ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

GENETIC DIVERSITY IN SORGHUM *(SORGHUM BICOLOR (L.)* MOENCH) GERMPLASM FROM ETHIOPIA AND ERITREA

By

Amsalu Ayana

A Thesis

Presented to the School of Graduate Studies of the Addis Ababa University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biology

Addis Ababa, Ethiopia

June 2001

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

Genetic Diversity in Sorghum (*Sorghum bicolor* (L.) Moench) Germplasm From Ethiopia and Eritrea.

By
Amsalu Ayana

A Thesis Presented to the School of Graduate Studies of the Addis Ababa University in Partial Fulfillment of the Degree of Doctor of Philosophy in Biology

Approved by Examining Board:

Dr Torbjörn Säll (External Examiner) _____

Dr Aberra Debelo (Internal Examiner) _____

Prof. Emeritus Waheeb Heneen (Advisor) _____

Prof. Tomas Bryngelsson (Advisor) _____

Prof. Endashaw Bekele (Advisor) _____

Dr Kifle Dagne (Chairman) _____

June 2001

Dedication

This dissertation is dedicated to my mother

Acknowledgements

I express my sincere gratitude and indebtedness to Professor Endashaw Bekele for accepting me as a Ph.D. student, for his invaluable help in designing and leading my Ph.D. research project as well as providing facilities. I acknowledge and value his competent guidance and unlimited encouragement throughout my study period. The illuminating discussions that I have had with him on many occasions have been very helpful. I am deeply grateful to him for providing helpful suggestions and comments on the manuscripts leading to this dissertation. His thoughtfully recommended seminar topics introduced me to the fascinating areas of population and molecular genetics. They helped me to strengthen my plant breeding background. His ability to facilitate working conditions and sense of understanding have been key factors, when I got stuck at times. His scholar responsibility and kind support did not end there. He successfully established our collaboration with two kind and competent professors — Professor Emeritus Waheeb Heneen and Professor Tomas Bryngelsson, both at the Department of Crop Science, Alnarp (previously at the Department of Plant Breeding Research, Svalöv), the Swedish University of Agricultural Sciences (SLU), Sweden. I do remember that this collaboration has been essential to reach this stage today and to translate my long-standing ambition to be materialized.

Therefore, I owe an enormous debt of gratitude to Professor Emeritus Waheeb Heneen and Professor Tomas Bryngelsson for their kind collaboration and unwavering encouragement. Their kind willingness to collaborate with us not only firmly established my admission to Ph.D. studies at AAU under a program popularly known as a "sandwich program" but also has given me the opportunity to enjoy the expertise and academic atmosphere of two nations. I am deeply grateful to both of them for their competent and

expert supervision, kind support, and the many insightful comments and suggestions that they have had made on the earlier versions of the published papers as well as on this dissertation. I appreciate their insight, kind consideration and patience, while giving me prolongation two times beyond the time allotted for me during my two study visits in Sweden. These two prolongation periods have meant a great deal to me. They were the periods during which I have harvested what I cultivated during the allotted times. The first prolongation period was specially remembered for entering into the ups and downs of publishing papers and generation of RAPD data. The results have been sweet! While I recall Professor Emeritus Heneen's inspiring advice towards publishing my Ph.D. works as fast I could and feel happy for getting sweet results, I regret that he is not a co-author on any of the published papers. I wish I had done my research on sorghum chromosome genetics, where he would have been one. But it was too late! The second period was specially assigned for writing the Ph.D. dissertation, the witness to get the long-desired Ph.D. In my case, not only that, this second prolongation period coincided with the movement of the Department of Plant Breeding Research from Svalöv to Alnarp. The fact that I had to live Burlöv, which is a bit far away from the Alnarp campus, forced me to learn riding a cycle. I count this also as one of my achievements in the course of craving for a Ph.D.

I am also deeply grateful to both Professor Emeritus Waheeb Heneen and Professor Bryngelsson for the maximum effort they have made for me to get treatment against my asthma. I am especially grateful to Professor Emeritus Waheeb Heneen and his family for the splendid hospitality that I have enjoyed on many occasions. Certainly, they are as generous as one's father and mother. Tack för allt!

I often hear sayings that there is no fun without money and if there is money there is a way. While agreeing to these sayings, I would like to emphasize the importance of human will for money to bring a success of any kind.

This study is fully supported by the Swedish Agency for Research Cooperation with Developing Countries/Swedish International Development Agency. The School of Graduate Studies (SGS) and Department of Biology, AAU, Ethiopia, and the Department of Crop Science, Alnarp, SLU, Sweden, are acknowledged for administering the finances and hosting the study. Similarly, I deeply acknowledge and appreciate the Jimma College of Agriculture (JCA), Jimma University, Jimma, Ethiopia, for providing me study leave with payment as part of its staff development program. I am also grateful to JCA for letting my family to stay in the College's apartment with nominal payment. That has given me the feeling that my family members are secured and supported, despite the extended loneliness they had to suffer during the period of my studies

I am indebted to the following institutions/organizations: Institute of Biodiversity Conservation and Research (IBCR), Ethiopia, for providing sorghum germplasm; Ethiopian Agricultural Research Organization (EARO) (previously Institute of Agricultural Research) in general, and to Nazareth/Melkassa, Jimma and Arsi Negele Agricultural Research Centers in particular for providing sorghum germplasm, research facilities and support during my field research; the National Herbarium, AAU, for providing facilities and support during my field trips for collecting wild sorghum; Research and Publication Office (RPO), AAU, for providing computer facilities to write my manuscripts and seminar papers. I remain very grateful to the Center Managers and staff members at Melkassa, Jimma and Arsi Negele Agricultural Research Centers, as well as to the many daily and contract laborers, who helped me in various aspects to successfully complete my field research at the respective centers. Here it is appropriate to extend my special thanks to my three field assistants — Mrs. Tirsit Tesfaye (at Melkassa), Mr. Hailu Kebet Yimer (at Arsi Negele), and Mr. Tesgaye Geremew (at Jimma) — and Mr. Neway Mengistu for overseeing my work at Melkassa in my absence.

Mrs. Britt Greén patiently taught me DNA extraction and introduced me to the use of the 'magic machine', Polymerase Chain Reaction (PCR), and thereby to the random amplified polymorphic DNA (RAPD) assay. Not only has Mrs. Greén been extremely helpful for swift acquisition of chemicals and PCR reagents during my laboratory work in Svalöv, but also for the laboratory work in Addis Ababa. Likewise, Mrs. Ann-Sofie Fält cheerfully introduced me to starch gel electrophoresis of isozymes. I sincerely acknowledge and appreciate their pedagogical and friendly approaches, patience, and always willingness to help. I am grateful to Mrs. Ann-Sofie Fält and her family for the superb hospitality that I enjoyed on several occasions. I am grateful to Mrs. Kerstin Brismar for photography and to Mrs. Helén Lindgren for ordering chemicals when Mrs. Greén was not in place. Mr. Göran Olsson has been rendering invaluable help in the greenhouse at Svalöv and has also been promptly responding to my repeated requests for help when I got a problem running computers, especially printing dendrograms from NTSYS program. Tack för allt, Göran! I am also indebted to Mr. Michael Tadesse for his technical assistance and help in transporting liquid nitrogen during my laboratory work in the Genetics Research Laboratory, Department of Biology, AAU. I am thankful to Mr. Melkau Wondafrash and Mrs. Altash Tefera for their help in the National Herbarium. I am also thankful to Mr. Abdul and Mr. Terefe from SGS, AAU and Mr. Zbigniew Karasiewiz at the Department of Crop Science, Alnarp, Sweden, for their driving services.

I have greatly benefited from the statistical discussions I have had with Mr. Girma Taye, EARO; Mr. Zerihum Tadesse, International Livestock Research Institute (ILRI), Addis Ababa; and Dr Jan-Erik Englund, SLU, Alnarp.

I acknowledge using the library facilities of AAU, EARO, IBCR, ILRI and Armuer Hanssen Research Institute, Addis Ababa, as well as those of SLU and Svalöv-Weibull, Sweden. My special thanks go to the library staff of SLU at Alnarp, who helped me to get in

touch with journal articles and books from different libraries in and outside Sweden. Without their sincere services, it would not have been possible to get the reference materials that have been of such importance for the write-up of several manuscripts and this dissertation. In line with this, I would also like to thank Dr. Zemede Asfaw, Department of Biology, AAU, who has been an excellent source to several articles and useful books.

Throughout my study period, I have received invaluable help from my fellow Ph.D. students, postdoctoral fellows and researchers in both Svalöv/Alnarp and Addis Ababa. Many times I took some minutes from their busy laboratory or office work to respond thoughtfully to my inquires and to enlighten my ambiguities. They all have taught me a great deal and their ideas and work have inspired my own. Many thanks to all of them. I am afraid that a listing of names would be too long! However, I would be remiss if I do not mention Dr. Genlou Sun for help in RAPD protocol development and Dr. Karin Persson for introducing me to BIOSYS for isozyme data analysis.

Thanks are due to many anonymous reviewers who have made insightful comments, suggestions, and given constructive criticism on our published papers. Though indirect, their ideas considerably improved this dissertation and I acknowledge that I have learnt much from them.

I am thankful to the past and present Heads and Graduate Student Advisors at the Department of Biology and past and present Deans of SGS, AAU, for facilitating my studies. My special thanks goes to Dr. Kifle Dagne and Dr. Seyoum Mengistou for exciting encouragement and understanding. I am also thankful to Professor Sebsibe Demissew, Drs Enserum Kelebssa and Sileshi Nemomessa for their valuable help and encouragement. The past and present secretaries at the Department of Biology, SGS and RPO, AAU, have been instrumental in terms of administrative tasks. Mrs. Cecilia Åkesson and Mrs. Marianne Varfeldt have played similar roles in Svalöv and later at Alnarp, Sweden. They all deserve my

special thanks. I wish to thank the finance staff of AAU at Sidist Kilo and Arat Kilo for their services regarding the release of my research money, although I get cross with them on several occasions.

My very special thanks go to my friends Mr. Bedassa Tadesse (now Ph.D. student in USA), Mr. Kaba Urgessa, Mr. Fikre Lemessa and the late Mr. Asmare Hagos for cheerful encouragement, considerable help and for taking care of my family throughout my study period. I am deeply grateful to Mrs. Adamitu Kumsa and Mrs. Mulu Shiferaw for their sisterly support and encouragement. I also gratefully acknowledge my colleagues at JCA for their encouragement and support for me as well as for my family throughout my study period. I am particularly indebted to Mr. Abegaz Beyene and Mr. Habtemariam Gebre, Drs Solomon Mogus, Abraham Bantirgu and Lulu Belete for taking my kids by their private cars to school and to hospital when they got sick. I also extend my thanks to Mr. Mekias Sintayehu and Mrs. Zewdie Sintayehu and their families as well as to Mrs. Rebeca Sintayehu and Mr. Biniyam Sintayehu for their kind support for myself and for my family during my study period. I am thankful to Dr. Lisanework Nigatu for introducing me to Professor Endashaw Bekele.

My stay during the first two years in Addis Ababa would have been boring and intolerable without my friend and roommate Mr. Tesfaye Awas. I do remember Mr. Tesfaye not only for his humorous personality but also for his nursing on several occasions when I was under asthma attack in Addis Ababa. Mr. Tilahun Teklehaimont and Mr. Menassie Gashaw, my fellow Ph.D. students in AAU, have been helpful to get access to computers, while writing my seminar papers and manuscripts: Many thanks to all of them.

Similarly, my stay in Sweden during the second visit (February to December 2000) has been pleasant and full of unforgettable events. I have had the nice company, enjoyable party and helpful academic discussions and exchanged many useful literature with my fellow Ph.D. students — Esayas Aga (Ethiopia), Karin Persson, Malin Carlsson and Fredric

(Sweden), Agnese Kolodinska (Latvia), Ramune Kuktaite (Lithuania), Beatrice Angiyo Were (Kenya), Linus Masumbuko (Tanzania), and Dr. Xin-Quan Zhang (China).

I cherish the love and encouragement of my mother, father and three of my nieces as well as my brother and sister and their families. I remain venerating their unlimited support for they are all the base and ladder for me to come up.

I am deeply obliged to my wife, Martha Sintayehu, for boundless patience. Not only has she been a source of encouragement and support throughout my study period, but also has cared for our children, who are at their very early ages. Without exaggeration, I could not have reached this stage without her absolute will to shoulder such a burden. I extremely appreciate her responsibility and seriousness of purpose.

My greatest debt is to my sons, Chali and Gudina, who missed my care at their very early childhood and tolerated my many years of absence from them. Dear Chali and Gudina! I now come home and stay with you. I promise that I will not disappear after a brief stay as I used to do before.

I encountered many events during the period of my Ph.D. study. So far I have recounted only the significantly positive ones. But I have also faced several bad events. Nothing is shocking and mournful as losing my father-in-law and mother-in-law for ever. I do remember their happy encouragement and goodwill for my success, but they failed to see my end. Let their souls rest in peace in heaven. Amen!

Contents

List of Tables

List of Tables

List of Figures

List of Figures

Appendices

Amsalu Ayana (2001). Genetic Diversity in Sorghum (*Sorghum bicolor* (L.) Moench) Germplasm From Ethiopia and Eritrea

Department of Biology, Science Faculty, Addis Ababa University, P. O. Box 1176, Addis Ababa, Ethiopia

Abstract

Information about the amount and distribution of genetic variation in germplasm collections is important for efficient management of germplasm collections and for effective utilization of such materials in plant breeding. Sorghum (*Sorghum bicolor* (L.) Moench) is one of the most important cereals worldwide and its tolerance to drought makes it particularly important in semiarid tropical regions, including Ethiopia and Eritrea. The significance of sorghum germplasm from these countries for worldwide improvement of sorghum is also well acknowledged. However, there is limited information about the amount and distribution of genetic variation in sorghum germplasm from these countries. In this thesis, the amount and patterns of distribution of genetic variation was determined in cultivated sorghum (*S. bicolor* ssp. *bicolor*) germplasm collections from Ethiopia and Eritrea using morphological traits, allozymes and random amplified polymorphic DNA (RAPD) markers. A similar investigation was made on wild sorghum (*S. bicolor* ssp. *verticilliflorum*) from Ethiopia using RAPD markers.

The results show that the materials exhibit a wide range of variation for quantitative traits and a high Shannon-Weaver diversity index ($\bar{H}$ = 0.90) for 10 qualitative traits. In contrast, the level of allozyme variation was found to be low (mean expected heterozygosity = 0.024). The level of RAPD variation was found to be intermediate both in cultivated ($\bar{H}$ =

0.53) and wild sorghum ($\bar{H}$ = 0.49). The lack of correspondence among the three levels of variation revealed by these methods points at the difficulty of making prediction of variation of one based on the other and explanations are provided for the possible causes of the disparity. Despite the disparity in the levels of variation from the three data sources (i.e., morphological traits, allozymes and RAPD markers), the structure of the variation was found to be similar for all the data sources. In all, the within accessions, regions and the within adaptation zones variation accounted for a large portion of the total variation compared with the between accessions, regions and the between adaptation zones variation. Based on these results, sampling more accessions within each regions and adaptation zones and more individuals within populations are suggested as future sampling strategies. Grouping of the materials into lowland, intermediate and highland elevations seems to be justified only for discriminant analysis of quantitative traits, otherwise, there was no convincing evidence that such materials are genetically substructured. Regions with overall high diversity are identified and suggested for *in situ* conservation. The results from quantitative traits indicate that accessions from Eritrea could be important sources of genes for early maturity. Detailed implications of the results for germplasm conservation and utilization are discussed and future lines of research are suggested.

Key words: allozymes; genetic variation; germplasm; morphological variation; RAPD; *Sorghum bicolor*

Chapter 1

General Introduction

Genetic diversity refers to the variation among alleles of genes in different individuals of populations of a species (IPGRI, 1993; Weir, 1996; Kremer, et al., 1998). While the ultimate source of genetic diversity is gene mutation, it is molded and shaped by selection, recombination, genetic drift, and migration in the face of heterogeneous environment in space and time (Falconer and Mackay, 1996; Hartl and Clark, 1997). Since natural selection chooses the best fit among the variants within a population based on the adaptability to the ever-changing environment, there can be no adaptive evolution without genetic variation. It follows that genetic diversity is an essential raw material for evolution, enabling populations of the species to survive, adapt, and evolve into new genetic variants that meet long-term changes in the environment (Avise and Hamrick, 1997; Hedrick, 2000).

Similarly, genetic diversity is a raw material in plant breeding for developing high yielding varieties and for maintaining the productivity of such varieties by incorporating genes for disease and insect resistance as well as tolerance to abiotic stresses, such as drought, cold, and salinity (Allard, 1999). Indeed, the amount of genetic diversity determines the evolutionary potential of a species and the rate as well as gain from artificial selection in breeder's materials. Consequently, a major focus of research in genetics has been to determine the amount of genetic variation in both natural and domestic populations and describing the possible mechanisms maintaining such variability (Weir, 1996).

Germplasm, also known as plant genetic resources, includes landraces, advanced or modern varieties, wild relatives of domesticated plants, wild species, genetic stocks and cloned genes (Chang, 1992; Frankel et al., 1995). Landraces, also known as traditional varieties, are cultivated forms of a crop species, which have evolved over generations of

selections by farmers (Harlan 1992; Frankel et al., 1995; Smale, 1997). According to these authorities, landraces are characterized by high genetic heterogeneity, good adaptation to local environmental conditions and low productivity. They are also noted as sources of useful genes required for further increment and maintenance of the productivity of modern varieties.

Studying the genetic diversity of sorghum (*Sorghum bicolor* (L) Moench) germplasm from Ethiopia and Eritrea attracts special interest for several reasons. Both countries are characterized by a diversity of climate, physiography, soils, vegetation, farming systems and soci-economic conditions. For instance, the Natural Resources Management and Regulatory Department of the Ministry of Agriculture of Ethiopia, divided the country into 18 main agro-ecological zones and 49 sub agro-ecological zones (MOA, 1998). The diversity in abiotic and biotic factors results in evolution and maintenance of different genetic variability of the various crops grown in the country, including sorghum. The presence of a wide range of agro-ecological conditions presents an opportunity and a challenge for germplasm conservationists and plant breeders.

Sorghum is the second preferred cereal after tef (*Eragrostis tef* (Zucc.) Trotter) for preparing '*injera*', which is the staple food in both Ethiopia and Eritrea (Gebrekidan and Gebrehiwot, 1982; Doggett and Prasada Rao, 1995). According to Central Statistics Authority of Ethiopia (Central Statistics Authority, 1999), sorghum ranks third after maize (*Zea mays* L.) and tef in total production, after maize and wheat (*Triticum spp.*) in yield per hectare and after tef and maize in area harvested. Its adaptation to stress environments such as drought, makes sorghum in a sense more important than the other cereals (Kebede and Menkir, 1987, Kebede, 1991; Debelo et al., 1995).

Moreover, Ethiopia and Eritrea come within the broad geographical range where sorghum is believed to have been first domesticated (Vavilov, 1951; Mann et al. 1983; Doggett, 1988, 1991) and where the greatest genetic diversity for both cultivated and wild

than one locus, where as allozyme is different molecular forms of an enzyme having the same catalytic function but coded by different alleles of a specific locus (Markert and Moller, 1959; Prakash et al., 1969). In other words, isozymes are the result of the expression of different genes, whereas allozymes are the result of the expression of different alleles of a single gene (Weeden, 1989). Both terms are used in this thesis. Molecular markers are those markers that reflect direct changes at the DNA sequence level.

Information on the extent and patterns of distribution of genetic variation of a crop species is essential for: (1) effective utilization of germplasm in plant breeding programs (Hayward and Breese, 1993; Moreno-Gonzalez and Cubero, 1993); (2) devising appropriate sampling procedures for germplasm collection and conservation (Allard, 1970; Marshall and Brown, 1975); (3) obtaining core collections for efficient germplasm management (Brown, 1989a, 1989b); (4) elucidating the taxonomy, evolution and origin of the crop species (de Wet et al., 1976).

Therefore, the objective of this study was to determine the amount and patterns of distribution of genetic diversity in large samples of sorghum germplasm from Ethiopia and Eritrea using morphological, isozyme and random amplified polymorphic DNA (RAPD) data.

The thesis is divided into 7 chapters. In chapter 2, literature germane to the thesis is reviewed. In chapters 3, 4 and 5, morphological, allozyme and random amplified polymorphic DNA (RAPD) diversity in cultivated sorghum (*Sorghum bicolor* ssp. *bicolor* (L) Moench) are presented, respectively. In chapter 6, RAPD diversity in wild sorghum (*S. bicolor* ssp. *verticilliflorum*) is presented. Chapter 7 deals with general discussion and conclusions. To avoid redundancy and save space, all references cited in the entire thesis are listed under one section and placed at the end of chapter 7 and appendices are put at the end of the reference section.

Chapter 2

Literature Review

Economic importance

In 1998 worldwide production of sorghum was over 63 million metric tons of grain from over 44 million hectares of land with an average yield of 1.4 metric tons per hectare (FAO, 1998). This places sorghum fifth among the major cereals of the world after wheat, rice, maize, and barley. More than 80% of the world area under sorghum is in Africa and Asia (FAO, 1998). The top ten worldwide major producers of sorghum include USA (13,207), India (9,000), Nigeria (7,516), Mexico (6,386), China (5,767), Sudan (4,891), Argentina (3,720), Ethiopia (1,083), Australia (1,068) and Burkina Faso (943) (FAO, 1998). Figures in parenthesis indicate production in thousands of metric tons.

The sorghum grain is staple food for more than 500 million people in the semi-arid tropics of Africa and Asia and is also used for preparing traditional beverages (Doggett, 1988; House, 1995b). Both the grain and the stover are also used as animal feed (Duncan, 1996). Industrial use of sorghum for making sugar, starch, syrup, alcohol and molasses is increasing (House, 1995b). The stalk is used for construction and fuel. For the traditional farmer, as is in the Harerge area of Ethiopia, every part of sorghum is utilized — the grain for food; the leaf for feed; the sweet stalk for chewing; the dry stalk for construction; and the root and the dry stalk for fuel (personal observation).

Adaptation

Sorghum is widely adapted, being grown between 40°N and 40°S of the equator (Doggett, 1988). Although sorghum is cultivated both in tropical and temperate climates, it is primarily adapted to the semi-arid parts of the world. Indeed, sorghum is best known for its good adaptation to the drought-prone semi-arid tropical (SAT) regions of the world (Doggett, 1988; Tuinstra et al., 1996, 1997; Unger and Baumhardt, 2000). SAT is defined as tropical areas where rainfall exceeds evapo-transpiration for two to seven months in the year, usually less than 1,200 mm, and where there is a long dry season (Doggett, 1976). Within the SAT, sorghum is dominant in those areas that receive between 600 and 1,000 mm rainfall per annum, but below and above this range pearl millet and maize, respectively, become more important than sorghum (House, 1995b).

Taxonomy

The genus *Sorghum* is highly diverse and comprises five sections (Garber, 1950). These are: *Chaetosorghum*, *Heterosorghum*, *Parasorghum*, *Sorghum*, and *Stiposorghum*. According to de Wet (1978), the section *Sorghum* includes (1) *S. halepense* (L.) Pers., a tetraploid (2n = 40) rhizomatous perennial species, (2) *S. Propinquum* (Kunth) Hitchc., a diploid (2n = 20) rhizomatous species, and (3) *S. bicolor* (L.) Moench, a diploid (2n = 20) annual species. Although generally treated as predominantly self-pollinating, sorghum is a wind pollinated crop with an outcrossing rate of 5-30% (Ellstrand and Foster, 1983; Doggett, 1988; Pedersen et al., 1998).

S. bicolor is further subdivided into the following three subspecies (Harlan and De Wet, 1971): (1) *S. bicolor* ssp. *bicolor*, which includes cultivated sorghum. (2) *S. bicolor* ssp.

arundinaceum including wild sorghum. (3) *S. bicolor* ssp. *drummondii* includes weedy types, which are hybrid derivatives produced wherever cultivated and wild sorghums are sympatric. They are commonly known as shattercanes (Harlan, 1992). The naming of the spontaneous subspecies *arundinaceum* was changed into *S. bicolor* ssp. *verticilliflorum* by de Wet and Prasada Rao (1986) and its consistent use is recommended by Doggett (1988). The subspecies *verticilliflorum* includes four races: *aethiopicum, arundinaceum, verticilliflorum* and *virgatum.* The races are known to differ in morphology, geographic distribution and ecological adaptation (de Wet et al., 1970, 1976). All the wild races freely hybridize with cultivated types (de Wet et al., 1970).

Harlan and De Wet (1972) partitioned *S. bicolor* ssp. *bicolor* into five basic races (*bicolor*, *caudatum*, *durra*, *guinea*, and *kafir*) and ten two-by-two combinations of the basic races, yielding a total of 15 races. The classification was based on spikelet and inflorescence morphology and proved to be simple and practical. Races differ in spikelet morphology, geographic distribution, ecological adaptation and ethnobotanical attributes (Harlan and De Wet, 1972; Shechter and de Wet, 1975). Ethiopia and India are the centers of diversity of *durra*, which is normally adapted to dry environment. The region from eastern Nigeria through Chad and the Sudan to western Ethiopia is a center of diversity for the *caudatum* race. The region from western Nigeria to Senegal is the center of diversity of *guinea*, which is largely adapted to wet and humid environments. The region from Tanzania to South Africa is the center for *kafir*. The *bicolor* lacks a specific center of diversity (Harlan and de Wet, 1972).

For practical purpose, it is relevant here to refer to Harlan and De Wet's (1971) classification of cultivated plants and their near relatives into primary, secondary and tertiary gene pools. The classification was based on extent of natural gene transfer, ease of crossability and subsequent hybrid fertility and, hence, it has direct application in plant breeding. Primary gene pool includes two subspecies, one for the cultivated races and the other for spontaneous

races. Crossing is easy, hybrids are generally fertile, genes are normally segregating and gene transfer is generally easy between the two subspecies. Secondary gene pool includes those species that can cross with the cultivated group but where recovery of fertile hybrid becomes difficult because of poor chromosome pairing. Tertiary gene pool includes those species, which may be crossed with the cultivated group but hybrids are almost completely sterile and hence gene transfer is nearly impossible with known traditional techniques. Modern techniques of molecular genetics are expected to access the tertiary gene pool.

Based on such system of classification, the primary gene pool of sorghum includes *S. bicolor* ssp. *bicolor*, *S. bicolor* ssp. *verticilliflorum*, *S. bicolor* ssp. *drummondii*, and *S. propinquum* (Acheampong et al., 1984). Available evidence indicates that the types of sorghum that form the primary gene pool of sorghum occur naturally in Ethiopia (Doggett, 1988, 1991). For instance, shattercanes are noted as the most serious weeds in the highlands of Ethiopia, where they are known as *keelo* or *sepo* in local languages and it means the fool (Damon, 1962; Harlan, 1975; Doggett, 1991). The secondary gene pool of sorghum includes only one species, i.e., *S. halepense*. The tertiary gene pool of sorghum includes the different sections of *Sorghum* listed earlier.

Origin

All lines of evidence indicate that cultivated sorghum originated from wild members of *S. bicolor* ssp. *verticilliflorum* (de Wet and Huckaby, 1967; de Wet and Harlan, 1971; Mann et al., 1983; Doggett, 1988; Duvall and Doebley, 1990; Aldrich and Doebley, 1992; Sun et al., 1994). However, unequivocal evidence is lacking about which members of the subspecies *verticilliflorum* were progenitors of the cultivated sorghum. Some authorities suggest that different races of the subspecies *verticilliflorum* gave rise to the different cultivated races. For

instance, de Wet and Huckaby (1967) suggested that the wild race *aethiopicum* gave rise to race *bicolor*; *arundinaceum* to race *guinea*; *verticilliflorum* to race *kafir* and, in turn, *kafir* to race *durra*. The suggestion was based largely on botanical similarity between the respective wild and cultivated races. However, this suggestion was discounted for two major reasons (de Wet et al., 1976). First, the morphological similarity could be the result of a considerable amount of introgression of genes between the wild and the cultivated sorghum (Doggett, 1965a, b).

Second, distribution patterns exclude *arundinaceum*, a forest grass, and *virgatum*, a desert grass, as probable progenitors of early-cultivated sorghum (Harlan 1976, 1989). This presents races *verticilliflorum* and *aethiopicum* as more probable progenitors of the early cultivated sorghum, with balance of evidence favoring the former race (de Wet et al., 1976; Harlan, 1976; Harlan and Stemler, 1976; Mann et al., 1983; Aldrich et al., 1992; Sun et al., 1994).

The distribution of these races coincides perfectly with the presumed area of early domestication of sorghum (Doggett, 1988). Thus, it is more likely that races *verticilliflorum* and/or *aethiopicum* first gave rise to the race bicolor-type. However, it is to be recalled that all wild races have greatly contributed to the evolution of cultivated sorghum through free gene flow (Doggett, 1965a, b). Harlan and de Wet (1972) described race *bicolor* as the most primitive of the five basic races of the subspecies *bicolor* and is widely distributed throughout sorghum growing regions, without marked center of diversity. With the possible exception of race *kafir*, the other races are thought to be modifications of the early bicolor-type (Mann et al., 1983).

Center of domestication

Early cultivated sorghum is believed to have originated in the African savanna (Doggett, 1988). However, convincing evidence is lacking to demonstrate where and when it was first domesticated in Africa (Wasylikowa and Dahlberg, 1999). Vavilov (1951) and Doggett and his coworkers (Doggett, 1965a, 1976, 1991; Doggett and Prasada Rao, 1995) considered Ethiopia as the center of domestication. However, the assertion that Ethiopia is the center of domestication has been criticized by Harlan and his co-workers (Harlan, 1969, 1971, 1975, 1992; de Wet et al., 1976; Stemler et al., 1975, 1977), suggesting that sorghum arose across a large area (extending from Sudan to Chad), where it was likely domesticated many times over a period of years. Specifically, Harlan (1971) presented sorghum as a typical example of non-centric crop plants.

Analysis on the origin and early cultivation of sorghum in Africa by Mann et al. (1983) tries to bring the two proposals into one by suggesting a more extended area for the origin of the early bicolor-type. Implicit in the suggestion is that the area for the early bicolor cultivation of sorghum might extend from Ethiopia through the Sudan to Chad, with the possibility of domestication taking place at different places at different times. Although specific evidence is lacking, all lines of evidence (biogeographic, isozyme and molecular data) point to the north-east quadrant of Africa as the probable center for domestication of sorghum.

It has been suggested that sorghum might have been domesticated some 3,000-5,000 before present (BP) (Doggett, 1965a; Mann et al., 1983; Bramel-Cox et al., 1995). Recent evidence from archeological investigations in southern Egypt (Wendrof et al., 1992; Wasylikowa and Dahlberg, 1999) and from ancient DNA sequencing (Rowley-Conwy et al., 1999) revealed exploitation of sorghum by humans as far back as 8000 BP. Nevertheless, such remains are found to be more similar to the present day wild than to the cultivated sorghum. Based on the

results, these authors conclude that sorghum domestication is a recent phenomenon, not much before 3000 BP.

Methods for estimating genetic diversity

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) and statistical analysis of quantitative agro-morphological traits along with eco-geographic information (de Wet et al., 1976; Murty, 1976). These methods, particularly statistical analysis of agro-morphological characters, are still widely used to quantify the amount and distribution of variation in large samples of sorghum germplasm collections (Zongo et al., 1993; Prasada Rao and Ramanatha Rao, 1995; Appa Rao et al., 1996; Teshome et al., 1997; Ayana and Bekele, 1998, 1999, 2000; Djè et al., 1998). Analysis of morphological variation is technically simple (Bretting and Widrlechner, 1995; Lee, 1995).

However, these methods of estimating genetic variation have several limitations. Discrete morphological traits, though they have high heritability, are limited in number, each being conditioned by a few genes (Karp et al., 1996, 1997; Berg and Hamrick, 1997). Thus, only a small portion of the genome could be covered. They are usually characterized by epistasis, pleiotropy and dominant-recessive relationships, further limiting their values as an ideal genetic marker (Smith and Smith, 1992; Bretting and Widrlechner, 1995).

The quantitative traits are influenced by environmental factors and show continuous variation, resulting in low heritability and high genotype by environment interactions. Consequently, it is difficult to accurately determine genetic diversity. Their complex

inheritance makes them also less amenable to statistical estimation of basic population genetic parameters (Gepts, 1995).

To overcome these problems, biochemical (primarily seed storage proteins and isozymes) and many molecular techniques have been widely used in genetic diversity assessment. As remarked by Karp et al. (1997), characterization of plant genetic resources for agro-morphological traits cannot, however, be replaced by any of the biochemical or molecular techniques or both. The results of agro-morphological, biochemical and molecular studies should be used in combination, supplementing and complementing one another.

Kafirin, the seed storage protein of sorghum, is found to be environmentally stable and provide high levels of variation (Smith and Smith, 1988). However, contrary to the situations found in other major cereals such as barley, maize and wheat, as well as grain legumes (Gepts, 1990), there are limited number of studies on variation of seed storage proteins in sorghum (Virupasha and Sastry, 1968; Shechter, 1975; Shechter and de Wet, 1975; Smith and Smith, 1988; Derose et al., 1989). This may be because of the high heterogeneity of these proteins and their complex molecular basis (Virupasha and Sastry, 1968). These studies have demonstrated that seed storage proteins can effectively discriminate the different races of cultivated sorghum. Derose et al. (1989) also demonstrated very high sequence homology between the kafirin and zein (maize seed storage protein) gene families, indicating that sorghum and maize have evolved from a common ancestor.

The most remarkable progress in attempts of measuring genetic variation was made in 1966, when gel electrophoresis of isozymes was successfully applied to reveal an enormous amount of genetic variation in *Drosophila pseudobscura* (Lewontin and Hubby, 1966) and in humans (Harris, 1966). In plants, the applicability of gel electrophoresis of isozymes for the study of genetic variation was soon demonstrated in cultivated barley (Khaler and Allard, 1970) and *Avena fatua* and *A. barbata* (Marshall and Allard, 1970). Since then, isozymes

have been widely used for estimating the amount and distribution of genetic variation in both cultivated and natural populations of a wide range of plant species and comprehensive reviews are found in Hamrick et al. (1979), Gottlieb (1981), Loveless and Hamrick (1984), Weeden (1989), Hamrick and Godt (1990, 1997) and Berg and Hamrick (1997). These reviews have drawn several general conclusions regarding the extent and patterns of distribution of isozyme variation in cultivated and natural plant populations. They have also identified important life history traits and ecological factors affecting the levels and patterns of distribution of isozyme variation. Among several factors, breeding system and geographical range have the most influential role at the population level.

As a result, predominantly self-pollinating crop plants, such as sorghum, have less genetic diversity than predominantly outcrossing crop plants, such as maize. Similarly, self-pollinating crop plants usually maintain more of the total genetic variation between populations than do outcrossing crop plants.

In sorghum, many studies have used isozyme or allozyme analysis to quantify the level of genetic variation within and between cultivated and wild populations (Tripathi et al., 1983; Warwick et al., 1984; Doebley et al., 1986; Morden et al., 1989, 1990; Ollitrault et al., 1989; Aldrich et al., 1992; Djè et al., 1998, 1999; Ayana et al., 2000c). The major conclusions that have emerged from these studies are that: (1) the extent of allozyme variation is low in sorghum compared to other cereals, (2) differentiation of accessions is more pronounced on the basis of geographical origin than on the basis of racial classification of Harlan and de Wet (1972), (3) the amount of genetic variation found within cultivated forms is less than that found in wild relatives, suggesting populations are subjected to bottleneck during domestication, (4) the level of genetic variation is higher in north-eastern Africa than in other regions both for cultivated and wild types, confirming the previous conclusions that this part of Africa is both the center for origin and diversity.

The major disadvantages of isozymes compared to molecular markers include low polymorphism and limited coverage of the genome (Weeden, 1989; Clegg, 1990). Gel electrophoresis of isozymes detects only variation for coding genes and still only a fraction of mutational events that change protein mobility (Lewontin and Hubby, 1966). Thus, differences in trinucleotide sequences that code for the same amino-acid are not detected. Similarly, changes in intron sequences that are post-transcriptionally removed from the mRNA are not detected. Thus, isozyme analysis may underestimate the actual genetic variation present in the population. Moreover, isozyme patterns may be influenced by developmental stages of the plant tissue (Karp et al., 1997; Hash and Bramel-Cox, 2000). The solution is to use complementary techniques that include one or more of the various molecular markers (Clegg, 1990; Tanksley et al., 1989; Breeting and Widrlechner, 1995; Karp et al., 1996, 1997).

The most remarkable development in search for more desirable methods for determining genetic variation is the advent of molecular markers in the late 1970s. The discovery of restriction enzymes (Linn and Arber, 1968; Meselson and Yuan, 1968) and a method for the separation of DNA restriction fragments electrophretically, followed by immobilization on nitrocellulose membranes and detection by radioactive probe labeling (Southern, 1975), have lead to the use of restriction fragment length polymorphism (RFLP) for measuring genetic variation more accurately than ever before (Tanksley et al., 1989; Paterson et al., 1991). Variation in RFLP arises because of length difference (i.e. molecular weight), which, in turn, arises because of mutation that alters nucleotide sequences at the restriction sites (Paterson et al., 1991). The mutation event could be gain or loss of restriction site, as well as inversion or insertion (Tanksley et al., 1989).

Depending on the sources of the probes (cDNA or genomic DNA), RFLPs could reveal variation in coding (low-copy number genes) or both coding and non-coding sequences

(high copy number sequences). Similarly, depending on the objective of the study, RFLP could be used to investigate variation in nuclear, chloroplast, and mitochondrial DNA (nDNA, cpDNA and mtDNA, respectively). Crop plants have diverged from their wild-weedy relatives only during the last ten to fifteen thousand years (Harris, 1989; Harlan, 1992). Therefore, nDNA RFLPs are suitable since nDNA generally evolves faster than the cpDNA and mtDNA, whereas the latter ones are preferred to resolve taxonomic relationships since they are relatively more conserved (Doebley and Wendel, 1989). In general, RFLPs have such desirable attributes as high reproduciability, high polymorphism (esp. relative to isozymes), co-dominant inheritance, selective neutrality and freedom from epistatic and pleiotropic effects (Clegg, 1990; Jones et al., 1997; Karp et al., 1997).

In sorghum, nDNA RFLPs (Aldrich and Doebley, 1992; Tao et al., 1993; Deu et al., 1994; Vierling et al., 1994; Cui et al., 1995; Oliveira et al., 1996), mtDNA RFLPs (Deu et al., 1995) and cpDNA RFLPs (Duvall and Doebley, 1990; Aldrich and Doebley, 1992) have been used to investigate the level of genetic variation and to establish genetic relationships between cultivated and wild forms. Results showed a higher level of polymorphism and a greater number of alleles per locus among accessions within races and among geographic groups than those obtained in isozyme studies. Moreover, these studies provided supporting evidence that: (1) northeastern Africa is the center of diversity both for cultivated and wild sorghum, (2) there is high level of differentiation among the accessions on the basis of geographic origin compared to racial classification, which is based on morphological traits, (3) sorghum is characterized by a low level of polymorphism compared to other cereals such as maize and barley, and (4) *S. bicolor* ssp. *verticilliflorum* is the progenitor of cultivated sorghum.

However, RFLP is expensive to be routinely applied for germplasm characterization at intraspecific level and it is also of generally low polymorphism (Waugh and Powell, 1992). Moreover, since RFLP is detected by probe-hybridization, it involves radioactivity, requiring

more expertise and safety conditions. It also requires relatively more sample DNA than polymerase chain reaction (PCR) based markers. Perhaps these might be the reasons why RFLP analysis has been much more widely used for construction of genetic linkage maps (Hulbert et al., 1990; Binelli et al., 1992; Whitkus et al., 1992; Chittenden et al., 1994; Xu et al., 1994; Bennetzen, 1995) than for estimation of genetic diversity in germplasm collections of sorghum. Generally RFLP also exhibits lower level of polymorphism than PCR-based markers.

The invention of PCR has led to the rapid development of increasing number of molecular markers (Saiki et al., 1985; Mullis et al., 1986). In PCR, a segment of the DNA molecule of known or unknown function is amplified many times by the enzyme DNA polymerase (Saiki et al., 1985). Three temperature stages are involved in the PCR assay. These are: (1) denaturation temperature, at which the double stranded DNA is separated apart, (2) annealing temperature, at which primers with complementary sequences are annealing or priming at two ends of the DNA segment to be amplified, and (3) polymerization temperature, at which DNA polymerase synthesizes a new double stranded DNA in the presence of free nucleotides and $MgCl_2$ in the PCR buffer system.

During the initial development of PCR, Klenow polymerase was used. This enzyme was thermo-sensitive (i.e., liable to denaturation by heat treatment) and a fresh enzyme had to be added after each cycle of amplification. A major breakthrough came when a thermostable DNA polymerase (i.e., resistant to denaturation by heat treatment) was isolated and purified from the bacterium *Thermophilus aquatics*, which lives in hot springs (Mullis and Faloona, 1987; Saiki et al., 1988). The use of TAQ polymerase allowed automation of the three temperature stages. In 30-45 cycles, a very large quantity of the target DNA segment would be amplified. The amplification products are separated electrophoretically, mostly in agarose

gels. After staining with ethidium bromide, the bands are detected under ultraviolet light (Saiki et al., 1985, 1988).

Unlike in the RFLP analysis, a small quantity of template DNA is required in PCR assays (Williams et al., 1990). Since the quantity of the amplified DNA is high enough to be detected by staining with fluorescent chemicals, such as ethidium bromide, there is no need to develop probes by cloning and detection by radioactive labeling (Williams et al., 1990). This makes the PCR assay, simple and safe compared to RFLP.

In the beginning, it was a requirement to have a prior knowledge of the sequences of DNA segment to be amplified and to design two complementary primers (Saiki et al., 1985). This limitation was overcome in 1990 with the advent of random amplified polymorphic DNA (RAPD) (Williams et al., 1990) and arbitrary-primed PCR (AP-PCR) (Welsh and McClelland, 1990). In RAPD, short primers (usually 10 base pairs (bp)), with a guanine-cytosine (i.e., G-C) content of at least 50%, are used and amplification products are separated in agarose gels and detected with ethidium bromide staining (Williams et al., 1990).

In contrast, in AP-PCR slightly longer primers are used and amplification products are separated in polyacrylamide gels and detected with silver staining (Welsh and McClelland, 1990). A slight modification of these two markers, DNA amplification fingerprinting (DAF) was developed in 1991 (Caetano-Anollis et al., 1991). In DAF, short random primers of 5-8 bp are used and amplification products are detected with silver staining after electrophoresis in polyacrylamide gels. RAPD is more widely used for assessing genetic diversity than AP-PCR or DAF (Karp et al., 1997).

In sorghum, RAPD has been widely used for assessing genetic variation (Tao et al., 1993; Vierling et al., 1994; Menkir et al., 1997; Ayana et al., 2000a, b). RAPD results demonstrated considerable differentiation among accessions on the basis of geographical regions than on the basis of racial classification, confirming the results obtained in the RFLP

analysis. Overall, the results showed a higher level of polymorphism than with allozymes analysis and comparable or more than that obtained in RFLP analysis. This high ploymorphism, the technical simplicity, safety and high speed of the RAPD assay all make RAPD a method of choice for analyzing a large sample of germplasm collections (Tao et al., 1993; Vierling et al., 1994; Menkir et al., 1997; Ayana et al., 2000a). However, the dominant inheritance of RAPD markers and the low reproducibility of results (due to the high sensitivity of PCR reaction conditions) are the two major limiting factors in the use of RAPD markers (Hadrys et al., 1992; Tao et al., 1993).

