doggett, H. 1988. Sorghum. Longmans (Second edition). Green & Co. Ltd., London.

Gebrekidan, B. 1973. The importance of the Ethiopian sorghum in the world sorghum collections. Econ. Bot. 27:442-445.

Gebrekidan, B. 1981. Salient features of the sorghum breeding strategies used in Ethiopia. Eth. J. Agri. Sci. 3:97-104.

House, L.R. 1985. A guide to sorghum breeding (Second Edition). Patancheru, A.P., ICRISAT, India.

Klingele, Ralph. 1998. Hararghe farmers on the cross-roads between subsistence and cash economy. http://www.telecom.net.et/~undp-eue/reports/hararghe.html.

Kumar, L.S. 1999. DNA markers in plant improvement: An overview. Biotechnology Advances 17:143-182.

Manjarrez-Sandoval, P., Carter, T.E., Webb, D.M. and Burton J.W. 1997. RFLP genetic similarity estimates and coefficient of parentage as genetic variance predictors for soyabean yield. Crop Sci. 37:698-703.

Maxted, N., Guarino, L., Myer, L., and Chiwona, E.A. 2002. Towards a methodology for on-farm conservation of plant gentic resources. Genet. Resour. Crop Evol. 49:31-46.

Menkir, A., Goldsbrough, P. and Ejeta, G. 1997. RAPD based assessment of genetic diversity in cultivated races of sorghum. Crop Sci. 37:564-569.

Morell, M.K., Peakall, R., Appels, R., Preston, L.R., and Lloyd, H.L. 1995. DNA profiling techniques for plant variety identification. Australian J. Exp. Agric. 35:807-819.

Paterson, A.H., Tanksley, S.D. and Sorrells, S.M. 1991. DNA markers in plant improvement. Advances in Agronomy 46:39-90.

Stemler, A.B., Harlan, J.R., and de Wet, J.M.T. 1977. The sorghums of Ethiopia. Eco. Bot. 31:446-460.

Subudhi, P.K., Nguyen, H.T., Gilbert, M.L., and Rosenow, D.T. 2002. Sorghum improvement: past achievements and future prospects. In: Kang, M.S. (Ed.). Crop improvement: Challenges in the twenty-first century. The Haworth Press, Inc., NY, pp 109-158.

Vavilov, N.I. 1951. The origin, variation, immunity and breeding of cultivated plants. Chronica Botanica 13:1-366.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Genetic diversity, the heritable portion of observable variation, is the raw material on which natural and artificial selection has acted to create earth's vast array of organisms. Estimation of genetic diversity in plant species can assist in the evaluation of different germplasm as possible sources of genes that can improve the performance of cultivars. As a result, qualitative and quantitative traits can be more efficiently introduced into plant breeding programmes. In search for diverse breeding material, landraces or farmer varieties (locally adapted populations bred through traditional methods of direct selection) are usually the major sources of genetic variation. As discussed by Jain (1975), within the same species, estimates of the amount of variation may vary widely, depending upon the area sampled, geographical scale of sampling, etc., presumably due to the complex interrelationships between the genetic, ecological as well as historical variables. Individual loci can also vary widely, due to both adaptive and/or non-adaptive reasons, in the geographical distribution of alleles. This, of course, adds a great deal to the variation and, therefore, uncertainty about the expected allelic frequency distribution at any specific locus, in any individual population. It appears that numerous complex environmental gradients and the high phenotypic plasticity characteristic of species often yield highly irregular variation patterns (both phenotypic and genetic criteria).

Sorghum is known for the ability to grow in harsh environments and has numerous mechanisms that allows it to survive and be productive in these conditions. Despite the importance of the sorghum crop, comprehensive genetic characterization has been limited (Subudhi and Nguyen, 2000a). Harlan and de Wet (1972) classified traditional sorghum cultivars into five

main races (bicolor, caudatum, durra, guinea, and kafir) and 10 intermediates, mainly on the basis of the morphology of spikelets and grains. Of the five basic races, four races (bicolor, caudatum, durra, and guinea) are reported found in Ethiopia (Stemler *et al.*, 1977). According to Harlan (1992), the intermediate races involving these four basic races also widely occur in Ethiopia.

Efforts have been made to identify the different accessions/cultivars of Ethiopian origin sorghum germplasm based on morphological characters (Ayana and Bekele, 1998, 1999; Teshome *et al.*, 1997; Geleta, 1997; Abebe and Wech, 1982; Gebrekidan and Menkir, 1979). However, quantitative traits are influenced by environmental factors and show variation, resulting in low heritability and high genotype by environment interactions. Consequently, it is difficult to accurately determine genetic diversity. However, a continued use of morphological data to describe cultivars indicates that these data retain popularity as descriptors (Smith and Smith, 1992). Due to their limited number of detectable loci, allozyme markers also did not clearly separate the various races of cultivated and wild sorghum accessions into distinct classes (Ayana, 2001).

Advances in molecular biology provided new methodologies, which extend the list of useful genetic markers (Paterson *et al.*, 1991). Determining the genetic diversity of the different accessions at the DNA level can hold many advantages for the plant breeder, since it may increase the efficiency of breeding efforts to improve crop species (Barrett *et al.*, 1998). This may explain the reason for the development of the different marker techniques. On the other hand, Karp *et al.* (1997) have pointed out that DNA markers should not be seen as a substitute for other agro-morphological or biochemical studies that provide researchers with the information they need. The results of molecular or biochemical studies should be considered as complementary to morphological characterisation. In this review, the main techniques available to analyse variation and their major features have been dealt with.

2.2 Morpho-agronomic traits as markers

2.2.1 Qualitative traits

Morphological traits, for which the variant allelic phenotypes are sufficiently discrete to allow their segregation to be followed, are the easiest and generally most economical of all markers to assay. Discrete morphological traits, though they have high heritability, are limited in number, each being conditioned by a few genes (Karp *et al.*, 1996, 1997). Thus, only a small portion of the genome could be covered. They are usually characterised by epistasis, pleiotropy and dominant-recessive relationships, further limiting their values as an ideal genetic marker (Smith and Smith, 1992). Besides, morphological characterisation requires mature plants, it usually displays dominant phenotype and there are too few available in single species (Koebner *et al.*, 1994). The method involves a lengthy survey of plant growth that is labour intensive and time consuming (CIAT, 1993).

In sorghum, as is true for other crop plants, the earliest methods for estimating genetic diversity include Mendelian analysis of discrete morphological traits (Doggett, 1988). By morphological trait study, earlier works have shown that eastern Ethiopian sorghum is believed to be predominantly race durra (Doggett, 1988; Stemler *et al.*, 1977; Brooke, 1958). Using *ex situ* conserved sorghum accessions from Ethiopia and Eritrea, Ayana and Bekele (1998) reported that high and comparable levels of phenotypic variation exists between the regions of origin. Nevertheless, *in situ* pattern of genetic diversity at country as well as regional scale has not been investigated and remains less understood.

2.2.2 Quantitative traits

Multivariate analysis such as clustering and principal component analysis of quantitative characters has been used previously to measure genetic relationships within cereal species. Examples include tef (*Ergrostis tef* (Zucc.) Trotter (Assefa *et al.*, 1999); barley (*Hordeum vulgare* L.), (Bekele,

1984); Ethiopian wheats (*T. aestivum* L.) (Negassa, 1986); durum wheats (*T. turgidum* L.) (Jain *et al.*, 1975). Statistical analysis of quantitative morpho-agronomical traits along with eco-geographic information (de Wet *et al.*, 1976) is one of the earliest methods used for estimating genetic diversity in sorghum. It is still widely used to quantify the amount and distribution of variation in large samples of sorghum germplasm collections (Prasada Rao and Ramanatha Rao, 1995; Teshome *et al.*, 1997; Ayana and Bekele, 1999). Using multivariate analysis procedures, Ayana and Bekele (1999) have revealed that the morphological variation in sorghum germplasm from Ethiopia and Eritrea was structured by environmental factors.

2.3 DNA-based molecular marker systems (DNA fingerprinting)

Accurate estimates of genetic diversity levels among and within crop plant species are becoming increasingly useful in crop improvement. During the past 20 years, DNA marker systems have become extremely useful tools for assessing genetic diversity levels within and between genotypes.

DNA fingerprinting involves the display of a set of DNA fragments from a specific DNA sample. Differences in DNA sequence are observed as the presence/absence of bands. These differences are characteristic and heritable. A number of problems in plant breeding can be addressed via a DNA-based molecular marker approach (Karp *et al.*, 1996). In addition to individual identification, DNA fingerprinting techniques can be used in tests of parentage, in genetic mapping of loci conditioning economic traits, in measurement of genetic diversity, and in discerning patterns of genetic diversity (Smith and Smith, 1992). The decision to exploit the possibilities opened up by the technology is more often influenced by economical or practical, rather than by technical considerations. The high variability of DNA fingerprinting described in humans, animals and plants allows the identification of different individual genotypes and species (Lin *et al.*, 1993).

Various DNA fingerprinting techniques have been successfully developed and put in use for estimation of genetic diversity in plant species, complementing the use of morphological markers. DNA techniques have the advantage over conventional methods in that the composition of DNA is consistent in similar tissue types and is not affected by environmental changes (Beeching *et al.*, 1993). The development of DNA markers provides an opportunity to detect, monitor and manipulate genetic variation (Yamamoto *et al.*, 1994) more precisely than in the case of morphological and expressed phenotypic markers, though results may be confounded by biased or incomplete genome coverage, detection of co-migrating non-homologous fragments, or high crossover frequency between markers used in the evaluation and linked genetic material (Barrett and Kidwell, 1998). These techniques include a variety of different methodologies commonly referred to as DNA fingerprinting (Nybom *et al.*, 1990). DNA molecular markers are potentially unlimited in number, are not affected by the environment and can be mapped on linkage maps (Soller and Beckmann, 1983; Winter and Kahl, 1995).

Morell *et al.* (1995) stated that DNA-based markers offer a number of advantages over isozymes and other biochemical methods for demonstrating distinctness. Firstly, the DNA sequence of an organism is independent of environmental conditions or management practices. Secondly, the presence of the same DNA in every living cell of the plant allows tests on any tissue at any stage of growth (provided that DNA of sufficient purity can be isolated). Thirdly, the recent advent of the polymerase chain reaction (PCR) has enabled the development of new DNA profiling techniques that are simply and quickly performed. These techniques offer a number of advantages over other DNA profiling techniques and conventional methods for identifying plants.

The DNA techniques have been used to investigate the extent of genetic diversity and genetic relationships within and between cultivars and elite materials of many plant species. In sorghum, molecular markers have been used to identify and characterise quantitative trait loci (QTL) associated

with several different traits including plant height and maturity (Pereira and Lee, 1995), characters concerned with plant domestication (Paterson *et al.*, 1995), disease resistance (Gowda *et al.*, 1995), and drought tolerance (Tuinstra *et al.*, 1996, 1997, 1998). In addition, several sorghum linkage maps (Hulbert *et al.*, 1990; Melake-Berhan *et al.*, 1993; Xu *et al.*, 1994; Chittenden *et al.*, 1994; Pereira *et al.*, 1994; Ragab *et al.*, 1994; Lin *et al.*, 1995; Dufour *et al.*, 1997; Boivin *et al.*, 1999) have been generated. Tao *et al.* (1998) constructed a sorghum map using a recombinant inbred line (RIL) population and a variety of probes, including sorghum genomic DNA, maize genomic DNA and cDNA, sugarcane genomic DNA and cDNA, cereal anchor probes, and eight SSR loci. Recently, Subudhi and Nguyen (2000b) completely aligned all the 10 linkage groups of all the major sorghum RFLP maps using common RIL populations and sorghum probes from all three sources (Chittenden *et al.*, 1994; Ragab *et al.*, 1994; Xu *et al.*, 1994) along with many cereal anchor and maize probes.

Over the past decade a number of DNA fingerprinting techniques have been developed to provide genetic markers capable of detecting differences among DNA samples across a wide range of scales ranging from individual/clone discrimination up to species level differences. Currently available techniques include: RFLPs (restriction fragment length polymorphisms, Liu and Furnier, 1993), DAF (DNA amplification fingerprinting, Caetano-Anolles and Gresshoff, 1994), AP-PCR (arbitrarily primed PCR, Welsh and McClelland, 1990), RAPDs (randomly amplified polymorphic DNAs, Williams *et al.*, 1990), microsatellites (Tautz, 1989), and most recently AFLPs (amplified fragment length polymorphisms, Zabeau and Vos, 1993; Vos *et al.*, 1995). At present, the information available on genetic diversity within cultivated sorghum utilised RFLPs (Aldrich and Doebley, 1992; Deu *et al.*, 1994; Cui *et al.*, 1995), RAPDs (Menkir *et al.*, 1997; Ayana et al., 2000; Dahlberg *et al.*, 2002), and SSRs (Smith *et al.*, 2000, Dje *et al.*, 2000, Grenier *et al.*, 2000, Brown *et al.*, 1996) techniques, with varying degrees of success.

2.3.1 Restriction Fragment Length Polymorphisms (RFLPs)

RFLP technology has pioneered the integration of DNA markers into molecular genetics and plant breeding. The evolution of chromosomal organization, taxonomic characterization, and the measurement of genetic diversity are some areas of study that have been greatly enhanced by the use of RFLPs (reviewed in Yang *et al.*, 1996). The first DNA profiling technique to be widely applied in the study of plant variation was the RFLP assay. In RFLP analysis, the complete digestion of genomic DNA with restriction endonucleases generates the detection of differences in the length of restriction fragments and the resultant fragments are separated by gel electrophoresis (Karp *et al.*, 1997; Beckman and Soller, 1983). RFLP analysis, as applied to other crops (Demissie *et al.*, 1998; Song *et al.*, 1988; Miller and Tanksley, 1990), as well as to sorghum (Aldrich and Doebley, 1992), has proven to be an additional, and more sensitive, tool for studying the amount of genetic diversity and the phylogenetic relationships among populations, accessions and species.

Prior RFLP diversity studies in sorghum found low frequencies of polymorphisms for 27 genotypes examined (Tao *et al.*, 1993), but much greater allelic diversity among RFLPs detected by maize probes than when isozymes were used to compare a set of 56 geographically and racially diverse accessions (Aldrich and Doebley, 1992), and Cui *et al.* (1995) reported that there was greater nuclear diversity in the wild subspecies than in the domestic accessions. Though exceptions were common, especially for the race bicolour, accessions classified as the same morphological race tended to group together on the basis of RFLP similarities (Cui *et al.*, 1995). In species such as maize, wheat, and soybeans, a large number of DNA probes are available, and extensive DNA profiling with RFLP analyses is feasible (Morell *et al.*, 1995). RFLP analysis requires relatively large amounts of DNA (often requiring destructive sampling) and produces relatively few bands or polymorphisms. And these conditions make RFLP a technique of lower priority.

2.3.2 Polymerase chain reaction (PCR)–based techniques

Saiki *et al.* (1985) indicated that the polymerase chain reaction (PCR) was invented by Kary B. Mullis in 1985 and has revolutionised many areas of biological science. The PCR relies on the use of a specific class of enzymes, DNA polymerase, which all living cells possess and use to copy their own DNA. DNA polymerase copies single-stranded DNA from the 3'OH end of double-stranded DNA. In PCR, the sample is first heated to separate the double-stranded DNA (denaturation step of three to five min. at 94-95°C) into single-stranded molecules. Next, the temperature is lowered to allow short synthetic DNA molecules called primers (typically 8-20 nucleotides in length) to anneal to complementary sequences (Rolfs *et al.*, 1992). These double-stranded complexes serve as starting points for the copying of single-stranded DNA polymerase. By flanking a region of DNA with specific DNA primers and cycling the temperature to facilitate strand separation, primer annealing, and primer extension, PCR can exponentially amplify a single copy of a DNA molecule to yield sufficient DNA for electrophoretic analysis (Morell *et al.*, 1995). The use of heat-stable DNA polymerases that survive the lengthy exposure to high temperatures, and the development of thermocyclers capable of cycling temperatures quickly and accurately, have facilitated the automation of this process. The most critical component for optimising the specificity of any PCR-based assay is the choice of the annealing temperature (Ruano *et al.*, 1991) until they find complementary annealing sites. Yu and Pauls (1992) concluded that the best results should be obtained by optimising for the shortest possible denaturing time. Too many cycles may result in primer depletion and subsequent priming by amplification products, which often leads to longer products and smears in the gel (Rolfs *et al.*, 1992).

The main advantage of the PCR-based technique over RFLP analysis is its inherent simplistic analysis and the ability to conduct PCR tests with extremely small quantities of tissue for DNA extraction (Edwards *et al.*, 1991). Currently, PCR is used worldwide in many areas of biology, agriculture, and medicine (Morell *et al.*, 1995).

2.3.2.1 Random amplified polymorphic DNA (RAPD) markers

The RAPD procedure is PCR based and allows a relatively large number of genetic loci to be assayed rapidly and inexpensively (Williams *et al.*, 1990). The assay has alleviated some of the technical problems associated with RFLP and has been widely used to resolve problems in plant breeding and genetics (Waugh and Powell, 1992). The DNA fragment patterns generated by this technique depend on the sequence of the primers and the nature of the template DNA. No prior sequence characterization of the target genome is needed and PCR is performed at low annealing temperatures to allow the primers to hybridise to multiple loci. RAPD markers have been used to estimate genetic diversity in several crops, including sorghum (Ayana *et al.*, 2000; Menkir *et al.*, 1997; Yang *et al.*, 1996). However, the need to repeat each PCR reaction multiple times and the inability to obtain identical banding patterns in different labs have limited the use of the RAPD technique (Bai *et al.*, 1999).

2.3.2.2 The amplified fragment length polymorphisms (AFLP's)

The AFLP technique, developed by Zabeau and Vos (1993) and Vos *et al.* (1995), is capable of detecting non-specific but many independent loci, with reproducible amplification (Pejic *et al.*, 1998). The AFLP fingerprints can be used to distinguish even very closely related organisms, including near isogenic lines. The technique involves a selective PCR amplification of restriction fragments from an endonuclease digest of total genomic DNA. The differences in fragment lengths generated by this technique can be traced to base changes in the restriction/adaptor site, or to insertions or deletions in the body of the DNA fragment. Dependence on sequence knowledge of the target genome is eliminated by the use of adaptors of known sequence that are ligated to the restriction fragments. The PCR primers are specific for the known sequences of the adaptors and restriction sites. In order to give reproducible results, several reaction components need to be optimised in the PCR reaction. Reaction components that should be optimised include template, primer, $MgCl_2$,

enzyme and dNTP concentration (Caetano-Anollés *et al.*, 1991). This usually relies on the sequential investigation of each reaction variable.

Most importantly, AFLPs have been shown to be reproducible and reliable. This is at least partially due to the fact that limited sets of generic primers are used and these are annealed to the target under stringent hybridisation conditions. The technique can be adjusted to generate consistent banding patterns from DNA of any origin or complexity, and no appreciable effect has been observed as a result of template concentration (in the range of 2.5 picogram to 25 nanogram). Typically, 50-200 bands are generated in a single lane after electrophoresis of the PCR amplified products on an analytical polyacrylamide gel. However, if the template concentration is too high then the PCR amplification often results in a smear without distinct bands (Yu *et al.*, 1993). Certain amplified fragments show continuous increase or decrease in band intensity depending on template concentrations (Hosaka and Hanneman, 1994). This implies that adjusting genomic DNA concentrations to the same level in all samples may be an initial step in obtaining reproducible and comparable banding patterns.

The AFLP data usually must be treated as dominant markers, since the identity of homo/heterozygotes cannot be established unless breeding/ pedigree studies are carried out to determine inheritance patterns of each band. However, the large number of bands gives an estimate of variation across the entire genome, thus giving a good general picture of the level of genetic variation. This type of information is generally more applicable to genotyping, forensics, and conservation biology than detailed information on variation at one or few loci (example, RFLPs, microsatellites, isozymes).

For DNA isolation, plants can be grown in a variety of environments and in different locations (Young, 1994). Any part of a plant can be used to extract DNA, however the most common starting material is young leaves. They can be fresh, lyophilised, dried in an oven or in some cases dried at room temperature (Kochert, 1994). Several methods for DNA extraction have

been developed; and simplicity, speed, and utilisation of a small amount of starting material are a common goal in all of them (Lamalay *et al.*, 1990).

The AFLP technique is rapidly becoming the method of choice for estimating genetic diversity in both cultivated and natural/rare populations (Karp *et al.*, 1997; Paul *et al.*, 1997; Qamaruz-Zaman *et al.*, 1997; Sharma *et al.*, 1996; Hill *et al.*, 1996; Lu *et al.*, 1996; Travis *et al.*, 1996). Additional characteristics of the AFLP technique are described as follows:

1. It is relatively fast (samples can be processed on automated thermocyclers and DNA sequencers).
2. The technique assays the entire genome for polymorphic markers.
3. It requires relatively small amounts of genomic DNA. Typically 0.05-0.5μg of DNA is required, depending upon the size of the genome.
4. It provides 10-100 times more markers and is thus more sensitive than other fingerprinting techniques (example, isozymes, RFLPs, microsatellites) (Lu *et al.*, 1996; Sharma *et al.*, 1996).
5. Unlike RAPDs, it is highly reproducible. Analyses performed by different workers or in different labs can be compared or reproduced. The bands (DNA fragments) can be run on an automated sequencer that resolves fragment length to single-base units. In addition, since each lane incorporates a set of size standards, fragment sizes can be estimated accurately thus facilitating comparison of data across gels.
6. Unlike microsatellites, no taxon-specific primer sets are required. Commercially available primers are available that work for most organisms.

There are many applications of AFLP markers, the genetic relationship studies being an important one (Incirli and Akkaya, 2001; Aggarwal *et al.*, 1999; Breyne *et al.*, 1999; Singh *et al.*, 1999; Schut *et al.*, 1997; Negash *et al.*, 2002). AFLP technology currently offers the fastest, most cost-effective way to generate high-density genetic maps for marker-assisted selection of desirable traits. It is also the ideal tool for determining varietal identity and assessing trueness to type (Perkin-Elmer, 1996).

In many species, AFLPs assay more loci per PCR than RAPD's, and have greater reproducibility (Russell *et al.*, 1997; Powell *et al.*, 1996), which has led to the increasing use of AFLPs for DNA profiling (Maheswaran *et al.*, 1997; Powell *et al.*, 1997; Maughan *et al.*, 1996). The suitability of AFLP analysis for cultivar identification, is demonstrated by the large number of reports published on the use of the technique for line identification in a variety of plant species, such as tomato, soybeans, brassicas, sunflower, pepper, sugar beet, lettuce (Perkin-Elmer, 1996), wheat (Donini *et al.*, 1997) and barley (Pakniyat *et al.*, 1997). However, AFLPs provide dominant markers in most cases and their distribution along the genome is not uniform (Subudhi and Nguyen, 2000a).

2.3.2.3 Microsatellites or simple sequence repeats (SSRs)

Microsatellites or simple sequence repeats (SSRs) are DNA sequences with repeat lengths of a few base pairs (2-6 bp). They are highly mutable loci and they are present at many sites throughout a genome. The flanking sequences at each of these sites are often unique. Variation in the number of repeats can be detected with PCR by developing primers for the conserved DNA sequence flanking the SSR. Specific primers can be designed according to the flanking sequences, which then result in single locus identification. Alleles that differ in length can be resolved using agarose gels or sequencing gels where single repeat differences can be resolved and all possible alleles detected. As molecular markers, SSR's combine many desirable marker properties including high levels of polymorphism and information content, unambiguous designation of alleles, even dispersal, selective neutrality, high reproducibility, co-dominance, and rapid simple genotyping assays (Jones *et al.*, 1997). For measuring genetic diversity, assigning lines to heterotic groups and genetic fingerprinting, microsatellites provide power of determination equal to or greater than that of RFLP in a more cost effective manner (Senior *et al.*, 1998; Smith *et al.*, 1997).

In actual fact, SSR markers are time consuming and costly to develop in that the genomic regions carrying them must be identified and sequenced. However, once the primers are developed, the technique is one of the most informative marker systems available. Even between closely related individuals, the number of repeat units at a locus is highly variable (Mazur and Tingey, 1995). SSR's are used to cluster lines into groupings (Liu and Wu, 1998; Senior *et al.*, 1998). SSR markers have shown high levels of polymorphism in many important crops including maize (Senior *et al.*, 1998), wheat (Devos *et al.*, 1995; Roder *et al.*, 1995), rice (Chen *et al.*, 1997), barley (Liu *et al.*, 1996), beans (Yu *et al.*, 1999), cowpea (Li *et al.*, 2001), soybean (Akkaya *et al.*, 1992), tomato (Smulders *et al.*, 1997), and grapevines (Thomas and Scott, 1993).

In sorghum, microsatellites were used to study genetic diversity (Ghebru *et al.*, 2002; Dje *et al.*, 1999, 2000; Grenier *et al.*, 2000; Smith *et al.*, 2000; Dean *et al.*, 1999; Brown *et al.*, 1996). Results from these studies have suggested that the microsatellite markers are suitable for applications relevant to conservation and use of sorghum germplasm.

2.4 Genetic distance analysis

Analyses of the extent and distribution of genetic variation in a crop are essential in understanding the evolutionary relationships between accessions and to sample genetic resources in a more systematic fashion for breeding and conservation purposes (Ejeta *et al.*, 1999). Menkir *et al.* (1997) suggested that molecular markers, in particular genetic distance estimates determined by molecular markers, are suitable to assess genetic diversity and to identify diverse sources in crop germplasm collections. Genetic distance is the extent of gene differences between cultivars, as measured by alleles frequencies at a sample of loci (Nei, 1987). Genetic similarity is the converse of genetic distances, i.e., the extent of gene similarities among cultivars. The measure of distance or similarity among cultivars is the covariance of allele frequencies summed for all characters (Smith, 1984).

Several genetic distance measures have been used to quantify genetic relationships among cultivars or germplasm accessions. Each variable of molecular bands such DNA-based marker bands are considered a locus so that every locus has two alleles. Banding profiles of each accession or cultivar can be scored as present (1) or absent (0). Generally, two approaches are used to deduce phylogenetic relationships from fingerprinting data. The first widely used approach involves the cluster analysis of pairwise genetic distances for the construction of dendrograms. Pairwise genetic distances are calculated directly from input data containing presence (1) or absence (0) values for all markers. One of the most commonly used genetic distance formulae is the Euclidean distance, which is the square root of the sum of squares of the distances between the multidimensional space values of the distances for any two cultivars (Kaufman and Rouseeuw, 1990) and it can be put as,

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

where, GD is the genetic distance between individual X and individual Y: i = 1 to N; N is the total number of bands, and X_i and Y_i are i^{th} band scores (1 or 0) for individual Xs and Ys. The process is repeated for all the possible pairwise groupings of individuals and the pairwise distance values tabled in a pairwise distance matrix. Genetic distance has also been calculated from several genetic similarity indices (GS) that can be calculated using either: D = 1-S or D = -ln (S). One useful similarity index is that of Nei and Li (1979): GD = 1-[$2N_{xy}/N_x+N_y$], here $2N_{xy}$ is the number of shared bands, and the N_x and N_y are the number of bands observed in individual x and individual y, respectively. Other similarity indices such as Jaccard's (Rohlf, 1993) and Gower's similarity coefficients (Gower, 1971) have been extensively used in genetic distance determination (Barrett and Kidwell, 1998).

The pattern of genetic relationship among accessions can be conveniently shown by multivariate techniques such as cluster or ordination analyses. Clustering is a useful tool for studying the relationships among closely

related cultivars or accessions. In cluster analysis, cultivars or accessions are arranged in hierarchy by agglomerative algorithm according to the structure of a complex pairwise genetic proximity measure. The hierarchies emerging from the cluster analysis are highly dependent on the proximity measures and clustering algorithm used (Kaufman and Rousseeuw, 1998).

In ordination analysis, the multidimensional variability in a pairwise, inter marker proximity is depicted in one to several dimensions through eigen structure analysis. Ordination is best suited to revealing interactions and associations among cultivars or accessions, which are described by continuous quantitative data (Bretting and Widrlechner, 1995). Principal component, principal coordinate, and linear discriminate analyses are the ordination techniques most commonly used in genetic relationships and cultivar classification studies (Schut *et al.,* 1997). Generally, statistical methods such as univariate, bivariate and multivariate analysis can be applied to analyse the data generated from germplasm accessions.

2.5 Comparison of major marker systems

In general, when morpho-agronomic and genetic marker data are available on a set of genotypes for studying their diversity and the formation of homogeneous groups, two types of hierarchical classifications are independently performed (Franco *et al.*, 2001). One is obtained based on the morpho-agronomic traits in which a standard metric distance (such as Squared Euclidean) is applied. The other classification is obtained based on the genetic marker attributes when genetic similarities (or dissimilarities) of n individuals are determined with molecular markers such as RFLPs, AFLPs, or SSRs. Using each fragment as an attribute (with values of 0 and 1 denoting the presence or absence of the fragment in that genotype, respectively), and applying any clustering strategy (such as single or complete linkage, UPGMA, the centroid method, etc.), genotypes can be clustered into groups that are as homogeneous as possible and heterogeneous among groups. Earlier findings have showed that groups

formed based on both continuous and categorical classifications had a low to medium consensus (Franco *et al.*, 2001).

Furthermore, availability of a large number of molecular markers necessitates a comparison of one marker technique with other commonly used markers. The range of DNA polymorphism assays for genome fingerprinting, investigating genetic relatedness, for genetic mapping and marker assisted plant breeding have expanded with the dramatic advances in molecular genetics (Karp *et al.,* 1997). These techniques include RFLP (Botstein *et al.*, 1980), RAPD (Welsh and McClelland, 1990; Williams *et al.*, 1990), AFLP (Zabeau and Vos, 1993) and SSR (Tautz, 1989, Weber and May, 1989). These methods detect polymorphism by assaying subsets of the total amount of DNA sequence variation in a genome. Polymorphisms detected with AFLP and RFLP assays reflect restriction size variation. RAPD polymorphisms result from DNA sequence variation at primer binding sites and from DNA length differences between primer binding sites (Williams *et al.*, 1993). This is also true for AFLPs. SSR loci differ in the number of repetitive di-, tri- or tetranucleotide units present (Tautz and Renz, 1984), and this length variation is detected with the PCR by utilizing pairs of primers flanking each simple sequence repeat.

Comparison of different classes of molecular markers was conducted in several crops including soybean, barley and wheat. In soybean, estimates were made for the information content (expected heterozygosity), the multiplex ratio (number of loci simultaneously analysed per experiment) and the effectiveness in assessing relationships between genotypes. Estimates of a single parameter, the marker index (product of expected heterozygosity and multiplex ratio), were also obtained to evaluate the overall utility of the marker. The use of this approach showed that SSR markers have the highest expected heterozygosity while the AFLP markers have the highest multiplex ratio and highest marker index. The utility of the recently developed AFLP markers has widely been reported in the literature. Since recently, SSR markers are also considered as markers of choice in plant species, including sorghum (Ghebru *et al.*, 2002; Smith *et*

al., 2000; Dje *et al.*, 2000; Brown *et al.*, 1996), maize (Pejic *et al.*, 1998, Senior *et al.*, 1998), wheat (Ahmad, 2002), soyabean (Akkaya *et al.*, 1992) and cowpea (Li *et al.*, 2001).

The choice of an appropriate DNA profiling technique is dependent on the aims of the testing. To facilitate selection of an appropriate technique for a given application, Powell *et al.* (1996) have utilized two metrices to compare different marker systems. The first metric was a good measure of information content/expected heterozygosity. Expected heterozygosity corresponds to the probability that two alleles taken at random from a population can be distinguished using the marker in question. SSR markers are known to have the highest expected heterozygosity (Pejic *et al.*, 1998). The second metric was the multiplex ratio of a marker system, which defines the number of loci (or bands) simultaneously analysed per experiment, for example in a single gel lane. Both AFLPs and RAPDs generally have higher multiplex ratios than RFLPs and SSRs. The practical considerations are that the test must also be inexpensive, technically straightforward, reliable, reproducible, and capable of unambiguous analysis. The cost of developing and conducting of the test must also be justified by the economic importance of the species or variety.

2.6 Food quality characteristics

Sorghum product quality is supposed to be determined by an important role played by two grain characteristics, endosperm texture and endosperm type (Pushpamma and Vogel, 1982). Endosperm type refers to either a horny or floury endosperm (Dewar *et al.*, 1993), while endosperm texture is the proportion of horny (hard) to floury (soft) endosperm (Cagampang and Kirleis, 1984).

Some consumers do not positively accept the visual appearance, mouth-feel and flavour of sorghum foods. The dark colour, pronounced flavour, grittiness of the flour, tannin content and palatability are some of the negative aspects associated with sorghum products (Sooliman, 1993). The

grittiness in mouth feel is caused by a high horny endosperm content. The starch in the horny endosperm, with high protein content, swells less tightly bound starch. Less swelling causes an underdeveloped jelly layer covering the particles, with a consequential harder and grittier mouth-feel (Novellie, 1982).

2.6.1 Physical properties and chemical composition

2.6.1.1 Physical properties

The major components of the seed are the pericarp or outer cover, the endosperm or storage organ, and the embryo or germ, which germinates to reproduce a plant. The endosperm forms the bulk of the kernel, generally being corneous on the outer extremes and floury toward the centre. Starch granules in the corneous outer portion are embedded in a protein matrix and are difficult to separate. Protein content in the floury endosperm is less than in the corneous types, and there can be voids in the structure contributing to a more opaque appearance of this portion of the endosperm. Starch is more easily recovered from the floury endosperm (Rooney and Miller, 1982). The embryo appears at the lower portion on one side of the seed. Most of the oil content of the seed is in the embryo.

2.6.1.2 Chemical composition

2.6.1.2.1 Protein content

The protein content of sorghum is an important quality-attribute in terms of consumer acceptability (Pushpamma and Vogel, 1982), and nutrition (Serna-Saldivar and Rooney, 1995).

From a nutritional view, sorghum is mainly utilised in developing countries where cereals are a staple food. This might cause nutritional problems, since sorghum and most other grains, when tested for albumin, glutelin and globulin proteins, are deficient in essential amino acids, especially lysine.

The breeding of high lysine sorghum varieties involves an increase in the levels of these three proteins, causing these varieties to contain approximately 50% more lysine and better amino acid profiles than regular varieties (Serna-Saldivar and Rooney, 1995).

The protein content is usually the most variable (Dendy, 1995). The average protein content of sorghum is 11 to 12%. In his review Lásztity (1996) reported that the protein content varies from 6 to 25%. The protein content and composition varies due to genotype, water availability, soil fertility, temperatures, and environmental conditions during grain development. Approximately 80, 16, and 3% of the sorghum protein is located in the endosperm, germ, and pericarp, respectively (Taylor and Schussler, 1986). Nitrogen fertilization significantly increases amounts of protein due to accumulation of prolamins (Warsi and Wright, 1973). The albumin-globulin and glutelin fractions are rich in lysine and other essential amino acids. Cultivars with improved protein quality usually contain higher amounts of albumins, glutelins, and globulins and correspondingly lower proportions of prolamins. The cooking process of sorghum–flour could decrease the protein content. Raw sorghum flour was found to contain 10.4% protein, while boiled and roasted flour contain 9.2 and 9.5% protein, respectively (Singh and Singh, 1991).

2.6.1.2.2 Lipid content

Lipids, which are minor components in cereal grains, are found primarily in the germ of sorghum. The whole grain consists of three types of lipids. The most abundant group, the nonpolar lipids, consists mainly of triglycerides, which serve as reserve nutrients during germination. The other two smaller groups, i.e. the polar (for example phospholipids, glycolipids) and unsaponifiable lipids (for example phytosterols, carotenoids and tocopherols) have other important biochemical fractions to fulfil (Serna-Saldivar & Rooney, 1995). The lipid content of sorghum ranges from 2.1 to 5.0% (Hoseney, 1994). Different sorghum cultivars show some variation, but these variations are not as extreme as in the case of other chemical

and physical properties of sorghum. Beta *et al.* (1995) found that 16 different sorghum cultivars had an average fat content of 3.7 ± 0.6%, while Yang and Seib (1995) found a fat content ranging from 3.2 to 4.1% for nine sorghum samples. Thus, lipid contents are significantly reduced when kernels are decorticated and/or de-germed.

Lipid content has been found to be positively correlated with protein content, so both traits can be selected for simultaneously (House et al., 1995). Milling plays an important role in the final lipid content of sorghum meal, because of the large part of the lipid fraction situated in the sorghum germ. Lipid content could also be used as a means of quality control of meal, to indicate whether proper separation of kernel parts took place during milling.

The fatty acid composition of sorghum oil is similar to that of maize and pearl millet, which contain higher levels of C18:1 fatty acids than for example barley and wheat (Hoseney, 1994). Fatty acid composition of sorghum oil is also similar to that of maize oil, with high concentrations of linoleic (49%), oleic (31%), and palmitic acids (14%). In addition, the oil contains 2.7% linolenic, 2.1% stearic acid, and 0.2% arachidic acid (Rooney, 1978).

2.6.1.2.3 Carbohydrate content

The carbohydrates of sorghum are composed of starch, soluble sugar, pentosans, cellulose, and hemicellulose. The quality and quantity of carbohydrates present in sorghum are important quality characteristics of sorghum and could influence consumer acceptance of the end product (Pushpamma and Vogel, 1982). Starch is the most abundant chemical component, while soluble sugars and crude fibre are low. In this study, only total starch content was examined.

The primary carbohydrate, starch, is the most abundant chemical component and makes up about 60 to 80% of the normal, non-waxy, kernels. The soluble sugars and crude fibre contents are low. This leads to the major role that starch properties play in the textural properties of cooked sorghum products (Cagampang and Kirleis, 1985), as well as the provision of fermentable sugars for beer brewing during malting (Taylor and Dewar, 1996). Sorghum starches have off-colours that are dependent on the pericarp colour and the presence of a black pigment in the glumes or other portions of the plant. Those containing black pigments have pinkish colours and those lacking this pigment have yellowish off-colours, while colour intensity is influenced by the pericarp colour (Watson and Hirata, 1955). The starch content of different sorghum cultivars shows wide variation. Buffo *et al.* (1997) found a starch content of 72.1% in the Dekalb hybrid, while Wankhede *et al.* (1989) found the CSH-1 hybrid to contain 64.5% starch. Klopfenstein and Hoseney (1995) also reported that starch makes up to 60 to 80% of normal (non-waxy) kernels.

Starches exist in highly organized granules in which amylose and amylopectin molecules are held together by hydrogen bonding. The amylose component plays a significant role in the rheological and shelf life properties of sorghum foods such as porridge, *tortillas* and *injera* (Ring *et al.*, 1982) and significantly correlated to the vitreousness of sorghum (Cagampang and Kirleis, 1984). Starch can be classified as waxy or non-waxy according to the amylose content. Waxy varieties contain up to 5% amylose, while non-waxy varieties were found to contain amylose levels from 24 to 30% (Ring *et al.*, 1982). A third group, the heterowaxy type has a lower amylose content than non-waxy starches. The waxiness of sorghum starch influences its rheological properties. Many of the properties of cereal starches that determine their suitability for particular end-uses are dependent upon their amylose / amylopectin ratios (Gibson *et al.*, 1997). These include gelatinisation and gelation characteristics, solubility, and the formation of resistant starch.

On the basis of chemical composition sorghum endosperm is classified as waxy (100% amylopectin in starch), normal (75% ampylopectin and 25% amylose), high lysine, sugary or yellow. Regular endosperm sorghum types contain from 23 to 30% amylose (Ring *et al.*, 1982). Waxy varieties contain up to 5% amylose.

The starch content and composition of sorghum are influenced by several factors. Firstly, the type of endosperm from which starch was extracted plays a significant role. Starch from the corneous endosperm of sorghum, exhibits a lower amylose content and higher gelatinisation temperatures, as well as a higher intrinsic viscosity than that from the floury endosperm (Cagampang and Kirleis, 1985). Environmental and genetic factors determine amylose levels in sorghum (Ring *et al.*, 1982), as was demonstrated with sorghum grown under supplementary irrigation and rainfed conditions. The starch of irrigated sorghum was shown to have significantly higher amylose contents than the rainfed ones (Taylor *et al.*, 1997). Environmental factors may affect the amylose content of starch more than genetic differences in the case of non-waxy varieties (Ring *et al.*, 1982).

2.6.1.2.4 Polyphenol / Tannins

All sorghums contain phenolic compounds, including phenolic acids and flavonoids (Klopfenstein and Hoseney, 1995). Some sorghum cultivars with a pigmented testa (B_1- B_2- genes) produce condensed polyphenols known as tannins (Butler, 1990). The compounds can affect colour, flavour, and nutritional quality of the grain and products prepared from it (Hahn *et al.*, 1984). Desirable agronomic characters of high-tannin sorghums are that they protect the grain against insects, birds, and weathering (Waniska *et al.*, 1989). The agronomic advantages are accompanied by nutritional disadvantages and reduced food qualities.

