Population Genetics and Ecological Studies in Wild Sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia: Implications for Germplasm Conservation

A Dissertation Submitted to the School of Graduate Studies
College of Natural Sciences
Microbial, Cellular and Molecular Biology Program Unit
Addis Ababa University
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Biology (Applied Genetics)

Asfaw Adugna Ageru

Addis Ababa

Advisors:

Prof. Endashaw Bekele
Microbial, Cellular and Molecular Biology Program Unit, Addis Ababa University
Prof. Allison A. Snow
Department of Evolution, Ecology and Organismal Biology, Ohio State University, USA

May 2012

ACKNOWLEDGEMENTS

Above all, I thank the almighty GOD for keeping me strong from the start to the end of this work. I thank Prof. Endashaw Bekele, my major advisor, for his professional and fatherly supervision. I take in the inadequacy of words to thank my co-advisor, Prof. Allison A. Snow, Distinguished Scholar of the Department of Evolution, Ecology and Organismal Biology (EEOB), Ohio State University for her professional guidance in every walk of the dissertation research activities. I would have never done this work without the sole opportunity created to me by Prof. Gebisa Ejeta, Distinguished Professor of Agronomy, Purdue University, Indiana, and the 2009 World Food Prize Laureate. I thank Ohio State University for offering me the visiting research scholar position in the Department of EEOB to learn and undertake the molecular work. I also have special words of thanks to Dr. Patricia M. Sweeney, the Snow lab ex-manager for professional guidance in the laboratory during molecular work. My thanks also go to Prof. Peter S. Curtis, Chair, Department of EEOB and Mrs. Faith Haleem, for administrative help during my stay in the department.

This study was fully supported by the grant from the Biotechnology and Biodiversity Interface (BBI) program of the United States Agency for International Development (USAID) to Prof. Allison A. Snow, Department of EEOB, Ohio State University, under a competitive grant entitled “Continued US-Africa Research and Capacity Building on Risk Assessment of Crop-wild Gene Flow in Sorghum.” The logistic support given by Ethiopian Institute of Agricultural Research (EIAR) is duly appreciated.

I have to thank my beloved wife W/ZO Tsige Abate, my daughters Hilina Asfaw, Maryamawit Asfaw and Veronica Asfaw for their love, care and patience, which have been always power to me. I also thank my sisters Atitegeb Adugna and Ejigitu Adugna and my brothers Tesfahun Adugna, Dereje Adugna and Yihun Adugna for their encouragement.

I thank Dr. Adefris Teklewold, Crops Research Director, EIAR, Dr. Tesfaye Tesso, Department of Agronomy, Kansas State University, Dr. Mesfin Tadesse, Curator, college of biological sciences museum, Ohio State University, Ato Taye Tadesse, PhD student, University of Queensland, and Ato Alemu Tirfessa, the national sorghum research project coordinator for their kind help and encouragement. My thanks also go to Ato Mesfin Bekele, Ato Hailye Tadesse, Ato Tsegaye G/Maryam, and Ato Hussen Harun-for their help during sample collection.The Snow lab group Amy Stottlemyer, Hiochi Chang (Chi), Samantha, Maria, Jenalle, for their assistance in DNA purification work. I also thank Dr. Evans Mutegi, post doctoral fellow at EEOB for the crop and ICRISAT accessions DNA analysis in the lab and for his comments on the thesis. I thank Dr. Jose Ziad of the Department of EEOB for his technical help with the genetic analyzer (sequencer).

My thanks extend to Prof. Kermit Ritland, Statistical Molecular Ecologist, University of California, Davis, for his advice in the number of SSR markers and the use of MLTR Software for estimating outcrossing rate. I also thank Ato Gizachew Legesse and Dr. Mezgebu Getinent for construction of maps and Dr. Kifle Dagne and Dr. Kassahun Tesfaye for their encouragement and comments on the thesis.

I thank the management, administration and finance staff at Melkassa Agricultural Research Center: Dr. Lemma Desalegne, Dr. Setegn Gebeyehu, Dr. Getachew Ayana, Ato Alemseged Mamuneh, Ato Mamuye Belayneh. I also thank Ato Solomon Leulseged, ex-farm Manager of MARC.

The following people have helped me in one way or another during the research work: Ato Tessema Weldu, Ato Feyera Liben, Ato Tesfaye Terefe, Ato Bedo Dugo and all the drivers at MARC, Ato Mohammed Salah, Ato Worku Beyene, Ato Adane W/Giorghis, Ato Amare Nega, Ato Sewmehon Siraw, Ato Tilahun Haddis and the sorghum research and technical group, Aster, Damenu, Genet, Tsige, Etenesh, Eshetu, Bayu. Last, but not least, I have great thanks and appreciation to all those, whose names I might have not listed.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABREVATIONS

AMOVA - Analysis of Molecular Variance

ANOVA - Analysis of variance

AW - Presence of awn at maturity

CSA - Central Statistical Agency

DNA - Deoxyribose Nucleic Acid

DTE - Days to emergence

DTF - Days to flowering

EDTA - Ethylenediaminetetraacetic acid

GLC - Glume color

GLH - Presence of glume hair

GRC - Grain color

GRCV - Grain covering

GRS - Grain size

IAM - Infinite Allele Model

IBD - Isolation by Distance

ICRISAT - International Crops Research Institute for the Semi-Arid Tropics

LL-Leaf length

LN-Leaf number

LW-Leaf width

NSPP - Number of seeds per panicle

OP - Osmotic Potential

PCA - Principal Component Analysis

PCR - Polymerase Chain Reaction

PCS - Panicle compactness and shape

PH - Plant Height

SH - Shattering

SMM- Stepwise Mutation Model

SSR - Simple-sequence Repeat

SSW-Single seed weight

TE - Tris EDTA

TILL-Number of tillers per plant

TSW - Thousand Seed Weight

UPGMA - Unweighted Pair-group Method Using Arithmetic Averages

ABSTRACT

An exploration was made from October through November in 2008 to five sorghum growing geographical regions to study the distribution, the range of co-occurrence and diversity in the crop-wild-weed sorghum complex. Data were recorded on *in situ* quantitative and qualitative phenotypic characters from 30 populations. Moreover, the genetic structure of 19 wild populations from the five regions, eight cultivar populations from three regions, and 10 wild sorghum accessions from ICRISAT was studied using nine SSR loci. The extent of outcrossing was investigated in seven wild/weedy sorghum populations using five polymorphic SSR markers. The study was also aimed at investigating the fitness of wild-crop sorghum hybrids for various juvenile survival, adult and fertility phenological and morphological characters. Further, included in the present study was investigation of ecotypic differences of wild sorghum for dormancy, longevity, and their seed germination requirements as part of a risk assessment of crop-wild gene flow. There was high diversity among the wild/ weedy sorghum populations for phenotypic traits. SSR diversity was greater in the Ethiopian wild sorghum populations than in the sampled cultivars or wild accessions. Analysis of molecular variance (AMOVA) showed that 41% of the genetic variation in the wild plants was partitioned among populations, indicating a high degree of differentiation and the average number of migrants per generation (N_m) was 0.43 indicating limited gene flow within the wild pool. Cluster analyses showed that some wild populations were grouped by geographic region, whereas others were not, presumably due to long-distance seed movement. There was moderate differentiation between the wild and the cultivated sorghum probably because of historical gene flow. Wild sorghums collected from different geographical regions

exhibited variation (range=0.31-0.65) in outcrossing rate. Most wild × crop hybrids didn't show any fitness costs with respect to the measured traits and in some cases they showed mid-parent heterosis. The study indicated that crop-to-wild gene flow is possible, and the existence of morphologically intermediate forms between cultivated and wild sorghum indicates that gene flow is likely to have occurred in Ethiopia. Moreover, the high outcrossing rates of wild/weedy sorghum populations in Ethiopia may indicate a high potential for the spread of crop genes into the wild pool. The probable gene flow from the cultivated sorghum to the wild sorghum and introgression may pose risk of loss of genetic diversity in the wild and genes from transgenic sorghum are expected to enter wild populations if transgenic sorghum is deployed in Eastern Africa. Therefore, effective risk management strategies may be needed if the introgression of crop genes (including transgenes) from improved cultivars into wild/ weedy populations is deemed to be undesirable.