Alternatively, other PCR-based markers, such as simple sequence repeats (SSR) or microsatellites (Brown et al., 1996; Djé et al., 1999, 2000; Smith et al., 2000) and inter-simple sequence repeats (ISSR) (Yang et al., 1996; Taramino et al., 1997), are increasingly being used in sorghum genetic diversity analysis. These markers are co-dominantly inherited, show high level of polymorphism and results are quite reproducible (Powell et al., 1996; Jones et al., 1997). Microsatellites are tandem repeats of short (2-6 bp) DNA sequences (Tautz and Rez, 1984; Tautz, 1989). Polymorphism results from the number of repeat units and, hence, length variation. However, prior sequence information, which is usually species specific, is a requirement to design SSR primers and, hence, initial cost is high (Gupta et al., 1999).

Another marker system gaining wide acceptance in genetic diversity analysis is the amplified fragment length polymorphism (AFLP) (Vos et al., 1995; Barrett and Kidwell, 1998). It is a robust and reliable molecular marker. The number of polymorphism detected per reaction is much higher than that revealed by RFLP and RAPD. The steps in AFLP analysis include digestion of genomic DNA with two endonucleases and site-specific adapters are then ligated to the DNA fragments. Primers complementary to the adapters and to the restriction sites are designed with two or three selective nucleotides added to the 3' ends of the primers. Thus, only DNA fragments with nucleotides flanking the restriction sites that match the

selective nucleotides of the primer are amplified during PCR (Karp et al., 1996; McGregor et al., 2000). The amplified fragments are radioactively or fluorescently labeled and separated on sequencing gels. AFLP is believed to combine the advantages of RFLP and RAPD.

However, AFLP is technically more demanding and require more DNA than RAPD and it generates dominant markers rather than co-dominant markers (Boivin et al., 1999; Mueller, 1999). In sorghum, AFLP is used for linkage mapping (Boivin et al., 1999), but not for diversity assessment, at least for my knowledge.

Comparison of marker systems

A number of studies involving two or more markers in sorghum (Vierling et al., 1994; Oliveira et al., 1996; Djé et al., 1999) and other crop plants (Kongkiatngam et al., 1995; Bjørnstad et al., 1997; Chan and Sun, 1997; Fernando et al., 1997; Russell et al., 1997; Schut and Stam, 1997; Yee et al., 1997; Bahrman et al., 1999; Bohn et al., 1999; Fuentes et al., 1999; McGregor et al., 2000) as well as many reviews based on a vast body of literature (Gepts, 1993, 1995; Bretting and Widrlechner, 1995; Jones et al., 1997; Karp et al., 1997; Mohan et al., 1997; Haussmann et al., 2000) concluded that each marker system has its own advantages and disadvantages. Hence, no single method is adequate to reveal the full picture of the genetic variation available in germplasm collections. Rather, a combination of complementing methods should be used depending on the objective of the study and availability of resources (Karp et al., 1997; Mueller et al., 1999). In the following studies, morphological characters, allozymes and RAPD markers are used.

Chapter 3

Morphological Variation in Sorghum (*S. bicolor* ssp. *bicolor* (L.) Moench)

Introduction

In light of the rapidly increasing human population and expansion of agriculture into marginal areas, the importance of sorghum in the semi-arid regions of the world will increase in the future (Gebrekidan, 1987; Andrews and Bramel-Cox, 1993; House, 1995a, b). Under such situations, there will be a continuous demand for broad genetic base sorghum varieties that are high yielding and stable under abiotic and biotic stresses (Appa Rao et al., 1996). The development of high yielding and stable varieties requires a continuous supply of new germplasm as a source of desirable genes and/or gene complexes. The primary sources of such genes are landraces, introductions, weedy, and wild relatives of crop plants (Brown and Munday, 1982; Brown, 1978; Prasada Rao et al., 1989; Harlan, 1992; Frankel et al., 1995). The utilization of such germplasm requires the identification of areas of diversity of various characters of agronomic importance, especially within the centers of diversity (Bekele, 1984; Beuselinck and Steiner, 1992; Brush, 1995). Therefore, the study of genetic diversity of Ethiopian and Eritrean sorghum is worthwhile, since both countries lie within the center of diversity of the crop (Doggett, 1976, 1988, 1991; Doggett and Prasada Rao, 1995). Both countries are also constantly facing severe environmental stresses, such as drought, to which sorghum is more tolerant than other major cereals (Doggett, 1988).

With regard to morphological variation of Ethiopian sorghum germplasm, some studies have been done in the past. However, these studies have two major limitations. First, the studies did not show the patterns of character variation and distribution in the different

agro-ecological zones of the country. Since sorghum is grown throughout Ethiopia and also in Eritrea with a wide range of variation in altitude, rainfall, temperature, agricultural systems, socio-economic factors, it is essential to assess the patterns of character variation and distribution in different regions and adaptation zones. The significance of such line of studies for efficient germplasm conservation and utilization has been emphasized in many studies on other Ethiopian major cereals such as tef (Bekele, 1996), tetraploid and hexaploid wheat (Bekele, 1984; Tesfaye et al., 1991; Bechere et al., 1996) and barley (Bekele, 1983; Negassa, 1985; Demissie and Bjørnstad, 1996).

Secondly, the previous studies were limited with either using only univariate statistics (Damon, 1962; Gebrekidan and Kebede, 1977, 1978; Gebrekidan and Menkir, 1979; Abebe and Wech, 1982) or using samples from a limited geographical range (Teshome et al., 1997). Multivariate methods are useful for characterization, evaluation and classification of plant genetic resources when a large number of accessions are to be assessed for many characters of agronomic and physiological importance (Peeters and Martinelli, 1989). The application of multivariate methods for handling morphological variation in germplasm collections has been demonstrated in many crop plants. Examples among cereals include sorghum (Zongo et al., 1993; Prasada Rao and Ramanatha Rao, 1995; Teshome et al., 1997; Alagarsmamy and Chandra, 1998; Ayana and Bekele, 1999), barley (Cross, 1992), finger millet (Hussaini et al., 1977), maize (Alika et al., 1993), oat (Souza and Sorrels, 1991), rice (Kanwal et al., 1983), tef (Bekele, 1996), and wheat (Bekele, 1984; Elings, 1991; Pecetti et al., 1992; Damania et al., 1996). In whole, knowledge of the amount and patterns of distribution of genetic variation within and among germplasm collections as well as across geographical and ecological ranges of germplasm collections is important for efficient management and utilization of the germplasm.

In this chapter, the extent and patterns of distribution of morphological variation is presented for 25 agro-morphological (15 quantitative and 10 qualitative) characters in 415 sorghum accessions using univariate, bivariate and multivariate approaches of data analyses.

Materials and methods

Plant material

A total of 415 sorghum (*Sorghum bicolor* (L.) Moench) accessions, consisting of 391 landraces, 8 standard varieties and 16 introduced elite breeding lines were used in this study (Appendix 3.1). The landraces were obtained from the Institute of Biodiversity Conservation and Research (IBCR), Ethiopia, along with their passport data. The improved varieties and introduced lines were obtained from the Ethiopian Agricultural Research Organization along with relevant data.

The 391 landraces and the 8 standard varieties were collections from Eritrea (54) and from the following regions of Ethiopia: Bale (2), Gamo Gofa (24), Gojam (10), Gonder (42), Harerge (56), Illubabor (38), Kafa (6) Shewa (38), Sidamo (25), Tigray (30), Welega (14), and Welo (60) (Fig. 3.1). The landraces were sampled from all regions of Ethiopia (except Arsi) and from Eritrea. The altitude of the collection sites for the landraces used in this study ranged from 400 to 2890 meters above sea level, while the altitude of the sorghum database in the IBCR ranged from 400 to 2940 meters above sea level. Thus the landraces reasonably represented the altitudinal distribution of sorghum germplasm collection base of the IBCR and presumably the distribution of the crop in Ethiopia. The introduced materials were one each from Burkina Faso and Somalia and two each from America, India, Kenya, South Africa, Sudan, Uganda, and Zimbabwe.

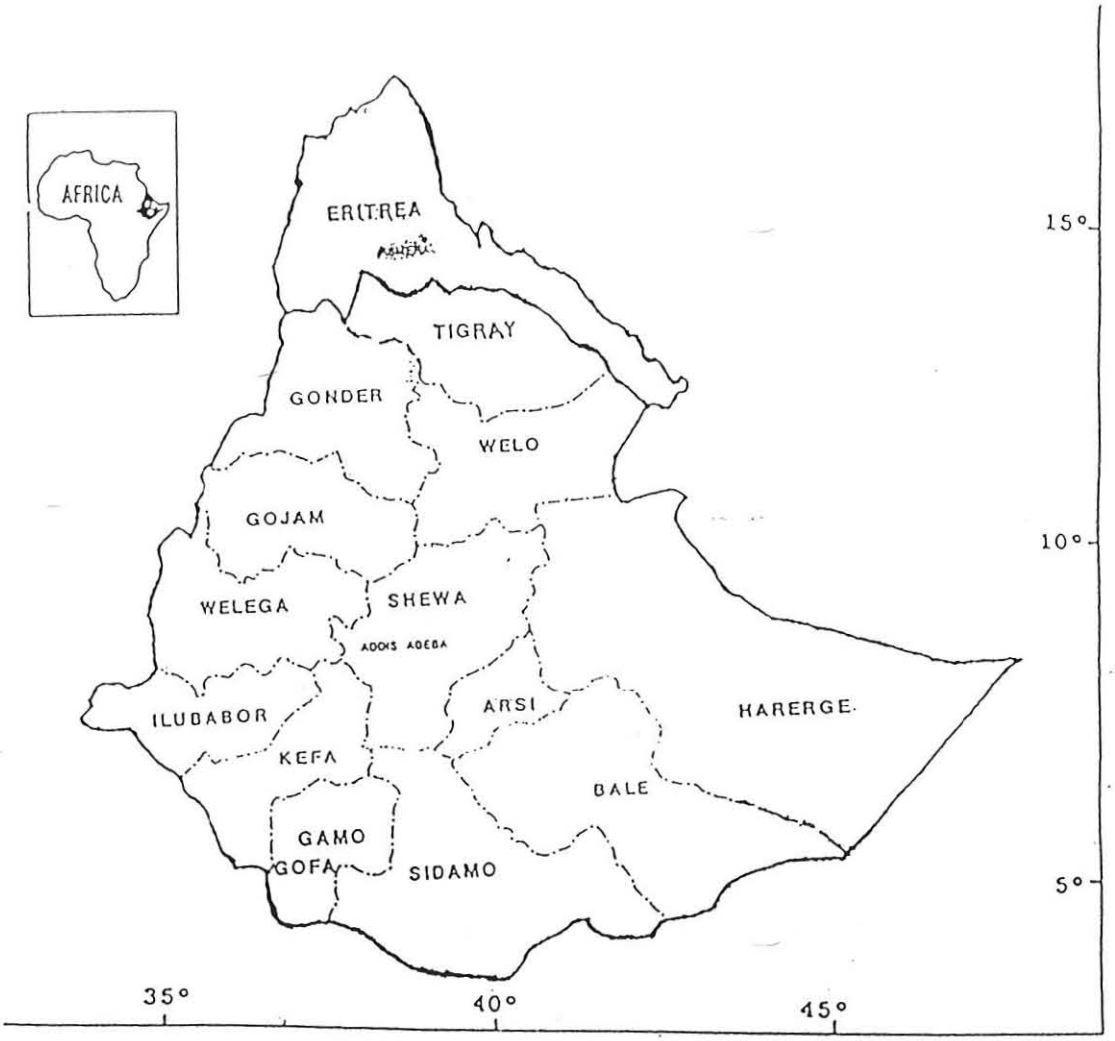


Fig. 3.1. Map of Ethiopia and Eritrea showing regions from where the sorghum germplasm are collected.

Solid lines represent international boundaries; broken lines represent boundaries between previous provinces within Ethiopia.

Methods

For data analysis, the standard varieties were assigned to the different regions of Ethiopia depending on their record of previous origin and/or the regions for which they are normally recommended for cultivation. Accessions from regions with sample size less than 12 were also included in adjacent regions to reduce experimental error due to small sample size. Hence, the two Bale accessions were included in Harerge; the ten Gojam accessions were included in Gonder and the six accessions of Kafa were placed in Illubabor. This reduced the 12 regions of Ethiopia from which the landraces were originally drawn to nine. With the Eritrea and the introduced materials included, 11 regions of origin were used in the statistical analyses.

Characterization and evaluation of germplasm collections must be made in an area of adaptation of a particular collection (Frankel, 1970). Recognizing the importance of this fact, the national sorghum improvement program of Ethiopia identified four representative characterization and evaluation sites largely based upon altitude and rainfall (Gebrekidan, 1981, 1982; Kebede and Menkir, 1987) (Table 3.1). Recently, the Natural Resources Management and Regulatory Department of the Ministry of Agriculture, Ethiopia, (MOA, 1998), classified the country into 18 agro-ecological zones and 49 subagro-ecological zones. Therefore, it is to be noted that the four agro-ecological zone classification of Gebrekidan (1981) for sorghum is quite broad. On the basis of altitude of the collection site of each accession, the 415 materials were divided into three sets. These were: (1) 196 lowland accessions (<1600 meter above sea level), (2) 109 intermediate accessions (1600-1900 masl), and (3) 110 highland accessions (>1900 masl). Since rainfall data were not available in the passport data of the accessions, the lowland dry and lowland wet were treated together as lowland accessions.

The lowland, the intermediate, and the highland accessions were grown at Nazareth, Jimma, and Arsi Negele research centers, respectively, during the 1996/97 growing season in Ethiopia. List of the three groups of accessions are given in Appendix 3.1. Details of the planting locations are shown in Table 3. 2.

Table 3.1. Adaptation zones for sorghum research in Ethiopia[a]

Adaptation zone	Altitude (masl)[b]	Rainfall (mm)	Growing period (days)	Recommended evaluation site (s)
Lowland dry	<1600	<600	90-130	Nazareth, Kobo, Mieso
Lowland wet	<1600	>1000	110-150	Gambella
Intermediate	1600-1900	>1000	150-180	Jimma, Bako
Highland	>1900	800	170-200	Arsi Negele, Alemaya

[a] After Gebrekidan (1981) and Kebede and Menkir (1987).
[b] masl = meters above sea level.

Table 3.2. Some meteorological data for the three evaluation locations[a]

Location	Latitude	Altitude (masl)[b]	Rainfall (mm)	Temperature (°C)	Growing period (days)
Nazareth/Melkassa	8°30´N	1500	<600	21.5	85-130
Jimma/Melko	7°46´N	1753	1469	18.8	150-180
Arsi Negele	7°20´N	1960	900	18.4	170-200

[a] Source: Meteorological Stations of respective locations.
[b] masl = meters above sea level.

Thirty-eight accessions, consisting of the eight standard varieties and 30 landraces (ten each from the lowland, intermediate and the highland altitudes) were also grown in all the three locations irrespective of their particular adaptation zone. Within a group, each landrace accession was chosen to represent different regions of origin. Data from these 38 accessions were used for comparison purposes and to help combining of data from the three groups of accessions planted in the three locations.

Each accession was grown in a single row plot of 3 m long, and 0.75 m between rows, with two replications in a randomized complete block design. Within every row, the spacing from one plant to another was 0.3 m. Fertilizers were applied at the rate of 100 kg ha^{-1} diammonium phosphate (DAP) and 100 kg ha^{-1} urea as recommended for sorghum production in Ethiopia. All DAP was applied at the time of planting, while urea was a split application.

For every accession, ten individuals were used for recording data for quantitative characters, except days for 50% flowering, which was recorded on plot basis (Table 3.3a). For the qualitative characters (Table 3.3b) the most frequent character state was recorded. For both the quantitative and qualitative characters, data was recorded based on Sorghum Descriptors (IBPGR/ICRISAT, 1993). For leaf characteristics, procedures developed by Ayyangar (1942) and Stickler et al. (1961) were used.

Table 3.3a. List of quantitative characters used for the study

Character	Code	Description
Leaf number (count)	LN	Count of total leaves per plant
Leaf length (cm)	LL	Length of the fourth leaf from the flag leaf
Leaf width (cm)	LW	Width of the fourth leaf from the flag leaf
Leaf area (cm^2)	LA	Area of the fourth leaf from the flag leaf
Length of leaf sheath (cm)	LLS	Length measured on the fourth internode from the surface
Internode length (cm)	IL	Length measured on the fourth internode measured from the surface
Stalk diameter (cm)	SD	Diameter measured on the fourth interonde from the surface
Plant height (cm)	PH	Height from the surface to the tip of the panicle
Peduncle exsertion (cm)	PE	Length between the base of flag leaf and the base of the panicle
Days for 50% flowering (count)	DF	Days from planting to when 50% plants in a plot flowered half way down the panicle
Panicle length (cm)	PL	Length of the panicle from its base to tip
Panicle width (cm)	PW	Width of the panicle in middle
Number of Primary branches per panicle (cm)	NPBP	Counts of primary branches per panicle
Length of primary branches per panicle (cm)	LPBP	Length of primary branches per panicle
Thousand seed weight (g)	TSW	Weight of 1000 seed counts

Table 3.3b. List of qualitative characters used for the study along with their codes and descriptions

Character	Code	Description
Leaf midrib color	1	White
	2	Dull green
	3	Yellow
	4	Brown
Waxy bloom	3	Slightly present
	5	Medium present
	7	Mostly present
	9	Completely present
Panicle compactness & shape	1	Very lax panicle
	2	Very loose erect primary branches
	3	Very loose drooping primary branches
	4	Loose erect primary branches
	5	Loose drooping primary branches
	6	Semi-loose erect primary branches
	7	Semi-loose drooping primary branches
	8	Semi-compact elliptic
	9	Compact elliptic
	10	Compact oval
	11	Half broom corn
	12	Broom corn
Awn at maturity	0	Absent
	1	Present
Glume color	1	White
	2	Yellow
	3	Brown
	4	Purple
	5	Black
	6	Gray
Grain covering	25	Glume covers 25% of the length of the grain
	50	Glume covers 50% of the length of the grain
	75	Glume covers 75% of the length of the grain
	100	Glume covers 100% of the length of the grain
Seed color	1	White
	2	Yellow
	3	Red
	4	Brown
	5	Buff
Grain plumpness	3	Dimple
	7	Plump
Grain sub-coat	0	Sub-coat color is absent
	1	Sub-coat color is present
Endosperm texture	1	Completely corneous
	3	Mostly corneous
	5	Intermediate
	7	Mostly starchy
	9	Completely starchy

Statistical analyses

Quantitative characters

Univariate analysis

Using the raw data of the 38 accessions, which were grown at the three locations, analysis of variance was made for the 15 quantitative characters for each location. Homogeneity of the error variances among the three locations was assessed by Bartlett's test for each of the 15 characters. The test established the homogeneity of the error variances ($P \leq 0.01$) for seven characters. These are LN, LL, LW, LA, LLS, SD, and PL. The data for these characters were subjected to site-accession wise analysis of variance to determine effects of test sites and genotype x environment (G x E) interaction. The results showed the non-significance of effects of test sites for LN and the non-significance of G x E interactions for LL. For the other five characters (i.e., LW, LA, LLS, IL, and SD) both test site effects and G x E interactions were significant ($p \leq 0.05$).

For the remaining eight characters, namely, IL, PH, PE, DF, PW, NPBP, LPBP, and TSW, the test indicated significant ($P \leq 0.01$) heterogeneity of the error variances across the three locations. Test site effects and G x E interactions were also significant ($p \leq 0.05$). For these characters logarithmic data transformation, which is recommendable for continuous data, also failed to homogenize the error variances. Heterogeneity of error variance and lack of appropriate data transformation system are expectable because of the wide range of phenotypic variation of the materials used and the nature of the data itself (Gomez and Gomez, 1984; Pecetti et al., 1992). For such characters like plant height and days for 50% flowering, values are expected to differ greatly from one location to another as well as from one accession to another in the same location (as happens with the present study), leading to heterogeneity of error variances that cannot be corrected by data transformation.

Bivariate analysis

Correlations between the characters were computed at three levels, following the procedures used by Thorpe (1976) and Bekele (1984). First correlations of the characters were assessed based on the 415 accession means. Then inter-region correlation was computed using the means of characters for each region. Lastly, a series of intra-region correlation coefficient matrices were obtained for each region using the accession means from that region for the characters.

Multivariate analysis

For multivariate analysis, the data were standardized to a mean of zero and a variance of unity to avoid differences in scales used for recording data on the different characters (Sneath and Sokal, 1973). Then a series of multivariate analyses were made using the appropriate procedures of SAS computer program (SAS Institute, 1988). Principal component analysis was performed using correlation matrix to define the patterns of variation both between accessions and between their regions of origin.

The 415 accessions were clustered using the PROC FASCLUS of SAS (SAS Institute, 1988), which grouped and sorted the accessions into clusters. The measure of dissimilarity was Euclidean distance and the clustering method was Unweighted Pair-Group Method Using Centroids (UPGMC) (Sneath and Sokal, 1973). This procedure is preferable for clustering a large number of accessions (SAS Institute, 1988). The number of clusters was determined in such a way that a cluster should contain at least two accessions to allow subsequent statistical analysis. Grouping into ten clusters met the criterion. The relationships among the clusters were assessed by measuring the inter-cluster distances using Mahalanobis distance (D^2).

For the regions of origin, hierarchical cluster analysis was run using the regional means for the 15 quantitative characters. The measure of dissimilarity was Euclidean distance and the clustering method was Unweighted Pair-Group Method Using Arithmetic Average (UPGMA) (Sneath and Sokal, 1973). Discriminant analysis was used to assess if the accession membership conformed to their regions of origin and adaptation zones.

Qualitative characters

Phenotypic frequency distributions of the characters were worked out for all the materials, the regions of origin and the adaptation zones. The significance of proportions across regions and adaptations zones for a particular character state was tested with chi-square using the formula given by Fleiss (1973, p. 93). The Shannon-Weaver diversity index (H′) was computed using the phenotypic frequencies to assess the phenotypic diversity for each character for all the entire sample, and the samples were grouped for each region of origin and for the adaptation zones. The Shannon-Weaver diversity index as described by Hutchenson (1970) is given as:

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

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 for a character. Each value of H′ was divided by its maximum value, $\log_e n$, and normalized in order to keep the values between zero and one. The partitioning of the phenotypic diversity into within and between regions of origin and within and between adaptation zones was made following the methods given by Wachira et al. (1995). Contingency chi-square test was made using the entire data set to assess whether the distribution of the characters studied was random or non-random.

Results

Quantitative characters

Univariate analysis

Analyses of variances revealed highly significant differences ($P \leq 0.01$) between the 11 regions of origin of the 415 sorghum accessions and between accessions pooled over the regions for the 15 characters studied (Table 3.4). Region-wise partitioning of the variance indicated significant within-region differences ($P \leq 0.05$) among the populations within Gamo Gofa, Gonder, Harerge, Illubabor, and Welo for all the 15 characters; for 14 characters within Tigray; for 13 characters within Eritrea, Shewa, and Welega; and for 12 characters for Sidamo (Table 3.4). For accessions within the introduced materials differences in only seven of the 15 characters were found to be significant. In general, within region variation was greater for plant height, peduncle exsertion, number of days for 50% flowering, thousand seed weight and for panicle characteristics than for leaf characteristics for all the regions.

The Duncan's multiple range testing for regional means for all the characters is shown in (Table 3.5). Apparently, much more regional differentiation was observed for mean days for 50% flowering and for mean plant height. The mean for number of days for 50% flowering for the accessions from Sidamo were significantly higher than those for the other regions ($P \leq 0.05$). The highest mean number of leaves per plant was also observed for the Sidamo accessions. The introduced elite breeding materials and the Eritrea accessions attained 50% flowering significantly earlier than those of other regions. The mean for plant height for the accessions from Welega was significantly higher than those for all the other regions.

Table 3.4. Analysis of variance for 15 quantitative characters in sorghum[a]

Source	df	LN	LL	LW	LA	LLS	IL	S
Regions	10	243.87**	4863.14**	44.60**	639016.55**	458.46**	960.53**	
Pooled accessions within regions	404	9.41**	185.41**	2.63	26810.74**	23.89**	82.49**	
Accessions within:								
Introduction	15	2.32	201.77**	2.26	18982.49	5.70	22.23	
Eritrea	53	3.55**	128.74**	2.46*	5676.40	9.75	37.80**	
Gamo Gofa	23	12.34**	239.75**	2.69*	31254.95**	36.73**	125.87**	
Gonder	51	9.72**	246.22**	2.67**	31497.81**	33.14**	99.74**	
Harerge	57	12.51**	289.67**	3.49**	42793.13**	26.31**	86.30**	
Illubabor	43	16.50**	237.12**	3.05**	36954.92**	23.21**	105.28**	
Shewa	37	9.66**	79.04	2.30	25235.68**	18.45**	52.23**	
Sidamo	24	8.69**	74.07	1.65	12691.77	20.27**	61.99**	
Tigray	29	7.28**	235.34**	2.50	37545.62**	32.54**	111.10**	
Welega	13	3.26	154.05**	2.30	28662.57**	22.85**	67.87**	
Welo	59	9.44**	114.39**	2.46*	20214.77**	27.25**	102.30**	
Pooled error within region	404	2.14	72.83	1.64	12318.89	7.52	17.18	

[a] Refer to Table 3.3a for the abbreviations of the characters.
* and ** significant at p = 5% and 1* levels, respectively.

Table 3.4 Continued[a]

Source	df	PH	PE	DF	PL	PW	NPBP	LPBP	TSW
Regions	10	290485.47**	317.20**	61165.73**	1885.29**	2390.00**	7463.00**	1343.90**	30.25**
Pooled accessions within regions	404	17981.45**	53.84*	1047.02**	74.07**	85.54**	464.79**	54.94**	4.36**
Accessions within:									
Introduction	15	2942.79**	41.23**	133.60**	46.37**	5.28	276.97**	4.40	0.43*
Eritrea	53	3540.12**	74.39**	468.17**	99.95**	23.01**	971.69**	114.68**	0.39**
Gamo Gofa	23	30752.13**	109.29**	2517.54**	52.19**	167.79**	302.72**	89.30**	6.16**
Gonder	51	15837.24**	62.38**	676.90**	88.19**	101.60**	329.16**	44.24**	7.12**
Harerge	57	16900.42**	27.74**	1202.58**	84.49**	45.76**	609.88**	21.90**	5.36**
Illubabor	43	55332.69**	30.98**	3337.75**	59.74**	44.12**	154.07**	82.67**	6.08**
Shewa	37	24978.84**	27.50**	522.16**	37.80**	171.16**	396.55**	56.32**	6.01**
Sidamo	24	17069.76**	85.13**	1713.14**	49.01**	142.55**	491.92**	32.46**	4.64**
Tigray	29	10507.05**	74.38**	88.86**	78.63**	77.06**	536.56**	84.87**	2.97**
Welega	13	14415.72**	88.86**	776.46**	57.79**	232.69**	348.75**	47.07**	8.60**
Welo	59	5917.14**	37.48**	314.89**	88.89**	79.68**	345.92**	17.02**	2.23**
Pooled error within region	404	1068.53	15.92	29.09	19.14	8.24	87.41	4.82	0.24

[a] Refer to Table 3.3a for the abbreviations of the characters.
* and ** significant at p = 5% and 1* levels, respectively.

Table 3.5. Regional means for 15 quantitative characters in sorghum[a]

Regions	LN	LL	LW	LA	LLS	IL	SD
Introduction	9.22^{d}	76.59^{ab}	8.50	431.55^{ab}	13.66^{c}	13.11^{c}	2.21$^{a\,b}$
Eritrea	9.54^{cd}	61.92^{b}	8.09	377.86^{b}	13.96^{c}	18.82$^{a\,b\,c}$	2.05$^{a\,b}$
Gamo Gofa	13.33^{ab}	78.66^{ab}	9.01	551.34^{ab}	18.96^{abc}	23.18$^{a\,b}$	2.34$^{a\,b}$
Gonder	12.60^{bc}	84.04^{a}	8.87	592.18$^{a\,b}$	19.03^{abc}	23.46$^{a\,b}$	2.19$^{a\,b}$
Harerge	13.96^{ab}	75.92^{ab}	9.99	589.86$^{a\,b}$	14.64^{bc}	16.40$^{b\,c}$	2.65^{a}
Illubabor	12.04^{bcd}	80.94^{ab}	10.34	654.63$^{a\,b}$	17.85^{abc}	20.61$^{a\,b\,c}$	2.37$^{a\,b}$
Shewa	13.28^{ab}	85.72^{a}	9.92	664.54$^{a\,b}$	20.46^{ab}	25.66$^{a\,b}$	2.17$^{a\,b}$
Sidamo	15.85^{a}	86.32^{a}	9.38	624.17$^{a\,b}$	18.92^{abc}	21.04$^{a\,b\,c}$	2.41$^{a\,b}$
Tigray	11.19^{bcd}	79.16^{ab}	8.59	533.97$^{a\,b}$	18.42^{abc}	21.98$^{a\,b\,c}$	1.82^{b}
Welega	13.57^{ab}	88.41^{a}	9.22	646.18$^{a\,b}$	21.70^{a}	27.28^{a}	2.12$^{a\,b}$
Welo	13.45^{ab}	74.76^{ab}	9.84	558.48$^{a\,b}$	16.09^{abc}	17.76$^{a\,b\,c}$	2.65^{a}
Entire data	12.56	77.50	9.33	563.33	17.17	20.41	2.32

Table 3.5. Continued[a]

Regions	PH	PE	DF	PL	PW	NPBP	LPBP	TSW
Introduction	144.68^{f}	7.21	71.68^{h}	23.83$^{a\,b}$	4.79^{e}	56.11$^{a\,b\,c}$	7.78^{e}	1.51^{d}
Eritrea	171.91^{ef}	3.89	68.83^{h}	14.81^{c}	3.86^{e}	75.82^{a}	5.64^{e}	2.06^{bcd}
Gamo Gofa	314.27^{bc}	7.64	126.94^{cd}	25.04$^{a\,b}$	15.07^{bc}	47.42^{c}	16.08$^{a\,b\,c}$	3.03$^{a\,b}$
Gonder	296.55^{bcd}	7.58	119.83^{de}	29.24$^{a\,b}$	11.31^{c}	60.36$^{a\,b\,c}$	13.20^{bcd}	3.62^{a}
Harerge	251.58^{cd}	2.14	122.59^{cde}	19.61^{c}	7.78^{de}	75.52^{a}	7.90^{e}	3.04$^{a\,b}$
Illubabor	302.15^{bc}	4.53	108.27^{g}	25.50$^{a\,b}$	9.40^{cde}	53.53$^{b\,c}$	12.97^{bcd}	3.13$^{a\,b}$
Shewa	317.69^{bc}	4.20	140.18^{b}	27.36$^{a\,b}$	21.46^{a}	63.56$^{a\,b\,c}$	16.64^{abc}	3.02$^{a\,b}$
Sidamo	335.81^{b}	2.61	181.19^{a}	23.35$^{a\,b}$	19.79^{a}	54.65$^{a\,b\,c}$	17.86$^{a\,b}$	2.50^{abcd}
Tigray	219.58^{de}	7.90	108.71^{fg}	28.13$^{a\,b}$	11.20^{cd}	65.80$^{a\,b\,c}$	13.05^{bcd}	3.07$^{a\,b}$
Welega	405.35^{a}	5.34	148.04^{b}	32.71^{a}	18.19^{ab}	58.45$^{a\,b\,c}$	19.02^{a}	2.85^{abc}
Welo	256.79^{d}	4.27	114.12^{efg}	25.78$^{a\,b}$	7.64^{de}	74.67$^{a\,b}$	9.38^{de}	1.92^{cd}
Entire data	266.47	4.87	115.77	24.18	10.75	65.20	11.64	2.74

[a] Refer to Table 3.3a for the abbreviations of the characters.
Means of each character followed by the same letter were not significantly different at $p \leq 0.05$ according to Duncan's multiple range test.

The accessions from Eritrea and Harerge were noted for having significantly shorter mean for panicle length than those of the other regions. The mean number of primary branches per panicle, a trait directly related to yield, was high for accessions from Eritrea, Harerge and Welo, though statistically non-significant compared with those of most of the rest of the regions. The accessions from Eritrea and the introduced materials were not significantly different from each other for all the characters studied, except for panicle length.

The range of variation of the accession means demonstrated wide variation between the regions and the accessions within the regions for the characters studied (Table 3.6). Peduncle exsertion showed complex range of variation between regions, ranging from 0 to 33 cm. For the other characters, the maximum score was 54 times the minimum for thousand seed weight, 42 for panicle width, 24 times for length of primary branches per panicle, 9 times for number of primary branches per panicle and for inter-node length, 8 times for plant height, 6 for panicle length and for leaf area, 4 times for days for 50% flowering, length of leaf sheath and stalk diameter and, 3 times for leaf number, length and width. The same trend was observed between accessions within a particular region. However, accessions within Gamo Gofa, Gonder, Harerge and Illubabor showed wider ranges of variation than those found in accessions within other regions for the majority of the characters.

Computed as a ratio of standard deviation of each character to the corresponding entire data mean and expressed as percentage, coefficient of variation is useful to compare different characters measured in different units. It is also useful to compare the same character in different groups of populations with different sample size, mean and variance, or different characters in different populations. In the present study, high coefficients of variation were observed between regions and within each region for peduncle exsertion, leaf area, panicle width, length of primary branches per panicle, plant height, hundred seed weight, days for 50% flowering and number of primary branches per panicle (Table 3.7). Though accessions from a particular region were more variable for a specific character compared to other regions, accessions from Gamo Gofa, Gonder, Harerge and Illubabor were more variable than accessions from other regions. The accessions from Eritrea and the introduced materials had low coefficients of variation for many characters, indicating relatively high within-region uniformity.

Table 3.6. Regional ranges of variation for 15 quantitative characters in sorghum[a]

Regions	LN	LL	LW	LA	LLS	IL	SD
Introduction	6.91-11.91	47.22-80.22	6.56-10.06	274.20-565.90	11.54-17.04	8.64-22.14	1.85-2.68
Eritrea	6.91-11.91	42.22-88.99	5.06-11.06	155.70-613.90	11.04-22.24	110.14-32.44	1.19-2.50
Gamo Gofa	8.41-17.13	45.22-91.29	6.52-11.52	321.10-771.60	8.04-24.72	6.14-34.22	1.23-3.58
Gonder	7.63-18.41	59.22-108.99	5.02-11.06	289.70-851.10	7.04-27.22	5.14-34.92	1.25-3.04
Harerge	6.96-19.13	51.22-98.49	6.96-13.46	302.00-998.74	6.54-21.22	4.64-28.92	1.13-4.04
Illubabor	7.91-17.63	53.72-101.49	7.06-12.56	307.50-898.70	8.54-24.74	6.64-34.92	1.51-3.01
Shewa	8.46-20.13	73.72-99.49	7.46-12.02	450.50-884.80	8.04-26.74	6.14-32.94	1.25-3.60
Sidamo	10.91-20.63	71.79-97.99	7.46-11.46	458.70-747.90	10.22-24.22	9.94-32.42	1.56-3.68
Tigray	7.41-16.13	53.79-105.49	6.52-11.52	272.00-969.60	8.04-29.74	6.14-39.92	1.16-2.89
Welega	10.63-15.46	70.22-101.99	6.96-10.52	442.60-836.80	15.04-27.24	13.14-34.92	1.54-2.78
Welo	9.46-18.41	55.22-92.49	7.06-13.46	333.80-826.90	8.04-22.72	6.64-32.92	1.57-3.86
Entire data	6.91-20.63	42.22-108.99	5.02-13.46	155.70-998.74	6.54-29.74	4.64-39.92	1.13-4.04

[a] Refer to Table 3.3a for the abbreviations of the characters.

Table 3.6. Continued[a]

Regions	PH	PE	DF	PL	PW	NPBP	LPBP	TSW
Introduction	80.49-227.49	1.24-15.24	58.08-89.58	11.21-34.21	2.88-9.38	43.99-86.99	3.28-9.28	0.78-2.20
Eritrea	93.83-357.18	0.00-23.26	49.08-136.50	9.21-40.71	1.38-24.73	31.49-127.99	2.28-50.10	1.13-3.72
Gamo Gofa	150.50-558.70	0.00-23.50	60.58-188.50	16.48-35.81	5.38-42.23	13.99-68.06	5.12-26.60	1.04-6.52
Gonder	155.30-453.70	0.00-26.00	77.08-168.92	17.31-51.31	2.88-40.73	31.43-91.56	6.60-29.60	0.28-6.50
Harerge	83.30-484.20	0.00-17.24	69.58-175.42	7.98-51.21	1.00-20.23	41.43-116.49	2.12-21.60	1.08-6.315
Illubabor	141.50-611.70	0.00-14.74	65.58-170.92	18.21-39.31	2.88-18.89	30.43-80.56	5.78-25.10	0.63-7.02
Shewa	160.30-547.70	0.00-12.00	100.42-191.92	17.48-37.21	7.88-42.23	27.43-103.56	7.12-27.10	1.00-6.45
Sidamo	197.30-541.70	0.00-23.50	106.08-214.00	13.98-31.81	2.88-32.23	23.49-99.56	10.62-24.62	0.94-5.98
Tigray	110.80-394.70	0.00-33.50	60.58-133.50	15.21-45.81	2.88-27.73	37.43-98.06	5.28-34.10	1.18-5.58
Welega	295.70-561.20	0.00-21.00	98.42-171.42	24.81-46.81	3.88-33.73	38.43-74.56	12.60-27.60	0.13-5.88
Welo	169.83-454.68	0.00-15.74	80.58-136.42	13.21-44.71	1.38-23.73	50.43-104.56	3.28-16.78	0.48-5.94
Entitre data	80.49-611.70	0.00-33.50	49.08-214.00	7.98-51.21	1.00-42.23	13.99-127.99	2.12-50.10	0.13-7.03

[a] Refer to Table 3.3a for the abbreviations of the characters.

Table 3.7. Regional percent coefficient of variation for 15 quantitative characters in sorghum[a]

Region	LN	LL	LW	LA	LLS	IL	SD	PH	PE	DF	PL	PW	NPBP	LPBP	TSW
Introduction	2.39	10.36	5.97	176.72	5.55	7.79	6.10	11.49	73.05	6.24	15.26	11.32	14.93	3.94	11.25
Eritrea	6.69	6.82	6.86	88.52	6.15	15.73	6.82	13.19	111.03	12.80	26.29	25.28	32.25	63.67	9.99
Gam Gofa	17.98	11.79	7.77	229.79	22.26	36.12	26.40	45.72	140.30	30.47	16.81	83.09	15.91	55.84	62.79
Gonder	15.50	12.01	7.69	230.35	20.5	31.48	14.93	32.25	98.97	15.55	24.30	63.56	16.86	38.14	67.69
Harerge	18.30	13.44	10.31	270.35	17.85	28.80	22.81	33.39	49.92	20.92	23.64	40.29	24.79	25.11	58.39
Illubabor	21.33	11.69	9.00	250.35	16.31	32.52	9.14	61.81	56.35	35.13	18.63	39.40	8.85	53.60	62.36
Shewa	15.44	2.27	6.16	205.48	13.62	20.51	23.61	41.03	49.41	13.56	12.63	83.96	19.07	43.59	61.99
Sidamo	14.41	1.02	0.76	142.02	14.71	23.19	21.77	33.57	120.79	25.06	15.98	76.23	21.81	31.94	54.13
Tigray	12.76	11.63	7.03	252.71	20.60	33.58	15.24	25.78	111.02	4.72	22.56	54.57	22.98	54.35	42.64
Welega	5.96	8.22	6.16	219.65	16.12	24.67	15.54	30.66	124.00	16.70	18.18	98.55	17.53	39.49	74.62
Welo	15.21	5.88	6.86	182.74	18.29	31.96	24.00	18.48	67.42	10.33	24.42	55.60	17.44	21.22	36.40
Entire data	20.18	13.71	10.70	21.91	21.33	32.04	21.32	40.67	96.19	30.59	29.02	75.82	25.35	54.75	56.20

[a] See the materials and methods for the abbreviations of the characters.

Bivariate analysis

Phenotypic correlation coefficients for 15 quantitative characters were computed for the entire data (Table 3.8). The majority of the correlation coefficients were positive and highly significant. Panicle length was positively and significantly correlated with panicle width and length of primary branches per panicle but negatively correlated with number of primary branches per panicle. Twelve of the 15 characters also showed significant positive correlations with altitude of the collection sites.

The trend for associations of characters for between regions (Table 3.9) and within regions (Appendix 3.2) was similar to that of the entire data. For instance, plant height was significantly and positively correlated with days for 50% flowering between regions and within seven of the eleven regions. Leaf area correlated significantly and positively with leaf length and width, which seemed to be due to functional relationship (leaf area = leaf length x leaf width x 0.75). Leaf number correlated significantly and positively with plant height and days for 50% flowering between regions and within most of the regions. Stalk diameter and length of leaf sheath were positively and significantly correlated with plant height within most of the regions. However, these characters were negatively and significantly correlated with each other in five of the eleven regions. Another similarity of trend was that while most correlation coefficients were positive, those involving peduncle exsertion and number of primary branches were negative.

Table 3.8. Simple correlation coefficients among 15 quantitative characters and between the characters and altitude of collection sites (based on the mean for 415 sorghum accessions[a])

	LN	LL	LW	LA	LLS	IL	SD	PH	PE	DF	PL	PW	NPBP	LPBP	TSW
LL	0.36**														
LW	0.39**	0.35**													
LA	0.45**	0.86**	0.75**												
LLS	0.06**	0.56**	0.00	0.40**											
IL	0.00	0.52**	-0.05	0.39**	0.84**										
SD	0.59**	0.04	0.44**	0.28**	-0.31**	-0.26**									
PH	0.64**	0.60**	0.23**	0.59**	0.48**	0.57**	0.37**								
PE	-0.30**	0.07	-0.29**	-0.08	0.19**	0.25**	-0.22**	0.06							
DF	0.67**	0.62**	0.24**	0.54**	0.49**	0.38**	0.13**	0.65**	-0.16**						
PL	0.19**	0.40**	0.08	0.31**	0.42**	0.29**	0.06	0.43**	0.33**	0.29**					
PW	0.16**	0.46**	0.05	0.31**	0.62**	0.49**	-0.37**	0.28**	-0.04	0.61**	0.26**				
NPBP	0.09	-0.25**	0.11*	-0.15**	-0.24**	-0.24**	0.08	-0.24**	-0.31**	-0.09	-0.30**	-0.13*			
LPBP	0.30**	0.60**	0.11*	0.49**	0.61**	0.58**	0.01	0.67**	0.21**	0.57**	0.62**	0.55**	-0.45**		
TSW	0.28**	0.53**	0.10*	0.52**	0.36**	0.54**	0.18**	0.66**	0.17**	0.36**	0.20**	0.10*	-0.28**	0.42**	
Altitude	0.18**	0.33**	-0.08	0.17**	0.45**	0.44**	-0.22**	0.24**	-0.03	0.58**	0.15**	0.61**	0.12*	0.35**	0.16**

[a] See the materials and methods section for the abbreviations of the characters.
* and ** significant at p = 5% and 1% levels, respectively.