The polyphonols (condensed tannins) are mainly situated in the pericarp and/ or testa of pigmented sorghum varieties (Deshpande *et al.*, 1982).

These compounds in sorghum grain can be characterised using several techniques (Dendy, 1995). Most tannin assays measure the level of phenols, which may or may not be condensed tannins. The absolute amount of tannin present in the sorghum kernel is virtually impossible to determine, because a significant proportion cannot be extracted and assayed (Butler, 1990; Hahn *et al.*, 1984; Hahn and Rooney, 1986), and a suitable standard for sorghum tannins is unavailable. Different assays are likely to yield different tannin values because they respond to different chemical parts of the tannin molecule (Hagerman and Butler, 1981). Hahn *et al.* (1984) indicated that a typical brown sorghum contains the highest amount of free phenolic acids. Waniska *et al.* (1989) also partitioned sorghum phenolic acids and concluded that white cultivars without pigmented testa contained the lowest amounts of phenolic acids.

Tannins offer advantages of supplying bird-proof sorghum varieties with bird and mould-resistance (Serna-Saldivar and Rooney, 1995; Menkir *et al.*, 1996). High tannin sorghums are cultivated in a number of places in the world including Ethiopia where birds are a serious problem. On the negative side, tannins are anti-nutritional factors, as they bind proteins with consequent precipitation, which causes a lower nutritional value. Tannins also cause astringent tastes (Beta, 1998). The second problem lies in the colour acceptance of sorghum products. White sorghum produces the most acceptable products to consumers in terms of colour (Serna-Saldivar and Rooney, 1995), which is a further shortcoming of bird-resistant varieties.

2.6.2 Food quality / Sensory evaluation

Traditionally, *injera* is the staple diet of Ethiopia and is prepared from either tef [*Eragrostis tef* (Zucc.) Trotter], a cereal unique to Ethiopia or sorghum, depending on regional preference and/or income level. Sorghum is the first choice in eastern regions of Ethiopia for *injera* making and also in other regions of the country. Sorghum *injera* quality depends mainly on cultivar and fermentation process. Sorghum cultivars with white or yellow grain colour are preferred, but local landrace sorghums with soft endosperm and

a red or a brown colour also make acceptable *injera*. In the past, studies on the performance of some Ethiopian sorghum cultivars for *injera* making indicated a remarkable variability among the cultivars (Wuhib and Tekabe, 1987; Gebrekidan and Gebrehiwot, 1982). However, information on the relationships between the physical and chemical characteristics of the grain and *injera* quality (end-use) is required for the efficient utilization and conservation of the genetic resources.

The traditional art of *injera* preparation is not standardised, and there are minor differences in details of recipes between households, communities, and regions due to individual preferences and needs. Sorghum flour is sometimes mixed with that of *tef* or barley to improve acceptance and storability. Traditional preparation of injera involves fermentation of the flour with a starter from a previous batch. The flour is fermented for about 48 hr. A fresh batch of flour is gelatinised by adding boiling water, and the gelatinised flour is added to the fermented batter to hasten a secondary fermentation. Alternatively, a small part of the fermenting flour is prepared into a gruel (*absit*) and added to the fermenting batter a day before *injera* preparation. The fermented batter is baked on a hot clay griddle called *mitad/ qibaba*. To bake *injera*, 0.5 l of batter is poured on the hot *mitad/ qibaba* in centrifugal motion from the edge of the griddle to the center. When small holes (eyes) start appearing on the top of *injera*, it is covered with the griddle lid and left to bake for 2-3 min. A typical *injera* is a circular pancake of about 60-cm diameter and 6-mm thick. The front side has uniformly spaced honeycomb like "eyes", each measuring about 4 mm in diameter and about four per square centimetre of *injera* surface. The underside is smooth, without any holes, and is non-sticky. The colour is white to brown, depending on the grain colour of the cultivar used for flour.

A good quality *injera* has a soft and pliable texture, enabling the consumer to pick up *wat* (stew) by hand in a piece of *injera*. A good *injera* tastes slightly sour, is spongy, can be folded or rolled without cracking, and keeps its pliable texture for three days. On the contrary, poor *injera* is dry and brittle, does not exhibit uniform "eyes," and cannot be stored overnight.

Insufficient fermentation produces a sweetish taste, and such *injera* is called *aflegna injera* or *bidena benni*. The quality of *injera/ bidena* is influenced to a large extent by the fermentation process and the length of fermentation time. In general, decorticated sorghum grain produces *injera* of better quality and a lighter colour. Soft endosperm sorghum types produce better quality *injera*, but such grains cannot be decorticated well. Earlier workers indicated that more information is required on the type of sorghums suited for *injera* preparation (Gebrekidan and Gebrehiwot, 1982; Klopfenstein and Hoseney, 1995). The colour of sorghum grain and flour plays an important role in its acceptance. In general, white sorghums produce the most acceptable food products (Serna-Saldivar and Rooney, 1995). However, in some countries, brown sorghum products are preferred, e.g., for opaque sorghum beer. Food colour is the result of factors such as grain colour and type (pericarp colour, pigmented testa, endosperm colour), degree of milling, and pH of the food system.

REFERENCES

Abebe, B. and Wech, H.B. 1982. The 1981 activities of the Plant Genetic Resources Centre/Ethiopia. In: Proceedings of the Regional Workshop on Sorghum Improvement in Eastern Africa, 17-22 October, Addis Ababa, Ethiopia, pp. 31-45.

Aggarwal, R.K., Brar, D.S., Nandi, S., Huang, N. and Khush, G.S. 1999. Phylogenetic relationships among *Oryza* species revealed by AFLP markers. Theor. Appl. Genet. 98:1320-1328.

Ahmad, M. 2002. Assessment of genomic diversity among wheat genotypes as determined by simple sequence repeats. Genome 45:646-651.

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

Aldrich, P.R. and Doebley, J. 1992. Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild *Sorghum bicolor*. Theor. Appl. Genet. 85:293-302.

Assefa, K., Ketema, S., Tefera, H., Nguyen, H.T., Blum, A., Ayele, M., Bai, G., Simane, B. and Kefyalew, T. 1999. Diversity among germplasm lines of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. Euphytica 106: 87-97.

Ayana, A. 2001. Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. Ph.D. dissertation. Department of Biology, Science Faculty, Addis Ababa University, Addis Ababa, Ethiopia.

Ayana, A. and Bekele, E. 1998. Geographical patterns of morphological variation in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from Ethiopia and Eritrea: qualitative characters. Hereditas 129:195-205.

Ayana, A. and Bekele, E. 1999. Multivariate analysis of morphological variation in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from Ethiopia and Eritrea. Genetic Resources and Crop Evolution 46: 273-284.

Ayana, A., Bryngelsson, T. and Bekele, E. 2000. Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). Genetic Resources and Crop Evolution 47(5): 471-482.

Bai, G., Ayele, M., Tefera, H., and Nguyen, H.T. 1999. Amplified fragment length polymorphism analysis of tef [*Eragorstis tef* (ZUCC.) Trotter]. Crop Sci. 39: 819-824.

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

Barrett, B.A., Kidwell, K.K. and Fox, P.N. 1998. Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. Crop Sci. 38: 1221-1278.

Beckman, J.S. and Soller, M. 1983. Restriction fragment length polymorphisms in genetic improvements: methodologies, mapping and costs. Theor. Appl. Genet. 67: 35-43.

Beeching, J.R., Marmey, P., Gavalda, M-C., Noirot, M., Hayson, H.R., Hughes, M.A. and Charrier, A. 1993. An assessment of genetic diversity within a collection of cassava (*Manihot esculenta* Crantz) germplasm using molecular markers. Annuals of Botany 72:515-520.

Bekele, E. 1984. Analysis of regional patterns of phenotypic diversity in the Ethiopian tetraploid and hexaploid wheats. Hereditas 100:131-154.

Beta, T. 1998. Polyphenols and tannins in sorghum. In: INSORMILL Workshop. Department of Food Science, University of Pretoria, South Africa.

Beta, T., Rooney, L.W. and Waniska, R.D. 1995. Malting characteristics of sorghum cultivars. Cereal Chem. 72:533-538.

Boivin, K., Deu, M., Rami, J-F., Trouche, G. and Hamon, P. 1999. Towards a saturated sorghum map using RFLP and AFLP markers. Theor. Appl. Genet. 98:320-328.

Botstein, D., White, R.L., Skolnick, M.H., Davis, R.W. 1980. Construction of a genetic map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

Bretting, P.k. and Widrlechner, M.P. 1995. Genetic markers and horticultural germplasm management. HortScience 37:1349-1356.

Breyne, P., Rombaut, D., Van Gysel, A., Van Montagu, M. and Gerats, M. 1999. AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. Mol. Gen. Genet. 261: 627-634.

Brooke, C. 1958. The durra complex in the central highlands of Ethiopia. Econ. Bot. 2:192-204.

Brown, S.M., Hopkins, M.S., Mitchell, S.E., Senior, M.L., Wang, T.Y., Duncan, R.R., Gonzalez-Candelas, F. and Kresovich, S. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Genet. 93:190-198.

Buffo, R.A., Weller, C.L. and Parkhurst, A.M. 1997. Optimisation of sulfur dioxide and lactic acid steeping concentrations for milling of grain sorghum. Transactions of the American Soc. of Agric. Engineers 40:1643-1648.

Butler, L.G. 1990. The nature and amelioration of the antinutritional effects of tannins in sorghum grain. In: Ejeta, G., Mertz, E.T., Rooney, L., Schaffert, R. and Yohe, J. (Eds.). Proc. Int. Conf. On Sorghum Nutritional Quality. Purdue University, W. Lafayette, Indiana, pp. 191-205.

Caetano-Anolles, G. and Gresshoof, P.M. 1994. DNA amplification fingerprinting of plant genomes. Methods in Molecular Cell Biology 5:62-70.

Caetano-Anollés, G.C., Bassam, B.J. and Gresshoff, P.M. 1991. DNA amplification fingerprinting: A strategy for genome analysis. Plant Molecular Biology Reporter 9:553-557.

Cagampang, G.B. and Kirleis, A.W. 1984. Relationship of sorghum grain hardness to selected physical and chemical measurements of grain quality. Cereal Chem. 61:100-105.

Cagampang, G.B. and Kirleis, A.W. 1985. Properties of Starches isolated from sorghum floury and corneous endosperm. Starch/Staerke 37:253-257.

Chen, X., Temnykh, S., Xu, Y., Cho, Y.G. and McCouch, S.R. 1997. Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). Theor. Appl. Genet. 95:553-567.

Chittenden, M.L., Schertz, K.F., Lin, Y-R., Wing, R.A. and paterson, A.H. 1994. A detailed RFLP map of *Sorghum bicolor* x *S. propinquum*, suitable for high density mapping, suggests ancestral duplication of sorghum chromosomes or chromosomal segments. Theor. Appl. Genet. 87:925-933.

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

Cui, Y.X., Xu, G.W., Magill, C.W. and Schertz, K.F. 1995. RFLP-based assay of *Sorghum bicolor* (L.) Moench genetic diversity. Theor. Appl Genet. 90:787-796.

Dahlberg, J.A., Zhang, X., Hart, G.E. and Mullet, J.E. 2002. Comparative assessment of variation among sorghum germplasm accessions using seed morphology and RAPD measurements. Crop Sci. 42:291-296.

de Wet, J.M.J., Harlan, J.R. and Price, E.G. 1976. Variability in *Sorghum bicolor*. In: Harlan, J.R., de Wet, J.M.J. and Stemler, A.B.L. (eds.). Origins of African Plant Domestication. Mouton, The Hague, pp. 453-463.

Dean, R.E., Dahlberg, J.A., Hopkins, M.S. Mitchell, S.E. and Kresovich, S. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39:1215-1221.

Demissie, A., Bjørnstad, Å. and Kleinhofs, A. 1998. Restriction fragment length polymorphisms in landrace barleys from Ethiopia in Relation to geographic, altitude, and agro-Ecological factors. Crop Sci. 38:237-243.

Dendy, D.A.V. 1995. Sorghum and Millets: Chemistry and Technology, American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA.

Deshpande, S.S., Sathe, S.K., Salunkhe, D.K. and Cornforth, D.P. 1982. Effects of dehulling on phytic acid, polyphenols, and enzyme inhibitors of dry beans (*Phaseoulus vulgaris* L.) flours. Journal of Food Sci. 47:1846-1849.

Deu, M., Gonzalez-De-Leon, D., Glaszmann, J.C., Degremont, I., Chantereau, J., Lanaud, C. and Hamon, P. 1994. RFLP diversity in cultivated sorghum in relation to racial differentiation. Theor. Appl. Genet. 88:838-844.

Devos, K.M., Bryan, G.J., Collins, A.J., Stephenson, P. and Gale, M.D. 1995. Application of two microsatellite sequences in wheat storage proteins as molecular markers. Theor. Appl. Genet. 90:247-252.

Dewar, J., Von Ascheraden, S.R.F. and Taylor, J.R.N. 1993. Analysis of hardness and other kernel characteristics of grain sorghum cultivars. CSIR Report of the Sorghum Board, Pretoria, South Africa.

Djè, Y., Forcioli, D., Ater, M., Lefebvre, C. and Vekemans, X. 1999. Assessing population genetic structure of sorghum landraces from north-western Morocco using allozyme and microsatellite markers. Theor. Appl. Genet. 99:57-163.

Djè, Y., Heuertz, M., Lefébvre, C. and Vekemans, X. 2000. Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theor. Appl. Genet. 100: 918-925.

Doggett, H. 1988. Sorghum. Longman Group U.K. Ltd. Essex, England.

Donini, P., Elias, M.L., Bougourd, S.M. and Koebner, R.M.D. 1997. AFLP fingerprinting reveals pattern differences between DNA extracted from different plant organs. Genome 40:521-526.

Dufour, P., Deu, M., Grivet, L., D'Hont, A., Paulet, F., Bouet, A., Lanuad, C., Glaszmann, J.C. and Hamon, P. 1997. Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid. Theor. Appl. Genet. 94:409-418.

Edwards, K., Johnstone, C. and Thompson, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucl. Acid. Res. 19:1349-1458.

Ejeta, G., Goldsbrough, P.B., Tuinstra, M.R., Grote, E.M., Menkir, A., Ibrahim, Y., Cisse, N., Weerasuriya, Y., Melakeberhan, A. and Shaner, C.A. 1999. Molecular marker applications in sorghum: Paper presented at Workshop on Application of Molecular Markers, Ibadan, Nigeria, 16-20 August.

Franco, J., Crossa, J., Ribaut, J.M., Betran, J., Warburton, M.L. and Khairallah, M. 2001. A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. Theor. Appl. Genet. 103:944-952.

Gebrekidan, B. and Menkir, A. 1979. Ethiopian Sorghum Improvement Project Progress Report No. 7. College of Agriculture, Addis Ababa University, Addis Ababa, Ethiopia.

Gebrekidan, B., and Gebrehiwot, B. 1982. Sorghum *injera* preparations and quality parameters. In: Rooney, L.W., and D.S. Murty, D.S. (Eds.). Proc. Int. symp. on sorghum grain quality. ICRISAT, Patancheru, A. P., India, pp.55-66.

Geleta, N. 1997. Variability and association of morpho-agronomic characters with reference to highland sorghum [*Sorghum bicolor* (L.) Moench] landraces of Hararghie, Eastern Ethiopia. Alemaya University of Agriculture, Alemaya, Ethiopia. (MSc. Thesis)

Ghebru, B., Schmidt, R.J., and Bennetzen, J.L. 2002. Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeats (SSR) markers. Theor. Appl. Genet. 105:229-236.

Gibson, T.S., Solah, V.A. and McCleary, B.V. 1997. A prcedure to measure amylose in cereal starches and flours with Concanavalin A. Journal of Cereal Science 25:111-119.

Gowda, P.S.B., Xu, G.W., Frederiksen, R.A. and Magill, C.W. 1995. DNA markers for downey mildew resistance genes in sorghum. Genome 38:823-826.

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27:857-871.

Grenier, C., Deu, M., Kresovich, S., Bramel-Cox, P.J. and Hamon, P. 2000. Assessment of Genetic Diversity in Three Subsets Constituted from ICRISAT Sorghum Collection Using Random vs Non-Random Sampling Procedures. Theor. Appl. Genet. 101:197-202.

Hagerman, A.E. and Butler, L.G. 1981. Choosing appropriate methods and standars for assaying tannin. J. Chem. Ecol. 15:1795-1810.

Hahn, D.H., Rooney, L.W. and Earp, C.F. 1984. Tannins and phenols of sorghum. Cereal Foods World 29:776-779.

Hahn, D.H. and Rooney, L.W. 1986. Effect of genotype on tannins and phenols of sorghum. Cereal Chem. 63:4-8.

Harlan, J.R. 1992. Crops and Man (second edition). American Society of Agronomy and Crop Science Society of America, Madison, WI.

Harlan, J.R. and de Wet, J.M.J. 1972. A simplified classification of cultivated sorghum. Crop Sci. 12:172-176.

Hill, M., Witsenboer, H., Zabeau, M., Vos, P., Kesseli, R. and Michelmore, R. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. Theor. Appl. Genet. 93:1202-1210.

House, L.R., Osmanzai, M., Gomez, M.I., Monyo, E.S. and Gupta, S.C. 1995. Agronomic principles. In: Dendy, D.A.V. (Ed.). Sorghum and millets: Chemistry and technology. American Association of Cereal Chemists, St. paul, MN, USA, pp. 27-67.

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

Hoseney, R.C. 1994. Principles of Cereal Science and Technology (Second edition). In: American Assoc. of Cereal Chemists, St. Paul, USA, pp.7-21, 82-98, and 130-135.

Hulbert, S.H., Richter, T.E., Axtell, J.D. and Bennetzen, J.L. 1990. Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proc. Natl. Acad. Sci. 87:4251-4255.

Incirli, A. and Akkaya, M.S. 2001. Assessment of genetic relationships in durum wheat cultivars using AFLP markers. Genetic Resources and Crop Evolution 48:233-238.

Jain, S.K. 1975. Population structure and the effects of breeding system. In: Frankel, O.H. and Hawkes, J.G. (Eds.). Crop genetic resources for today and tomorrow. Cambridge University Press, London, pp.15-36.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Van De Wiel, C., Bredemeijer, G., Vosman, B., Mathes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Mamiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vaazquez, A. and Karp, A. 1997. Reproducibilty

testing of RAPD, AFLP, and SSR markers in plant by a network of European laboratories. Molecular Breeding 3:381-390.

Karp, A., Kresovich, S., Bhat, K.V., Ayad, W.G. and Hodgkin, T. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy.

Karp, A., Seberg, O. and Buiatti, M. 1996. Molecular techniques in the assessment of botanical diversity. Ann. Bot. 78:143-149.

Kaufman, L. and Rousseeuw, P.J. 1990. Finding groups in data: An introduction to cluster analysis. A Wiley-Interscience Publication, NY.

Klopfenstein, C.F. and R.C. Hoseney. 1995. Nutritional Properties of Sorghum and Millets. In: Dendy, D.A.V. (Ed.). Sorghum and Millets: Chemistry and Technology. American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA, pp.125-184.

Kochert, G. 1994. RFLP technology. In: Phillips, R.L. and Vasil, I.K. (Eds). DNA-based markers in plants. Kluwer Academic Publishers, Netherlands, pp.9-35.

Koebner, R.M.D., Devos, K.M. and Gale, M.D. 1994. Advances in the application of genetic markers in plant breeding. Asian Seed, Chiang Mai, Thailand.

Lásztity, R. 1996. The chemistry of cereal proteins. Bocaraton, Fla.:CRC Press.

Li, C.D., Fatokun, C.A., Ubi, B., Singh, B.B. and Scoles, G.J. 2001. Determining genetic similarities and relationships among cowpea breeding lines and cultivars by microsatellite markers. Crop Sci. 41:189-197.

Lin, D., Hubbes, M. and Zsuffa, L. 1993. Differentiation of poplar and willow clones using RAPD fingerprints. Tree Physiology 14:1097-1105.

Lin, Y.-R., Shertz, K.F. and Paterson, A.H. 1995. Comparative analysis of QTL affecting plant height and maturity across the *Poaceae*, in reference to an interspecific sorghum population. Genetics 140: 391-411.

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

Liu, Z. and Furnier, G.R. 1993. Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and big tooth aspen. Theor. Appl. Genet. 87:97-105.

Liu, Z.W., Biyashev, R.M. and Saghai Maroof, M.A. 1996. Development of simple sequence repeat markers and their integration into a barley linkage map. Theor. Appl. Genet. 93: 869-876.

Lu, J., Knox, M.R., Ambrose, M.J., Brown, J.K.M. and Ellis, T.H.N. 1996. Comparative analysis of genetic diversity in pea assessed by RFLP and PCR-based methods. Theor. Appl. Genet. 93:1103-1111.

Maheswaran, M., Subudhi, P.K., Nandi, S., Xu, J.C., Parco, A., Yang, D.C., Huang, N. 1997. Polymorphism, distribution, and segregation of AFLP markers in a double haploid rice population. Theor. Appl. Genet. 94:39-45.

Maughan, P.J., Saghai-Maroof, M.A., Buss, G.R., Huestis, G.M. 1996. Amplified fragment length polymorphism (AFLP) in soyabean: species diversity, inheritance, and near-isogenic line analysis. Theor. Appl. Genet. 93:392-401.

Mazur, B.J. and Tingey, S.V. 1995. Genetic mapping and introgression of genes of agronomic importance. Current Opinion in Biotechnology 6:175-182.

Melake-Berhan, A., Hulbert, S.H., Butler, L.G. and Bennetzen, J.L. 1993. Structural and evolution of the genomes of *Sorghum bicolor* and *Zea mays*. Theor. Appl. Genet. 86:598-604.

Menkir, A., Ejeta, G., Butler, L. and Melakeberhan, A. 1996. Physical and chemical kernel properties associated with resistance to grain mold in sorghum. Cereal Chem. 73:613-617.

Menkir, A., Goldsbrough, P. and Ejeta, G. 1997. RAPD based assessment of genetic diversity in cultivated races of sorghum. Crop Sci. 37:564-569.

Miller, J.C. and Tanksley, S.D. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor. Appl. Genet. 80:437-448.

Morell, M.K., Peakall, R., Appels, R., Preston, L.R., and Lloyd, H.L. 1995. DNA profiling techniques for plant variety identification. Australian J. Exp. Agric. 35:807-819.

Negash, A., Tsegaye, A., Treuren, R.van, and Visser, B. 2002. AFLP analysis of *Enset* clonal diversity in south and southwestern Ethiopia for conservation. Crop Sci. 42:1105-1111.

Negassa, M. 1986. Estimates of phenotypic diversity and breeding potential of Ethiopian wheats. Hereditas 104: 41-48.

Nei, M. 1987. Molecular evolutionary genetics. Colombia University Press, New York.

Nei, M. and Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. (USA) 79:5269-5273.

Novellie, L. 1982. Fermented porridges. In: Rooney, L.W. and Murty, D.S. (Eds.). Int. Symp. on Sorghum Grain Quality. ICRISAT, Patancheru, India, pp. 121-128.

Nybom, H.Y., Rogstad, S.H. and Schaal, B.A. 1990. Genetic variation detected by use of the M13 'DNA fingerprint' probe in *Malus, Prunus* and *Rubus* (Rosacea). Theor. Appl. Genet. 79:153-156.

Pakniyat, H., Powell, W., Baird, E., Handley, L.L., Robinson, D., Scrimgeour, C.M., Nevo, E., Hackett, C.A., Caligra, P.D.S. and Forster, B.P. 1997. AFLP variation in wild barley with reference to salt tolerance and associated ecogeography. Genome 40:332-341.

Paterson, A.H., Tanksley, S.D. and Sorrells, M.E. 1991. DNA markers in plant improvement. Adv. Agron. 46:39-90.

Paterson, A.H., Lin, Y., Li, Z., Schertz, K.F., Doebley, J.F., Pinson, S.R.M., Liu, S., Stanzel, J.W. and Irvine, J.E. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269:1714-1718.

Paul, S., Wachira, F.N., Powell, W. and Waught, R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis* (L.) O. kuntze) revealed by AFLP markers. Theor. Appl. Genet. 94:255-263.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G. and Motto, M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor. Appl. Genet. 97: 1248-1255.

Pereira, M.G. and Lee, M. 1995. Identification of genomic regions affecting plant height in sorghum and maize. Theor. Appl. Genet. 90:380-388.

Pereira, M.G., Lee, M., Bramel-Cox, P., Woodman, W., Doebly, J. and Whitkus, R. 1994. Construction of a RFLP map in sorghum and comparative mapping in maize. Genome 37:236-243.

Perkin-Elmer. 1996. AFLP Plant Mapping Kit–the PCR marker of choice for plant mapping.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2:225-238.

Powell, W., Thomas, W.T.B., Baird, E., Lawrence, P., Booth, A., Harrower, B., McNicol, J.W. and Waugh, R. 1997. Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisims. Heredity 79: 48-59.

Prasada Rao, K.E. and Ramanatha Rao, V. 1995. Use of characterization data in developing a core collection of sorghum. In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (Eds.). Core collection of plant genetic resources. IBPGRI, Sayce Publishing, and John Wiley and Sons, Chichester, pp. 109-115.

Pushpamma, P. & Vogel, S.M. 1982. Consumer acceptance of sorghum and sorghum products. In: Rooney, L.W. and Murty, D.S. (Eds.). International Symposium on Sorghum Grain Quality. ICRISAT, Patancheru, India, pp.341-353.

Qamaruz-Zaman, F., Fay, M.F., Parker, J.S. and Chase, M.W. 1997. The use of AFLP fingerprinting in conservation genetics: a case study of *Orchis simia* (*Orchidaceae*). Preliminary report. Royal Botanical Gardens, Kew.

Ragab, R.A., Dronavalli, S., Saghai Maroof, M.A. and Yu, Y.G. 1994. Construction of sorghum RFLP linkage map using sorghum and maize DNA probes. Genome 37:590-594.

Ring, S.H., Akingbala, J.O. and Rooney, L.W. 1982. Variation in amylose content among sorghums. In: Rooney, L.W. and Murty, D.S. (Eds.). International Symposium on Sorghum Grain Quality. ICRISAT, Patencheru, India, pp.269-279.

Roder, M.S., Plaschke, J., Konig, S.U., Borner, A., Sorrells, M.E., Tanksley, S.D. and Ganal, M.W. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. Mol. Gen. Genet. 246:327-333.

Rohlf, F.J. 1993. NTSYS-pc numerical taxonomy and multivariate analysis system. Exerer Software, Setauket, NY.

Rolfs, A., Schuller, I., Finckh, U. and Weber-Rolfs, I. 1992. PCR: Clinical Diagnostics and Research. Springer-Verlag, New York.

Rooney, L.W. 1978. Sorghum and pearl millet lipids. Cereal Chem. 55:584-590.

Rooney, L.W. and Miller, F.R. 1982. Variation in the structure and kernel characteristics of sorghum. In: Rooney, L.W., and Murty, D.S. (Eds.). International symposium on sorghum grain quality. ICRISAT, Patancheru, pp. 143-162.

Ruano, G., Brash, D.E. and Kidd, K.K. 1991. The first few cycles. Amplifications: A Forum for PCR users 4:5-7.

Russel, J.R., Fuller, J.D., Macaulay, M., Hatz, B.G., Jahoor, A., Powell, W. and Waugh, R. 1997. Direct Comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95: 714-722.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, KB., Horn, G.T., Erlich, H.A. and Arnheim, N. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350-1354.

Schut, J.W., Qi, X. and Stam, P. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor. Appl. Genet. 95:1161-1168.

Senior, M.L., Murphy, J.P., Goodman, M.M. and Stuber, C.W. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci. 38:1088-1098.

Serna-Saldivar and Rooney, L.W. 1995. Structure and chemistry of sorghum and millets. In: Dendy, D.A.V. (Ed.). Sorghum and Millets Chemistry and Technology. American Assoc. of Cereal Chemists, St. Paul, USA, pp. 69-124.

Sharma, S.K., Knox, M.R. and Ellis, T.H.N. 1996. AFLP analysis of the diversity and phylogeny of Lens and its comparison with RAPD analysis. Theor. Appl. Genet. 93:751-758.

Singh, A., Negi, M.S., Rajagopal, J., Bhatia, S., Tomar, U.K., Srivastava, P.S. and Lakshmikumaran, M. 1999. Assessment of genetic diversity in *Azadirachta indica* using AFLP markers. Theor. Appl. Genet. 99:272-279.

Singh, U. and Singh, B. 1991. Functional properties of sorghum-peanut composite flour. Cereal Chem. 68:460-463.

Smith, J.S.C. 1984. Genetic variability within U.S. hybrid maize: multivariate analysis of isozyme data. Crop Sci. 24: 1041-1046

Smith, J.S.C. and Smith, O.S. 1992. Fingerprinting crop varieties. Adv. Agron. 47:85-141.

Smith, J.S.C., Chin, E.C.L., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovitch, S. and Ziegle, J. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPs and pedigree. Theor. Appl. Genet. 95:163-173.

Smith, J.S.C., Kresovich, S., Hopkins, M.S., Mitchell, S.E., Dean, R.E., Woodman, W.L., Lee, M. and Porter, K. 2000. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci. 40:226-232.

Smulders, M.J.M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P. and Vosman, B. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. Theor. Appl. Genet. 97:264-272.

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

Song, k.M., Osborn, T.C. and William, P.H. 1988. Brassica taxonomy based on nuclear restriction fragment polymorphisms (RFLPs). Theor. Appl. Genet. 75:784-794

Sooliman, A.O.G. 1993. Sorghum processing at the crossroads in South Africa. In: Taylor, J.R.N., Randall, P.G. and Viljoen, J.H. (Eds.). Cereal Science and Technology: Impact on a Changing Africa. The CSIR, Pretoria, pp.497-507.

Stemler, A.B.L., Harlan, J.R. and de Wet, J.M.J. 1977. The sorghums of Ethiopia. Econ. Bot. 31:446-460.

Subudhi, P.K. and Nguyen, H.T. 2000a. New Horizons in Biotechnology. In: Wayane Smith, C. (Ed.). Sorghum: Origin, History, Technology, and production. John Wily and Sons, Inc., pp.349-397.

Subudhi, P.K. and Nguyen, H.T. 2000b. Linkage group alignment of sorghum RFLP maps using a common RIL mapping population. Genome 43:240-249.

Tao, Y., Manners, J.M., Ludlow, M.M. and Henzell, R.G. 1993. DNA polymorphisms in grain sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Genet. 86:679-688.

Tao, Y.Z., Jordan, D.R., Henzell, R.G. and McIntyre, C.L. 1998. Construction of a genetic map in sorghum RIL population using probes from different sources and its comparison with other sorghum maps. Aust. J. Agric. Res. 49:729-736.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucl. Acids Res. 17:6463-6471.

Tautz, D. and Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucl. Acids Res. 12:4127-4138.

Taylor, J.R.N. and Dewar, J. 1996. Recent developments in sorghum malting quality assessment. In: Randall, P.C. (Ed.). Grain Quality Assessment. Paper presented at the ICC-SA Seminar. The Maize Board, Pretoria, pp. 65-69.

Taylor, J.R.N. and Schussler, L. 1986. The protein composition of the different anatomical parts of sorghum grain. J. Cereal Sci. 4:361-369.

Taylor, J.R.N., Dewar, J., Taylor, J. and Von Ascheraden, R.F. 1997. Factors affecting the porridge-making quality of South African Sorghums. Journal of the Science of Food and Agriculture 73:464-470.

Teshome, A., Baum, B.R., Fahrig, L., Torrance, J.K., Arnason, T.J. and Lambert, J.D. 1997. Sorghum [*Sorghum bicolor* (L.) Moench] landrace variation and classification in North Shewa and South Welo, Ethiopia. Euphytica 97:255-263.

Thomas, M.R. and Scott, N.S. 1993. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theor. Appl. Genet. 70:1-6.

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

Tuinstra, M.R., Ejeta, G. and Goldsbrough, P.B. 1998. Evaluation of near-isogenic sorghum lines contrasting for QTL markers associated with drought tolerance. Crop Sci. 38:835-842.

Tuinstra, M.R., Grote, E.M., Goldsbrough, P.B. and Ejeta, G. 1997. Genetic analysis of post-flowering drought tolerance and components of grain development in sorghum. Mol. Breed. 3:439-448.

Tuinstra, M.R., Grote, E.M., Goldsbrough, P.B., and Ejeta, G. 1996. Identification of quantitative trait loci associated with pre-flowering drought tolerance in sorghum. Crop Sci. 36:1337-1344.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. 23:4407-4414.

Waniska, R.D., Poe, J.H. and Bandyopadhhyay, R. 1989. Effects of growth conditions on grain molding and phenols in sorghum caryopsis. J. Cereal Sci. 10:217-225.

Wankhede, D.M., Deshpahde, H.W., Gunjal, B.B., Bhosale, M.M., Patil, H.B., Gahilod, A.T., Sawate, A.R. and Walde, S.G. 1989. Studies on physiochemical pasting characteristics and amyolytic susceptibility of starch from sorghum (*S. bicolor* (L.) Moench) grains. Starch/Staerke 41:123-127.

Warsi, A.S. and Wright, B.C. 1973. Effects of rates and methods of nitrogen application on the quality of sorghum grain. Indian J. Agric. Sci. 43:722-726.

Watson, S.A., and Hirata, Y. 1955. The wet milling properties of grain sorghum. Agron. J. 47:11-15.

Waugh, R. and Powell, W. 1992. Using RAPD markers for crop improvement. Trends Bio/technol. 10:186-191.

Weber, J. and May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J. Hum Genet. 44:388-396.

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

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

Williams, J.G.K., Rafalski, J.A and Tingey, S.V. 1993. Genetic analysis using RAPD markers. Meth. Enzymol. 218:704-780.

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

Wuhib, E. and Tekabe, F. 1987. Sorghum Utilization Project Final Report, Ethiopian Nutrition Institute, Addis Ababa, Ethiopia.

Xu, G.W., Magill, C.W., Schertz, K.F. and Hart, G.E. 1994. A RFLP linkage map of *Sorghum bicolour* (L.) Moench. Theor. Appl. Genet. 89:139-145.

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

Yang, P. and Seib, P.A. 1995. Low-input wet-milling of grain sorghum for readily accessible starch and animal feed. Cereal Chem. 72:498-503.

Yang, W., de Oliveira, A.C., Godwin, I., Schertz, K. and Bennetzen, J.L. 1996. Comparison of DNA marker technologies in characterizing plant genome diversity: variability in Chinese sorghums. Crop Sci. 36:1669-1676.

Young, N.D. 1994. Constructing a plant genetic linkage map with DNA markers. In: Ronald, L.P. and Indra, K.V. (Eds.). Kluwer Academic Publishers, Netherlands, pp.39-58.

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

Yu, K., Park, S.J. and Poysa, V. 1999. Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus and Vigna*). Genome 42:27-34.

Yu, K., Van Deynze, A. and Pauls, K.P. 1993. Random amplified polymorphic DNA (RAPD) analysis. In: Methods in plant molecular biology and biotechnology. pp. 287-301. CRC Press, Inc.

Zabeau, M. and Vos, P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European patent Application EP 534858A1.

CHAPTER 3

PHENOTYPIC DIVERSITY FOR MORPHO-AGRONOMICAL TRAITS IN SORGHUM

Abstract

Forty-five sorghum accessions growing in the eastern highlands of Ethiopia were evaluated for phenotypic diversity. Ten qualitative and 16 quantitative traits were recorded for all the accessions. Phenotypic frequencies between the accessions from each of the 10 *aanaas* / *woredas* and AU, grouped in 10 localities were tabulated. Phenotypic diversity index, H', was analysed and the result indicated the between localities component of diversity to be relatively smaller than the variation in H' among characters within localities. Multivariate methods, including clustering and principal component analyses were used on quantitative traits data to estimate the patterns and distribution of phenotypic variation. Cluster analysis grouped the accessions into seven cluster groups. Further among the 16 principal components involved, the first eight principal components explained 78% of the total variation between the accessions. Grain yield, panicle weight, kernel number per panicle and number of primary branches being the most important traits in the first principal component with 19% of the total variation, and leaf traits in the second with 17% total variation. The results both in qualitative and quantitative traits data showed that there is a wide morpho-agronomical diversity among the accessions studied. The sorghum improvement programme of AU can use this information for its future breeding strategy and the conservation of these resources.

3.1 Introduction

Information on the genetic diversity within and among closely related crop species is essential for rational use and management of genetic resources. It is particularly useful in characterising individual accessions and cultivars, in detecting genetic materials with novel genes and thereby rescuing them

from erosion, and as a general guide in selecting parents for crossing in breeding programmes. Most of the genetic diversity of food crops in Ethiopia is traditionally maintained by farmers *in situ* (Worede, 1988). Local farmers deliberately, and unconsciously too, grow several varietal forms together to add variety to their diet and also to reduce the risk of economic loss from new parasites or insect pests or in the event of unusual environmental conditions (Worede, 1988; Bekele, 1984, Brooke, 1958). Stemler *et al.* (1977) reported the presence of four of the primary races (except Kafir) in Ethiopia and described their distribution within the country. However, many valuable landraces of sorghum either have been lost or under serious risk due to various human and environmental factors (Teshome, 2001). Investigating the degree of character variation and distribution of sorghum in Ethiopia, where it was domesticated (Vavilov, 1951) and diversified (Kebede, 1991; Doggett, 1988; Harlan, 1969) is essential.

Categorizing germplasm accessions into morphologically similar, and presumably genetically similar, groups is most useful when the population structure in a collection is unknown (Marshal and Brown, 1975). Genetic relationships among a large number of accessions can be summarised using cluster analysis to place similar accessions into groups. Phenotypic diversity index of morphological characters and/or multivariate analysis of quantitative characters has been used previously to measure genetic relationships within cereal crop species. Examples include tef (Assefa *et al.*, 1999), barley (Demissie and Bjornstad, 1996; Negassa, 1985; Bekele, 1984; Tolbert *et al.*, 1979), tetraploid wheat (Tesfaye *et al.*, 1991; Bechere *et al.*, 1996) and Ethiopian wheats (Negassa, 1986).

In sorghum, morphological characters were used to estimate phenotypic variation among landrace accessions grown in north Shewa and south Welo regions of Ethiopia (Teshome *et al.*, 1997). Based on morphological variation assessment of 415 sorghum accessions collected from different regions of Ethiopia and Eritrea have shown significant levels of variation within the regions of origin and within the adaptation zones (Ayana and

Bekele, 1998, 1999). In a previous investigation, we analysed the level of morpho-agronomic traits variability in sorghum landraces from the eastern highland regions of Ethiopia based on quantitative traits data (Geleta, 1997). However, the extent and patterns of phenotypic variation that exist among and within sorghum accessions of the regions have not been assessed using Shannon-Weaver diversity index and multivariate analyses. In this study, both qualitative and quantitative traits were used to estimate the levels of variation among the sorghum accessions grown in the eastern highlands of Ethiopia. The main objectives of the study were to: (1) estimate the extent of genotypic diversity among sorghum accessions based on 10 qualitative and 16 quantitative traits data, and (2) assess the regional patterns of phenotypic diversity using qualitative and quantitative traits data.

3.2 Materials and methods

3.2.1 Qualitative traits

3.2.1.1 Plant materials

Forty-five accessions of sorghum were used and the information on the collection sites is given (Table 3.1). Among the 45 sorghum accessions 34 were collected from farmers' fields in 10 *Aanaas/ Woredas* of the highlands of eastern Ethiopia. Five each were elite breeding lines and improved cultivars (through selection), and one local variety acquired from the Sorghum Improvement Program (SIP), Alemaya University (AU). The accessions/cultivars were collected and chosen on representation basis, stratified systematic sampling method to a given range of geographic area, a range of morpho-agronomic traits and material under development, and potential interest in SIP.

3.2.1.2 Methods

The evaluation was conducted at the glasshouse of Plant Breeding, University of the Free State during 2001. Two plants of each accession were grown in an 8 ℓ pot, replicated three times. To categorise each accession morphologically, Sorghum Descriptors (IBPGR/ICRISAT, 1993) was employed. Table 3.2 lists the qualitative traits, their descriptors and the codes used in the analyses. Each accession was scored for the most frequent character-state. Seed colour, glume colour and leaf midrib colour were examined and scored using the Munsell colour firm (1990).