Key words: Gene flow, genetic diversity, mating systems, population structure, sorghum, transgenic

1. INTRODUCTION

All the major crop species of today are believed to have been domesticated by humans between 5000 and 10,000 years ago (Smith, 2001) from their wild ancestors through conscious and unconscious selection for agronomically desirable traits. The majority of them originated in today's developing countries, which hold a high diversity of traditional crop varieties and crop wild relatives and thus could serve as important sources for new genes. Although the antiquity of sorghum is not known (Doggett, 1965) it is assumed to have originated from around 4000-3000 B.C. in Ethiopia and surrounding countries from wild forms (Dillon *et al*., 2007).

Sorghum is among the most important grain crops in the world including Ethiopia. Because of its multiple purposes and its ability to cope up with unfavorable growing conditions, sorghum will continue to feed the world's expanding populations. Moreover, sorghum will be the crop of the future due to the changing global climatic trends and increase in use of marginal lands for agriculture (Paterson *et al*., 2008). Even though wild relatives and traditional varieties could serve as sources of new genes and offer an untapped wealth of novel traits against new pests or changing abiotic conditions (Tanksley and McCouch, 1997; Dillon *et al*., 2007), little attention has been paid to wild sorghum conservation in Ethiopia. Moreover, very limited research has been done to study the geographical distribution, the extent of genetic diversity and population structure, and mating systems yet these are important steps on the effort towards the use of these valuable genetic resources. The presence of wild and weedy types of sorghum

(Amsalu Ayana *et al*., 2001; Tesfaye Tesso *et al*., 2008) and heterogeneous environmental conditions in Ethiopia make the country an ideal place for conducting population genetic studies of wild sorghum.

Population genetic studies are useful for determining how genetic diversity is partitioned within and among populations of native plant species and for inferring levels of gene flow among populations. This baseline knowledge provides a starting point for designing *in situ* and *ex situ* conservation programs (Rice, 2004; Ellstrand and Elam, 1993). Determining the level of gene flow between crops and their wild relatives also is critical for conserving wild germplasm (Ellstrand *et al*., 1999; Jarvis and Hodgkin, 1999).

Conservation of genes in the wild has become a concern especially since the past two decades due to the development and official release of transgenic crop plants (Gepts and Papa, 2003). The concerns are to the most part associated with the trepidation that transgenic plants containing herbicide tolerance gene(s) (for example) may escape to the environment and create increased weediness or even "super weeds", which can be difficult to control. Due to the increasing deployment of genetically modified (GM) crops in developing countries (James, 2011), scientific risk assessment of gene flow from transgenic plants and their impact on conventionally bred crops and crop wild relatives is needed to establish adequate biosafety regulations and measures. For example, it is important to understand the level of outcrossing under current field conditions. Assessment of gene flow and its consequences should also include the current crop-specific mitigation strategies by studying the species biology such as sexual

compatibility, presence of related species, seed dispersal, fecundity and dormancy (Van Deynze, 2011) because seeds and pollen are the vehicles of gene dispersal and movements in plant species.

Although centers of diversity of crop species are often mentioned in discussions of risk assessment and regulatory policies for GM crops (Gepts and Papa, 2003; Haygood *et al*., 2003), little work has been done so far to address the issue of gene flow and outcrossing rates under field conditions in sorghum in the center of origin. To date, no transgenic sorghum is under commercial production, but it is under development for improved nutritional traits (Zhao, 2007). Even if studies have been conducted to estimate outcrossing rate in culivated sorghum (Dje *et al*., 2004; Barnaud *et al*., 2008), and in wild sorghum (Muraya *et al*., 2011b), such work is lacking in the center of origin. In Ethiopia, it is likely that natural cross-fertilization between wild and cultivated species occurs since the two forms grow together (Tesfaye Tesso *et al*., 2008). Determining the level of gene flow between crops and their wild relatives is critical for sustainable management of genetic resources in the wild (Ellstrand *et al*., 1999; Jarvis and Hodgkin, 1999) and to develop specific biosafety regulation for GM sorghum.

1.1. Underlying hypotheses and objectives of the study

The study was carried out with the following hypotheses:

1. Wild sorghum adapted to different geographical and climatic regions might show different patterns of genetic diversity
2. Genetic diversity differs among and within populations of wild sorghum
3. The population genetics of the species can be explained by various evolutionary forces
4. The genetic variation (gene flow) of the species follows isolation-by-distance
5. Wild and cultivated sorghums co-exist in sympatric range and their co-occurrence might have resulted in gene flow
6. Wild sorghum could have high outcrossing rate which could lead to wild to wild gene flow so that genes (including transgenes) transferred from crop to wild could spread in the wild pool
7. Crop genes (including transgenes) could have fitness advantage in the crop-wild hybrids and could lead to increased weediness
8. Wild sorghum seeds carrying crop genes could persist in the soil bank and serve as reservoir for future gene flow and this persistence could differ among ecotypes

The study had the following objectives

General objective

To study the population genetics and ecology of wild sorghum and the factors associated with biosafety and genetic resources conservation

Specific objectives:

i. To investigate the distribution and *in situ* phenotypic diversity in wild sorghum in areas where sorghum is the major crop;

ii. To determine the distribution of microsatellite variation within and among different populations of wild sorghum to aid efforts to conserve genetic diversity;

iii. To estimate *in situ* outcrossing rates of wild sorghum populations using microsatellite markers;

iv. To quantify the levels of gene flow within and among populations of wild and cultivated sorghum using indirect population genetic methods;

v. To estimate the fitness of wild × crop sorghum hybrids compared to their parents for various life history traits;

vi. To assess ecotypic variation in wild/weedy sorghum populations for dormancy, longevity and germination requirements.

2. REVIEW OF LITERATURE

2.1. Origin, domestication and distribution of *Sorghum bicolor*

The antiquity of sorghum is not clearly known (Doggett, 1976) and there is no consensus among the different authorites with regard to its origin. There is a speculation that the domestication has its origins 3000-5000 years ago (Doggett, 1976; Dillon *et al*., 2007; Laidlaw and Godwin, 2008). There is also anthropological evidence that hunter-gatherers consumed the crop during the early 8000BC (Smith and Frederickson, 2000). The primary centers of origin and diversity for sorghum are believed to be in sub-Sahara Africa extending from the extreme East to West Africa (Laidlaw and Godwin, 2008). Most literatures share the speculation that this region is in the North eastern quadrant of Africa (Doggett, 1965; 1988). However, at present the most accepted probable place seems to be the area extending from Ethiopia to Lake Chad, which was also supported by allozyme evidence (Aldrich *et al*., 1992). Members of the wild *S. bicolor* ssp. *verticilliflorum* are reported to be the immediate progenitors of the domesticated sorghum (Snowden, 1936; Doggett, 1965; de Wet and Huckabay, 1967; de Wet, 1978) and are dry savanna plants that were most probably domesticated west of the Ethiopian highlands (Stemler *et al*., 1977).