Table 3.9. Inter-region simple correlation coefficients among 15 quantitative characters (based on the mean of the 11 regions of origin)[a]

	LN	LL	LW	LA	LLS	IL	SD	PH	PE	DF	PL	PW	NPBP	LPBP
LL	0.62*													
LW	0.62**	0.42												
LA	0.77**	0.84**	0.79**											
LLS	0.56	0.83**	0.26	0.76**										
IL	0.42	0.61*	0.12	0.61*	0.94**									
SD	0.53	-0.01	0.67*	0.27	-0.26	-0.38								
PH	0.78**	0.78**	0.50	0.86**	0.88**	0.80**	0.15							
PE	-0.45	0.11	-0.53	-0.24	0.19	0.17	-0.57	-0.15						
DF	0.94**	0.78**	0.47	0.80**	0.74**	0.59	0.25	0.85**	-0.32					
PL	0.29	0.82**	0.21	0.63*	0.79**	0.64*	-0.23	0.62*	0.48	0.44				
PW	0.70*	0.81**	0.32	0.74*	0.88**	0.79**	-0.10	0.82**	-0.08	0.87**	0.56			
NPBP	-0.16	-0.58	-0.02	-0.32	-0.49	-0.33	0.13	-0.42	-0.45	-0.33	-0.45	-0.47		
LPBP	0.64*	0.87**	0.27	0.75**	0.95**	0.83**	-0.16	0.88**	0.12	0.83**	0.71**	0.94**	-0.64*	
TSW	0.40	0.51	0.31	0.64*	0.63*	0.66*	-0.13	0.57	0.16	0.43	0.43	0.45	-0.23	0.50

[a] See the materials and methods section for the abbreviations of the characters.
* and ** significant at p = 5% and 1% levels, respectively.

The distribution of the 11 regions of origin of the accessions along the first two axes of the principal components is shown in Figure 3.2. The regions of origin of the accessions were dispersed in all four quadrants. The first principal component was much more important than the second one in separating the regions of origin of the accessions. On one hand, the extremes of the first axis were occupied, by Eritrea and the introduced materials with low negative principal scores and, on the other hand, by Welega, Sidamo and Shewa with high positive scores (see Table 3.11). Apparently the first principal component differentiated accessions that were short in stature and early flowering from those that were tall and late flowering. On the second axis, Tigray and Welega had relatively high principal positive scores, while Harerge and Sidamo had low negative scores. This axis differentiated accessions mainly based on peduncle exsertion and stalk diameter. Illubabor was the closest region to the center, while Eritrea was the most far away from the center, suggesting that accessions from Eritrea were relatively more divergent from those of the other regions.

The 415 sorghum accessions were grouped into 10 clusters (Table 3.12). The number of accessions per cluster varied from nine accessions in cluster X to 67 accessions in cluster VI. Clusters I, II, III, V and VIII contained mostly lowland accessions from Eritrea, Gamo Gofa, Gonder, Harerge, Illubabor, Tigray, Welo and the introduced materials. Cluster IV encompassed mostly highland accessions from Gamo Gofa, Gonder, Shewa, Sidamo and Welega. Clusters VI and VII were rather mixed encompassing nearly equal proportions of lowland, intermediate and highland accessions, which were mostly drawn from Welo, Gonder, Illubabor, Harerge and Tigray and with some accessions from Gamo Gofa, Shewa and Sidamo. Clusters IX and X contained mainly intermediate elevation accessions, which were mostly from Illubabor, Gonder, Shewa, Gamo Gofa, Sidamo and Welega.

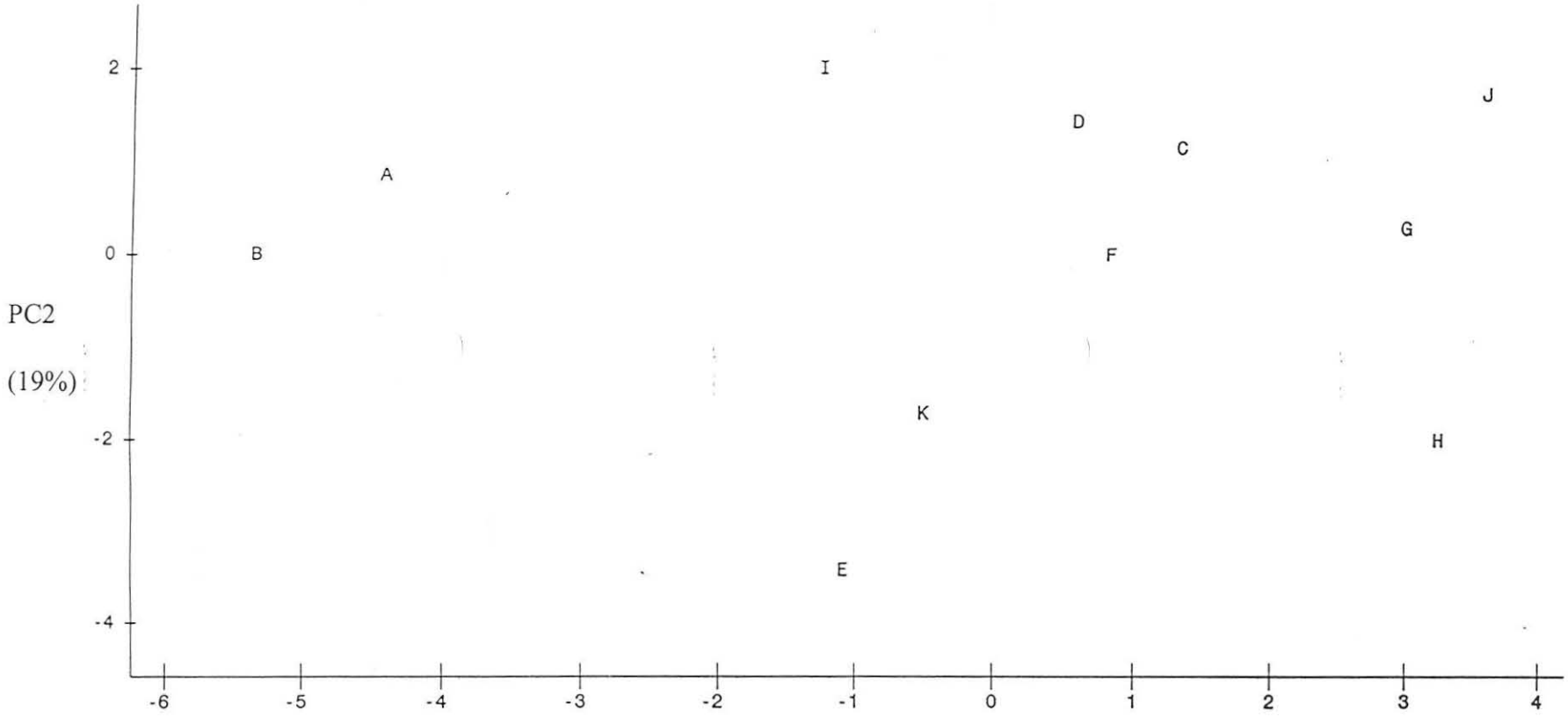


Fig. 3.2. Plot of principal component based regional means 15 quantitative characters in sorghum.

A = Introduction, B = Eritrea, C = Gamo Gofa, D = Gonder, E = Harerge, F = Illubabor, G = Shewa, H = Sidamo, I = Tigray, J= Welega, and K = Welo.

Table 3.12. Distributions of 415 sorghum accessions over 10 clusters by region of origin and adaptation zones[a]

Regions	Zones	I	II	III	IV	V	VI	VII	VIII	IX	X	Total For Zone	Total For Region
Introduction	1	1	6	6		3						16	
	2											0	
	3											0	16
Eritrea	1	33	17									50	
	2										2	2	
	3			1							1	2	54
Gamo Gofa	1		2	3	1	2	1					9	
	2						2			4	1	7	
	3				3	1	2		1	1		8	24
Gonder	1		1	7	1	1	1	2				13	
	2		1	1		1	11	3		8	1	26	
	3			1	7		4				1	13	52
Harerge	1	3	2	5		2	1	1	16			30	
	2			2			1	7	7	1		18	
	3	1						7	2			10	58
Illubabor	1		3	5	1		12	6				27	
	2				1					15		16	
	3			1								1	44
Shewa	1				1							1	
	2						2	2		9		13	
	3	1			18	2		3				24	38

[a] 1 = lowland (< 1600 m above sea level), 2 = intermediate elevation (1600-1900 masl), 3 = highland (> 1900 masl).

Table 3.12 Continued[a].

Regions	Zones	I	II	III	IV	V	VI	VII	VIII	IX	X	Total For Zone	Total For Region
Sidamo	1						1					1	
	2			1						3	1	5	
	3			1	14			2	2			19	25
Tigray	1	2	1			1	2		1			7	
	2			1			8	1			1	11	
	3	1	1	4		3	3					12	30
Welega	1				2							2	
	2				1	1	1			4		7	
	3				5							5	14
Welo	1			7	1		5	7	19	1	1	40	
	2							3				3	
	3	1		2	1		10	3				17	60
Total		43	34	48	57	17	67	47	48	45	9	415	415

[a] 1 = lowland (< 1600 m above sea level), 2 = intermediate elevation (1600-1900 masl), 3 = highland (> 1900 masl).

On the whole, lowland accessions from Gamo Gofa, Gonder and Illubabor showed close similarity in their clustering patterns, while the highland accessions from Shewa, Sidamo and Welega clustered together. On the other hand, the accessions from Gonder and Tigray grouped together in most cases and it was also noticed that lowland, intermediate and highland accessions from these two regions appeared in the same cluster. Similarly, the accessions from Harerge and Welo appeared in the same cluster in most cases. From the clustering pattern, clear-cut differentiation of the accessions into lowland, intermediate and highland was apparent for Eritrea, Gamo Gofa, Illubabor, Shewa, Sidamo, and Welega. However, such differentiation was not apparent for accessions from Gonder, Harerge, Tigray and Welo.

Compared with accessions of the other regions, most of the accessions from Eritrea, Welega and the introduced materials were grouped into a few clusters, suggesting that accessions from these regions were relatively less variable than those from the other regions. For most of the remaining regions, the accessions were distributed patchily over many clusters, reflecting wide variation among accessions within a particular region. The overlapping of the clustering patterns of accessions of many of the regions was an indication of lack of strong regional differentiation, which could be partly ascribed to gene flow.

It was noteworthy that the clustering pattern of the standard varieties was consistent with what was expected. For instance, the two early maturing and dwarf standard varieties, Seredo and 76T1#23, were placed in cluster II, in which lowland accessions (mostly from Eritrea, the introduced materials, Gamo Gofa, Gonder, Illubabor, Tigray and Welo) were most frequent. Similarly, the two standard varieties for the highland areas, Alemya-70 and ETS-2752, were included in cluster VII, which also included mainly highland landrace accessions.

Based on the Mahalanobis distance (D^2), all the clusters were different from each other (Table 3.13). The lowest inter-cluster distance (9.03 units) was between cluster III and cluster V, both of which were composed of mainly lowland accessions from Welo, Gonder, Harerge, Tigray, Illubabor and the introduced materials. The largest inter-cluster distance (75.70 units) was between cluster II (accommodating lowland accessions mostly from Eritrea, Welo and the introduced materials) and cluster IX (containing intermediate elevation accessions mostly from Illubabor, Shewa, Gamo Gofa, Welega, and Sidamo). This result suggested that rainfall and growing season are the most important environmental factors in differentiating the sorghum accessions since the latter group of regions are normally characterized by greater annual rainfall and a longer growing period than the former ones.

Cluster analysis, based on means for regions of origin for 15 quantitative characters, was used to obtain a dendrogram of the regions of origin (Fig. 3.3). The dendrogram clearly showed the close relationship between accessions from Gonder and Tigray, Shewa and Welega, Harerge and Welo, Gamo Gofa and Illubabor, and Eritrea and the introduced materials. The cluster analysis thus substantially confirmed the distinctiveness of the accessions from Eritrea from accessions of the other regions as a whole, which was already shown by the bi-plot of the principal components (see Fig. 3.2).

Discriminant analysis, using the regions of origin of the accessions as a grouping variable, revealed that 265 of the 415 (64%) accessions were correctly classified to their respective regions of origin (Table 3.14). The percentage of accessions correctly classified was relatively high for the introduced materials and accessions from Eritrea and Welega, while it was relatively low for accessions from Gonder, Gamo Gofa and Illubabor. These results agreed with the hypothesis made by Holcomb et al. (1977) and Pecetti and Damania (1996) that the higher the diversity of the group, the higher is the probability of

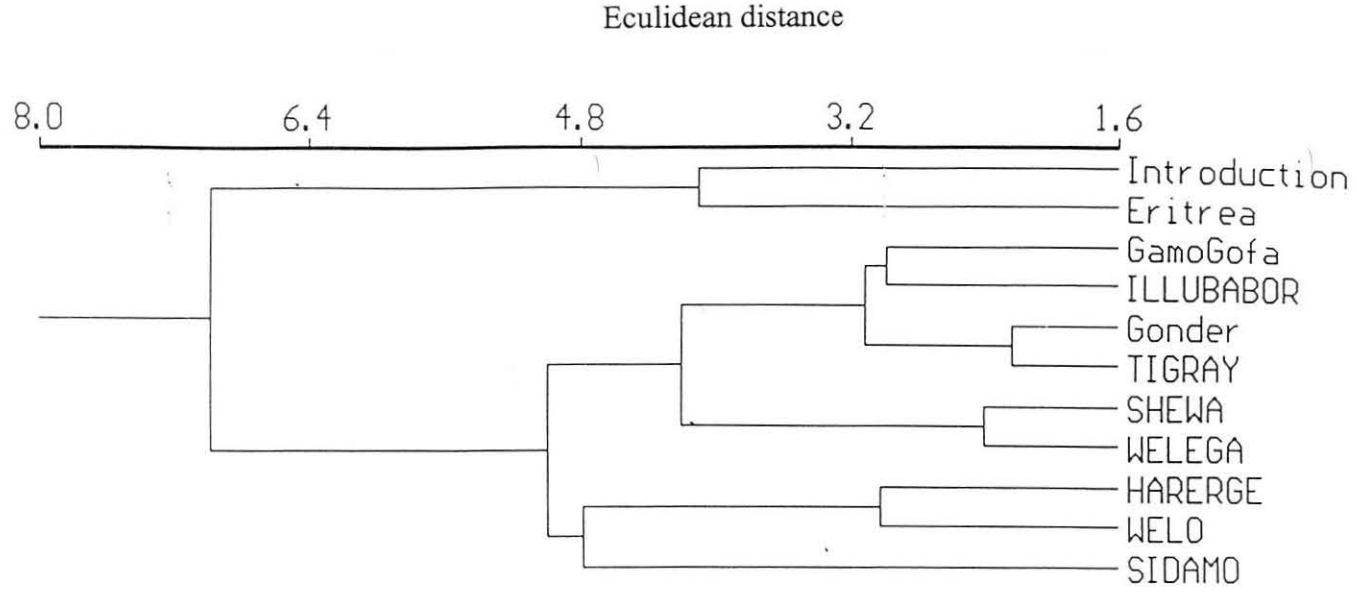


Fig. 3.3. Dendrogram showing the clustering patterns of eleven regions of origin of sorghum.

Table 3.14. Summary of discriminant analysis for 415 sorghum accessions by regions of origin

Put into region	No	True region										
		1	2	3	4	5	6	7	8	9	10	11
1 Introduction	16	15	1									
2 Eritrea	54	4	46		1					3		
3 Gamo Gofa	24	2		12	2	1	1	3	2			1
4 Gonder	52	1	1	1	24	2	4	5		7	1	6
5 Harerge	58	4	1	1	3	40	2			2		5
6 Illubabor	44	5		3		1	23	2			9	1
7 Shewa	38		2	2	2	2	1	23	2		1	3
8 Sidamo	25			3		1		4	16			1
9 Tigray	30	1		2	3	1	2	1		20		
10 Welega	14			1				1			12	
11 Welo	60	1			6	11		4		3		35
% correct		94	85	50	46	69	52	60	64	67	86	58

Table 3.15. Summary of discriminant analysis for 415 sorghum accessions by adaptation zones[a]

Put into zone	No	True zone		
		1	2	3
1 Lowland	196	188	3	5
2 Intermediate	109	13	93	3
3 Highland	110	6	6	98
% correct		96	85	89

[a] 1 = lowland (< 1600 m above sea level), 2 = intermediate elevation (1600-1900 masl), 3 = highland (> 1900 masl).

Misclassification and vice versa. Close examination of the predicted membership for the misclassified accessions of each region revealed that most of such accessions were either placed in the adjacent regions or in regions having similar climatic conditions. Discriminant analysis, using the three zones of adaptation of sorghum in Ethiopia as a classifying variable, revealed that 379 of the 415 (91%) accessions were correctly placed in their respective zones of adaptation (Table 3.15).

Qualitative characters

Regional distribution of characters

The regional frequency distributions for the ten characters are presented in Table 3.16. The most remarkable result was achieved with regard to the distributions of different panicle compactness and shapes. Out of the 415 accessions, 101 (24%) had the compact oval panicle with goose peduncle or crooked head. They comprised 42, 40 and 16 accessions from Harerge, Eritrea and Welo, respectively, and one each from Gonder, Tigray and the introduced materials. The distributions of the other panicle types also showed interesting patterns. The loose and semi-loose panicle types with erect primary branches were more frequent among the accessions from Gonder, Tigray and Welo, while the loose and very loose panicle types with dropping primary branches were more frequent among the accessions from Shewa, Sidamo and Welega. Both the semi-compact and loose to very loose panicle types were found more or less equally among the accessions from Gamo Gofa and Illubabor. The number of accessions with awn was more or less equal to the number of awnless accessions for all the entire samples combined. The same trend was maintained for accessions from Harerge and Gamo Gofa. However, most accessions from Eritrea, Shewa, Sidamo, Tigray, Welega and Welo had awn, while most of the accessions from Gonder, Illubabor and introduced materials were awnless. The extent of the seed covered by the glumes varied from 25% to 100%. For all the materials, the number of accessions in which the glumes had covered 25, 50, 75 and 100% of the grain decreased in the order given and a similar trend was followed for accessions from Gamo Gofa, Gonder, Sidamo and Welo. The opposite trend was exhibited by accessions from Welega. The most abundant glume colors were brown, black and gray. White glume was more frequent among the accessions from Tigray and Sidamo, while black glume was more frequent among the accessions from Welega.

Table 3.16. Proportions in different phenotypic classes for 10 qualitative characters in sorghum by regions of origin

Regions	Leaf midrib color				Waxy bloom				Panicle compactness and shape											
	1	2	3	4	3	5	7	9	1	2	3	4	5	6	7	8	9	10	11	12
Introduction	13	62	0	25	7	31	31	31	0	0	0	0	7	0	0	56	31	6	0	0
Eritrea	37	11	52	0	22	21	35	22	0	0	0	4	5	4	0	9	4	74	0	0
Gamo Gofa	17	21	29	33	29	25	25	21	0	0	21	4	0	4	8	13	17	0	25	8
Gonder	29	17	35	19	21	23	40	16	0	2	6	21	11	25	0	19	12	2	2	0
Harerge	12	46	28	14	36	33	12	19	0	0	0	4	2	0	0	7	12	72	3	0
Illubabor	4	24	25	47	9	38	31	22	7	0	2	2	7	7	4	51	2	0	16	2
Shewa	37	16	18	29	21	48	26	5	0	3	10	10	5	5	32	5	3	0	24	3
Sidamo	12	12	36	40	4	36	24	36	4	0	12	0	4	0	12	4	0	0	36	28
Tigray	10	7	70	13	30	47	10	13	3	0	3	10	7	44	3	17	10	3	0	0
Welega	29	14	36	21	29	43	28	0	0	0	7	0	57	0	29	0	0	0	7	0
Welo	10	5	65	20	37	28	22	13	0	0	0	3	2	12	0	25	31	27	0	0
All	19	20	39	22	24	32	26	18	1	1	4	6	7	10	6	18	12	24	8	3
Chi-square*	37	60	59	43	25	15	19	16	18	7	32	29	66	68	77	68	46	217	67	73

* = Chi-square value > 18 is significant at $p \leq 0.05$ and indicates that column (i.e., regional) proportions are significantly different for a particular character state.

Table 3.16. Continued.

Regions	Awn		Glume color						Grain covering			
	0	1	1	2	3	4	5	6	25	50	75	100
Introduction	94	6	7	12	25	0	6	50	25	69	6	0
Eritrea	30	70	0	6	48	2	33	11	37	43	18	2
Gamo Gofa	58	42	8	4	29	8	38	13	71	25	4	0
Gonder	63	37	6	15	19	0	41	19	44	31	19	6
Harerge	50	50	9	17	7	3	7	57	26	29	40	5
Illubabor	67	33	2	20	29	18	27	4	38	42	11	9
Shewa	21	79	8	3	26	32	21	10	32	34	24	10
Sidamo	28	72	24	0	8	36	16	16	40	28	16	16
Tigray	33	67	34	13	20	3	10	20	23	30	44	3
Welega	14	86	0	0	0	7	86	7	7	14	36	43
Welo	33	67	3	12	37	15	8	25	68	20	10	2
All	44	56	8	11	25	11	23	22	40	32	21	7
Chi-square*	61	61	45	18	40	55	68	67	47	22	37	43

* = Chi-square value > 18 is significant at $p \leq 0.05$ and indicates that column (i.e., regional) proportions are significantly different for a particular character state.

Table 3.16. Continued.

Regions	Seed color					Grain plumpness		Grain sub-coat		Endosperm texture				
	1	2	3	4	5	3	7	0	1	1	3	5	7	9
Introduction	69	0	0	25	6	63	37	44	56	0	13	31	37	19
Eritrea	41	7	6	46	0	87	13	65	35	0	9	22	43	26
Gamo Gofa	33	0	17	33	17	88	12	79	21	0	12	8	42	38
Gonder	48	10	17	23	2	52	48	67	33	0	11	12	44	33
Harerge	53	21	10	16	0	74	26	90	10	0	22	38	23	17
Illubabor	58	2	11	29	0	51	49	67	33	0	18	20	49	13
Shewa	11	10	32	42	5	45	55	87	13	0	16	5	53	26
Sidamo	28	0	52	20	0	52	48	76	24	0	0	0	68	32
Tigray	47	10	0	37	6	70	30	93	7	0	17	23	40	20
Welega	14	14	22	50	0	21	79	79	21	0	29	14	21	36
Welo	9	23	25	25	18	62	38	87	13	0	0	5	18	77
All	37	11	17	30	5	63	37	77	23	0	13	17	38	32
Chi-square*	63	28	48	22	40	46	46	35	35	0	23	40	34	73

* = Chi-square value > 18 is significant at $p \leq 0.05$ and indicates that column (i.e., regional) proportions are significantly different for a particular character state.

Among the different seed colors considered, the white and brown were the most frequent for all materials and for all the regions of origin, except for Welo where the frequency of white seed color was low. Yellow seed color was more frequent among the accessions from Harerge and Welo. Most of the accessions had dimple seeds and a similar trend was found for accessions from most of the regions. Exceptions were Illubabor in which accessions with both dimple and plump seed forms were equally frequent and Welega in which accessions with plump seeds were more frequent. On the whole, starchy endosperm texture was more frequent than corneous one for all the accessions and regions. Most of the accessions had no grain sub-coat color. The distributions of the different leaf midrib colors and waxy bloom did not show any apparent trend in general or relative to a particular region. It is interesting to mention here that we found one rare character, i.e., twin seeds. Only three accessions out of 415 (0.7%) had twin seeds, all from Eritrea.

Distribution of characters in the three adaptation zones

The compact and semi-compact panicles were most frequent in the lowland areas, while the loose panicles were more frequent in the intermediate and highland areas (Table 3.17). Presence of awn increased as one moves from lowland to highland areas (Table 3.17). The distribution of grain covering by the glume was also obvious. The proportion of the 50% grain covering was more frequent in the lowland than in the intermediate and highland, while the 100% covering was more in the highland than in the intermediate and lowland (Table 3.17). The 25% and 75% grain coverings were evenly distributed in the three adaptation zones.

Table 3.17. Proportions in different phenotypic classes for 10 qualitative characters in sorghum by adaptation zones

Adaptation zones	Leaf midrib color				Waxy bloom				Panicle compactness and shape											
	1	2	3	4	3	5	7	9	1	2	3	4	5	6	7	8	9	10	11	12
Lowland	16	26	38	21	24	27	28	21	0	0	1	4	4	7	0	25	18	38	2	1
Intermediate	28	15	44	13	16	30	36	18	3	1	11	6	12	15	8	5	7	16	15	1
Highland	16	16	35	33	33	43	13	11	1	1	4	10	6	11	14	19	5	8	14	7
All	19	20	39	22	34	36	35	31	1	1	4	6	7	10	6	18	12	24	8	3
Chi-square*	7	7	2	13	24	11	28	38	6	2	18	4	7	5	26	19	14	40	22	10

* = Chi-square value > 6 is significant at $p \leq 0.05$ and indicates that column (i.e., across adaptation zones) proportions are significantly different for a particular character state.

Table 3.17. Continued.

Adaptation zones	Awn		Glume color						Grain covering			
	0	1	1	2	3	4	5	6	25	50	75	100
Lowland	61	39	3	15	32	7	20	23	38	42	17	3
Intermediate	39	61	3	10	23	9	32	23	43	24	28	5
Highland	19	81	21	5	14	20	21	19	41	24	22	13
All	44	56	8	11	25	11	23	22	40	32	21	7
Chi-square*	52	52	37	7	12	13	6	0.7	0.8	15	5	12

* = Chi-square value > 6 is significant at $p \leq 0.05$ and indicates that column (i.e., across adaptation zones) proportions are significantly different for a particular character state.

Table 3.17. Continued.

Adaptation zones	Seed color					Grain plumpness		Grain sub-coat		Endosperm texture				
	1	2	3	4	5	3	7	0	1	1	3	5	7	9
Lowland	45	10	10	30	5	71	29	68	32	0	9	17	36	38
Intermediate	37	11	17	33	2	56	44	79	21	0	24	22	38	16
Highland	25	12	29	26	8	55	45	92	8	0	8	12	43	37
All	37	11	17	30	5	63	37	77	23	0	13	17	38	32
Chi-square*	12	0.3	18	1	4	11	11	23	23	0	17	4	1	17

* = Chi-square value > 6 is significant at $p \leq 0.05$ and indicates that column (i.e., across adaptation zones) proportions are significantly different for a particular character state.

Among the different seed colors, white and red showed clinal variation. White increased from highland to lowland, while red increased from lowland to highland (Table 3.17).

Estimates of diversity

The estimates of the Shannon-Weaver diversity index, H′, by region of origin and by adaptation zones for the ten characters are presented in Table 3.18. For all accessions, the value of H′ varied from 0.78 for grain sub-coat to 0.98 for waxy bloom and awn with an overall mean of 0.90. Apparently there was high diversity for all the characters. Pooled over characters within regions, the mean of H′ ranged from 0.69 for the introduced materials and Welega to 0.87 for Gonder. Individual characters showed different levels of diversity index in different regions. The diversity estimates of all characters were also high for the three adaptation zones (Table 3.19). Pooled over characters within adaptation zones, the mean of H′ ranged from 0.83 for the highland to 0.88 for the intermediate. It seems that the distribution of variability for the characters is more or less uniform in the three adaptation zones.

Partitioning of the phenotypic diversity into within and between the regions revealed that 86% of the total variation was found within regions of origin, while only 14% found between regions (Table 3.20). Panicle compactness and shape contributed relatively more (31%) to the regional differentiation and this agrees well with our field observations. Similarly, 96% of the variation was found within adaptation zones, while only 4% was found between adaptation zones (Table 3.20). Though the differentiation among the adaptation zones was weak, grain sub-coat and awn contributed relatively more to the observed differentiation.

Table 3.18. Estimates of the Shannon-Weaver diversity index, H′, for ten qualitative characters in sorghum by regions of origin

Regions	Leaf midrib color	Waxy bloom	Panicle Compactness & shape	Awn	Glume color	Grain covering	Seed color	Grain plumpness	Grain sub-coat	Endosperm texture	Mean ± S.E
Introduction	0.80	0.92	0.41	0.33	0.67	0.56	0.48	0.95	0.98	0.80	0.69 ± 0.07
Eritrea	0.68	0.98	0.39	0.88	0.62	0.81	0.67	0.55	0.94	0.78	0.73 ± 0.06
Gamo Gofa	0.97	1.00	0.76	0.98	0.78	0.52	0.82	0.55	0.75	0.74	0.79 ± 0.05
Gonder	0.97	0.94	0.75	0.95	0.73	0.86	0.81	0.99	0.91	0.75	0.87 ± 0.03
Harerge	0.89	0.93	0.39	1.00	0.57	0.89	0.74	0.84	0.46	0.83	0.75 ± 0.06
Illubabor	0.84	0.92	0.66	0.91	0.80	0.85	0.61	0.99	0.91	0.77	0.83 ± 0.04
Shewa	0.97	0.85	0.77	0.75	0.81	0.94	0.83	0.99	0.55	0.70	0.82 ± 0.04
Sidamo	0.89	0.87	0.65	0.86	0.76	0.94	0.63	0.99	0.79	0.38	0.78 ± 0.06
Tigray	0.67	0.86	0.69	0.91	0.82	0.84	0.69	0.88	0.37	0.83	0.76 ± 0.05
Welega	0.96	0.78	0.43	0.58	0.26	0.79	0.76	0.75	0.75	0.83	0.69 ± 0.07
Welo	0.71	0.95	0.59	0.91	0.80	0.64	0.97	0.97	0.55	0.41	0.75 ± 0.07

Table 3.19. Estimates of the Shannon-Weaver diversity index, H′, for ten qualitative characters in sorghum by adaptation zones

Adaptation zones	Leaf midrib color	Waxy bloom	Panicle Compactness & shape	Awn	Glume color	Grain covering	Seed color	Grain plumpness	Grain sub-coat	Endosperm texture	Mean ± S.E
Lowland	0.97	0.99	0.67	0.97	0.82	0.81	0.83	0.86	0.89	0.78	0.86 ± 0.03
Intermediate	0.91	0.97	0.90	0.97	0.82	0.87	0.84	0.98	0.75	0.83	0.88 ± 0.02
Highland	0.94	0.88	0.90	0.68	0.88	0.95	0.94	0.99	0.40	0.73	0.83 ± 0.06
All	0.96	0.98	0.86	0.98	0.87	0.90	0.88	0.95	0.78	0.80	0.90 ± 0.02

Table 3.20. Partitioning of phenotypic diversity into within and between regions of origin and adaptation zones

Characters	Region				Adaptation zone		
	H_{sp}	H_r	H_r/H_{sp}	$(H_{sp}-H_r)/H_{sp}$	H_z	H_z/H_{sp}	$(H_{sp}-Hz)/H_{sp}$
Leaf midrib color	0.96	0.85	0.89	0.11	0.94	0.98	0.02
Waxy bloom	0.98	0.91	0.93	0.07	0.95	0.97	0.03
Panicle compactness and shape	0.86	0.59	0.69	0.31	0.83	0.96	0.04
Awn	0.98	0.82	0.84	0.16	0.87	0.89	0.11
Glume color	0.87	0.69	0.79	0.21	0.84	0.97	0.03
Grain covering	0.90	0.78	0.87	0.13	0.88	0.98	0.02
Seed color	0.88	0.73	0.83	0.17	0.87	0.99	0.01
Grain plumpness	0.95	0.86	0.91	0.09	0.94	0.99	0.01
Grain sub-coat	0.78	0.72	0.92	0.08	0.68	0.87	0.13
Endosperm texture	0.80	0.71	0.89	0.11	0.78	0.98	0.02
Mean	0.90	0.77	0.86	0.14	0.86	0.96	0.04

H_{sp} = Diversity index for each character calculated from the entire data set; H_r and H_z= Average diversity index of each character for the eleven regions and the three adaptation zones, respectively; H_r/H_{sp} and H_z/H_{sp} = Proportion of diversity within regions and within adaptation zones, respectively; $(H_{sp}- H_r)/H_{sp}$ and ; $(H_{sp}- H_z/H_{sp}$ = Proportion of diversity between regions and between adaptation zones, respectively, in relation to the total variation.

Table 3.21. Associations among 10 qualitative characters in sorghum

Characters	Waxy bloom	Panicle compactness and shape	Awn	Glume color	Grain covering	Seed color	Grain plumpness	Grain sub-coat	Endosperm texture
	DF = 3	DF = 11	DF = 1	DF = 6	DF = 3	DF = 4	DF = 1	DF = 1	DF = 4
Leaf midrib color	8.86	87.64*	16.49**	27.07	20.22**	22.47**	1.94	0.74	16.54
Waxy bloom		52.95*	2.24	22.34	14.93	31.72**	3.70	4.39	5.80
Panicle compactness and shape			39.00**	124.80**	127.50*	139.73*	52.25**	9.26	57.92**
Awn				17.86**	17.98**	43.43**	1.56	24.34**	2.43
Glume color					60.15**	110.55**	11.82*	6.18	46.70**
Grain covering						21.17*	18.04**	6.20	14.10
Seed color							23.42**	26.80**	61.09**
Grain plumpness								0.59	1.02
Grain sub-coat									1.24

* and ** significant P = 0.05 and P = 0.01, respectively.

There was low and non-significant correlation (r = 0.35) between the mean diversity index for the regions pooled over the characters and the number of accessions from each region. The correlation between mean diversity index for the adaptation zones pooled over characters and the number of accessions from each adaptation zone was also low and non-significant (r = 0.10).

Chi-square heterogeneity analysis revealed that seed color was significantly associated with all the other characters (Table 3.20). This result suggests that the other characters might have their influences on seed color. Awn, glume covering and glume color were also significantly associated with other characters more than chance would permit.

Discussion

Quantitative characters

Univariate analysis

Phenotypic diversity in 415 sorghum accessions was studied for 15 quantitative traits. The results suggest existence of significant phenotypic variation among the accessions as a whole. Assuming that a significant portion of the phenotypic variation is genetic, it would be possible to make selection for such important characters like days for 50% flowering, plant height and panicle characteristics within a particular region. It was apparent that between regions variance was greater than between accessions pooled over regions and the latter was greater than that between accessions in some regions. Appa Rao et al. (1996) studied morphological diversity in sorghum germplasm from India and reported substantially higher differences among different states than within states. Since significant variation was found between

regions and between accessions within regions in the present study, it would be necessary to collect from as many regions as possible and to adequately sample the variable populations from different localities in a region in order to sample the variation.

In general, accessions from Shewa, Sidamo and Welega were characterized by tall plant height and late flowering. In areas with relatively high rainfall and high relative humidity, like in the south-west of Ethiopia, late flowering types mature after the cessation of rains and hence avoid grain-weathering problems. On the other hand, the introduced materials and accessions from Eritrea and Tigray were characterized by short plant height and early flowering, suggesting the possibility of obtaining genes for early flowering and short stature from the landraces of the latter regions. To be more specific, the results clearly indicated that sorghum accessions from Eritrea could be a good source of early flowering and short plant height genes for which there is a strong need in Ethiopia. Gebrekidan (1981) pointed out that early flowering and short stature sorghum types are not widely found among the Ethiopian sorghum collections and he emphasized that they are the kinds of sorghum most suitable for lowland areas with a limited amount of rainfall and a short growing season. This is also the main reason for introducing sorghum germplasm from other countries (Kebede, 1986; Debelo et al., 1995). The studies of Gebrekidan (1981) probably did not include materials from Eritrea because the Eritrean sorghum collections were not available for morphological description until recently (Doggett, 1988). The strong positive correlation between plant height and days for 50% flowering revealed in the present study, as well as in previous studies (Kumar and Singhania, 1984; Zongo et al., 1993) suggest that transfer of maturity and height genes into a desired genotype would be successful. Recent analysis of quantitative trait loci (QTL) showed linkage between QTL for height and maturity in sorghum (Lin et al., 1995).

The Eritrea, Harerge and Welo accessions had relatively short and narrow panicles with a higher number of primary branches per panicle, which are attributes of race durra (Doggett, 1988; Doggett and Prasad Rao, 1995). On the other hand, long and fewer primary branches, which are characteristic features of races bicolor and guinea (Doggett, 1988; Doggett and Prasad Rao, 1995), were recorded for accessions from Gamo Gofa, Illubabor, Shewa, Sidamo, and Welega. Taken together, there was similarity between the accessions from Eritrea and introduced materials; between accessions from Harerge, Tigray and Welo; between accessions from Gamo Gofa, Illubabor and Gonder; and between accessions from Shewa, Sidamo and Welega

The high coefficients of variation observed for most of the characters agreed well with those reported by Abu-el-Gasim and Kambal (1975) for indigenous sorghum of the Sudan. The different levels of regional variability of a particular character could be due to differences in forces of selection and/or differences in the intensity of a particular selecting force. Similar results were reported in tetraploid and hexaploid wheat (Bekele, 1984; Pecetti and Damania, 1996).

The wide range of variation detected by the various univariate statistics for the characters studied was in agreement with previous studies on Ethiopian sorghum germplasm (Gebrekidan, 1973; Kebede, 1991). Kwolek et al. (1986) also reported large differences in seed size in sorghum ranging from 0.70 to 6.10 g per 100 seeds, which is within the range (0.13-7.03 g per 100 seeds) observed in the present study. Large seed size in sorghum has been reported to be associated with increased germination percentage, improved stand establishment and increased grain yield (Kwolek et al., 1986; Kebede and Menkir, 1987). Available evidence also indicates that it is possible to improve seed size in sorghum by mass selection (Lothrop et al., 1985).

Information on geographical pattern of variation in Ethiopian sorghum germplasm has been mostly based on visual observations and personal experiences (Mengesha, 1975). Our results provided experimental evidence on occurrence of significant geographical variation and corroborated the idea that regions such as Gamo Gofa, Harerge, Illubabor, Gonder and Welo have high variation for sorghum in Ethiopia. The overall patterns of similarity or difference between regions seemed to depend on environmental factors such as rainfall, temperature, length of growing season and altitude. Regional mean for plant height (See Table 3.5 in relation to Fig. 3.1), for example, increased from north to south and from east to west, which obviously followed the rainfall, temperature and growing season patterns in Ethiopia (Tato, 1964; Stemler et al., 1977). Though its clinal variation was not as sharp as that for plant height, the mean for number of days for 50% flowering also followed the same trend. In this case temperature was implicated as a major factor since accessions from Gamo Gofa and Illubabor, which both included many accessions from hot lowland areas, did not follow the expected clinal variation. Similarly, Appa Rao et al. (1996) studied regional and seasonal variation of morphological characteristics of sorghum in India and attributed the considerable regional and seasonal variations for mean plant height and for mean number of days for 50% flowering to differences in growing season, temperature and day length. The clinal variation for plant height and days for 50% flowering could provide good opportunity for breeding and selection, particularly for designing breeding programs. Breeders may be able to predict the performance of segregating lines or hybrids based upon the height and maturity of the parental lines and design their breeding plans to obtain "ideal" types for a particular region. All in all, the results supported the previous studies (Gebrekidan, 1973) that Ethiopia, with its unique geographic and climatic features, possesses a tremendously high degree of morphological variation for sorghum.

Bivariate analysis

The significant positive correlation of panicle length with panicle width but negatively with number of primary branches per panicle may be explained by the "multiplication and condensation" hypothesis of Harlan et al. (1973). They defined multiplication as an increase of number of branches per inflorescence and condensation as shortening of internodes and branches in cereals during domestication. In sorghum, increase of number of branches results in shortening of the primary branches and the central axis of the inflorescence.

The correlation coefficients between altitude of the collection sites and most of the characters are significant and positive. As described by Harlan et al. (1973), ecological characteristics have considerable influence on the genotypic constitution of landraces during domestication and hence a relationship exists between the agro-ecology of the collection site and the morphological characteristics of the landraces. Thus, the significant positive correlation between the altitude of the collection site and the various plant characteristics would suggest that the variation between accessions is related to agro-ecological variations among the collection sites. The correlation coefficients of altitude with leaf width and peduncle exsertion were negative and non-significant, indicating that other environmental factors (i.e., other than altitude) and/or non-environmental factors might account for the variation for these particular characters. Appa Rao et al. (1996) suggested that the variation in peduncle exsertion is more in improved varieties than in landraces and it is also related to the type of sorghum race. Information on the relationships between environmental factors of the collecting sites and morpho-physiological variation of germplasm could enhance the understanding of evolutionary adaptive patterns, which could assist germplasm collectors and users (Elings, 1991; Annicchiarico et al., 1995). For instance, Stemler et al. (1977) pointed

out that the distribution of different races of sorghum in Ethiopia is largely determined by environmental factors such as temperature, rainfall and altitude. Since temperature decreases with an increase of altitude in Ethiopia (Tato, 1964; Stemler et al., 1977), it is more likely that temperature has exerted strong selection pressure on highland Ethiopian sorghum. It is to be noted that many cold tolerant sorghum lines have been identified from sorghum germplasm collections from the highlands of Ethiopia and Uganda (Singh, 1985; Doggett, 1988). Such relationship also calls for the need to have enough data on environmental variables of the collection sites such as temperature, rainfall, etc., if germplasm collection is required to meet its objectives effectively and efficiently. Unfortunately, data on such useful environmental variables like temperature and rainfall are lacking in the passport data of Ethiopian sorghum collections. Damania et al. (1996) and Pecetti and Damania (1996) also found environmental factors of the collection sites as very important determinants in structuring morpho-physiological variations in tetraploid wheat.

The significant positive correlation of leaf number with plant height and days for 50% flowering is in agreement with the results of Quinby (1967) and Hesketh et al. (1969). These characters also determine the competitive ability of a genotype when grown with other varieties or species (Hesketh et al., 1969). Stalk diameter and length of leaf sheath are traits that are related to lodging resistance (Esechie et al., 1977) and their positive associations with height would help in reducing the chance of lodging as height increases. In the present study, nearly all correlation coefficients involving panicle exsertion are negative. Abu-el-Gasim and Kambal (1975) also reported negative correlation coefficients for combinations involving peduncle exsertion.

Knowledge of the magnitude and the direction of correlation coefficients between quantitative characters would be beneficial for the interpretation of the patterns of variation.

Within the limit of experimental error and environmental effects, high correlation coefficients between characters may show that the characters share some common element of genetic control (i.e., pleiotropy, linkage) between genes or else from independently controlled characters responding similarly to geographic variation in selection pressures (Thorpe, 1976; Bekele, 1984). The between region (also called inter-region) correlation coefficient between the characters measures the concordance of their patterns of regional variation, while the within region (also called intra-region) correlation coefficient measures the association arising from genetic factors but is not affected by regional variation (Thorpe, 1976). Since this study showed significant positive correlations intra-regionally for some character combinations, it seems that common genetic control is playing a role in bringing about correlations between the various characters. However, it appears that similar responses to regional variation are playing a greater role than common genetic control as shown by the many more significant and moderate to high correlation coefficients inter-regionally than intra-regionally. Moreover, correlations among characters are of interest to plant breeders because they help in the identification of easily measured characters that could be used as indicators of more important (but more complex to score) characters. They are also useful in pointing out the possibility and limitation of simultaneous selection of desirable characters (Amurrio et al., 1993).