3.2.1.3 Data analysis

Phenotypic frequency distributions of the characters were worked out for all the accessions and *aanaas/ woredas*. The Shannon-Weaver diversity index (H′) was computed using the phenotypic frequencies to assess the phenotypic diversity for each character for all accessions. The Shannon-Weaver diversity index as described by Perry and McIntosh (1991) is given as:

$$H' = 1-\sum_{i=1}^{n} p_i \log_e p_i$$

where, p_i is the proportion of accessions in the i^{th} class of an n-class character and n is the number of phenotypic classes of traits. Each H′ value was divided by its maximum value ($\log_e n$) and normalised in order to keep the values between 0 and 1. By pooling various characters across the collection sites, the additive properties of H' were used to evaluate diversity of localities and characters within the population.

3.2.2 Quantitative traits

3.2.2.1 Location of the study

To study the quantitative traits variability the sorghum accessions were grown at the Department of Plant Sciences Experimental field, AU, Alemaya during the 2000 and 2001-growing seasons. Alemaya is located at 9°26´N latitude, 43°03´E longitude with an altitude of 1980 m above sea level. In Ethiopia, Alemaya is one of the recommended testing sites for the highland sorghum germplasm characterisation. The area receives an average annual rainfall of about 800 mm and has a temperature ranging between 10 to 24°C (Alemaya Weather Station). The accessions were grown in three rows of 5m long and 0.75 m between rows plot, with three replications in randomized complete block design. Within every row, plant-to-plant spacing was 0.20 m.

3.2.1.4 Plant materials

The 45 accessions listed in Table 3.1 were used for this study.

3.2.2.3 Parameters measured

Data were collected for 14 morpho-agronomic characters (Table 3.3). Two derived characters were also estimated for each accession from the above measurements (i) grain number per panicle, and (ii) threshing percentage.

For every accession, data was recorded from the middle row on five randomly selected individual plants, except for days to 50% flowering, which was recorded on plot basis (Table 3.3). For leaf characteristics measurement, a procedure developed by Stickler *et al.* (1961) was used.

Table 3.1 Local / cultivar name, collection site, altitude and sample status of sorghum samples used in the study.

Accession number.	Local / cultivar name	Collection site (*Aanaa/ Woreda*)[1]	Altitude (m)	Sample status[2]
1	Wagare	Chinhakssen	1950	LR
2	Wagare	Chinhakssen	1970	LR
3	Wagare	Haro Maya	2120	LR
4	Muyra adi	Haro Maya	2120	LR
5	Muyra	Kurfa Challe	2220	LR
6	Fandisha duudaa	Kurfa Challe	2310	LR
7	Fandisha faca'a	Bedeno	2050	LR
8	Wagare	Meta	2050	LR
9	Hamedaya	Meta	2180	LR
10	Muyra	Meta	2180	LR
11	Abedelota	Deder	2080	LR
12	Ambajeetee	Deder	2080	LR
13	Muyra	Tulo	2000	LR
14	Dassile	Tulo	1900	LR
15	Hanchiro	Tulo	2100	LR
16	Wagare	Tulo	1940	LR
17	Key Fendisha	Tulo	2140	LR
18	Sharif	Tulo	1960	LR
19	Alaa guuraacha	Tulo	1940	LR
20	Suuta naqaaphu	Tulo	1900	LR
21	Qirendaye	Tulo	2200	LR
22	Alegid	Tulo	1900	LR
23	Fandisha	Doba	1900	LR
24	Fandisha gababa	Doba	2230	LR
25	Bulo	Doba	2010	LR
26	Janga	Doba	2000	LR
27	Harka basi	Doba	2000	LR
28	Shafare	Doba	2000	LR
29	Shafare	Chiro	2270	LR
30	Gababe	Chiro	2240	LR
31	Warabi	Chiro	1900	LR
32	Zangada	Habro	1940	LR
33	Zangada	Habro	1940	LR
34	Dassile	Habro	1900	LR
35	ETS 721	AU	-	BE
36	ETS 993	AU	-	BE
37	ETS 789	AU	-	BE
38	ETS 804	AU	-	BE
39	Wotet begunche	AU	-	BE
40	AL-70	AU	-	IC
41	ETS 2752	AU	-	IC
42	Chiro	AU	-	IC
43	ETS 1005	AU	-	IC
44	ETS 576	AU	-	IC
45	Long muyra	AU	-	LC

[1] Administrative unit
[2] LR = Landrace, BE = Breeding entry, IC = Improved cultivar, LC = Local cultivar

Ch = Chinhakssen
Hm = Haro Maya
Kc = Kurfa Challe
Bn = Bedeno
Mt = Meta
Dd = Deder
Db = Doba
Tl = Tulo
Cr = Chiro
Hb = Habro

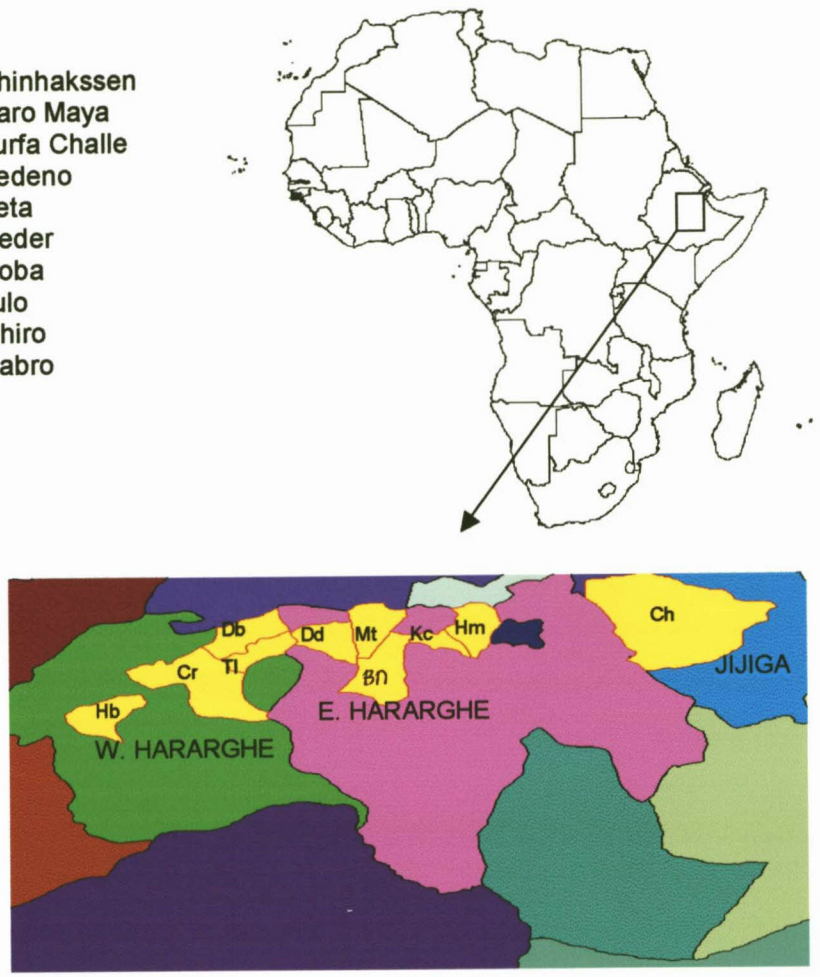


Figure 3.1 Geographic locations where the sorghum landrace accessions used in the study were collected. (The sites are located in Oromiya and Somali Regional States of Ethiopia.)

Table 3.2 Character, descriptor and codes used for characterisation of sorghum accessions.

Character	Descriptor & Code
Plant colour	Pigmented (1) and Tan (2)
Stalk juiciness	Dry (0) and Juicy (1)
Leaf midrib colour	White (1), Dull green (2) & Yellow (3)
Inflorescence exsertion	Slightly exserted (1), Exserted (2), Well-exserted (3), and Peduncle re-curved/ goose (4)
Panicle compactness & shape	Very lax panicle (1), Very loose erect primary branches (2), Loose erect primary branches (4), Semi-loose erect primary branches (6), Semi-compact elliptic (8), Compact elliptic (9), and Compact oval (10)
Awns (at maturity)	Absent (0) and Present (1)
Glume colour	White (1), Yellow (2), Grey-orange group (3), Orange-red (4), Purple (5), Black (6), and Grey (7)
Grain covering	25% grain covered (1), and 50% grain covered (3)
Grain colour	White (1), Yellow (2), Red (3), Light brown (4), Brown (5), Red brown (6), Dark brown (7), Grey (8), and Straw (9)
Endosperm texture	Completely corneous (1), Mostly corneous (3), Intermediate (5), Mostly starchy (7), and Completely starchy (9).

Table 3.3 List, code and descriptions of the quantitative characters recorded in the study.

Character	Code	Description
Days to 50% flowering (count)	DF	From emergence to when 50% of plants have started flowering.
Leaf number (count)	LN	Count of total number of leaves per plant (main stalk).
Leaf length (cm)	LL	Length of the third or fourth leaf from the flag leaf.
Leaf width (cm)	LW	Width of the third or fourth leaf from the flag leaf.
Leaf area (cm^2)	LA	Area of the third or fourth leaf from the flag leaf, computed as (LL x LW x 0.747) suggested by Stickler et al. (1961).
Internode length (cm)	INL	Length of the third internode counted from the ground surface.
Leaf sheath Length (cm)	LSL	Length measured on leaf sheath found on the third internode from the ground.
Plant height (cm)	PHt	Height of the main stalk from the ground to the tip of the panicle.
Panicle length (cm)	PL	Length of panicle from its base to tip.
Panicle width (cm)	PW	Width of panicle in natural position at the widest part.
Number of primary branches per panicle (count)	NPBP	Number of branches arising directly from the rachis of the panicle.
Head weight (g)	HWt	Weight of head (panicle) before threshing.
Grain yield per panicle (g)	GYPP	Weight of grain per panicle.
1000-seed weight (g)	TSWt	Weight of 1000 seed counts.
Threshing percent (%)	TP	The ratio of grain weight per panicle to the head weight of the same multiplied by 100.
Grain number per panicle (count)	GNPP	The ratio of grain weight per panicle to the average 1000-seed weight per panicle of the same multiplied by 1000.

3.2.2.4 Statistical analysis

Raw data was entered into a Microsoft Excel spreadsheet and mean values were calculated for each accession, and the mean values were imported into the Number Cruncher Statistical Systems (NCSS, 2000) program. The 16 quantitative traits data were used as columns and the 45 accessions as rows. Pair-wise genetic distance estimates were obtained between accessions and used to group them using UPGMA method. Results were obtained from UPGMA analysis as dendrogram.

Principal component analysis (PCA) was also carried out on the same data employing the multivariate analysis option of the NCSS program. PCA is a data analysis tool that is usually used to reduce the dimensionality (number of variables) of a large number of interrelated variables, while retaining as much of the information (variation) as possible. PCA calculates an uncorrelated set of variables (factors or pc's). These factors are ordered so that the first few retain most of the variation present in all of the original variables. The PCA was performed on the correlation matrix rather than covariance matrix.

3.3 Results and discussion

3.3.1 Qualitative traits

The amount of phenotypic diversity estimates based on Shannon-Weaver diversity index (H′) and its partitioning within and between collection sites are shown in Table 3.4. The 10 characters differed in their distribution as well as the amount of variation. The overall average phenotypic diversity (H′) among accessions was 0.71, varying from 0.36 (grain covering) to 0.95 (grain colour). Leaf midrib colour, stalk juiciness, awns, and grain covering were relatively monomorphic, while plant colour was intermediate, and inflorescence exsertion, panicle compactness and shape, glume colour, grain colour, and endosperm texture were highly polymorphic (Table 3.4).

The distribution of phenotypic frequencies for stalk juiciness and grain covering at collection site level showed only weak clinal variation among *aanaas/ woredas*. The majority of the accessions (42 out of 45) were non-juicy/dry.

The H' pooled across characters by region of collection ranged from 0.38 to 0.73 (Table 3.5). The *aanaas* that had the highest H' were Kurfa Challe, Chiro, and accessions from AU. The lowest values of H' are from Doba (0.38), Habro (0.45) and Meta (0.46). Here, low H' values were not necessarily associated with lack of adequate sample size. The importance of the sample size for phenotypic diversity suggests, therefore that the best way to sustain biodiversity in sorghum fields depends on the many options to be offered to the local farmers by national government and/or international organisations.

Table 3.4 Estimates of H', partitioning into within and between collection sites for 10 qualitative characters in 45 sorghum accessions.

Character	H'	H_{cl}	H_{cl}/H'	$(H'-H_{cl})/H'$
Leaf midrib colour (LMC)	0.55	0.36	0.65	0.35
Plant colour (PC)	0.73	0.38	0.52	0.48
Stalk juiciness (SJ)	0.37	0.10	0.27	0.73
Inflorescence exsertion (IE)	0.90	0.63	0.70	0.30
Panicle compactness & Shape (PCS)	0.90	0.89	0.99	0.01
Awns	0.57	0.34	0.60	0.40
Glume colour (GLC)	0.85	0.81	0.95	0.05
Grain covering (GC)	0.36	0.23	0.64	0.36
Grain colour (GRC)	0.95	0.93	0.98	0.02
Endosperm texture (ET)	0.90	0.85	0.94	0.06
Average	0.71	0.55	0.72	0.28

H' = Diversity index for each character calculated from entire data set; **H_{cl} =** Average diversity index of each character for the 10 localities; **H_{cl}/H' =** Proportion of diversity within localities; and **$(H'-H_{cl})/H'$** = Proportion of diversity between localities in relation to the total variation.

Table 3.5 Estimates of the Shannon-Weaver diversity index, H', for 10 qualitative characters in sorghum accessions by location of collection.

	Character										
Location	LMC‡	PC	SJ	PE	PCS	Awn	GLC	GC	GCl	ET	Mean±S.E.
Ch†	0.00	0.00	0.00	1.0	0.99	0.00	0.99	0.00	1.00	1.00	0.50±0.17
Hm	0.00	1.00	0.00	1.0	0.99	0.00	0.99	0.00	1.00	1.00	0.60±0.16
Kc	0.00	0.92	0.00	0.92	0.92	0.92	0.92	0.92	0.82	0.92	0.73±0.12
Mt	0.00	0.00	0.00	0.92	0.83	0.00	0.92	0.00	0.99	0.92	0.46±0.15
Dd	1.00	0.00	0.00	0.00	0.83	0.00	0.00	0.00	1.00	1.00	0.38±0.16
Tl	0.00	0.47	0.00	0.86	0.86	0.00	0.85	0.47	0.76	0.86	0.51±0.12
Db	0.99	0.00	0.00	0.65	0.83	0.65	0.90	0.00	0.96	0.92	0.59±0.13
Cr	0.92	0.00	0.00	0.00	0.92	0.92	0.82	0.92	0.99	0.92	0.64±0.14
Hr	0.00	0.92	0.00	0.00	0.92	0.92	0.82	0.00	0.92	0.00	0.45±0.15
AU	0.68	0.47	0.85	0.95	0.83	0.00	0.85	0.00	0.83	0.91	0.64±0.11

‡ Character abbreviations as defined in Table 3.4

† Location abbreviations as defined in Figure 3.1.

3.3.2 Quantitative traits

3.3.2.1 Clustering

The UPGMA dendrogram (Figure 3.2) shows the clustering of sorghum accessions based on quantitative data. First two major cluster groups were formed at a genetic distance of about 1.90: the first group included only three accessions (two from Chinhakssen and one Tulo) and the second contained all the other accessions. The second major group was further split into two subgroups at a genetic distance of 1.45. The first of this subgroup was further split into two sub-subgroups, and the first comprised

a single accession (#24), and the second consisted of all the rest. In total, seven clusters were formed, where two of the clusters (cluster V and VI) were made only of a single accession each. The constitution of each cluster is given in Table 3.6. The second cluster contained about 45% of the accessions, and a wide intra-cluster variation is observable.

Table 3.7 showed differences among clusters by summarising cluster means for 16 characters. Cluster I was characterised by lowest values in most of the variables except days to 50% flowering, leaf number, number of primary branches, and threshing percentage. Cluster II comprises the maximum number of accessions from various localities and it is characterised by low to medium mean values in all variables. Longest leaf sheath, tallest plant height, higher number of primary branches, highest head weight, and highest grain yield characterised cluster III. The days to 50% flowering in cluster III is medium. Cluster IV includes four of the five accessions known as *Fandisha* and was identified mainly by longest days to 50% flowering, lowest threshing percentage, widest leaf area and panicle. The grain yield was second to cluster II.

Cluster V consisted of only a single accession from Kurfa Challe aanaa, and it has recorded the longest leaf length, largest leaf number, broadest leaf, and heaviest 1000-seed weight. Cluster VI also consisted of a single accession from Doba, and it has recorded the shortest number of days to 50% flowering, lowest 1000 seed weight, highest threshing percentage and highest number of grains per panicle. The last cluster, cluster VII, is characterised by lowest number of leaves, longest internode length, longest panicle length, and fewest number of branches.

Accessions from the same source/ collection site were clustered together. Nevertheless, some exceptions were observed. For example, out of the three accessions from Habro, one (#34) was clustered in cluster II, while two (#32 and 33) clustered in cluster VII. Two previous studies reported using qualitative and quantitative data supports the absence of clear grouping of accessions based on geographical origin (Ayana and Bekele,

1999; Teshome *et al.*, 1997). There was evidence that both natural selection for adaptation to environment and farmers' selection for specific use or specific cropping systems accounted for most of the morpho-agronomic diversity (Grenier *et al.*, 2001).

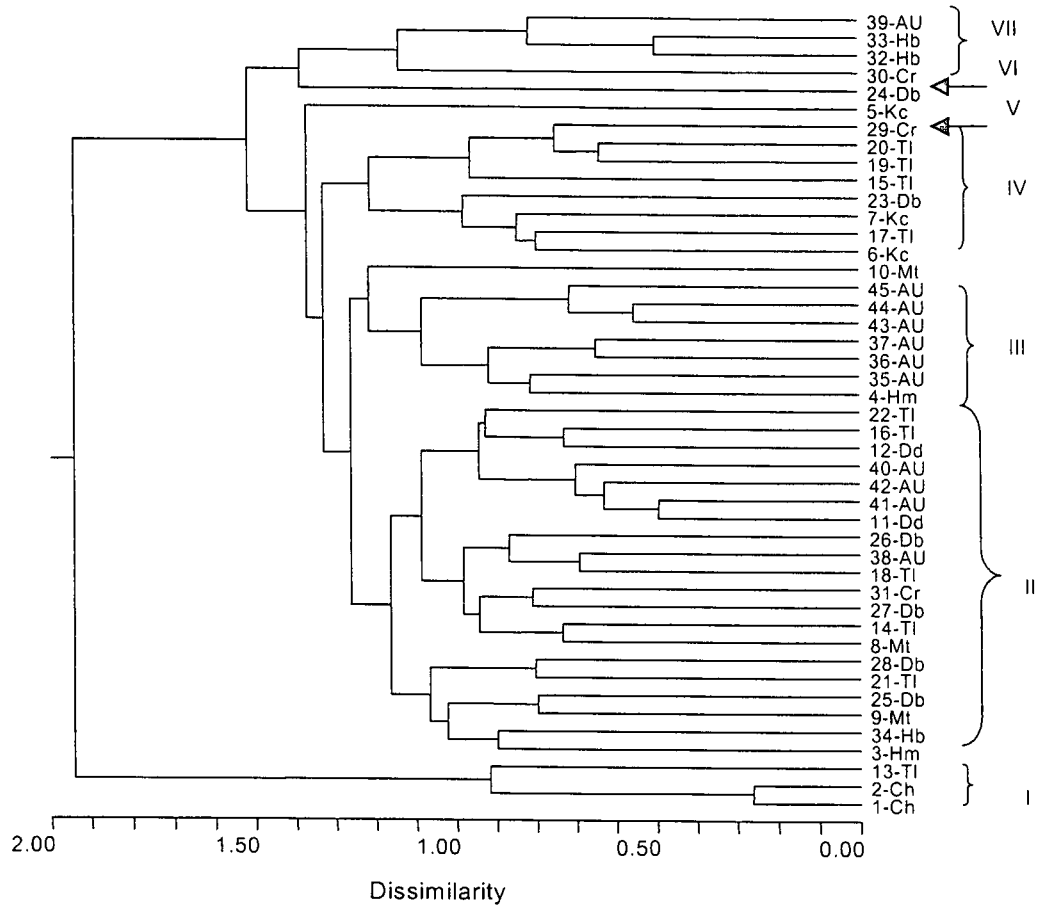


Figure 3.2 Dendrogram showing cluster groups among the 45 sorghum accessions based on 16 quantitative traits data.

Clustering from this study provides a structure for sampling accessions for further germplasm collection; for genetic, breeding, or agronomic studies where information on multiple quantitative traits is required. Cluster analysis based on quantitative traits also may aid parental selection by providing specific trait information that would allow the breeder to focus on sampling of accessions from specific locations for cultivar improvement.

Table 3.6 Distribution of the 45 sorghum accessions into seven clusters by location of origin using average values of quantitative characters.

Location	Cluster							Location
	I	II	III	IV	V	VI	VII	Total
Ch†	2							2
Hm		1	1					2
Kc				2	1			3
Mt		2	1					3
Dd		2						2
Tl	1	5		4				10
Db		4		1		1		6
Cr		1		1			1	3
Hb		1					2	3
AU		4	6				1	11
Total	3	20	8	8	1	1	4	45

† Location abbreviations as defined in Figure 3.1

3.3.2.2 Principal component analysis

The principal component analysis (PCA) grouped the 16 variables (Appendix III) into 16 components, which accounted for the entire (100%) variability evident among the test accessions. It also showed that the first 10 eigenvectors explained about 88% of the total variance. Of these, the first eight (as only the ones with eigenvalues over 1 accounted for a cumulative of 77.9% of the entire variability apparent among the accessions were shown in Table 3.9. Each trait was an important source of variation in at least one PC axis. Because each of the PC axes was given equal weight in the cluster analysis, each trait contributed to the information used to group accessions; however, some characters may have greater importance in determining plant phenotype than others.

Grain yield per panicle was significantly correlated ($P \leq 0.01$) to eight of the 15 characters (Table 3.8). Internode length, threshing percentage, panicle

width and panicle length didn't show significant correlation with grain yield and the other components. The highest degree of correlation of grain yield, head weight and grain number per panicle with other characters was supported by the PC analysis where these three are the primary source of variation with coefficient values of –0.34, -0.35 and –0.30, respectively (Table 3.8). In this study, the first two PC axes accounted for 36% of the multiple variations among genotypes (Table 3.9), indicating a high degree of correlation among characters for these accessions.

The existence of wider morpho-agronomical diversity among sorghum accessions studied was further substantiated by the PCA plot (Figure 3.3) analysed based on 16 quantitative traits. The first principal component alone explained 19% of the gross variability among the accessions. This has been due chiefly to variations in grain yield per panicle, head weight, grain number per panicle, and number of primary branches. Similarly, 17% of the overall variability of the accessions comes from the second principal component that originated primarily from variations in leaf characteristics (leaf length, leaf area, leaf width and leaf number).

The PC analysis didn't completely place the sorghum accessions into distinct groups. They remained scattered in all four quadrants, showing greater variability among them. Overlap between accessions with similar names was not seen. In some cases, the samples collected/obtained from similar locations fall close to each other (for example, accessions number 1 and 2, 32 and 33). Like in clustering, accession # 24 occupied an extreme of the first axis. Jain *et al.* (1975) indicated that variation either in terms of individual genotypic frequencies or overall measures of genetic or phenotypic diversity may not show obvious geographic patterns. Rather, the variation was higher within the same location than between locations. In a similar study, Ayana and Bekele (1999) reported that the within regions variation accounted for a large portion of the total variation compared with the between regions variation. The trend of higher diversity within than between regions also has been reported in barley germplasm (Kebebew *et al.*, 2001).

Table 3.7 Cluster means for 16 quantitative characters in 45 sorghum accessions.

Cluster	Character															
	DF†	LN	LL	LW	LA	INL	LSL	PHt	PL	PW	NPBP	HWt	GYPP	TSWt	TP	GNPP
C_I	92.67	9.57	59.3*	7.15*	316.73*	22.06*	16.77*	197.5*	11.2*	9.2*	69.6	88.4*	63.37*	27	71.83	2377.65*
C_{II}	96.75	11.74	73.04	9.53	521.31	25.05	21.39	287.55	19.45	10.32	88.75	151.13	107.14	27.05	71.07	3997.02
C_{III}	100.75	13.1	71.93	10.24	555.15	26.56	22.41**	320.38**	14.71	10.6	124.14**	178.46**	131.59**	29.71	73.67	4435.6
C_{IV}	106**	12.95	81.16	11.67**	707.43	26.09	21.83	282.47	18.59	11.05**	78.94	165.51	112.22	23.75	68.09*	4739.56
C_V	106	13.3**	89.2**	10.8	719.6**	24	19	251.3	13.6	10.6	116.3	123.66	87.56	31.7**	70.81	2762.15
C_{VI}	89*	10.3	66.4	9.85	488.6	22.67	19.2	230	17.5	10.2	69.95	131.63	102.04	18*	77.52**	5668.89**
C_{VII}	95	9.44*	72.93	8.80	478.58	27.23**	21.7	280.25	21.28**	10.08	67.21*	91.27	65.14	21.7	71.5	3025.14

† Character abbreviations as defined in Table 3.4.

*,** Lowest and highest values, respectively.

Table 3.8 Correlation coefficients (n = 45) between quantitative traits and grain yield per panicle, head weight and grain number per panicle in sorghum.

	GYPP†	HWt	GNPP
HWt	0.94**	-	-
GNPP	0.74**	0.73**	-
DF	0.25	0.29	0.22
PHt	0.49**	0.43**	0.35*
LN	0.56**	0.62**	0.46**
LL	0.15	0.26	0.25
LW	0.48**	0.56**	0.57**
LA	0.34*	0.44**	0.44**
INL	0.18	0.15	0.21
LSL	0.46**	0.44**	0.44**
NPBP	0.60**	0.58**	0.21
TSWt	0.43**	0.37**	-0.27
TP	0.17	-0.16	0.08
PW	0.10	0.14	0.23
PL	-0.13	0.00	0.10

*, ** r values significant at p = 0.05 and p = 0.01 probability level, respectively. † Character abbreviations as defined in Table 3.3.

Table 3.9 Principal component (PC) analysis of 16 quantitative traits in 45 sorghum accessions showing eigenvectors, eigenvalues and proportion of variations explained with the first eight PC axes.

Character	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
DF[†]	-0.20	0.09	-0.12	0.19	0.17	0.68	0.34	0.21
LN	-0.33	-0.04	-0.21	0.03	0.04	0.14	-0.22	-0.35
LL	-0.24	0.32	-0.25	0.29	-0.12	-0.23	-0.22	0.05
LW	-0.35	0.23	-0.13	0.03	-0.08	0.06	-0.05	0.24
LA	-0.32	0.30	-0.23	0.17	-0.11	-0.05	-0.15	0.16
INL	-0.17	0.03	0.48	0.41	0.15	0.12	0.19	-0.27
LSL	-0.27	-0.00	0.43	0.11	0.13	-0.05	0.12	0.03
PHt	-0.26	-0.16	0.28	0.30	-0.01	-0.32	-0.11	-0.04
PL	-0.04	0.39	0.25	-0.24	0.25	-0.38	0.18	0.44
PW	-0.11	0.14	0.26	-0.40	0.39	0.34	-0.59	0.07
NPBP	-0.23	-0.36	-0.08	0.01	0.27	-0.17	-0.32	-0.10
HWt	-0.35	-0.18	-0.09	-0.32	0.02	-0.10	0.27	-0.07
GYPP	-0.34	-0.29	-0.00	-0.25	-0.14	-0.03	0.18	0.12
TSWt	-0.06	-0.44	-0.28	0.11	0.39	-0.08	0.12	0.38
TP	0.05	-0.33	0.27	0.16	-0.49	0.19	-0.31	0.53
GNPP	-0.30	0.02	0.15	-0.38	-0.45	0.07	0.11	-0.15
Eigenvalues	2.97	2.74	1.16	1.07	1.34	1.09	1.06	1.04
Individual%	18.55	17.10	7.24	6.71	8.39	6.82	6.60	6.49
Cumulative%	18.55	35.64	42.88	49.59	57.98	64.80	71.40	77.89

[†] Character abbreviations as defined in Table 3.3.

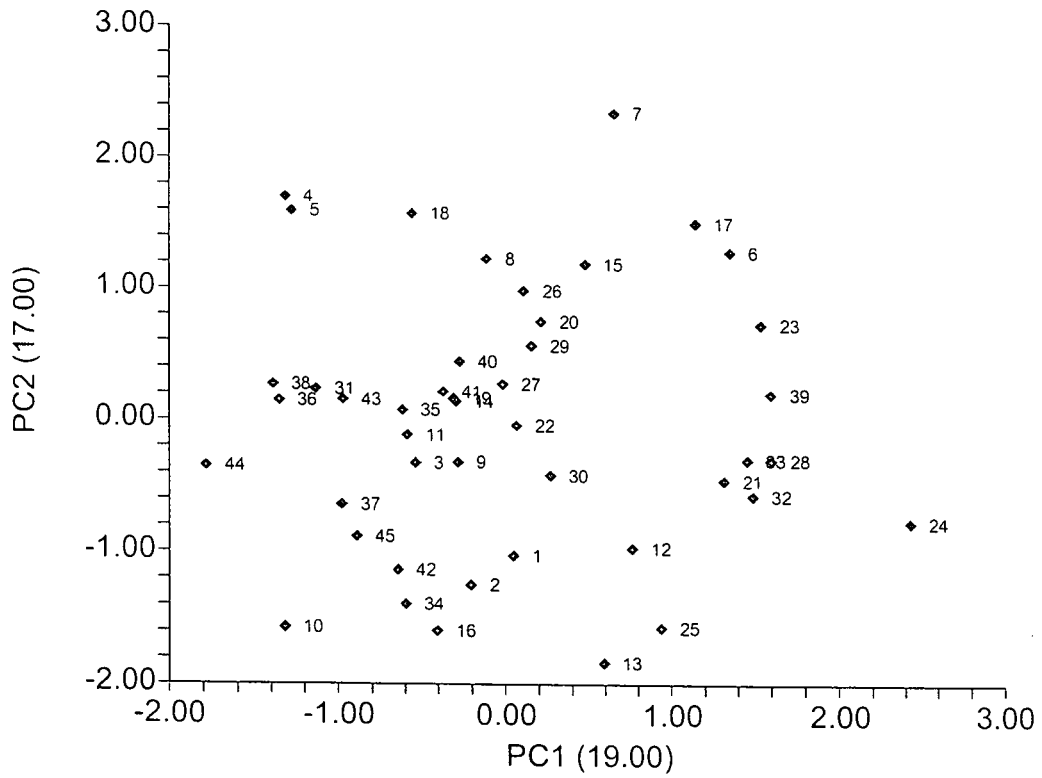


Figure 3.3 Principal component plot of the 45 sorghum accessions, estimated using 16 quantitative traits data. (Each accession is designated by number given in Table 3.1).

In general, the name given by the farmers to landraces does reflect morpho-agronomically different sorghum landraces. For example, the two *Zangadas* (#32 and 33) from Habro were clustered in cluster VII, while Dassile (#34) from the same collection site was grouped in cluster II. On the other hand, between identically named landraces grown in the same locality or different localities, a greater level of dissimilarity was observed. Accessions with the local name *Muyra* (#4, 5, 10, 13 and 45) were grouped into three distinct cluster groups (Figure 3.2).

The implications of these results for genetic resource collection and maintenance are immense. Firstly, as expected, geographical separation is likely to be associated with overall genetic differences. However, as confirmed from these results large differences may also be found within quite small regions. Secondly, although seven well-differentiated clusters were obtained from the 16 quantitative traits based data appreciable intra-

cluster diversity was found. Phenotypic diversity estimates, aided by molecular marker analysis may help to establish the amount of genetic diversity still available for conservation and for further sorghum crop improvement.

3.4 Conclusions

Among the 10 qualitative traits measured in the 45 sorghum accessions, highest diversity index (H') was obtained for grain colour, inflorescence exsertion, panicle compactness and shape, endosperm texture, and glume colour. On the other hand, the traits that showed the lowest average diversity index for the 10 localities were stalk juiciness, grain covering, awns and leaf midrib colour (Table 3.4). Furthermore, it was found that the proportion of total diversity obtained within collection sites were larger than between the collection sites. Pooled over characters within localities, the mean of H' ranged from 0.38 for Deder to 0.73 for Kurfa Challe. Panicle compactness and shape, and glume colour showed high diversity index in all the localities (Table 3.5).

Clustering of the accessions based on dissimilarity of quantitative traits produced seven groups. In most cases, accessions from one collection site appeared in more than one cluster, indicating the existence of phenotypic variation within the locations. The exceptions are accessions from Chinhakssen and Deder, which fell in cluster I and II, respectively (Table 3.6). The existence of wider agro-morphological diversity among the sorghums accessions implies the potential to improve the crop and the need to conserve these resources.

REFERENCES

Assefa, K., Ketema, S., Tefera, H., Nguyen, H.T., Blum, A., Ayele, M., Bai, G., Simane, B. and Kefyalew, T. 1999. Diversity among germplasm lines of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. Euphytica 106: 87-97.

Ayana, A. and Bekele, E. 1998. Geographical patterns of morphological variation in sorghum (*Sorghum bicolor* (L)Moench) germplasm from Ethiopia and Eritrea: qualitative characters. Hereditas 129: 195-205.

Ayana, A. and Bekele, E. 1999. Multivariate analysis of morphological variation in sorghum [*Sorghum bicolour* (L.) Moench] germplasm from Ethiopia and Eritrea. Genetic Resources and Crop Evolution 46: 273-284.

Bechere, E., Belay, G., Mitiku, D., and Merker, A. 1996. Phenotypic diversity of tetraploid wheat landraces from north-central regions of Ethiopia. Hereditas 124:165-172.

Bekele, E. 1984. Analysis of regional patterns of phenotypic diversity in the Ethiopian tetraploid and hexaploid wheats. Hereditas 100:131-154.

Brooke, C. 1958. The durra complex in the central Highlands of Ethiopia. Econ. Bot. 2:192-204.

Demissie, A. and Bjørnstad, Å. 1996. Phenotypic diversity of Ethiopian barley in relation to geographical regions, altitudinal range, and agro-ecological zones: as an aid to germplasm collection and conservation strategy. Hereditas 124: 17-29.

Doggett, H. 1988. Sorghum (Second edition). Longman, UK.

Geleta, N. 1997. Variability and association of morpho-agronomic characters with reference to highland sorghum [*Sorghum. bicolour* (L.) Moench] landraces of Hararghe, eastern Ethiopia. M.Sc. Thesis, Alemaya University of Agriculture, Ethiopia.

Grenier, C., Bramel-Cox, P.J. and Hamon, P. 2001. Core collection of sorghum: I. Stratification based on eco-geographical data. Crop Sci. 41:234-240.

Harlan, J.R. 1969. Ethiopia: A center of diversity. Economic Botany 23:309-314.

IBPGR/ICRISAT, 1993. Descriptors for Sorghum [*Sorghum bicolour* (L) Moench]. International Board of Plant Genetic Resources. Rome, Italy/ International Crop Research Institute for Semi-Arid Tropics, Patancheru, India.

Jain, S.K., Qualset, C.O., Bhatt, G.M. and Wu, K.K. 1975. Geographical patterns of phenotypic diversity in a world collection of durum wheat. Crop Sci. 15:700-704.

Kebebew, F., Tsehaye, Y. and McNeilly, T. 2001. Morphological and farmers cognitive diversity of barley (*Hordeum vulgare* L. [Poaceae]) at Bale and North Shewa of Ethiopia. Genetic Resources and Crop Evolution 48:467-481.

Kebede, Y. 1991. The role of Ethiopian sorghum germplasm resources in the national breeding programme. In: Engels, J.M., Hawakes, J.G. and Worede, M. (Eds.). Plant Genetic Resources of Ethiopia, Cambridge University Press. Cambridge, pp.315-322.

Marshall, D.R. and Brown, A.H.D. 1975. Optimum sampling strategies in genetic conservation. In: Frankel, O.H. and Hawkes, J.G. (Eds.). Crop genetic resources for today and tomorrow. Cambridge Univ. Press, London, pp 53-80.

Munsell Colour Firm. 1990. The Munsell book of colour: Matte collection. New Windsor, New York.

NCSS, 2000. Number Cruncher Statistical Systems, Dr. Jerry L. Hintze, 329 North 1000 East, Kaysville, Utah 84037, Canada.

Negassa, M. 1985. Patterns of phenotypic diversity in an Ethiopian barley collection and the Arsi-Bale highland as a centre of origin of barley. Hereditas 102:139-150.

Negassa, M. 1986. Estimates of phenotypic diversity and breeding potential of Ethiopian wheats. Hereditas 104:41-48.

Perry, M.C. and McIntosh, M.S. 1991. Geographical patterns of variation in the USDA soybean germplasm collection: I. Morphological traits. Crop Sci. 31:1350-1355.

Stemler, A.B.L., Harlan, J.R. and de Wet, J.M.J. 1977. The sorghums of Ethiopia. Economic Botany 31:446-460.

Stickler, F.C., Weaden, S., and Pauli, A.W. 1961. Leaf area determination in grain sorghum. Agron. J. 53:187-188.

Tesfaye T., Getachew, B. and Worede, M. 1991. Morphological diversity in tetraploid wheat landrace populations from central highlands of Ethiopia. Hereditas 114: 171-176.

Teshome, A. 2001. Long-term sustainability of Ethiopian landraces at risk. In: Eberlee, J. (Ed.) International Development Research Centre Reports Online. http://www.idrc.ca/reports

Teshome, A., Baum, B.R., Fahrig, L., Torrance, J.K., Arnason, T.J. and Lambert, J.D. 1997. Sorghum [*Sorghum bicolour* (L.) Moench] landrace variation and classification in North Shewa and South Welo, Ethiopia. Euphytica 97: 255-263.

Tolbert, D.M., Qualset, C.O., Jain, S.K., and Craddock,J.C. 1979. A diversity analysis of a world collection of barley. Crop Sci. 19:789-794.

Vavilov, N.I. 1951. The origin, variation, immunity and breeding of cultivated plants. Chronica Botanica 13:1-366.

Worede, M. 1988. Diversity and genetic resource base. Eth. J. Agric. Sci. 10:39-52.

CHAPTER 4

ANALYSIS OF GENETIC DIVERSITY BASED ON DNA MARKERS IN SORGHUM

Abstract

Sorghum [*Sorghum bicolor* (L.) Moench] is one of the most important cereal crops providing food in Ethiopia, which is believed to be its centre of origin and diversity. Genetic diversity among landrace sorghum accessions from 10 localities in the eastern highlands of Ethiopia and breeding cultivars obtained from Alemaya University, Ethiopia, was assessed using amplified fragment length polymorphism (AFLP) and microsatellite (SSR) markers. The extent of genetic diversity among the different sorghum accessions was as high as 85% based on AFLPs and 90% using microsatellite data. The accessions studied separated into five clusters for both AFLP and microsatellite data. AFLPs provided more detail discrimination than microsatellite data. Both cluster and PCA analyses failed to distinctly group accessions from the same locality or presumably the same name. Despite the relatively low discrimination level observed for SSR marker data in this study, both marker techniques proved to be useful for characterising and identifying genetic diversity in sorghum.

4.1 Introduction

Sorghum is one of the world's most important crop plants, ranking fifth in acreage among cereals (Doggett, 1988) and first in the eastern regions of Ethiopia (CSA, 2000). Northeastern Africa, which includes Ethiopia, is believed to be the centre of origin and domestication for sorghum (Vavilov, 1951; Doggett, 1988; Stemler *et al.*, 1977). The sorghum grown by small-scale farmers in eastern Ethiopia is observably diverse for characters including plant height, panicle types and seed colour. This is attributed to adaptation to a wide range of geographical and ecological niches and climatic regimes. On the basis of survey observations and current

germplasm collection, accessions *Muyra, Fandisha, Wagare*, and *Dassile* are the most frequently represented in the AU/SIP collection. However, it is impossible to determine whether a specific accession represents a single genotype due to continued adaptation and selection. In order to establish a sound basis and strategy for germplasm utilisation and maintenance, sorghum accessions need to be characterised genetically.

Molecular markers are considered to be an efficient method for estimating genetic diversity due to the abundance of markers that are not affected by environmental or epistatic interactions that may affect morphological traits (Schut *et al.*, 1997; Gepts, 1993). Techniques to generate markers based on the polymerase chain reaction provide new opportunities for characterising and describing germplasm and assessing genetic diversity within and between different crop varieties. These techniques include RAPD (Welsh and McClelland, 1990), AFLP (Zabeau and Voss, 1993) and SSR (Tautz, 1989). Previous reports to assess genetic variation in sorghum are based on RFLPs (Ahnert *et al.*, 1996), RAPDs (Ayana *et al.*, 2000; Menkir *et al.*, 1997) and SSRs (Ghebru *et al.*, 2002; Smith *et al.*, 2000; Djè *et al.*, 1999; Brown *et al.*, 1996) and have shown that these techniques are suitable for applications relevant to conservation and utilisation of sorghum germplasm.