Stemler *et al*. (1977) reported the existence of all races of *S. bicolor* ssp. *verticilliflorum* except race *virgatum* in Ethiopia. Amsalu Ayana *et al*. (2001) reported the presence of wild sorghum *(Sorghum bicolor* ssp. *verticilliforum* (L.) Moench) in five districts of

western, north western and south western Ethiopia. Likewise, survey results of Tesfaye Tesso *et al*. (2008) indicated that wild sorghum has been found in all the surveyed areas in eastern, north eastern and northern Ethiopia. There are no other reports regarding the classification and distribution of wild sorghum in the country.

2.2. Taxonomy in the genus *Sorghum*

Formal taxonomy

The taxonomy of sorghum is among the most unsettled despite the publication of a number of papers by different authors with continued modification (e.g., Snowden, 1936; Garber, 1950; De Wet, 1978). Sorghum belongs to the family poacea, tribe Andropogoneae, subtribe Sorghinae, genus Sorghum Moench (Clayton and Renvoize, 1986). Moench (1794) coined the name sorghum by changing the original name "*Holcus*" that was given by Linnaeus in 1753 (House, 1985). Piper (1916), Stapf (1917), Kuwada (1915, 1919), Marchal (1920), and Vinall (1926), all contributed to the classification of sorghum (as cited in House, 1985). However, the significant changes in the taxonmic classification were made after the publication of Snowden's (1936) book, containing a comprehensive grouping within the genus. He classified the genus Sorghum into *Halepensia* and *Arundinacea* and further the *Arundinacea* into 48 taxa including cultivated, wild and, hybrids between them. The *Halepensia* includes four rhizomatous taxa including Johnson grass (*S. halepense*) (De Wet and Harlan, 1971).

Cytogenetically Johnson grass (*S. halepense*) is tetraploid (2n = 40). Later they all were included under the species *S. bicolor* (L.) Moench (de Wet and Huckabay, 1967). Garber

(1950) and Celarier (1959) divided the genus in to five subgenera: Eusorghum (sorghum), Chaetosorghum, Heterosorghum, Parasorghum, and Stiposorghum. The sorghum genus consists of 25 species (USDA-ARS, 2007). Species of the genus have chromosome numbers of 2n = 10, 20, 30 or 40 (Garber, 1950; Lazarides *et al*., 1991). Kuwada (1915) first reported the diploid chromosome number of *Sorghum bicolor* (the then *Andropogon sorghum*) to be 2n=20.

The most recent classification was made by de Wet (1978) based on experimental data accumulated over two decades of biosystematic studies of herbarium specimens and seed collected across the range of section Sorghum in Europe, Africa and southern Asia. He changed the five "subgenera" classification given by Garber (1950) into "sections". According to such a classification, section Eusorghum includes the species *S. bicolor* which in turn includes three sub-species: 1) ssp. *bicolor* which includes 5 basic and 10 hybrid cultivated races; 2) sub-species *drummondii*, which includes the hybridized forms between the cultivated and the races of ssp. *verticilliflorum*; and 3) sub-species *arundinaceum*, which was first divided into 3 races/varieties: *arundinaceum*, *verticilliflorum*, and *aethiopicum* by de Wet and Huckabay (1967) and later race *virgatum* was added as a fourth race within the sub-species by De Wet *et al*. (1970). Later Dogget (1988) named the sub-species as *verticilliflorum*. *S. bicolor* ssp *drummondii* (also known as shatter cane) is a well known aggressive weed and causes much damage to the crop (Harlan and de Wet, 1974; Ejeta and Grenier, 2005).

Sorghum bicolor ssp. *arundinaceum* and ssp. *verticilliflorum* have very similar morphological appearance in that both of them have large and open inflorescences, the difference being the flexuous and/ or spreading inflorescence branches that are undivided near the base in the case of ssp. *arundinaceum* but are usually divided near the base in ssp. *verticilliflorum*. (Harlan and De Wet, 1971). Ssp. *aethiopicum* and ssp. *virgatum* are also characterized by their small, contracted inflorescences with suberect branches. ssp. *virgatum* is typically known for its narrowly linear leaf blades (De Wet and Harlan, 1971; 1978).

Folk Taxonomy

Farmers in different parts of Ethiopia recognize each sorghum type commonly grown in their area by various vernacular/ folk names (Brhane Gebrekidan, 1982; Awgchew Teshome *et al*., 1997; Beyene Seboka and van Hintum, 2006; Firew Mekbib, 2007). These names may be reflections of their morphological appearance e.g., Mishinga Worabisa (Oromo language, meaning hyena like sorghum, due to the awn covering of the panicle), Cherekit (Amharic, moon like, due to its very white, shiny seeds it lights as a moon) or the special purposes they offer, e.g., Marye (Amharic, my honey, due to its sweet taste), Gan seber (Amharic, pot breaker, due to its potency in making local beer). Moreover, names are given due to their ability in surviving biotic and abiotic stresses (e.g., *Kitign ifere*, unafraid of syphilis (*Kitign*/striga, mean to say tolerant to striga). They also recognize wild sorghums by different vernacular names. However, unlike the cultivated sorghum, most of the naming is related to their noxious nature. For instance, in Wello and Pawe areas farmers call them “Killo” (the fool, as it fools farmers by

mimicry), “Yeseyitan ageda” (Satan’s cane), “yeseyitan mashila” (Satan’s sorghum) in Amharic to express their abhorrence to this weed. In Oromifa, it is called *Seepo*. In Tigray it is called “Kancha Seyitan” (Alamata area), “Kuchuye” (Adenkel area), “Seyitan zer” (in Subday), and “Aadar” (in Hagereselam and Humera areas). “Aadar” is similar to the naming in the adjacent region of Sudan (Gebisa Ejeta and Grenier, 2005).

2.4. Importance of cultivated and wild sorghum

Sorghum is the fifth most important food crop in the world and a staple food crop in many regions of Africa and Asia (Ejeta and Grenier, 2005). In Ethiopia, it covers 16% of the total area allocated to grains (cereals, pulses, and oil crops) and 19.5% of the area covered by cereals (CSA, 2011). The crop stands third next to tef and maize in area of production (CSA, 2011). It is the second most important crop for *injera* making quality next to tef. The grain is used for the preparation of traditional foods, distilled and undistilled beverages and the stalk is used for construction, fuel and animal feed. In Ethiopia the crop grows almost exclusively during the main rainy season.

The primary gene pools of many crop plants are depleted in genetic variability and breeders often rely on wild relatives for sources of important agronomic traits (e.g., Langridge and Chalmers, 2004). Therefore, wild relatives and locally adapted cultivars (landraces) are often viewed as sources of "genetic insurance" against emerging pests or changing abiotic conditions (Tanksley and McCouch, 1997) including climate change.