The large variation observed in this study and previous studies (Gebrekidan, 1973; Teshome et al., 1997) in Ethiopian sorghum germplasm could be ascribed to many factors. One important factor is the fact that sorghum is grown in many different environmental conditions. These include rainfall, temperature, altitude, growing period, and edaphic factors, resulting in 18 main agro-ecological zones and 49 subagro-ecological zones (MOA, 1998). Other factors are linguistic, cultural, historical and economic system differences among the people who are cultivating sorghum (Stemler et al., 1977), which contribute to its variation. The various physical, biological and human factors as well as complex interactions among

such factors all seem to have contributed to the wide range of variation of the crop in the country. Another source of variation must come from gene flow between the cultivated sorghum and its wild weedy relatives (Doggett and Majisu, 1968). Shattercanes (derivatives of wild sorghum x cultivated sorghum) are often noted as the most serious weeds on the highlands of Ethiopia, where they are known as *kello* or *sepo* (which means the fool in local languages) (Harlan, 1975; Doggett, 1988, 1991). As described by Harlan (1992), the shattercanes mimic the type of cultivated sorghum race with which they are associated in Ethiopia.

It is interesting to note here that many collectors of sorghum germplasm confuse shattercane with wild sorghum because of such camouflaging ability of the shattercanes (Harlan, 1984). There were certain accessions that the local farmers identified as *kello* or *sepo* from the materials that we used for this study. If farmers' folk taxonomy was right (Teshome et al., 1997), then the confusion holds also true the other way round, confusing the shattercanes with cultivated sorghum. In any case, gene flow occurs between the cultivated sorghum and its weedy relatives. The rate of outcrossing can be significant in sorghum depending upon genotypes and environmental factors. Outcrossing rate ranging from 5 to 7% is common in the race durra (Doggett, 1988) and from 10 to 30% in the race guinea (Ollitrault, 1987 cited in Deu et al., 1994). Commonly, farmers grow mixtures of genotypes or races of sorghum in the same field to meet immediate food requirements and for different purposes (Appa Rao and Mushonga, 1987), which provide a good opportunity for natural outcrossing. Since the bulk of sorghum production in Ethiopia is based on landraces (i.e., mixture of genotypes and even races) and since sorghum is grown under very variable environmental conditions, it is reasonable to presume outcrossing as one of the sources of the observed wide range of variation. For the future, it would be worthwhile to investigate the extent of outcrossing rate in contrasting environments within Ethiopia and Eritrea.

In conclusion, the results showed that there was a wide range of variation residing in the materials studied at regional and within region levels. Therefore, future germplasm collection should concern all levels of variation. Plant height and maturity are very important characters in sorghum adaptation and breeding. The clinal variation observed for these characters indicates that future germplasm sampling should pay more attention to populations along gradients of growing period, rainfall and temperature. The enormous variation would continue to provide breeders with good opportunities for breeding and selection.

Multivariate analysis

Results of multivariate analyses for 15 quantitative traits in 415 sorghum accessions suggest that the structure of morphological variation in the sorghum accessions studied is strongly influenced by environmental factors. The greater importance of such adaptive characters like plant height and days to 50% flowering in differentiating accessions of different regions in principal component analysis and the grouping together of accessions belonging to regions having similar agro-climatic conditions in cluster analysis, all indicate the relevance of environmental factors in ordering the structure of variation in the sorghum accessions. The relative positions of the regions of origin in Figures 3.2 and 3.3 are consistent with geographic variation of Ethiopia in terms of amount and distribution of rainfall, relative humidity and length of growing season. These environmental factors also determine the distribution of different races of cultivated sorghum in Ethiopia (Stemler et al., 1977). Thus the clustering pattern of the regions of origin also reflects the distribution patterns of different races of cultivated sorghum in Ethiopia (Stemler et al., 1977). This could be the most plausible reason

for the relatedness of the accessions from Gamo Gofa and Illubabor as revealed by cluster analysis, since the occurrence of similar races (for example, guinea, caudatum) in both regions has been reported by Prasada Rao and Mengesha (1981) and Harlan (1992). In general, the sorghum accessions are highly variable being clustered over ten clusters that are significantly different from each other. However, the lack of strong regional differentiation observed by the cluster and discriminant analyses could be partly ascribed to gene flow between regions (Doggett, 1988; Teshome et al., 1997).

The greater discrimination power of adaptation zones compared to the regions of origin of the accessions in the present study clearly indicated the greater importance of environmental factors (altitude, mean annual rainfall, growing season, etc.) than regions of origin in discriminating the sorghum accessions. It also corroborated previous studies (Gebrekidan, 1981) that Ethiopian sorghum germplasm could be classified into lowland, intermediate and highland materials. As human and natural selection factors affect morphological traits related to adaptation of a population, genetic distances from quantitative data allow inferences about the adaptation of populations (Camussi et al., 1985). Thus, classification using multiple agronomic characters identifies adaptation of a genotype and would improve the evaluation of genotype for potential adaptation (Souza and Sorrells, 1991). Moreover, as Zhong and Qualset (1995) have suggested, the clustering together of accessions from geographically similar areas is an indication for the evolution of co-adaptive association of quantitative characters. This is substantiated by clustering of standard varieties for lowlands with lowland landraces and standard varieties for highlands with highland landraces in the present study.

All parameters that were used invariably point to the fact that the accessions from Eritrea are closely related to the introduced materials. More importantly, both groups are characterized by early flowering and short stature (Ayana and Bekele, 2000). Two standard

varieties, Seredo and 76T1#23, which are early maturing and dwarf are also clustered in cluster III together with a great proportion of accessions from Eritrea. Lin and Binns (1985) and Lin et al (1986) showed the advantages of hierarchical cluster analysis in identifying useful germplasm particularly by including reference cultivars. The importance of early maturing and short stature types in national sorghum improvement has been pointed out (Gebrekidan, 1981; Kebede, 1986; Ayana and Bekele, 1999, 2000) and also discussed in the previous section. These authors emphasized that early maturing and short stature sorghum types are the most suitable for lowland areas of Ethiopia, where rainfall is limited and the growing season is short. From the results of the present study, it is concluded that the accessions from Eritrea could be a good source of early maturing and short plant height genes. The strong positive correlation between plant height and days to 50% flowering (Ayana and Bekele, 2000) suggests that the transfer of early maturity and short plant height genes into desired genotypes could be possible. The results also indicate that additional collection of short stature and early maturing types of sorghum should be made in the lowland areas of Gamo Gofa, Gonder, Harerge, Illubabor (Gambella), Tigray and Welo.

Categorizing germplasm accessions into morphologically similar, and presumably genetically similar groups (Souza and Sorrells, 1991), is useful for selecting parents for crossing. Crossing accessions belonging to different clusters of wide Mahalanobis distance could maximize opportunities for transgressive segregation. Because there is a higher probability that unrelated genotypes would contribute unique desirable alleles at different loci (Hussaini et al., 1977; Kanwal et al., 1983; Shamusddin, 1985; Peeters and Martinell, 1989; Beer et al., 1993). Therefore, the grouping of accessions by multivariate methods in the present study would be of practical value to sorghum breeders in that representative accessions may be chosen from different clusters for crossing programs.

The number of germplasm collections of major crops continues to grow both in national and international gene banks (Frankel and Brown, 1984; Brown, 1989a, b; Greener et al., 2000a, b). Although holding very large and heterogeneous germplasm collections is advantageous to preserve the genetic diversity of cultivated species before much is lost forever, the effective and efficient management and utilization of the germplasm collection is inversely related to its size (Brown, 1989a, b). As a solution to this paradox, Frankel (1984) proposed the setting up of core collections. A core collection would represent the genetic diversity of a crop species with a minimum repetitiveness, where in-depth evaluation and effective utilization could be made from such a collection, while the remaining collection is maintained as a reserve in a long-term storage (Frankel and Brown, 1984). Brown (1989a, b) outlined the genetic basis of the core collection concept. He suggested that different data sources (passport data, characterization and evaluation data of morphological traits, isozymes and molecular markers data such as restriction fragment length polymorphism and random amplified polymorphic DNA) could be used for establishing a core collection for a particular crop species. Based on cluster analysis of morphological data in the world collection of red clover (*Trifolium pratense* L.), Kouamé and Quesenberry (1993) suggested that a stratified random sampling of 10% of the accessions in each cluster accessions could constitute a core collection, representing the ecological and morphological variation of the germplasm. Prasada Rao and Ramanatha Rao (1995) used multivariate analysis of morphological data and Grenier et al. (2000a, b) used both morphological and molecular markers for establishing a core collection from the world collection of sorghum. The data from the present study together with data from isozymes (Aayna et al., 2000c) and random amplified polymorphic DNA (RAPD) data (Ayana et al., 2000a), can form a baseline data towards establishing a core collection for Ethiopian sorghum germplasm.

Results of the present study also can help to define strategies for further collection. Since our results indicate that the structure of observed morphological variation is governed by environmental factors, future germplasm collections should aim to explore as many geographically and climatically different areas within the country as possible, instead of collecting extensively within individual regions.

Qualitative characters

In this section, results obtained with regard to the amount of variation and patterns of distribution of the variation for 10 categorical traits in 415 sorghum accessions are discussed. The results show that the regional distribution of the different panicle types are not random, which could be explained in two ways. First, the patterns of distribution of the different panicle types appeared to follow the temperature, humidity and rainfall patterns of Ethiopia and Eritrea. The compact and semi-compact panicle types were more frequent in Eritrea and in relatively hot and dry regions of Ethiopia such as Harerge, Tigray and Welo. The loose panicle types with dropping branches occurred abundantly in relatively cool and wet regions of Ethiopia such as Welega, Illubabor, Shewa and Sidamo. This finding corroborated the suggestions of Stemler et al. (1975) and Prasada Rao and Mengesha (1981) that the open panicle of sorghum is an adaptive trait, which facilitates quick drying of the panicle in areas of high rainfall and humidity, thereby minimizing grain weathering due to fungal diseases such as grain mould.

Secondly, the distribution of the different panicle types reflects the distribution of different races of sorghum (Harlan and de Wet 1972; Acheampong et al. 1984). The compact oval panicle with goose peduncle, characteristic of the durra race (Doggett 1988), was limited

to the east and northeastern regions of Ethiopia and Eritrea. Brooke (1958) and Stemler et al. (1977) reported the predominant occurrence of the durra race in the east of the Rift Valley in Ethiopia. According to Stemler et al. (1977), race durra-bicolor (with an intermediate morphological structure between the races durra and bicolor) is grown in the southwest of Ethiopia because it does better under cold and high rainfall conditions than the race durra. The fact that accessions from Gamo Gofa and Illubabor had the semi-compact and loose to very loose panicle types in this study supported the occurrence of different races such as guinea (often with open panicle) and caudatum (often with moderately compact panicle) in these regions (Stemler et al., 1977; Prasada Rao and Mengesha, 1981; Harlan, 1992).

From the results of the distribution of glume covering and glume color, it was possible to deduce the predominant occurrence of race bicolor in Welega, because enclosing black glumes and loose panicle types are often characteristics of bicolor (Harlan and de Wet, 1972; Doggett, 1988). The abundance of brown, black and gray glumes in our materials was in disagreement with that of Kebede (1991), who reported that the most predominant glume color in Ethiopian sorghum germplasm was purple, perhaps due to their lumping of color groups. Many workers (Pluncknett et al., 1983; Bekele, 1984; Clawson, 1985; Harlan, 1992) emphasized the importance of the practice of intra- and inter-specific diversity in traditional tropical agriculture as a means of ensuring harvest security, adding variety to diet and preserving genetic diversity. Most farmers in Ethiopia grow different types of sorghum in the same field, which are frequently distinguished on the basis of seed color. In addition to differing in resistance to environmental stresses, bird damage, as well as in appearance and taste, the seed color-based varieties often differ in time of maturity and uses. The high variation for seed color observed in the present study thus could be the result of both human and natural selection.

to the east and northeastern regions of Ethiopia and Eritrea. Brooke (1958) and Stemler et al. (1977) reported the predominant occurrence of the durra race in the east of the Rift Valley in Ethiopia. According to Stemler et al. (1977), race durra-bicolor (with an intermediate morphological structure between the races durra and bicolor) is grown in the southwest of Ethiopia because it does better under cold and high rainfall conditions than the race durra. The fact that accessions from Gamo Gofa and Illubabor had the semi-compact and loose to very loose panicle types in this study supported the occurrence of different races such as guinea (often with open panicle) and caudatum (often with moderately compact panicle) in these regions (Stemler et al., 1977; Prasada Rao and Mengesha, 1981; Harlan, 1992).

From the results of the distribution of glume covering and glume color, it was possible to deduce the predominant occurrence of race bicolor in Welega, because enclosing black glumes and loose panicle types are often characteristics of bicolor (Harlan and de Wet, 1972; Doggett, 1988). The abundance of brown, black and gray glumes in our materials was in disagreement with that of Kebede (1991), who reported that the most predominant glume color in Ethiopian sorghum germplasm was purple, perhaps due to their lumping of color groups. Many workers (Pluncknett et al., 1983; Bekele, 1984; Clawson, 1985; Harlan, 1992) emphasized the importance of the practice of intra- and inter-specific diversity in traditional tropical agriculture as a means of ensuring harvest security, adding variety to diet and preserving genetic diversity. Most farmers in Ethiopia grow different types of sorghum in the same field, which are frequently distinguished on the basis of seed color. In addition to differing in resistance to environmental stresses, bird damage, as well as in appearance and taste, the seed color-based varieties often differ in time of maturity and uses. The high variation for seed color observed in the present study thus could be the result of both human and natural selection.

The predominant occurrence of starchy endosperm in the present study was in complete agreement with previous studies (Kebede, 1991). The type of sorghum endosperm texture is related to utilization and storage potential of the grain (Bello et al., 1990; Attere, 1994). Accordingly, the corneous endosperm is preferred to the starchy one for making *tô* or *ugali* (traditional thick African porridge) and for long storage period of the grain. On the other hand, starchy endosperm is preferred for making bread such as *injera* (a traditional pancake like bread in Ethiopia) and where the grain is stored for only a few months. Since a great bulk of sorghum grain is used for making *injeria* in Ethiopia, it could be more likely that the frequency of starchy endosperm texture increased by human selection. The presence of grain sub-coat color, which is an undesirable trait for preparing any sort of food from sorghum grain, occurred in several accessions of all regions often, deceitfully, associated with white seed. In the absence of bird damage and/or grain mold, good food quality is usually associated with white seed (Doggett, 1982).

The distributions of the different leaf midrib colors and waxy bloom did not show any apparent trend in general or relative to a particular region. Certainly, there is no conscious human selection for these characters in subsistence agriculture though farmers often distinguish sweet sorghum by its leaf midrib color (Teshome et al., 1997). Available evidence indicates that sorghum lines with brown leaf midrib color are desirable for forage (Duncan, 1996).

It is interesting to mention here that we found one rare character, i.e., twin seeds. Only three accessions out of 415 (0.7%) had twin seeds, all from Eritrea. This character is also reported by Harlan quoted in Frankel and Soule (1981, p. 195) to be rare among the world collections of sorghum and it is said to have medicinal value for whooping cough among pastoralists in the bordering areas of Ethiopia and Sudan.

Distribution of characters in the three adaptation zones

The distribution of panicle compactness and shape also showed clear trends in the three adaptation zones. The compact panicle types occurred more frequently in the lowland than in the intermediate and highland, while the occurrence of the loose types generally followed the opposite trend, indicating further the adaptive significance of panicle compactness and shape. Asfaw (1989) also reported significant association between altitude and spike morphology of Ethiopian barley (*Hordeum vulgare* L.).

The more abundance of awnless in the lowland than in the intermediate and highland areas needs explanation. Under warm and dry climatic conditions and where leaf diseases affect leaves, the presence of awn is advantageous in small cereals like wheat and barley (Patterson and Ohm, 1975 and references therein; Negassa, 1986), serving as the closest photosynthetic apparatus to the seed. The situation of sorghum seems to contradict the adaptive significance of awn in dry areas. It is likely that the awnless character is favored in sorghum to reduce evapo-transpiration in dry lowland areas. The need for immediate photosynthetic products might be satisfied by the relatively large number of glumes on a typical sorghum panicle as compared to a typical spike of wheat or barley. The validity of this speculation warrants further investigation. An alternative explanation could be the more prevalence of certain leaf diseases in sorghum in intermediate and highland than in lowland areas. Thus it may be possible that the awn compensates for reduction of photosynthesis that may result from loss of leaf tissues due to leaf diseases in the intermediate and highland areas. The apparent increase of the frequency of red seed color from low to high and the opposite trend for white seed color suggests that seed color might have adaptive significance. It seems that the white grain sorghum lacks certain polyphenolic compounds that serve to protect the sorghum grain from pre-harvest germination in humid regions in the intermediate and

highland areas, while red and brown colored seeds are rich in polyphenolic compounds (Asante, 1995).

Estimates of diversity

Characters with two to four phenotypic classes (i.e., descriptor states) generally had higher diversity index than those characters with more than four phenotypic classes. For example, nearly all regions had low diversity index for panicle compactness and shape compared to other characters. This should be attributed to unequal distributions of the accessions of every region over the 12 phenotypic classes of the character rather than lack of variation for the character. As Negassa (1985) remarked, this result signified the caution that should be made while interpreting the estimates of diversity index of different characters with different phenotypic classes as measured by the Shannon-Weaver diversity index. Thus, it might be misleading to compare the values of H′ from characters having different classes.

There is more differentiation on the regional level than on the adaptation zones for the categorical traits, which contrasts with the results from discriminant analysis of quantitative characters. This could be because of the nature of the two sets of characters used for the study. Taken together, panicle types, awn, glume and seed colors and grain sub-coat are useful in discriminating accessions of different regions or adaptation zones. The classification of Ethiopian sorghum into three adaptation zones (i.e., into lowland, intermediate and highland) has been made largely on altitude, length of growing period and amount and distribution of rainfall (Gebrekidan, 1981; Kebede and Menkir, 1987). Characters, such days to 50% flowering, maturity and plant height, are more important in determining sorghum adaptation to a particular ecological zone (Doggett, 1988). In a parallel study involving these and other

quantitative characters, Ayana and Bekele (1999) found that the discrimination of sorghum accessions was more pronounced when discriminant analysis was based on adaptation zones rather than regions of origin.

Since the level of diversity was high and more or less similar for all regions and adaptation zones pooled over characters, further collections should give equal weight for all regions and adaptation zones regardless of the extent of acreage of sorghum production in a particular region and adaptation zone. This suggestion was further supported by the low and non-significant correlation ($r = 0.35$) between the mean diversity index for the regions pooled over the characters and the number of accessions from each region. The correlation between mean diversity index for the adaptation zones pooled over characters and the number of accessions from each adaptation zone is also low ($r = 0.10$). Lack of a significant correlation between the number of accessions per region and the mean diversity index pooled over characters was also reported in wheat (Jain et al., 1975; Jaradat, 1992 and Pecetti et al., 1992) and in barley (Demissie and Bjørnstad, 1996).

Chi-square contingency table analysis revealed that seed color was significantly associated with all the other characters. This result suggests that the other characters might have their influences on seed color. Awn, glume covering and glume color were also significantly associated with other characters more than chance would permit. Olsen and Miller (1958) hypothesized that the degree of inter-dependence among morphological characters is directly related to the degree of phenotypic morphological integration among the characters. The degree of integration was measured by the intensity of statistical association in the phenotype using the chi-square test, which revealed a very intense character association.

In conclusion, the overall high diversity index ($\bar{H}$ = 0.90) observed for the qualitative in this study is consistent with Doggett's long standing hypothesis that Ethiopia is not only one of the centers of diversity but also the center of domestication of sorghum (Dogget, 1991; Doggett and Prasada Rao, 1995). Knowledge of adaptive traits linked to certain ecological conditions help in choosing sites for *in situ* conservation that need to be integrated with *ex situ* conservation. Our results clearly revealed that the patterns of distribution of panicle compactness and shape followed the temperature, humidity and rainfall patterns of Ethiopia and Eritrea, indicating that panicle compactness and shape have adaptive significance and could serve as criteria in choosing sites for sorghum *in situ* conservation strategy. The localized distributions of the different panicle types observed in this study and the adaptive and taxonomic significance of panicle morphology (Harlan and de Wet, 1972) would also indicate that samples collected from one or few regions do not represent samples from other regions. Therefore, future study should include adequate samples from all regions and analysis of variation should consider the finely classified main and subagro-ecological zones (MOA, 1998). This would help to obtain better-refined information on the impact of environmental variables on the distribution of genetic variation of sorghum in Ethiopia.

Summary

A total of 415 sorghum accessions consisting of 391 landraces collected from different geographical regions in Ethiopia and Eritrea, 8 varieties and 16 introduced elite breeding lines were evaluated for 15 quantitative and 10 qualitative characters to determine the extent and geographical patterns of morphological variation. The extent of variation was highly

portion of the total variation was found within regions (86%) of origin and within adaptation zones (96%). Panicle compactness and shape contributed relatively more to regional differentiation. These characters were also found to be disproportionately distributed within the regions of origin with the compact panicles frequently distributed in relatively dry regions, while the loose panicle types were widely found in relatively wet and humid regions. The differential distribution of the different panicle types indicated the adaptive significance of panicle compactness and shape and at the same time reflected the distribution patterns of different races of sorghum in Ethiopia.

Based on the observed patterns of variation both for the quantitative and qualitative characters, it is concluded that the morphological variation in the material studied is structured by environmental factors. The implications of the results in plant breeding, germplasm collection and conservation as well as the probable sources of the wide range of variation are discussed.

Chapter 4

Allozyme Variation in Sorghum (*S. bicolor* ssp. *bicolor* (L.) Moench)

Introduction

Assessing the amount and distribution of genetic variation maintained in germplasm collections would enhance effective management and efficient utilization of the collections (Bretting and Widrlechner, 1995; Lee, 1995; Karp et al, 1997). In the previous chapter (see also Ayana and Bekele, 1998, 1999, 2000), it has been shown that sorghum germplasm from Ethiopia and Eritrea is characterized by a wide range of variation for agro-morphological characteristics, confirming earlier studies (Gebrekidan and Kebede, 1977; Teshome et al., 1997). Furthermore, using 415 sorghum accessions for 15 quantitative characters, Ayana and Bekele (1999) showed that the discrimination of these accessions is more pronounced when the analysis was made based on the adaptation zones rather than the regions of origin. Partitioning of the total variation ($\overline{H} = 0.90$) in 415 sorghum accessions for 10 categorical traits showed that a large proportion of the variation (86% and 96%) is found within the regions of origin and within the adaptation zones, respectively. Apparently, relative differentiation was more on the basis of regions of origin than on the basis of the adaptation zones.

Assessment of genetic variation in germplasm collections for agro-morphological traits is interesting for a plant breeder who looks for adaptive traits. However, the fact that these traits are polygenically inherited and hence highly influenced by environmental factors limits their use for accurate measurement of genetic variation. In the succeeding two chapters, 80 accessions of cultivated sorghum were analyzed for 20 RAPD primers (see also Ayana et

al., 2000a) and 11 populations of wild sorghum for 9 RAPD primers (see also Ayana et al., 2000b). The results showed intermediate levels of variation both for the cultivated and wild sorghum. Similar to the categorical morphological characters, partitioning of the total RAPD variation showed that a large proportion of the variation is found within regions of origin and within adaptation zones (Ayana et al., 2000a, b). The major limitation of RAPD for assessment of genetic variation is the dominant inheritance of these markers (Williams et al., 1990, 1993).

Primarily because of its co-dominant inheritance, allozyme analysis is one of the best candidates for analyzing genetic variation in germplasm collections (Weeden, 1989). A number of studies have already used allozyme analysis to estimate the amount and distribution of genetic variation in sorghum (Morden et al., 1989, 1990; Aldrich et al., 1992; Dje et al., 1998, 1999). However, these studies did not consider the allozyme variation in relation to altitude, which influences largely the adaptation of sorghum to different eco-geographic conditions in the tropics.

In this study, the extent and distribution of allozyme variation was determined in 48 sorghum accessions in relation to their geographical origin and adaptation zones. An estimate of the level of gene flow between sorghum accessions in different regions and adaptation zones was also provided.

Materials and Methods

Plant material

A total of 48 accessions of sorghum (*Sorghum bicolor* (L.) Moench) was sampled from 80 accessions previously used for RAPD study (Ayana et al., 2000a). The sampling was made in such a way that each region and each adaptation zone within a region is well represented relative to the original sample. The 48 accessions were grouped into three equal adaptation zones, i.e. lowland, intermediate elevation, and highland accessions (Appendix 4.1). Seeds maintained from previous morphological and agronomic traits characterization (Ayana and Bekele, 1998, 1999, 2000) were used. Ten individual plants were assayed per accession.

Enzyme assay

A total of 11 enzyme systems were analyzed with horizontal starch gel electrophoresis. They are acid phosphatase (ACP, E.C. 3.1.3.2), aspartate aminotransferase (AAT, E.C. 2.6.1.1), catalase (CAT, E.C. 1.11.1.6), esterase (EST, E.C. 3.1.1.-), glucose-6-phosphate isomerase (GPI, E.C. 5.3.1.9), isocitrate dehydrogenase (IDH, E.C. 1.1.1.42), malate dehydrogenase (MDH, E.C. 1.1.1.37), peroxidase (PRX, E.C. 1.11.1.7), phosphoglucomutase (PGM, E.C. 2.7.5.1), phosphogluconate dehydrogenase, (PGD, E.C. 1.1.1.44) and shikimate dehydrogenase (SKD, E.C. 1.1.1.25).

The ACP, GPI, IDH, MDH, PGM, PGD, PRX and SKD enzyme systems were resolved using a discontinuous buffer system with 0.005 M L-histidine-HCl, pH 8.0, gel buffer and 0.410 M trisodium citrate, pH 8.0 electrode buffer (Wendel and Weeden, 1989).

This system was set to run at a constant current of 50 mA and approximately 300 volt for six hours. The AAT, CAT and EST enzyme systems were also resolved in a discontinuous buffer system with 0.076 M Tris, pH 8.6, gel buffer and 0.3 M boric acid, pH 8.0, electrode buffer (Wendel and Weeden, 1989). This system was set to run at a constant current of 120 mA and approximately 300 volt for 8 hours.

For both systems, 11% starch was boiled in 425 ml of gel buffer and poured into a 10 mm gel-casting tray a day before electrophoresis was to be run. The gel surface was covered with plastic film, wrapped with moist paper, put in a plastic bag and stored at 4°C overnight. For extraction, dark grown 7 to 12 days old plant tissue was used. Plant material was cut and placed in an ice-cooled 1.5 ml eppendorf tube to which 100 μl extraction buffer (10 mM Tris-HCl, pH 7.2 and 0.05% β-mercaptoethanol) was added (Persson and Bothmer, 2000). Homogenization was made using a glass rod mounted onto a rotatory homogenizer and the homogenized sample was kept on ice until loading.

One highland commercial variety of sorghum, ETS-2752, was used in every gel to aid as reference while comparing migration distance of bands of different occasions. The gel was kept cool during electrophoresis by a thermo-circulator machine set at 2°C.

Following electrophoresis, the gel was sliced into three slices of 3 mm thickness, excluding the top thin slice, and stained for a particular enzyme system. For the AAT, IDH, and SKD enzyme systems, staining recipes given by Morden et al. (1987) were used. For the ACP, CAT and EST, (colorimetric) enzyme systems, staining recipes given by Wendel and Weeden (1989) were used. For the GPI, MDH, PGD, PGM and the PRX enzyme systems, staining recipes given by Persson and Bothmer (2000) were used.

Knowledge of the genetic and structural bases of the enzymes would be essential for accurate interpretation of the bands scored. These aspects were inferred from Morden et al.

(1987), Morden et al. (1989, 1990), Kephart (1990), and Aldrich et al. (1992). By so doing, it was possible to determine the putative number of loci present for each enzyme assayed. The following designation was used: AAT (enzyme), Aat-1, 2, etc., (locus), Aat-1a, b, etc. alleles per locus. For enzyme systems with more than one locus, the most anodal band was designated as locus one, and successive cathodal loci designated locus two and three, respectively. Similarly, the most anodal band (i.e. allele) within a locus was designated as "a".

Statistical analysis

For each of the 48 sorghum accessions, genotype frequencies were calculated for each locus. The genotype frequencies were further analyzed using BIOSYS-1 release 1.7 software package (Swofford and Selander, 1989) to obtain estimates for the following standard population genetics parameters: mean number of alleles per locus (A), mean number of alleles per polymorphic locus (A_p), effective number of alleles per locus (A_e), percent polymorphic loci (P), observed heterozygosity (H_o) and expected heterozygosity (H_e) (Nie, 1978). Differences in allele frequencies among accessions was examined using contingency chi-square analyses for polymorphic loci. Since the number of individuals (10) analyzed per accession was small, test for Hardy-Weinberg expectations and computation of fixation indices were not made.

The genetic population structure was investigated for polymorphic loci using the gene diversity statistics of Nei (1973, 1977). For each polymorphic locus, the total gene diversity (H_T) was partitioned into its components, i.e., gene diversity within (H_S) and among (D_{ST}) sorghum accessions. Hence, $H_T = H_S + D_{ST}$. The coefficient of gene differentiation among the

accessions was quantified by the gene diversity statistics, $G_{ST} = D_{ST}/H_T$. The value of G_{ST} indicates the proportion of total gene diversity that resides among accessions.

An indirect estimate of gene flow was obtained using the formula suggested by Wright (1965).

$$Nm = \frac{(1 - F_{ST})}{4F_{ST}}$$

where Nm is the absolute number of migrants between populations per generation and F_{ST} is the proportional deviation of overall observed heterozygosity from expected heterozygosity in the subpopulations. Theoretically, the value for F_{ST} ranges from 0 to 1, where 0 indicates that all samples are fixed for the same allele and 1 indicates that all populations are monomorphic, but not all of them are fixed for the same allele. The significance of F_{ST} from zero was tested using the x^2 test.

$$X^2 = 2NF_{ST}(k\text{-}1);\ df = (k\text{-}1)(s\text{-}1),$$

where 'N' is the total number of individuals sampled, 'k' is the number of alleles at a locus and 's' is the number of population subdivisions.

Nei's (1978) unbiased genetic identities (I) and unbiased genetic distances (D) were calculated from allele frequencies at all loci (i.e. both monomorphic and polymorphic) for pair-wise comparison of the accessions. The genetic distances matrix was used to perform an hierarchical cluster analysis using the unweighted pair group method with arithmetic averages (UPGMA) (Sneath and Sokal, 1973). A dendrogram was constructed from the cluster analysis to further illustrate the patterns of the relationship of the sorghum accessions.

All the above analyses were repeated for pooled accessions within regions of origin and within, as well as between adaptation zones. In addition, multivariate ordination of the regions of origin was performed by subjecting the variance-covariance matrix of allele frequencies at polymorphic loci to principal component analysis.

Results

A total of 23 putative loci encoding 11 enzyme systems was analyzed. A maximum of 27 alleles were scored across the 48 sorghum accessions studied. Nineteen loci, each with one allele per locus, were found to be monomorphic and fixed for the same allele in the 48 accessions. Only 4 loci (Acp-3, Cat, Est-3 and Prx), each with 2 alleles per locus, were found to be polymorphic across the 48 accessions.

Allele frequency variation

The allele frequencies at the polymorphic loci are presented in Table 4.1. None of the alleles was unique for a particular accession. Instead, the common allele of each locus was the same in most of the accessions. Moreover, no locus was polymorphic in all the accessions. The Prx locus was the most variable locus, being polymorphic in 33 of the 48 accessions. The other 15 accessions were fixed for the most frequent allele at this locus. The Acp-3 locus was the least variable as it was polymorphic only in 8 accessions, while the remaining 38 and 2 accessions were fixed for the most frequent and less frequent allele, respectively. The Cat and Est-3 loci were polymorphic in 16 accessions. For both loci, the remaining 32 accessions were fixed for the most common allele, except for three accessions from Welega, which were fixed for the less frequent allele at the Cat locus. In terms of accessions, only two accessions, one each from Bale and Shewa, were polymorphic for all four loci. Polymorphism for three, two and one loci prevailed in 8, 13, and 2 accessions, respectively. Nine of the accessions were fixed for all loci.

Table 4.1. Allele frequencies at polymorphic loci in 48 accessions (sample size 10)

Adaptation zone	No	Acp-3		Cat		Est-3		Prx	
		A	b	a	b	a	b	a	b
Lowland	1	0.50	0.50	0.85	0.15	0.85	0.15	0.70	0.30
	2	1.00	0.00	1.00	0.00	1.00	0.00	0.50	0.50
	3	1.00	0.00	0.65	0.35	1.00	0.00	0.75	0.25
	4	1.00	0.00	0.90	0.10	1.00	0.00	0.40	0.60
	5	1.00	0.00	0.50	0.50	0.65	0.35	0.90	0.10
	6	1.00	0.00	1.00	0.00	1.00	0.00	0.85	0.15
	7	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	8	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	9	0.90	0.10	1.00	0.00	0.40	0.60	0.75	0.25
	10	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	11	1.00	0.00	0.65	0.35	1.00	0.00	0.65	0.35
	12	1.00	0.00	0.50	0.50	1.00	0.00	0.80	0.20
	13	1.00	0.00	1.00	0.00	0.80	0.20	0.75	0.25
	14	0.50	0.50	0.00	1.00	1.00	0.00	1.00	0.00
	15	1.00	0.00	1.00	0.00	0.85	0.15	1.00	0.00
	16	1.00	0.00	0.65	0.35	0.80	0.20	1.00	0.00
Intermediate	17	1.00	0.00	1.00	0.00	0.50	0.50	0.60	0.40
	18	1.00	0.00	1.00	0.00	0.80	0.20	0.80	0.20
	19	1.00	0.00	0.50	0.50	0.85	0.15	0.25	0.75
	20	0.50	0.50	0.40	0.60	1.00	0.00	0.75	0.25
	21	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	22	1.00	0.00	1.00	0.00	1.00	0.00	0.60	0.40
	23	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	24	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	25	1.00	0.00	1.00	0.00	1.00	0.00	0.75	0.25
	26	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	27	1.00	0.00	1.00	0.00	1.00	0.00	0.95	0.05
	28	1.00	0.00	1.00	0.00	0.70	0.30	0.85	0.15
	29	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
	30	0.60	0.40	0.00	1.00	1.00	0.00	1.00	0.00
	31	1.00	0.00	1.00	0.00	0.90	0.10	1.00	0.00
	32	1.00	0.00	0.50	0.50	1.00	0.00	1.00	0.00
Highland	33	1.00	0.00	1.00	0.00	0.75	0.25	0.85	0.15
	34	1.00	0.00	1.00	0.00	1.00	0.00	0.50	0.50
	35	0.00	1.00	1.00	0.00	1.00	0.00	0.50	0.50
	36	1.00	0.00	1.00	0.00	1.00	0.00	0.40	0.60
	37	0.55	0.45	0.75	0.25	1.00	0.00	0.75	0.25
	38	1.00	0.00	1.00	0.00	1.00	0.00	0.90	0.10
	39	1.00	0.00	0.75	0.25	0.70	0.30	0.60	0.40
	40	0.70	0.30	0.50	0.50	0.85	0.15	0.75	0.25
	41	1.00	0.00	1.00	0.00	1.00	0.00	0.05	0.95
	42	1.00	0.00	0.50	0.50	0.85	0.15	0.80	0.20
	43	1.00	0.00	1.00	0.00	1.00	0.00	0.90	0.10
	44	1.00	0.00	0.65	0.35	1.00	0.00	0.60	0.40
	45	1.00	0.00	0.85	0.15	1.00	0.00	0.05	0.95
	46	0.00	1.00	0.00	1.00	1.00	0.00	1.00	0.00
	47	1.00	0.00	0.50	0.50	0.90	0.10	0.95	0.05
	48	0.50	0.50	1.00	0.00	1.00	0.00	0.40	0.60
Overall mean		0.89	0.11	0.80	0.20	0.92	0.08	0.76	0.24

Allele frequencies at the 4 polymorphic loci are summarized by regions of origin (Table 4.2). The most common alleles were the anodal bands for loci for pooled accessions within the respective regions of origin, exceptions are the Cat and the Acp-3 loci in Welega, where the cathodal bands were more frequent than the anodal ones. All the four loci were polymorphic for pooled sorghum accessions within Bale, Harerge, Shewa and Welo. Only the Prx and the Acp-3 loci were polymorphic for pooled accessions within Kafa and Welega, respectively. Although no locus was polymorphic in all the regions, the Prx locus was the most variable locus being polymorphic in 12 of the 13 regions (except in Welega), while Acp-3 was the least variable, being polymorphic in 6 regions. Taken together, the mean allele frequency data showed that Bale, Harerge, Shewa, and Welo were the most variable regions, while Kafa and Welega were the least variable ones.

Table 4.2. Sample size and mean allele frequencies at polymorphic loci for pooled accessions within regions of origin and adaptation zones

Category		N	Acp-3		Cat		Est-3		Prx	
			A	b	a	b	A	b	a	b
Regions	Bale	20	0.75	0.25	0.92	0.08	0.93	0.07	0.60	0.40
	Eritrea	20	1.00	0.00	0.78	0.22	1.00	0.00	0.58	0.42
	Gamo Gofa	60	1.00	0.00	0.92	0.08	0.78	0.22	0.75	0.25
	Gojam	20	1.00	0.00	0.75	0.25	0.93	0.07	0.63	0.37
	Gonder	50	0.70	0.30	0.88	0.12	1.00	0.00	0.73	0.27
	Harerge	60	0.91	0.09	0.96	0.04	0.90	0.10	0.83	0.17
	Illubabor	30	1.00	0.00	0.72	0.28	1.00	0.00	0.82	0.18
	Kafa	20	1.00	0.00	1.00	0.00	1.00	0.00	0.88	0.12
	Shewa	40	0.93	0.07	0.81	0.19	0.89	0.11	0.59	0.41
	Sidamo	30	1.00	0.00	0.82	0.18	0.85	0.15	0.85	0.15
	Tigray	40	1.00	0.00	0.88	0.12	0.95	0.05	0.60	0.40
	Welega	30	0.37	0.63	0.00	1.00	1.00	0.00	1.00	0.00
	Welo	60	0.92	0.08	0.78	0.22	0.91	0.09	0.89	0.11
Weighted mean		37	0.89	0.11	0.80	0.20	0.92	0.08	0.76	0.24
Adaptation zone	Lowland	160	0.93	0.07	0.79	0.21	0.90	0.10	0.82	0.18
	Intermediate	160	0.94	0.06	0.84	0.16	0.92	0.08	0.85	0.15
	Highland	160	0.80	0.20	0.78	0.22	0.94	0.06	0.62	0.38
Mean		160	0.89	0.11	0.80	0.20	0.92	0.08	0.76	0.24

All the four loci were polymorphic for pooled accessions within the adaptation zones (Table 4.2). Here also the common allele was the same at all loci in the three adaptation zones. There was a slight trend for the Prx locus to be more variable among the highland accessions than among the lowland and intermediate elevation accessions.

Genetic variation

The genetic variation observed in the 48 sorghum accessions was within a very narrow range (Table 4.3). For the individual accessions, from each of which 10 plants were studied, the mean number of alleles per locus (A) varied from 1.0 to 1.2 (mean 1.1), while the mean number of alleles per polymorphic locus (A_P) varied from 1.0 to 2.0 (mean 1.4). The mean effective number of alleles per locus (A_e) ranged from 1.00 to 1.08 (mean 1.03). The percent polymorphic loci (P) varied from 0.0 to 17.4 (mean 6.7). The mean observed heterozygosity (H_o) and expected heterozygosity (H_e) ranged from 0.000 to 0.091 (mean 0.026) and from 0.000 to 0.071 (mean 0.026), respectively. Considering the pooled data for the 48 accessions as species level variation, A_s, A_{ps}, A_{es}, P_s, H_{os}, and H_{es} were 1.2, 2.0, 1.04, 17.4, 0.023, and 0.042, respectively.

For the regions of origin (Table 4.4), A varied from 1.0 for Kafa to 1.2 for Bale, Harerge, Shewa and Welo (weighted mean 1.13), while A_P varied from 1.3 for Kafa to 2.0 for Bale, Harerge, Shewa and Welo (weighted mean 1.8). A_e ranged from 1.01 for Kafa to 1.05 for Bale, Gojam, Gonder and Shewa (weighted mean 1.04). P varied from 4.3 for Kafa and Welega to 17.4 for Bale, Shewa and Welo (weighted mean 12.8). H_o ranged from 0.000 for Welega to 0.048 for Bale (weighted mean 0.026), while H_e ranged from 0.010 for Kafa to 0.051 for Bale (weighted mean 0.037).

Table 4.3. Estimates of genetic variation in 48 accessions: Mean number of alleles per locus (A); mean number of alleles per polymorphic locus (AP); effective number of alleles per locus (Ae); percent polymorphic loci (P); observed heterozygosity (H_o); and expected heterozygosity (H_e)

Adaptation zone	Acc. No	A	Ap	Ae	P	Ho	He
Lowland	1	1.2	2.0	1.07	17.4	0.087 ± 0.048	0.065 ± 0.032
	2	1.0	1.3	1.02	4.3	0.009 ± 0.009	0.023 ± 0.023
	3	1.1	1.5	1.04	8.7	0.052 ± 0.037	0.038 ± 0.026
	4	1.1	1.5	1.03	8.7	0.009 ± 0.009	0.030 ± 0.023
	5	1.1	1.8	1.06	13.0	0.083 ± 0.052	0.052 ± 0.031
	6	1.0	1.3	1.02	4.3	0.013 ± 0.013	0.017 ± 0.012
	7	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	8	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	9	1.1	1.8	1.05	13.0	0.030 ± 0.019	0.047 ± 0.028
	10	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	11	1.1	1.5	1.04	8.7	0.043 ± 0.033	0.042 ± 0.029
	12	1.1	1.5	1.04	8.7	0.052 ± 0.044	0.038 ± 0.027
	13	1.1	1.5	1.03	8.7	0.004 ± 0.004	0.032 ± 0.022
	14	1.0	1.3	1.02	4.3	0.000 ± 0.000	0.023 ± 0.023
	15	1.0	1.3	1.01	4.3	0.013 ± 0.013	0.012 ± 0.012
	16	1.1	1.5	1.04	8.7	0.048 ± 0.034	0.035 ± 0.025
Intermediate	17	1.1	1.5	1.05	8.7	0.052 ± 0.044	0.045 ± 0.031
	18	1.1	1.3	1.03	8.7	0.017 ± 0.017	0.029 ± 0.020
	19	1.1	1.8	1.06	13.0	0.061 ± 0.045	0.052 ± 0.030
	20	1.1	1.8	1.07	13.0	0.091 ± 0.055	0.067 ± 0.034
	21	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	22	1.0	1.3	1.02	4.3	0.000 ± 0.000	0.022 ± 0.022
	23	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	24	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	25	1.0	1.3	1.02	4.3	0.004 ± 0.004	0.017 ± 0.017
	26	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	27	1.0	1.3	1.00	4.3	0.004 ± 0.004	0.004 ± 0.004
	28	1.1	1.5	1.03	8.7	0.030 ± 0.026	0.031 ± 0.022
	29	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	30	1.0	1.3	1.02	4.3	0.000 ± 0.000	0.022 ± 0.022
	31	1.0	1.3	1.01	4.3	0.009 ± 0.009	0.008 ± 0.008
	32	1.0	1.3	1.02	4.3	0.043 ± 0.043	0.023 ± 0.023
Highland	33	1.1	1.5	1.03	8.7	0.035 ± 0.025	0.029 ± 0.020
	34	1.0	1.5	1.02	4.3	0.000 ± 0.000	0.023 ± 0.023
	35	1.0	1.3	1.02	4.3	0.017 ± 0.017	0.023 ± 0.023
	36	1.0	1.3	1.02	4.3	0.017 ± 0.017	0.022 ± 0.022
	37	1.1	1.8	1.06	13.0	0.057 ± 0.32	0.057 ± 0.032
	38	1.0	1.3	1.01	4.3	0.000 ± 0.000	0.008 ± 0.008
	39	1.1	1.8	1.06	13.0	0.057 ± 0.034	0.058 ± 0.032
	40	1.2	2.0	1.08	17.4	0.070 ± 0.046	0.071 ± 0.034
	41	1.0	1.3	1.00	4.3	0.004 ± 0.004	0.004 ± 0.004
	42	1.1	1.8	1.05	13.0	0.057 ± 0.045	0.049 ± 0.028
	43	1.0	1.3	1.01	4.3	0.000 ± 0.000	0.008 ± 0.008
	44	1.1	1.5	1.05	8.7	0.039 ± 0.031	0.043 ± 0.030
	45	1.1	1.5	1.02	8.7	0.017 ± 0.014	0.016 ± 0.012
	46	1.0	1.0	1.00	0.0	0.000 ± 0.000	0.000 ± 0.000
	47	1.1	1.8	1.04	13.0	0.057 ± 0.044	0.035 ± 0.024
	48	1.1	1.5	1.05	8.7	0.061 ± 0.046	0.045 ± 0.031
Mean population		1.1	1.4	1.03	6.7	0.026 ± 0.012	0.026 ± 0.019
Species level		1.2	2.0	1.04	17.4	0.023 ± 0.013	0.042 ± 0.017

For the three adaptation zones (Table 4.4), there was no difference in A and A_P, while P was 13.0 for intermediate elevation accessions and 17.4 for the lowland and highland accessions (mean 15.9). H_o varied from 0.014 for the intermediate to 0.030 for the highland (mean 0.023), while H_e varied from 0.030 for the intermediate to 0.054 for the highland (mean 0.041).