The potential of semi-automated, robust, cost-effective molecular genetic markers, specifically AFLPs, allows the evaluation of quality in germplasm collections for enhanced breeding (Mitchell *et al.*, 1997; Morell *et al.*, 1995). A previous study based on qualitative and quantitative characters has identified significant variability among sorghum accessions taken from the same region as the present study (Geleta, 1997). However, there is no molecular data for comparison of these accessions. The objectives of the study were (1) to assess the extent and pattern of genetic diversity among sorghum accessions from the selected region using AFLP and SSR markers, (2) examine the distribution patterns of genetic diversity in different localities, and (3) verify the applicability of AFLP and SSR markers in accession identification.

4.2 Materials and methods

4.2.1 Plant material

A total of 45 accessions (Table 3.1) were used in this study. Three to four plants were grown in 8 L size pots, containing soil (Bainsvlei soil type), under standard glasshouse conditions at the University of the Free State, Bloemfontein, South Africa, during March through August 2001. The growth temperature was set at 14±2°C night and 28±2°C day.

4.2.2 DNA extraction

Leaf material was taken from 10, four-to six–week–old, plants of each accession. Single-plant samples were ground to a powder in liquid nitrogen using a mortar and pestle. A modified monocot extraction procedure (Edwards *et al.*, 1991) was followed to isolate the DNA. Extraction buffer (10 ml) (1M Tris-HCl pH 8: 0.25M EDTA, and 1.25% (w/v) SDS) and 1 ml (10% w/v) Cetyl triethyl ammonium bromide (CTAB) was added. The homogenate was vortexed and incubated at 65°C for 60 min, with periodic shaking. Chloroform extraction was performed to remove cellular debris and proteins by the addition of 10 ml chloroform-isoamyl alcohol (24:1v/v) followed by centrifugation for 15 min at 10 krpm. Thereafter, the DNA was precipitated by the addition of two volumes of cold absolute ethanol. The precipitate was spooled using a sterile Pasteur pipette and washed twice in 70% ethanol. The DNA was dissolved in 250 µl sterile distilled water and stored at -20°C.

4.2.3 DNA concentration determination

The concentration and protein content of individual DNA sample was determined spectro-photometrically (U-2000), by taking readings at 260 nm and 280 nm. The DNA concentration was calculated using the formula, [DNA] = Optical density (OD_{260}) x dilution x constant (50 µg/ml).

The DNA was diluted to a working concentration of 100 ng/µl in sterile distilled water. Equal quantities (100 ng) of genomic DNA from 10 plants for each accession were bulked and used in AFLP and SSR analyses.

4.2.4 AFLPs

4.2.4.1 Restriction endonuclease digestion and ligation of adaptors

Genomic DNA (250 ng of the bulked DNA) was double digested with 5 units each of *EcoRI* and *MseI* endonuclease, at 37°C for 2 hr. The digested DNA fragments were ligated to *EcoRI* and *MseI* adaptors (Table 4.1) with T4 DNA ligase for 2 h at 20 ± 2°C. The ligated DNA was diluted 1:10 In TE buffer (10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA) and stored at -20°C.

4.2.4.2 PCR amplification reactions

PCR was performed in two consecutive reactions: a pre-selective and selective PCR, following the protocol supplied by the manufacturer (GIBCO BRL). In the pre-selective reaction, genomic DNA was amplified using an AFLP primer pair, each having one selective nucleotide (Table 4.1). Accordingly, a 5 µl diluted ligation product, 40 µl pre-amplification primer mix, 5 µl 10X PCR buffer mixed with $MgCl_2$ for AFLP and 1 unit / 0.2 µl of Taq polymerase were mixed for the pre-selective reaction. The pre-selective reactions were performed as follows: 20 cycles of 30 sec at 94°C, 60 sec at 56°C and 2 min at 72°C. Pre-selective PCR amplification was confirmed by gel electrophoresis and the amplified product diluted to 1:50 in TE buffer (10 mM Tris-HCl [pH 8.0] and 0.1 mM EDTA) and used as template for the selective amplification using AFLP primers, each containing three selective nucleotides.

Selective PCR amplification was performed in 20 µl reactions consisting of, 5 µl pre-selective template DNA (1:50 dilution), 4.5 µl (6.7 ng/µl, dNTPs) *Mse* primer with selected nucleotide extensions, 1 µl (1 µM) Eco primer with selected nucleotide extensions (Eco-ACA and Eco-AAC labelled with

FAM and NED, respectively (PE Biosystems), 2 µl 10x PCR buffer (200 mM Tris-HCl [pH 8.4], 15 mM $MgCl_2$, 500 mM KCl) and 1 0.1 µl *Taq* polymerase (5 units / µl). Selective PCR amplification reactions were performed for 35 cycles, with 30 sec at 94°C, and 30 sec at 65°C, followed by 2 min at 72°C. The annealing temperature was lowered 0.7°C in each subsequent cycle during the first 12 cycles down to 56°C. All amplification reactions were performed in a PCR System 2700 (Applied Biosystems).

Table 4.1 Adaptors and primers used for pre-selective and selective AFLP amplification reactions.

Primer / adaptor code	Sequence
1. Adaptor:	
EcoRI adaptor	5'-CTCGTAGACTGCGTACC-3'
	3'-CATCTGACGCATGGTTAA-5'
MseI adaptor	5'-GACGATGAGTCCTGAG-3'
	3'-TACTCAGGACTCAT-5'
2. Primer:	
EcoRI primer	
E-AAC	5'-GATCTGCCTACCAATTCAAC-3' (NED)
E-ACA	5'-GATCTGCGTACCAATTCACA-3' (FAM)
MseI primer	
M-CAA	5'-GATGAGTCCTGAGTAACAA-3'
M-CAT	5'-GATGAGTCCTGAGTAACAT-3'
M-CTA	5'-GATGAGTCCTGAGTAACTA-3'
M-CAG	5'-GATGAGTCCTGAGTAACAG-3'

Following selective amplification, 5 µl of amplification product was mixed with 24 µl formamide (deionised) and 1 µl GeneScan™ 500 ROX™ size standard marker (PE Biosystems), denatured at 94°C for 10 min and quick cooled in ice slurry and resolved according to size on a Perkin-Elmer ABI310 Automated Capillary Sequencer (PE Biosystems).

4.2.4.3 AFLP analysis

AFLP analysis was performed using GeneScan® software. Only clear and unambiguous bands were included in the analyses. AFLP fragments larger than or equal to 60 bp with a peak height above or equal to 45 RFUs were scored. A visual comparison was used to correlate the binary output of electropherograms.

4.2.4.4 Statistical analysis

Each AFLP fragment was treated as a unit character and all the accessions were scored for presence or absence of AFLP fragments against the eight primer combinations used. Data was entered into a binary matrix as discrete variables ("1" for presence and "0" for absence of a 'homologous' fragment). The number of polymorphic and monomorphic fragments was determined from the amplified fragments for each primer pair (Appendix I). Monomorphic loci were excluded from further data analyses. The bivariate 1-0 data were used to estimate genetic distance for all possible pair-wise comparisons between accessions using the Euclidean distance method. The genetic distance matrix was used for cluster analysis using the Unweighted Pair Group Method with arithmetic Averages (UPGMA) algorithm. Principal component analysis (PCA) was performed to visualise the dispersion of individual accessions in relation to the first two principal axes of variation. All statistical analyses were done using the NCSS statistical package (NCSS, 2000).

4.2.5 Microsatellites (SSRs)

4.2.5.1 SSR primers

Fifteen SSR sorghum primer pairs (Brown *et al.,* 1996) were used in this study. Primers were excluded from the study if banding patterns were too complex to score accurately from agarose gels and/or if the primers failed to amplify consistently in all 45 accessions. The SSR markers were

selected across all the sorghum linkage groups (A to I). Ten SSR primer pairs were used in the final analysis (Table 4.2).

4.2.5.2 SSR amplification

A standard PCR method was used to amplify microsatellites. PCR conditions were optimised for each primer pair by adapting the annealing temperature (Tm). The PCR reaction was performed by taking 0.5 µl of bulked DNA, 2.5 µl 10x PCR buffer (200 mM Tris – HCl (pH 8.4), 500 mM KCl), 0.75 µl $MgCl_2$ (50 mM), 0.5 µl dNTP's (40 mM), 0.2 µl *Taq* polymerase (5 units / µl) and 18.25 µl sterile distilled water in a total reaction volume of 25 µl. PCR cycling conditions were: 2 min initial denaturation at 95°C; followed by 30 cycles of 30 s at 94°C, 45 s at either 45°C (Sb4-32, sb5-85 and sb6-84), 50°C (sb4-22) or 55°C (sb1-10, sb4-15, sb5-236, sb6-36, sb6-57 and sb6-342) and 1 min elongation at 72°C, followed by a final elongation of 10 min at 72°C. All PCR reactions were performed on a PCR System 2700 (Applied Biosystems).

4.2.5.3 SSR locus visualization / gel analysis

The result of the PCR amplification was analyzed by electrophoresis on a 2% agarose gel (Molecular Screening agarose Roche) for resolution of fragments ranging from 50 to 1000 bp in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) run at 80V for 2.5 hours. Amplified fragments were visualised and sized using the Gel Doc 1000^{TM} image analysis system (Biorad) after ethidium bromide (0.5 µg/ml) staining.

4.2.5.4 Data collection and statistical analyses

The presence or absence of each fragment was coded as 1 or 0, respectively, and scored in a binary data matrix. The polymorphism information content (PIC) was determined according to the formula described in Smith *et al.* (2000):

$$PIC = 1 - \sum_{i=1}^{n} (f_i)^2$$

where f_i is the frequency of the *i*th allele carried by the population, calculated for each SSR locus. PIC values range from 0 (monomorphic) to 1 (highly discriminative). The PIC values provide an estimate of the discriminatory power of a marker by taking into account not only the number of alleles at a locus, but also the relative frequencies of those alleles in the population under study.

Table 4.2 Summary of the SSR-primer pairs used in this study.

SSR Locus	Primer sequence	Repeat motif	Linkage group	Size range (bp)[1]
Sb1-10	F:GTGCCGCTTTGCTCGCA	$(AG)_{27}$	D	242-488
	R:TGCTATGTTGTTTGCTTCTCCCTTCTC			
Sb4-15	F:GCTGCTAAGCCGTGCTGA	$(AG)_{16}$	E	120-134
	R:TTATTTGGGTGAAGTAGAGGTGAACA			
Sb4-22	F:TGAGCCGAAAACCGTGAG	$(ACGAC)_4$ $(AG)_6$	NA	270-300
	R:CCCAAAACCAAGAGGGAAGG			
Sb4-32	F:CTCGGCGGTTAGCACAGTCAC	$(AG)_{15}$	E	160-216
	R:GCCCATAGACAGACAGCAAAGCC			
Sb5-85	F:AGACGCTTTTCTCTCTCTCTCTCTCTCTCT	$(AG)_{12}$	NA	200-225
	R:TAGCCCTGCCGCATACTGAATG			
Sb5-236	F:GCCAAGAGAAACACAAACAA	$(AG)_{20}$	G	162-222
	R:AGCAATGTATTTAGGCAACACA			
Sb6-36	F:GCATAATGACGGCGTGCTC	$(AG)_{19}$	C	155-199
	R:CTTCCAAGTGAAAGAAACCATCA			
Sb6-57	F:ACAGGGCTTTAGGGAAATCG	$(AG)_{18}$	C	283-313
	R:CCATCACCGTCGGCATCT			
Sb6-84	F:CGCTCTCGGGATGAATGA	$(AG)_{14}$	F	170-212
	R:TAACGGACCACTAACAAATGATT			
Sb6-342	F:TGCTTGTGAGAGTGCCTCCCT	$(AC)_{25}$	A	250-320
	R:GTGAACCTGCTGCTTTAGTCGATG			

[1] Data from Brown *et al.*, 1996; Dean *et al.*, 1999; Ghebru *et al.*, 2002.

NA = Not available

Genetic distances between pairs of accessions for SSR data were calculated from comparisons of the band scores using the Euclidean distance method in the NCSS program. The 45 accessions were clustered

based on genetic distance, using the UPGMA clustering method, and relationships between accessions visualised as dendrograms. Principal component analysis (PCA) based on correlation matrix of the allelic frequencies was performed with NCSS statistical package (NCSS, 2000).

4.3 Results and discussion

4.3.1 AFLP markers

4.3.1.1 Level of AFLP polymorphism

AFLP analysis of bulked DNA of 45 sorghum accessions from the eastern highlands of Ethiopia using eight primer combinations identified a total of 651 fragments of which 85% were polymorphic (Table 4.3). A typical AFLP electropherogram of two accessions illustrating band sizes and peak heights is shown in Figure 4.1. The number of scorable fragments detected by an individual primer combination ranged from 55 (for primer pair E-AAC/M-CAT) to 114 (for primer pair E-ACA/M-CAA). The number of polymorphic fragments for each primer pair varied from 45 (for primer pair E-AAC/M-CTA) to 102 (for primer pair E-ACA/M-CAA) with an average of 69 per primer pair (Table 4.3).

Based on AFLP data percentage polymorphism identified per primer pair ranged from 69% (E-AAC/M-CTA) to 94% (E-ACA/M-CAT). The polymorphisms generated by any single primer combination were able to uniquely distinguish all accessions. The two accessions (#25 and #26) most closely related based on AFLP data differed by 2 AFLP fragments. Both accessions were from Doba but are known by different names. In 49 accessions of *tef,* four or more primer pair combinations were required to separate each accession distinctly (Bai *et al.*, 1999)

The frequency distribution of the 990 pair-wise genetic distance (GD) estimate values generated among 45 sorghum accessions for the combined AFLP data is presented in Figure 4.2. The average GD among

all the accessions was 0.62 with values ranging from 0.41 (accession #1 vs #5) to 0.75 (accession #5 vs #36). Most (60%) of the GD values ranged from 0.61 to 0.70, indicating the absence of a close genetic relationship between those pairs of accessions. Accession pairs #1 and #2, #44 and #45, and #1and #3 were more closely related to each other with GD ranging from 0.41 to 0.44.

Table 4.3 The number of AFLP fragments and degree of polymorphism determined for 45 sorghum accessions using eight AFLP primer combinations.

Primer combinations	Total bands	Polymorphic bands	Polymorphism rate (%)
E-AAC/M-CAA	90	84	93
E-ACA/M-CAA	114	102	90
E-AAC/M-CAT	55	49	89
E-ACA/M-CAT	79	74	94
E-AAC/M-CTA	65	45	69
E-ACA/M-CTA	96	75	78
E-AAC/M-CAG	68	57	84
E-ACA/M-CAG	84	66	79
Total	651	552	
Average	81	69	85

Cluster analysis using the UPGMA method grouped the 45 accessions into two major clusters (Figure 4.3). At a genetic distance of 0.62 the first major cluster was split into two sub-clusters and the second into three sub-clusters. The cophenetic correlation coefficient between the dissimilarity and the original matrix for data generated from eight primer combinations was high (0.82), indicating a good fit of the cluster analysis performed. In most cases, the accessions were clustered in general agreement with their site of collection and geographic distribution. For example, seven of the 11 accessions from AU were clustered in the second major cluster, while four of them grouped in the first major cluster.

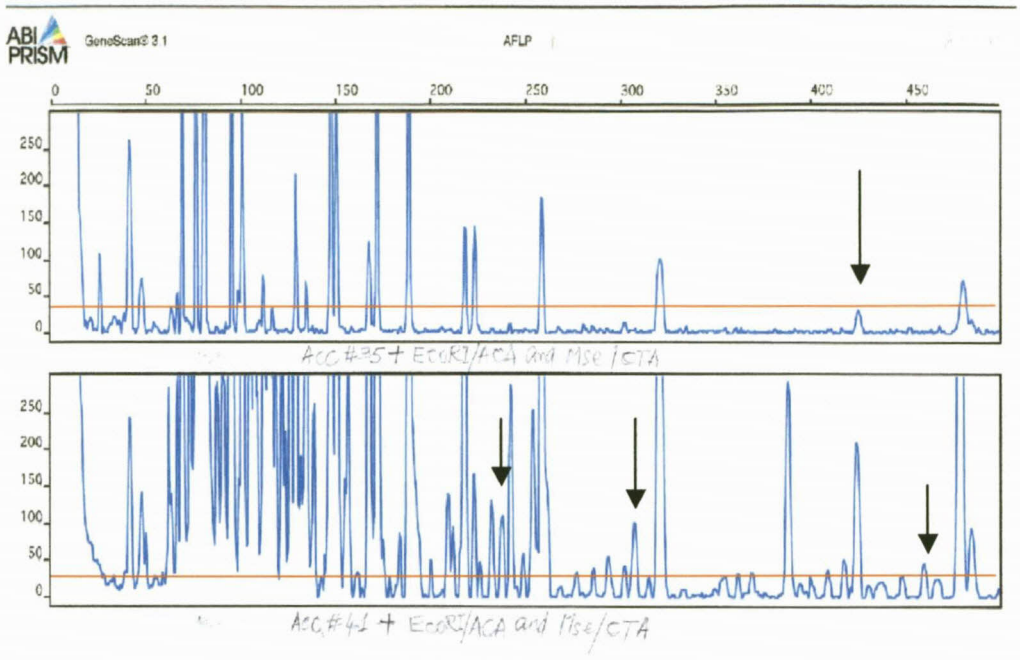


Figure 4.1 Genescan electropherograms showing part of the AFLP runs of bulked DNA of two accessions (ETS 721 and ETS 2752) using *EcoRI* / ACA and *MseI* / CTA primers in the present study. Arrows indicate some of the peaks that are polymorphic between the two accessions.

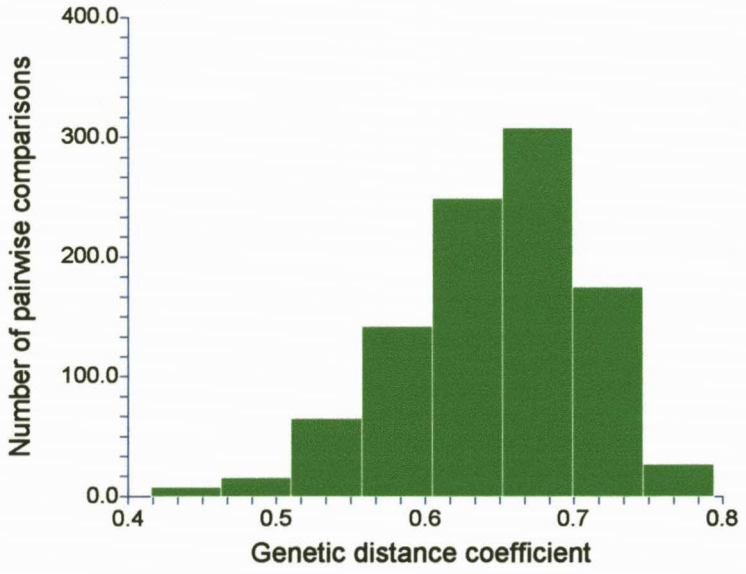


Figure 4.2 Frequency distribution of pair-wise genetic distance coefficients obtained for 45 sorghum accessions using AFLP data.

Similarly, eight of the 10 accessions from Tulo were grouped together in cluster II, whereas two accessions (accessions #15 and 20) were grouped in cluster V together with accession #12 from Deder.

It is interesting to note that clustering based on similarity of local or cultivar name was not common. As an example, accessions with the name *Wagare* (# 1, 2, 3, 8 and 16) were clustered in three different clusters, where # 1, 2, and 3 grouped in cluster I, and #8 and #16 were grouped in clusters III and II, respectively. Similarly, the five accessions with the local name *Fandisha* (# 6, 7, 17, 23 and 24) were grouped in the first major cluster, but in two clusters, #6 and #7 in cluster I, whereas #17, 23 and 24 grouped in cluster II. The accessions named *Muyra* (# 4, 5, 10, and 13) grouped within the first major cluster, but subsequently accession #4 and 5 were included in cluster I, while #10 & 13 were grouped in cluster II. The two accessions collected from Habro with the local name *Zangada* (#32 & 33), also first grouped under the first major cluster, and then split into two different sub-clusters.

The results this study confirms that AFLP markers are applicable for DNA fingerprinting sorghum accessions to determine their genetic relationships and for genotype identification studies. Furthermore, the high level of genetic diversity identified among the different sorghum accessions indicates the potential for sorghum improvement. In previous studies, sorghum accessions from wider regions of Ethiopia and Eritrea have been analysed using RAPDs and identified only intermediate levels of genetic variation (Ayana *et al.*, 2000). In this study, a different approach was taken by analysing accessions from a smaller region, including the eastern highlands of Ethiopia using the AFLP marker technique. The findings clearly demonstrate the usefulness and efficiency of AFLPs in analysing genetic diversity and discriminating the accessions studied as well as the importance of studying the genetic diversity within smaller regions.

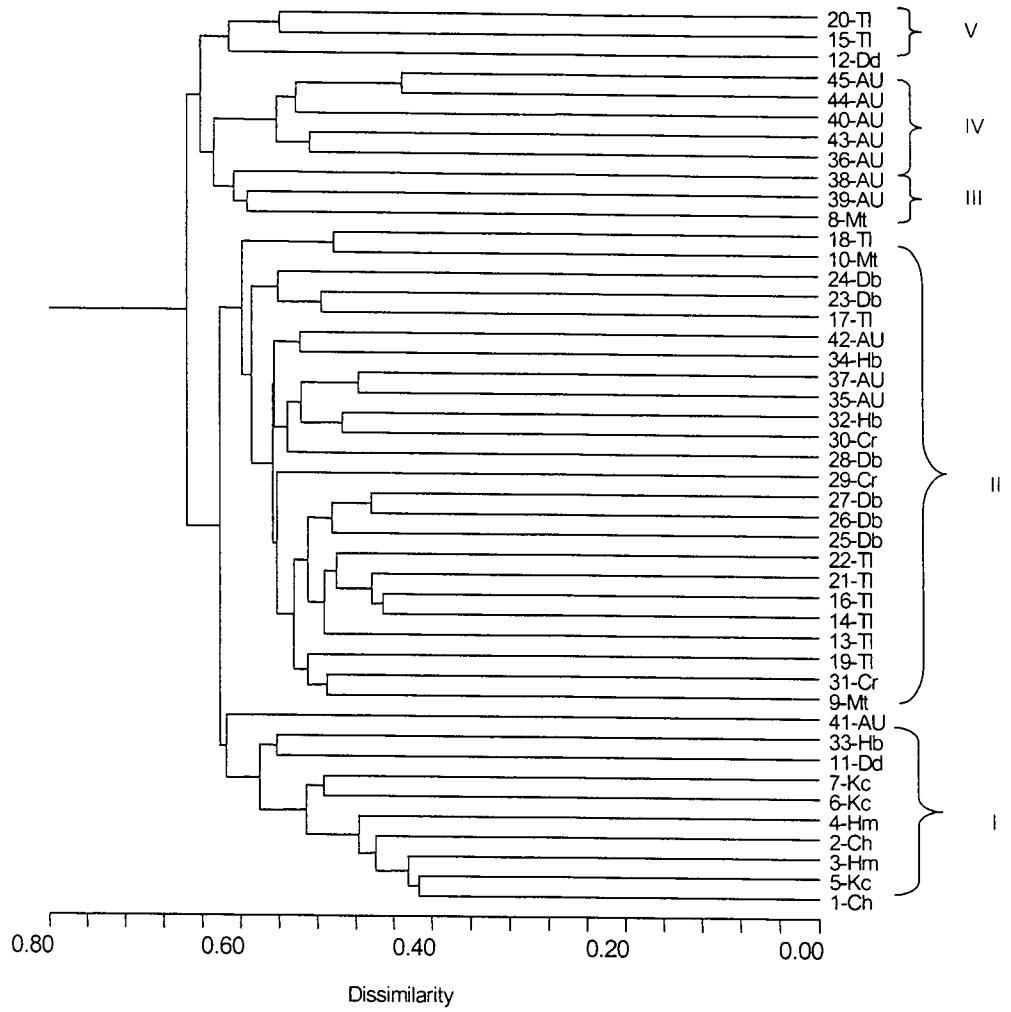


Figure 4.3 Dendrogram constructed based on AFLP data, showing genetic distance and cluster groups among 45 sorghum accessions. Each horizontal line represents a separate sorghum accession, and the number-letter after each line indicates sample number and the locality from which the accession was collected/ obtained (refer Figure 3.1).

4.3.1.3 Principal component analysis (PCA)

Principal component analysis (PCA) is a useful way to determine the effect of different variables on the data set. The PCA plots analysed using a combined AFLP data set are shown in Figure 4.4. The first and second components (PC1 and PC2) explained 45.62% of the variance. The first axis (PC1) accounted for 23.24% of the variation and was influenced by primer combination E-ACA/M-CAT and E-AAC/M-CAT (with factor loading 0.81 and 0.78, respectively). The second axis (PC2) explained 22.38% of the variation and was positively influenced by primer combination E-AAC/ M-CAG (loading 0.88) and E-ACA/ M-CAG (loading 0.81). The analysis confirmed the wider genetic diversity among the accessions; as they were randomly distributed all over the four quadrants. The AU accessions, for example, had a wider distribution of genetic variation, as seen from the first two axes.

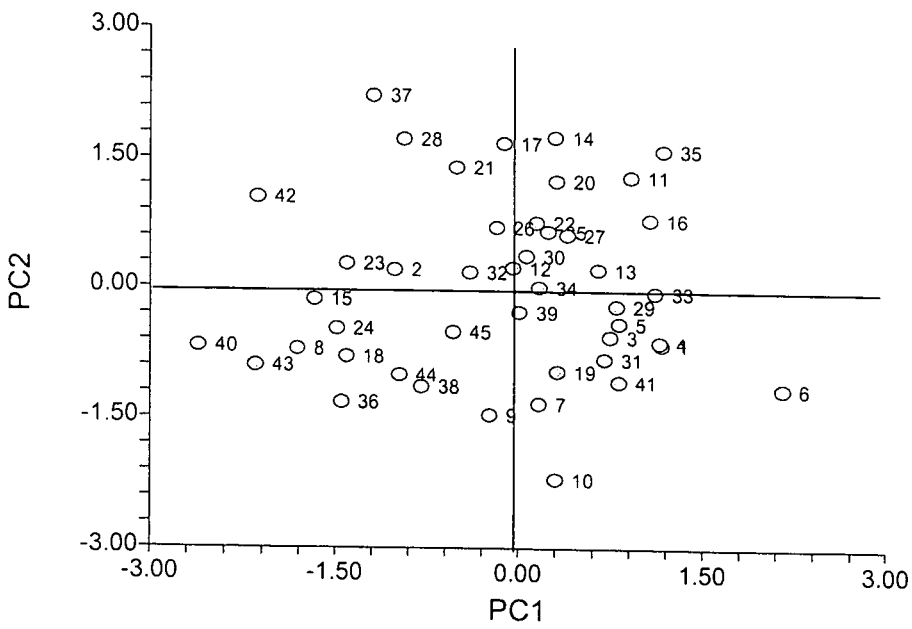


Figure 4.4 Plot of the 45 sorghum accessions against the first two principal components (PC1 and PC2) based on AFLP data. (Individual accessions are shown by numbers as described in Table 3.1).

4.3.2 Microsatellite markers

4.3.2.1 Polymorphism of microsatellites in sorghum accessions

The 45 sorghum accessions were analysed using 10 microsatellite primer pairs, which amplified bands distinguishable on agarose gels. An example of typical microsatellites polymorphism for a single primer pair (sb6-36) is shown in Figure 4.5. The level of polymorphism was estimated from the number of alleles and their frequency was analysed. The primers, their repeat number, allele number, and polymorphism information content (PIC) are given in Table 4.4. In total, the 10 primer pairs detected 48 polymophic and monomorphic alleles. The number of alleles scored per primer pair varied from one (for sb4-15) to nine (for sb4-32), with an average of 4.8 alleles per primer pair. The number of alleles observed for most of the loci was in agreement with the range reported in sorghum (Dean *et al.*, 1999; Brown *et al.*, 1996), maize (Senior *et al.*, 1998) and wheat (Ahmad, 2002), but lower than in other studies reported in sorghum (Djè *et al.*, 2000; Ghebru *et al.*, 2002).

The PIC values for SSR loci ranged from 0.52 for sb4-32 and sb6-342 loci to 0.79 for sb4-22, while one primer pair (sb4-15) was found to be monomorphic. The mean PIC value was 0.645. Smith *et al.* (2000) reported similar observations of PIC in sorghum detected by SSR markers. In this study, however, no significant correlation was detected between the repeat number and the allele number ($r = -0.31$) or polymorphic information content ($r = -0.07$). The 43 polymorphic SSR alleles collectively yielded unique genotypes for each of the 45 accessions. A wide range of fragment sizes (differences between the shortest and longest alleles ranged from 51 to 421bp was obtained, with most within the ranges previously reported in studies with different sorghum germplasm (Ghebru *et al.*, 2002; Dean *et al.*, 1999; Djè *et al.*, 2000).

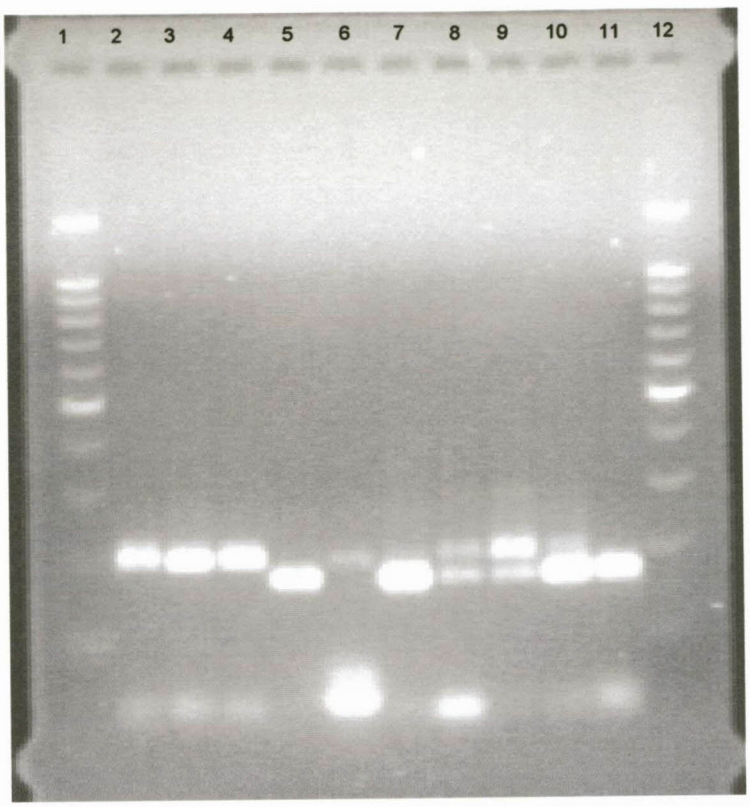


Figure 4.5 Agarose gel electrophorsis of SSR-PCR products amplified using primer Sb6-36 $(AG)_{19}$. Lanes 1 and 12 contain 100 bp size standard markers. Lanes 2 to 11 contains amplification products using bulk DNA from sorghum accessions 1 to 10, according to the accessions list in Table 3.1.

4.3.2.2 Genetic diversity

The distribution of pair-wise genetic distance among 45 accessions for nine microsatellite loci is shown in Figure 4.6. Genetic distances ranged from 0.15 (for ETS 993 x ETS 804) to 0.76 (for Fandisha faca'a x Warabi), with the majority (63%) of pair-wise comparisons having a value from 0.60 to 0.70. Cluster analysis identified two major clusters consisting of 5 subclusters at a genetic distance of 0.58 (Figure 4.7). Most of the accessions from the same collection area clustered together. Cluster I consists of accession #1, 2 and 3, the first two from Chinhakssen and the third from Haro Maya all having the same local name, *Wagare*. Cluster II comprised a large grouping of accessions predominantly from Tulo, and

some AU germplasm, and also accessions from Metta and Deder. Cluster III contained accessions collected from *Aanaas* in the Western Hararghe zone, Doba, Chiro and Tulo. Cluster IV was a group of accessions mainly from Kurfa Challe. Accession #4 from Haro Maya also fell within this cluster. A large grouping of AU germplasm, as well as the three accessions from Habro and one Chiro collected accession were found in Cluster V.

Table 4.4 Number of alleles, size range (in base pairs), and PIC value for SSR loci found in 45 accessions of sorghum.

SSR Locus	Repeat Number	No. of alleles	Size range	PIC value
Sb1-10	$(AG)_{27}$	4	110-510	0.717
Sb4-15	$(AG)_{16}$	1	135	0.000
Sb4-22	$(ACGAC)_4$ /$(AG)_6$	6	254-492	0.793
Sb4-32	$(AG)_{15}$	9	112-533	0.521
Sb5-85	$(AG)_{12}$	7	109-510	0.594
Sb5-236	$(AG)_{20}$	4	110-265	0.578
Sb6-36	$(AG)_{19}$	5	165-279	0.773
Sb6-57	$(AG)_{18}$	4	292-343	0.683
Sb6-84	$(AG)_{14}$	3	173-295	0.630
Sb6-342	$(AC)_{25}$	5	254-634	0.520
Average		4.8		0.645

From Brown *et al.*, 1996

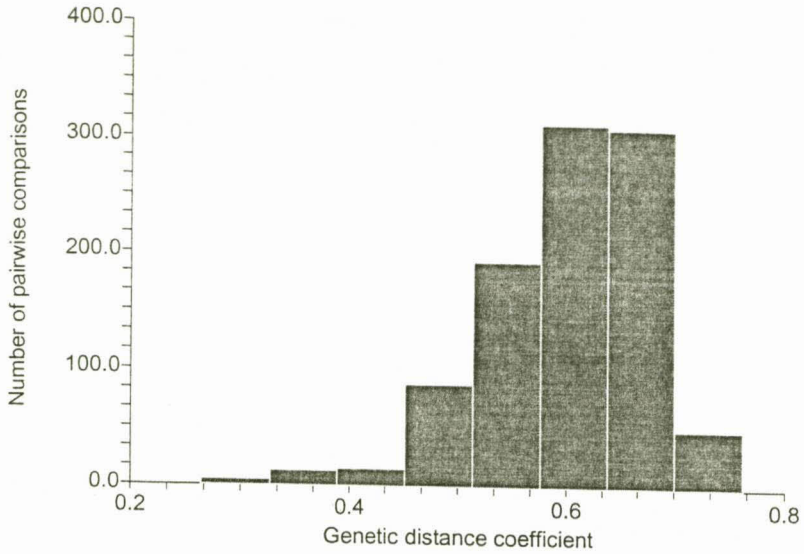


Figure 4.6 Frequency distribution of pair-wise genetic distance coefficients among 45 sorghum accessions based on SSR data.

A principal component analysis was performed on the basis of allele frequencies, and individual accessions were plotted against the first two principal components (PC1 and PC2) that were responsible for 36.8% of the total variance (Fig. 4.8). The first axis accounted for 19.2% of the variance and the principal component scores were positively influenced by primer pairs sb1-10, sb6-84 and sb6-36 (with factor loading 0.82, 0.82 and 0.56, respectively). The second axis described 17.6% of the variance and was largely influenced by sb6-57 (loading 0.88) and sb4-32 (loading 0.75). The majority of the accessions clustered very closely in the bottom left quadrant, showing the existence of a bias in terms of selection pressure for the microsatellites. Accessions #14, 25 and 43 occupied the most extreme of the first, third and fourth quadrants, respectively. The two accessions from Chinhakssen (#1 and 2) together fell in the lower left quadrant.

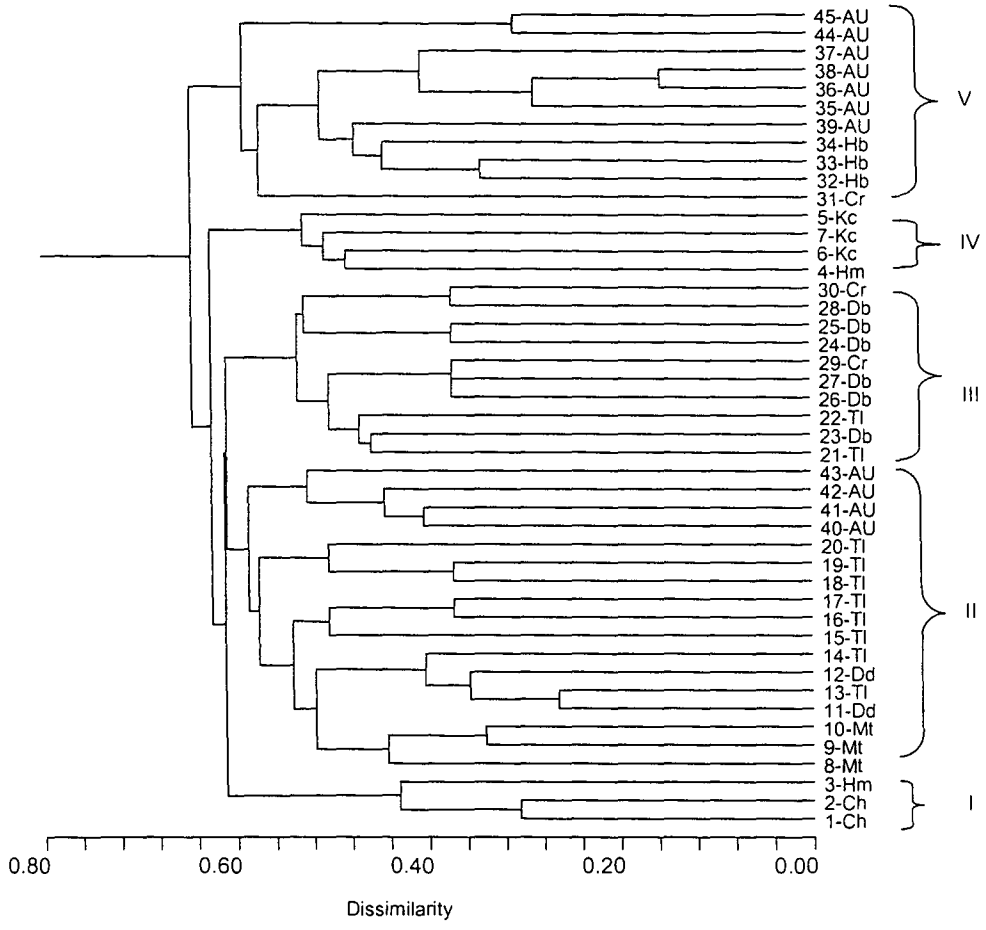


Figure 4.7 Dendrogram constructed for 45 accessions of sorghum based on data from 43 polymorphic SSR alleles. [Each horizontal line represents a separate sorghum population sample, and the number-letter after each line indicate sample number and the locality from which the accession was collected (abbreviations as in Figure 3.1)].

The use of microsatellite markers in the study of genetic diversity within and between sorghum germplasm have already been demonstrated in previous studies (Brown *et al.*, 1996; Dean *et al.*, 1999; Djè *et al.*, 2000; Ghebru *et al.*, 2002). SSR markers identified two to six fragment sizes per polymorphic locus with a total diversity index ranging from 0.21 to 0.73 on 17 temperately and tropically adapted lines of sorghum. The SSR loci are widely spread over sorghum genome with 14 of these 10 sorghum different linkage groups (Dean *et al.*, 1999). With the same SSR set, Dean *et al.* (1999) assessed the diversity among 95 'Orange' accessions and found from three to 11 alleles per locus and a genetic diversity ranging from 0.16

to 0.77. Furthermore, using the same SSR set, Grenier *et al.* (2000) reported from seven to 33 alleles per locus and a genetic diversity range of 0.71 to 0.93.

The present study is the first report of the application of the microsatellite markers for genetic diversity estimation and cultivar identification in Ethiopian originated sorghums. It is interesting to note the bias identified through PCA. This suggests that SSR loci are under selection pressure. The use of additional microsatellite loci may collectively allow the sorghum genome to be surveyed more comprehensively, and include the molecular mapping of important traits.