Thus, there has been continued interest in studying the wild relatives of domesticated species as reservoirs of genetic variation for crop improvement (Escalante *et al*., 1994).

As Ethiopia is one of the centers of origin for sorghum (De Wet and Harlan 1971; Doggett 1988), diversified forms of the crop and its wild relatives represent possible sources of germplasm for crop improvement. For example, sources of resistance have been identified from wild sorghum species to ergot and green bug (Reed *et al*., 2002) and to Striga (Rich *et al*., 2004). Singh and Axtell (1973) identified two high-lysine Ethiopian sorghum varieties. Moreover, recent studies identified sources of "stay-green" drought tolerance derived from sorghum lines native to Ethiopia (Borrel, 2001). As a center of origin and diversity for domesticated sorghums, Ethiopia may harbor unique wild germplasm that is worthy of further conservation efforts.

2.5. Genetic diversity in sorghum

Over the years, a number of studies have been carried out in many countries including Ethiopia to estimate genetic diversity in cultivated sorghum using phenotypic traits (e.g., Amsalu Ayana and Endashaw Bekele, 1998; 2000; Zongo *et al*., 1993; Appa Rao *et al*., 1996; Dahlberg *et al*., 2002; Shehzad *et al*., 2009), RAPDs (Abebe Menkir *et al*., 1997; Amsalu Ayana *et al*., 2000; Agrama and Tuinistra, 2003; Nkongolo and Nsapato, 2003; Uptmoor *et al*., 2003), RFLPs (Cui *et al*., 1995; Yang *et al*., 1996; Jordan *et al*., 1998), AFLPs (Uptmoor *et al*., 2003; Menz *et al*., 2004; Nemera Geleta *et al*., 2006; Ritter *et al*., 2007), and SSRs (Dean *et al*., 1999; Dje *et al*., 1999; 2000; Smith *et al*., 2000; Grenier *et al*., 2000; Ghebru *et al*., 2002; Agrama and Tuinistra, 2003; Uptmoor *et al*., 2003; Menz

et al., 2004; Abu Assar *et al*., 2005; Casa *et al*., 2005; Dillon *et al*., 2005; Folkertsema *et al*., 2005; Nemera Geleta *et al*., 2006; Manzelli *et al*., 2007; Barnaud *et al*., 2007; Ali *et al*., 2008; Deu *et al*., 2008; Zhan *et al*., 2008; Wang *et al*., 2009; Shehzad *et al*. 2009; Mutegi *et al*., 2010). Some of these studies were based on world and national accessions from gene banks, while others were based on field collections and most of them reported high diversity.

Few studies dealt with the genetic diversity of wild sorghum accessions (Aldrich *et al*., 1992; Cui *et al*., 1995; Casa *et al*., 2005). Amsalu Ayana *et al*. (2001) 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 and reported low genetic diversity. There were no other reports on wild sorghum diversity in the country. Recently, in *situ* population genetic studies done on wild sorghum in Kenya and Mali using SSR markers indicated high diversity (Mutegi *et al*., 2010; 2011; 2012; Muraya *et al*. 2011; Sagnard *et al*., 2011).

The earlier reports on molecular marker based studies used randomly amplified polymorphic DNA (RAPD) markers, which lack many of the desirable properties a molecular marker has to fulfil. However, simple sequence repeat (SSR) markers are currently the markers of choice for population genetic studies due to their high polymorphism even between closely related individuals within a species (Edwards *et al*., 1996), high reproducibility, codominance, abundance, and fairly evenly distributed throughout the euchromatic region of the genome (e.g., Schlotterer, 2004). However, they

have limitations in genotyping errors associated with their potential null alleles, large allele drop out and homoplasy (Estoup et al., 2002; Dewoody et al., 2006).

Todate, only two studies by Nemera Geleta *et al*. (2006) and Alemu Tirfessa (2009) used these markers to study genetic diversity in Ethiopian cultivated sorghum.

2.6. Estimation of population genetic structure

Population structure is defined as "the amount of genetic variability and its distribution within and among local populations and individuals within a species" (Templeton, 2006). Genetic structure has an impact in the ecological adaptation and evolution of plant species (Loveless and Hamrick, 1984). Genetic diversity can be studied using phenotypic traits, biochemical, or molecular markers. The materials used for the study can be accessions from gene banks, or field collections. Phenotypic traits do not give reliable estimate of genetic diversity as these traits are limited in number and due to environmental effect. On the contrary, molecular diversity data can potentially bridge conservation and use when employed as a tool for mining germplasm collections for genomic regions associated with adaptive or agronomically important traits (Casa *et al*., 2005).

For a better understanding of the evolutionary forces responsible for genetic variation, a geographic component should be included in population genetic structure studies (Gillespie, 1998). This is true because the geographical distribution of crops and their wild relatives determines the probability of gene exchange between them and this is

expected to be more in the centers of origin and/or centers of diversity (Simmonds, 1995).

Different coefficients can be used to study genetic structure among populations. Wright (1951) first introduced fixation index, F_{ST} for use in population genetic studies and most studies involving population genetic structure have used it as a measure of population differentiation. Fixation index (F_{ST}) is computed as (Weir and Cockerham, 1984):

$$F_{ST} = \frac{\sigma_p^2}{\bar{p}(1-\bar{p})}$$

Where, p is the frequency of allele A, $\bar{P}$ is the mean frequency of one of the alleles (with two alleles) in the population, σ^2 is the variance in frequencies among populations (demes). The interpretation of population differentiation based on F_{ST} was given by Hartl (1980). Based on the magnitude of their F_{ST}, population pairs can be considered as divergent or related. Accordingly, $F_{ST} < 0.05$ can be considered as low, $F_{ST} = 0.05 - 0.15$ shows moderate differentiation, $F_{ST} = 0.15 - 0.25$ can be considered as great differentiation, and $F_{ST} > 0.25$ indicates very great differentiation. However, $F_{ST} < 0.05$ and even lower should not be considered as negligible (Wright, 1978, cited in Hartl, 1980).

2.7. Gene flow and its consequences in natural plant populations

Rittner and McCabe (2004) defined gene flow as "The exchange of genes between different but usually related populations." Domesticated plants and their wild relatives often hybridize and give rise to viable and fertile progenies if requirements for successful mating are met (Ellstrand *et al*., 1999). Such hybridization may lead to gene flow. Gene flow can occur when an individual or groups of individuals migrate from one population to another and interbreed with its members (Rittner and McCabe, 2004). Gene exchange between crops and their wild relatives is a natural phenomenon and has a great importance in the study of ecology and biodiversity. Effective crop-to-wild gene flow depends on the completion of a number of successive steps required for gene dispersal (Ellstrand *et al*., 1999; Jenczewski *et al*. 2003, Gepts and Papa, 2003): (i) the presence of crops and close relatives growing nearby, (ii) their biology and phenology in relation to pollen and seed dispersal, (iii) the production of viable and fertile F_1 hybrids, (iv) the production of fertile successive generations, (v) the opportunity for gene transmission, chromosome recombination and crop gene introgression into the wild genetic background, and ultimately (vi) the persistence of introgressed crop genes in natural communities. The crop sorghum and its wild relatives are sympatric virtually everywhere the crop is cultivated, and the flowering time of the two plants overlap (Holm *et al*., 1977).