Heterogeneity of allele frequency and genetic population structure

Chi-square tests for heterogeneity of allele frequencies at polymorphic loci are presented in Table 4.5. There were significant differences in allele frequencies at all loci for the entire accessions and for regions of origin as well as for the within and among the adaptation zones, although there was not diagnostic alleles in any of the categories.

The genetic population structure is summarized in Table 4.5. The mean total gene diversity for the 48 sorghum accessions was 0.25; partitioning of it into its component parts gave 59% within and 41% among accessions. For the regions of origin, 80% of the total gene diversity was found among accessions within the regions of origin, while only 20% was found among accessions between regions of origin. Partitioning of the total genetic variation within each of the 3 adaptation zones revealed that 68%, 58%, and 57% of the variation was found within the lowland, intermediate elevation and highland accessions, respectively. The corresponding variation among accessions within adaptation zones was 32%, 42%, and 43%, respectively. A very large portion of the total gene diversity (97%) was found among accessions within adaptation zones, while only the remaining 3% was found among accessions between the adaptation zones.

Table 4.4. Estimates of genetic variation by regions of origin and adaptation zones: Mean number of alleles per locus (A); mean number of alleles per polymorphic locus (A_P); effective number of alleles per locus (A_e); percent polymorphic loci (P); observed heterozygosity (H_o); and expected heterozygosity (H_e)

Category		A	Ap	A_e	P	H_o	H_e
Regions	Bale	1.2	2.0	1.05	17.4	0.048 ± 0.026	0.051 ± 0.027
	Eritrea	1.1	1.5	1.04	8.7	0.030 ± 0.022	0.037 ± 0.026
	GamoGofa	1.1	1.8	1.04	13.0	0.033 ± 0.021	0.038 ± 0.022
	Gojam	1.1	1.8	1.05	13.0	0.030 ± 0.022	0.044 ± 0.026
	Gonder	1.1	1.8	1.05	13.0	0.025 ± 0.014	0.045 ± 0.026
	Harerge	1.2	2.0	1.03	13.0	0.014 ± 0.007	0.031 ± 0.016
	Illubabor	1.1	1.5	1.03	8.7	0.032 ± 0.025	0.031 ± 0.022
	Kafa	1.0	1.3	1.01	4.3	0.002 ± 0.002	0.010 ± 0.010
	Shewa	1.2	2.0	1.05	17.4	0.034 ± 0.020	0.050 ± 0.026
	Sidamo	1.1	1.8	1.04	13.0	0.029 ± 0.019	0.035 ± 0.019
	Tigray	1.1	1.8	1.04	13.0	0.015 ± 0.012	0.035 ± 0.023
	Welega	1.0	1.3	1.02	4.3	0.000 ± 0.000	0.021 ± 0.021
	Welo	1.2	2.0	1.04	17.4	0.038 ± 0.022	0.038 ± 0.019
Weighted mean		1.13 ± 0.08	1.79 ± 0.02	1.04 ± 0.07	12.75 ± 1.30	0.026 ± 0.030	0.037 ± 0.031
Adaptation zone	Lowland	1.2	2.0	1.04	17.4	0.025 ± 0.013	0.039± 0.019
	Intermediate	1.2	2.0	1.03	13.0	0.014 ± 0.007	0.030 ± 0.015
	Highland	1.2	2.0	1.06	17.4	0.030 ± 0.016	0.054 ± 0.028
Mean		1.2. ±00	2.0 ± 0.00	1.04 ± 0.01	15.9 ± 1.80	0.023 ± 0.005	0.041 ± 0.007

Table 4.5. Heterogeneity chi-square (X^2) test and degrees of freedom (df), total gene diversity (H_T), gene diversity within (H_S) and among (D_{ST}), coefficient of gene differentiation (G_{ST}) and an indirect estimate of gene flow (N_m) for the entire sorghum accessions, regions of origin, and adaptation zones

Category	Locus	X^2	df	H_T	H_S	D_{ST}	G_{ST}	$F_{(ST)}$	N_m
Accessions	Acp-3	593.00**	47	0.19	0.08	0.12	0.60	0.62	0.17
	Cat	500.27**	47	0.31	0.16	0.16	0.50	0.52	0.25
	Est-3	249.05**	47	0.15	0.11	0.03	0.22	0.26	0.89
	Prx	331.60**	47	0.36	0.25	0.11	0.31	0.35	0.56
	Mean			0.25	0.15	0.10	0.41	0.44**	0.36
Among Regions	Acp-3	270.75**	12	0.20	0.14	0.06	0.30	0.33	0.58
	Cat	298.67**	12	0.34	0.23	0.11	0.32	0.35	0.53
	Est-3	63.41**	12	0.12	0.11	0.01	0.08	0.07	2.88
	Prx	83.03**	12	0.37	0.34	0.03	0.08	0.10	2.88
	Mean			0.26	0.21	0.05	0.20	0.21**	1.00
Within Lowland	Acp-3	135.69**	15	0.13	0.08	0.05	0.62	0.42	0.15
	Cat	148.97**	15	0.33	0.18	0.15	0.55	0.47	0.20
	Est-3	94.57**	15	0.18	0.14	0.04	0.78	0.30	0.07
	Prx	71.63**	15	0.30	0.25	0.05	0.83	0.22	0.05
	Mean			0.24	0.16	0.07	0.69	0.35**	0.11
Within Intermediate	Acp-3	135.39**	15	0.11	0.06	0.05	0.55	0.42	0.20
	Cat	211.25**	15	0.27	0.10	0.17	0.37	0.66	0.43
	Est-3	87.43**	15	0.14	0.11	0.03	0.79	0.27	0.07
	Prx	107.55**	15	0.26	0.18	0.08	0.69	0.34	0.11
	Mean			0.20	0.11	0.09	0.55	0.45**	0.20
Within Highland	Acp-3	232.58**	15	0.32	0.09	0.23	0.28	0.73	0.64
	Cat	146.80**	15	0.34	0.19	0.15	0.56	0.46	0.20
	Est-3	54.11**	15	0.11	0.10	0.01	0.91	0.17	0.02
	Prx	111.36**	15	0.47	0.32	0.15	0.68	0.35	0.12
	Mean			0.31	0.18	0.14	0.61	0.43**	0.16
Among adaptation Zones	Acp-3	43.56**	2	0.20	0.19	0.01	0.05	0.04	4.75
	Cat	8.45*	2	0.29	0.29	0.00	0.00	0.01	0.00
	Est-3	8.60*	2	0.13	0.13	0.00	0.00	0.01	0.00
	Prx	50.97**	2	0.36	0.34	0.02	0.06	0.05	3.92
	Mean			0.25	0.24	0.01	0.03	0.03**	8.08

* and ** = significant at 5% and 1%, respectively.

Gene flow

An indirect estimate of the levels of gene flow between accessions, regions and within and between adaptation zones is given in Table 4.5. The results showed low levels of gene flow (mean $N_m \leq 1$) between the accessions, regions of origin and between accessions within the adaptation zones. On the other hand, the level of gene flow was high among the adaptation zones. The high significance of mean F_{ST} for the sorghum accessions, regions of origins as well as within and among adaptation zones supplemented the chi-square tests, which revealed that allele frequencies were different for all the categories.

Genetic distance and cluster analysis

The mean unbiased genetic distance among the 48 sorghum accessions was 0.02 and varied narrowly from 0.00 to 0.09 (data not shown). A dendrogram constructed from the corresponding genetic distance matrix failed to group the sorghum accessions either on the basis of region of origin or adaptation zone (Fig. 4.1).

For the regions of origin, the genetic distance varied from 0.00 to 0.06 with a mean of 0.01 (Table 4.6). Cluster analysis clearly showed three groups (Fig. 4.2). The mean genetic distance among the adaptation zones ranged from 0.000 to 0.003 and averaged 0.002 (Table 4.7). For both the regions of origin and adaptation zones, the mean within genetic distance was greater than the between genetic distance, suggesting that variation was more pronounced from accession to accession than from region to region or from adaptation to adaptation zone.

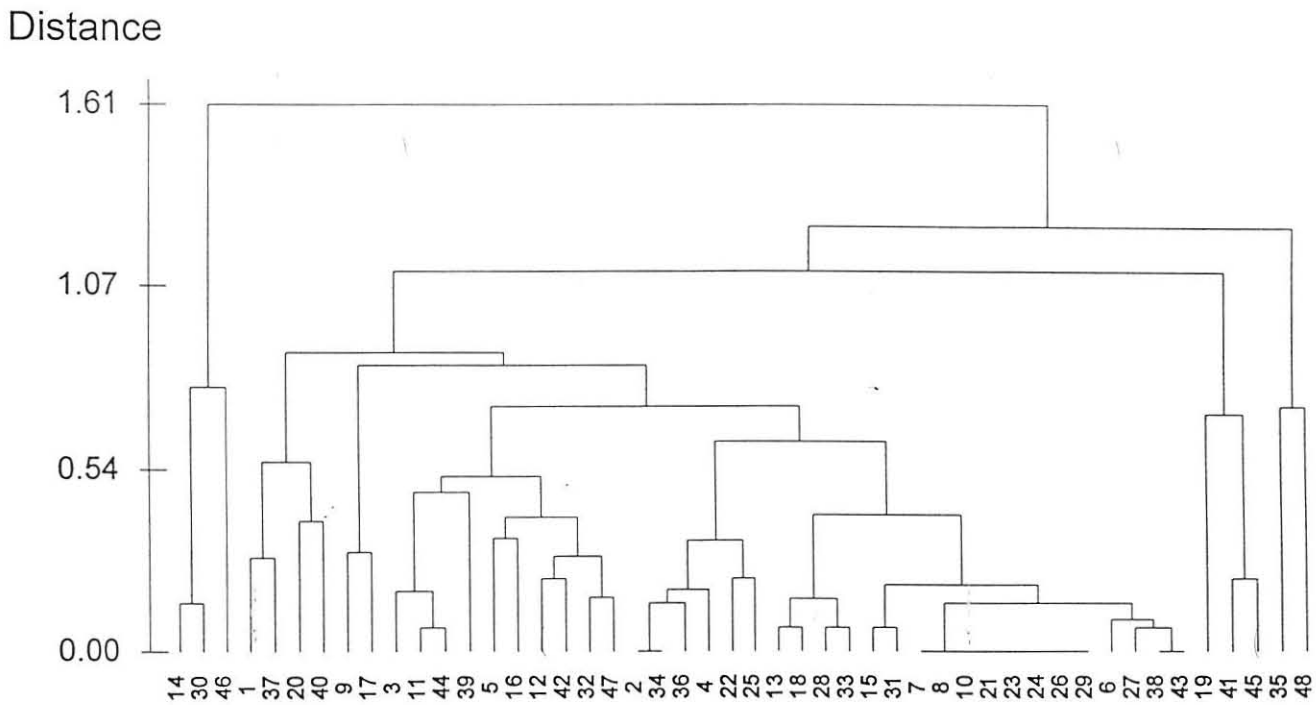


Fig. 4.1. Dendrogram generated based on unbiased genetic distance for 23 allozyme loci for 48 sorghum accessions

Table 4.6. Unbiased genetic distance (above diagonal), mean within region unbiased genetic distance (diagonal) and genetic identity (below diagonal) for 13 regions of origin of sorghum based both on monomorphic and polymorphic loci

Region	1	2	3	4	5	6	7	8	9	10	11	12	13
Bale	0.013	0.003	0.004	0.003	0.000	0.003	0.006	0.006	0.001	0.005	0.002	0.053	0.005
Eritrea	0.997	0.007	0.004	0.000	0.005	0.005	0.002	0.006	0.000	0.004	0.000	0.054	0.005
GamoGofa	0.996	0.996	0.001	0.002	0.006	0.001	0.004	0.003	0.002	0.000	0.002	0.062	0.003
Gojam	0.997	1.000	0.998	0.036	0.005	0.004	0.001	0.005	0.000	0.002	0.000	0.050	0.003
Gonder	1.000	0.995	0.994	0.995	0.055	0.003	0.005	0.005	0.003	0.005	0.004	0.044	0.004
Harerge	0.997	0.995	0.999	0.996	0.997	0.015	0.003	0.001	0.003	0.001	0.003	0.057	0.001
Illubabor	0.994	0.998	0.996	0.999	0.995	0.997	0.007	0.003	0.003	0.001	0.003	0.043	0.001
Kafa	0.994	0.994	0.997	0.995	0.995	0.999	0.997	0.003	0.005	0.002	0.004	0.064	0.003
Shewa	0.999	1.000	0.998	1.000	0.997	0.997	0.997	0.995	0.021	0.003	0.000	0.053	0.004
Sidamo	0.995	0.996	1.000	0.998	0.995	0.999	0.999	0.998	0.997	0.008	0.003	0.052	0.000
Tigray	0.998	1.000	0.998	1.000	0.996	0.997	0.997	0.996	1.000	0.997	0.017	0.061	0.004
Welega	0.948	0.948	0.940	0.951	0.957	0.945	0.958	0.938	0.949	0.949	0.941	0.008	0.042
Welo	0.995	0.995	0.997	0.997	0.996	0.999	0.999	0.997	0.996	1.000	0.996	0.959	0.014

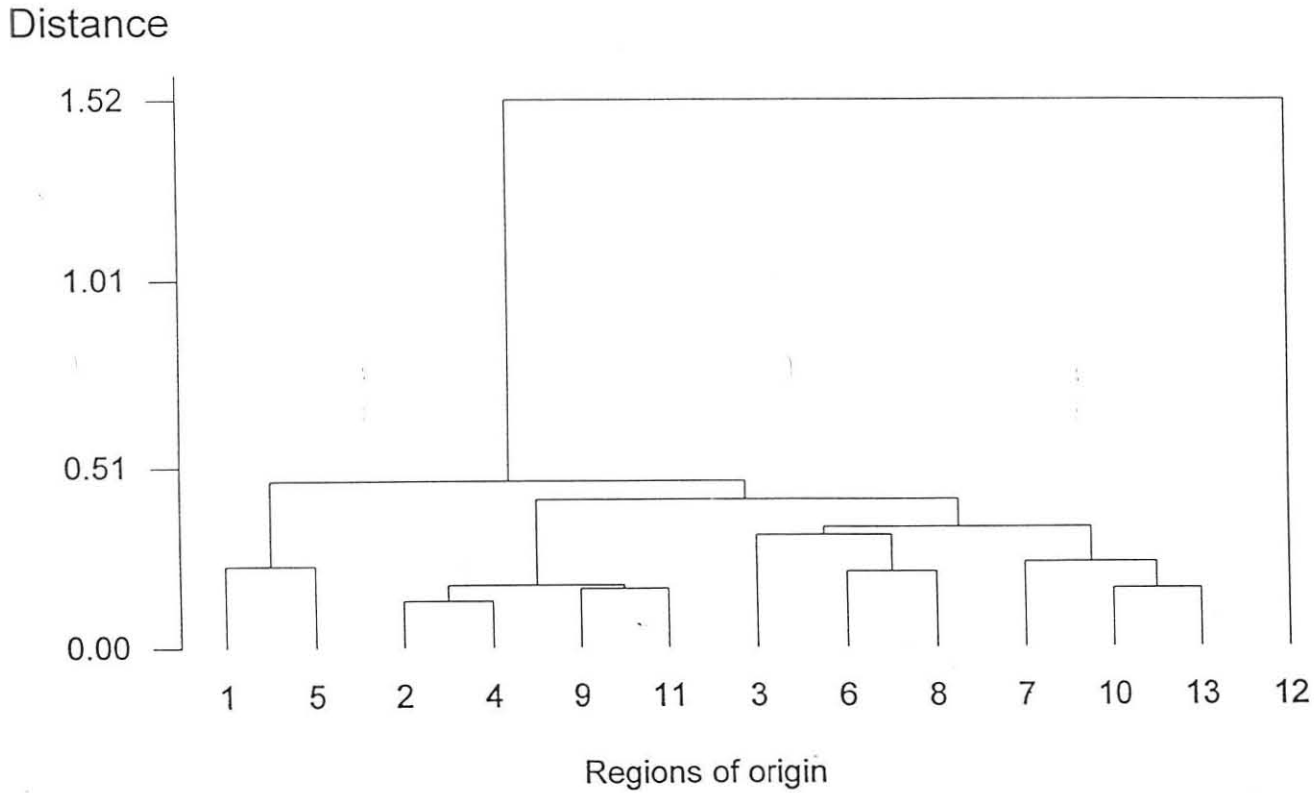


Fig 4.2. Dendrogram for regions of origin of sorghum

1 = Bale, 2 = Eritrea, 3 = Gamo Gofa, 4 = Gojam, 5 = Gonder, 6 = Harerge, 7 = Illubabor, 8 = Kafa, 9 = Shewa 10 = Sidamo, 11 = Tigray, 12 = Welega, and 13 = Welo.

Table 4.7. Unbiased genetic distance (above diagonal), mean within adaptation genetic distance (diagonal) and unbiased genetic identity (below diagonal) based on both monomorphic and polymorphic loci for the adaptation zones

Adaptation zones	1	2	3
Lowland	0.026	0.000	0.002
Intermediate	1.000	0.016	0.003
Highland	0.998	0.997	0.026

Principal component analysis

Principal component analysis revealed that the first three principal components (PC) accounted for 92% of the total variation (Table 4.8). For the regions of origin, the first three PCs explained 96% of the total variation (Table 4.8). The first PC alone, with an eigenvalue of 0.05, explained 97% of the total variation found among the adaptation zones (Table 4.8).

A plot of the regions' principal scores against the first two principal axes distinguished three groups (Fig. 4.3): (1) Gamo Gofa, Harerge, Illubabor, Sidamo and Welo with high scores on PC1 and PC2, (2) Bale, Eritrea, Gojam, Gonder, Shewa and Tigray with high scores on PC1 but low scores on PC2 and (3) Welega with low scores on both axes.

Correlation of allele frequencies with altitude

The correlation between allele frequencies at the polymorphic loci and altitude was low and non-significant for Acp-3, Cat and Est-3 (Table 4.9). Significant ($P< 0.05$) but weak negative correlation ($r = -0.29$) was found between the most common allele of the Prx locus and vice versa for the least common allele ($r = 0.29$).

2000; potato, Rio et al., 1997). This calls for a comparative study of *in situ* and *ex situ* materials to draw firm conclusions about the level of isozyme variation in sorghum. An alternative explanation for discrepancy among studies could be the use of different enzyme systems. Enzymes such as acid phosphatase, esterase, peptidase, and peroxidase are usually more polymorphic than other enzymes (Weeden, 1989), an observation well supported by the present study.

At the population level, the amount and distribution of allozyme variation is primarily influenced by breeding system. Hamrick and Godt (1990) reviewed a large number of allozyme diversity studies in plant species in general and in cultivated crops in particular (Hamrick and Godt, 1997) and concluded that self-pollinating plants and cultivated crops have lower levels of allozyme variation than their outbreeding counterparts. Their conclusions also showed that self-pollinating plants maintain more of their genetic variation among populations, while the outbreeding plants maintain more genetic variability within populations. Sorghum is predominately a self-pollinating crop (Doggett, 1988). All the measures of genetic variation obtained in the present study and in previous studies (Morden et al., 1989; Aldrich et al., 1992; Djè et al., 1998, 1999) are less than the mean corresponding values reported for self-pollinating plants (Hamrick and Godt, 1990, 1997). The results confirm the earlier conclusions made by Morden et al. (1989) that sorghum is depauperated in allozymic variation compared to other cereals such as barley (self-pollinating) (Brown and Munday, 1982) and maize (outbreeding but belonging to the same tribe as sorghum) (Doebley et al., 1985).

Analysis of population genetic structure showed that 41% of the variation is found among sorghum accessions, which lies between the mean values reported by Djè et al. (1999) (G_{ST} = 7%) and Morden et al. (1989) (G_{ST} = 91%) and relatively closer to the value of G_{ST} for self-pollinating plant species (G_{ST} = 51%) (Hamrick and Godt, 1990) and self-pollinating

cultivated crops (G_{ST} = 59%) (Hamrick and Godt, 1997). The discrepancy among the studies can be due to sample size differences as well as differences in the number of informative loci and the number of alleles per locus. Our results also showed that the proportion of genetic variation maintained among regions (G_{ST} = 20%) is the same as the corresponding value among sorghum fields reported by Djè et al. (1998). Similar results have also been obtained in Ethiopian and Eritrean sorghum germplasm for qualitative morphological characters (Ayana and Bekele, 1998) and random amplified polymorphic DNA (Ayana et al., 2000a).

According to Wright's (1978) classification of the level of genetic differentiation, the values of F_{ST} obtained for the sorghum accessions and for the regions are very great and great, respectively. The x^2 was also significant for F_{ST}. Fundamentally, F_{ST} and G_{ST} are equivalent (Slatkin and Barton, 1989; Berg and Hamrick, 1997). By analogy, G_{ST} is also significant. This finding plus the highly significant heterogeneity in allele frequencies both among sorghum accessions and among the regions of origin, the low frequency of some of the alleles and the low levels of gene flow among sorghum accessions and among regions, all lead us to recommend that future collection should consider sampling as many accessions as possible from all parts of Ethiopia and Eritrea.

One of the goals of this study was to examine the extent and patterns of distribution of genetic variation among the three adaptation zones of sorghum in Ethiopia and Eritrea. The correlation of allele frequencies with the altitude of the collection sites was poor. This contrasts with maize allozyme allele frequencies, which significantly associated with altitude (Doebley et al., 1984, 1985). Furthermore, there was low level of genetic differentiation (G_{ST} = 3%) among the adaptation zones, despite high levels of genetic differentiation among accessions within adaptation zones. This may be explained by the observed high level of gene flow, which leads to genetic homogeneity among the adaptation zones.

However, the fact that we found low levels of gene flow among the sorghum accessions as a whole and among the regions of origin, makes gene flow an unlikely explanation. Alternatively, the low genetic differentiation among the adaptation zones can be explained by invoking a suggestion that the altitudinal range of sorghum has expanded recently to produce the appearance of a high level of gene flow. In other words, sorghum might have radiated recently from lowland, its natural habitat, into intermediate and highland elevations through human activity within the age of sorghum. Perhaps the idea of recent radiation is viable because very recently DNA sequence data showed no differences among ancient and modern races of sorghum (Rowley-Conwy et al., 1999). The authors concluded that sorghum is domesticated recently in northeast Africa, although several studies indicated the use of sorghum seeds by humans as late as 8000 before present (Wendorf et al., 1992; Wasylikowa, 1997; Wasylikowa and Dahlberg, 1999). It is to be noted that Harlan and de Wet (1972) classified cultivated sorghum into five basic (*bicolor*, *caudatum*, *durra*, *guinea* and *kafir*) and ten intermediate races based largely on panicle morphology. However, recent allozyme and molecular data do not support the classification. For instance, Morden et al. (1989) found higher levels of differentiation on the basis of regions of origin than on the basis of racial classification, which could be explained by invoking the hypothesis that the modern races of sorghum radiated recently from the ancient bicolor-type. If the hypothesis that sorghum is domesticated recently is proved, then the evolutionary time may not be enough for sorghum to accumulate enough gene mutations that could lead to high allozyme variation. However, such an argument is difficult to test at this stage. Another possibility is that although the bands with same migration distance are considered as similar alleles in allozyme assays, they could differ in their amino-acid sequences and what one considers as a single allele per locus could be constellation of alleles.

The fact that many of the loci assayed were monomorphic and the low level of polymorphism in the present study and also in previous studies dealing with allozyme variation in sorghum, suggest that the wide range of morphological variation is caused by phenotypic plasticity or by genetic variation at a few loci. Lack of congruence between morphological variation and genetic variation has also been observed in crop plants such as cultivated groundnut (*Arachis hypogaea* L.) (Subramanian et al., 2000 and references therein) and melon (*Cucumis melo* L.) (Dane, 1983; Shattuck-Eidens et al., 1990), as well as in wild plants such as *Poa annum* L. (Poaceae) (Frenot et al., 1999) and arctic *Potentilla* (Rosaceae) (Hansen et al., 2000). The latter two groups of authors attributed the conspicuous morphological variation to phenotypic plasticity or genetic variation at limited number of loci.

Summary

The amount and distribution of genetic variation was investigated in 48 sorghum landrace accessions, representing 13 regions of origin and three adaptation zones (lowland, intermediate and highland elevation) in Ethiopia and Eritrea. Assaying 11 enzyme systems, 23 putative loci were scored for a total of 27 alleles. Nineteen loci were monomorphic and fixed for the same allele, while the remaining four loci, each with two alleles, were polymorphic across the 48 accessions. The results show significant differences in allele frequencies among the accessions, regions of origin and the adaptation zones. However, all measures of genetic variation used show that accessions maintained much lower levels of variation than the corresponding mean values for self-pollinating crop plants, confirming previous conclusions that sorghum is depauperated in allozymic variation. The total gene diversity was 0.25, which partitioned 59% within and 41% among accessions. Similarly, most of the total gene diversity was found within the regions of origin (80%) and within the adaptation zones (97%). Both the dendrogram (constructed from Nei's unbiased genetic distance) and the plot of the first two principal components (based on variance-covariance matrix of allele frequency at

polymorphic loci) distinguished three groups of regions. The level of gene flow was low among accessions, regions of origin and among accessions within adaptation zones, but high among adaptation zones. The results are discussed with emphasis on genetic resources conservation and utilization.

In conclusion, the sorghum accessions used in the present study maintained low levels of allozyme variation. The results indicate that altitudinal variation does not contribute substantially to gene differentiation among the sorghum accessions. All in all there is no compelling genetic (allozyme plus RAPD) evidence that the low, intermediate and highland sorghums are different, despite the wide range of differences for quantitative agro-morphological traits. Hence, they can be considered as the same gene pool in breeding varieties for the different adaptation zones.

polymorphic loci) distinguished three groups of regions. The level of gene flow was low among accessions, regions of origin and among accessions within adaptation zones, but high among adaptation zones. The results are discussed with emphasis on genetic resources conservation and utilization.

In conclusion, the sorghum accessions used in the present study maintained low levels of allozyme variation. The results indicate that altitudinal variation does not contribute substantially to gene differentiation among the sorghum accessions. All in all there is no compelling genetic (allozyme plus RAPD) evidence that the low, intermediate and highland sorghums are different, despite the wide range of differences for quantitative agro-morphological traits. Hence, they can be considered as the same gene pool in breeding varieties for the different adaptation zones.

Chapter 5

Random Amplified Polymorphic DNA (RAPD) Variation in Sorghum (*S. bicolor* ssp *bicolor* (L.) Moench)

Introduction

Information about genetic variation in cultivated plant species is relevant for plant breeding and for germplasm management (Allard, 1996; Grenier et al., 2000b). In chapter three, it was shown that sorghum germplasm collections from Ethiopia and Eritrea are highly variable for morphological and agronomic characteristics (also refer to Ayana and Bekele, 1998, 1999, 2000). Similar conclusions are also made previously based on some morphological and agronomic traits (Gebrekidan, 1973; Teshome et al., 1997). However, morphological variation does not reliably reflect the actual genetic variation because of genotype-environment interaction and the largely unknown genetic bases of polygenically inherited morphological and agronomic traits (Smith and Smith, 1992).

In contrast to the wide range of morphological variation, the allozyme analysis in chapter four revealed low level of variation among 48 sorghum accessions drawn from different ecological regions of Ethiopia and Eritrea (see also Ayana et al., 2000c), confirming previous conclusions that sorghum is characterized by a low level of allozyme variation (Morden et al., 1989; Aldrich et al., 1992; Djè et al., 1998, 1999). Low level of polymorphism and, thus, inadequate coverage of the genome, limit the use of allozymes as an adequate technique for genetic diversity analysis (Clegg, 1990; Karp et al., 1997).

This calls for use of DNA-based markers, which are generally more polymorphic and cover a larger portion of the genome than isozymes and which measure genetic variation at molecular level unlike morphological characters. Moreover, many studies have also demonstrated that no single method (morphological, biochemical or molecular) is adequate for assessing genetic variation in germplasm collections (Singh et al., 1991; Beer et al., 1993; Liu and Furnier, 1993; Zhang et al., 1993; Oliveira et al., 1996). The reason is that different methods sample genetic variation at different levels and, hence, differ in their power of genetic resolution as well as the quality of information. Because of their generally high polymorphism and high resolution, molecular markers should be considered only as complementary to morphological characterization and isozyme analysis.

In sorghum, as is true for other crop plants, efforts to estimate genetic variation have been greatly enhanced with the advent of molecular markers. These include restriction fragment length polymorphism (RFLP) (Aldrich and Doebley, 1992), random amplified polymorphic DNA (RAPD) (Tao et al., 1993; Oliveira et al., 1996; Menkir et al., 1997; Ayana et al., 2000a), microsatellites or simple sequence repeats (Brown et al., 1996; Dean et al., 1999; Djè et al., 1999, 2000; Grenier et al., 2000b), and inter-simple sequence repeats (Yang et al., 1996; Taramino et al., 1997). The results from these studies have demonstrated that molecular markers are providing additional and complementary information regarding the amount and distribution of genetic variation in germplasm collections and elite breeding lines, enhancing the efficiency of germplasm management and utilization. Of all the DNA-based techniques, the RAPD technique is preferable for initial assessment of genetic diversity in germplasm collections owing to its technical simplicity, speed, cost effectiveness and the wide availability of universal primers (Tao et al., 1993; Williams et al., 1993), although there are many other DNA-based markers with several advantages over RAPD markers. As with other DNA-based markers, advantages of RAPD

Table 5.1. Sample size of sorghum accessions used for the study by regions of origin and adaptation zones

Region	Adaptation zone			Total
	Lowland	Intermediate	Highland	
Bale	2	-	-	2
Eritrea	6	-	-	6
Gamo Gofa	2	2	2	6
Gojam	1	4	-	5
Gonder	2	2	2	6
Harerge	2	2	3	7
Illubabor	5	1	-	6
Introduction	4	-	-	4
Kafa	-	2	-	2
Shewa	-	3	3	6
Sidamo	-	3	2	5
Tigray	2	2	2	6
Variety	4	-	2	6
Welega	2	2	2	6
Welo	3	2	2	7
Total	35	25	20	80

DNA extraction

Ten individuals represented an accession for DNA extraction. Seed samples were germinated in greenhouse pots. From each of the 10 individuals, fresh leaves were harvested from 7-day old seedlings for DNA extraction. Total genomic DNA was extracted from 300 mg leaf samples according to the method of Junghans and Metzlaff (1990) with minor modifications.

The DNA quality was assessed by electrophoresis in 1% agarose gels and the concentration was estimated by fluorescence spectroscopy using Hoechst 33258 dye with calf thymus DNA as standard (Brunk et al., 1979). An equal quantity of DNA from the 10 individuals was pooled to a working concentration of 25 ng/ml in ddH_2O for each accession.

The frequency of amplified products generated with 20 oligonucleotide primers across 80 sorghum accessions is grouped into four classes and is given in Table 5.3. The result showed that 61% of the bands have intermediate frequency (0.26-0.75).

The percentage of polymorphic RAPD bands for the 15 regions of origin is presented in Table 5.4. The percentage of polymorphic bands ranged from 32% (for accessions from Kafa) to 91% (for accessions from Harerge) with a mean of 77%. In general, sorghum accessions from Harerge, Welo and Illubabor manifested the highest polymorphism.

Estimation of RAPD variation

Mean regional estimates of RAPD variation, computed from band frequencies using the Shannon-Weaver diversity index, is given in Table 5.4 (refer to Appendix 5.2 for details). Pooled over the 20 primers, the mean genetic diversity ranged from 0.21 for Kafa to 0.48 for Illubabor, with an overall variation of 0.53, showing an intermediate level of RAPD variation.

For the adaptation zones the mean estimates of the Shannon-Weaver diversity index, pooled over primers within adaptation zones, varied over a very narrow range (0.49-0.52), with an overall mean of 0.50 (Table 5.5). Apparently the adaptation zones did not differ from each other in their overall RAPD variability.

Table 5.2. Sequences, number of amplified products (bands) and approximate molecular size range in base pairs (bp) for 20 primers used to generate RAPD markers in 80 sorghum accessions

Primer	Sequence 5' to 3'	bands	Bp
OPA-01	CAGGCCCTTC	10	400-1900
OPA-02	TGCCGAGCTG	10	500-1600
OPA-03	AGTCAGCCAC	8	400-1800
OPA-04	AATCGGGCTG	7	600-1900
OPA-05	AGGGGTCTTG	6	400-2000
OPA-08	GTGACGTAGG	4	800-1800
OPA-10	GTGATCGCAG	8	700-1700
OPA-11	CAATCGCCGT	10	400-2400
OPA-13	CAGCACCCAC	6	900-2100
OPA-14	TCTGTGCTGG	6	400-1800
OPA-15	TTCCGAACCC	7	400-1900
OPA-16	AGCCAGCGAA	8	400-2000
OPA-17	GACCGCTTGT	10	500-2300
OPA-18	AGGTGACCGT	10	400-1600
OPA-19	CAAACGTCGG	7	300-1800
OPA-20	GTTGCGATCC	7	500-1700
OPC-01	TTCGAGCCAG	3	700-2000
OPC-02	GTGAGGCGTC	8	500-1500
OPC-05	GATGACCGCC	8	500-2000
OPC-19	GTTGCCAGCC	4	800-2300
Total		147	
Range		3-10	300-2400
Average		7.35	

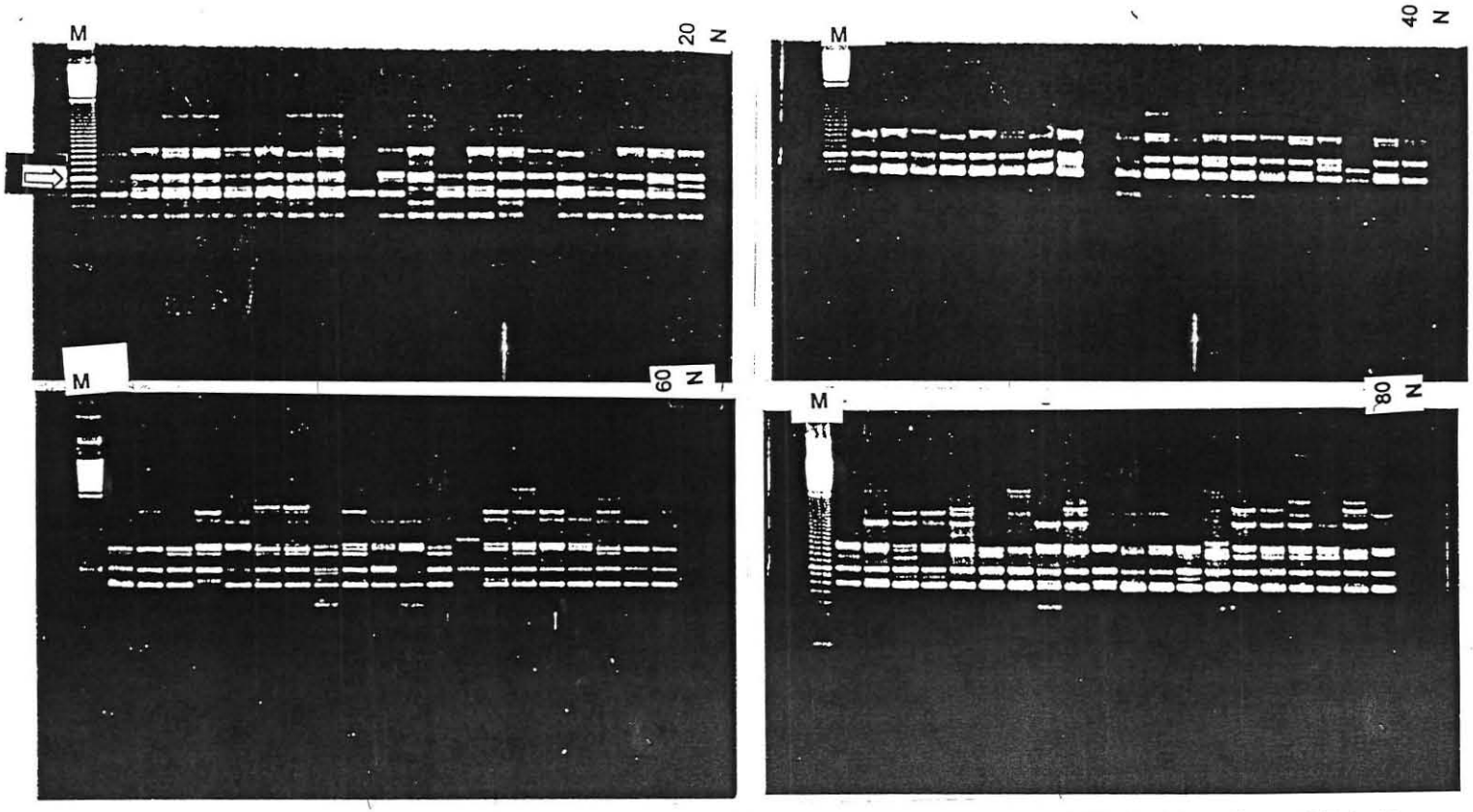


Fig. 5.1. Amplification products from genomic DNA of 80 sorghum accessions amplified with primer OPA-11.

The molecular weight marker (100 bp ladder) is shown in the left lane (M) and the arrow indicates 800 bp band. The control lane is designated as (N).

Other Lane numbers represent accessions listed in Appendix 5.1.

Table 5.3. Frequency of amplified products generated with 20 primers in 80 sorghum accessions

Frequency of RAPD bands	Number of bands	Proportion
0.01 - 0.25	37	0.25
0.26 - 0.50	38	0.26
0.51 - 0.75	52	0.35
0.76 - 0.99	20	0.14
Total	147	1.00

Table 5.4. Percentage of polymorphic bands (P), mean estimates of the Shannon-Weaver diversity index (Mean ± S.E) for 20 primers for 15 regions of origin of 80 sorghum accessions

Region	P	Mean ± S.E
Bale	48.35	0.28 ± 0.04
Eritrea	70.00	0.39 ± 0.03
Gamo Gofa	85.00	0.43 ± 0.03
Gojam	74.15	0.40 ± 0.04
Gonder	80.80	0.43 ± 0.03
Harerge	90.85	0.46 ± 0.03
Illubabor	89.20	0.48 ± 0.02
Introduction	80.00	0.43 ± 0.03
Kafa	31.65	0.21 ± 0.04
Shewa	82.50	0.44 ± 0.02
Sidamo	78.35	0.43 ± 0.03
Tigray	85.80	0.43 ± 0.03
Variety	77.50	0.44 ± 0.03
Welega	85.80	0.43 ± 0.04
Welo	90.00	0.46 ± 0.02
Entire data	76.61	0.53 ± 0.02

Table 5.5. Estimates of the Shannon-Weaver diversity index for three adaptation zones of 80 sorghum accessions

Primer	Adaptation zones		
	Lowland	Intermediate	Highland
OPA-01	0.44	0.32	0.38
OPA-02	0.60	0.61	0.52
OPA-03	0.58	0.56	0.53
OPA-04	0.64	0.66	0.63
OPA-05	0.58	0.31	0.43
OPA-08	0.50	0.58	0.62
OPA-10	0.58	0.45	0.51
OPA-11	0.43	0.30	0.34
OPA-13	0.54	0.43	0.45
OPA-14	0.63	0.63	0.57
OPA-15	0.51	0.66	0.61
OPA-16	0.58	0.46	0.55
OPA-17	0.53	0.38	0.44
OPA-18	0.53	0.51	0.51
OPA-19	0.50	0.45	0.48
OPA-20	0.55	0.59	0.48
OPC-01	0.19	0.35	0.42
OPC-02	0.64	0.57	0.55
OPC-05	0.44	0.44	0.53
OPC-09	0.43	0.50	0.47
Mean	0.52 ± 0.02	0.49 ± 0.03	0.50 ± 0.02

Partitioning of RAPD variation

A large portion of the total variation (77%) was found within the 15 regions of origin (Table 5.6). Similarly, partitioning of the variation into within and among adaptation zones revealed that 94% of the total variation was found within and the remaining 6% among the adaptation zones (Table 5.6), indicating very little differentiation among the adaptation zones.

Table 5.6. Partitioning of genetic variation into within and between regions of origin and adaptation zones of 80 sorghum accessions

Primer	Total	Region			Adaptation zones		
	H_{sp}	H_r	H_r/H_{sp}	$H_{sp}-H_r)/H_{sp}$	H_z	Hz/H_{sp}	$(H_{sp}-H_z)/H_{sp}$
OPA-01	0.41	0.30	0.74	0.25	0.38	0.93	0.07
OPA-02	0.61	0.47	0.77	0.23	0.58	0.95	0.05
OPA-03	0.58	0.46	0.79	0.21	0.57	0.96	0.04
OPA-04	0.65	0.51	0.79	0.21	0.64	0.98	0.02
OPA-05	0.50	0.37	0.74	0.26	0.44	0.88	0.12
OPA-08	0.57	0.48	0.84	0.16	0.57	1.00	0.00
OPA_10	0.54	0.43	0.79	0.20	0.51	0.94	0.06
OPA-11	0.40	0.29	0.73	0.27	0.37	0.90	0.10
OPA-13	0.51	0.40	0.78	0.22	0.47	0.92	0.08
OPA-14	0.66	0.51	0.77	0.23	0.61	0.92	0.08
OPA-15	0.60	0.50	0.83	0.16	0.59	0.98	0.02
OPA-16	0.56	0.45	0.80	0.19	0.53	0.95	0.05
OPA-17	0.49	0.36	0.74	0.26	0.45	0.92	0.08
OPA-18	0.53	0.41	0.77	0.22	0.52	0.98	0.02
OPA-19	0.51	0.39	0.76	0.23	0.48	0.94	0.06
OPA-20	0.58	0.44	0.76	0.24	0.54	0.93	0.07
OPC-01	0.31	0.20	0.63	0.36	0.32	1.00	0.00
OPC-02	0.63	0.48	0.76	0.24	0.59	0.94	0.06
OPC_05	0.44	0.33	0.75	0.25	0.40	0.91	0.09
OPC-19	0.48	0.38	0.75	0.25	0.47	0.92	0.08
Mean	0.53 ± 0.02	0.41± 0.02	0.77 ± 0.01	0.23 ± 0.01	0.50 ± 0.02	0.94 ± 0.01	0.06 ± 0.01

H_{sp} = Genetic variation for each primer calculated from the entire data set; H_r and H_z = Mean genetic variation index for each primer for the regions of origin and the adaptation zones, respectively; H_r/H_{sp} and H_z/H_{sp} = Proportion of genetic variation within regions of origin and within adaptation zones, respectively; $(H_{sp}-H_r)/H_{sp}$ and $(H_{sp}-H_z)/H_{sp}$ = Proportion of genetic variation within regions of origin and within adaptation zones, respectively, in relation to the total genetic variation.