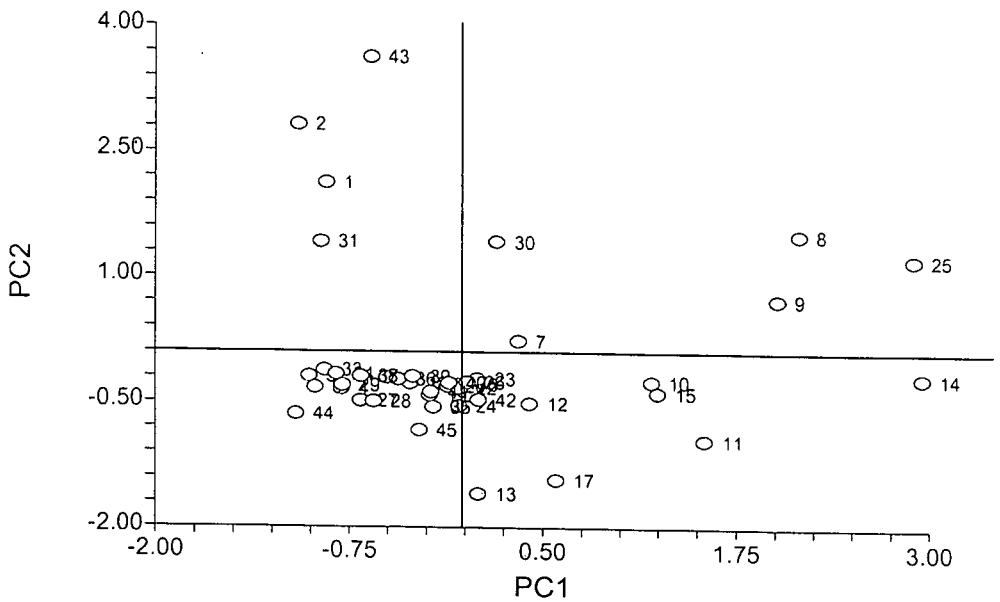


Figure 4.8 Plot of the 45 sorghum accessions against the first two principal components analysis (PC1 and PC2) computed using the SSR data. Individual accessions are designated using numbers as described in Table 3.1.

4.3.3 Comparison of AFLP and SSR markers

Both PCR-based DNA marker techniques generated high levels of polymorphisms to distinguish between all the studied sorghum accessions. The percentage of polymorphic bands was slightly higher in SSR (90%)

than AFLP (85%) data. However, the AFLP technique was 14 times more efficient in detecting polymorphism per assay, as it has a higher multiplex ratio. The average polymorphic information content (PIC) values among the 45 accessions for AFLP and SSR markers were 0.464 and 0.645, respectively (Table 4.5).

Although there are some similarities between the dendrograms from AFLP (Figure 4.3) and SSR (Figure 4.7) data, both techniques produced different results. Genetic clustering of sorghum accessions, depending on the site of collection / geographical origin, was clearly observed in SSR based dendrogram compared to AFLP data. For example, the three accessions from Metta (#8, 9 and 10) that classified together based on SSR markers, accession #8 was clustered in a different cluster based on AFLP markers. However, some accessions from the same collection site (#1 and 2) and sharing the same local name, *Wagare* (#, 1, 2 and 3), were genetically distinct.

Forty-five sorghum accessions were analysed for variability using eight selective AFLP primer combinations and 10 microsatellite loci. Both AFLP and SSR markers were highly efficient in detecting polymorphisms among the studied accessions, showing their usefulness in characterisation of sorghum germplasm accessions. The overall results confirm the high variability that can be found among landrace populations, underlining the value of landraces for future breeding programmes, which require the flexibility offered by a wide gene pool, in order to improve grain quality, neutralize the effects of environmental stresses, such as drought.

Table 4.5 Level of polymorphism and comparison of the amount of information obtained with AFLP and SSR markers in 45 sorghum accessions.

Parameters	Marker system	
	AFLP	SSR
Total number of assays	8 (primer combinations)	10 (primer pairs)
Total number of bands	651	48
Number of polymorphic bands	552	43
Mean number of bands per assay	81.4	4.8
% polymorphic bands	85.8	89.6
PIC	0.46	0.65

4.4 Conclusions

Forty-five sorghum accessions consisting of landraces, elite breeding entries and improved cultivars were analysed for variability using eight AFLP primer combinations and 10 SSR loci. The UPGMA clustering algorithms grouped the accessions into five clusters for each marker technique. However, the results indicated that the accessions could be more clearly discriminated with AFLP markers than microsatellite markers. This may result from the relatively low microsatellite alleles used, and availability of more SSR loci would facilitate a better identification and discrimination of sorghum accessions. This study has also identified a selection bias in SSR data that was not observed in AFLP data for the accessions studied.

This study also provided information for evaluating how well the sorghum landrace collection has been managed by SIP/AU. The breeding entries / improved cultivars included in the study were grouped in four and two, different clusters based on AFLP and microsatellite markers, respectively.

REFERENCES

Ahmad, M. 2002. Assessment of genomic diversity among wheat genotypes as determined by simple sequence repeats. Genome 45: 646-651.

Ahnert, D., Lee, M., Austin, D.F., Livini, C., Woodman, W.L., Openshaw, S.J., Smith, J.S.C., Porter, K. and Dalton, G. 1996. Genetic diversity among elite sorghum inbred lines assessesd with DNA markers and pedigree information. Crop Sci. 36:1385-1392.

Ayana, A., Bryngelsson, T. and Bekele, E. 2000. Genetic variation of Ethiopian and Eritrean sorghum [*Sorghum bicolor* (L.) Moench] germplasm assessed by random amplified polymorphic DNA (RAPD). Genetic Resources and Crop Evolution 47:471-482.

Bai, G., Ayele, M., Tefera, H., and Nguyen, H.T. 1999. Amplified fragment length polymorphism analysis of tef [*Eragrostis tef* (Zucc.) Trotter]. Crop Sci. 39:819-824.

Brown, S.M., Hopkins, M.S., Mitchell, S.E., Senior, M.L., Wang, T.Y., Duncan, R.R., Gonzalez-Candelas, F. and Kresovich, S. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Genet. 93:190-198.

Central Statistical Authority (CSA). 2000. The Federal Democratic Republic of Ethiopia: Agricultural sample survey report on area and production for major crops, Statistical Bulletin No 227. Addis Ababa, Ethiopia, pp 16-17.

Dean, R.E., Dahlberg, J.A., Hopkins, M.S. Mitchell, S.E. and Kresovich, S. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39:1215-1221.

Djè, Y., Forcioli, D., Ater, M., Lefebvre, C. and Vekemans, X. 1999. Assessing population genetic structure of sorghum landraces from

north-western Morocco using allozyme and microsatellite markers. Theor. Appl. Genet. 99:57-163.

Djè, Y., Heuertz, M., Lefébvre, C. and Vekemans, X. 2000. Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theor. Appl. Genet. 100: 918-925.

Doggett, H. 1988. Sorghum. Longman Group U.K. Ltd. Essex, England.

Edwards, K., Johnstone, C. and Thompson, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucl. Acid. Res. 19:1349-1458.

Geleta, N. 1997. Variability and association of morpho-agronomic characters with reference to highland sorghum [*Sorghum bicolor* (L.) Moench] landraces of Hararghie, Eastern Ethiopia. MSc. Thesis. Alemaya University of Agriculture, Alemaya, Ethiopia.

Gepts, P. 1993. The use of molecular and biochemical markers in crop-evaluation studies. In: Hecht, M.K. (Ed.). Evolutionary biology, vol. 27. Plenum press, New York, pp 51-94.

Ghebru, B., Schmidt, R.J., and Bennetzen, J.L. 2002. Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. Theor. Appl. Genet. 105:229-236.

Grenier, C., Deu, M., kresovich, S., Bramel-Cox, P.J. and Hamon, P. 2000. Assessment of Genetic Diversity in Three Subsets Constituted from ICRISAT Sorghum Collection Using Random vs Non-Random Sampling Procedures. Theor. Appl. Genet. 101:197-202.

Menkir, A., Goldsbrough, P. and Ejeta, G. 1997. RAPD based assessment of genetic diversity in cultivated races of sorghum. Crop Sci. 37:564-569.

Mitchell, S.E., Kresovich, S., Jester, C.A., Hernández, J. and Szewc-McFadden, A.K. 1997. Application of multiplex and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci. 37:617-624.

Morell, M.K., Peakall, R., Appels, R., Preston, L.R., and Lloyd, H.L. 1995. DNA profiling techniques for plant variety identification. Australian J. Exp. Agric. 35:807-819.

NCSS, 2000. Number Cruncher Statistical Systems, Dr. Jerry L. Hintze, 329 North 1000 East, Kaysville, Utah 84037, Canada.

Schut, J.W., Qi, X. and Stam, P. 1997. Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor. Appl. Genet. 95:1161-1168.

Senior, M.L., Murphy, J.P., Goodman, M.M. and Stuber, C.W. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci. 38:1088-1098.

Smith, J.S.C., Kresovich, S., Hopkins, M.S., Mitchell, S.E., Dean, R.E., Woodman, W.L., Lee, M. and Porter, K. 2000. Genetic Diversity among Elite Sorghum Inbred Lines Assessed with Simple Sequence Repeats. Crop Sci. 40:226-232.

Stemler, A.B.L., Harlan, J.R. and de Wet, J.M.J. 1977. The sorghums of Ethiopia. Econ. Bot. 31:446-460.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucl. Acids Res. 17: 6463-6471.

Vavilov, N.I. 1951. The origin, variation, immunity and breeding of cultivated plants. Chronica Botanica 13:1-366.

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

Zabeau, M. and Vos, P. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. European patent Application EP 534858A1.

CHAPTER 5

COMPARISON OF AFLP, SSR AND MORPHO-AGRONOMICAL MARKERS FOR ESTIMATING GENETIC DIVERSITY IN SORGHUM

Abstract

A comparison of the results of the different methods of the estimation of genetic diversity is important to evaluate their utility as a tool in plant breeding and germplasm conservation. Amplified fragment length polymorphisms (AFLP's), microsatellites or (SSR's) and morpho-agronomical markers were used to evaluate 45 sorghum accessions in terms of genetic diversity assessment and discrimination power. The diversity index for SSR markers was a higher (0.645) than for AFLPs (0.464). The average pair-wise genetic distance estimates were 0.566 for morpho-agronomical markers, 0.604 for SSR and 0.615 for AFLP based data. There was no congruency between dendrograms constructed from the three different data matrices and the combined matrix. In some cases, accessions from the same locality and/or with the same name were found to exhibit particularly high levels of association overall the different methods used. Both AFLP and SSR-based data matrices differentiated between the 45 accessions more distinctly than morpho-agronomical trait data and genetic diversity estimates from morpho-agronomic traits was not well suited for elucidating more complex relationships but was adequate for estimating the overall pattern of genetic variation among the accessions. Although relationships determined by molecular data are different to those determined by morpho-agronomical traits, this remains a useful way to assess diversity for breeding purposes even though the more detailed genetic relationships may be misrepresented. Therefore the strategy of combining molecular and morpho-agronomic traits would be best to study genetic diversity of sorghum accessions.

5.1 Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is one of the most important cereals of the semi-arid tropics. It is the third most important cereal crop after *tef* [*Eragrostis tef* (Zucc.) Trotter] and maize, and first in the eastern regions of Ethiopia, in terms of cultivation area and production (CSA, 2000). Current models of sorghum race and variety distribution differentiate the main S. *bicolor* races as Bicolor, Caudatum, Durra, Guinea, and Kafir (Harlan and de Wet, 1972). All of these (except Kafir) are found in Ethiopia (Stemler *et al.*, 1977; Teshome *et al.*, 1997), and have a broad agro-ecological variation, which has resulted in the accumulation of genetic diversity in this crop species.

Estimation of genetic diversity to identify groups with similar genotypes is important for conserving, evaluating and utilising genetic resources, for studying the diversity of different germplasm as possible sources of genes that can improve the performance of cultivars, and for determining the uniqueness and distinctness of the phenotypic and genetic constitution of genotypes with the purpose of protecting the breeder's intellectual property rights (Franco *et al.*, 2001; Subudhi *et al.*, 2002). In the past, plant breeders made selections of breeding material on the basis of morphological characteristics that were readily observable and that were co-inherited with the desired trait. Although these methods remain effective, morphological comparisons have limitations, including the influence of environment or management practice, subjectivity in the character evaluation are also linked to developmental stage (Morell *et al.*, 1995). In addition, morphology is not efficient in discrimination between closely related lines that differ for the trait of interest due to the scarcity of phenotypic markers (Perkin-Elmer, 1996).

The use of DNA markers for evaluating genetic diversity has improved dramatically since the advent of the polymerase chain reaction. Different techniques are used to generate DNA based markers that result in different estimates of genetic similarity depending on the number of markers generated and the genome coverage. SSR marker technique has been used

to characterize genetic diversity represented by elite inbred genotypes and cultivated races of sorghum (Brown *et al.*, 1996; Dean *et al.*, 1999; Djé *et al.*, 2000; Smith *et al.*, 2000).

Although DNA markers have been compared in the assessment of sorghum genetic diversity (Yang *et al.*, 1996; de Oliveira *et al.*, 1996; Smith *et al.*, 2000), both AFLPs and SSRs are more recent techniques, and have not been evaluated for use in discriminating between different sorghum accessions. The objective of the study was to compare the use of AFLPs, SSRs and morpho-agronomical markers to assess genetic diversity in different accessions of sorghum in Ethiopia.

5.2 Materials and methods

5.2.1 Plant material

The sorghum accessions listed in Table 3.1 were used.

5.2.2 Methods

5.2.2.1 Morpho-agronomical traits

The data collection methods described in Chapter Three, Sections 3.2.1.2 and 3.2.2.3 were used. Ten qualitative and 16 quantitative traits data were coded as presence or absence (on a 1/0) basis to compare them with DNA markers data. For qualitative traits, the presence or absence of the different variants was coded as 1 or 0, respectively (Appendix IV). In the case of the quantitative traits, data were transformed to binary data (Appendix V) by using a statistical significance test after running an Analysis of Variance (ANOVA) on the original data.

5.2.2.2 DNA markers

5.2.2.2.1 AFLP's

The generation of AFLP markers is described in Chapter Four, Section 4.2.3. AFLP analysis was performed using GeneScan® software. Only clear and unambiguous bands were included in the analyses. AFLP fragments larger than or equal to 60 bp with a peak height above or equal to 45 RFUs were scored.

5.2.2.2.2 Microsatellites (SSR's)

Ten microsatellite sorghum primer pairs were used in this study. A standard PCR method was used to amplify microsatellites, the detail is described in Chapter Four, Section 4.2.4. The result of the PCR amplification was analyzed by electrophoresis on a 2% agarose gel (Molecular Screening agarose Roche) in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) run at 80V for 2.5 hours. Amplified fragments were visualised and sized using the Gel Doc 1000^{TM} image analysis system (Biorad) after ethidium bromide (0.5 ug / ml) staining.

5.2.3 Data analysis

A combined binary matrix included the morpho-agronomical matrix with the AFLP and SSR data matrices. The NCSS computer program (NCSS, 2000) was used for all data analyses. The dissimilarity values were used to graphically represent genetic relationships between the accessions by the unweighted pair group method using an arithmetic average (UPGMA) clustering algorithm the NCSS computer programme (NCSS, 2000), and relationships between accessions visualised as dendrograms.

5.3 Results and discussion

5.3.1 Level of polymorphism detection

The levels of polymorphism detected by the different marker approaches showed wide differences (Table 5.1). The eight AFLP primer pairs used in this study generated 552 polymorphic bands (average = 81 per pair), and the 10 microsatellite loci used produced 43 polymorphic bands (4.4 per pair) across 45 sorghum accessions. The highest polymorphism level was obtained for AFLP markers compared to microsatellite and morpho-agronomical markers. These can be probably attributed to the small data set for SSR and morpho-agronomic techniques compared to the AFLPs. Polymorphism detection efficiency among sorghum accessions by AFLPs and SSRs compared favourably with other available marker systems. Yang *et al.* (1996) detected 55%, 25%, 44% polymorphic bands for RFLP, RAPD, and ISSR techniques, respectively, in a selection of 34 Chinese sorghums.

5.3.2 Genetic diversity estimation

The summary of genetic distance (GD) coefficients calculated using Euclidean distance type from the individual marker data matrices, pair of data matrices and from the combined (AFLP, SSR and morpho-agronomic traits) data matrix is shown in Table 5.1. The GDs were estimated using 552 polymorphic AFLP fragments, 43 SSR polymorphic alleles, and 26 morpho-agronomical traits with 96 variants. The average GD estimates were 0.615 (for AFLP), 0.604 (for SSR), 0.566 (for morpho-agronomical traits) and 0.609 (for the combined data set). Range-wise, AFLP data produced lower (0.413 to 0.745) GD estimates compared to SSR (0.152 to 0.762), but the average GDs were very close in both. Furthermore, the genetic diversity estimates resulted with AFLP and microsatellite data sets were comparable despite the differences in the number of markers generated by each technique. An example of the pair-wise GDs estimates for the three markers were shown in (Appendices VI to VIII).

Table 5.1 Number of polymorphic bands, average and range of pair-wise genetic distance (GD) estimates among 45 sorghum accessions based on AFLP, SSR, morpho-agronomical traits data.

Marker system	No. of polymorphic markers	Average	Range
AFLP	552	0.615	0.413-0.745
SSR	43	0.604	0.152-0.762
Morpho-agro	96	0.566	0.354-0.707
AFLP + SSR	595	0.615	0.423-0.741
Combined	691	0.609	0.426-0.723

GD estimates can be affected by several factors such as, the distribution of markers in the genome (genome coverage) and the nature of evolutionary mechanisms underlying the variation measured (Powell *et al.*, 1996). AFLPs are believed to detect mainly point mutations while SSRs are specific to hypervariable loci (Giancola *et al.*, 2002). Another important factor is the influence of individual loci used for the analysis. While the SSR loci were based on availability, AFLP loci were randomly distributed whereas morpho-agronomical traits were selected.

Correlations between genetic distance matrices based on the two molecular marker techniques and the morpho-agronomical traits were significant (Table 5.2). In a similar study, Tatineni *et al.* (1996) reported a high correlation between RAPD and morphological characters. However, a lower correlation between AFLP and SSR genetic distance estimates has been reported (Giancola *et al.*, 2002). Powell *et al.* (1996) also reported that SSR similarity estimates were not significantly correlated to RFLPs, RAPDs, or AFLPs in soybean. Renganayaki *et al.* (2001) indicated that if comparisons are restricted within a species, then overall correlations between marker systems are significantly lower.

Table 5.2 Correlation coefficients between genetic distance values estimated for the three marker techniques (AFLP, SSR and morpho-agronomical traits), with sample size of 990.

	AFLPs	SSRs
SSRs	0.281**	
Morpho-agronomical traits	0.084*	0.188**

*, ** significant at p = 0.05 and p = 0.001, respectively.

5.3.3 Clustering based on AFLP, SSR and morpho-agronomical markers

Dendrograms resulting from UPGMA cluster analyses of morpho-agronomical, AFLP, SSR, AFLP and SSR, and combined data are illustrated in Figure 5.1a to e. On the whole, the dendrograms separated the 45 sorghum accessions into five to seven cluster groups. However, the output of each cluster tree was rather unique with some evident similarities; for instance, three of the five accessions with the name *Wagare* showed a high level of similarity across all the dendrograms, while the fourth and fifth were found to be quite different. In a similar study, Yang *et al.* (1996) found differences among RFLP, RAPD, and Inter-simple sequence repeat amplification (ISSR) markers in the discrimination of 34 Chinese sorghum lines. The lack of correspondence between levels of genetic diversity obtained from different techniques has already been reported (Djè *et al.*, 1999). Similarly, after studying sorghum accessions from Ethiopia and Eritrea, Ayana (2001) indicated a higher level of variation for morphological traits than those obtained for allozymes and RAPD markers. This situation, according to Yang *et al.* (1996), could have originated by contamination of the original material, by independent selection of very different inbreds from an originally diverse line, or by different materials receiving the same name. On the other hand, the similarity obtained in the dendrograms produced by the UPGMA analytical method proved the clustering to be acceptable.

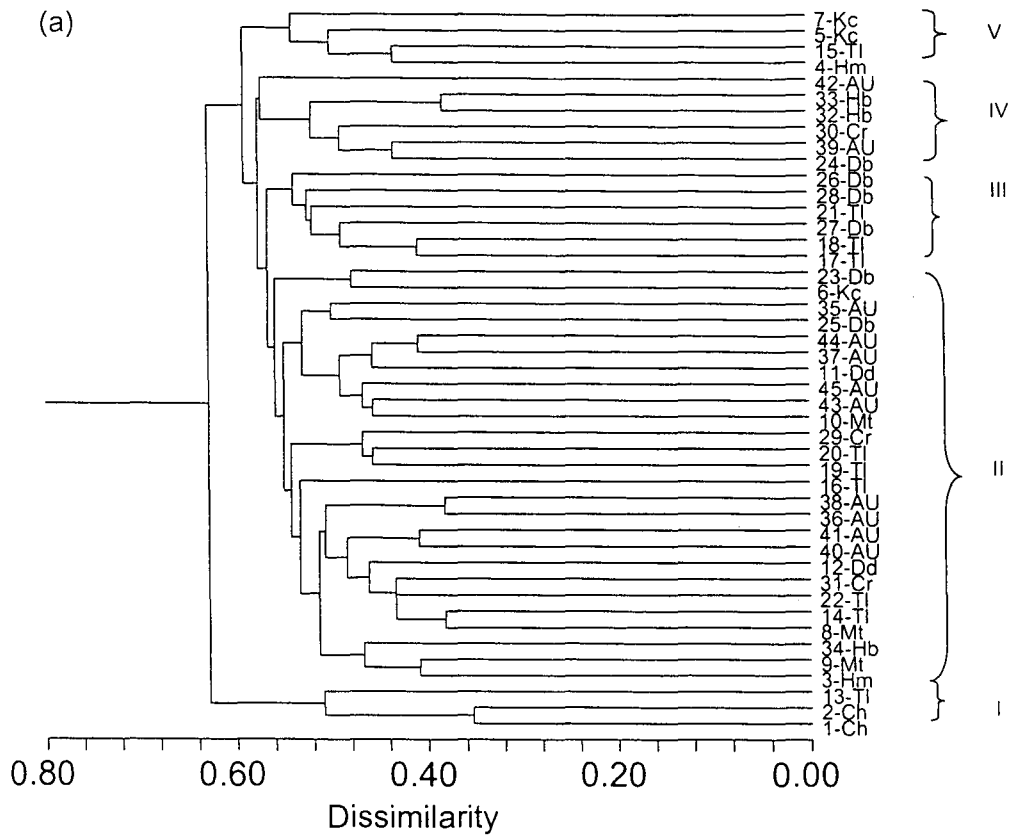
Clustering based on morpho-agronomical data analysis produced five clusters (Figure 5.1a), and the two accessions (1 and 2) from Chinhakssen and one (13) from Tulo were found the most different from the rest accessions. Cluster II contained the most accessions from different localities, nine from AU, five from Tulo, three each from Metta and Doba, two each from Deder and Chiro, and one each from Habro and Kurffa Challe. AFLP based clustering (Figure 5.1b) also resulted in five groups, where cluster II constituted the largest accessions from various localities. In cluster I, accession 41 was the most distinct. Cluster IV was formed by five accessions only from AU.

In the dendrogram constructed from SSR markers analysis (Figure 5.1c), it can be observed that there was more grouping based on the site of collection than observed in the morphological and AFLP based clustering. For instance, all accessions from Doba were grouped together with two Tulo and two Chiro collections. The groupings for accessions from Kurffa Challe and Habro also support the clustering together of accessions from the same site. But, AFLP and SSR based cluster analysis (Figure 5.1d) showed more dissimilarity among accessions from the same site. The clustering based on AFLP + SSR data analysis largely resembles that of AFLP data alone, and it is probably due to the large data point obtained from AFLPs when compared to SSRs. As the result greater degree of discrimination was obtained in AFLPs and AFLPs + SSRs data than SSRs alone. In addition, DNA markers may be affected by selection, drift, and mutation (Senior *et al.*, 1998). The same author further indicated that incongruities can result from the clustering process whenever clusters are non-overlapping.

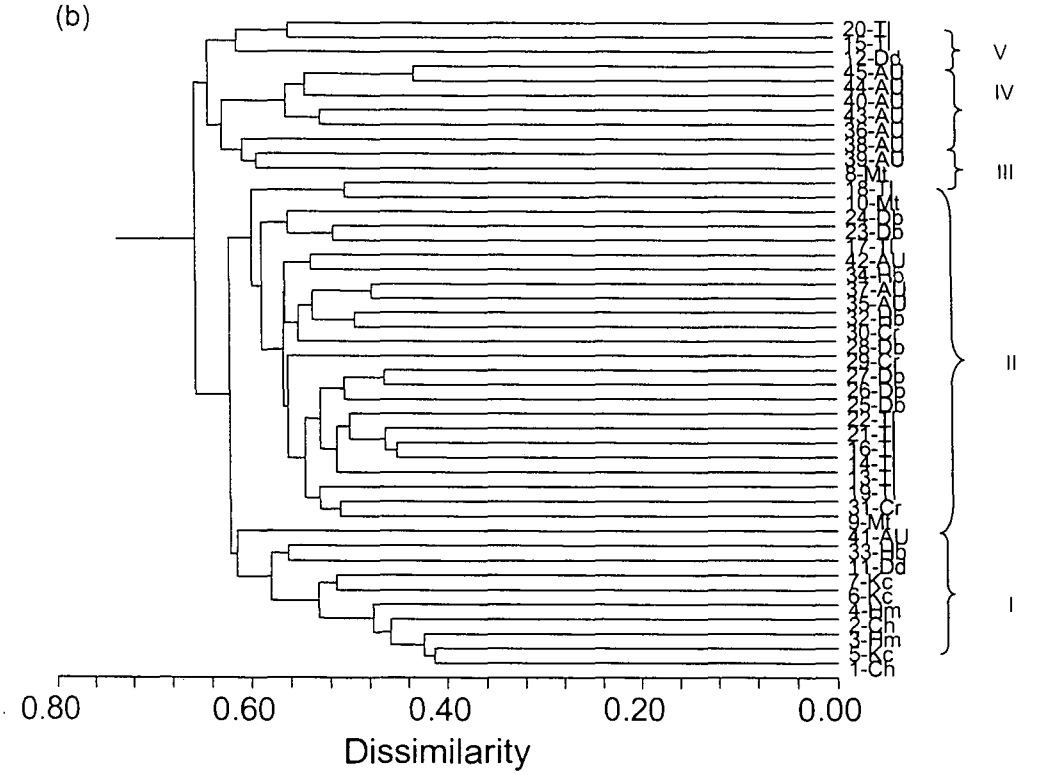
A clustering performed on the combined data is shown in Figure 5.1e. Cluster I contained a mixed group of accessions from Chinhakssen, Haro Maya, Habro and AU. All of the accessions from Doba were found in cluster II. Accession 11 and 12, both from Deder clustered separately in cluster III. Similarly, accessions 15 and 20 exhibited the most dissimilarity from the other Tulo accessions. Except accessions 41, 42 and 39 accessions from AU consistently showed closer relationships.

The lack of strong collection site based differentiation observed in cluster and principal component analyses could be partly ascribed to gene flow between the sites. The clustering together of accessions from the same sites is an indication of the evolution of co-adaptive association of quantitative characters (Zhong and Qualset, 1995). The consistent clustering of most breeding entries / improved cultivars from SIP/AU together in present study apparently substantiates this. Individual characters differ in their patterns of distribution and amount of variation. In general, the present results showed that by using the AFLP or SSR DNA technique, a large set of informative data could be generated in less time than with morpho-agronomical analysis. Also when simultaneously using DNA markers and morpho-agronomic traits to classify genotypes, it is possible to obtain a relevant minimum subset of marker-fragments that can be used in conjunction with available morpho-agronomic data to better classify genotypes compared to using only the quantitative or only the qualitative traits. The present results imply that although morpho-agronomical characterisation is influenced by the environment and is time consuming in general, among other disadvantages in relation to AFLP's and SSR's, it can still be an important and practical means of making progress in germplasm evaluation by conservationists and breeders.

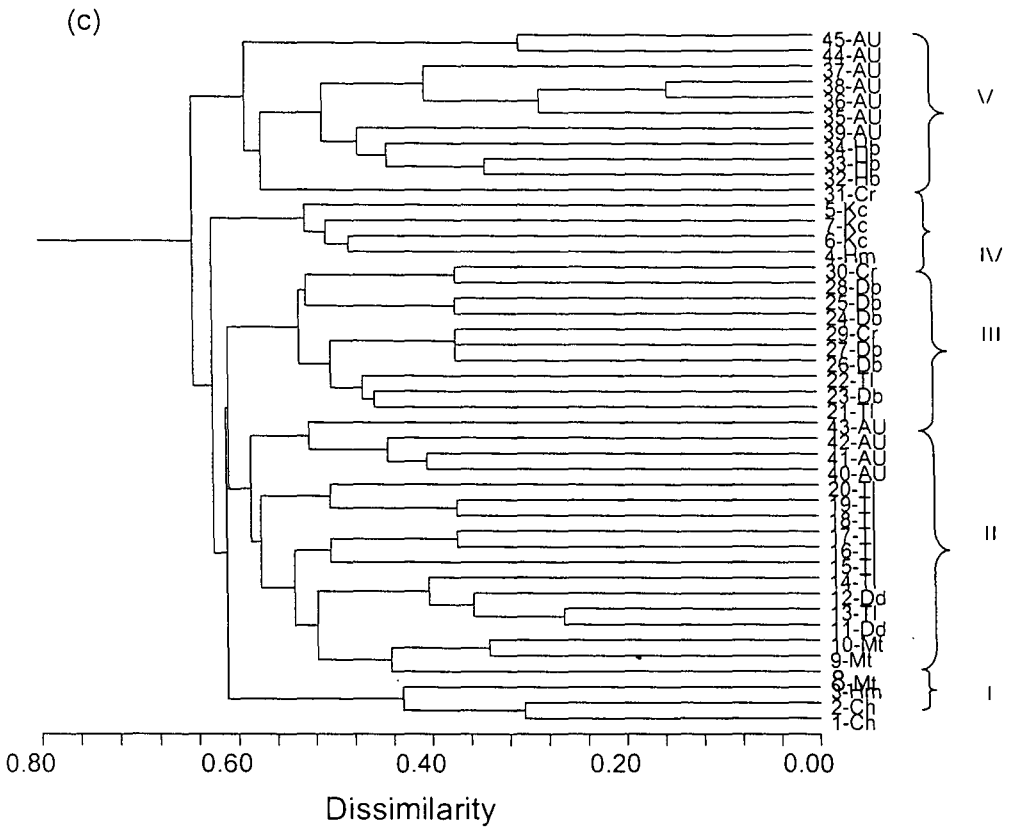
Morpho-agronomic



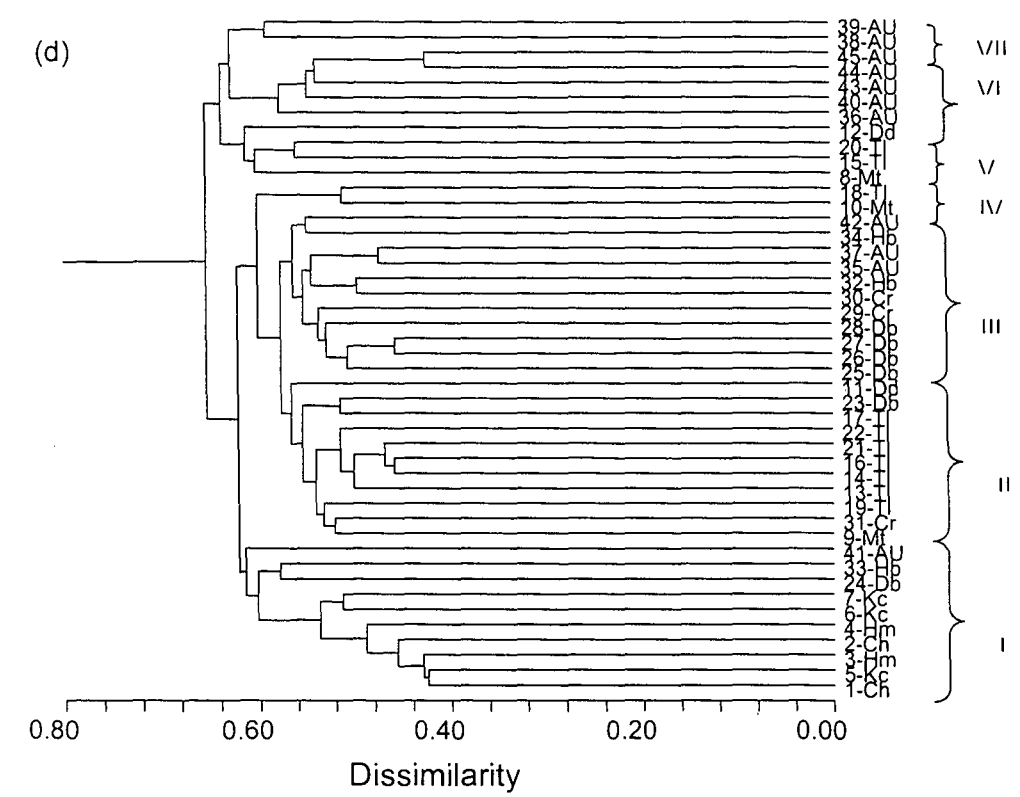
AFLP



SSR


(c)
45-AU
44-AU
37-AU
38-AU
36-AU
35-AU
39-AU
34-Db
33-Hb
32-Hb
31-Cr
5-Kc
7-Kc
6-Kc
4-Hm
30-Cr
28-Db
25-Db
24-Db
29-Cr
27-Db
26-Db
22-Tl
23-Db
21-Tl
43-AU
42-AU
41-AU
40-AU
20-Tl
19-Tl
18-Tl
17-Tl
16-Tl
15-Tl
14-Tl
12-Dd
13-Tl
11-Dd
10-Mt
9-Mt
8-Mt
3-Hm
2-Ch
1-Ch
V
IV
III
II
I
0.80
0.60
0.40
0.20
0.00
Dissimilarity

AFLP + SSR


(d)
39-AU
38-AU
45-AU
44-AU
43-AU
40-AU
36-AU
12-Dd
20-Tl
15-Tl
8-Mt
18-Tl
10-Mt
42-AU
34-Db
37-AU
35-AU
32-Hb
30-Cr
29-Cr
28-Db
27-Db
26-Db
25-Db
11-Dd
23-Db
17-Tl
22-Tl
21-Tl
16-Tl
14-Tl
13-Tl
19-Tl
31-Cr
9-Mt
41-AU
33-Hb
24-Db
7-Kc
6-Kc
4-Hm
2-Ch
3-Hm
5-Kc
1-Ch
VII
VI
V
IV
III
II
I
0.80
0.60
0.40
0.20
0.00
Dissimilarity

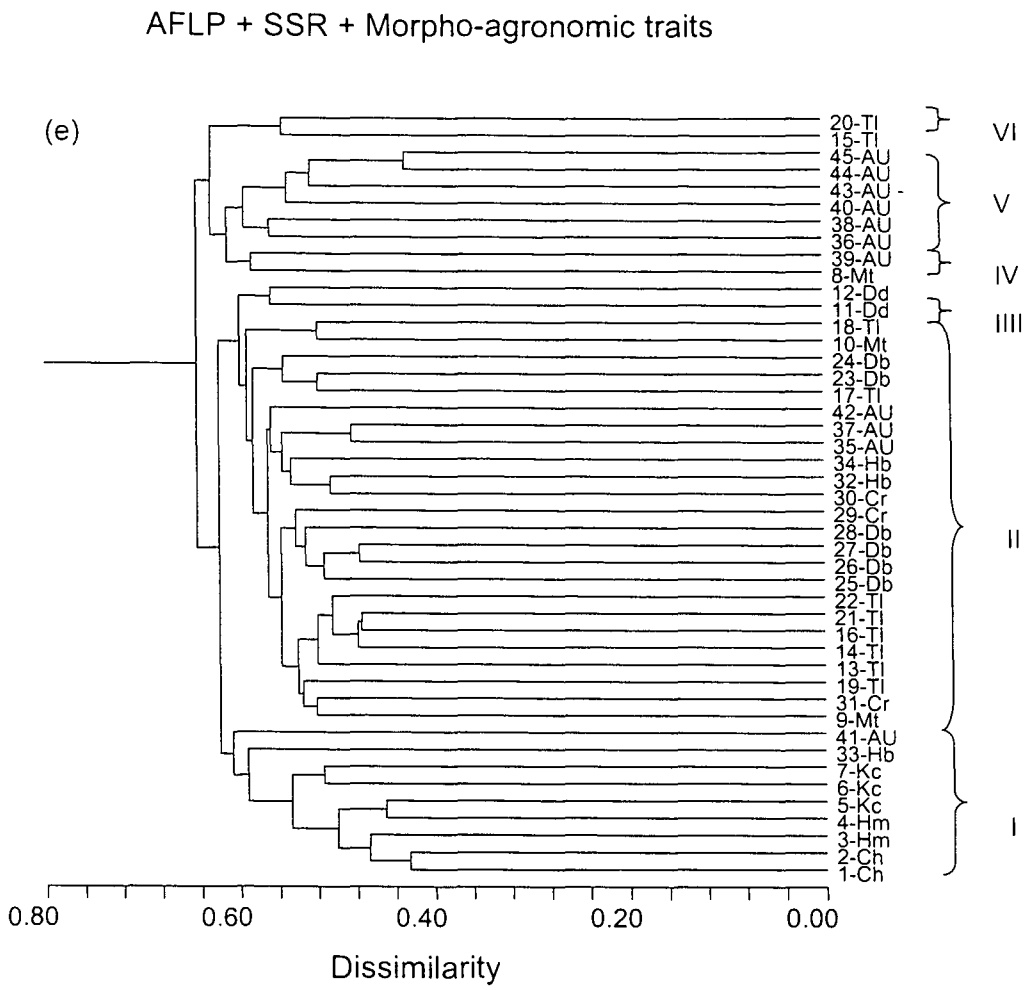


Figure 5.1 Dendrograms of 45 sorghum accessions constructed using dissimilarity matrix from (a) morpho-agronomic, (b) AFLP, (c) SSR, (d) AFLP + SSR and (e) combined data.

5.4 Conclusions

Characterisation of sorghum accessions at the DNA level can help identify genetically representative, non-redundant sets of germplasm for sorghum breeding and conservation purposes. As observed from significant correlation coefficients obtained between genetic distance values from the three marker techniques, all have shown a comparable genetic diversity level. AFLP markers were more powerful than SSRs in distinguishing accessions that were collected from the same site. Overall, from this result, using DNA-based markers clearly suggests that conventional methods have been effective in selecting unique collections, and further indicate that diversity assessed by molecular markers may efficiently represent the genetic diversity in morpho-agronomic traits. Although the relationships determined by molecular data are different to those identified by morpho-agronomical traits, the latter is still useful in assessing genetic diversity for the purpose of breeding selection on condition that the genotypes under investigation are not too closely related. Molecular techniques have a clear advantage over morpho-agronomical traits in elucidating complex relationships, especially of genotypes sharing morpho-agronomic traits or coming from the same geographic location.

REFERENCES

Ayana, A. 2001. Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. Ph.D. dissertation. Department of Biology, Science Faculty, Addis Ababa University, Addis Ababa, Ethiopia.

Brown, S.M., Hopkins, M.S., Mitchell, S.E., Senior, M.L., Wang, T.Y., Duncan, R.R., Gonzalez-Candelas, F. and Kresovich, S. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Genet. 93:190-198.

Central Statistical Authority (CSA). 2000. The Federal Democratic Republic of Ethiopia: Agricultural sample survey report on area and production for major crops, Statistical Bulletin No 227. Addis Ababa, Ethiopia, pp 16-68.

de Olivera, A.C., Richter, T. and Bennetzen, J.L. 1996. Regional and racial specificities in sorghum germplasm assessed with DNA markers. Genome 39:579-587.

Dean, R.E., Dahlberg, J.A., Hopkins, M.S. Mitchell, S.E. and Kresovich, S. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39:1215-1221.

Djè, Y., Forcioli, D., Ater, M., Lefebvre, C. and Vekemans, X. 1999. Assessing population genetic structure of sorghum landraces from north-western Morocco using allozyme and microsatellite markers. Theor. Appl. Genet. 99:57-163.

Djè, Y., Heuertz, M., Lefébvre, C. and Vekemans, X. 2000. Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theor. Appl. Genet. 100: 918-925.

Franco, J., Crossa, J., Ribaut, J.M., Betran, J., Warburton, M.L. and Khairallah, M. 2001. A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. Theor. Appl. Genet. 103:944-952.

Giancola, S., Marcucci Poltri, S., Lacaze, P. and Hopp, H.E. 2002. Feasibility of integration of molecular markers and morphological descriptors in a real case study of a plant variety protection system for soybean. Euphytica 127:95-113.

Harlan, J.R., and de Wet, J.M.J. 1972. A simplified classification of cultivated sorghum. Crop Sci. 12:172-176.

Morell, M.K., Peakall, R., Appels, R., Preston, L.R., and Lloyd, H.L. 1995. DNA profiling techniques for plant variety identification. Australian J. Exp. Agric. 35:807-819.