While gene flow from the wild species to the crop might lead to reduced yield and loss of the genetic purity of the cultivated varieties (Felber *et al*., 2007), gene flow in the reverse direction and introgression of crop alleles into wild germplasm could lead to the

contamination of wild populations, which in turn could lead to increased weediness and/or the erosion of genetic diversity of wild populations. Crop-to-weed gene flow may lead to significant evolutionary flux in the recipient populations (Anderson, 1949). Gene flow will most likely affect the organization of genetic diversity in a crop's primary gene pool (Harlan and de Wet, 1971). Because crops are typically genetically impoverished compared to their wild relatives (Ladizinsky, 1985), extensive gene flow from the crop is expected to deplete genetic diversity in wild populations (Ellstrand *et al*., 1999; Bartsch *et al*. 1999). In the worst case scenario this can lead to the extinction of certain wild crop relatives (Small, 1984) in plants mainly due to outbreeding depression and swamping (Ellstrand *et al*., 1999). Also, it may have important practical and economic consequences if it promotes the evolution of more aggressive weeds (Barrett, 1983) perhaps facilitated by fitness-enhancing transgenes (Snow *et al*., 2005). Ellstrand *et al*. (1999) found that crop-to-weed gene flow has been implicated in the evolution of enhanced weediness in wild relatives of seven of the world's 13 most important crops. Consequently, the escape of transgenes through hybridization of crops with wild relatives in their centers of origin has been cited as one of the potential environmental risks of transgenic crops (Hails, 2000).

Despite its obvious evolutionary significance, crop-to-weed gene flow has received little experimental attention (Jenczewski *et al*., 2003). Identification and quantification of the risks associated with gene flow, details of how to measure rare hybrid events and the ecological fitness costs require knowledge of the mechanisms of pollen dispersal and hybridization events. Population genetic studies are useful for determining how genetic

diversity is partitioned within and among populations of native plants and for inferring levels of gene flow among populations. This baseline knowledge provides a starting point for designing *in situ* and *ex situ* conservation programs (Rice, 2004; Ellstrand and Elam, 1993).

Analyzing six non-GE maize hybrids, Mellon and Rissler (2004) found transgenes in 50% of the samples at levels from 0.1 to 1%. Messeguer *et al*. (2004) studied the frequency of pollen-mediated gene flow from a transgenic rice line to the red rice weed and conventional rice in the Spanish *japonica* cultivar Senia and found gene flow in conventional rice planted at 1, 2, 5 and 10 m distance and there was a clear decrease with increasing distance.

So far, only limited studies dealt with investigating the risk of gene flow in sorghum. Schmidt and Bothma (2006) carried out direct gene flow studies from crop to crop using non-transgenic male sterile plants as pollen recipients. Arriola and Ellstrand (1997) studied crop-wild gene flow. Tesfaye Tesso *et al*. (2008) surveyed crop and wild sorghum growing regions in parts of Ethiopia. Mutegi *et al*. (2011) also studied indirect estimates of gene flow from crop to wild sorghum in Kenya. Muraya *et al*. (2010, 2011) investigated the risk of trans-gene flow through estimation of outcrossing rate, pollen competition, and wild-crop sorghum hybrid fitness in Kenya and Mali. All of the above studies reported that gene flow from cultivated to wild sorghum is likely to occur.

2.7.1. Estimation of gene flow

Knowledge about gene flow is primarily needed to evaluate whether novel transgenes could unintentionally enter wild or weedy populations and confer fitness-enhancing traits that could exacerbate weed problems. Generally, there is interest in knowing whether transgenes are likely to spread from crop-to-crop, crop-to-weeds, and crop-to-wild relatives, especially where landraces and wild or weedy relatives represent valued germplasm for future crop breeding. Regardless of whether transgenic or nontransgenic cultivars are grown, information about the potential for crop-to-wild gene flow is needed to understand its effects on the genetic diversity of wild populations (Ellstrand, 2003).

The extent of gene flow is used to determine the extent to which different populations of a species are independent evolutionary units (Slatkin and Maddison, 1989). Determining the level of gene flow between crops and their wild relatives also is critical for conserving wild germplasm (Ellstrand *et al*., 1999; Jarvis and Hodgkin, 1999). Gene flow can be estimated using direct and indirect methods. While direct methods, which use estimates of dispersal distances and breeding success of dispersers, can be used to assess the ongoing gene flow, the indirect methods, which are mainly based on spatial distributions of allele frequencies and DNA sequence variations, can be used to infer past gene flow (Slatkin, 1987). The direct methods use direct observation of dispersal by estimating paternity from progeny seed set (Ellstrand *et al*., 1999). An indirect method of estimating gene flow is one that uses observed spatial distribution of alleles, chromosomal segments, or phenotypic traits to draw inferences about the level or pattern

of gene flow in a population (Slatkin, 1985). Two indirect methods of estimating gene flow and population structure are widely used in the literature. One uses Wright's (1951) F_{ST} and the other uses its analogue Slatkin's (1995) R_{ST} (Balloux and Lugon-Moulin, 2002). Fixation index (F_{ST}), computed from molecular marker allele frequency data is commonly used to investigate the role of migration in the genetic structure of a species. Wrigh (1951) introduced the following relationship between fixation index (F_{ST}) and the number of migrant per generation (N_m) under the infinite allele model:

$$F_{ST} = \frac{1}{1 + 4Nm}$$

Which could be rewritten as:

$$N_m = \frac{1 - F_{ST}}{4F_{ST}}$$

F_{ST}=0.2 is the turning point from gene flow to drift, which is equivalent to one migrant per generation (Nm =1). When F_{ST} is greater than 0.2 (N_m<1), it is genetic drift and not gene flow which governs the genetic structure of populations (Halliburton, 2004).

Slatkin and Barton (1989) compared F_{ST}, rare alleles, and maximum likelihood as the three indirect estimates of gene flow and concluded that F_{ST} is more useful under realistic conditions. F_{ST} can be computed from data on relative frequencies of molecular marker alleles in adjacent populations. Both F_{ST} and R_{ST} based methods of computing the

number of migrants per population use similar formula. The distinction between them is that F_{ST} assumes infinite allele model (IAM) whereas R_{ST} assumes the step-wise mutation model (SMM) (Balloux and Lugon-Moulin, 2002). The limitation of R_{ST} is its high variance while the limitation of F_{ST} is its bias (Balloux and Goudet, 2002).

2.7.2. Mating systems as determinants of gene flow in *Sorghum bicolor*

Around 40% of the world production comes from crops pollinated by the wind (Ramsay, 2005). Information on the mating systems and probable introgression of transgene into wild relatives and its economic and ecological consequences is still rare despite increasing study (Schmidt and Bothma, 2006). Mating system is one of the most important life history traits of a plant species that determines gene exchange (Hamrick and Godt, 1990) and plays a major role in shaping plant population genetic structure (Brown and Allard, 1970; Ritland, 1983; Enjalbert and David, 2000). It is also a key parameter in plant evolutionary studies (Brown, 1979). The dynamics of genetic diversity may be influenced by mating system, which in turn is influenced by different farming practices and environmental factors (Barnaud *et al*., 2008). Cross pollination stimulated most attention perhaps because of the uncontrolled and involuntary nature of the process, and this source of impurity has received much publicity in scientific, public and political forums in recent years (Ramsay, 2005). Thus, worldwide, government regulatory agencies require information about the extent to which a transgenic crop will cross-pollinate with its wild or weedy relatives (Muraya *et al*., 2011b).