Cluster analysis

Cluster analysis of the 80 sorghum accessions failed to show close relationships among accessions of the same region or adaptation zone (Appendix 5.3). Instead, the dendrogram clearly showed three groups at 48% similarity, each including accessions from a wide range of regions of origin and adaptation zones. The coefficient of genetic similarity ranged from 32.7% (between a landrace accession from Welega and an improved variety known as Bakomash) to 92.5% (between two landrace accessions from Shewa and Tigray) with a mean genetic similarity of 64%.

Cluster analysis based on regional frequencies of amplified bands showed that accessions from Gojam and Sidamo; Gamo Gofa, Gonder and Illubabor; Shewa and Welega; Tigray, Welo and Eritrea are closely related (Fig. 5.2). The coefficient of dissimilarity among the 15 regions of origin ranged from 6.6% (between Sidamo and Welega) to 34.3% (between Bale and a set of improved varieties) with an average of 13.3% (Table 5.7).

Table 5.7. Genetic dissimilarity among 15 regions of origin of 80 sorghum accessions

Region	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Bale	0.00														
Eritrea	0.29	0.00													
Gamo Gofa	0.14	0.17	0.00												
Gojam	0.20	0.20	0.10	0.00											
Gonder	0.21	0.16	0.08	0.12	0.00										
Harerge	0.22	0.16	0.10	0.12	0.15	0.00									
Illubabor	0.17	0.18	0.07	0.10	0.07	0.09	0.00								
Introduction	0.25	0.16	0.12	0.17	0.14	0.11	0.14	0.00							
Kafa	0.28	0.34	0.18	0.17	0.14	0.28	0.16	0.25	0.00						
Shewa	0.24	0.17	0.09	0.15	0.16	0.10	0.12	0.15	0.27	0.00					
Sidamo	0.20	0.19	0.07	0.07	0.12	0.10	0.09	0.15	0.16	0.10	0.00				
Tigray	0.25	0.12	0.12	0.18	0.15	0.10	0.12	0.13	0.32	0.13	0.16	0.00			
Variety	0.34	0.15	0.18	0.17	0.18	0.13	0.15	0.10	0.23	0.19	0.16	0.15	0.00		
Welega	0.22	0.16	0.07	0.07	0.11	0.07	0.08	0.13	0.18	0.09	0.06	0.12	0.13	0.00	
Welo	0.26	0.10	0.10	0.18	0.12	0.10	0.10	11.00	0.27	0.10	0.14	0.07	0.15	0.10	0.00

Discussion

The mean number of bands per primer obtained in the present study is high compared to most of the previous studies in sorghum using RAPDs (Vierling et al., 1994; Oliveira et al., 1996; Menkir et al., 1997) and is comparable with that reported by Tao et al. (1993). The observed level of RAPD polymorphism is also higher than those reported (25-77%) in previous studies (Tao et al., 1993; Vierling et al., 1994; Yang et al., 1996; Menkir et al., 1997). The higher level of polymorphism and average number of bands per primer may partly be due to our sampling technique, which took into account not only differences in geographical range and adaptation zones but also results of morphological variation (Ayana and Bekele, 1999). Sampling germplasm collections across diverse environments and based on morphological variation has been considered to be a more effective means for capturing genetic diversity (Bogyo et al., 1980). Additional explanations may be differences in DNA extraction protocols and bias in scoring bands of different intensity. Our protocol involved RNase treatment. RNase-treatment and additional phenol/chloroform/isoamyl (25:24:1) extraction yields purer DNA than methods of DNA extraction that do not involve RNase-treatment. When the same concentration of primer is used, amplification is more efficient when RNase-treated DNA is used as a template compared to the untreated DNA template (Williams et al., 1993; Pammi et al., 1994). These authors added that constitution of amplification mix, thermal cycling, sensitivity of gel staining system and purity of DNA also may affect the number of amplification products.

The fact that sorghum accessions from Harerge, Welo and Illubabor manifested the highest RAPD polymorphism is congruent with previous studies (Mengesha, 1975; Ayana and Bekele, 1998). These authors reported that sorghum collections from these regions are

highly variable for morphological and agronomic characteristics. However, the level of variation observed for RAPDs in this study is lower than the one observed for qualitative characters (Ayana and Bekele, 1998) but more than that observed for allozymes (Ayana et al., 2000c; refer also to chapter four in this thesis). The results imply that Ethiopian and Eritrean sorghum germplasm has a lower level of variation for RAPD markers than for morphological characters (ca. overall mean of H = 0.90) and still lower for allozymes than for RAPD.

The fact that in the present study much more RAPD variation is found among accessions within regions of origin than among regions of origin suggests weak regional differentiation of the sorghum germplasm. The result is consistent with previous studies using qualitative morphological characters (Ayana and Bekele, 1998) and allozyme analysis (Ayana et al., 2000c),), in which 86% and 80% of the total variation was found within regions, respectively. Several studies have also reported limited regional differentiation for the world collection of sorghum using allozyme data (Morden et al., 1989), RFLP (Aldrich and Doebley, 1992; Deu et al., 1994; Cui et al., 1995) and RAPD data (Tao et al., 1993; Menkir et al., 1997).

Similarly, much more RAPD variation is found among accessions within the adaptation zones than among the adaptation zones. This result is also in agreement with previous studies using qualitative morphological characters (Ayana and Bekele, 1998) and allozyme analysis (Ayana et al., 2000c), in which 96% and 97% of the total variation was found within adaptation zones, respectively. The classification of Ethiopian sorghum into different adaptation zones (Gebrekidan, 1981) has been based mainly on altitude, which determines length of the growing period. It is concluded that sorghums of the three adaptation zones do not seem to possess significant genetic differences but differ only in their adaptation to agro-ecological zones, which differ in length of growing period.

Since much of the RAPD variation is found within regions of origin and adaptation zones, it is tempting to suggest a germplasm sampling strategy in which many accessions from a few regions or adaptation zones may capture a large portion of the genetic variation. However, sampling from as many regions as possible, with a wide range of ecological conditions, is still recommended because significant variation has been observed for quantitative characters that have adaptive significance (Ayana and Bekele, 1999). In addition, rare types can also be captured. Variation in RAPD patterns may not necessarily have adaptive significance as both coding and non-coding regions of the genome are accessed (Williams et al., 1993).

The result from the pair-wise similarity (dissimilarity) analysis suggests medium RAPD variation among the 80 sorghum accessions studied. Both the range and average similarity observed in the present study are very similar to those reported by other investigators for sorghum using RAPD, RFLP and simple sequence repeats (SSR) markers either singly or combined (Deu et al., 1994; Oliveira et al., 1996; Yang et al. 1996).

The results from the present study also showed that limited variation exists among the regions of origin, almost twice lower than what was found among the accessions as a whole. The overall close similarity of accessions from Gamo Gofa and Gonder as well as from Shewa and Welega concurred with results obtained using multivariate analysis of morphological variation of quantitative characters (Ayana and Bekele, 1999). Deu et al. (1994) also found congruence between morphological classification and classification based on RFLP for sorghum. These authors have hypothesized that such correspondence could arise from multi-trait associations, which in turn result from founder effect during domestication and restricted recombination after domestication.

Cluster analysis revealed that patterns of variation are not congruent with regions of origin or adaptation zones. The fact that a large portion of the variation is found among

accessions within region of origin and within the adaptation zones also confirms that differentiation of the sorghum material on the basis of region of origin and adaptation zone is weak. Results obtained using analysis of morphological and agronomic characters (Ayana and Bekele, 1998) also showed weak regional and agro-ecological zone differentiation. The low levels of genetic differentiation among sorghum accessions sampled from different regions and adaptation zones may be due to: (1) the high rate of outcrossing (5-30%) in cultivated sorghum (Ollitrault, 1987 cited in Deu et al., 1994), and, (2) hybridization of cultivated sorghum with its wild and weedy relatives (Aldrich et al., 1992; Harlan, 1992).

Although genetic differentiation is low both on regional and adaptation zone bases, regional differentiation is greater than differentiation on the basis of adaptation zone. Ayana and Bekele (1998) also reported similar results using qualitative morphological data. However, applying discriminative analysis to quantitative morphological and agronomic data, they found that adaptation zone is more discriminative than regions of origin for 415 sorghum accessions (Ayana and Bekele, 1999). Apparently regions of origin become more important in differentiating sorghum accessions when qualitative morphological characters and RAPD markers are used, whereas adaptation zones do when quantitative morphological and agronomic characters are used. This may be due to the fact that quantitative characters like days to flowering and maturity and plant height are important in determining sorghum adaptation to a particular agro-ecological zone (Doggett, 1988). On the other hand, some of the qualitative characters and RAPD markers, having no or less adaptive significance, tend to be less discriminative on the basis of adaptation zones. The significance of region of origin for determining the relatedness or differences among sorghum accessions has been documented in several studies (Morden et al., 1989; Tao et al., 1993; Deu et al., 1994; Oliveira et al., 1996).

The results reported here and those of others (Tao et al., 1993; Vierling et al., 1994; Yang et al., 1996; Menkir et al., 1997) have shown that RAPD markers are useful for measuring the extent and distribution of genetic variation in sorghum. However, there are some drawbacks associated with the use of RAPD markers for studying genetic variation. The dominant inheritance of RAPD markers (Williams et al., 1990, 1993) makes them less informative for estimating population genetic parameters than co-dominant markers such as isozymes and RFLPs. Since it is not possible to distinguish between heterozygote and homozygote genotypes, it is impossible to calculate allelic frequencies for individual loci. Consequently, the fraction of shared RAPD bands between two genotypes is used as the measure of similarity between accessions. But the fraction of shared bands may not be equivalent to the fraction of shared genes (Lynch, 1988).

The solution to the dominant inheritance of RAPD markers is to analyze a great number of fragments by using many primers and thereby increase the statistical power (Lu et al., 1995). Nei (1978) stated the advantages of using a large number of loci for estimation of genetic distance and recommended that at least 50 loci should be used. In the present study we used 147 RAPD bands (putative loci), which is almost three times the recommended number. Therefore, it is likely that the values of genetic similarity obtained and the dendrograms constructed from them are reliable indications of the relationship between the sorghum accessions and their regions of origin.

Whether RAPD bands of similar size (i.e., co-migrating bands) are homologous remains to be verified by hybridization experiments. This may be a serious problem when comparison is made at higher taxonomic levels (Newbury and Ford-Lloyd, 1993; Rieseberg, 1996). However, it is likely that shared RAPD bands between different genotypes belonging to a lower taxonomic group (i.e., below species level) reflects sequence homology (Kazan et al., 1993; Tingey and del Tufa, 1993; Williams et al., 1993). Since our study is based on

comparison of accessions of *Sorghum bicolor* ssp *bicolor*, it is thus presumed that shared bands are homologous.

The molecular basis of the bands observed in a RAPD pattern is not yet clearly understood. However, it has been suggested that RAPD polymorphism could arise because of nucleotide changes at a primer binding site as well as from deletions or insertions that change the size of the amplified product or render the priming sites too distant to support amplification (Williams et al., 1990, 1993). As a result of this lack of genetic basis of RAPD bands, analysis and interpretation are commonly limited to similarity and distance comparisons between individuals rather than genetic analysis, unlike the results from isozymes. Certainty whether individual RAPD fragments represent independent characters is also lacking. Instead, they may be linked or allelic to each other (Lynch, 1988). However, Tinker et al. (1993) provided evidence for non-allelism of individual RAPD bands and it is considered valid to treat each RAPD band as an independent locus carrying two alleles, either presence or absence of the band (Lynch and Milligan, 1994; Nevo et al., 1998).

Additional limitations of RAPDs include low level of reliability and poor reproducibility because of PCR sensitivity to minor changes in reaction conditions and competition for primer banding sites (Newbury and Ford-Lloyd, 1993; Skroch and Nienhuis, 1995). Rigorous standardization of the reaction conditions and care may reduce the number of unreliable amplification products and increase reproducibility (Pammi et al., 1994; Skroch et al., 1998). Our experience also supports this fact.

In conclusion, a large number of polymorphic RAPD bands are generated by each primer among the 80 sorghum accessions. Hence RAPDs are very useful for estimating genetic variation and investigating relationships among sorghum germplasm collections. Information about genetic similarity (distance) between germplasm collections would be helpful for plant breeders to choose diverse parents for crossing.

Chpter 6

Random Amplified Polymorphic DNA (RAPD) Variation in Wild Sorghum (*S. bicolor* ssp. *verticilliflorum* (L.) MOENCH) from Ethiopia

Introduction

Wild sorghum (*Sorghum bicolor* ssp. *verticilliflorum* (L.) Moench) is the progenitor of cultivated sorghum (*S. bicolor* ssp. *bicolor* (L.) Moench) (Doggett, 1988; Wiersema and León, 1999). *S. bicolor* ssp. *verticilliflorum* includes four races, namely *aethiopicum*, *arundinaceum*, *verticilliflorum* and *virgatum*. They differ from each other mainly in inflorescence morphology and distribution and they are fully inter-fertile with the cultivated sorghum (de Wet et al., 1970; Doggett, 1991). The hybrid derivative is known as shattercane (*S. bicolor* ssp. *drummondii*), which is a weedy type of sorghum. Shattercane, known as '*keelo*' or '*seepo*'(literally means the fool in local languages), are weeds in cultivated sorghum fields in Ethiopia (Damon, 1962; Doggett, 1991; Harlan, 1992). Doggett (1965a, b) and Doggett and Majisu (1968) demonstrated prevalence of extensive genetic interaction between wild, weedy and cultivated races of sorghum in Africa wherever they are sympatric. Being in the primary gene pool with cultivated sorghum, these wild and weedy relatives could be important sources of genes for sorghum improvement and are also used as forage for animal feed (Wiersema and León, 1999). Wild and weedy relatives of crop plants have been known to be sources of useful genes, such as genes for resistance to diseases and pests as well as genes for tolerance to different stress conditions like drought and cold (Hoyt, 1988; Tanksley and McCouch, 1997; Ellis et al., 2000). A constant flow of genes creates new

diversity in the cultivated, weedy and wild types (Doggett, 1965b, 1988; Doggett and Majsu, 1968; Harlan, 1992). Indeed, sorghum is one of the crop plants that has extensive genetic exchange with its wild and weedy relatives.

In terms of natural distribution, wild sorghum is strictly African (de Wet and Harlan, 1971; Doggett, 1991) and continues to survive in the wild until today across many African countries, including Ethiopia (Damon, 1962; Gebrekidan, 1970; Doggett, 1988, 1991), although naturalized forms of wild sorghum can be found in the Americas, Australia, and India (Wiersema and León, 1999). As is the case elsewhere, however, wild sorghum in Ethiopia is rapidly disappearing, largely because more and more of its natural habitat is used for agriculture, grazing, human settlement, construction, etc. Old literature indicates the presence of wild sorghum in places where none or few currently exist (Damon, 1962; Gebrekidan, 1970). Thus, it is prudent to collect and conserve the rapidly disappearing wild sorghum germplasm in Ethiopia before it is lost forever.

Information about the extent and distribution of genetic variation within a species is an important requirement for efficient collection and conservation of plant germplasm as well as for utilization of such a germplasm in crop improvement (Bretting and Widrlecher, 1995). The present study reports the extent and distribution of genetic variation among eleven populations of wild sorghum collected from different geographical regions of Ethiopia using random amplified polymorphic DNA (RAPD) (Williams et al., 1990). RAPD markers have been used successfully for characterization of genetic variation in germplasm collections of sorghum (Menkir et al., 1997; Ayana et al., 2000a).

Materials and methods

Plant material

Eleven populations of wild sorghum (*Sorghum bicolor* ssp. *verticilliflorum*) were collected from five regions of Ethiopia (Fig. 6.1) during the period October-December 1998. For each population, 10 to 15 panicles were collected from independent plants (i.e., excluding tillers) across a transverse path throughout the selected site within a region. During the collecting trip it was observed that the size of the standing population is quite variable. It was high (> 100 plants) per unit area of sampling (4 m x 4 m) for the Ghibe populations, moderate (50-100 plants) for the Gambella populations and low for the Assosa, Gamo Gofa and Pawe populations (< 50 plants).

RAPD analysis

A total of 93 individual plants representing 11 populations were assayed. The number of individuals per population varied from 4 to 10 (Table 6.1). For each individual, approximately 10 seeds were germinated in plastic pots placed on the laboratory tables. Total genomic DNA was extracted from about 300 mg fresh leaves harvested from 7-10 day old seedlings according to the method of Junghans and Metzlaff (1990), with some modifications (Ayana et al., 2000a).

Nine decamer primers (Operon Technology, Ca., USA), which gave high and reproducible bands in a previous study of cultivated sorghum (Ayana et al., 2000a; refer also to Table 5.2 in this thesis), were used (Table 6.2).

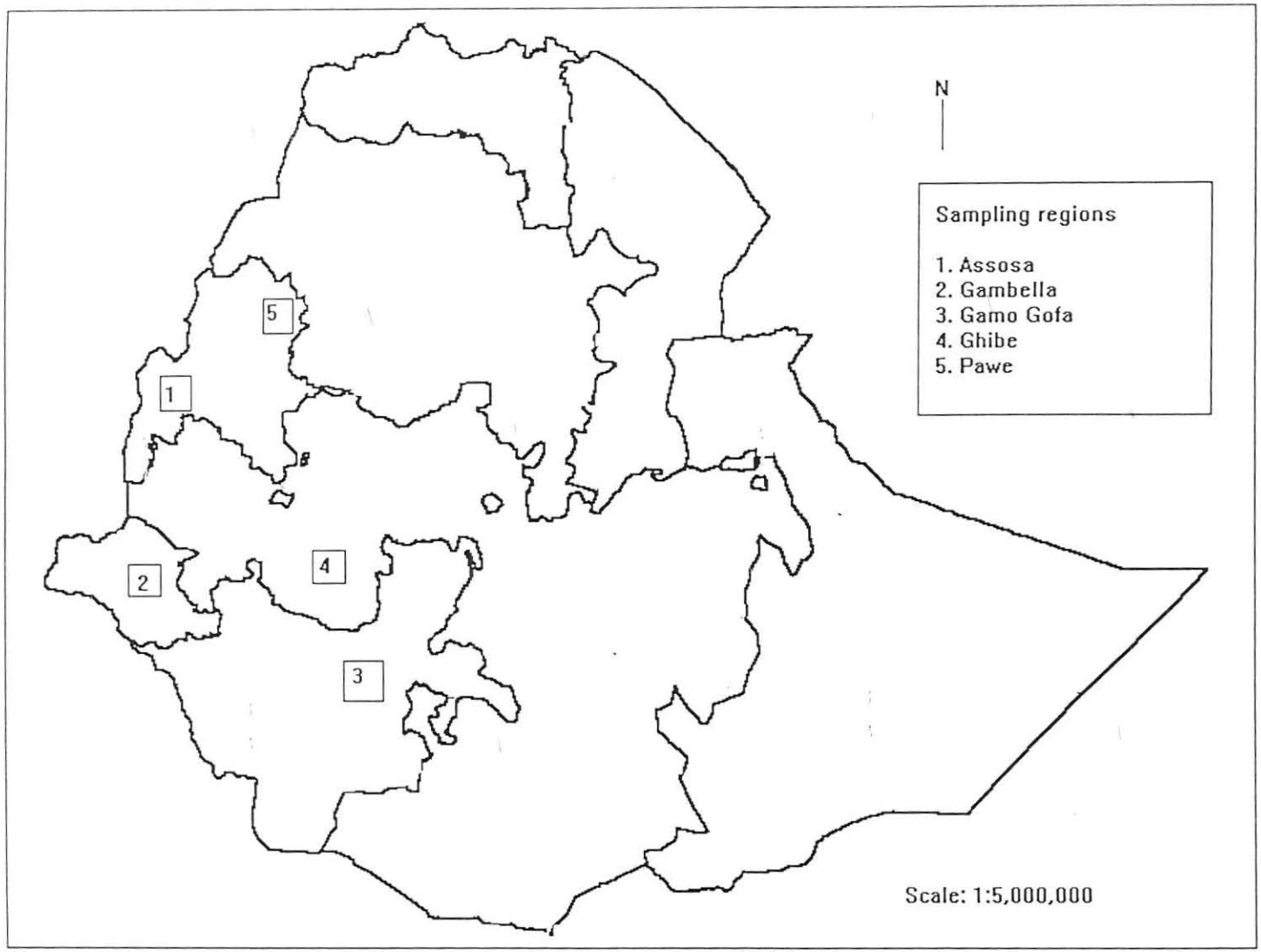


Fig 6.1. Map of Ethiopia showing the geographical regions from where the wild sorghum populations are collected.

Table 6.1. List of wild sorghum *(S. bicolor* ssp. *verticilliflorum)* populations used for the study

Population designation	Sample size	Geographical origin	Locality	Altitude (m)
Assosa	4	Assosa	Banibis river bank. 43 km away from Assosa town on the way to Banbis Settlement Site	1050
Gambella-1	9	Gambella	Chebo Ukuna Settlement Site. 27 km away from Gambella town on the way to Abobo town	630
Gambella-2	9	Gambella	Abobo area. 54 km away from Gambella town on the way to Abobo town	605
Gambella-3	9	Gambella	Mazoria. Itang-Larae junction. 46 km away from Gambella town on the way to Itang town	570
Gamo Gofa-1	10	Gamo Gofa	Silie State Cotton Farm. 18 km away from Arba Minich town on the way to Konso town	1000
Gamo Gofa-2	10	Gamo Gofa	Chano Mili. 15 km away from Araba Minich town on the way to Sodo town	1020
Gamo Gofa-3	9	Gamo Gofa	Yayika. 58 km away from Araba Minich town on the way to Sodo town	1010
Ghibe-1	9	Ghibe	On the descending slope to Ghibe river bridge. 26 km from Walkite town to the bridge	1385
Ghibe-2	9	Ghibe	On stream bank after crossing the Ghibe river bridge. 1 km away from the bridge on the way to Jimma town	1185
Ghibe-3	8	Ghibe	On ascending slope from Ghibe river bridge. 7 km away from the bridge on the way to Jimma town	1490
Pawe	7	Pawe	Around Pawe abandoned Air Port	1240
	93			

The PCR amplification reactions were performed in a final volume of 20 µl, consisting of 1x PCR reaction buffer (Sigma), 0.20 µM primer, 100 µM each of dATP, dCTP, dGTP and dTTP, 2 mM $MgCl_2$, 0.5 units of Taq polymerase and about 25 ng of genomic DNA. The amplifications were performed using a Genius Thermocycler (Techne Ltd, Cambridge), with hot lid accessory. The amplifications were programmed for 3 min at 94 °C for initial denaturation, followed by 45 cycles of 1 min at 94 °C; 1 min at 37 °C and 2 min at 72 °C, using the fastest possible transition times between each temperature. A final extension was programmed for 10 minutes at 72 °C and this was followed by hold time at 4 °C until samples were collected.

Amplification products were resolved by gel electrophoresis in 1.2% agarose gels run in 1x TAE (10 mM Tris-HCl and 1 mM EDTA) buffer, pH 8.0, for 2½ hr at 90 voltage. The gel was stained with ethidium bromide (10 mg/ml) and the DNA fragments (bands) were detected by UV trans-illumination and photographed under UV using Polaroid 667 films.

Each RAPD band was considered as an independent character or locus and assigned numbers in order of decreasing molecular weight. The size of each band was estimated against the DNA molecular weight marker (100 bp ladder). A band was scored as present (1) or absent (0).

The amount of genetic variation was determined using the Shannon-Weaver diversity index. The Shannon-Weaver diversity index as described by Hutchenson (1970) is given as:

$$H = -\sum p_i \ln p_i$$

where p_i is the proportion of amplified and unamplified bands among individuals of a population or among populations in a region of origin. The proportion of variation attributable to differences between and within populations as well as between and within regions of origin were estimated from the Shannon-Weaver diversity index following the method of King and

Schaal (1989) as used in several RAPD data (Wachira et al., 1995; Menkir et al., 1997; Ayana et al., 2000a).

Pair-wise genetic similarity matrix was generated among the 93 individuals using the Jaccard's similarity coefficient (Jaccard, 1908 cited in Mumm and Dudley, 1995). The Jaccard similarity coefficient is given as

$$S_{ij} = \frac{a}{a+b+c}$$

where 'a' is the total number of bands shared between individuals 'i' and 'j'; 'b' is the total number of bands present in individual 'i' but not in individual 'j'; and 'c' is the total number of bands present in individual 'j' but not in individual 'i'. The similarity matrix was then used for cluster analysis.

Following the assumption of Lynch and Milligan (1994) that there are two alleles per locus when analyzing genetic population structure from RAPD data, we computed the proportion of amplified bands (dominant alleles) and non-amplified ones (recessive alleles) among individuals within every population as well as among populations within every region of origin. Then pair-wise genetic distances among populations and regions of origin were computed using Nei's (1972) coefficient of genetic distance from the corresponding allele frequency data. The formula for genetic distances (Nei, 1972) among populations and regions of origin is given as:

$$D = \log_e\left(\frac{J_{XY}}{J_X J_Y}\right)$$

where J_{XY}, J_X and J_Y are the arithmetic means of j_{XY}, j_X and j_Y, respectively, and $j_{xy} = \sum x_i y_i$; $j_x = \sum x_i^2$ and $j_Y = \sum y_i^2$ where x_i and y_i represent the frequencies of the i^{th} allele in populations X and Y.

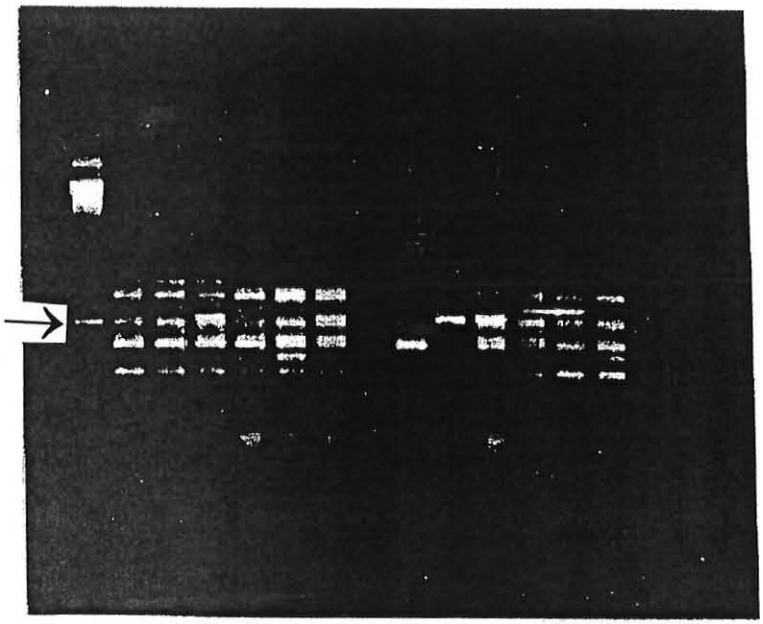


Fig. 6.2. PCR amplification products for genomic DNA of 14 sorghum samples using OPA-11 primer.

From left, lane 1 molecular marker (bold band is 800 bp ladder), lanes 2-4 Gamo Gofa, lanes 5-7 Gambella, lanes 8-9 Ghibe, lanes 10-11 Assosa, lanes 12-14 Pawe, and lane 15 negative control.

Table 6.2. Sequences, number of amplified products (bands) and approximate molecular size range (bp) for 9 primers used to generate RAPD markers in 93 wild sorghum individuals

Primer	Sequence (5' to 3')	Bands	bp
OPA-01	CAGGCCCTTC	11	400-2000
OPA-02	TGCCGAGCTG	12	400-2400
OPA-03	AGTCAGCCAC	9	400-1900
OPA-05	AGGGGTCTTG	9	400-2300
OPA-10	GTGATCGCAG	8	600-1500
OPA-11	CAATCGCCGT	8	400-2300
OPA-17	GACCGCTTGT	9	400-1900
OPA-18	AGGTGACCGT	8	400-2000
OPA-19	CAAACGTCGG	9	400-1900
Total		83	
Mean		9	

Table 6.3. Frequency of amplified products generated with 9 primers in 93 wild sorghum individuals

Frequency	Number of bands	Proportion
0.00-0.25	52	0.63
0.26-0.50	23	0.28
0.51-0.75	5	0.06
0.76-1.00	3	0.04
Total	83	1.00

Estimation of genetic variation

The extent of genetic variation for the 11 populations was determined from band frequencies using the Shannon-Weaver diversity index (Table 6.5). Pooled over the 9 primers, the mean genetic variation varied from 0.20 for the Ghibe-3 populations to 0.47 for the Ghibe-2 and the Pawe populations, with an overall variation of 0.49 for the entire data. The estimate of genetic variation, computed for the 5 regions of origin, is also given in Table 6.5. The mean for genetic variation varied from 0.30 for Assosa to 0.50 for Gambella.

Table 6.4. Number and percentage of polymorphic bands for 11 populations of wild sorghum and for their five regions of origin

Category		Number of Polymorphic bands	% Polymorphic bands
Population	Assosa	40	48
	Gambella-1	53	64
	Gambella-2	67	81
	Gambella-3	43	52
	Gamo Gofa-1	60	72
	Gamo Gofa-2	72	87
	Gamo Gofa-3	49	59
	Ghibe-1	56	67
	Ghibe-2	72	87
	Ghibe-3	32	39
	Pawe	67	81
Mean		56	67
Entire data		83	100
Regions	Assosa	40	48
	Gammbella	80	96
	Gamo Gofa	83	100
	Ghibe	77	93
	Pawe	67	81
Mean		69	84

Partitioning of genetic variation

Partitioning of the genetic variation into between and within populations revealed that 75% of the variation was found within the population, while the remaining 25% was found between populations (Table 6.6). Similarly, 88% of the genetic variation was found within the regions of origin of the populations, while only 12% was found between the regions of origin of the populations (Table 6.6).

Table 6.6. Partitioning of genetic variation into within and between populations as well as within and between regions of origin of the wild sorghum populations

	Population				Region		
Primer	H_{Total}	H_{Pop}	H_{Pop}/H_{Total}	$(H_{Total}-H_{Pop})/H_{Total}$	H_r	Hr/H_{Total}	$(H_{Total}-H_r)/H_{Total}$
OPA-01	0.50	0.35	0.69	0.31	0.40	0.80	0.20
OPA-02	0.47	0.38	0.80	0.20	0.45	0.96	0.04
OPA-03	0.52	0.40	0.76	0.24	0.42	0.81	0.19
OPA-05	0.46	0.33	0.72	0.28	0.48	0.96	0.04
OPA-10	0.44	0.32	0.73	0.27	0.35	0.80	0.20
OPA-11	0.65	0.51	0.78	0.22	0.49	0.75	0.25
OPA-17	0.41	0.28	0.69	0.31	0.42	0.98	0.02
OPA-18	0.54	0.44	0.82	0.18	0.50	0.93	0.07
OPA-19	0.46	0.36	0.78	0.22	0.48	0.96	0.04
Mean	0.49 ± 0.02	0.37 ± 0.02	0.75 ± 0.01	0.25 ± 0.01	0.44 ± 0.01	0.88 ± 0.01	0.12 ± 0.01

H_{Total} = Mean for total genetic variation computed from the entire data set; H_{Pop} and H_r = Mean genetic variation for the populations and regions, respectively; H_{Pop}/H_{Total} and H_r/H_{Total} = Proportion of genetic variation within the populations and regions, respectively; $(H_{Total}-H_{Pop})/H_{Total}$ and $(H_{Total}-H_r)/H_{Total}$ = Proportion of genetic variation between the populations and regions, respectively.

Cluster analysis

Cluster analysis of the 93 wild sorghum individuals failed to completely group the individuals either on population basis or regions of origin, although many individuals of the same population or region formed close similarity at different points of the dendrogram (Appendix 6.1). Pair-wise genetic distances among the 11 populations is presented in Table 6.7. The lowest genetic distance (0.03) was found between Gambella-3 and Gamo Gofa-3, while the highest coefficient of genetic distance (0.13) was found between Gamo Gofa-1 and Gamo Gofa-2. The average genetic distance was 0.0789, indicating the existence of a low genetic variation among the populations. Likewise, the average genetic distance for the regions of origin was 0.04 (Table 6.8). The lowest genetic distance (0.02) was found between Gambella and Ghibe, while the highest (0.10) was found between Assosa and Pawe, with an average of 0.03.

Cluster analysis of the 11 populations revealed four groups at a half way cut-point (Fig. 6.3). While the Assosa population formed a separate group at this cut-point, the populations of Gambella, Gamo Gofa and Pawe failed to be differentiated. Similarly, clustering of the regions of origin showed two groups at a half way cut-point (Fig. 6.4). Still Assosa appeared to be distinct, while the other four regions were in the same group. A close observation of this second group revealed that the Gambella, Ghibe and Gamo Gofa populations were closer to each other than either was to the Pawe population.

Table 6.7. Genetic distances among 11 populations of wild sorghum

Populations	1	2	3	4	5	6	7	8	9	10	11
1. Assosa											
2. Gambella-1	0.10										
3. Gambella-2	0.11	0.10									
4. Gambella-3	0.12	0.11	0.11								
5. Gamo Gofa-1	0.09	0.05	0.11	0.11							
6. Gamo Gofa-2	0.13	0.13	0.06	0.10	0.13						
7. Gamo Gofa-3	0.10	0.11	0.12	0.03	0.10	0.11					
8. Ghibe-1	0.10	0.05	0.10	0.11	0.04	0.12	0.11				
9. Ghibe-2	0.10	0.08	0.04	0.10	0.10	0.05	0.10	0.08			
10. Ghibe-3	0.11	0.12	0.12	0.03	0.11	0.10	0.04	0.12	0.10		
11. Pawe	0.10	0.11	0.09	0.10	0.11	0.07	0.09	0.10	0.07	0.10	

Table 6.8. Genetic distances among the five regions of origin of wild sorghum populations

Regions	1	2	3	4	5
1. Assosa					
2. Gambella	0.09				
3. Gamo Gofa	0.08	0.02			
4. Ghibe	0.08	0.02	0.02		
5. Pawe	0.10	0.08	0.07	0.07	

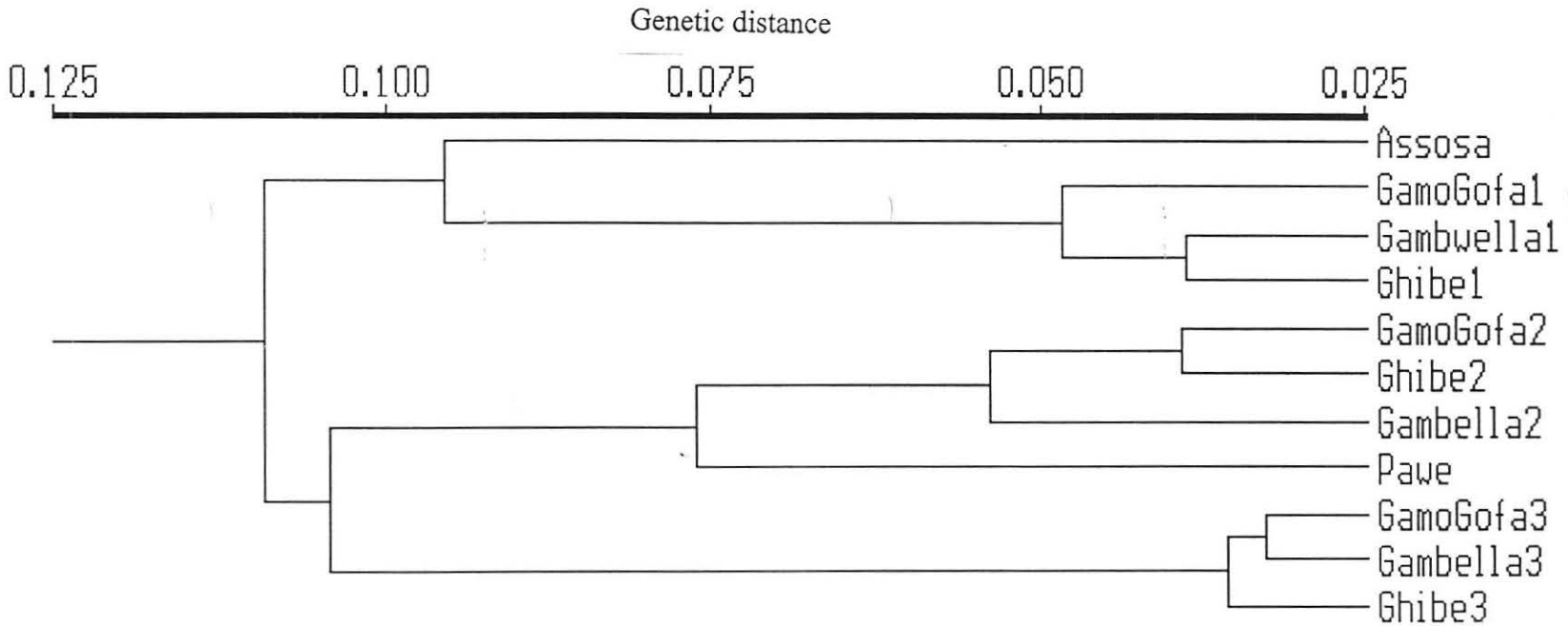


Fig. 6.3. Dendrogram for 11 populations of wild sorghum collected from five geographical regions in Ethiopia.

Discussion

We have analyzed RAPD variation in 93 wild sorghum individuals representing 11 populations from five geographical regions of Ethiopia to determine the extent and distribution of genetic variation. The number of polymorphic bands generated per primer in the present study was high and similar results are obtained for cultivated sorghum (Ayana et al., 2000a). Based on our observation during the collecting trip, we witness that the population size of the wild sorghum in Ethiopia is highly reduced because of rapid natural habitat change for human settlement, agriculture, construction, etc. The high degree of RAPD polymorphic bands observed in the present study is appreciable compared with the highly reduced population size in all the five regions. It is also our observation that human population is rapidly increasing in these regions with concomitant disappearance of wild sorghum from these regions. Construction works, settlement projects and large farms are rapidly expanding in the regions, particularly in Gambella and Ghibe, where relatively good stand of wild sorghum is observed at the moment. This calls for strong need to collect and conserve this invaluable genetic resource before it is lost forever. In several of his works, Doggett pointed out that the Gamo Gofa and the Gambella regions of Ethiopia are the probable places where sorghum was domesticated (Doggett, 1991; Doggett and Prasada Rao, 1995). The fact that the populations from these two regions showed higher percentage of polymorphic bands and higher genetic variation than populations of the other regions, did not only render support to Doggett's and others long standing argument (Doggett, 1976, 1991; Prasada Rao and Mengesha, 1981; Doggett and Prasada Rao, 1995) but also forces us to focus on these regions for future germplasm collecting missions.

reported by Ayana et al. (2000a, c) using RAPD and allozyme data and by Ayana and Bekele (1998) using qualitative morphological characters.

Although both *ex situ* and *in situ* methods of germplasm conservation are possible and complementary to each other, at this perilous time, conservation of wild sorghum germplasm from Ethiopia in its natural habitat seems impractical as the natural habitat is rapidly changing under various human activities. Moreover, existing socio-economic conditions do not seem to allow allocating vast acreage for conserving wild sorghum in its natural habitat. Therefore, we strongly recommend urgent rescue collection and *in situ* conservation. In this regard the Institute of Biodiversity Conservation and Research, Ethiopia, which is mandated for germplasm collection and conservation, should be heightened for the collection and *ex situ* conservation. Eventually it might be possible to extend to *in situ* conservation i.e., in the form of a botanical garden. Abobo Agricultural Research Centre, which is situated in Gambella, could play a significant role by allocating area for the establishment and maintenance of the botanical garden.

Summary

The extent and distribution of genetic variation in wild sorghum (*S. bicolor* ssp. *verticilliflorum* (L.) Moench) collected from five different geographical regions in Ethiopia were analyzed using random amplified polymorphic DNA (RAPD) markers for 93 genotypes representing 11 populations. Nine decamer primers generated a total of 83 polymorphic bands, across the 93 genotypes, with 8-12 bands per primer and an average of 9 bands. The amount of genetic variation within the populations ($\bar{H} = 0.37$) and within the geographical

region ($\bar{H} = 0.44$) was low to moderate, despite the high degree of polymorphic bands per primer. Similarly, the average genetic distance (0.08) between populations as well as between regions of origin (0.04) of the population was found to be low. The low genetic variation may be due to the reduced population size of the wild sorghum in Ethiopia because of habitat change. Partitioning of the genetic variation into between and within the population as well as between and within the regions of origin revealed that 75% and 90% of the variation was found within the populations and within the regions, respectively. Cluster analysis of genetic distance estimates further confirmed the low level of differentiation of wild sorghum populations both on population and regional bases. The implications of the results for genetic conservation purposes are discussed.

Chapter 7

General Discussion and Conclusions

Genetic variation is a prerequisite for short- and long-term survival of a species and for crop improvement through plant breeding (Allard, 1996) and the significance of preserving the genetic diversity of domesticated and wild plant species is widely acknowledged (Frankel et al., 1995). As a result, many national, regional and international gene banks are maintaining a large number of germplasm collections of economically important crop plants. For instance, Ethiopia, being one of the Vavilovian centers of genetic diversity, is repository of germplasm collections of worldwide important crop plants, including sorghum. The collections include predominantly landraces, which are invaluable sources of useful genes for further increasing crop yield and quality. Landraces are particularly noted as rich sources of genes for disease and insect resistance as well as for stress tolerance against abiotic factors such as drought (Frankel et al., 1995; Zeven, 1998). Incorporating such useful genes into breeding materials helps to maintain the stability of crop productivity. Knowledge about the amount and patterns of distribution of genetic variation in germplasm collections is of great importance for designing future collection strategies and management of genetic resources as well as for utilization of such collections for plant breeding (Frankel et al., 1995; Ortiz et al., 1998).

In this thesis, we have presented the amount and patterns of distribution of genetic variation in cultivated sorghum (*Sorghum bicolor* ssp. *bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea as well as some introduced elite breeding materials using morphological, allozyme and RAPD data. We also investigated the extent and patterns of distribution of genetic variation in wild sorghum (*S. bicolor* ssp. *verticilliflorum* (L.) Moench) germplasm

from Ethiopia using RAPD markers. The central objective of the studies has been to provide comprehensive information about the magnitude and patterns of distribution of genetic diversity in sorghum germplasm from Ethiopia and Eritrea that would help enhance the management and utilization of the extant genetic resources.