NCSS, 2000. Number Cruncher Statistical Systems, Dr. Jerry L. Hintze, 329 North 1000 East, Kaysville, Utah 84037, Canada.

Perkin-Elmer. 1996. AFLP Plant Mapping Kit–the PCR marker of choice for plant mapping.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breed. 2:225-238.

Renganayaki, K., Read, J.C. and Fritz, A.K. 2001. Genetic diversity among Texas bluegrass genotypes (*Poa arachnifera* Torr.) revealed by AFLP and RAPD markers. Theor. Appl. Genet. 102:1037-1045.

Smith, J.S.C., Kresovich, S., Hopkins, M.S., Mitchell, S.E., Dean, R.E., Woodman, W.L., Lee, M. and Porter, K. 2000. Genetic Diversity among Elite Sorghum Inbred Lines Assessed with Simple Sequence Repeats. Crop Sci. 40:226-232.

Stemler, A.B.L., Harlan, J.R. and de Wet, J.M.J. 1977. The sorghums of Ethiopia. Econ. Bot. 31:446-460.

Subudhi, P.K., Nguyen, H.T., Gilbert, M.L. and Rosenow, D.T. 2002. Sorghum improvement: Past achievements and future prospects. In: Kang, M.S. (Ed.). Crop improvement: Challenges in the Twenty-First Century. The Haworth Press, Binghamton, NY, pp. 109-159.

Tatineni, V., Cantrell, R.G. and Davis D.D. 1996. Genetic diversity in elite cotton germplasm determined by morphological characters and RAPDs. Crop Sci. 36:186-192.

Teshome, A., Baum, B.R., Fahrig, L., Torrance, J.K., Arnason, T.J. and Lambert, J.D. 1997. Sorghum [*Sorghum bicolor* (L.) Moench] landrace variation and classification in North Shewa and South Welo, Ethiopia. Euphytica 97:255-263.

Yang, W., de Oliveira, A.C., Godwin, I., Schertz, K. and bennetzen, J.L. 1996. Comparsion of DNA marker technologies in characterizing plant genome diversity: variability in Chinese sorghums. Crop Sci. 36:1669-1676.

Zhong, G.Y. and Qualset, C.O. 1995. Quantitative genetic diversity and conservation strategies for an allogamous annual species, dasypyrum villosum (L.) Candargy (Poaceae). Theor. Appl. Genet. 91:1064-1073.

CHAPTER 6

PHYSICO-CHEMICAL ANALYSIS OF SORGHUM AND SENSORY EVALUATION OF *INJERA*

Abstract

Samples of 13 genetically diverse sorghum accessions were analysed for physical properties and chemical composition. The food (*Injera)* quality and the phenolic (condensed tannins) content of six sorghum samples were also analysed. The accessions showed a wide variation in protein (7.99 to 17.8%), lipids (2.52 to 3.72%), starch (51.88 to 85%), and amylose (12.30 to 28.38%) content. Grain weight ranged from 19.5 to 33.7 g/1000 seeds, and endosperm texture varied from intermediate to soft (completely starchy). Linoleic acid (18:2) and oleic acid (18:1) were found to be the major fatty acid constituents of sorghum lipids. Only a few significant correlations were obtained among the physical and chemical properties, indicating that these properties could not be predicted from other properties. The principal component analysis showed that protein and lipid contents, and endosperm texture largely contributed to grouping the accessions in PC1, and grain colour and amylose content in PC2.

Three out of the six accessions evaluated for sensory analysis, namely *Ambajeettee*, AL-70 and ETS 2752 were chosen for their desirable properties in *injera* making. The chemical and physical properties of the selected accessions were characterised by having high protein content, low tannins, intermediate endosperm texture, and white and yellow seed colours. Red sorghum and white sorghum with pigmented testa were found to be less desired. From the sensory analysis study, it was observed that two of the accessions chosen, *Ambajeettee* and ETS 2752 have shown a unique DNA profile and cluster group. Further investigation may be necessary to validate the result by replicating the trial over seasons and/or locations.

6.1 Introduction

Grain sorghum is one of the major crops grown for human consumption in arid and semi-arid regions of Africa and Asia. Asia, northern and Central America, and Africa contribute most of the world production of sorghum. Nigeria, Mali, Niger, Burkina Faso, Chad, Cameroon, Sudan, Ethiopia, Botswana, and Rwanda are the major African countries in which sorghum production is of critical importance (Subudhi *et al.*, 2002). According to Vavilov (1951) the Ethiopian area is considered as the centre of its origin; and the contribution of Ethiopian origin sorghums in world sorghum improvement is well recognised. For example, extensive use of the zerazera sorghums of Ethiopian origin in the U.S. hybrid sorghums has made major contributions to disease resistance, yield potential, and quality (Rosenow and Dahlberg, 2000). High lysine was initially found in two sorghum accessions from Ethiopia, IS 11167 and IS 11758 (House *et al.*, 1995). In Ethiopia as a whole, sorghum ranks fourth, next to maize (*Zea mays* L.), *tef* [*Eragrostis tef* (Zucc.) Trotter] and wheat (*Triticum spp.*), in total production and area under cultivation, and represents 13% of the total cereal production (CSA, 2000). In the eastern regions of Ethiopia, sorghum is the most important food crop (accounting for 45% in total area and production of cereals). It is traditionally cultivated by small-scale farmers, and is used either alone or in mixtures with other grains for different types of food preparations.

Injera, fermented pancake-like bread from sorghum or *tef* flour is the staple diet in Ethiopia and Eritrea. It is well known that *injera* quality depends mainly on cultivar and fermentation process. From the point of view of the nutritive value of sorghum grain, high tannin is the primary nutrient-limiting component (Dendy, 1995). A wide variability among some Ethiopian sorghum cultivars on their performance for *injera* making was reported (Wuhib and Tekabe, 1987; Gebrekidan and Gebrehiwot, 1982). However, a detailed understanding of the physical and chemical basis for cultivar differences in sorghum *injera* quality is lacking. Furthermore, previous workers (Murty and Kumar, 1995; Gebrekidan and Gebrehiwot, 1982) indicated that more information is required on the type of sorghums suited for *injera* preparation.

For industrial use and various food products, many quality parameters have been determined (Rooney *et al.*, 1997). Information on nutritional and quality characteristics of sorghum of the region, in terms of genetic control and the influence of environmental factors, need to be studied. With enhanced characterisation and improved end-use utilization, sorghum could maintain, or increase, its contribution to the region's food security in particular, and world agriculture in general. In most cases, since food quality varies by region and use, quality criteria have been somewhat difficult to define and therefore to use (Subudhi *et al.*, 2002). Use of multiple-trait selection criteria may enhance development of genotypes with a combination of desirable physical and chemical kernel attributes to improve sorghum *injera* quality.

Exploration of the available genetic variation in landraces and improved cultivars for chemical and physical grain attributes and their association with end-uses, such as *injera* quality, would require the screening of germplasm for quality evaluation before subsequent inclusion in breeding programs. The objectives of the study in this chapter were (1) to assess variability among the accessions for physical and chemical parameters, and (2) to understand relationships between the different physical and chemical parameters and sorghum *injera* quality.

6.2 Materials and methods

6.2.1 Selection of sorghum accessions

Thirteen of the sorghum accessions used for morpho-agronomical traits and molecular markers, based on genetic diversity analyses, were chosen for the study (Table 6.1). The selection was based on representing accessions from the clustering result formed from DNA (AFLPs and SSRs) data. In addition to the clustering result, consideration was also given to include accessions with variable physical traits (namely, grain colour, 1000-grain weight and endosperm texture). Data generated under sections 3.2.1.2 (on grain colour and endosperm texture) and 3.2.2.3 (on 1000-kernel weight) were used.

Table 6.1 List of sorghum accessions included in the chemical composition determination.

№	Local / cultivar name	1000 Kernel weight (g)	Grain colour	Endosperm texture
1	*Wagare*	29.0	Light brown	9[t]
2	*Fandisha*	23.3	Yellow	5
3	*Muyra*	33.7	Red	7
4	*Ambajeette*	25.0	Yellow	5
5	*Suuta naqaphu*	24.3	White	7
6	*Zangada*	19.5	Dark brown	9
7	ETS 721	25.0	Light brown	9
8	*Wotet begunche*	20.3	Red	7
9	AL-70	30.0	White	5
10	ETS 2752	30.0	White	5
11	ETS 1005	31.0	Red	7
12	ETS 576	33.0	White	7
13	Long *muyra*	28.7	Red	9

[t] Indicates endosperm texture (5 = intermediate; 7 = mostly starchy; and 9 = completely starchy).

6.2.2 Methods

Kernels were cleaned manually and ground in a small Braun sample mill, and the flour was stored at 4°C. Analysis was conducted in triplicate, and averaged results were expressed on a dry matter basis.

6.2.3 Chemical composition

6.2.3.1 Moisture content

Moisture content was determined in a Memmert UL 80 drying oven, according to the method described by Gomez *et al.* (1997). Pre-weighed samples (2 g) were dried at 110°C for 5 hr and re-weighed.

6.2.3.2 Protein content

Flour sample was weighed, oven-dried and protein content was determined by a combustion method (Leco ®, model FP-528, St. Joseph, MI) in the Nutritional Laboratory, Department of Animal, Wildlife and Grassland Sciences, University of the Free State, South Africa.

6.2.3.3 Lipid extraction and methylation

Lipids were extracted by using chloroform: methanol (2:1 v/v) with 0.001% BHT from 1 g samples of milled flour (Folch *et al.*, 1957) in the Lipid Chemistry Laboratory, Department of Microbial, Biochemical and Food Biotechnology, University of the Free State.

Methylation. Fatty acid methyl esters were prepared using BF_3 – Methanol (14%), according to Slover and Lanaza (1979).

6.2.3.4 Fatty acid analysis

Fatty acids were quantified by flame ionisation gas chromatography (Varian GX 3400), with a Chrompack CPSIL 88 fused silica capillary column (100 m length, 0.25 μm ID, 0.2 μm film thickness). Identification of sample fatty acids was made by comparing the relative retention times of fatty acid methyl ester peaks from samples with those of standards obtained from SIGMA (cat. no. 189-19).

6.2.3.5 Starch content

The starch was hydrolysed enzymatically according to Horn *et al.* (1992). To solubilize the starch, the procedure proposed by Thomas *et al.* (1996) was followed. Accordingly, about 200 mg of the homogenized sample was weighed into a 100 ml Erlenmeyer flask, and added 20 ml dimethyl sulphoxide (DMSO) and 5 ml hydrochloric acid (HCl, 8 mol/l). This was followed by the addition of 5 ml sodium hydroxide (NaOH, 8 mol/l). The Erlenmeyer flask was covered with parafilm and the sample was incubated for 30 min at 60°C in a water bath. After immediate cooling in a cold water bath, 500 μl of the sample (aliquot of the culture supernatant) was taken and adjusted to a pH of 6.0 by adding 1.25 ml of 0.5 M citrate-phosphate (pH 6.6) buffer plus 0.5ml of 0.5 M citrate buffer (pH 4.5), and 10 μl of a thermo-stable α-amylase was added and boiled for 10 min. Then 1 ml of 0.5 M citrate buffer (pH 4.5) was added, followed by 10 μl of glucoamylase and incubated for 15 min at 60°C. The mixture was then cooled in a cold water bath and, the starch released (as glucose equivalent) was determined colorimetrically, using 3, 5 – dinitrosalicylic acid (DNS).

A standard glucose curve was obtained by diluting a 1 mg/ml pure glucose solution to 0.25, 0.5, 1.0, 1.25 and 1.5 mg glucose per ml. The standard curve determined with the glucose standards is shown in Figure 6.1.

Absorbance measurements were carried out as follows:

a) Standard glucose solution: One ml distilled H_2O was pipetted into a blank test tube, and into five other labelled test tubes, 1 ml of each glucose solution (0.25 to 1.5 mg/l). Then 1 ml DNS reagent and 2 ml water were added to each tube. Absorbance of each solution was determined at 540 nm.

b) Hydrolysate: One ml of hydrolysate prepared was pipetted into a test tube and 2ml water and 1ml DNS reagent was added.

c) Standards and hydrolysates: To allow the reaction between glucose and DNS to occur, all tubes were heated in a boiling water bath for 5 min. Then, the tubes were cooled down, and each volume was adjusted to 20 ml with distilled water and mixed well. The absorbance of each solution was then read at 540 nm. The percentage total starch was calculated as follows:

$$\text{Total starch (\%)} = \frac{(A_{540}/m) \times 20 \times 3.27 \times 60 \times 10}{\text{mass (mg)}} = (A_{540}/0.06) \times 39240$$

where, A_{540} = Absorbance at 540 nm, and m = slope.

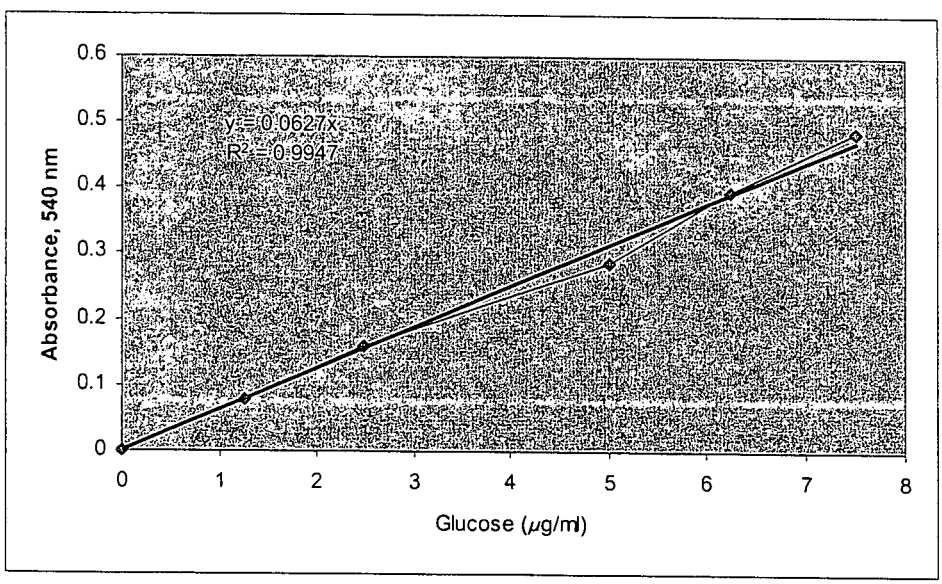


Figure 6.1 Standard graph for glucose content determined by duplicate data points.

6.2.3.6 Amylose content

Amylose content of starch was determined by the method of Knutson and Grove (1994). A standard curve was prepared with duplicate samples of amylose concentrations ranging from 5 to 20 μg/ml (dry weight basis). Absorbance was measured at 600 nm and the concentrations were used to develop the amylose standard curve (Figure 6.2). Finally, percentage amylose was calculated using a formula,

$$\text{Amylose\%} = \frac{A_{600} \times 9.1 \times 55 \times 100}{20.60 \times \text{mass (mg)}} = \frac{A_{600} \times 2429.61}{\text{mass (mg)}}$$

where, A_{600} = Absorbance at 600 nm.

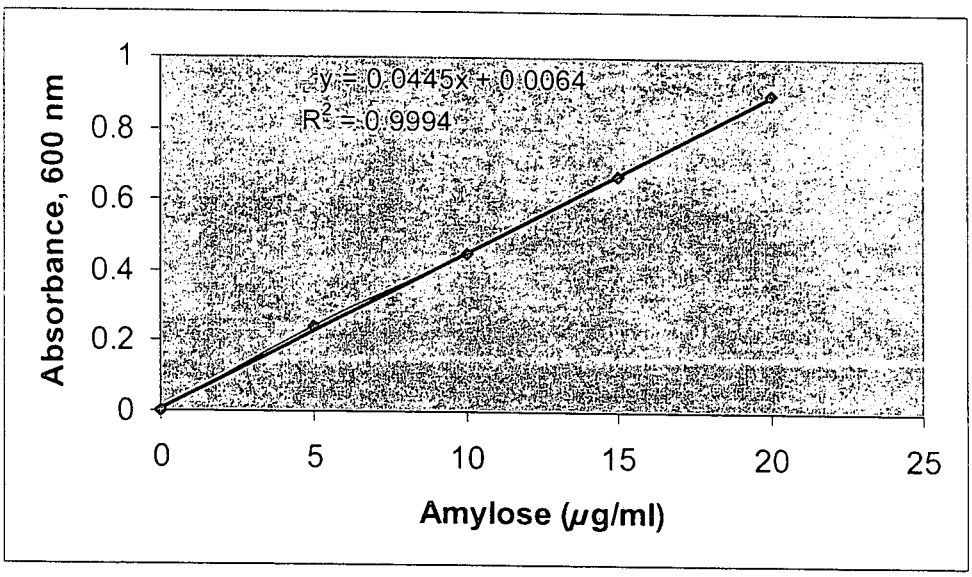


Figure 6.2 Standard amylose curve determined by duplicate data points.

6.2.3.7 Polyphenol / Tannin content

The polyphenol (condensed tannin) content was analysed for six accessions, which were also included in the sorghum *injera* sensory evaluation. A modified phenolic compounds determination method (Menkir *et al.*, 1996) was used. A 250 mg flour sample was extracted in 15 ml of 0.5% HCl in methanol for 20 min. The suspension was centrifuged for 5 min at 5000 rpm and the supernatant was

carefully removed. A 1 ml aliquot was taken from the supernatant and mixed with 14 ml of 30% HCl in 1-butanol. A blank was prepared by mixing 0.5 ml of acidic methanol extract with 7 ml of methanol, 0.1N acetic acid and butanol (15:15:70, v/v). The sample and the blank were vortexed and left to stand for 1 hr. Test tubes containing the supernatant were then maintained in boiling water for 2 hr. After cooling at room temperature and correcting for blanks, the concentrations of tannins (proanthocyanidins) were measured at 550 nm. Results of the concentrations were expressed as ug/g of dry sample.

6.2.4 Sensory evaluation of *Injera*

6.2.4.1 Sorghum samples

Six out of the 13 accessions (Table 6.1) studied for the chemical composition analysis were selected and used for the evaluation. The selection was done mainly on the basis of the cluster groups formed by using DNA marker based data. In addition, their chemical components were considered together with grain colour variability and endosperm texture. The accessions used include, *Fandisha* (A), *Ambajeette* (B), ETS 1005 (C), ETS 576 (D), ETS 2752 (E) and AL-70 (F). The English letter in parenthesis was used to represent the accession in the analysis and discussion.

Samples were coded using three digit numbers picked from a table of random numbers. Evaluations were performed at room temperature (25 -30°C) in individual testing booths in the Sensory Evaluation Laboratory, Department of Microbial, Biochemical and Food Biotechnology, University of the Free State. Clear white lights were used to facilitate the evaluation of the colour.

6.2.4.2 Methods of sensory evaluation

After cleaning, samples were milled in a falling number mill, 8 mm screen at Bethlehem Grain Institute, Bethlehem, South Africa. *Injera* was made following the steps described by Umeta and Faulks (1988) with some modifications. *Injera* batter was prepared from the flour by mixing the necessary ingredients, and fermented for 48 hr. The breads were baked in a heated clay griddle/ *mitad* one day before testing and stored in glass containers, covered with cling film, to prevent it from drying out. For evaluation, 60 x 25 mm samples were cut and presented on white polystyrene trays.

Visual and taste quality characteristics of sorghum *injera* were evaluated using the preference ranking test (Basker, 1988). Fifty-one panellists (25 - 49 years of age) were recruited for the panel session from Ethiopian and Eritrean students (47 males; 4 females) presently enrolled at the University of the Free State, South Africa. A partial view of the *injera* evaluation practice is shown in Appendix IX.

Three visual and one tactile characteristic were evaluated along with taste. The visual characteristics included colour, eye quality (the presence, size and amount of coagulated bubbles on the surface of the *injera*) and underside appearance. Texture was evaluated inside the mouth as well as by manipulation by hand, since the *injera* is often used as eating utensils to soak up a *wat/* stew. Panellists assigned ranks by using a scale of 1 to 6, with 1 = like/prefer the most and 6 = like/prefer the least. No ties were allowed, i.e., no two samples were awarded the same numerical value. Tap water at room temperature was provided for rinsing the palate between samples during taste sessions.

The values for each sample were added to obtain a rank sum for each sample. Differences in the preference rank sums between all possible pairs of samples were calculated and considered according to Basker (1988).

6.2.5 Statistical analyses

Univariate analyses

All analyses were performed in triplicate and data from individual chemical component experiments were analysed by analysis of variance (ANOVA) using the ACB procedure of Agrobase (Agrobase, 2000). The least significant difference (LSD) test was used for mean separation where the F-test was found significant for accessions. Pearson's correlation coefficients were calculated to determine relationships between the physical, chemical and *injera* properties.

Multivariate analysis

Principal component analysis (PCA) was computed from the correlation matrix generated from the chemical and physical properties using an NCSS Program (NCSS, 2000). The first and second PCs were used to show the patterns of the data.

6.3 Results and discussion

6.3.1 Chemical composition

The average moisture content of the 13 sorghum samples included in the study was 10.31±0.46%. The average values and LSD test for content of protein, lipid, starch and amylose are presented in Table 6.2.

Protein content varied between 7.99 and 17.80% and was within the range cited in the literature (Lasztity, 1996; Martin, 1984; Stemler *et al.*, 1976). The protein content was highest for AL-70 (17.8%), followed by *Ambajeette* (15.88%). *Wotet begunche*, a high lysine content sorghum, was also high in protein content. The lowest protein content was obtained in ETS 1005, ETS 576, and Long *muyra*; all three are of the *Muyra* sorghums, indicating the probability of genetic influence. In his review Làsztity (1996) reported that the protein content of sorghum varies from 6 to 25%. This wide variation in protein content and composition was indicated to be due to genotype, water availability, soil fertility, temperature, and environmental conditions during grain development. Nevertheless, according to (Hulse *et al.*, 1980) genotypic effects on protein content were greater than environmental effects.

Among the accessions evaluated, the total starch content ranged between 52 and 85%. It can be seen that the starch content of ETS 2752, AL-70 and *Fandisha* was the highest. *Zangada* and Long *muyra* had the lowest starch content and can be classified as hetrowaxy. Previous workers (Buffo *et al.*, 1997; Klopfenstein & Hoseney, 1995; Làsztity, 1996; Wankhede *et al.*, 1989) have reported that starch makes up to 60 to 80% of normal (non-waxy) sorghum kernels.

Amylose content was the highest for accessions ETS 2752, ETS 721, *Suuta naqaphu*, AL-70, and ETS 576. In contrast, *Zangada*, *Wotet begunche* and Long *muyra* had the lowest amylose content. The result obtained in this study is within the range reported for sorghum by Serna-

Saldivar and Rooney (1995). According to Ring *et al.* (1982), most of the accessions in this study can be grouped as non-waxy (contain 23 to 30% amylose) type sorghums, and some heterowaxy ones, but no waxy types (those that contain up to 5% amylose). It was indicated in the literature (Taylor *et al.*, 1997; Ring *et al.*, 1982) that environmental and genetic factors determine amylose levels in sorghum.

Table 6.2 Average values and mean separation of the chemical compositions for 13 sorghum accessions.

Local / cultivar name	Protein	Lipid	Starch	Amylose
	%			
Wagare	11.18^{ft}	3.17^{cd}	73.68^{e}	20.71^{de}
Fandisha	12.10^{de}	2.52^{e}	81.60^{ab}	24.15^{c}
Muyra	11.67^{def}	3.22^{cd}	77.51^{bc}	20.97^{d}
Ambajeette	15.88^{b}	3.00^{cd}	71.99^{e}	19.63^{de}
Suuta naqaphu	11.17^{f}	3.22^{cd}	65.83^{e}	27.01^{ab}
Zangada	12.19^{d}	2.97^{d}	51.88^{f}	12.30^{f}
ETS 721	11.46^{ef}	3.34^{bc}	71.51^{d}	27.45^{ab}
Wotet begunche	15.55^{bc}	2.97^{d}	74.59^{cd}	14.55^{f}
AL-70	17.80^{a}	2.94^{d}	82.73^{ab}	26.65^{abc}
ETS 2752	14.97^{c}	3.09^{cd}	85.01^{a}	28.38^{a}
ETS 1005	7.99^{h}	3.66^{ab}	82.22^{ab}	25.04^{bc}
ETS 576	10.40^{g}	3.72^{a}	63.27^{e}	26.17^{abc}
Long *muyra*	10.37^{g}	3.66^{ab}	53.44^{f}	17.96^{e}
Average	12.52	3.19	71.94	22.38
LSD (P= 0.01)	0.64	0.36	5.46	2.84
CV (%)	2.50	5.59	3.77	6.24

† Means for each trait factor followed by different superscript letters are significantly different at $p< 0.01$.

The lipid content ranged from 2.52 to 3.72%, with an average of 3.19%. ETS 576 had the highest lipid content, but not significantly different from

ETS 1005 and Long *muyra*. On the other hand, *Fandisha* was found to be the lowest container of lipid. Results, which are similar (a lipid content that ranged between 2.1 to 5.0%) to the results obtained in the present study, were reported in previous studies (Hoseney, 1994; Beta *et al.*, 1995; Yang and Seib, 1995). Table 6.3 shows the composition of fatty acids of 10 sorghum accessions.

Table 6.3 Fatty acid compositions (%) of sorghum accessions.

Local/ cultivar name	16: 0	18: 0	18: 1	18: 2	18: 3
Wagare	11.92^{eft}	1.15^{f}	29.23^{h}	53.65^{a}	2.50^{b}
Fandisha	16.34^{a}	1.41^{a}	29.88^{g}	46.13^{f}	4.52^{a}
Muyra	12.17^{de}	1.33^{bc}	31.02^{f}	52.69^{b}	1.37^{de}
Ambajeette	12.18^{de}	1.19^{ef}	29.60^{gh}	53.67^{a}	1.83^{c}
Suuta naqaphu	11.68^{f}	1.24^{de}	34.35^{e}	49.67^{c}	1.48^{d}
Zangada	13.62^{b}	1.37^{ab}	43.82^{a}	36.28^{h}	2.53^{b}
ETS 721	12.25^{de}	1.27^{cd}	36.56^{c}	47.12^{e}	1.23^{ef}
ETS 1005	12.45^{cd}	1.44^{a}	35.64^{d}	47.83^{d}	1.23^{ef}
ETS 576	12.69^{c}	1.41^{a}	38.26^{b}	44.79^{g}	1.32^{e}
Long *muyra*	12.49^{cd}	1.33^{bc}	36.91^{c}	46.47^{ef}	1.17^{f}
Average	12.78	1.32	34.53	47.83	1.92
LSD (p = 0.01)	0.41	0.07	0.60	0.65	0.14
CV%	1.55	2.64	0.83	0.65	3.56

† Means for each trait factor followed by different superscript letters are significantly different at $p< 0.01$.

The major fatty acid was linoleic acid (18:2), followed by oleic acid (18:1) and palmitic acid (16:0). Serna-Saldivar and Rooney (1995) have indicated that the fatty acid composition of sorghum oil is similar to that of maize and is dominated by linoleic and oleic acids. Accession *Zangada* contained oleic acid (18:1) as its major fatty acid constituent. The accession with the lowest lipid content, *Fandisha,* has an interesting fatty acid composition, with the highest content of palmitic acid (16:0) and linolenic acid (18:3). Linoleic, oleic, and palmitic acids were reported as the main constituents in

three species of small millets (Sridhar and Lakshminarayana, 1994). The average polyphenol (condensed tannins) content and mean separation for the six sorghum samples is presented in Table 6.4. Among the samples, *Ambajeette* (yellow seeded), ETS 2752 (white seeded) and AL-70 (white seeded) contained lower amount of tannins, while *Fandisha* (yellow seeded) showed a significantly higher tannin content than ETS 2752, but similar to AL-70. ETS 1005 (red sorghum) scored the highest tannin content, but statistically not different from ETS 576 (a white sorghum with pigmented testa).

In general, the result from this study and Menkir *et al.* (1996) showed that white sorghum accessions without a pigmented testa were characterised by low tannin content.

Table 6.4 Average tannin content and mean separation for six sorghum accessions.

Local / cultivar name	Tannin (%)
Fandisha	1.61^{bt}
Ambajeette	1.55^{c}
ETS 1005	1.83^{a}
ETS 576	1.81^{a}
ETS 2752	1.52^{c}
AL-70	1.58^{bc}
Mean	1.65
LSD (p = 0.05)	0.14
CV%	0.02

† Means followed by different letters are significantly different ($p < 0.05$).

The correlation coefficients among the different properties are presented in Table 6.5. There was a negative correlation coefficient between protein and lipids, and protein and endosperm (soft) texture. House *et al.* (1995) found a negative correlation between protein and lipid contents. A positive

correlation was obtained between starch and endosperm (soft) texture, lipids and 1000 seed weight (TSWt), soft endosperm texture (ET) and dark grain colour (GCl). The other attributes were not significantly related to one another; therefore, they could not be predicted from other properties. The correlation obtained between protein and lipids in this study disagrees with what House *et al.* (1995) found. In maize, lipid content was positively correlated ($r > 0.5$) with starch (Dorsey-Redding *et al.*, 1991).

Table 6.5 Correlation coefficients (n = 39) between chemical and physical properties for 13 sorghum accessions.

	Protein	Lipid	Starch	Amyolse	TSWt	ET
Lipid	-0.62*					
Starch	0.32	-0.35				
Amylose	-0.27	0.37	-0.05			
TSWt	-0.25	0.56*	0.34	0.39		
ET	-0.58*	0.47	0.67*	-0.01	-0.14	
GCl	-0.33	0.02	-0.39	-0.37	-0.37	0.76**

*, ** Significant at p = 0.05 and p = 0.01, respectively.

The PC analysis (Table 6.6) showed that the first three eigenvectors explained 85% of the total variance apparent among the 13 accessions. The first two PC axes accounted for 55% of the multiple variations among the accessions (Figure 6.3), indicating a high degree of association among the parameters studied. The first PC explained 30% of the gross variability, and this was due mainly to variations in lipids and protein content, and endosperm texture. Similarly, 27% of the overall variability of the accessions comes from the second PC that originated mainly from variations in grain colour and amylose content. Eigenvectors of the PC1 had large positive weights for protein, starch and amylopectin, large negative weights for lipids, endosperm texture and grain colour, and a. smaller, but positive weight for 1000 kernel weight (Table 6.6).

The PC analysis grouped the accessions into groups over the four quadrants (Figure 6.3). The different groups of accessions could be seen as under the right bottom quadrant (Long *muyra* and ETS 576) with similar protein, lipid, and amylose contents. The left bottom quadrant contained the red and brown accessions (ETS 1005, *Muyra*, ETS 721 and *Wagare*) with significant intra-cluster variability for protein, lipid and amylose contents. The accessions *Zangada* and *Wotet begunche* occupied the left top quadrant and both had similar lipid contents, but differed in their content of protein, starch and amylose. The top right quadrant contained two closely grouped yellow seeded accessions (*Ambajeette* and *Fandisha*) and three white sorghums (*Suuta naqaphu,* ETS 2752, and AL-70). ETS 2752 and AL-70 showed different protein contents; but they are statistically similar in other physical and chemical properties.

Table 6.6. Principal component (PC) analysis of eight physicochemical parameters in 13 sorghum accessions with eigenvectors, eigenvalues and proportion of variations explained by the first three PC axes.

Parameter	Eigenvectors		
	PC1	PC2	PC3
Protein	0.38	0.24	-0.35
Lipids	-0.31	-0.51	0.11
Starch	0.45	-0.11	0.40
Amylose	0.01	-0.49	-0.35
Amylopectin	0.43	0.06	0.51
1000 kernel wt.	0.13	-0.56	0.32
Endosperm texture	-0.48	0.06	0.26
Grain colour	-0.35	0.35	0.40
Eigenvalues	2.44	1.98	2.39
Individual%	30.52	24.74	29.83
Cumulative%	30.52	55.25	85.08

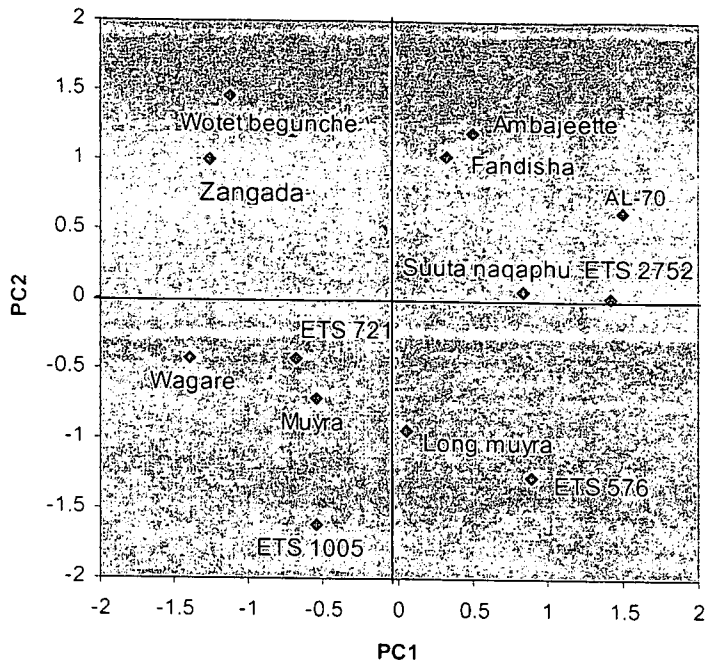


Figure 6.3 PC plots of 13 sorghum accessions analysed using eight physicochemical parameters.

6.3.2 Sensory evaluation of *injera*

Injera samples showing the different physical variations are shown in Figure 6.4 and Figure 6.5. The evaluation values for each sample were added to obtain a rank sum for each sample, and are presented from Table 6.7 through 6.11. Differences in the preference rank sums between all possible pairs of samples were calculated and considered, for example, between sample A and B (Table 6.7) it was 84, and between sample C and D it was 40. If any of these (absolute) differences exceeded a critical value, then the preferences for that pair of samples differed from one another at the stated statistical significance level.

Basker (1988) compiled a set of tables with critical values of differences among rank sums for multiple comparisons. From the significance tables, a significance level of $p = 0.05$ is attained when the rank sum differences are greater than or equal to 53.8 (for 51 panellists and 6 products), and a significance level of $p = 0.01$ is attained when the rank sum differences are greater than or equal to 63.6.

Table 6.7 Rank sums and significance tests for *injera* colour.

Sample		A	B	C	D	E	F
Rank sum		203	119	247	287	105	110
Difference vs	A		84	44	84	98	93
	B			128	168	14	9
	C				40	142	137
	D					182	177
	E						5
Significance level			p = 0.05			p = 0.01	
Critical difference			53.8			63.6	
Sample	E		α			α	
	F		α_β			α_β	
	B		α_β			α_β	
	A		β			β	
	C		γ			γ	
	D		δ			γ	

In the table above, the evaluation of the colour of the sorghum *injera*, the results are arranged with the products in decreasing order of preference, that is, in increasing order of rank sums. Sample E (ETS 2752) is put first, because it had the lowest rank sum and was preferred "the most". It is followed by sample F (AL-70), which had the second lowest rank sum, then by B (*Ambajeette*), A (*Fandisha*), C (ETS 1005) and finally by sample D (ETS 576), which had the highest rank sum and was preferred "the least". Lowercase Greek letters were used to indicate samples whose rank sums did not differ significantly. From this result it is clear that at both significance levels, sample E (ETS 2752) was significantly preferred over samples A, C and D, but not over samples F and B.

It can be seen from Table 6.8 that samples B (*Ambajeette*) and E (ETS 2752) had the same rank sums for 'eye' quality. At p = 0.05 samples B (*Ambajeette*), and E (ETS 2752) were significantly preferred over samples A (*Fandisha*), D (ETS 576) and C (ETS 1005), but not over sample F (AL-70). However, at p =0.01, samples B and E were significantly preferred only over samples D and C in regard to 'eye' quality, but not over samples F and A.

Figure 6.4 *Injera / bidena* prepared from two sorghum accessions, *Ambajeette* (A) and ETS 1005 (B) showing the colour variation and the distribution of the 'eyes.'

Table 6.8 The rank sums and differences between products along with the significance levels for 'eye' quality of *injera.*

Sample		A	B	C	D	E	F
Rank sum		192	138	220	215	138	168
Difference vs	A		54	28	23	54	24
	B			82	77	0	30
	C				5	82	52
	D					77	47
	E						30
Significance level				p = 0.05		p = 0.01	
Critical difference				53.8		63.6	
Sample	B			α		α	
	E			α		α	
	F			α_β		α_β	
	A			β		α_β	
	D			γ		β	
	C			γ_δ		β_γ	

For underside appearance (Table 6.9), sample E (ETS 2752) was significantly preferred over samples A (*Fandisha*), C (ETS 1005) and D (ETS 576) at both significant levels, but not over samples F (AL-70) and B (*Ambajeette*).

Table 6.9 Summarised results for underside appearance of *injera* from six sorghum samples.

Sample		A	B	C	D	E	F
Rank sum		203	160	216	249	119	124
Difference vs	A		43	13	46	84	79
	B			56	89	41	36
	C				33	97	92
	D					130	125
	E						5

Significance level		p = 0.05	p = 0.01
Critical difference		53.8	63.6
Sample	E	α	α
	F	α_β	α_β
	B	α_β	α_β
	A	β	β
	C	β_γ	β_γ
	D	γ	γ

When evaluating texture (Table 6.10), samples E (ETS 2752) and F (AL-70) scored the same rank sum totals, as can be seen from Table 6.10. Samples E and F were significantly preferred over samples C (ETS 1005), A (*Fandisha*) and D (ETS 576) at p = 0.05, but not over sample B (*Ambajeette*). However, at p = 0.01, samples E (ETS 2752) and F (AL-70) were significantly preferred over only samples A (*Fandisha*) and D (ETS 576).

Table 6.10 Summarised rank sums and significance tests for texture of *injera*.

Sample		A	B	C	D	E	F
Rank sum		196	158	188	265	129	129
Difference vs	A		38	8	69	67	67
	B			30	107	29	29
	C				77	59	59
	D					36	136
	E						0

Significance level		p = 0.05	p = 0.01
Critical difference		53.8	63.6
Sample	E	α	α
	F	α	α
	B	α_β	α_β
	C	β	α_β
	A	β_γ	β
	D	γ	γ

Table 6.11 Rank sums and significance tests for *injera* taste results.

Samples		A	B	C	D	E	F
Rank sum		195	150	222	237	131	133
Difference vs	A		45	27	42	64	62
	B			72	87	19	17
	C				15	91	89
	D					106	104
	E						2

Significance level		p = 0.05	p = 0.01
Critical difference		53.8	63.6
Sample	E	α	α
	F	α_β	α_β
	B	α_β	α_β
	A	β	β
	C	γ	γ
	D	γ_δ	γ_δ

The final characteristic evaluated was taste and the results are presented in Table 16.11. There was a significant statistical difference among accessions. ETS 2752 (E) was desired over *Fandisha* (A), ETS 1005 (C) and ETS 576 (D) at both significant levels, but not over accessions AL-70 (F) and *Ambajeette* (B).

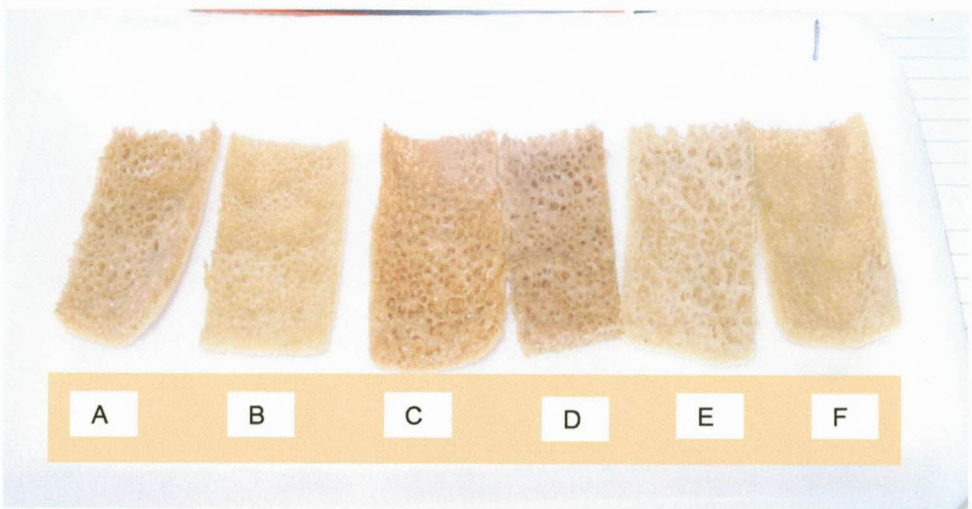


Figure 6.5 The six sorghum *Injera / bidena* samples evaluated by the panellists. (A = *Fandisha*, B = *Ambajeette*, C = ETS 1005, D = ETS 576, E = ETS 2752 and F = AL-70)

From the result, it can be said that *Ambajeette*, ETS 2752 and AL-70 appear to be preferred for making *injera*, since they scored the lowest rankings in all the tests. *Fandisha* was less preferred than the previous three, but more liked than ETS 1005 and ETS 576, which were found as least desired for *injera* making.