Attention is given to the extent of crossability between cultivated plants and their wild relatives as the crosses have been extensively used in plant breeding (Jenczewski *et al*., 2003). Spontaneous hybridization between diploid *S. bicolor* and tetraploid *S. halepense* is frequent (Arriola and Ellstrand, 1996). The introgression of crop sorghum genes into johnsongrass (*S. halepense* (L.) Pers.) has been implicated in the increased weediness of Johsongrass (de Wet and Harlan, 1975; Holm *et al*., 1977; Pedersen *et al*., 1998). Crop-weed hybridization has been involved in the evolution of numerous weeds (Barrett 1983, Small, 1984), and weedy sorghum subspecies such as shattercane (*S. bicolor* ssp *drmmondii*) already exist (Defelice, 2006; Gebisa Ejeta and Grenier, 2005). It has been suggested that all *S. bicolor* ssp *bicolor* races can freely hybridize with wild *S. bicolor* ssp. *verticilliflorum* to produce shatter canes (*S. bicolor* spp *drummondii*) (Defelice, 2006; Dahlberg, 2000).

In general, prior to widespread release of new transgenic germplasm, the potential of outcrossing between cultivated sorghum and its wild relatives should be assessed (Polowick *et al*. 2000). Brown and Allard (1970) first applied isozyme polymorphisms as markers for the estimation of mating system parameters in maize and later Ritland and Jain (1981) developed a computer program to analyze them. Since then the method of computing mating systems has been employed in many plant species following the advent of more polymorphic genetic markers such as microsatellites. Mixed mating model is based on various assumptions. For instance Brown and Allard (1970) suggested three basic assumtions: 1) deviation of the mating system from random mating is solely caused by excessive self fertilization, 2) multilocus outcrossing rate is uniform across the

maternal genotypes and pollen alleles (no incompatibility and independent mating events), and 3) pollen is uniformly distributed over the maternal genotypes. Mating can also vary among genotypes and among environments and can be influenced by genotypic frequencies (e.g., Harding and Tucker, 1964). However, as it is difficult to estimate all of these factors, Brown and Allard (1970) suggested to estimate t=1-s (where t=outcrossing rate, s=selfing rate) as the simplest and more practical procedure, which summarizes the complex set of biological phenomena in terms of random outcrossing and selfing.

Two strategies have been used to estimate the outcrossing rates of a mixed mating population from genetic markers: measuring the frequency of heterozygous genotypes and the analysis of progeny arrays (Ritland 1983). Using the first approach, Wright (1969) introduced a simple relationship between the within population inbreeding coefficient (f) and outcrossing rate (tf) as$f = \frac{(1-t)}{(1+t)}$. Such relationship holds under assumptions that selfing is constant for sufficient number of generations; the population is in inbreeding equilibrium; and the major cause of departure from Hardy-Weinberg equilibrium is selfing (Enjalbert and David, 2000).

Knowledge of the level of outcrossing under current field conditions is highly desirable for designing effective *in situ* conservation of wild and domesticated plants (Ritland and Jain, 1981) and because it can be modified by phenology, field size and planting practices (Barnaud *et al.*, 2008).

2.7.3. Seed as a vehicle for gene flow

Although both pollen and seed are the two major vehicles of gene flow, they are basically related to each other as volunteer seeds can grow and contribute to pollen mediated gene flow (Nielson *et al*., 2009). Spatial seed dispersal by natural means is often considered to be low because most crops have lost their ability for independent seed dispersal (De Wet and Harlan, 1975). However, seed dispersal can be aided by animals, water, wind, humans via agricultural implements, seed marketing, application of cattle manure, etc. This dispersal can be an important step in the dynamics of crop-wild gene exchange (Jenczewski *et al*., 2003). The initial level of seed dormancy, longevity and the response of seed germination with respect to environmental flux can influence the potential temporal dispersal of transgenes (Linder and Schmitt, 1994). If gene dispersal is known to have occurred via seed, the next step is to study the persistence and further establishment (Jenczewski *et al*., 2003). The role of soil seed bank dynamics in seed mediated gene flow is quite substantial (Nielson *et al*., 2009) as soil seed banks are the reservoirs of genetic material upon which natural selection can act (Simpson *et al*. 1989). Therefore, studying this dynamics is crucial for risk assessment of trans-gene flow.

Studies in natural plant populations have documented that the response of genotypes to different environments include the formation of genetically distinct populations (genetic differentiation), changes in response (phenotypic plasticity) or both (Hermanutz and Weaver, 1996). If populations have an adaptive response to their native environment, then local specialization will occur and phenotypic and genetic differentiation will be observed. Populations of a plant species that are exposed to different environments may

develop intraspecific genetic variation or ecotypic differences over time (Darwin, 1859). Ecotypes are plants of the same species having distinctive heritable characteristics that are often associated with environmental conditions, and are widely regarded as a mechanism to cope with heterogeneous environments (Vaughton and Ramsey, 2001). Allard (1965) suggested that self-fertilizing species are comprised of a large number of local populations with distinctive morphological and physiological characteristics, such that sharp differences may occur within a short distance in relation to environmental change. This was thought to be due to predominance of homozygous loci, but with some heterozygosity being present in a population which allowed for the segregation of better adapted genotypes (Bennett, 1999).

One of the important features of seed dormancy is to maintain a reservoir of soil seed bank for future re-establishment following unfavorable conditions (Monaghan 1979). It is one of the major invasion and persistence mechanisms of weeds in agricultural fields (Narwal *et al*., 2008). An important aspect frequently overlooked in studies of seed dormancy is the possibility that populations in different environments will possess differing seed dormancy strategies (Bennet, 1999). So far, no study has assessed whether wild/weedy sorghum has been able to evolve differences in their seed dormancy and longevity patterns enabling adaptation to particular environments in its center of origin, yet such knowledge may be useful to understand the weediness potential of different wild sorghum populations and also to design effective management methods. Environmental factors affecting germination, particularly whether or not exposures to high soil temperatures or soil moisture regimes are required for germination have not been

determined. In Ethiopia, the range of environments where wild sorghum persists is wide. Sorghum grows in the highlands and midlands as well as in the lowlands, where rainfall is short and erratic. It is, thus, proposed that wild sorghum in different environments has adapted differently in their response to different seed germination requirements at that site.