Amount of genetic variation

For cultivated sorghum, analysis was made on 415 accessions for 15 quantitative characters and 10 qualitative characters having a total of 46 character states; 48 accessions for 27 alleles of 23 allozyme loci encoding 11 enzyme systems; and 80 accessions for 147 RAPD bands generated using 20 oligonucleotide PCR primers. The results show that the materials exhibited a wide range of variation for the 15 quantitative characters and a high level of variation for the 10 categorical morphological traits (mean Shannon-Weaver diversity index, $\bar{H} = 0.90 \pm 0.02$). However, the level of variation was found to be intermediate for the RAPD markers (mean Shannon-Weaver diversity index, $\bar{H} = 0.53 \pm 0.02$) and low for the allozymes (expected heterozygosity, $H_e = 0.042 \pm 0.017$). More recently, studying the patterns of morphological and allozyme variation in sorghum landraces of Northwest Morocco, Djè et al. (1998) also reported that the level of variation for morphological traits is more than that for allozymes. Similar conclusions have been reached in earlier studies (Morden et al., 1989, 1990; Ollitrault et al., 1989; Duncan et al., 1991). Lack of congruence between levels of genetic variation obtained from different techniques is also common in sorghum literature (Vierling et al., 1994; Yang et al., 1996; Djé et al., 1999).

Thus, despite considerable phenotypic diversity, studies of allozyme variation suggest that genetic diversity in sorghum is limited. The low allozyme polymorphism in sorghum could be attributed to its recent origin (Aldrich et al., 1992; Rowley-Conwy et al., 1999) and to its self-pollination (Doggett, 1988) as well as to population bottleneck during domestication (Morden et al., 1989; Gepts, 1993). Recently, Gupta et al. (1999) reviewed molecular markers and their applications in wheat breeding and concluded that one of the reasons for the low RFLP polymorphism in wheat is its recent origin. Self-pollinating species such as sorghum are also generally characterized by low levels of allozyme variation compared to their outcrossing counterparts (Hamrick and Godt, 1990, 1997).

The difference between the levels of allozyme and RAPD variations could be related to the number of polymorphic loci/bands analyzed. We analyzed about 37 times more number of polymorphic RAPD bands (147) than polymorphic allozyme loci (only 4 of the 23 loci were polymorphic). This result provides support to the claim made by Yee et al. (1999) that the ability to resolve genetic variation among accessions may be more directly related to the number of polymorphisms detected with each marker technique rather than a function of which technique is employed. Moreover, rates and modes of evolutionary change may differ greatly between coding and non-coding regions of the genome, because of differential operation of selection (Schaal et al., 1991; Brown and Schoen, 1992). It has been suggested that RAPD loci may, in certain instances, be more selectively neutral with respect to natural selection than isozyme loci (Dawson et al., 1995; Allard, 1996). There is a possibility that RAPD can access the more variable portion of the genome, i.e., the non-coding region (Williams et al., 1990, 1993).

In contrast, allozyme assay detects variation arising only in the coding region of the genome, which may be conserved to maintain the function of enzymes (Gottlieb, 1982). Still only a fraction of the mutational events, those changing amino-acid composition and

consequently charge and the electrophoretic mobility in a gel, can be resolved in allozyme assays (Clegg, 1990; Hartl and Clark, 1997). In other words, many variations at the DNA level can remain hidden at the protein level because several single mutations do not change the amino-acid composition or the charge of proteins. Hence, estimates of genetic variation that are based on allozyme may under estimate the overall amount of genetic variation actually present in a plant species.

On the other hand, all mutational changes in DNA can be potentially detected by DNA-based markers such as RAPD (Dubreuil and Charcosset, 1997; Karp et al., 1997). RAPD markers have also shown more levels of genetic variability than isozymes in red clover (*Trifolium pratense* L.) (Kongkiatngam et al., 1995), trembling aspen (*Populus tremuloides* Michx.) and bigtooth aspen (*P. grandidentata* Michx.), which are outcrossing tree species widely distributed in North America (Liu and Furnier, 1993).

Taking together the results of the present study and those of previous studies with similar objectives, there remains an intriguing point that needs explanation about the level of genetic variation in sorghum. Sorghum is noted as one of the most morphologically variable crop plants (Doggett, 1988) and this high variability has been a major problem in sorghum systematics (Snowden, 1936; Harlan and de Wet, 1972; de Wet, 1978). Our results apparently show higher level of variation for morphological traits than those obtained for allozymes and RAPD markers.

Similarly, if one critically reviews the literature for levels of genetic variation in sorghum obtained by analyses of allozymes (Morden et al., 1989 1990; Aldrich et al., 1992; Djé et al., 1998, 1999; Ayana et al., 2000c), RFLP (Aldrich and Doebley, 1992; Tao et al., 1993; Vierling et al., 1994), RAPD (Tao et al., 1993; Menkir et al., 1997; Ayana et al., 2000a), one would easily reach a conclusion that the level of genetic variation does not seem to match with the enormous morphological variation for which sorghum is well noted.

Relatively high levels of variation have been reported when hypervariable markers with a characteristic high mutation rates, such as SSR (Brown et al., 1996; Taramino et al., 1997; Dean et al., 1999; Grenier et al., 2000b; Djè et al., 2000) and ISSR (Yang et al., 1996) are used and when RFLP probes are carefully selected to cover the different sorghum linkage groups (Ahnert et al., 1996).

Such discrepancy between the level of morphological variation and the levels of protein and DNA polymorphisms has been observed in other crop plants such as cultivated peanut (*Arachis hypogaea* L.) (Kochert et al., 1991; Halward et al., 1992; Stalker et al., 1994; and Subramanian et al., 2000 for reviews) and melon (*Cucumis melo* L.) (Dane, 1983; Shattuck-Eidens et al., 1990). Reviewing a large body of literature in crop evolution studies, Gepts (1993) also concluded that there is a general reduction in diversity at the biochemical and molecular levels during domestication for majority of crop plants.

The discrepancy between the levels of morphological variation and that obtained through allozymes and various DNA-based markers is not limited to domesticated species. It is also common in wild species of crop plants. For instance, Black-Samuelsson (1997) studied the genetic variation and phenotypic plasticity in the rare plant species *Vicia pisiformis* L. and *V. dumentorum* L. (Fabaceae) and found higher level of variation for quantitative characters than for isozyme and RAPD and still relatively higher for RAPD than isozyme. Similar results have been obtained in other wild plants such as *Poa annum* L. (Poaceae) (Frenot et al., 1999) and arctic *Potentilla* (Rosaceae) (Hansen et al., 2000).

Possible explanations for the discrepancy between the levels of variation for morphological traits on one hand, and the levels of variation for allozymes and molecular markers on the other hand, are as follows:

(1) That the variation for morphological characters has both genetic, environmental and genotype by environment interaction components (Falconer and Mackay,

1996), and the relative contribution of each of these is not easy to separate in genetic diversity assessment of germplasm collections (because of lack of controlled crosses) using morphological and agronomic traits. If the magnitude of environmentally induced variation is large in comparison to genetic variation, diversity estimates based on morphological data may poorly reflect the actual levels of genetic diversity among accessions (Messmer et al., 1993; Yee et al., 1999). But the fact that quantitative characters are mostly adaptive and their variations are correlated with ecological factors, increase their value in plant breeding and conservation genetics (Lewontin, 1984; Hamrick, 1989; Black et al., 1995). Similarly, the present study on sorghum revealed the adaptive significance of several quantitative characters such as panicle compactness and shape, days for 50% flowering, and plant height (refer also to Ayana and Bekele, 1998, 1999, 2000).

(2) That the reduction in diversity at the biochemical and molecular levels during domestication is accompanied by an increase in diversity at the phenotypic level, perhaps being caused by a few genes having major effects (Gepts, 1993; Paterson, 1997).

(3) That perhaps such discrepancy is much more pronounced in those species that are of recent origin (Gepts, 1993; He and Prakash, 1997). This possibility could hold for sorghum if the proposal of its origin is accepted (Aldrich et al., 1992; Rowley-Conwy et al., 1999).

(4) That a species employ two alternative, but not mutually exclusive, adaptive strategies — genetic and adaptive phenotypic plasticity (Bradshaw and Hardwick, 1989; Sultan, 1995). If the species possesses no genetic variation or if its environment changes very fast, then, adaptive phenotypic plasticity becomes

The overall mean estimate of genetic variation for 83 individuals of wild sorghum representing 11 populations from five regions in Ethiopia was (mean Shannon-Weaver diversity index, $\bar{H} = 0.49$). Comparing this diversity index to that of the cultivated sorghum ($\bar{H} = 0.53$), the wild material seems to be slightly less variable than the cultivated one. It seems that the low level of genetic variation in wild sorghums from Ethiopia is due to reduction in population size because of natural habitat destruction through human activities (Ayana et al., 2000b). Previous studies indicate that wild relatives of sorghum maintain more genetic variation than domesticated one for allozymes (Morden et al., 1990) and for RFLPs (Aldrich and Doebley, 1992; Cui et al., 1995).

Regional levels of variation

Assessment of genetic diversity is important for identifying geographical regions where diversity is high (Tuwafe et al., 1988; Aldrich et al., 1992; Menkir et al., 1997). In the present study, examination of the levels of variation from the three sources of data all taken together (i.e., morphological, allozyme and RAPD) revealed that Gamo Gofa, Gonder, Harerge, Illubabor, Shewa and Welo rank as the top six regions with high amounts of variation. Furthermore, cluster analyses and/or ordination techniques of the three sources of data revealed that two or more of these regions are clustering together (refer to Figures 3.2, 3.3, 4.2, 4.3 and 5.3). However, perfect matching of the dendrograms or the plots is lacking, which could be due to several reasons, such as the use of different algorithms and different sample sizes (both number of regions and number of accessions per region), as well as or due

to the nature of the genetic basis of variation for morphological characters, allozymes and RAPD markers.

The results for morphological characters suggested that there were regional patterns of diversity for traits such as days for 50% flowering, plant height and panicle compactness and shape. Accessions from relatively dry regions were characterized by early flowering, short plant height and compact panicle types (refer also Ayana and Bekele, 1998, 1999, 2000). The clinal variation for plant height, days for 50% flowering, and the localized distribution of panicle compactness and shape, as well as the clustering together of regions with similar climatic and ecological conditions in morphological analysis all suggested that environmental factors play significant roles in shaping the amount and patterns of distribution of variation and thereby the adaptive significance of some of the morphological traits.

Levels of variation among adaptation zones

The distinctiveness of the 415 cultivated sorghum accessions was much more pronounced when discriminant analysis for quantitative morphological characters was made on the basis of adaptation zones than on the regions of origin. Multivariate analysis of morphological traits is known to indicate the adaptation of populations (Camussie et al., 1985; Souza and Sorrels, 1991; Zhong and Qualset, 1995; Ayana and Bekele, 1999). However, the levels of variation for categorical characters, allozymes and RAPD markers were found to be of similar magnitude for the three adaptation zones. Thus, there is no compelling reason to assert that the sorghum materials from the three adaptation zones are genetically

substructured. Then, they are of equal choice in breeding programs for either of the adaptation zones. However, as already indicated, future studies aimed at diversity analysis in sorghum in Ethiopia should consider the 18 main agro-ecological zones and 49 subagro-ecological zones of the country (MOA, 1998), instead of the grossly classified three or four adaptation zones for sorghum (Gebrekidan, 1981).

Patterns of distribution of variation

Information about how genetic variation is partitioned among accessions as well as among regions and adaptation zones is essential for effective germplasm management (Dean et al., 1999). The grouping of the accessions into different clusters using the three different data sources for cultivated sorghum was not perfectly related to either their geographic origin or adaptation zones, suggesting lack of complete differentiation. Moreover, all lines of evidence (qualitative morphological traits, allozymes and RAPD markers) obtained for cultivated sorghum indicated that a larger proportion of the variation is found within than among regions and within than among adaptation zones. In addition, for the allozyme data, 59% of the total variation was found within population, while the remaining 41% of the variation was between populations. The latter was found to be statistically significant, suggesting that the accessions are structured, confirming previous findings that the germplasm collection of sorghum is highly structured for allozymes (Morden et al., 1989; Djé et al., 1998, 1999).

Observation of the relative differentiation of the materials revealed that the levels of regional differentiation were more than those for the adaptation zones for categorical traits, allozymes and RAPD markers. This contrasts with the higher discriminative power of

adaptation zones than regions of origins for quantitative morphological characters (Ayana and Bekele, 1999), suggesting the adaptive nature of quantitative characters as compared to the presumably neutral molecular markers (Ohta, 1992). The fact that geographical origin has significant influence in sorghum germplasm differentiation has already been noted in many studies (Morden et al., 1989; Ollitrault et al., 1989; Aldrich et al., 1992; Menkir et al., 1997; Djé et al., 2000). Interestingly, these authors have shown that the level of genetic differentiation on the basis of geographical origin or individual accessions is higher than the level of differentiation obtained on the basis of racial classification. Note that the five basic races of sorghum and the 10 intermediate races are defined on the basis of spikelet morphology (Harlan and de Wet, 1972).

Sorghum is largely self-pollinating species (Doggett, 1988). Accordingly, a large portion of the variation must have occurred among populations than within populations (Hamrick and Godt, 1997). This has been supported by results obtained for sorghum in different studies using different techniques. Using allozymes data, Morden et al., (1989) reported that 91% of the total variation was found among accessions, while only 9% was found within accessions. Using simple sequence repeats, Dean et al. (1999) found that 90% of the total variation is found among accessions, while 10% of the variation was found within accessions of U.S. National sorghum collections. Similarly, using microsatellite markers, Djé et al. (2000) reported 67% and 33% of the total variation to occur among and within accessions, respectively. However, their results revealed relatively low level of geographical and racial differentiation for the 25 accessions that they studied.

In the present study, partitioning of the total variation for the wild sorghum also revealed that most of the variation was found within populations (75%) and within regions of origin (88%) rather than between populations (25%) and between regions of origin (12%).

Similar to the cultivated sorghum, genetic differentiation of the wild sorghum was also found to be low.

Implications for conservation

Identification of rich centers of diversity is necessary for undertaking *in situ* conservation of plant genetic resources (Marshall, 1990; Brush, 1995). The overall view of our results revealed that Gamo Gofa in the south, Gonder in the northwest, Harerge in the east, Illubabor (including the Gambella region) in the southwest, Shewa in the central and Welo in the northeast of Ethiopia are rich in sorghum genetic diversity and hence could serve as satellite regions for *in situ* conservation. Although considering all these regions for *in situ* conservation appears to be expensive, respective regional states can easily accommodate in their research and biodiversity conservation programs. In this way efforts can be shared but for similar goal. Gambella, Gamo Gofa and the Ghibe valley also need due attention for further exploration and *in situ* conservation of wild sorghum (refer to Ayana et al., 2000b).

Moreover, future sampling strategy should consider sampling as many as possible samples within regions and within adaptation zones and as many as possible individuals since a large portion of the variation has been found within categories compared with that between categories. This strategy is also useful to capture genes for adaptive traits and rare alleles. Further justifications for this sampling strategy are given in Ayana and Bekele (1998, 1999, 2000) and in Ayana et al. (2000a, b, c).

Implications for utilization in sorghum breeding

Estimation of genetic diversity helps in the evaluation of different germplasm as possible sources of genes that can be used in plant breeding programs (Cowen and Frey, 1987; Melchinger et al., 1991; Gilzlice et al., 1993; Schut et al., 1997). The results of the present studies show that sorghum accessions from Eritrea are early flowering and hence can be useful sources of genes for early maturity, for which there is a great need in Ethiopia and other African countries (Gebrekidan, 1981; Kebede, 1986; Ayana and Bekele, 1998, 1999, 2000). The results show similar materials could be found, although in low frequencies, among lowland accessions of Gamo Gofa, Gonder, Harerge, Illubabor (Gambella), Tigray, and Welo. Therefore, additional collection of short stature and early maturing types of sorghum should be made in the lowland areas of Gamo Gofa, Gonder, Harerge, Illubabor (Gambella), Tigray and Welo.

Likewise, choice of parents is a crucial step in planned hybridization because it largely determines the outcome of subsequent selection steps in segregating populations (Allard, 1999). Sorghum breeders produce many potentially useful crosses and evaluate subsequent segregating generations (Debelo et al., 1995). It would be more likely that a wide range of segregating materials are produced if crossing is made by choosing materials that appear significantly different in cluster analyses for both morphological, allozymes and RAPD data. So we suggest that future crossing programs consider the use of some of the divergent accessions identified in cluster analyses of morphological traits (Ayana and Bekele, 1999), RAPD (Ayana et al., 2000b) and allozymes (Ayana et al., 2000c).

Conclusions

The following major conclusions are made from the present studies.

(1) The sorghum materials displayed a wide range of variation for quantitative characters and a high level of variation for categorical morphological traits.

(2) The distributions of some of the characters such as panicle compactness and shape suggest that such characters have adaptive significance in sorghum.

(3) The clinal patterns of distribution of some of the morphological characters such as plant height and days for 50% flowering suggest that environmental factors influence the patterns of distribution and amount of variation for morphological traits.

(4) The level of genetic variation is low for allozymes.

(5) The level of genetic variation is intermediate for RAPD markers.

(6) The classification of the sorghum materials into three adaptation zones, i.e., lowland, intermediate and highland seems to be justified only for quantitative traits, otherwise the level of genetic differentiation appeared to be low.

(7) The pattern of distribution of the variation for morphological traits, allozymes and RAPD markers is similar in that more of the variation is found within than between regions, adaptation zones and populations.

(8) The level of RAPD variation in wild sorghum is low probably due to reduction in population size.

(9) The level of differentiation of the sorghum materials either on regional or on adaptation zones seems to be low.

(10) The relative differentiation is found to be higher on the basis of regions of origin than on the basis of the adaptation zones.

Suggestions for future research

We suggest that future studies to address the following areas.

(1) Continue the current lines of investigations (i.e., determine the amount and patterns of distribution of genetic variation) using more sample size and the 18 main agro-ecological and 49 sub agro-ecological zones of Ethiopia (MOA, 1998) as well as using more polymorphic markers such as SSR.

(2) Compare degree of heterosis and amount and nature of genetic variation in crosses obtained from similarity or genetic distance matrices of morphological, allozymes and RAPD markers.

(3) Compare the amount of allozyme variation among *in situ* and *ex situ* materials regenerated at a specified years of intervals. This study would be necessary to reach a firm conclusion whether the low level of allozyme variation is due to loss of genetic variation during regeneration or due to its evolutionary consequences such as recent origin (see arguments in chapter four or Ayana et al., 2000c).

(4) Study the amount and patterns of genetic variation in relation with the amount and patterns of phenotypic plasticity in sorghum. Results from such lines of investigations would help to get insight about the major strategy — genetic differentiation or adaptive phenotypic plasticity — of adaptation of sorghum to its highly variable environments.

(5) Determine the rate of outcrossing in different ecological zones of Ethiopia.

(6) Establish core collection of Ethiopian sorghum germplasm using morphological and passport data and subject the core collection to detailed genetic diversity assessment using both morphological and suitable molecular markers.

(7) Study the evolutionary consequences of the five basic races of cultivated sorghum (Harlan and de Wet, 1972) by transplant experiments. The aim would be to see, for example, whether characteristic features of guinea could divert into those of durra and vice versa when both are grown in appropriate ecological conditions for many generations. This is an experiment which requires long time and commitment.

We think that information obtained from the above lines of investigations would have both theoretical and practical significance for sorghum improvement.

References

Abebe, B. and Wech, H.B. (1982). The 981 activities of the Plant Genetic Resources Centre/Ethiopia. In: *Proceedings of the Regional Workshop on Sorghum Improvement in Eastern Africa, 17-22 October 1982*, pp. 31-45, (Anonymous ed). Addis Ababa, Ethiopia.

Abu-el-Gasim, E.H. and Kambal, A.E. (1975). Variability and inter-relations among characters in indigenous grain sorghums of the Sudan. East Afr. Agric. For. J. **41:** 125-133.

Acheampong, E., Anishetty, N.M. and Williams, J.T. (1984). *A World Survey of Sorghum and Millets Germplasm: International Board for Plant Genetic Resources First Decade of Service*, 1974-1984. IBPGR, Rome, Italy.

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

Alagarswamy, G. and Chandra, S. (1998). Pattern analysis in international sorghum multi-environment trials for grain yield adaptation. Theor. Appl. Genet. **96:** 397-405.

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.

Aldrich, P.R., Doebley, J. Schertz, K.F. and Stec, A. (1992). Patterns of allozyme variation in cultivated and wild *Sorghum bicolor*. Theor. Appl. Genet. **85:** 451-460.

Alika, J.E., Akenóva, M.E. and Fatokun, C.A. (1993). Variation among maize (*Zea mays* L.) accessions of Bendel State, Nigeria. Multivariate analysis of agronomic data. Euphytica **66:** 65-71.

Allard, R.W. (1970). Population structure and sampling methods. In: *Genetic Resources in Plants- Their Exploration and Conservation, IBP Handbook No. 11*, pp. 97-107, (Frankel, O.H. and Bennett, E., eds). F.A. Davis Company, Philadelphia, PA.

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

Allard, R.W. (1999). *Principles of Plant Breeding, 2nd edition.* John Wiley & Sons, New York.

Amurrio, J.M., de Ron, A.M. and Escribano, M.R. (1993). Evaluation of *Pisum sativum* landraces from the Northwest of Iberian Peninsula and their breeding value. Euphytica **66:** 1-10.

Andrews, D.J. and Bramel-Cox, P.J. (1993). Breeding cultivars for sustainable crop production in low input dry land agriculture in the tropics. In: *International Crop Science*, pp. 211-223, (Buxton, D.R., Shibles, R., Forsberg, R.A., Blad, B.L., Asay, K.H., Paulsen, G.M. and Wilson, R.F., Eds). Crop Sci. Soc. Amer. Inc., Madison, Wisconsin, USA.

Annicchiarico, P., Pecetti, L. and Damania, A.B. (1995). Relationships between phenotypic variation and climatic factors in durum wheat landraces. Hereditas **122:** 163-167.

Appa Rao, S. and Mushonga, J.N. (1987*). A Catalogue of Passport and Characterization Data of Sorghum, Pearl Millet and Finger Millet Germplasm from Zimbabwe.* IBPGR, Rome, Italy.

Appa Rao, S., Prasada Rao K.E., Mengesha, M.H. and Gopal Reddy, V. (1996). Morphological diversity in sorghum germplasm from India. Genetic Resources and Crop Evolution. **43:** 559-567.

Asante, S.A. (1995). Sorghum grain quality and utilization. African Crop Sci. J. **3:** 231-240.

Asfaw, Z. (1989). Relationships between spike morphology, hordeins and altitude within Ethiopian barley, *Hordeum vulgare* L. (Poaceae). Hereditas **110**: 203-209.

Atterre, A.F. (1994). Conservation of plant genetic resources in sub-Saharan Africa. In: *Safeguarding the Genetic Basis of Africa's Traditional Crops*, pp. 65-72, Putter, A., ed). CTA, The Netherlands/IBPGRI, Rome, Italy.

Avise, J.C. and Hamrick, J.L. (1997). *Conservation Genetics: Case Histories from Nature.* Chapman & Hall Cop., New York.

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. and Bekele, E. (2000). Geographical patterns of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea: Quantitative characters. Euphytica **115:** 91-104.

Ayana, A., Bryngelsson, T. and Bekele, E. (2000a). 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.

Ayana, A., Bekele, E. and Bryngelsson, T. (2000b). Genetic variation in wild sorghum (*Sorghum bicolor* ssp. *verticilliflorum* (L.) Moench germplasm from Ethiopia assessed by random amplified polymorphic DNA (RAPD). Hereditas **132:** 249-254.

Ayana, A., Bryngelsson, T. and Bekele, E. (2000c). Geographic and altitudinal allozyme variation in sorghum (*Sorghum bicolor* (L.) Moench) landraces from Ethiopia and Eritrea. Hereditas (submitted).

Ayyangar, G.N.R. (1942). The description of crop plant characters and their ranges of variation. IV. Variability of Indian sorghum. Indian J. Agric. Sci. **12:** 528-563.

Bahrman, N., Gous, J.L.E., Hariri, D., Luilbaud, L. and Jestin, L. (1999). Genetic diversity of old French six-rowed winter barley varieties assessed with molecular, biochemical and morphological markers and its relation to BaMMW resistance. Heredity **83:** 568-574.

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.

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.

Beer, S.C., Goffreda, J. Phillips, T.D., Murphy, J.P. and Sorrells, M.E. (1993). Assessment of genetic variation in *Avena sterilis* using morphological traits, isozymes, and RFLPs. Crop Sci. **33:** 1386-1393.

Bekele, E. (1983). Some measures of gene diversity analysis on landrace populations of Ethiopian barley. Hereditas **98:** 127-143.

Bekele, E. (1983b). The neutralist-selectionist debate and estimates allozyme multilocus structure in conservation genetics of the primitive land races of Ethiopian barley. Hereditas **99:** 73-88.

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

Bekele, E. (1996). Morphological analysis of *Eragrostis tef*: detection for regional patterns of variation. Sinet: Eth. J. Sci. **19:** 117-140.

Bello, A.B., Rooney, L.W. and Waniska, R.D. (1990). Factors affecting quality of sorghum *tô*, thick porridge. Cereal Chem. **67:** 20-35.

Bennetzen, J.L. (1995). Biotechnology for sorghum improvement. African Crop Sci. J. **3:** 161-170.

Berg, E.E. and Hamrick, J.L. (1997). Quantification of genetic diversity at allozyme loci. Can. J. For. Res. **27:** 415-424.

Beuselinck, P.R. and Steiner, J.J. (1992). A proposed framework for identifying, quantifying, and utilizing plant germplasm resources. Field Crops Res. **29:** 261-272.

Binelli, G., Gianfranceschi, L., Pe, M.E., Taramino, G., Busso, C., Stenhouse, J. and Ottaviano, E. (1992). Similarity of maize and sorghum genomes as revealed by maize RFLP probes. Theor Appl. Genet. **84:** 10-16.

Bjørnstad, Å., Demissie, A., Kilian, A. and Kleinhofs, A. (1997). The distinctness and diversity of Ethiopian barleys. Theor. Appl. Genet. **94:** 514-521.

Black, S., Eriksson, G., Gustafsson, L., and Lundkvist, K. (1995). Ecological genetics of the rare species *Vicia pisiformis*: Quantitative genetic variation and temperature response in biomass and fecundity. Acta Oecol. **16:** 261-275.

Black-Samuelsson, S. (1997). *Genetic Variation and Phenotypic Plasticity in the Rare Plant Species Vicia pisiformis L. and V. dumetorum L. (Fabaceae)*. Ph.D. Thesis, Swedish University of Agricultural Sciences, Department of Forest Genetics, Uppsala, Sweden.

Bogyo, T.P., Porceddu, E. and Perrino, P. (1990). Analysis of sampling strategies for collecting genetic material. Econ. Bot. **34:** 160-174.

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

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.

Bradshaw, A.D. and Hardwick, K. (1989). Evolution and stress—genotypic and phenotypic components. Biol. J. Linn. Soc. **37:** 137-155.

Bramel-Cox, P.J. Kumar, K.A., Hancock, J.D. and Andrews, D.J. (1995). Sorghum and millets for forage and feed. In*: Sorghum and Millets Chemistry and Technology*, pp.325-364, (Dendy, A.V., ed). Amer. Assoc. Cereal Chemists, Inc., St. Paul, Minni, USA.

Bretting, P.K. and Widrlechner, M.P. (1995). Genetic markers and plant genetic resources management. Plant Breeding Rev. **13:** 11-86.

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

Brown, A.H.D. (1978). Isozymes, plant population genetic structure, and genetic conservation. Theor. Appl. Genet. **52:** 145-157.

Brown, A.H.D. (1989a). The case of core collections. In: *The Use of Plant Genetic Resources*, pp. 136-156, (Brown, A.H.D., Frankel, O.H., Marshall, D.R. and Williams, T.J., eds). Cambridge Univ. Press, Cambridge.

Brown, A.H.D. (1989b). Core collections: a practical application of genetic resources management. Genome 31:818-824.

Brown, A.H.D. (1990). The role of isozyme studies in molecular systematics. Aust. Syst. Bot. **3:** 39-46.

Brown, A.H.D. and Munday, J. (1982). Population genetic structure and optimal sampling of landraces of barley from Iran. Genetica **58**: 85-96.

Brown, A.H.D. and Schoen, D.J. (1992).Plant population genetic structure and biological conservation. In: *Conservation of of Biodiversity for Sustainable Development*, pp. 88-104, (Sandlund, O.T., Hindar, K., and Brown, A.H.D., eds). Scandinavian University Press, Oslo.

Brown, J.S. (1991). Principal component and cluster analyses of cotton cultivar variability across the U.S. Cotton Belt. Crop Sci. **31:** 915-922.

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.

Brunk, C.F., Jones, K.C. and James, T.W. (1979). Assay for nanogram quantities of DNA in cellular homogenates. Analytical Biochem. **92:** 497-500.

Brush, S.B. (1995). *In situ* conservation of landraces in centers of crop diversity. Crop Sci **53:** 346-354.

Caetano-Anollis, G., Bassam, B.J., and Gresshoff, P.M. (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Bio/Technol. **9:** 553-557.

Camussi, A., Ottaviano, E., Calinski, T. and Kaczmarek, Z. (1985). Genetic distances based on quantitative traits. Genetics **111:** 945-962.

Central Statistics Authority, 1999. *Agricultural Sample Survey (1998/99), Report on Area and Production for Major Crops (Private Peasant Holdings, Main Season), Statistical Bulletin No. 2000.* Addis Ababa, Ethiopia.

Chan, K.F. and Sun, M. (1997). Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species of *Amaranthus*. Theor. Appl. Genet. **95:** 865-873.

Chalmers, K.J., Waugh, R., Sprent, J.I., Simons, A.J. and Powell, W. (1992). Detection of genetic variation between and within populations of *Gliricidia sepium* and *G. maculata* using RAPD markers. Heredity **69:** 465-472.

Chang, T.T. (1992). Availability of plant germplasm for use in crop improvement. In: *Plant Breeding in the 1990s, Proc. Symp. on Plant Breeding in the 1990s*, pp. 17-35, (Stalker, H.T. and Murphy, J.P., eds). CAB International.

Chittenden, L.M., 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.

Clawson, D.L. (1985). Harvest security and intra-specific diversity in traditional tropical agriculture. Econ. Bot. **39:** 56-76.

Clegg, M.T. (1990). Molecular diversity in plant populations. . In: *Plant Population Genetics, Breeding and Genetic Resources*, pp. 98-115, (Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S., eds). Sinauer Assoc., Sunderland, Massachsetts.

Cowen, N.M. and Frey, K.J. (1987). Relationships between three measures of genetic distance and breeding methods in oat (*Avena sativa* L.). Genome **29:** 97-106.

Cross, R.J. (1992). A proposed revision of the IBPGR barley descriptor list. Theor. Appl. Genet. **84:** 501-507.

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

Damania, A.B., Pecetti, L., Qualset, C.O. and Humeid, B.O. (1996). Diversity and geographic distribution of adaptive traits in *Triticum turgidum* L. (durum group) wheat landraces from Turkey. Genetic Resources and Crop Evolution **43:** 409-422.

Damon, E.G. (1962). *Cultivated Sorghums of Ethiopia, Experiment Station Bulletin, No. 6.* College of Agriculture, Haile Sellasie I University, Addis Ababa, Ethiopia.

Dane, F. (1983). Cucurbits. In: *Isozymes in Plant Genetics and Breeding, Part B.*, pp. 369-390, (Tanksely, S.D. and Orton, T.L., eds). Elsevier Sci. Publ., Amsterdam.

Dawson, I.K., Simons, A.J., Waugh, R. and Powell, W. (1995). Diversity and genetic differentiation among subpopulations of *Gliricidia sepium* revealed by PCR-based assays. Heredity **74:** 10-18.

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

Debelo, A., Mulatu, T., Degu, E. and Gutema, Z. (1995). Progress and Prospects of the National Sorghum Breeding Program. In: *Institute of Agricultural Research Center (IAR), Nazareth Research Center: 25 Years of Research Experience in Lowland Crops Improvement, 20-23 September 1995*, , p. 37, (Anonymous, ed). Nazareth Research Center, Melkassa, Ethiopia.

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.

DeRose, R.T., Ma, D-P., Kwon, I-S., Hasnain, S.E., Klassy, R.C., Hall, T.C. (1989). Characterization of the kafirin gene family from sorghum reveals extensive homology with zein from maize. Plant Mol. Biol. **12:** 245-256.

de Wet, J.M.J. (1978). Systematics and evolution of Sorghum sect. *Sorghum* (Gramineae). Am. J. Bot. **57:** 704-707.

de Wet, J.M.J. and Huckabay, J.P. (1967. The origin of *Sorghum bicolor*. II. Distribution and domestication. Evol. **21:** 787-802.

de Wet, J.M.J. and Harlan, J.R. (1971). The origin and domestication of *Sorghum bicolor*. Econ. Bot. **25:** 128-131.

de Wet, J.M.J. and Prasada Rao, K.E. (1986). Wild sorghums and their significance in crop improvement. In: *Proceeding of Sorghum Conference*, pp. 28-33, (Anonymous, ed). Shenyang, China.

de Wet J.M.J., Harlan, J.R. and Price, E.G. (1970). Origin of variability in the Spontanea complex of *Sorghum bicolor*. Amer. J. Bot. **57:** 704-707.

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

Deu, M., González-De-León, D., Glaszmann, J.C., Dégremont, 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.

Deu, M., Hamon, P., Chantereau, J., Dufour, P., D'hont, A. and Lanaud, C. (1995). Mitochonddrial DNA diversity in wild and cultivated sorghum. Genome **38:** 635-645.

Djè, Y., Ater, M., Lefèbvre, C. and Vekemans, X. (1998). Patterns of morphological and allozyme variation in sorghum landraces of North-western Morocco. Genetic Resources and Crop Evolution **45:** 541-548.

Djè, Y., Forcioli, D., Ater, M., Lefèbure, 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:** 157-163.

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

Doebley, J.F. and Wendel, J.F. (1989). Application of RFLPs to plant systematics. In: *Development and Applications of Molecular Markers to Problems in Plant Genetics. Current: Communications in Molecular Biology*, pp. 57-68, (Helentjaris, T. and Burr, B., ed). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Doebley, J.F., Goodman, M.M. and Stuber, C.W. (1984). Isozyme variation in *Zea* (Gramineae). Syst. Bot. **9:** 203-218.

Doebley, J.F., Goodman, M.M. and Stuber, C.W. (1985). Isozyme variation in the races of maize from Mexico. Amer. J. Bot. **72:** 629-639.

Doebley, J.F., Morden, C.W. and Schertz, K.F. (1986). A gene modifying mitochondrial malate dehydrogenase isozyme in sorghum (Gramineae). Biochem. Genetics **24:** 813-819.

Doggett, H. (1965a). The development of cultivated sorghums. In: *Essays on Crop Plant Evolution*, pp. 50-69, (Hutchinson, J.B., ed). Cambridge Univ. Press, Cambridge.

Doggett, H. (1965b). Disruptive selection in crop development. Nature **206:** 279-280.

Doggett, H. (1976). Sorghum: In: *Evolution of Crop Plants*, pp. 112-16, (Simmonds, N.W:, ed). Longman Group Limited, UK.

Doggett, H. (1982). The importance of food quality in sorghum improvement programs. In: *Proceedings of the International Symposium on Sorghum Grain Quality, 28-31 October 1981*, pp. 5-8, (Rooney, L.W. and Murty, D.S., eds). Patancheru, A.P., India.

Doggett, H. (1988). *Sorghum*, 2nd edition. Longman Group Limited, UK.

Doggett, H. (1991). Sorghum history in relation to Ethiopia. In: *Plant Genetic Resources of Ethiopia*, pp. 140-159, (Engels, J.M.M., Hawkes, J.G. and Worede, M., eds). Cambridge Univ. Press, Cambridge.

Doggett, H. and Majisu, B.N. (1968). Disruptive selection in crop development. Heredity **23:** 1-23.

Doggett, H. and Prasada Rao, K.E. (1995). Sorghum. In: *Evolution of Crop Plants, 2nd edition*, pp. 140-159, (Smartt; J. and Simmonds, N.W., eds). Cambridge Univ. Press, Cambridge.

Dubreuil, P. and Charcosset, A. (1997). Genetic diversity within and among maize populations: A comparison between isozyme and nuclear RFLP loci. Theor. Appl. Genet. **96:** 577-587.

Duncan, R.R. (1996). Breeding and improvement of forage sorghums for the tropics. Adv. Agron. **57:** 161-185.

Duncan, R.R., Bramel-Cox, P.J. and Miller, F.R. (1991). Contributions of introduced sorghum germplasm to hybrid development in the USA. In: *Use of Plant Introductions in Cultivar Development, Part 1*, pp. 69-101, (Shands, H.L. and Wesner, L.E., eds). CSSA Spec. Publ. 17., CSSA, Madison, WI., USA.

Duvall, M.R. and Doebley, J.F. (1990). Restriction site variation in chloroplast genome of *Sorghum* (Poaceae). Syst. Bot. **15:** 472-480.

Ebba, T. (1978). Plant Genetic Resources Center/Ethiopia: Activities and programs. In*: Crop Genetic Resources in Africa, Proceedings of a Workshop Jointly Organized by the Association for the Advancement of Agricultural Sciences in Africa and the International Institute of Tropical Agriculture, 4-6 January 1978*, pp. 25-30, (Anonymous, ed). IITA, Ibadan, Nigeria.

Elings, A. (1991). Durum wheat landraces from Syria. II. Patterns of variation. Euphytica **54:** 231-243.

Ellis, R.P., Foster, B.P., Robinson, D., Handley, L.L., Gordon, D.C., Russell, J.R., and Powell, W. (2000). Wild barley: a source of genes for crop improvement in the 21st century? J. Exptl. Bot. **51:** 9-17.

Ellstrand, N.C. and Foster, K.W. (1983). Impact of population structure on the apparent outcrossing rate of grain sorghum (*Sorghum bicolor*). Theor. Appl. Genet. **66:** 323-327.

Esechie, H.A., Maranville, J.W. and Ross, W.M. (1977). Relationship of stalk morphology and chemical composition to lodging resistance in sorghum. Crop Sci. **17:** 609-612.

Falconer, D.S. and Mackay, T.F.C. (1996). *Introduction to Quantitative Genetics, 4th edition.* Longman, Essex, UK.

FAO (1998). *Production Yearbook.* Vol. 52. Food and Agricultural Organization of the United Nations, Rome, Italy.

Faris, M.A., Lira, M.de. A. and Veiga, A.F.S.L. (1979). Stability of sorghum midge resistance. Crop Sci. **19:** 577-588.

Fernando, W.M.U., Hayward, M.D. and Kearsey, M.J. (1997). Isozyme and quantitative traits polymorphism in European provenances of perennial ryegrass (*Lolium perenne* L.). Euphytica **93:** 263-269.

Fleiss, J.L. (1973). *Statistical Methods for Rates and Proportions.* John Wiley & Sons, New York.

Frankel, O.H. (1970). Evaluation and utilization—Introductory remarks. In: *Genetic Resources in Plants—Their Exploration and Conservation, IBP Handbook No. 11*, pp. 395-401, (Frankel, O.H. and Bennett, E., eds). F.A. Davis Company, Philadelphia. PA., USA.

Frankel, O.H. (1984). Genetic perspectives of germplasm conservation. In: *Genetic Manipulation: Impact on Man and Society*, pp. 161-170, (Arber W.K., Limensee, K., Peacock, W.J. and Starlinker, P., eds). Cambridge Univ. Press, Cambridge.

Frankel, O.H. and Soule, M.E. (1981). *Conservation and Evaluation*. Cambridge Univ. Press, Cambridge, p. 195.

Frankel, O.H. and Brown, A.H.D. (1984). Current plant genetic resources — a critical appraisal. In: *Genetic New Frontiers Vol. IV*, pp. 3-13, (Chopra, V.L., Joshi, B.C., Sharma, R.P. and Bansal, H.C., eds). Oxford & IBH Publishing Co., New Delhi, India.

Frankel, O.H., Brown, A.H.D. and Burdon, J.J. (1995). *The Conservation of Plant Biodiversity*. Cambridge Univ. Press, Cambridge.

Frenot, Y., Aubry, M., Misset, M.T., Gloaguen, J.C., Gourret, J.P. and Lebouvier, M. (1999). Phenotypic plasticity and genetic diversity in *Poa annua* L. (Poaceae) at Crozet and Kergulen Islands (subantarctic). Polar Biol. **22:** 302-310.

Fuentes, J.L. and Escobar, F., Alvarez, A., Gallego, G., Cristina, D., Ferrer, M., Deus, J.E. and Tohme, J.M. (1999). Analysis of genetic diversity in Cuban rice varieties using isozyme, RAPD and AFLP markers. Euphytica **109:** 107-115.

Garber, E.D. (1950). Cytotaxonomic studies in the genus *sorghum*: Univ. Calif- Pub. Bot. **23:** 283-362.

Gebrekidan B. (1970). *A Report on Crop Collecting Trip to Konso Wereda in Gamo Gofa, 24 December 1969 - January 5, 1970*. Addis Ababa Univ. Press, Addis, Ababa, Ethiopia.

Gebrekidan, B. (1973). The importance of the Ethiopian sorghum germplasm in the world sorghum collection. Econ. Bot. **23:** 442-445.

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

Gebrekidan, B. (1982). Sorghum improvement in eastern Africa. In: *Proceedings of the Regional Workshop on Sorghum Improvement in Eastern Africa, 17-21 October 1982, Nazareth and Debre Zeit*, pp. 26-38, (Gebrekidan, B., ed). Institute of Agricultural Research, Addis Ababa, Ethiopia.

Gebrekidan, B. (1987). Sorghum improvement and production in eastern Africa. In: *Proceedings of an International Drought Symposium held at the Kenyatta Conference Center, 19th-23rd May 1986, Nairobi, Kenya*, pp. 35-51, (Menyonga, J.M., Bezuneh, T. and Youdeowei, A., eds). Nairobi, Kenya.

Gebrekidan, B. and Kebede, Y. (1977). *Ethiopian Sorghum Improvement Project Progress Report No. 5,* College of Agriculture, Addis Ababa University, Addis Ababa, Ethiopia.

Gebrekidan, B. and Kebede, Y. (1978). *Ethiopian Sorghum Improvement Project Progress Report No. 6.* College of Agriculture, Addis Ababa University, Addis Ababa, Ethiopia.

Gebrekidan, B. and Kebede, Y. (1979). The traditional culture and yield potentials of the Ethiopian high lysine sorghums. Eth. J. Agric. Sci. **1:** 29-40.

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* preparation and its quality parameters. In: *Proceedings of the International Symposium on Sorghum Grain Quality, 28-31 October 1981* pp. 335-345, (Anonymous ed) .ICRISAT, Patancheru P.O, A.P., India.

Gepts, P. (1990). Genetic diversity of seed storae proteins in plants. In: *Plant Population Genetics, Breeding and Genetic Resources*, pp. 64-82, (Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S., eds). Sinauer Assoc., Sunderland, Massachsetts.

Gepts, P. (1993). The use of molecular markers and biochemical markers in crop evolution studies. Evol. Biol. **27:** 51-94.

Gepts, P. (1995). Genetic markers and core collections. In: *Core Collections of Plant Genetic Resources*, pp. 127-146, (Hodgkin, T., Brown, A.H.D., Hintum, Th.J.L. van. and Morales, E.A.V., eds). International Plant Genetic Resources Institute (IPGRI), A Wiley-Sayce Publication.