6.3.3 Relationships between physicochemical properties and *injera* quality

The physical properties included in the quality study were 1000 kernel weight, grain colour and endosperm texture. These traits were variable over the accessions and their relationships with chemical components (Table 6.5) and *injera* quality shown in Table 6.12.

Protein content positively and significantly ($r = 0.94$) correlated with *injera* quality. Sorghum *injera* quality was negatively correlated with tannin content ($r = -0.97$) and soft or floury textured endosperm ($r = -0.92$). In a previous study, it was reported that porridges made from flours of corneous sorghums were less sticky than those made from floury sorghums (Cagampang *et al.*, 1982). Stickiness of cooked sorghum flour was a function of starch gelatinisation (Hoseney, 1986). Starch in the corneous endosperm portion of sorghum grain is embedded in protein, which has been shown to influence starch gelatinisation in sorghum (Chandrashekar and Kirleis, 1988).

The three accessions that performed the best in the *injera* quality evaluation were all intermediate (heterowaxy) types. In relation to amylose content, high amylose sorghum products dry and become hard upon cooling. In contrast, the low amylose types are moist and sticky when cooled under optimum conditions. The intermediate types cook moist and tender and do not become hard upon cooling (Ring *et al.*, 1982).

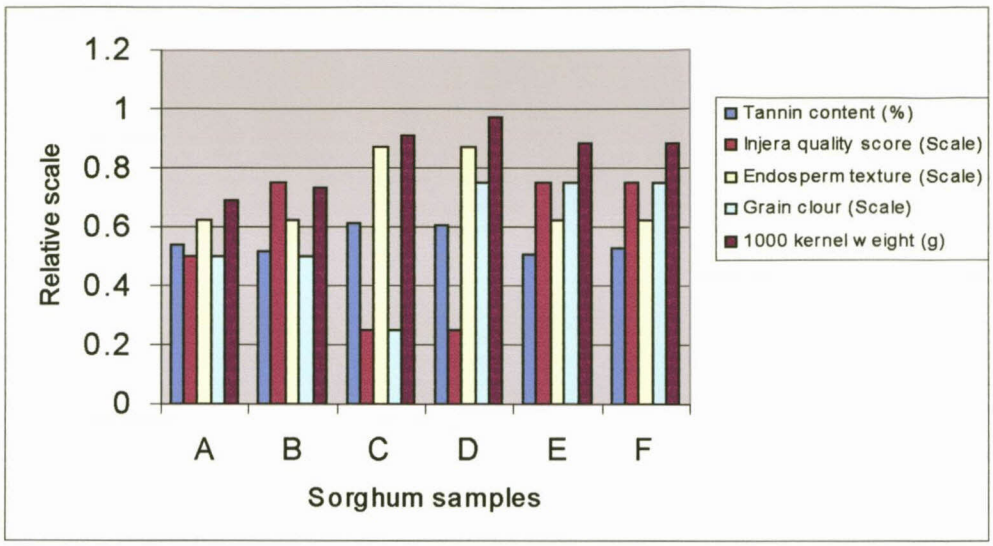


Figure 6.6 Bar graph showing the relationship between tannin content, *injera* quality, endosperm texture, grain colour and 1000-kernel weight in six sorghum samples. (A = Fandisha, B = Ambajeette, C = ETS 1005, D = ETS 576, E = ETS 2752 and F = AL-70)

Table 6.12 Correlation coefficients between *injera* quality and some physico-chemical parameters in sorghum.

Parameters	Correlation coefficient
Protein	0.94**
Lipid	-0.67
Starch	0.16
Amylose	-0.11
Tannin	-0.97**
1000 kernel weight	-0.42
Endosperm texture	-0.92**
Grain colour	-0.41

** Significant at p = 0.01 probability

6.4 Conclusions

Variability in physical and chemical composition was observed among sorghum accessions sampled. While the flours from all six accessions could be made into *injera*, the differences in chemical compositions and physical properties of the kernels resulted in differences in its quality. Of these, lower tannin and high protein contents, white or yellow seed colour, and intermediate endosperm texture were correlated highly with the quality of *injera*. Thus, future selection for best *injera* quality can combine high protein content, medium to high amylose content, intermediate endosperm texture, low tannin content attributes as effective screening criteria.

It is interesting to note that accessions 12 (*Ambajeette*), 40 (AL-70) and 41 (ETS 2752) had desirable *injera*-making characteristics. Data from morpho-agronomical traits grouped these accessions together. However, according to molecular data, these accessions were genetically distinct and were clustered in different genetic groups. This indicates that a good basis of genetic variability exists for the improvement of the sorghum cultivars using these accessions.

REFERENCES

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

Basker, D. 1988. Critical values of differences among rank sums for multple comparisons. Food Technology 42: 80-83.

Beta, T., Rooney, L.W. and Waniska, R.D. 1995. Malting characteristics of sorghum cultivars. Cereal Chem. 72:533-538.

Buffo, R.A., Weller, C.L. and Parkhurst, A.M. 1997. Optimisation of sulfur dioxide and lactic acid steeping concentrations for milling of grain sorghum. Transactions of the American Soc. of Agric. Engineers 40:1643-1648.

Cagampang, G.B., Griffith, J.E. and Kirleis, A.W. 1982. Modified adhesion test for measuring stickiness of sorghum porridges. Cereal Chem. 59:234-3-235.

Central Statistical Authority (CSA), 2000. The Federal Democratic Republic of Ethiopia: Agricultural sample survey report on area and production for major crops, Statistical Bulletin No 227. Addis Ababa, Ethiopia, pp 16-68.

Chandrashekar, A. and Kirleis, A.W. 1988. Influence of protein on starch gelatinisation in sorghum. Cereal Chem. 65:457-462.

Dendy, D.A.V. 1995. Sorghum and Millets: Chemistry and Technology, American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA.

Dorsey-Redding, C., Hurburgh, C.R.Jr., Johnson, L.A. and Fox, S.R. 1991. Relationships among maize quality factors. Cereal Chem. 68: 602-605.

Folch, J., Lees, M. and Sloane-Stanley, G.H. 1957. A simple method for the isolation and purification of total lipids from animal tissue. J. Biological Chem. 226: 497-509.

Gebrekidan, B., and Gebrehiwot, B. 1982. Sorghum *injera* preparations and quality parameters. In: Rooney, L.W. and Murty, D.S. (Eds.). Proc. Int. symp. on sorghum grain quality. ICRISAT, Patancheru, AP, India, pp. 55-66.

Gomez, M.I., Obilana, A.B., Martin, D.F., Madzvamuse, M. and Monyo, E.S. 1997. Quality evaluation of sorghum and pearl millet. ICRISAT, Patancheru, India.

Horn, C.H., du Preez, J.C. and Kilian, S.G. 1992. Fermentation of grain sorghum starch by co-cultivation of *Schwanniomyces occidentalis* and *Saccharomyces cerevisiae*. Bioresource Technology 42:27-31.

Hoseney, R.C. 1986. Principles of Cereal Science and Technology. In: American Assoc. of Cereal Chemists, St. Paul, MN, USA, pp. 47-54.

Hoseney, R.C. 1994. Principles of Cereal Science and Technology (Second edition). In: American Assoc. of Cereal Chemists, St. Paul, MN, USA, pp.82-98.

House, L.R., Osmanzai, M., Gomez, M.I., Monyo, E.S. and Gupta, S.C. 1995. Agronomic principles. In: Dendy, D.A.V. (Ed.). Sorghum and millets: Chemistry and technology. American Association of Cereal Chemists, St. paul, MN, USA, pp. 27-67.

Hulse, J.H., Laing, E.M. and Pearson, O.E. 1980. Sorghum and the millets: Their composition and nutritive value. Academic Press, NY.

Klopfenstein, C.F. and R.C. Hoseney. 1995. Nutritional Properties of Sorghum and Millets. In: Dendy, D.A.V. (Ed.). Sorghum and Millets: Chemistry and Technology. American Association of Cereal Chemists, St. Paul, MN, USA, pp.125-184.

Knutson, C.A. and Grove, M.J. 1994. Rapid method for estimation of amylose in maize starches. Cereal Chem. 71:469-471.

Lásztity, R. 1996. The chemistry of cereal proteins. Bocaraton, Fla.:CRC Press.

Martin, F.W. 1984. Handbook of Tropical Crops. CRC Press, Boca Raton, FL.

Menkir, A., Ejeta, G., Butler, L. and Melakeberhan, A. 1996. Physical and chemical kernel properties associated with resistance to grain mold in sorghum. Cereal Chem. 73:613-617.

Murty, D.S. and Kumar, K.A. 1995. Traditional uses of sorghum and millets. In: Dendy, D.A.V. (Ed.). Sorghum and millets: Chemistry and technology. American Association of Cereal Chemists, Inc., St. paul, MN, USA, pp. 185-223.

NCSS, 2000. Number Cruncher Statistical Systems, Dr. Jerry L. Hintze, 329 North 1000 East, Kaysville, Utah 84037, Canada.

Ring, S.H., Akingbala, J.O. and Rooney, L.W. 1982. Variation in amylose content among sorghums. In: Rooney, L.W. and Murty, D.S. (Eds.). International Symposium on Sorghum Grain Quality. ICRISAT, Patencheru, India, pp.269-279.

Rooney, L.W., Waniska, R.D. and Subramanian, R. 1997. Overcoming constraints to utilization of sorghum and millets. In Proceedings of the International Conference on Genetic Improvement of Sorghum and Pearl Millet, Sep. 23-27, 1996 (pp. 549-557). Lubbock, TX, INTSORMIL Publ. No. 97-5.

Rosenow, D.T. and Dahlberg, J.A. 2000. Collection, conversion, and utilization of sorghum. In: Smith, C.W. (Ed.). Sorghum: Origin, history, technology, and production. John Wiley & Sons, Inc. pp. 309-328.

Serna-Saldivar, and Rooney, L.W. 1995. Structure and chemistry of sorghum and millets. In: Dendy, D.A.V. (Ed.). Sorghum and millets chemistry and technology. American Association of Cereal Chemists, St. Paul, MN, USA, pp. 69-124.

Slover, H.T. and Lanza, E. 1979. Quantitative analysis of food fatty acids by capillary gas chromatography. Journal of the American Oil chemists Society. 56: 933-943.

Sridhar, R. and Lakshminarayana, G. 1994. Contents of total lipids and lipid classes and composition of fatty acids in small millets: Foxtail

(*Setaria italica*), Proso (*Panicum miliaceum*), and Finger (*Eleusine coracana*). Cereal Chem. 71:355-359.

Stemler, A.B.L., Collins, F.I., De Wet, J.M.J. and Harlan, J.R. 1976. Variation in the levels of lipid components and protein in ecogeographic races of sorghum bicolour. Biochem. Syst. Ecol. 4:43-45.

Subudhi, P.K., Nguyen, H.T., Gilbert, M.L. and Rosenow, D.T. 2002. Sorghum improvement: Past achievements and future prospects. In: Kang, M.S. (Ed.). Crop improvement: Challenges in the Twenty-First Century. The Haworth Press, Binghamton, NY, pp. 109-159.

Taylor, J.R.N., Dewar, J., Taylor, J. and von Ascheraden, R.F. 1997. Factors affecting the porrige-making quality of South African Sorghums. J. Sci. Food Agric. 73:464-470.

Thomas, K.C., Hynes, S.H. and Ingledew, W.M. 1996. Practical and theoretical considerations in the production of high concentrations of alcohol by fermentation. Process Biochem. 31:321-331.

Umeta, M. and Faulks, R.M. 1988. The effect of fermentation on the carbohydrates in tef (*Eragrostis tef*). Food Chemistry 27:181-189.

Vavilov, N.I. 1951. The origin, variation, immunity and breeding of cultivated plants. Chronica Botanica 13:1-366.

Wankhede, D.M., Deshpahde, H.W., Gunjal, B.B., Bhosale, M.M., Patil, H.B., Gahilod, A.T., Sawate, A.R. and Walde, S.G. 1989. Studies on physiochemical pasting characteristics and amyolytic susceptibility of starch from sorghum (*S. bicolor* (L.) Moench) grains. Starch/Staerke 41:123-127.

Wuhib, E. and Tekabe, F. 1987. Sorghum Utilization Project Final Report. Addis Ababa: ENI. (Proj. No. 3-P-80-0010)

Yang, P. and Seib, P.A. 1995. Low-input wet-milling of grain sorghum for readily accessible starch and animal feed. Cereal Chem. 72:498-503.

CHAPTER 7

GENERAL CONCLUSIONS

Molecular and morpho-agronomical marker techniques each have distinct advantages for assessing genetic relationships. Studies that combine the two approaches can thereby maximize both information content and usefulness. In this study, combinations of molecular and morpho-agronomical markers, and evaluation for chemical composition and food quality were undertaken to provide a comprehensive view of the genetic variation among sorghum accessions from eastern highlands of Ethiopia.

A total of 45 sorghum accessions were sampled and analysed for molecular and morpho-agronomical markers. The morpho-agronomic data was obtained by measuring 10 qualitative and 16 quantitative traits. Eight selective AFLP and 10 sorghum microsatellite primer pairs were used for DNA amplification generating a total of 651 and 48 bands, respectively. And also, a chemical composition determination and sensory analysis were conducted on 13 and six sorghum samples, respectively.

The phenotypic, chemical and molecular diversity estimates showed a high level of variation among accessions, and indicated that sorghum accession populations studied are a mixture of a large number of distinct genotypes. Based on qualitative data, the estimates of H', individually and pooled over characters and localities was highest for kurffa challe, AU and Chiro. The lowest H' were from Doba, Habro and Metta. The average GD estimates from the individual markers and combined data produced comparable results. Hence, it is possible to make predictions about the level of variation for one technique based on the level of the other.

From the cluster and / principal component analyses of the three sets of data and the combined data matrix, it was observed that though some regions are clustering together, perfect matching of the dendrograms or the plots is lacking, which could be due to many reasons, such as the nature of the genetic basis of variation for agro-morphological, AFLP and mocrosatellite markers, or may be related to the number of polymorphisms detected with each marker technique rather than a function of which technique is employed.

Furthermore, the grouping of the accessions into different clusters using the three different data sets and the combined data was not perfectly related to either their sites of collection or the same naming, suggesting the existence variability within the region of collection and same name. Overall, the results from these studies showed that it is possible to both classify the genetic diversity of the sorghum accessions for the highest genetic diversity using AFLP and microsatellite markers. However, AFLP based cluster analysis has more distinctly differentiated among the accessions than both morpho-agronomical and microsatellite based clusters. The difference between the levels of AFLPs and microsatellites in detecting distinctness could be related to the number of polymorphic bands/ loci analysed.

A significant variability was obtained among the sorghum samples analysed for chemical components and sensory evaluation, and this shows the potential to improve the traits through breeding. Protein and tannin contents influenced the *injera*-making quality significantly. Three accessions were preferred for *injera*-making, and two of these have shown a unique DNA profile in this study.

CHAPTER 8

SUMMARY

Sorghum is the most important cereal crop providing food to millions of people in the world. It is well known for its adaptation to harsh environments, specifically to drought and heat stresses, which accounts for its success throughout the semi-arid regions of the world. Africa, specifically the northeast quadrant of Africa is believed to be the primary centre of origin and domestication of the crop. In these parts of Africa, genetic variability is available both in cultivated races and the wild progenitors. In regard to sorghum utilization in general, developing countries use it primarily as food, whereas developed countries use it as feed. A wide variety of traditional foods are used from sorghum in the semi-arid tropics.

Despite its importance, the genetic characterization of sorghum is very limited. The accurate estimation of genetic diversity of the species is important for conservation of valuable resources and possible future use in its improvement. Farmers' varieties or landraces (locally adapted populations bred through traditional methods of direct selection) are usually the major sources of genetic variation. Cultivated sorghums in Ethiopia show diverse morpho-agronomic diversity, and have not been studied using the recently developed molecular markers. The aim of this study was to estimate genetic diversity by using DNA markers, morpho-agronomical traits and food quality attributes in sorghum accessions.

Forty-five accessions, including landrace collections, breeding materials and improved cultivars from the eastern highlands of Ethiopia were used. A total of 552 and 43 polymorphic AFLP and microsatellite alleles, respectively were scored and used to calculate pair-wise genetic distances and clustering. In addition, 10 qualitative and 16 quantitative traits with 96 variants were scored

and used to analyse the genetic distances and clustering. The physical and chemical composition and food (*injera*)-making qualities of selected sorghum samples were also investigated. A high phenotypic, chemical and genetic variability among the accessions was observed.

The resulting knowledge of genetic distance and discrimination of sorghum accessions, chemical composition variability and *injera*-making quality in this study will contribute towards sorghum improvement programmes in Ethiopia, and conservation of novel genotypes. It permits an organization of germplasm resources and identification of parents for crossing blocks. This will enable the breeder/ improvement scientist to make more scientific based choices. These findings have shown that both AFLP and microsatellite techniques can be successfully used and that they are informative in estimation of genetic diversity and identification of sorghum accessions. The result from morpho-agronomical traits analysis generally agreed with the molecular marker results in estimating diversity, hence it can be used in the management of sorghum genetic resources. A further extensive investigation of Ethiopian sorghum genetic diversity including wider areas and more samples is recommended.

OPSOMMING

Sorghum is die mees belangrike graan gewas wat voedsel verskaf aan miljoene mense in die wêreld. Dit is bekend vir goeie aanpassing in marginale omgewings, onder veral hitte en droogte stremming, wat die sukses van die gewas verklaar dwarsoor semi-ariede areas van die wêreld. Afrika, spesifiek die noordoostelike kwadrant van Afrika, word beskou as die primêre sentrum van oorsprong en die plek van eerste verbouing van die gewas. In hierdie dele van Afrika is genetiese variabiliteit beskikbaar vir beide verboude rasse sowel as wilde verwantes. Vir sorghum gebruik in die algemeen, word dit in ontwikkelende lande gebruik primêr as voedsel, en in die ontwikkelde lande as voer. 'n Wye verskeidenheid tradisionele kos word gemaak van sorghum in die semi-ariede trope.

Ten spyte van die belangrikheid daarvan, is die genetiese karakterisering van sorghum baie beperk. Die akkurate vasstelling van genetiese diversiteit van die spesie is belangrik vir bewaring van waardevolle bronne en moontlike toekomstige gebruik vir verbetering van die gewas. Boere variëteite of landrasse (plaaslik aangepaste populasies wat geteel is deur tradisionele metodes van direkte seleksie) is gewoonlik die hoof bron van genetiese variasie. Verboude sorghums in Etiopië wys diverse morfo-agronomiese diversiteit, en is nog nie bestudeer met onlangs ontwikkelde molekulêre merkers nie. Die doel van hierdie studie was om genetiese diversiteit te bepaal met DNA merkers, morfo-agronomiese eienskappe en voedsel kwaliteit eienskappe in sorghum inskrywings.

Vyf en veertig inskrywings, insluitend landras versamelings, teelmateriaal en vebeterde cultivars van die oostelike hooglande van Etiopië is gebruik. 'n Totaal van 552 en 43 polimorfiese AFLP en mikrosatelliet allele, onderskeidelik, is geëvalueer en gebruik om paarsgewyse genetiese afstande en groepering te bepaal. Verder is 10 kwalitatiewe en 16 kwantitatiewe

eienskappe van 96 variante geëvalueer om genetiese afstande en groeperings te bepaal. Die fisiese en chemiese samestelling en voedsel (*Injera)* kwaliteit van geselekteerde sorghum monsters is ook bepaal. 'n Hoë fenotipiese, chemiese en genetiese variabiliteit is tussen inskrywings gesien.

Die data gegenereer in terme van genetiese afstande en diskriminasie van sorghum inskrywings, chemiese samestelling variabiliteit en *injera*-bakkwaliteit in hierdie studie sal bydra tot sorghum verbeterings programme in Etiopië, sowel as die bewaring van unieke genotipes. Dit dra by tot die organisasie van kiemplasma bronne en die identifisering van ouers vir kruisings blokke. Dit sal die teler meer wetenskaplike keuses toelaat. Hierdie bevindings het getoon dat beide AFLP en mikrosatelliet tegnieke suksesvol gebruik kan word en dat hulle bruikbaar is vir die bepaling van genetiese diversiteit en identifiksaie van sorghum inskrywings. Die resultate van die morfo-agronomiese eienskap analise het in die algemeen ooreengestem met molekulêre merkers in die bepaling van diversiteit, dus kan dit gebruik word in die bestuur van sorghum genetiese bronne. 'n Verdere uitgebreide ondersoek na Etiopiese sorghum genetiese diversiteit, insluitend wyer areas en meer inskrwyings, word aanbeveel.

Appendix I AFLP fragment scores for primer combination *EcoI*+ACA / *MseI*+CAA

	Fragment size (bp)																		
ACC.	60	62	64	69	71	77	80	82	86	87	89	93	95	96	100	102	105	106	108
1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1
2	1	1	0	1	1	1	1	1	0	1	1	1	1	0	1	0	1	0	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1
5	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0	1	1
6	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	1
7	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	1
8	1	0	0	0	1	1	0	1	0	1	1	1	1	1	1	1	1	0	0
9	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1
10	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1
11	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
14	0	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	0
15	1	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	0	0
16	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0
17	0	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	0
18	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0
19	0	0	1	0	1	1	1	0	1	1	1	1	1	0	1	0	0	1	1
20	1	1	0	0	1	1	1	1	0	1	1	1	1	1	0	1	1	0	0
21	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	0	0
22	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	0	0
23	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1
24	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0
25	1	1	0	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	1
26	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0
27	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0
28	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0
29	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	1	0	0
30	1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0
31	0	0	0	1	1	1	1	1	1	0	1	1	0	1	1	0	0	0	0
32	1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0
33	1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0
34	1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	0	1	0	1
35	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1
36	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	1	0	0
37	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0
38	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0
39	1	1	0	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	0
40	0	1	0	0	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0
41	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0
42	1	1	0	0	1	1	0	1	0	1	1	1	1	1	1	1	0	0	0
43	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	0	0	1
44	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0
45	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0

Appendix I cont.

	Fragment size (bp)																		
ACC.	113	117	119	123	125	127	129	133	135	137	140	143	145	147	150	153	155	157	160
1	1	0	0	1	0	0	1	1	0	1	1	1	1	0	0	1	1	1	1
2	1	0	0	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	1
3	1	0	0	1	0	0	1	1	0	0	1	0	1	1	1	0	0	1	1
4	1	0	0	1	0	0	1	0	0	0	1	1	1	1	0	1	1	1	0
5	1	0	0	1	0	0	1	1	0	1	1	1	1	1	0	0	1	1	0
6	1	1	0	1	0	0	1	1	0	1	1	1	0	1	0	0	1	0	1
7	1	0	0	1	0	0	1	1	0	0	1	1	0	1	0	1	1	1	1
8	1	1	0	1	1	0	1	1	0	1	1	1	0	1	0	1	0	1	0
9	1	0	0	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	1
10	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1
11	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	0
12	1	1	1	1	0	0	1	1	0	0	0	1	1	1	0	0	0	0	1
13	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1
14	1	0	0	1	0	0	1	1	0	0	1	1	0	1	0	0	1	1	0
15	1	1	0	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1
16	1	0	0	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1
17	1	0	0	1	0	0	1	1	0	0	1	1	0	1	0	0	1	0	1
18	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	1
19	0	0	1	1	0	0	1	0	0	1	1	1	0	1	0	1	1	0	1
20	1	0	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	0	1
21	1	0	0	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	1
22	1	0	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0
23	1	1	1	1	0	0	1	1	0	0	1	0	1	1	1	0	1	0	1
24	1	0	0	1	0	0	1	1	0	0	1	1	0	1	1	0	1	0	0
25	1	0	0	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1
26	1	0	0	1	0	1	0	1	0	0	1	1	1	0	0	0	1	1	0
27	1	0	0	1	0	1	1	1	0	0	1	1	1	0	0	0	1	1	0
28	1	0	0	1	0	0	1	1	0	0	1	1	1	0	1	0	1	0	0
29	1	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	1	1
30	1	0	0	1	0	0	1	1	0	0	1	1	1	0	0	0	1	0	0
31	1	0	0	1	0	0	1	1	1	1	1	1	0	1	0	0	1	0	1
32	1	0	0	1	0	1	0	1	1	0	1	1	1	0	1	0	1	1	0
33	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	0	1	0	0
34	1	0	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	0
35	1	0	1	1	0	1	1	1	0	0	1	1	1	0	1	0	1	1	0
36	1	1	0	1	0	0	0	1	0	1	1	1	0	1	0	1	1	0	1
37	1	0	1	1	0	1	0	1	1	0	1	1	1	0	1	0	1	1	0
38	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	1	0	1
39	1	1	1	1	0	1	1	1	1	0	1	0	1	1	0	0	0	1	1
40	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	1
41	1	1	0	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1
42	1	1	0	1	0	1	1	1	1	0	1	1	1	0	1	0	1	1	1
43	1	0	0	1	1	0	1	1	0	1	1	1	1	1	0	1	1	0	1
44	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1
45	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1

Appendix I cont.

	Fragment size (bp)																			
ACC.	161	163	167	169	173	175	177	181	186	190	192	194	195	198	200	202	204	206	209	213
1	1	0	0	1	1	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1
2	0	0	0	1	1	0	1	1	1	0	0	0	1	1	1	0	0	1	0	1
3	1	0	0	1	1	0	1	1	1	1	0	0	0	0	1	0	0	1	0	1
4	0	1	0	0	1	0	1	1	1	0	0	1	0	0	1	0	0	1	1	1
5	0	1	0	1	1	0	0	1	1	0	0	0	0	1	1	0	0	1	0	1
6	0	1	0	0	0	1	0	1	1	0	1	0	0	0	1	0	0	1	0	1
7	0	1	0	0	1	1	1	1	1	0	0	0	0	1	1	0	0	1	0	1
8	1	0	0	0	1	1	1	1	1	0	0	0	0	1	0	1	0	0	1	0
9	1	0	0	0	1	1	0	1	1	0	0	1	0	0	1	0	0	1	0	1
10	0	1	0	1	1	1	0	1	1	0	0	1	1	0	1	0	0	1	1	1
11	1	1	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	0	0	1
12	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0
13	1	0	1	1	0	1	1	1	1	0	0	1	0	0	1	1	0	1	1	1
14	0	0	1	0	0	1	0	1	1	0	0	0	1	0	1	0	0	1	1	1
15	1	0	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1
16	1	0	0	1	1	1	1	1	1	0	0	0	1	0	1	0	0	1	0	1
17	0	1	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0	1	0	1
18	0	1	0	1	1	0	1	1	1	0	0	1	0	0	1	0	1	1	0	1
19	0	1	0	0	1	1	1	0	1	0	1	0	0	0	1	0	0	1	0	1
20	1	0	0	1	1	1	1	0	0	1	0	0	0	1	0	1	0	0	1	0
21	1	0	1	0	0	1	0	1	0	0	0	0	1	0	0	1	0	1	1	1
22	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1
23	0	1	1	1	0	0	1	1	1	0	0	0	0	0	1	0	0	1	0	1
24	1	1	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	1	0	1
25	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	1
26	1	0	0	0	1	1	1	1	1	0	0	1	1	0	1	0	0	1	0	1
27	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	0	0	1	0	1
28	1	0	1	0	1	0	1	1	0	0	1	0	0	0	1	0	0	1	1	1
29	0	1	1	1	0	0	1	1	1	0	0	1	1	1	0	1	0	0	0	1
30	0	1	0	1	1	0	1	1	1	0	1	0	0	0	1	0	0	1	0	1
31	0	0	1	0	0	1	0	1	0	0	0	0	1	0	1	1	0	1	0	1
32	1	1	1	0	1	0	1	1	1	0	0	1	0	1	0	0	0	1	1	0
33	1	0	1	1	1	0	1	1	1	0	1	0	0	1	0	0	0	1	0	1
34	1	0	1	1	0	1	1	1	1	0	0	1	1	1	1	0	0	1	0	1
35	1	0	0	0	1	1	1	1	0	0	1	0	0	1	0	0	0	1	0	1
36	0	0	0	1	0	1	0	0	0	1	0	1	0	1	0	1	0	0	1	1
37	1	0	0	0	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	1
38	0	0	0	1	1	1	1	1	1	0	1	0	1	0	0	1	0	1	1	1
39	0	0	0	1	0	0	1	1	0	1	0	0	1	0	0	1	0	0	0	1
40	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1
41	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	0	1
42	0	1	0	1	0	1	1	1	1	0	0	1	1	0	1	0	0	1	0	1
43	0	0	0	1	1	1	1	1	1	1	0	0	0	1	0	1	0	0	1	1
44	1	1	0	1	1	1	1	1	0	1	1	0	0	1	0	1	1	1	0	1
45	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1

Appendix I cont.

	Fragment size (bp)																		
ACC.	217	219	221	224	228	230	232	236	238	241	247	251	254	258	264	266	270	272	276
1	0	0	0	1	0	1	0	0	1	0	1	0	1	0	1	0	1	1	0
2	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	0	1	1	1
3	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	0	1	1	0
4	0	0	0	1	1	0	0	0	1	0	1	1	0	0	0	1	1	1	1
5	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	0	1	1	0
6	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	0	1	1	0
7	0	0	1	1	0	1	1	0	1	1	1	1	0	0	1	0	1	1	0
8	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0	0
9	0	0	0	1	0	1	1	0	1	1	1	0	0	0	1	0	1	1	0
10	1	1	1	1	0	1	1	0	1	0	1	1	1	0	1	0	0	0	0
11	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0
12	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0
13	1	0	0	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	0
14	0	0	0	1	1	1	0	0	1	0	1	0	0	0	1	0	1	1	0
15	1	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0
16	0	0	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	1	0
17	0	0	1	1	0	1	1	0	1	1	0	0	0	0	0	1	0	0	0
18	0	0	1	1	0	1	1	0	1	1	1	1	0	0	0	1	1	0	0
19	0	0	0	1	0	1	0	0	1	0	1	1	0	0	1	0	1	1	0
20	1	0	1	0	1	0	0	1	0	0	0	1	0	0	1	0	1	0	0
21	0	0	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0
22	0	0	0	1	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0
23	0	0	0	1	0	1	1	0	1	1	1	1	1	0	0	0	0	1	0
24	0	0	0	1	0	0	1	0	1	1	1	1	0	0	1	0	1	0	0
25	1	0	1	1	1	0	0	0	1	0	1	0	1	0	1	0	1	1	0
26	1	1	1	1	1	0	0	1	1	0	1	0	1	1	1	0	1	1	0
27	0	0	0	1	0	1	0	0	1	1	1	0	1	0	1	0	1	1	0
28	0	1	0	1	0	0	0	0	1	1	0	1	0	0	1	0	0	1	0
29	1	1	0	1	1	0	0	1	1	1	0	1	1	1	1	0	1	1	1
30	0	0	1	1	0	1	0	0	1	1	0	0	1	0	0	0	0	1	0
31	0	0	0	1	1	1	0	0	1	0	0	0	0	0	1	0	1	1	0
32	0	0	0	1	1	0	0	0	1	1	0	1	0	0	0	1	0	0	0
33	0	1	0	1	0	1	0	0	1	0	1	1	1	0	1	1	1	0	0
34	1	0	0	1	1	1	0	0	1	1	1	1	0	0	1	0	1	1	1
35	1	0	1	1	0	1	0	0	1	0	1	0	1	0	1	0	1	1	0
36	1	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
37	0	0	1	1	0	1	0	0	1	0	1	1	0	0	0	1	1	1	0
38	1	0	1	1	1	0	0	1	0	0	1	1	0	0	1	1	0	1	0
39	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	1	1	0
40	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1
41	1	0	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	0
42	1	1	1	1	1	1	0	0	1	0	1	0	0	1	1	0	1	1	0
43	1	0	1	1	1	0	0	1	1	0	0	1	1	0	1	0	1	0	0
44	1	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	0	0	0
45	1	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1	0	0	0

Appendix I cont.

	Fragment size (bp)																
ACC.	279	283	287	294	300	303	305	309	319	325	330	336	342	346	352	355	358
1	0	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0
2	0	0	0	0	1	0	0	1	0	1	1	0	0	1	1	0	0
3	0	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0
4	0	0	0	0	0	1	0	1	1	0	1	0	0	1	1	0	1
5	0	1	0	0	0	1	0	1	0	0	1	0	0	1	1	0	1
6	0	0	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1
7	0	1	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0
8	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0
9	0	0	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0
10	0	0	0	1	1	0	0	1	0	0	1	0	1	1	1	0	0
11	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	1	0	1	1	0	0	0	1	1	1	1	0	0	1	1	0	0
14	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
15	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0
16	0	1	0	0	1	1	0	1	0	1	1	0	0	1	1	0	0
17	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0
18	0	1	1	0	1	1	0	1	1	0	1	0	0	1	1	0	0
19	0	0	0	0	0	1	0	1	0	0	1	0	0	1	1	0	0
20	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	0
21	0	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	0
22	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0
23	0	0	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0
24	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0
25	0	1	0	1	0	1	0	1	1	0	1	1	1	1	1	0	0
26	1	1	0	0	0	1	0	1	1	1	1	1	0	1	1	0	0
27	0	0	1	0	1	1	0	1	0	1	1	0	1	0	1	0	0
28	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0
29	0	1	0	1	0	1	1	1	0	1	1	0	1	1	1	0	0
30	0	1	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0
31	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
32	0	1	0	0	0	1	0	1	0	0	1	0	0	0	1	0	0
33	0	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	1
34	1	0	0	0	0	1	1	1	1	0	1	0	0	1	0	1	0
35	0	0	1	0	0	0	0	1	1	0	1	0	0	1	0	1	0
36	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
37	0	1	1	0	0	0	0	1	1	0	1	0	0	1	1	0	0
38	0	1	1	0	1	0	0	1	1	0	0	1	1	0	0	1	1
39	1	1	0	0	1	0	0	1	0	0	0	1	1	0	0	1	1
40	0	0	0	0	0	1	1	1	0	1	1	0	0	1	1	1	0
41	0	1	1	0	1	1	0	1	1	0	1	0	0	1	1	0	1
42	0	1	0	0	0	1	1	1	0	0	1	0	0	1	1	0	0
43	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	1	0
44	0	1	0	0	1	0	0	1	1	0	1	0	0	1	1	0	1
45	1	1	0	0	1	0	1	1	1	0	1	0	0	1	1	0	1

Appendix I cont.

	Fragment size (bp)																	
ACC.	364	366	367	372	375	388	393	403	411	418	432	447	454	465	473	480	487	493
1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1
2	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1
3	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
4	0	1	0	1	1	0	0	0	1	0	0	0	1	0	0	0	1	0
5	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
6	0	1	0	0	0	1	0	1	1	0	0	1	1	0	0	1	1	0
7	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
8	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
9	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1
10	0	1	0	0	1	0	1	0	1	1	0	0	1	0	0	0	1	0
11	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0	1
14	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
15	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	0
16	0	1	0	0	1	0	1	0	1	1	0	1	1	1	0	0	1	0
17	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
18	0	1	0	1	1	0	0	1	1	0	0	0	1	0	0	0	1	0
19	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
20	1	0	0	0	0	0	0	1	0	1	0	1	0	0	1	1	0	0
21	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
22	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
23	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
24	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
25	0	1	0	1	0	0	1	0	1	0	0	0	1	0	0	0	1	0
26	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0
27	0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0
28	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
29	0	1	0	1	1	0	0	1	1	0	0	0	1	1	0	0	1	0
30	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
31	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
32	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	1	0
33	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
34	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
35	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
36	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0
37	0	1	1	1	0	0	1	0	1	0	0	0	1	0	0	0	1	0
38	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1
39	0	0	0	0	1	1	0	1	0	0	1	0	0	0	1	0	0	0
40	0	1	0	0	1	0	0	0	1	0	1	1	1	0	1	0	0	0
41	1	1	0	1	1	1	1	0	1	1	0	0	1	0	0	0	1	1
42	0	1	0	1	1	1	0	0	1	1	0	0	1	0	0	0	1	1
43	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
44	0	1	1	0	1	1	0	0	0	1	0	0	1	0	0	0	1	1
45	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0

Appendix II Amplified fragments score for 10 microsatellite loci: sb1-10, sb4-15, sb4-22, sb4-32, sb5-85, sb5-236, sb6-36, sb6-57, sb6-84 and sb6-342.

	Fragment size (bp)											
	sb1-10				sb4-15	sb4-22						
Acc	110	275	304	510	135	254	314	330	347	405	492	
1	0	1	1	0	1	0	1	1	0	0	0	
2	0	1	1	0	1	1	0	0	0	0	0	
3	0	1	0	0	1	0	0	1	0	0	0	
4	0	1	0	0	1	0	0	1	0	0	0	
5	0	1	0	0	1	0	0	1	0	0	0	
6	0	0	1	0	1	0	1	0	0	0	0	
7	0	0	1	0	1	0	0	1	0	0	0	
8	0	1	1	0	1	0	0	1	1	0	0	
9	0	1	1	0	1	0	0	1	0	0	0	
10	0	0	1	0	1	0	0	1	0	0	0	
11	0	1	0	0	1	0	0	1	0	0	0	
12	0	1	0	0	1	0	0	0	1	0	0	
13	1	0	0	0	1	0	0	1	0	0	0	
14	0	0	1	0	1	0	0	0	1	0	0	
15	0	0	0	1	1	0	0	1	0	0	0	
16	0	1	0	0	1	0	0	1	0	0	0	
17	0	0	1	0	1	0	0	1	0	0	0	
18	0	0	1	0	1	0	0	0	1	0	0	
19	0	1	0	0	1	0	0	0	1	0	0	
20	0	1	0	0	1	0	0	0	1	0	0	
21	0	1	0	0	1	0	0	1	0	0	0	
22	0	1	0	0	1	0	0	1	0	0	0	
23	0	0	1	0	1	0	0	1	0	0	0	
24	0	0	0	1	1	0	0	1	1	0	0	
25	0	1	0	1	1	0	0	1	1	0	0	
26	0	1	0	0	1	0	0	1	1	0	0	
27	0	1	0	0	1	0	0	1	0	0	0	
28	0	0	0	1	1	0	0	1	0	0	0	
29	0	1	0	0	1	0	0	1	0	0	0	
30	0	0	0	1	1	0	0	1	0	0	0	
31	0	1	0	0	1	0	1	0	0	0	0	
32	0	0	0	1	1	0	1	0	0	0	0	
33	0	0	0	1	1	0	1	0	0	0	0	
34	0	0	1	0	1	0	1	0	0	0	0	
35	0	1	0	0	1	0	1	0	1	0	0	
36	0	1	0	0	1	0	1	0	0	0	0	
37	0	1	0	0	1	0	0	0	1	0	0	
38	0	1	0	0	1	0	1	0	0	0	0	
39	0	1	0	0	1	0	0	0	1	0	0	
40	0	1	0	0	1	0	0	1	0	0	0	
41	0	1	0	0	1	0	1	0	0	0	0	
42	0	1	0	0	1	0	0	0	1	0	0	
43	0	1	1	0	1	0	0	0	1	0	0	
44	0	1	0	0	1	0	0	0	1	1	1	
45	0	1	0	0	1	0	0	0	1	1	1	

Appendix II cont.