Weed seeds buried in the soil have been of concern to agriculturalists (Omami *et al*., 1999). Soil seed banks constitute a key component of plant population dynamics, community structure and annual life-cycles because they influence population size within and between years. Also, seed characteristics such as longevity and dormancy enable populations to survive for several years (Baker, 1974). The seed dormancy properties can vary among ecotypes (Monaghan, 1979). Reports of researches done in different species and ecotypes showed variable results for persistence of seed in the soil. The duration of loss of viability in buried seeds may range from days in the seeds of the composites (Marks and Nwachukut, 1986) to nearly 20 years in American dragonhead, common lambsquarters, Pennsylvania smartweed, and shepherd's-purse (Conn *et al*., 2006). In between reports also vary with species and depth of burial. Narwal *et al*. (2008) found total loss of viability of ryegrass seeds after 16 months of burial at 5 and 10 cm depth in different soil and simulated rainfall conditions. Noldin *et al*. (2006) compared the seed longevity of several ecotypes of wild rice from 4 US states and found <1% viable seeds in 3 ecotypes at 5cm depth, but 9 ecotypes had similar proportion of viable seeds at a depth of 25cm after 2 years of burial at 2 locations in Texas, but within each depth there was minor variation among ecotypes. Seed banks can serve as pools of genetic material,

enabling a range of responses to environmental variability, and can buffer populations against temporarily extreme adverse conditions (Teo-Sherrell *et al*., 1996). Seed-bank persistence between years also allows plants to average their success over time, by reducing both opportunity and risk (Cohen, 1966). Dormancy and the timing of germination are also dependent on the environment because they are highly responsive to environmental conditions and may strongly influence the evolution of post-germination life-history characters beyond their role in reducing risk (Donohue, 2002). Benvenuti (2003) found that germination inhibition due to burial depth was found to be directly proportional to clay content and inversely proportional to sand content. Soil temperature is known to strongly influence both percentage and time of germination of sorghum seed (e.g., Kanemasu *et al*., 1975). These authors suggested 23°C to be an optimum soil temperature for germination of sorghum seed.

The study of seed dormancy and longevity is useful because it is a fundamental element for determining seed bank dynamics (Van Mourik *et al*., 2005). As a result, characterization of the germination ecology of species and ecotypes would help to indicate their potential to become noxious weeds in different systems of cultivation (Buckley *et al*., 2004). Therefore, knowledge of seed longevity and dormancy can help weed management (Martinez-Ghersa *et al*., 1997). Dunbabin and Cocks (1999) reported large differences in the seed dormancy pattern of two ecotypes of capeweed (*Arctotheca calendula*) in Western Australia. Studies related to ecotypic variation in wild plants are considered to be essential for conservation of biodeiversity (Crandall *et al*., 2000). Conservation of this diversity should, thus be maintained in order to avoid extinction

associated with environmental change (Fraser, 2000). Dormancy is known to be controlled by many genes (Li and Foley, 1997).

2.8. Transgenics in sorghum

Since the early development of transgenic plants in the 1980s, tremendous accomplishments associated with transgenic biotechnology have been achieved (Huang *et al.* 2002). A great number of genetically modified (GM) crops have been released and some have been commercialized (Huang *et al*., 2002). In 2011, 160 million hectares of land was covered by genetically modified crops in 29 countries, of which 19 were developing and herbicide tolerance has been consistently the dominant trait of the genetically modified crops since the first commercialization of transgenic crops in 1996 (James, 2011). For the past 16 years, about 99% of the total GM crop land was dominated by 4 major crops: Soybean (47%), Maize (32%), Cotton (15%), and Canola (5%) (James, 2011).

Because sorghum is the 5th most economically important cereal in the world, it is one of the candidates for genetic transformation (Arriola and Ellstrand, 2002). Transformation of sorghum was started in the early 1990s via protoplast electroporation (Hagio *et al*., 1991) and shotgun (Battraw and Hall, 1991) methods. The first successful fertile sorghum transformation results were reported by Casas *et al*. (1993) and Zhao *et al*. (2000) via particle bombardment and Agrobacterium-mediated transformation systems, respectively. Unlike most cereals which were successfully transformed, the inefficiency in the genetic

transformation of sorghum has been suggested to be its recalcitrant nature for transformation (O'Kennedy *et al*., 2006).

Genetic transformation of sorghum has been underway for different characters using *Agrobacterium tumefaciens* mediated (e.g., Jeoung *et al*., 1999; Zhao *et al*., 2000) and biolistic bombardment (e.g., Zhu *et al*., 1998) methods. Liang and Gao (2001) successfully generated transgenic sorghum plants with a gene from rice, *chi 11* encoding chitinase for resistance to Fusarium stalk rot, and another gene, also from rice, *tlp*, encoding thaumatin –like protein for drought tolerance and disease resistance. The Africa Biofortified Sorghum (ABS) project developed the first generation transgenic sorghum line (ABS #1) that possesses grain with a 50% increase in lysine (Zhao *et al*., 2003). Another nutritionally enhanced sorghum named as ABS #2 which overcomes the nutritional deficiencies by substantially improving grain digestibility, by delivering vitamins, the essential amino acids lysine, threonine and tryptophan, and by improving the bioavailability of iron and zinc is under development (Zhao, 2007).

3. MATERIALS AND METHODS

3.1. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (l.) Moench] in Ethiopia

3.1.1 Geographical distribution and *in situ* phenotypic diversity

3.1.1.1 Germplasm collection sites

Ethiopia is divided into 32 agroecologies based on the amount of annual rainfall, temperature and length of growing period (MOARD, 2005). Oromiya, Amhara and Tigray regional states are the three major producers of sorghum covering 86.4% of the total area and 89% of the total production in the last 9 years (CSA 2003-2011). To study the geographical distribution and phenotypic diversity of wild and weedy sorghum, survey was made from October through November 2008 to five geographical regions representing the above regional states and the Benishangul Gumuz regional state (Metekel zone), belonging to different agroecologies. Collection and *in situ* measurements were made at 31 survey sites which fell in 8 of the 32 agroecologies (Table 3.1). These were A2=Warm arid lowland plains, SA2=Warm semi-arid lowlands, SM2=Warm sub-moist lowlands, SM3=Tepid sub-moist mid highlands, M2=Warm moist lowlands, M3=Tepid moist mid highlands, SH2=Warm sub-humid lowlands, and SH3=Tepid sub-humid mid highlands. The major characteristics of the collection sites are described in Table 3.1.

Table 3.1 Geographical characteristics of the sites where wild sorghum collection was made in 2008, and cultivated sorghum was collected in 2009

Geographical region[§]	Population code*		Name of location	Long. (N)	Lat. (E)	Alt. (masl)	Agro-ecology
	Wild	Cultivated					
Ghibe	G-1(M)		Bhede-Dhero	8º 10’	37º 33’	1686	SH2
(Oromiya)	G-2(I)		Berha/Sateri	8º 11’	37º 33’	1425	SH2
	G-3 (I)		Ghibe-Aqorbach	8º 12’	37º 33’	1261	SH2
	G-4(I)	GC2	Along Ghibe river bridge	8º 13’	37º 34’	1115	SH2
	G-5(M)	GC1	Ghibe-ILRI	8º 14’	37º 34’	1149	SH2
	G-6(M)		Boqata and Serti	8º 15’	37º 37’	1607	SH2
Hararghe	H-1(M)		Bisidimo	9º 12’	42º 13’	1386	M3
(Oromiya)	H-2(M)		Kara (Gerbi)	8º 52’	40º 42’	1709	SH3
	AW(I)		Awash National Park	8º 56’	40º 05’	1044	A2
Metekel/Pawe	P-1(I)		Kanensan	11º 01’	36º 27’	1682	SH3
(Benishangul Gumuz)	P-2(M)		Pawe Agricultural Research Center	11º 18’	36º 24’	1118	M2
	P-3(M)	PC2	Village-5	11º 16’	36º 23’	1040	M2
	P-4(I)		Village-3	11º 13’	36º 21’	1045	M2
	P-5(M)		Village-7	11º 21’	36º 23’	1076	M2
	P-6(M)	PC1	Mandura	11º 05’	36º 25’	1404	SH2
Tigray	T-1(M)	TC1	Alamata-Gerjele	12º 26’	39º 36’	1462	SM2
(Tigray)	T-2(M)		Bidajira	12º 40’	39º 38’	1633	SM3
	T-3(M)		Adenkel	13º 22’	39º 25’	1925	SM3
	T-4(I)		Subday	14º 21’	38º 10’	1808	SM3
	T-5(M)		Humera Agricultural Research Center	14º 15’	36º 37’	609	SA2
	T-6(M)		Humera-Donkey	14º 14’	36º 41’	631	SA2
	T-7(M)		Golel	14º 08’	36º 53’	730	SA2
	T-8(M)		Hagereselam-Idris	14º 05’	36º 58’	820	SA2
Wello	W-1(M)	WC1	Jara Kechema	10º 30’	39º 56’	1433	M3
(Amhara)	W-2(I)		Hijira	10º 38’	39º 55’	1425	M3
	W-3(M)		Milamile	10º 45’	39º 49’	1436	M3
	W-4(M)	WC2	Abuare	12º 05’	39º 39’	1426	SM2
	W-5(M)		Aradom	12º 06’	39º 38’	1460	SM2
	W-6(M)		Zobel	12º 11’	39º 46’	1712	SM3
	W-7(M)	WC3	Kobo town side	12º 08’	39º 37’	1500	SM2
	W-8(M)		Ayub	12º 14’	39º 36’	1467	SM2