Halward, T., Stalker, T., LaRue, E., and Kochert, G. (1992). Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis Hypogaea* L.). Plant Mol. Biol. **18:** 315-325.

Hamrick, J.L. (1989). Isozymes and analysis of genetic structure in plant populations. In: *Isozymes in Plant Biology*, pp. 87-105, (Soltis, D.E., Solits, P.S., eds). Dioscorides Press, Portland, Oregon, USA.

Hamrick, J.L. and Godt, M.J.W. (1990). Allozyme diversity in plant species. In: *Plant Population Genetics, Breeding and Genetic Resources*, pp. 43-63, (Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S., eds). Sinauer Assoc., Sunderland, Massachsetts.

Hamrick, J.L. and Godt, M.J.W. (1997). Allozyme diversity in cultivated crops. Crop Sci. **37:** 26-30.

Hamrick, J.L., Linhart, Y.B. and Mitton, J.B. (1979). Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Ann. Rev. Ecol. Syst. **10:** 173-200.

Hansen, K.T., Elven, R. and Brochmann, C. (2000). Molecules and morphology in concert: tests of some hypotheses in arctic *Potentilla* (Rosaceae). Amer. J. Bot.87:1466-1479.

Harlan, J. R. (1969). Ethiopia: A center of diversity. Econ. Bot. **23:** 309-314.

Harlan, J.R. (1971). Agricultural origins: Centers and non-centers. Sci. **174:** 468-473.

Harlan, J. R. (1975). Geographic patterns of variation in some cultivated plants. J. Heredity **66:** 182-191.

Harlan, J.R. (1976). Plant and animal distribution in relation to domestication. Philo. Trans. Roy. Soc. London B. **275:** 13-25.

Harlan, J. R. (1984). Evaluation of wild relatives of crop plants. In: *Crop Genetic Resources: Conservation and Evaluation*, pp. 213-222, (Holden, J.W.H. and Williams, J.T., eds). IBPGR, London.

Harlan, J.R. (1989). The tropical African cereals. In: *Foraging and Farming: The Evolution of Plant Exploitation*, pp. 335-343, (Harris, D.R. and Hillman, G.C., eds). Unwin Hyman Ltd, London.

Harlan, J.R. (1992). *Crops and Man, 2nd edition*. Amer. Soc. Agron. and Crop Sci. Soc. Amer., Madison, Wisconsin.

Harlan, J.R. and de Wet, J.M.J. (1971). Toward a rational classification of cultivated plants. Taxon **20:** 509-517.

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

Harlan, J.R. and Stemler, A.B.L. (1976). The races of sorghum in Africa. In: *Origins of African Plant Domestication*, pp. 465-478, (Harlan, J.R., de Wet, J.M.J. and Stemler, A.B.L., eds). Mouton Press, The Hague.

Harlan, J.R., de Wet, J.M.J. and Price, G. (1973). Comparative evolution of cereals. Evolution **27:** 325-325.

Harris, D.R. (1989). An evolutionary continuum of people-plant interaction. In: *Foraging and Farming: The Evolution of Plant Exploitation*, pp. 11-26, (Harris, D.R. and Hillman, G.C., eds). Unwin Hyman Ltd, London.

Harris, H. (1966). Enzyme polymorphism in man. Proc. Roy. Soc. Ser. B. **164:** 298-310.

Hartl, D.L. and Clark, A.G. (1997). *Principles of Population Genetics, 3rd edition*. Sinauer Assoc., Sunderland, Massachsetts.

Hash, C.T. and Bramel-Cox, P.J. (2000). Survey of marker applications. In: *Applications of Molecular Markers in Plant Breeding. Training Manual for a Seminar held at IITA,*

Ibadan, Nigeria, from 16-17 August 1999, pp. 3-12, (Haussmann, B.I.G , Geiger, H.H., Hess, D.E., Hash, C.T. and Bramel-Cox, P.J., eds). International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Andhra Pradesh, India.

Haussmann, B.I.G., Obilana, A.B., Ayiecho, P.O., Blum, A., Schipprack, W. and Geiger, H.H. (1999). Qantitative-genetic parameters of sorghum (*Sorghum bicolor* (L.) Moench) grown in semi-arid areas of Kenya. Euphytica **105:** 3109-118.

Haussmann, B.I.G., Obilana, A.B., Ayiecho, P.O., Blum, A., Schipprack, W. and Geiger, H.H. (2000). Yield and yield stability of four population types of grain sorghum in a semi-arid area of Kenya. Crop Sci. **40:** 319-328.

Hayward, M.D. and Breese, E.L. (1993). Population structure and variability. In: *Plant Breeding: Principles and Prospects*, pp. 16-29, (Hayward, M.D., Bosemark, N.O. and Romagosa, I., eds). Chapman & Hall.

He, G. and Prakash, C.S. (1997). Identification of polymorphic DNA markers in cultivated peanut (*Arachis Hypogaea* L.). Euphytica **97:** 143-149

Hedrick, P.W. (2000). *Genetics of Populations, 2nd edition.* Jones and Bartlett Pub. New York.

Hesketh, J.D., Chase, S.S. and Nanda, D.K. (1969). Environmental and genetic modification of leaf number in maize, sorghum and Hungarian millet. Crop Sci. **19:** 460-467.

Holcomb, J., Tolbert, D.M. and Jain, S.K. (1977). A diversity analysis of genetic resources in rice. Euphytica **26:** 441-450.

House, L.R. (1995a). Sorghum: One of the world's great cereals. African Crop Sci. J. **3:** 135-142.

House, L.R. (1995b). Sorghum and millets: History, taxonomy, and distribution. In: *Sorghum and Millets: Chemistry and Technology*, pp. 1-10, (Dendy, D.A.V., ed). Amer. Assoc. Cereal Chem., St. Paul, Minnestoa, USA.

Hoyt E. (1988). *Conserving the Wild Relatives of Crops.* IPGR, IUCN and WWF.

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.

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

Hutchenson, K. (1970). A test of comparing diversities based on the Shannon formula. J. Theor. Biol. **29:** 151-154.

IPGRI (1993). *Diversity and Development. The Strategy of the International Plant Genetic Resources Institute.* IPGRI, Rome, Italy.

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

Jain, S.K. (1979). Adaptive strategies: Polymorphism, plasticity, and homeostasis. In: *Topics in Population Biology*, pp. 160-187, (Solbrig, O.T., Jain, S.K., Johnson, G.B. and Raven, P.H., eds). Columbia Univ. Press, New York.

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.

Jaradat, A.A. (1992). Breeding potential of durum wheat landraces from Jordan. Phenotypic diversity. Hereditas **116:** 301-304.

Johnson, R.A. and Wichern, D.W. (1988). *Applied Multivariate Statistical Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

Jones, N., Ougham, H. and Thomas, H. (1997). Markers and mapping: We are all geneticists now. New Phytologist 137: 165-177.

Junghans, H. and Metzlaff, M. (1990). A simple and rapid method for the preparation of total DNA. Biotechniques **8:** 176.

Kahler, A.L. (1983). Inheritance and linkage of acid phosphatase locus Acp-4 in maize. J. Heredity **74:** 239-246.

Kahler, A.L. and Allard, R.W. (1970). Genetics of enzyme variants in barley. I. Esterases. Crop Sci. **10:** 444-448.

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

Kaplan, N.L., Hudson, R.R. and Langley, C.H. (1989). The hitchhiking effect revisited. Genetics **123:** 887-899.

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

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 Techniques, IPGRI Technical Bulletin No 2*. International Plant Genetic Resources Institute, Rome, Italy.

Kazan, K., Manners, J.M. and Cameron, D.F. (1993). Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus *Stylosanthes*. Genome **36:** 50-56.

Kebede, Y. (1986). The role of Ethiopian sorghum germplasm resources in the national breeding program. In: *The Conservation and Utilization of Ethiopian Germplasm, Proceedings of an International Symposium, 13-16 October 1986*, pp. 223-230, (Engels, J.M.M., ed). Addis Ababa, Ethiopia.

Kebede, Y., (1991). The role of Ethiopian sorghum germplasm resources in the national breeding program. In: *Plant Genetic Resources of Ethiopia*, p. 315-322, (Engels, J M., Hawkes, J.G. and Worede, M., eds). Cambridge Univ. Press, Cambridge.

Junghans, H. and Metzlaff, M. (1990). A simple and rapid method for the preparation of total DNA. Biotechniques **8:** 176.

Kahler, A.L. (1983). Inheritance and linkage of acid phosphatase locus Acp-4 in maize. J. Heredity **74:** 239-246.

Kahler, A.L. and Allard, R.W. (1970). Genetics of enzyme variants in barley. I. Esterases. Crop Sci. **10:** 444-448.

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

Kaplan, N.L., Hudson, R.R. and Langley, C.H. (1989). The hitchhiking effect revisited. Genetics **123:** 887-899.

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

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 Techniques, IPGRI Technical Bulletin No 2*. International Plant Genetic Resources Institute, Rome, Italy.

Kazan, K., Manners, J.M. and Cameron, D.F. (1993). Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus *Stylosanthes*. Genome **36:** 50-56.

Kebede, Y. (1986). The role of Ethiopian sorghum germplasm resources in the national breeding program. In: *The Conservation and Utilization of Ethiopian Germplasm, Proceedings of an International Symposium, 13-16 October 1986*, pp. 223-230, (Engels, J.M.M., ed). Addis Ababa, Ethiopia.

Kebede, Y., (1991). The role of Ethiopian sorghum germplasm resources in the national breeding program. In: *Plant Genetic Resources of Ethiopia*, p. 315-322, (Engels, J M., Hawkes, J.G. and Worede, M., eds). Cambridge Univ. Press, Cambridge.

Kebede, Y. and Menkir, A, (1987). Sorghum improvement for moisture-stress regions of Ethiopia. In: *Food Grain Production in Semi-Arid Africa*, pp. 131-139, (Menyonga, J.M., Bezuneh, T. and Youdeowei, A.). OAU7/STRC-SAFGRAD.

Kephart, S.R. (1990). Starch gel electrophoresis of plant isozymes: A comparative analysis of techniques. Amer. J. Bot. **77:** 693-712.

Kimura, M. (1968). Evolutionary rate at the molecular level. Nature **217:** 624-626.

King, L.M. and Schaal, B.A. (1989). Ribosomal-DNA variation and distribution in *Rudbeckia missouriensis*. Evol. **43**: 1117-1119.

Kochert, G., Halward, T., Branch, W.D. and Simpson, C.E. (1991). RFLP variability in peanut (*Arachis hypogea* L.) cultivars and wild species. Theor. Appl. Genet. **81:** 565-570.

Kongkiatngam, P. Waterway, M.J., Fortin, M.G. and Coulman, B.E. (1995). Genetic variation within and between two cultivars of red clover (*Trifolium pratense* L.): Comparisons of morphological, isozyme and RAPD markers. Euphytica **84:** 237-246.

Kouamé, C.N. and Quesenberry, K.H. (1993). Cluster analysis of a world collection of red clover germplasm. Genetic Resources and Crop Evolution. **40:** 39-47.

Kremer, A., Petit, R.G. and Pons, O. (1998). Measures of Polymorphism Within and Among Populations. In: *Molecular Tools for Screening Biodiversity. Plants and Animals*, pp. 301-311, (Karp, A., Isaac, P.G. and Ingram, D.S., eds). Chapman and Hall, London.

Kumar, S. and Singhania, D.L. (1984). Characters association and analysis in grain sorghum. Sorghum Newsletter **27:** 16-17

Kwolek, T.F., Atkins, R.E. and Smith, O.S. (1986). Comparisons of agronomic characteristics in C0 and C4 of IAP3BR(M) random-mating grain sorghum population. Crop Sci **26:** 1127-1131.

Labate, J.A., Lamkey, K.R., Lee, M. and Woodman, W.L. (1999). Temporal changes in allele frequencies in two reciprocally selected maize populations. Theor. Appl. Genet. **99:** 1166-1178.

Lewontin, R.C. (1984). Detecting population differences in quantitative characters as opposed to gene frequencies. Am. Nat. **13:** 115-124.

Lewontin, R.C. and Hubby, J.L. (1966). A molecular approach to the study of genic heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics **54:** 595-609.

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

Lin, C.S. and Binns, M.R. (1985). Procedural approach for assessing cultivar-location data: Pairwise genotype-environment interactions of test cultivars with checks. Can. J. Plant Sci. **65:**1065-1071.

Lin, C.S., Binns, M.R. and Lefkovitch, L.P. (1986). Stability analysis: Where do we stand? Crop Sci. **26:** 894-900.

Lin, Y-R., Shertz, K.F. and Paterson, A.H. (1995). Comparative analysis of QTLs affecting plant height and maturity across the Poaceae in reference to an interspecific sorghum population. Genetics **141:** 391-411.

Linn, S. and Arber, W. (1968). Host specificity of DNA produced by *Escherichia coli*. X. In vitro restriction of phage fd replicative form. Proc. Natl. Acad. Sci. USA **59:** 1300-1306.

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

Lothrop, J.E., Atkins, R.E. and Smith, O.S. (1985). Variability for yield and yield components in IAR1R grain sorghum random-mating population. II. Correlations, estimated gains from selection, and correlated responses to selection. Crop Sci. **25:** 240-244.

Loveless, M. and Hamrick, J.L. (1984). Ecological determinants of genetic structure in plant introductions. Ann. Rev. Ecol. Syst. **15:** 65-95.

Lu, M.Z., Szmidt, A.E. and Wang, X.R. (1995). Inheritance of RAPD fragments in haploid and diploid tissues of *Pinus sylvestris* (L.). Heredity **74:** 582-589.

Lynch, M. (1988). Estimation of relatedness by DNA fingerprinting. Mol. Biol. Evol. **5:** 584-599.

Lynch, M. and Milligan, M.G. (1994). Analysis of population genetic structure with RAPD markers. Mol. Ecol. **3:** 91-99.

McGregor, C.E., Lambert, C.A., Greyling, M.M., Louw, J.H. and Warnich, L. (2000). A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solannum tubersum* L.) germplasm. Euphytica **113:** 135-144.

Mann, J.A., Kimber, C.T. and Miller, F.R. (1983). *The Origin and Early Cultivation of Sorghums in Africa,* Texas Agric. *Exp. Stn. Bull. No 1454.* Texas A & M university, Texas.

Markert, C.L. and Moller, F. (1959). Multiple forms of enzymes: tissue, ontogenetic and species specific patterns. Proc. Natl. Acad. Sci. USA **45:** 753-763.

Marshall, D.R. (1990). Crop genetic resources: Current and emerging issues. In: *Plant Population Genetics, Breeding and Genetic Resources*, pp. 367-388, (Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S., eds). Sinauer Assoc., Sunderland, Massachsetts.

Marshall, D.R. and Jain, S.K. (1968). Phenotypic plasticity of *Avena fatua* and *A. barbata.* Amer. Naturalists **102:** 457-467.

Marshall, D.R. and Allard, R.W. (1970). Isozyme polymorphism in natural populations of *Avena fatua* and *A. barbata.* Heredity **29:** 373-382.

Marshall, D.R. and Brown, A.H.D. (1975). Optimum sampling strategies in genetic conservation. In: *Crop Genetic Resources for Today and Tomorrow*, pp. 53-80, (Frankel, O.H. and Hawkes, J.G., eds). Cambridge Univ. Press, Cambridge.

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

Mengesha, M. H. (1975). Crop germplasm diversity resources of Ethiopia. In: *Crop Genetic Resources for Today and Tomorrow*, pp. 449-453, (Frankel, O.H. and Hawkes, J.G., eds). Cambridge Univ. Press, Cambridge.

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

Meselson, M. and Yuan, R. (1968). DNA restriction enzyme from *E. coli.* Nature **217:** 1110-1114.

Messmer, M.M., Melchinger, A.E., Herrmann, R.G., Boppenmaier, J. (1993). Relationships among early European maize inbreds: II. Comparison of pedigree and RFLP data. Crop Sci. **33:** 944-950.

MOA (1998). *Agro-ecological Zones of Ethiopia.* Natural Resources and Management and Regulatory Department, Ministry of Agriculture, Addis Ababa, Ethiopia.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G. and Yano, M. (1997). Genome mapping, molecular markers, and marker-assisted selection in crop plants. Mol. Breeding **3:** 87-103.

Morden, C.F., Doebley, J.F. and Schert, K.F. (1987). *A Manual of Techniques for Starch Gel Electrophoresis of Sorghum Isozymes, Texas Agricultural Experiment Station MP-1635.* Texas College Station, Tx 77843.

Morden, C.W., Doebley, J.F. and Schertz, K.F. (1989). Allozyme variation in Old World races of *Sorghum bicolor* (Poaceae). Am. J. Bot. **76:** 247-255.

Morden, C.W., Doebley, J.F. and Schertz, K.F. (1990). Allozyme variation among the spontaneous species of *Sorghum* section Sorghum (Poaceae). Theor. Appl. Genet. **80:** 296-304.

Moreno-Gonzalez, J. and Cubero, J.I. (1993). Selection strategies and choice of breeding methods. In: *Plant Breeding, Principles and Prospects*, pp. 16-29, (Hayward, M.D., Bosemark, N.O. and Romagosa, I., eds). Chapman & Hall.

Mueller, U.G., and Wolfenbarger, L.L. (1999). AFLP genotyping and fingerprinting. Trends Ecol. and Evol. **14:** 389-394.

Mullis, K.B. and Faloona, F. (1987). Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. Meth. Enzymol. **155:** 335-350.

Mullis, K.B., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. **51:** 263-273.

Mumm, R.H. and Dudley, J.W. (1995). A PC SAS computer program to generate a dissimilarity matrix for cluster analysis. Crop Sci. **35:** 925-927.

Murty, B.R., Arundachlam, V. and Saxena, M.B.L. (1967). Classification and catalogue of a world collection of cultivated sorghums and pennisetums. Indian J. Genet. Plant Breeding. **27(s.):** 1-74.

Negassa, M. (1985). Patterns of phenotypic diversity in an Ethiopian barley collection and the Arusi-Bale highland as a center 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.

Nei, M. (1972). Genetic distance between populations. American Naturalists **106:** 283-292.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA **70:** 3321-3323.

Nei, M. (1977). F-statistics and analysis of gene diversity in subdivided populations. Ann. Hum. Genet. **41:** 225-233

Nei, M. (1978). Estimation of average heterozygosity and genetic distance froma fewindividuals. Genetics **89:** 583-590.

Newbury, H.J. and Ford-Lloyd, B.V. (1993). The use of RAPD for assessing variation in plants. Plant Growth Reg. **12:** 413-451.

Nevo, E., Baum, B., Beiles, A. and Johnson, D.A. (1998). Ecological correlates of RAPD diversity of wild barley, *Horedum spontaneum*, in the Fertile Crescent. Genetic Resources Crop Evolution **45:** 151-159.

Nevo, E., Beiles, A., Kaplan, D., Golenberg, E.M., Olsvig-Whittaker, L. and Naveh, Z. (1986). Natural selection of allozyme polymorphisms: A microsite test revealing ecological genetic differentiation. Evol. **40:** 13-20.

Ohta, T. (1992). Role of random genetic drift in the evolution of interactive systems. J. Mol. Evol. **44(S):** S9-S14.

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

Ollitrault, P., Arnaud, M., and Chantereau, J. (1989). Enzyme polymorphism in sorghum. II. Genetic and evolutionary constitution of cultivated sorghum. Agron. Trop. 44: 211-222.

Olsen, E.C. and Miller, R.L. (1958). *Morphological Integration.* Univ. of Chicago Press, Chicago.

Ortiz, R., Madssen, S. and Andersen, S.B. (1998). Diversity in Nordic spring wheat cultivars (1901-93). Acta. Scand. Sect. Soil and Plant Sci. **48:** 229-238.

Pammi, S., Schertz, K.F., Xu, G., Hart, G. and Mullet, J.E. (1994). Random-amplified-polymorphic DNA markers in sorghum. Theor. Appl. Genet. **89:** 80-88.

Parzies, H.K., Spoor, W. and Ennos, R.A. (2000). Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in *ex situ* gene banks. Heredity **84:** 476-486.

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

Paterson, A. H., Schertz, K.F., Lin, Y-R., and Li, Z. (1997). Case history in plant domestication, sorghum, an example of cereal evolution. In: *Molecular Dissection of Complex Traits*, pp. 187-196, (Paterson A.H., ed). CRC Press, Boca Raton, FL.

Patterson, F.L. and Ohm, H.W. (1975). Compensating ability of awns in Soft Red Winter Wheat. Crop Sci. **15:** 403-407.

Pecetti, L. and Damania, A.B. (1996). Geographic variation in tetraploid wheat (*Triticum turgidum* ssp. *turgidum convar. durum*) landraces from two provinces in Ethiopia. Genetic Resources and crop Evolution **43:** 395-407.

Pecetti, L., Annicchiarico, P. and Damania, A.B. (1992). Biodiversity in a germplasm collection of durum wheat. Euphytica **60:** 229-238.

Pedersen, J.F., Toy, J.J. and Johnson, B. (1998). Natural outcrossing of sorghum and sudangrass in the central great plains. Crop Sci. **38:** 937-939.

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

Persson, K. and Bothmer, R. von. (2000). Assessing the allozyme variation in cultivars and Swedish landraces of rye (*Secale cereale* L.). Hereditas **132:** 7-17.

Phillips, T.D. and Murphy, J.P. (1993). Distribution and analysis of isozyme polymorphism in North American cultivated oat germplasm. Crop Sci. **33:** 460-469.

Pluncknett, D.L., Smith, J.H., Williams, J.T. and Anishetty, N.M. (1983). Crop germplasm conservation and developing countries. Sci. **220:** 163-169.

Pollak, L.J., Gadner, C.O. and Parkhurst, A.M. (1984). Relationships between enzyme marker loci and morphological traits in two mass selected maize populations. Crop Sci. **24:** 1174-1179.

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. Breeding **2:** 225-238.

Prakash, S., Lewontin, R.C. and Hubbay, J.L. (1969). A molecular approach to the study of genic heterozygosity in natural populations. VI. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. Genetics **61:** 841-858.

Prasada Rao, K.E. and Mengesha M.H. (1981*).A Pointed Collection of Zera-zera Sorghum in the Gambella Area of Ethiopia, Genetic Resources Unit Progress Report, No. 33.* ICRISAT, Patancheru P.O., A.P., India.

Prasada Rao, K.E. and Ramanatha Rao, V. (1995). The use of characterization data in developing a core collection of sorghum. In: *Core Collections of Plant Genetic Resources*, pp. 49-67, (Hodgkin, T., Brown, A.H.D., Hintum, Th.J.L. van. and Morales, E.A.V., eds.). IBPGR, A Wiley-sayce.

Prasada Rao, K.E., Mengesha, M.H. and Reddy, V.G. (1989). International use of a sorghum germplasm collection. In: *The Use of Plant Genetic Resources*, pp. 49-67, (Brown, A.H.D., Marshall, D.R. Frankel, O.H. and Williams, T.J., eds). Cambridge Univ. Press, Cambridge.

Quinby, J.R. (1967). The maturity genes of sorghum. Adv. Agron. **19:** 267-305.

Rieseberg, L.H. (1996). Homology among RAPD fragments in inter-specific comparisons. Mol. Ecol. **5:** 99-105.

Rio, A.H. del., Bamberg, J.B., Huaman, Z., Salas, A. and Vega, S.E. (1997). Assessing changes in the genetic diversity of potato gene banks. 2. *In situ* vs *ex situ*. Theor. Appl. Genet. **95:** 199-204

Rohlf, F.J. (1993). *NTSYS-PC Numerical Taxonomy and Multivariate Analysis System, version 1.80.* Applied Biostatistics Inc., New York.

Rowley-Conwy, P. Deakin, W. and Shaw, C.H. (1999). Ancient DNA from sorghum: The evidence from Qasr Ibrim, Egyptian Nubia. In*: The Exploitation of Plant Resources in Ancient Africa*, pp. 55-61, (Veen, M. van. de.). Kluwer Academic/Plenum Publishers, New York.

Russell, J.R., Fuller, J.D., Macaulay, M., Hats, 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, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Sci. **230:** 1350-1354.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Sci. **239:** 487-491.

SAS Institute, 1988. *SAS/STAT™ Users Guide, Release 6.03 edition.* SAS Institute Inc., Cary, NC., USA.

Schaal, B.A., Leverich, W.J., and Rogstad, S.H. (1991). Comparisons of methods assessing genetic variation in plant conservation biology. In: *Genetic and conservation of Rare Plants*, pp. 123-134, (Falk, D.A. and Holsinger, K.E., eds). Oxford Univ. Press. Oxford and New York.

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.

Shamsudin, A.K.M. (1985). Genetic diversity in relation to heterosis and combining ability in spring wheat. Theor. Appl. Genet. **70:** 306-308.

Shattuck-Eidens, D.M., Bell, R.N., Neuhausen, S.L. and Helentjaris, T. (1900). DNA sequence variation within maize and melon: Observations from polymerase chain reaction amplification and direct sequencing. Genetics **126:** 207-217.

Shechter, Y. (1975). Biochemical systematic studies in *Sorghum bicolor*. Bull. Torrey Bot. Club **102:** 334-339.

Shechter, Y. and de Wet, J.M.J. (1975). Comparative electrophoresis and isozyme analysis of seed proteins from cultivated races of sorghum. Amer. J. Bot. **62:** 254-261.

Silander, J.A. (1985). The genetic basis of the ecological amplitude of *Spartina patens*. II. Variance and correlation analysis. Evol. **39:** 1034-1052.

Singh, R. and Axtell, J.D. (1973). High lysine mutant gene (hl) that improves protein quality and biological value of sorghum. Crop Sci. **13:** 535-539.

Singh, R.K and Chaudhary, B.D. (1985). *Biometrical Methods in Quantitative Genetic Analysis*. Kalyani Publ., New Delhi, India.

Singh, S. P. (1985). Sources of cold tolerance in grain sorghum. Can. J. Plant Sci. **65:** 251-257.

Singh, S.P., Gutierrez, J.A., Molina, A., Urrea, C. and Gepts, P. (1991). Genetic diversity in cultivated common bean: II. Marker-based analysis of morphological and agronomic traits. Crop Sci. **31:** 23-29.

Skroch, P. W. and Nienhuis, J. (1995). Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. Theor. Appl. Genet. **91:** 1086-1091.

Skroch, P.W., Nienhuis, J., Beebe, S., Tohme, J. and Pedraza, F. (1998). Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. Crop Sci. **38:** 488-496.

Slatkin, M. and Barton, N.H. (1989). A comparison of three indirect methods for estimating average levels of gene flow. Evol. **43:** 1349-1368.

Smale, M. (1997). The green revolution and wheat genetic diversity: Some unfounded assumptions. World Development **25:** 1257-1269.

Smith, J.S.C. and Smith, O.S. (1988). The description and assessment of distance between inbred lines of maize: I. The use of morphological traits as descriptors. Maydica **34:** 141-150.

Smith, J.S.C. and Smith, O.S. (1989). The description and assessment of distances between inbredlines of maize. II. The utility of morphological, biochemical and genetic descriptors and a scheme for testing of distinctiveness between inbred lines. Maydica **34:** 151-161.

Smith, J.S.C. and Smith, O.S. (1992). Fingerprinting crop varieties. Adv. Agron. **47:** 85-140.

Smith, J.S.C., Paszkiewicz, O.S., Smith, O.S. and Schaffer, J. (1987). Electrophoretic, chromatographic, and genetic techniques for identifying associations and measuring genetic diversity among corn hybrids. Ann. Corn and Sorghum Res. Conf. Proc. **42:** 187-203.

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

Sneath, P.H.A. and Sokal, R.R. (1973). *Numerical Taxonomy.* W. H. Freeman, San Francisco.

Snowden, J.D. (1936). *The cultivated races of sorghum.* Allard and Son, London.

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

Souza, E. and Sorrells, M.E. (1991). Relationships among 70 American oat germplasm. I. Cluster analysis using quantitative characters. Crop Sci. **31:** 599-605.

Stalker, H.T., Phillips, T.D., Murphy, J.P. and Jones, T.M. (1994). Variation of isozyme patterns among *Arachis* species. Theor Appl. Genet. **87:** 746-755.

Stemler, A.B.L., Harlan, J.R. and de Wet, J.M.J. (1975). Evolutionary history of cultivated sorghums (*Sorghum bicolor* (L.) Moench) of Ethiopia. Bull. Torrey Bot. Club. **102:** 325-333.

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

Stephan, W., Wiehe, T.H.E. and Lenz, M.W. (1992). The effect of strongly selected substitutions on neutral polymorphism: Analytical results based on diffusion theory. Theor. Pop. Biol. **41:** 237-340.

Stickler, F.C., Wearden, S. and Pauli, A.W. (1961). Leaf area determination in sorghum. Agron. J. **53:** 187-188.

Stuber, C.W. and Moll, R.H. (1972). Frequency changes of isozyme alleles in a selection experiment for grain yield in maize (*Zea mays* L.). Crop Sci. **12:** 337-340.

Subramanian, V., Gurtu, S., Nageswara Rao, R.C., and Nigam, S.N. (2000). Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome **43:** 656-660.

Sultan, S.E. (1995). Phenotypic plasticity and plant adapttion. Acta Bot. Neerl. **44:** 363-383.

Sun, Y., Skinner, D.Z., Liang, G.H. and Hulbert, S.H. (1994). Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. **89:** 26-32.

Swofford, D.L. and Selander, R.B. (1989). *BIOSYS-1: A Computer Program for the Analysis of Allelic Variation in Population Genetics and Biochemical Systematics*, Release 1.7. Natural History Survey, Champaign, Illinois, USA.

Tanksley, S.D. and McCouch, S.R. (1997). Seed banks and molecular maps: Unlocking genetic potential from the wild. Sci. **277:** 1063-1066.

Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierable, M.W. (1989). RFLP mapping in plant breeding: new tools for an old science. Biotech. **7:** 257-264.

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

Taramino, G., Tarchini, R., Ferrario, S., Lee, M. and Pé, M.E. (1997). Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor. Appl. Genet. **95:** 66-72.

Tato, K. (1964). Rainfall in Ethiopia. Eth. Geog. J. **2:** 28-36.

Tautz, D. (1989): Hypervariability of simple sequence as a general source for polymorphic DNA markers. Nucl. Acids Res. **7:** 6463-6470.

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

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., 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.

Thorpe, R.S. (1976). Biochemical analysis of geographical variation and racial affinities. Biol. Rev. **51:** 407-452.

Vierling, R.A., Xiang, Z., Joshi, C.P., Gilbert, M.L. and Nguyen, H.T. (1994). Genetic diversity among elite sorghum lines revealed by restriction fragment length polymorphisms and random amplified polymorphic DNAs. Theor. Appl. Genet. **87:** 816-820.

Virupasha, T.K. and Sastry, L.V.S. (1968). Studies on the protein content and amino acid composition of some varieties of grain sorghum. J. Agric. Food Chem. **16:** 199-203.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kupier, M., Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting . Nucleic Acids Res. **23:** 4407-4414.

Wachira, F.N.R., Waugh C.A., Hackett, C.A. and Powell, W. (1995). Detection of genetic variation in tea (*Camellia sinensis*) using RAPD markers. Genome **38:** 201-210.

Warwick, S.I., Thompson, B.K. and Black, L.D. (1984). Population variation of *Sorghum halepense*, Johnson grass, at the northern limits of its range. Can. J. Bot. **62:** 1781-1790.

Wasylikowa, K. (1997). Flora of the 8,000 years old archaeological site E-75-6 at Nabta Playa, Western Desert, southern Egypt. Acta Palaeobotanica **37:** 99-205.

Wasylikowa, K. and Dahlberg, J.A. (1999). Sorghum in the early economy of the early neolithic nomadic tribes at E-75-6 at Nabta Playa, Western Desert, southern Egypt. In: *The Exploitation of Plant Resources in Ancient Africa*, pp. 11-31, (Veen, M van de). Kluwer Academic/Plenum Publishers, New York.

Waugh, R. and Powell, R. (1992). Using RAPD markers for crop improvement. Trends Biotechnol. **10:** 186-191.

Weeden,. N.F. (1989). Applications of isozymes in plant breeding. Plant Breeding Rev. **6:** 11-54.

Weir, B.S. (1996). *Methods for Discrete Population Genetic Data.* Sinauer Associates, Sunderland, Massachusetts.

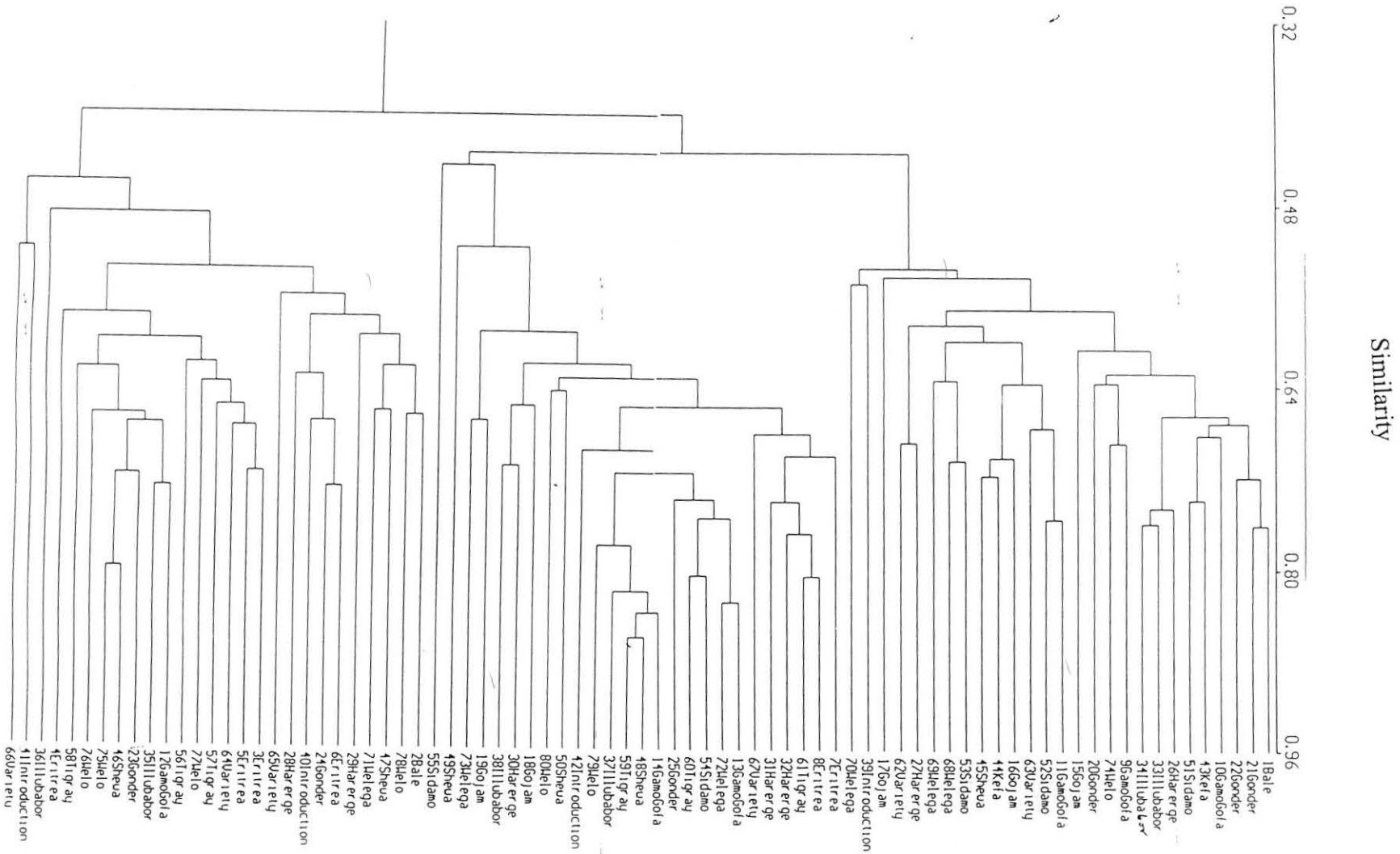
Appendix 3.1. Continued.

No	Acc.No	Altitude	Origin	Status	No	Acc.No	Altitude	Origin	Status
249	223589	1600	Shewa	Landrace	279	226072	2370	Gonder	Landrace
250	225831	2260	Gamo Gofa	Landrace	280	226073	2345	Gonder	Landrace
251	225831	2260	Gamo Gofa	Landrace	281	226074	1905	Gonder	Landrace
252	225832	2260	Gamo Gofa	Landrace	282	226075	1905	Gonder	Landrace
253	225833	2260	Gamo Gofa	Landrace	283	226076	1820	Gonder	Landrace
254	226046	1600	Welo	Landrace	284	226077	1820	Gonder	Landrace
255	226047	1580	Welo	Landrace	285	226078	1855	Gonder	Landrace
256	226048	2570	Welo	Landrace	286	226079	1855	Gonder	Landrace
257	226049	2485	Welo	Landrace	287	226080	1760	Gonder	Landrace
258	226050	2485	Welo	Landrace	288	226082	1760	Gonder	Landrace
259	226051	2450	Welo	Landrace	289	226083	1980	Gonder	Landrace
260	226052	2450	Welo	Landrace	290	226085	1955	Gonder	Landrace
261	226054	1500	Welo	Landrace	291	226087	1660	Gojam	Landrace
262	226055	1450	Welo	Landrace	292	227742	1480	Welo	Landrace
263	226056	2540	Welo	Landrace	293	228108	1540	Welo	Landrace
264	226057	2450	Welo	Landrace	294	228109	1520	Welo	Landrace
265	226058	2400	Welo	Landrace	295	228111	1500	Welo	Landrace
266	226059	2400	Welo	Landrace	296	228114	1500	Welo	Landrace
267	226060	2420	Welo	Landrace	297	228115	1500	Welo	Landrace
268	226061	2590	Welo	Landrace	298	228116	1500	Welo	Landrace
269	226062	2535	Welo	Landrace	299	228251	1550	Welo	Landrace
270	226063	2530	Welo	Landrace	300	228252	1550	Welo	Landrace
301	228253	1580	Welo	Landrace	331	231186	1500	Harerge	Landrace
302	228254	1580	Welo	Landrace	332	231187	1550	Harerge	Landrace
303	228736	1500	Shewa	Landrace	333	231188	1500	Harerge	Landrace
304	228738	1750	Shewa	Landrace	334	231189	1570	Harerge	Landrace
305	228821	1690	Harerge	Landrace	335	231190	1550	Harerge	Landrace
306	228824	1810	Harerge	Landrace	336	231191	1500	Harerge	Landrace
307	228826	1480	Harerge	Landrace	337	231192	1540	Harerge	Landrace
308	228830	1550	Harerge	Landrace	338	231193	1500	Harerge	Landrace
309	228831	1550	Harerge	Landrace	339	231194	1710	Harege	Landrace
310	228832	1550	Harerge	Landrace	340	231195	1960	Harerge	Landrace
311	228833	1920	Harerge	Landrace	341	231196	1650	Harege	Landrace
312	228836	1690	Harerge	Landrace	342	231196	1500	Harerge	Landrace
313	228839	1660	Harerge	Landrace	343	231197	2000	Harerge	Landrace
314	228846	1550	Harerge	Landrace	344	231198	2000	Harerge	Landrace
315	228849	1550	Harerge	Landrace	345	231199	1600	Harege	Landrace
316	228850	1580	Harerge	Landrace	346	231200	1720	Harege	Landrace
317	228851	1400	Harerge	Landrace	347	231202	1700	Harege	Landrace
318	228917	1800	Illubabor	Landrace	348	231203	1700	Harege	Landrace
319	228922	1800	Illubabor	Landrace	349	231204	1670	Harege	Landrace
320	229240	2550	Shewa	Landrace	350	231206	1700	Harege	Landrace
321	229839	1450	Gojam	Landrace	351	231207	1700	Harege	Landrace
322	230065	1300	Bale	Landrace	352	231208	1550	Harerge	Landrace
323	230242	1750	Eritrea	Landrace	353	231209	1550	Harerge	Landrace

Appendix 3.1. Continued.

No	Acc.No	Altitude	Origin	Status	No	Acc.No	Altitude	Origin	Status
324	230243	1750	Eritrea	Landrace	354	231459	1420	Harerge	Landrace
325	230780	1200	Sidamo	Landrace	355	231461	1450	Harerge	Landrace
326	230875	1800	Harege	Landrace	356	234089	2110	Tigray	Landrace
327	231179	1600	Harege	Landrace	357	234096	1550	Tigray	Landrace
328	231181	1800	Harege	Landrace	358	235442	1850	Gonder	Landrace
329	231182	1950	Harerge	Landrace	359	235443	1850	Gonder	Landrace
330	231185	2000	Harerge	Landrace	360	235445	1880	Gonder	Landrace
361	235447	1060	Tigray	Landrace	390	235932	620	Gonder	Landrace
362	235448	1600	Tigray	Landrace	391	235935	620	Gonder	Landrace
363	235450	1710	Tigray	Landrace	392	Alemaya-70	1980	Variety	Variety
364	235453	1650	Tigray	Landrace	393	Bakomash	1700	Variety	Variety
365	235457	1680	Tigray	Landrace	394	Birmash	1650	Variety	Variety
366	235458	1740	Tigray	Landrace	395	Dinkmash	1426	Variety	Variety
367	235460	1700	Tigray	Landrace	396	ETS-2752	1920	Variety	Variety
368	235462	1500	Tigray	Landrace	397	Gambella-1107	1500	Variety	Variety
369	235465	1490	Tigray	Landrace	398	76T1#23	1400	Variety	Variety
370	235466	1490	Tigray	Landrace	399	Seredo	1500	Variety	Variety
371	235467	1460	Tigray	Landrace	400	P9401	1500	America	Introduction
372	235468	1700	Tigray	Landrace	401	P9408	1500	America	Introduction
373	235469	1700	Tigray	Landrace	402	ISCV112BF	1500	Burkina Faso	Introduction
374	235469	1700	Tigray	Landrace	403	94MKSb#81	1500	India	Introduction
375	235477	2260	Tigray	Landrace	404	94MKSb#82	1500	India	Introduction
376	235478	1620	Tigray	Landrace	405	KAT/369-1	1500	Kenya	Introduction
377	235908	1620	Tigray	Landrace	406	KAT/369-2	1500	Kenya	Introduction
378	235910	1430	Gonder	Landrace	407	IS2284	1500	Somalia	Introduction
379	235913	1640	Tigray	Landrace	408	IS9302	1500	South Africa	Introduction
380	235914	1640	Tigray	Landrace	409	IS9308	1500	South Africa	Introduction
381	235920	1150	Gonder	Landrace	410	CR:35:5	1500	Sudan	Introduction
382	235921	1150	Gonder	Landrace	411	SRN-39	1426	Sudan	Introduction
383	235922	1000	Gonder	Landrace	412	94MK#84	1500	Uganda	Introduction
384	235923	950	Gonder	Landrace	413	94MK#86	1500	Uganda	Introduction
385	235924	910	Gonder	Landrace	414	3443-2-OP	1500	Zimbabwe	Introduction
386	235925	960	Gonder	Landrace	415	RSAVTENT#8	1500	Zimbabwe	Introduction
387	235926	930	Gonder	Landrace					
388	235927	820	Gonder	Landrace					
389	235928	850	Gonder	Landrace					

Appendix 5.3. Dendrogram for 80 sorghum accessions (generated from Jaccard similarity matrix)



Appendix 6.1. Dendrogram for 93 sorghum accessions (generated from Jaccard similarity matrix)