Acc	Fragment size (bp)															
	sb4-32									sb5-85						
	112	174	190	240	293	357	439	497	533	109	159	186	215	336	408	510
1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
2	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1
3	0	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	1	0	0	1	1	0	1	0	0	0
5	0	0	0	0	0	1	0	0	0	1	0	0	1	0	1	1
6	0	0	0	0	1	0	1	0	0	1	1	0	1	0	1	1
7	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	1
8	0	0	0	1	1	1	1	1	1	1	0	0	1	0	0	1
9	0	0	1	1	1	1	1	1	1	1	0	0	1	0	0	1
10	0	0	0	1	1	1	1	1	1	1	0	0	1	0	0	1
11	0	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1
12	0	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1
13	0	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1
14	0	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1
15	0	0	1	1	1	1	1	1	1	1	0	1	1	0	0	1
16	0	0	1	1	1	1	1	1	0	1	1	0	1	0	0	1
17	0	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1
18	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0
19	0	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0
20	0	0	1	1	0	0	0	1	1	1	0	0	1	0	0	0
21	0	0	1	1	1	1	0	0	0	1	1	0	0	1	0	0
22	0	0	1	1	0	1	0	1	1	1	1	1	1	0	0	0
23	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	0
24	0	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0
25	0	0	1	1	1	0	1	1	1	1	0	0	1	0	0	0
26	0	1	0	1	1	0	0	1	0	1	0	0	1	0	0	0
27	0	0	0	1	0	0	1	1	0	1	1	0	1	0	0	0
28	0	0	1	1	1	0	0	0	0	1	0	0	1	0	0	0
29	0	1	0	1	1	0	0	0	0	1	1	0	1	0	0	0
30	0	1	0	1	1	0	0	0	0	1	0	0	1	0	0	0
31	0	1	0	1	1	0	1	0	0	1	0	1	0	0	0	0
32	1	0	1	1	1	0	1	0	0	1	0	0	1	0	0	0
33	0	0	1	1	1	0	0	0	0	1	0	1	1	0	0	0
34	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0
35	1	0	1	0	1	0	0	1	0	1	1	0	1	0	0	0
36	1	0	1	0	1	0	0	1	0	1	1	0	1	0	0	0
37	1	0	1	0	0	0	0	1	0	1	1	0	1	0	0	0
38	1	0	1	0	0	0	0	1	0	1	1	0	1	0	0	0
39	0	0	1	1	1	0	1	0	1	1	0	0	1	0	0	0
40	0	1	0	1	1	0	1	0	1	1	0	0	1	0	0	0
41	0	1	0	1	1	0	1	0	1	1	0	0	1	0	0	0
42	0	1	0	1	1	0	1	1	1	1	1	0	1	0	0	0
43	0	1	0	1	0	0	1	1	1	1	1	0	1	0	1	0
44	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0
45	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0

Appendix II cont.

Acc	Fragment size (bp)													
	sb5-236				sb6-36					sb6-57				
	110	185	207	265	165	183	198	249	279	292	310	320	343	
1	1	1	0	0	0	0	1	0	0	0	0	1	0	
2	0	1	0	0	0	0	1	0	0	0	0	1	0	
3	1	1	0	0	0	0	1	0	0	0	0	1	0	
4	0	0	1	0	1	0	0	0	0	0	0	0	1	
5	0	0	1	0	0	0	1	0	0	0	0	0	1	
6	0	0	1	0	1	0	0	0	0	1	0	0	0	
7	0	0	1	0	1	0	1	0	0	0	0	0	1	
8	0	0	1	0	1	0	1	0	1	0	1	1	0	
9	0	1	0	0	1	0	0	0	0	0	1	1	0	
10	0	1	0	1	1	0	0	0	0	0	1	1	0	
11	0	1	0	0	1	0	0	0	0	1	1	0	0	
12	0	1	0	0	0	1	0	0	0	1	0	1	0	
13	0	1	0	0	0	0	1	0	0	1	0	0	0	
14	0	1	0	0	0	0	1	0	0	1	0	1	1	
15	0	0	1	1	0	1	0	0	0	1	0	1	0	
16	0	1	0	1	0	0	1	0	0	1	0	0	0	
17	0	1	0	1	0	1	0	0	0	1	0	0	1	
18	0	1	0	0	0	1	0	0	0	1	0	0	0	
19	0	1	0	0	0	1	0	0	0	0	0	1	0	
20	0	0	1	0	0	0	1	0	0	0	0	1	0	
21	0	1	0	0	0	0	0	1	0	0	0	1	0	
22	0	1	0	0	0	0	0	0	1	0	0	0	1	
23	0	1	0	0	0	0	0	1	0	0	0	0	1	
24	1	1	1	0	0	0	0	1	0	0	0	0	1	
25	1	1	1	0	0	0	0	1	1	1	1	0	1	
26	0	1	0	0	0	0	0	1	1	0	0	0	1	
27	0	1	0	0	0	0	0	0	1	0	0	0	1	
28	0	0	1	0	0	0	0	1	0	0	0	0	1	
29	1	1	0	0	0	0	0	0	1	0	0	0	1	
30	0	0	1	0	0	0	0	1	0	0	0	1	1	
31	1	0	0	1	0	1	0	0	0	0	0	1	0	
32	1	0	1	1	1	0	0	0	0	0	0	1	0	
33	1	0	1	1	1	0	0	0	0	0	1	0	0	
34	1	0	0	1	0	0	1	0	0	0	1	0	0	
35	1	0	0	1	1	0	0	0	0	0	1	0	0	
36	1	0	0	1	1	0	0	0	0	1	0	0	0	
37	1	1	0	0	1	0	0	0	0	1	0	0	0	
38	1	0	0	1	1	0	0	0	0	1	0	0	0	
39	1	0	1	1	1	0	0	0	0	1	0	0	0	
40	0	0	1	1	0	0	1	0	0	0	1	0	0	
41	1	0	0	1	0	0	1	0	0	1	0	0	0	
42	0	0	1	0	0	0	1	0	0	1	0	0	0	
43	0	1	1	1	1	0	1	0	1	0	1	0	0	
44	1	1	1	1	1	0	0	0	0	1	0	0	0	
45	1	0	1	0	1	0	0	0	0	1	0	0	0	

Appendix II Cont.

Acc	Fragment size (bp)							
	sb6-84			sb6-342				
	173	200	295	254	280	382	544	634
1	0	1	1	0	1	1	0	1
2	0	1	1	0	1	0	0	1
3	0	1	0	0	1	0	0	1
4	0	0	1	0	0	1	1	1
5	0	0	1	0	1	0	0	1
6	0	0	1	0	1	1	0	1
7	0	0	1	0	1	1	0	0
8	0	1	0	0	1	1	0	1
9	0	1	0	0	0	1	0	1
10	1	0	0	0	0	1	0	1
11	1	0	0	0	1	0	0	1
12	1	0	0	0	1	1	0	1
13	1	0	0	0	1	1	0	1
14	1	0	0	1	0	0	0	1
15	1	0	0	0	1	1	0	1
16	0	1	0	0	1	0	0	1
17	0	1	0	0	1	1	0	1
18	1	0	0	0	1	0	0	1
19	1	0	0	0	1	1	0	1
20	1	0	0	0	1	0	0	1
21	0	1	0	0	1	0	0	1
22	0	1	0	0	1	1	0	0
23	0	1	0	0	0	0	0	1
24	0	1	0	0	0	1	0	1
25	0	1	0	0	0	1	0	1
26	0	1	0	0	1	1	1	1
27	0	1	0	0	1	1	1	1
28	0	1	0	0	1	1	1	1
29	0	1	0	0	0	1	1	1
30	1	1	0	0	0	1	0	1
31	1	1	0	0	0	1	0	1
32	1	0	0	0	0	1	0	1
33	1	0	0	0	0	1	0	1
34	1	0	0	0	0	1	0	1
35	1	0	0	0	0	1	0	1
36	1	0	0	0	0	1	0	1
37	1	0	0	0	1	0	0	1
38	1	0	0	0	0	1	0	1
39	1	0	0	0	0	1	0	1
40	0	1	0	0	0	1	0	1
41	0	1	0	0	1	1	0	1
42	0	1	0	0	1	1	0	1
43	1	1	0	0	1	1	0	1
44	0	1	0	0	0	1	0	1
45	0	1	0	0	0	1	0	1

Appendix III Average values calculated for 16 quantitative traits in 45 sorghum accessions.

No	DF	LN	LL	LW	LA	INL	LLS	PHt	PL	PW	NPB	HWt	GYPP	TSWt	TP	KNPP
1	94.0	9.9	60.0	7.3	324.9	20.8	17.3	185.0	11.1	8.5	56.5	87.4	60.2	28.0	69.0	2151.1
2	93.0	9.3	59.2	7.1	311.8	20.8	16.0	192.0	10.7	8.8	56.3	94.4	65.6	30.0	69.5	2186.7
3	98.0	13.1	72.2	8.7	469.2	28.9	23.4	250.5	16.9	10.3	85.3	121.5	84.6	29.0	69.6	2915.5
4	94.0	13.8	89.0	11.8	781.2	24.3	20.9	308.0	12.9	10.7	138.6	179.2	123.6	28.3	69.0	4367.1
5	106.0	13.3	89.2	10.8	719.6	24.0	19.0	251.3	13.6	10.6	116.3	123.7	87.6	31.7	70.8	2762.2
6	100.0	13.9	80.4	11.4	684.7	27.5	22.8	305.0	13.9	9.6	66.5	152.6	107.3	20.7	70.3	5183.1
7	111.0	14.7	89.1	11.7	775.4	26.3	20.0	330.5	14.9	9.1	82.1	131.7	96.8	23.3	73.5	4152.8
8	92.0	13.8	86.2	10.0	643.9	21.9	21.0	240.0	17.5	9.4	83.0	169.0	115.6	26.7	68.4	4328.8
9	97.0	13.2	76.2	9.3	529.4	30.4	23.7	285.0	18.2	11.2	94.6	176.7	124.3	27.3	70.4	4553.1
10	113.0	12.2	61.6	8.3	381.9	26.3	20.5	285.0	12.4	10.7	103.1	180.6	138.5	33.7	76.7	4109.5
11	96.0	11.3	78.4	9.4	547.6	23.6	22.2	315.5	15.4	10.7	109.2	167.7	124.2	28.3	74.1	4389.1
12	94.0	12.2	72.2	8.2	442.3	22.4	22.0	275.0	17.2	11.3	69.2	146.5	113.3	25.0	77.3	4530.0
13	91.0	9.5	58.7	7.2	313.5	24.7	17.0	215.5	11.8	10.3	96.0	83.5	64.3	23.0	77.0	2795.2
14	94.0	14.0	72.6	8.9	482.7	23.6	20.5	290.5	20.1	9.4	87.8	176.4	114.3	24.8	64.8	4607.3
15	109.0	13.3	83.8	12.7	791.7	24.2	21.2	220.3	26.0	12.9	71.9	161.2	108.2	26.7	67.1	4052.4
16	92.0	11.4	63.9	8.1	386.6	22.0	21.0	265.0	17.5	12.0	84.7	121.0	91.7	29.7	75.8	3087.5
17	112.0	12.0	84.3	11.1	699.0	25.8	20.5	290.0	17.5	10.2	71.5	152.1	96.7	20.3	63.6	4762.1
18	106.0	11.5	83.0	11.1	685.1	26.5	21.0	285.0	18.0	8.7	87.7	155.3	101.7	29.0	65.5	3505.9
19	105.0	12.0	70.7	11.8	620.6	25.1	23.4	294.0	19.3	11.7	93.7	187.3	118.2	26.0	63.1	4546.2
20	109.0	13.0	80.0	12.5	747.0	26.1	24.3	260.0	21.0	12.5	101.3	196.7	128.1	24.3	65.1	5272.0
21	104.0	11.5	64.7	10.7	515.7	24.3	20.0	270.0	24.7	11.3	81.0	128.7	89.7	20.0	69.7	4483.5
22	93.0	11.5	68.0	9.5	482.6	19.1	18.0	275.0	17.6	8.6	82.3	141.0	104.5	26.7	74.1	3913.9
23	108.0	11.7	80.2	11.3	674.0	28.1	20.5	270.0	15.2	9.6	61.9	167.5	133.0	24.7	79.4	5384.6
24	89.0	10.3	66.4	9.9	488.6	22.7	19.2	230.0	17.5	10.2	70.0	131.6	102.0	18.0	77.5	5668.9
25	98.0	11.0	61.3	9.5	435.2	28.3	22.9	263.0	22.0	11.7	76.2	168.7	122.7	24.0	72.8	5112.5
26	98.0	11.5	78.7	10.5	618.8	28.0	24.7	290.0	23.0	8.3	84.7	153.3	102.6	25.7	66.9	3992.6
27	92.0	10.5	80.7	9.3	560.4	26.6	19.5	285.0	25.1	11.0	76.6	153.2	99.4	26.7	64.9	3721.0
28	106.0	11.5	77.3	9.8	567.6	26.4	23.4	295.0	20.7	12.6	79.0	115.7	87.3	20.0	75.4	4362.5
29	94.0	13.0	80.8	11.1	667.0	25.5	21.9	290.0	20.9	12.8	82.6	175.2	109.5	24.0	62.5	4563.3
30	105.0	9.1	67.7	9.3	467.8	23.0	21.3	220.0	21.3	9.5	75.2	106.7	74.2	27.0	69.6	2749.6
31	90.0	11.0	73.0	9.0	490.8	22.4	19.5	296.0	23.2	8.6	112.3	163.0	104.6	28.0	64.2	3735.7
32	89.0	9.7	70.0	8.8	461.7	30.4	22.8	310.0	18.1	10.5	64.3	84.5	61.7	20.0	73.0	3082.5
33	100.0	10.0	73.0	8.3	454.2	29.7	21.9	316.0	20.7	10.2	63.6	90.0	63.4	19.5	70.4	3248.7
34	101.0	11.4	62.7	9.1	425.1	28.5	20.9	320.0	19.5	12.9	100.3	125.3	87.5	27.7	69.8	3157.0
35	97.0	12.3	77.3	10.1	580.3	25.6	20.0	295.0	16.0	11.4	127.6	187.6	125.5	25.0	66.9	5021.6
36	101.0	14.4	72.7	10.9	589.2	25.9	22.3	320.0	17.5	11.7	111.5	160.1	112.0	30.7	69.9	3646.6
37	98.0	13.6	67.4	10.2	511.0	23.3	23.0	330.0	15.6	11.5	134.5	156.6	118.9	27.3	75.3	4354.6
38	106.0	12.4	71.5	9.6	512.7	24.8	20.2	300.0	16.8	8.9	102.5	170.6	114.7	31.3	67.2	3663.3
39	86.0	9.0	81.0	8.8	530.6	25.8	20.8	275.0	25.0	10.1	65.7	83.9	61.3	20.3	73.1	3019.7
40	95.0	10.5	77.5	10.2	587.6	24.8	22.6	340.0	21.2	9.0	89.3	141.3	109.6	30.0	77.6	3654.7
41	93.0	11.5	75.3	10.3	576.8	23.1	20.7	315.0	17.6	9.6	90.7	161.7	124.7	30.0	77.1	4156.0
42	90.0	10.0	65.3	9.6	466.1	25.4	20.6	295.5	16.9	10.8	98.7	166.2	126.2	31.0	75.9	4070.7
43	99.0	13.4	72.8	11.2	606.4	30.2	24.8	360.0	14.8	9.4	125.1	175.7	143.0	31.0	81.4	4612.6
44	101.0	12.6	69.7	10.1	525.9	30.1	22.9	355.0	14.1	9.7	132.0	196.5	148.0	33.0	75.3	4486.1
45	103.0	12.5	64.9	9.6	465.4	26.7	24.9	310.0	14.4	9.7	120.7	191.4	143.3	28.7	74.8	4886.8

Appendix IV Binary scores of qualitative traits for 45 sorghum accessions.

	Leaf midrib colour			Panicle exsertion				Panicle compactness & shape						Awn	
Acc	1	2	3	1	2	3	4	1	4	6	8	9	10	0	1
1	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0
2	1	0	0	0	0	0	1	0	0	0	0	1	0	1	0
3	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1
4	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1
5	1	0	0	0	0	0	1	0	0	0	0	1	0	1	0
6	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1
7	1	0	0	1	0	0	0	0	0	0	0	1	0	0	1
8	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1
9	1	0	0	0	0	1	0	0	1	0	0	0	0	0	1
10	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1
11	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1
12	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1
13	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1
14	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1
15	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1
16	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1
17	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1
18	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1
19	1	0	0	0	1	0	0	0	0	0	0	1	0	0	1
20	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1
21	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1
22	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1
23	0	1	0	1	0	0	0	0	0	0	0	1	0	0	1
24	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1
25	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1
26	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1
27	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0
28	0	0	1	1	0	0	0	0	1	0	0	0	0	0	1
29	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1
30	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1
31	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0
32	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1
33	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1
34	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0
35	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1
36	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1
37	1	0	0	0	0	0	1	0	0	1	0	0	0	0	1
38	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1
39	0	0	1	0	0	1	0	0	0	1	0	0	0	0	1
40	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1
41	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1
42	0	0	1	0	0	1	0	0	0	0	1	0	0	0	1
43	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1
44	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1
45	1	0	0	0	0	0	1	0	0	0	0	0	1	0	1

Appendix IV cont.

Acc	Glume colour						Grain covering		Grain colour						
	1	2	3	4	5	6	1	3	1	2	3	4	5	7	9
1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0
2	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0
3	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1
5	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1
6	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0
7	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0
8	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0
9	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0
10	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
11	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0
12	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0
13	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
14	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
15	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0
16	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0
17	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0
18	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
19	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0
20	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
21	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1
22	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
23	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
24	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0
25	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0
26	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0
27	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
28	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
29	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0
30	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0
31	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
32	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0
33	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0
34	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
35	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0
36	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
37	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
38	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0
39	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
40	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0
41	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
42	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
43	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
44	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0
45	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0

Appendix IV cont.

	Stalk juiciness		Plant colour		Endosperm texture		
Acc	D	J	1	2	5	7	9
1	1	0	0	1	1	0	0
2	1	0	0	1	0	0	1
3	1	0	0	1	0	0	1
4	1	0	1	0	0	1	0
5	1	0	1	0	0	1	0
6	1	0	0	1	1	0	0
7	1	0	1	0	1	0	0
8	1	0	0	1	1	0	0
9	1	0	0	1	1	0	0
10	1	0	0	1	0	1	0
11	1	0	0	1	0	1	0
12	1	0	0	1	1	0	0
13	1	0	0	1	0	0	1
14	1	0	0	1	0	0	1
15	1	0	1	0	0	1	0
16	1	0	0	1	1	0	0
17	1	0	0	1	0	1	0
18	1	0	0	1	0	0	1
19	1	0	0	1	0	1	0
20	1	0	0	1	0	1	0
21	1	0	0	1	0	0	1
22	1	0	0	1	1	0	0
23	1	0	0	1	1	0	0
24	1	0	0	1	0	1	0
25	1	0	0	1	0	0	1
26	1	0	0	1	0	0	1
27	1	0	0	1	0	1	0
28	1	0	0	1	0	1	0
29	1	0	0	1	1	0	0
30	1	0	0	1	0	1	0
31	1	0	0	1	0	1	0
32	1	0	1	0	0	0	1
33	1	0	0	1	0	0	1
34	1	0	0	1	0	0	1
35	0	1	0	1	0	0	1
36	1	0	0	1	0	1	0
37	1	0	0	1	0	1	0
38	1	0	0	1	0	1	0
39	1	0	0	1	0	1	0
40	1	0	0	1	1	0	0
41	1	0	0	1	1	0	0
42	0	1	1	0	0	0	1
43	1	0	0	1	0	1	0
44	1	0	0	1	0	1	0
45	0	1	0	1	0	0	1

Appendix V Binary scores of quantitative traits for 45 sorghum accessions.

Acc	DF					LN			LL			
	86-89	90-92	93-95	96-101	106-112	9:10	11:12	13 - 14	50 - 60	61 - 70	71-80	81-90
1	0	0	1	0	0	1	0	0	1	0	0	0
2	0	0	1	0	0	1	0	0	1	0	0	0
3	0	0	0	1	0	0	0	1	0	0	1	0
4	0	0	1	0	0	0	0	1	0	0	0	1
5	0	0	0	0	1	0	0	1	0	0	0	1
6	0	0	0	1	0	0	0	1	0	0	1	0
7	0	0	0	0	1	0	0	1	0	0	0	1
8	0	1	0	0	0	0	0	1	0	0	0	1
9	0	0	0	1	0	0	0	1	0	0	1	0
10	0	0	0	0	1	0	1	0	0	1	0	0
11	0	0	0	1	0	0	1	0	0	0	1	0
12	0	0	1	0	0	0	1	0	0	0	1	0
13	0	1	0	0	0	1	0	0	1	0	0	0
14	0	0	1	0	0	0	0	1	0	0	1	0
15	0	0	0	0	1	0	0	1	0	0	0	1
16	0	1	0	0	0	0	1	0	0	1	0	0
17	0	0	0	0	1	0	1	0	0	0	0	1
18	0	0	0	0	1	0	1	0	0	0	0	1
19	0	0	0	0	1	0	1	0	0	1	0	0
20	0	0	0	0	1	0	0	1	0	0	1	0
21	0	0	0	0	1	0	1	0	0	1	0	0
22	0	0	1	0	0	0	1	0	0	1	0	0
23	0	0	0	0	1	0	1	0	0	0	1	0
24	1	0	0	0	0	1	0	0	0	1	0	0
25	0	0	0	1	0	0	1	0	0	1	0	0
26	0	0	0	1	0	0	1	0	0	0	1	0
27	0	1	0	0	0	0	1	0	0	0	0	1
28	0	0	0	0	1	0	1	0	0	0	1	0
29	0	0	1	0	0	0	0	1	0	0	0	1
30	0	0	0	0	1	1	0	0	0	1	0	0
31	0	1	0	0	0	0	1	0	0	0	1	0
32	1	0	0	0	0	1	0	0	0	1	0	0
33	0	0	0	1	0	1	0	0	0	0	1	0
34	0	0	0	1	0	0	1	0	0	1	0	0
35	0	0	0	1	0	0	1	0	0	0	1	0
36	0	0	0	1	0	0	0	1	0	0	1	0
37	0	0	0	1	0	0	0	1	0	1	0	0
38	0	0	0	0	1	0	1	0	0	0	1	0
39	1	0	0	0	0	1	0	0	0	0	0	1
40	0	0	1	0	0	0	1	0	0	0	1	0
41	0	0	1	0	0	0	1	0	0	0	1	0
42	0	1	0	0	0	1	0	0	0	1	0	0
43	0	0	0	1	0	0	0	1	0	0	1	0
44	0	0	0	1	0	0	1	0	0	1	0	0
45	0	0	0	0	1	0	0	1	0	1	0	0

Appendix V cont.

Acc	LW			LA				INL			LSL		
	7-8	9-10	11-13	300-400	401-600	601-700	701-800	19-24	25-27	28-30	16-18	19-23.1	23.2-25
1	1	0	0	1	0	0	0	1	0	0	1	0	0
2	1	0	0	1	0	0	0	1	0	0	1	0	0
3	0	1	0	0	1	0	0	0	0	1	0	0	1
4	0	0	1	0	0	0	1	1	0	0	0	1	0
5	0	0	1	0	0	0	1	1	0	0	0	1	0
6	0	0	1	0	0	1	0	0	0	1	0	1	0
7	0	0	1	0	0	0	1	0	1	0	0	1	0
8	0	1	0	0	0	1	0	1	0	0	0	1	0
9	0	1	0	0	1	0	0	0	0	1	0	0	1
10	1	0	0	1	0	0	0	0	1	0	0	1	0
11	0	1	0	0	1	0	0	1	0	0	0	1	0
12	1	0	0	0	1	0	0	1	0	0	0	1	0
13	1	0	0	1	0	0	0	0	1	0	1	0	0
14	0	1	0	0	1	0	0	1	0	0	0	1	0
15	0	0	1	0	0	0	1	1	0	0	0	1	0
16	1	0	0	1	0	0	0	1	0	0	0	1	0
17	0	0	1	0	0	1	0	0	1	0	0	1	0
18	0	0	1	0	0	1	0	0	1	0	0	1	0
19	0	0	1	0	0	1	0	0	1	0	0	0	1
20	0	0	1	0	0	0	1	0	1	0	0	0	1
21	0	0	1	0	1	0	0	1	0	0	0	1	0
22	0	1	0	0	1	0	0	1	0	0	1	0	0
23	0	0	1	0	0	1	0	0	0	1	0	1	0
24	0	1	0	0	1	0	0	1	0	0	0	1	0
25	0	1	0	0	1	0	0	0	0	1	0	1	0
26	0	0	1	0	0	1	0	0	0	1	0	0	1
27	0	1	0	0	1	0	0	0	1	0	0	1	0
28	0	1	0	0	1	0	0	0	1	0	0	0	1
29	0	0	1	0	0	1	0	0	1	0	0	1	0
30	0	1	0	0	1	0	0	1	0	0	0	1	0
31	0	1	0	0	1	0	0	1	0	0	0	1	0
32	0	1	0	0	1	0	0	0	0	1	0	1	0
33	1	0	0	0	1	0	0	0	0	1	0	1	0
34	0	1	0	0	1	0	0	0	0	1	0	1	0
35	0	1	0	0	1	0	0	0	1	0	0	1	0
36	0	0	1	0	1	0	0	0	1	0	0	1	0
37	0	1	0	0	1	0	0	1	0	0	0	1	0
38	0	1	0	0	1	0	0	0	1	0	0	1	0
39	0	1	0	0	1	0	0	0	1	0	0	1	0
40	0	1	0	0	1	0	0	0	1	0	0	1	0
41	0	1	0	0	1	0	0	1	0	0	0	1	0
42	0	1	0	0	1	0	0	0	1	0	0	1	0
43	0	0	1	0	0	1	0	0	0	1	0	0	1
44	0	1	0	0	1	0	0	0	0	1	0	1	0
45	0	1	0	0	1	0	0	0	1	0	0	0	1

Appendix V cont.

Acc	PHt				PL				PW			
	185-195	200-294	300-322	323-373	10-12	13-16	17-20	21-25	8-8.5	8.6-10	10.5-11.5	12-14
1	1	0	0	0	1	0	0	0	1	0	0	0
2	1	0	0	0	1	0	0	0	1	0	0	0
3	0	1	0	0	0	0	1	0	0	1	0	0
4	0	0	1	0	0	1	0	0	0	0	1	0
5	0	1	0	0	0	1	0	0	0	0	1	0
6	0	0	1	0	0	1	0	0	0	1	0	0
7	0	0	0	1	0	1	0	0	0	1	0	0
8	0	1	0	0	0	0	1	0	0	1	0	0
9	0	1	0	0	0	0	1	0	0	0	1	0
10	0	1	0	0	0	1	0	0	0	0	1	0
11	0	0	1	0	0	1	0	0	0	0	1	0
12	0	1	0	0	0	0	1	0	0	0	1	0
13	0	1	0	0	1	0	0	0	0	1	0	0
14	0	1	0	0	0	0	1	0	0	1	0	0
15	0	1	0	0	0	0	0	1	0	0	0	1
16	0	1	0	0	0	0	1	0	0	0	0	1
17	0	1	0	0	0	0	1	0	0	0	1	0
18	0	1	0	0	0	0	1	0	0	1	0	0
19	0	1	0	0	0	0	1	0	0	0	0	1
20	0	1	0	0	0	0	0	1	0	0	0	1
21	0	1	0	0	0	0	0	1	0	0	1	0
22	0	1	0	0	0	0	1	0	0	1	0	0
23	0	1	0	0	0	1	0	0	0	1	0	0
24	0	1	0	0	0	0	1	0	0	1	0	0
25	0	1	0	0	0	0	0	1	0	0	0	1
26	0	1	0	0	0	0	0	1	1	0	0	0
27	0	1	0	0	0	0	0	1	0	0	1	0
28	0	1	0	0	0	0	1	0	0	0	0	1
29	0	1	0	0	0	0	0	1	0	0	0	1
30	0	1	0	0	0	0	0	1	0	1	0	0
31	0	1	0	0	0	0	0	1	0	1	0	0
32	0	0	1	0	0	0	1	0	0	0	1	0
33	0	0	1	0	0	0	0	1	0	1	0	0
34	0	0	1	0	0	0	1	0	0	0	0	1
35	0	1	0	0	0	1	0	0	0	0	1	0
36	0	0	1	0	0	0	1	0	0	0	0	1
37	0	0	0	1	0	1	0	0	0	0	0	1
38	0	0	1	0	0	0	1	0	0	1	0	0
39	0	1	0	0	0	0	0	1	0	1	0	0
40	0	0	0	1	0	0	0	1	0	1	0	0
41	0	0	1	0	0	0	1	0	0	1	0	0
42	0	1	0	0	0	0	1	0	0	0	1	0
43	0	0	0	1	0	1	0	0	0	1	0	0
44	0	0	0	1	0	1	0	0	0	1	0	0
45	0	0	1	0	0	1	0	0	0	1	0	0

Appendix V cont.

Acc	NPB			HWt				GYPP			
	50-77	80-100	101-135	80-100	110-125	130-150	160-200	60-75	80-100	105-120	125-150
1	1	0	0	1	0	0	0	1	0	0	0
2	1	0	0	1	0	0	0	1	0	0	0
3	0	1	0	0	1	0	0	0	1	0	0
4	0	0	1	0	0	0	1	0	0	1	0
5	0	0	1	0	1	0	0	0	1	0	0
6	1	0	0	0	0	1	0	0	0	1	0
7	0	1	0	0	0	1	0	0	1	0	0
8	0	1	0	0	0	0	1	0	0	1	0
9	0	1	0	0	0	0	1	0	0	1	0
10	0	0	1	0	0	0	1	0	0	0	1
11	0	0	1	0	0	0	1	0	0	1	0
12	1	0	0	0	0	1	0	0	0	1	0
13	0	1	0	1	0	0	0	1	0	0	0
14	0	1	0	0	0	0	1	0	0	1	0
15	1	0	0	0	0	0	1	0	0	1	0
16	0	1	0	0	1	0	0	0	1	0	0
17	1	0	0	0	0	1	0	0	1	0	0
18	0	1	0	0	0	1	0	0	1	0	0
19	0	1	0	0	0	0	1	0	0	1	0
20	0	0	1	0	0	0	1	0	0	0	1
21	0	1	0	0	1	0	0	0	1	0	0
22	0	1	0	0	0	1	0	0	0	1	0
23	1	0	0	0	0	0	1	0	0	0	1
24	1	0	0	0	0	1	0	0	1	0	0
25	1	0	0	0	0	0	1	0	0	1	0
26	0	1	0	0	0	1	0	0	1	0	0
27	1	0	0	0	0	1	0	0	1	0	0
28	1	0	0	0	1	0	0	0	1	0	0
29	0	1	0	0	0	0	1	0	0	1	0
30	1	0	0	0	1	0	0	1	0	0	0
31	0	0	1	0	0	0	1	0	0	1	0
32	1	0	0	1	0	0	0	1	0	0	0
33	1	0	0	1	0	0	0	1	0	0	0
34	0	1	0	0	1	0	0	0	1	0	0
35	0	0	1	0	0	0	1	0	0	0	1
36	0	0	1	0	0	0	1	0	0	1	0
37	0	0	1	0	0	1	0	0	0	1	0
38	0	0	1	0	0	0	1	0	0	1	0
39	1	0	0	1	0	0	0	1	0	0	0
40	0	1	0	0	0	1	0	0	0	1	0
41	0	1	0	0	0	0	1	0	0	0	1
42	0	1	0	0	0	0	1	0	0	0	1
43	0	0	1	0	0	0	1	0	0	0	1
44	0	0	1	0	0	0	1	0	0	0	1
45	0	0	1	0	0	0	1	0	0	0	1

Appendix V cont.

Acc	TSWt				TP			KNPP			
	20-22	24-26	27-29	30-34	60-70	71-75	76-85	20	30	40	50
1	0	0	1	0	1	0	0	1	0	0	0
2	0	0	0	1	1	0	0	1	0	0	0
3	0	0	1	0	1	0	0	1	0	0	0
4	0	0	1	0	1	0	0	0	0	1	0
5	0	0	0	1	0	1	0	1	0	0	0
6	1	0	0	0	1	0	0	0	0	0	1
7	1	0	0	0	0	1	0	0	0	1	0
8	0	0	1	0	1	0	0	0	0	1	0
9	0	0	1	0	1	0	0	0	0	1	0
10	0	0	0	1	0	0	1	0	0	1	0
11	0	0	1	0	0	1	0	0	0	1	0
12	0	1	0	0	0	0	1	0	0	1	0
13	1	0	0	0	0	0	1	1	0	0	0
14	0	1	0	0	1	0	0	0	0	1	0
15	0	0	1	0	1	0	0	0	0	1	0
16	0	0	0	1	0	0	1	0	1	0	0
17	1	0	0	0	1	0	0	0	0	1	0
18	0	0	1	0	1	0	0	0	1	0	0
19	0	1	0	0	1	0	0	0	0	1	0
20	0	1	0	0	1	0	0	0	0	0	1
21	1	0	0	0	1	0	0	0	0	1	0
22	0	0	1	0	0	1	0	0	1	0	0
23	0	1	0	0	0	0	1	0	0	0	1
24	1	0	0	0	0	0	1	0	0	0	1
25	0	1	0	0	0	1	0	0	0	0	1
26	0	1	0	0	1	0	0	0	1	0	0
27	0	0	1	0	1	0	0	0	1	0	0
28	1	0	0	0	0	1	0	0	0	1	0
29	0	1	0	0	1	0	0	0	0	1	0
30	0	0	1	0	1	0	0	1	0	0	0
31	0	0	1	0	1	0	0	0	1	0	0
32	1	0	0	0	0	1	0	0	1	0	0
33	1	0	0	0	1	0	0	0	1	0	0
34	0	0	1	0	1	0	0	0	1	0	0
35	0	1	0	0	1	0	0	0	0	0	1
36	0	0	0	1	1	0	0	0	1	0	0
37	0	0	1	0	0	1	0	0	0	1	0
38	0	0	0	1	1	0	0	0	1	0	0
39	1	0	0	0	0	1	0	0	1	0	0
40	0	0	0	1	0	0	1	0	1	0	0
41	0	0	0	1	0	0	1	0	0	1	0
42	0	0	0	1	0	0	1	0	0	1	0
43	0	0	0	1	0	0	1	0	0	1	0
44	0	0	0	1	0	1	0	0	0	1	0
45	0	0	1	0	0	1	0	0	0	1	0

Appendix VI An example of the pair-wise genetic distances estimated between some of the accessions based on morho-agronomical data.

First Row	Second Row	Actual Distance	Dendrogram Distance	Actual Difference	Percent Difference
1	2	0.354	0.354	0.000	0
1	3	0.629	0.628	0.001	0.14
1	4	0.629	0.628	0.001	0.14
1	5	0.661	0.628	0.033	5.01
1	6	0.629	0.628	0.001	0.14
1	7	0.677	0.628	0.049	7.2
1	8	0.612	0.628	-0.016	-2.6
1	9	0.612	0.628	-0.016	-2.6
1	10	0.645	0.628	0.017	2.67
1	11	0.645	0.628	0.017	2.67
1	12	0.595	0.628	-0.033	-5.57
1	13	0.540	0.509	0.031	5.68
1	14	0.645	0.628	0.017	2.67
1	15	0.629	0.628	0.001	0.14
1	16	0.612	0.628	-0.016	-2.6
1	17	0.645	0.628	0.017	2.67
1	18	0.645	0.628	0.017	2.67
1	19	0.677	0.628	0.049	7.2
1	20	0.629	0.628	0.001	0.14
1	21	0.645	0.628	0.017	2.67
1	22	0.595	0.628	-0.033	-5.57
1	23	0.677	0.628	0.049	7.2
1	24	0.645	0.628	0.017	2.67
1	25	0.677	0.628	0.049	7.2
1	26	0.661	0.628	0.033	5.01
1	27	0.595	0.628	-0.033	-5.57
1	28	0.692	0.628	0.064	9.24
1	29	0.629	0.628	0.001	0.14
1	30	0.520	0.628	-0.108	-20.73
1	31	0.612	0.628	-0.016	-2.6
1	32	0.629	0.628	0.001	0.14
1	33	0.577	0.628	-0.051	-8.82
1	34	0.604	0.628	-0.024	-4.05
1	35	0.692	0.628	0.064	9.24
1	36	0.661	0.628	0.033	5.01
1	37	0.629	0.628	0.001	0.14
1	38	0.661	0.628	0.033	5.01
1	39	0.629	0.628	0.001	0.14
1	40	0.645	0.628	0.017	2.67
1	41	0.595	0.628	-0.033	-5.57
1	42	0.707	0.628	0.079	11.15
1	43	0.677	0.628	0.049	7.2
1	44	0.661	0.628	0.033	5.01
1	45	0.677	0.628	0.049	7.2

Appendix VII An example of the pair-wise genetic distances estimated between some of the accessions based on AFLP data.

First Row	Second Row	Actual Distance	Dendrogram Distance	Actual Difference	Percent Difference
1	2	0.445	0.459	-0.014	-3.06
1	3	0.429	0.424	0.004	1
1	4	0.477	0.477	0.000	0.01
1	5	0.413	0.413	0.000	0
1	6	0.570	0.532	0.039	6.78
1	7	0.522	0.532	-0.010	-1.83
1	8	0.703	0.656	0.047	6.7
1	9	0.501	0.622	-0.121	-24.09
1	10	0.642	0.622	0.021	3.24
1	11	0.566	0.580	-0.014	-2.52
1	12	0.654	0.656	-0.002	-0.38
1	13	0.578	0.622	-0.043	-7.47
1	14	0.612	0.622	-0.010	-1.57
1	15	0.677	0.656	0.021	3.07
1	16	0.624	0.622	0.002	0.35
1	17	0.640	0.622	0.018	2.81
1	18	0.641	0.622	0.019	3.03
1	19	0.606	0.622	-0.016	-2.57
1	20	0.698	0.656	0.042	6.01
1	21	0.608	0.622	-0.014	-2.32
1	22	0.632	0.622	0.011	1.72
1	23	0.628	0.622	0.007	1.04
1	24	0.624	0.622	0.002	0.35
1	25	0.636	0.622	0.014	2.25
1	26	0.637	0.622	0.016	2.47
1	27	0.613	0.622	-0.008	-1.32
1	28	0.648	0.622	0.026	4.09
1	29	0.659	0.622	0.038	5.7
1	30	0.637	0.622	0.015	2.38
1	31	0.608	0.622	-0.014	-2.32
1	32	0.662	0.622	0.040	6.1
1	33	0.582	0.580	0.002	0.27
1	34	0.628	0.622	0.007	1.04
1	35	0.651	0.622	0.029	4.5
1	36	0.702	0.656	0.046	6.53
1	37	0.690	0.622	0.069	9.94
1	38	0.663	0.656	0.007	1.09
1	39	0.697	0.656	0.041	5.83
1	40	0.685	0.656	0.029	4.2
1	41	0.624	0.615	0.009	1.44
1	42	0.669	0.622	0.047	7.06
1	43	0.698	0.656	0.042	6.01
1	44	0.708	0.656	0.052	7.38
1	45	0.706	0.656	0.050	7.04

Appendix VIII An example of the pair-wise genetic distances estimated between some of the accessions based on microsatellites data

First Row	Second Row	Actual Distance	Dendrogram Distance	Actual Difference	Percent Difference
1	2	0.305	0.305	0.000	0
1	3	0.403	0.430	-0.027	-6.69
1	4	0.647	0.624	0.023	3.5
1	5	0.550	0.624	-0.074	-13.55
1	6	0.647	0.624	0.023	3.5
1	7	0.629	0.624	0.004	0.71
1	8	0.591	0.611	-0.020	-3.44
1	9	0.571	0.611	-0.040	-7.07
1	10	0.629	0.611	0.018	2.84
1	11	0.647	0.611	0.036	5.57
1	12	0.647	0.611	0.036	5.57
1	13	0.591	0.611	-0.020	-3.44
1	14	0.647	0.611	0.036	5.57
1	15	0.715	0.611	0.104	14.59
1	16	0.591	0.611	-0.020	-3.44
1	17	0.665	0.611	0.054	8.09
1	18	0.665	0.611	0.054	8.09
1	19	0.629	0.611	0.018	2.84
1	20	0.610	0.611	-0.001	-0.15
1	21	0.506	0.611	-0.105	-20.79
1	22	0.591	0.611	-0.020	-3.44
1	23	0.647	0.611	0.036	5.57
1	24	0.699	0.611	0.088	12.58
1	25	0.699	0.611	0.088	12.58
1	26	0.610	0.611	-0.001	-0.15
1	27	0.571	0.611	-0.040	-7.07
1	28	0.629	0.611	0.018	2.84
1	29	0.571	0.611	-0.040	-7.07
1	30	0.629	0.611	0.018	2.84
1	31	0.610	0.644	-0.034	-5.51
1	32	0.665	0.644	0.021	3.18
1	33	0.682	0.644	0.038	5.63
1	34	0.550	0.644	-0.094	-17.05
1	35	0.699	0.644	0.055	7.91
1	36	0.682	0.644	0.038	5.63
1	37	0.665	0.644	0.021	3.18
1	38	0.665	0.644	0.021	3.18
1	39	0.699	0.644	0.055	7.91
1	40	0.610	0.611	-0.001	-0.15
1	41	0.550	0.611	-0.061	-11.11
1	42	0.665	0.611	0.054	8.09
1	43	0.682	0.611	0.071	10.42
1	44	0.647	0.644	0.003	0.53
1	45	0.647	0.644	0.003	0.53

Appendix IX Partial view of the sensory evaluation of the *injera*. (Ethiopian and Eritrean students evaluating sorghum *injera* at the Sensory Laboratory, UFS, South Africa).