[§] Names of regional states in parentheses, *I= Isolated (>500m away from the crop), M= Mixed with the crop (intermixed or found <500m from the crop)

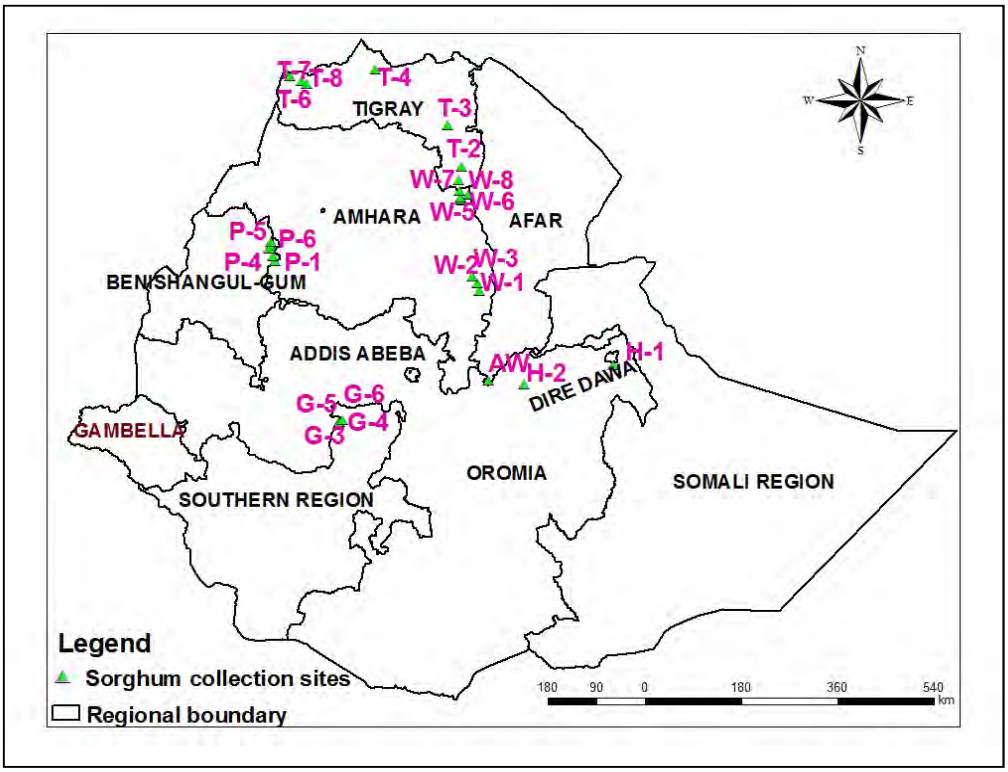


Fig.3.1 Regional map of Ethiopia showing the wild sorghum collection sites

Table 3.2 Agro-climatic characteristics of the five regions where collection was made*

Geographical region	Average temperature (ºC)	Annual rainfall (mm)	Dominant soil type	Dominant crops grown
Ghibe	25	1000-2000	Lithosols, eutric regosols, orthic acrisols	Sorghum, maize
Hararghe	21	800	Rendzina	Sorghum, tobacco, mango, chat
Metekel/Pawe	32.5	1591	Vertisols, Fluvisols, Regosols, Leptosols	Sorghum, finger millet, maize
Tigray	21-33	700-1200	Eutric cambisols, vertic luvisols, vertisols	Sorghum, finger millet, sesame, cotton
Wello	16-20	850	Vertisols, Eutric cambisols	Sorghum, maize, tef

* Part of the data were adapted from MOARD (1998)

Table 3.3 List of wild sorghum accessions acquired from ICRISAT

Subspecies	Race	Source of accession	Accession No.
Verticilliflorum	*arundinaceum*	South Africa	IS-14300
Verticilliflorum	*arundinaceum*	Swaziland	IS-14315
Verticilliflorum	*arundinaceum*	Angola	IS-14232
Verticilliflorum	*virgatum*	USA	IS-18804
Verticilliflorum	*virgatum*	Egypt	IS-18810
Verticilliflorum	*verticilliflorum*	Sudan	IS-18864
Verticilliflorum	*aethiopicum*	Sudan	IS-14485
Verticilliflorum	*aethiopicum*	Sudan	IS-18822
Drummondii		Sudan	IS-18920
Drummondii		Ethiopia	IS-11490

3.1.1.2 Germplasm collection, acquisition and study of *in situ* phenotypic diversity

In this study the 3 sub-species recognized by de Wet (1978): ssp. *bicolor*, ssp. *drumonddii*, and ssp. *verticilliflorum* were used and all non-cultivated sorghum were refered to as “wild” as in Aldrich *et al*. (1992) and Mutegi *et al*. (2011). A total of 600 wild/ weedy plants from 30 populations, representing the five regions of collection (Fig. 3.1 and Table 3.1) were measured *in situ* to estimate phenotypic diversity. There are no reports about the existence of wild species other than the close relatives belonging to the same species *S. bicolor* in Ethiopia. Hence, during the collection trip the roots of at least two plants in each population was excavated and checked for the presence of perennial rhizomatous species. Of the 31 populations 23 were found intermixed or within 500m

distance from cultivated sorghum (labeled as M in Table 3.1). The remaining 8 populations were found far from any crop (Awash and Ghibe-Bereha populations) or weeds in maize, barley or tef (labeled as I in Table 3.1).

Sampling was done at 2 to 8 sites in each geographical region (Table 3.1), and each site was considered to represent a population. At Ghibe, six populations were collected. Collection was also made at two locations in South Wello and six locations in North Wello. The weak culm, small grass-like wild sorghums were found in North Wello from Kobo to Alamata. This kind of wild sorghum plants were also collected at three other locations in Tigray (T-1, T-2, T-3) and one location in Hararghe (H-2). They are also distributed in Central and East Shewa (also Brhane Gebrekidan, 1981, cited in Dogget, 1988). Such kind of wild sorghums were found either in tef or in barley fields and not in sorghum or in maize fields. Moreover the compact head weedy sorghums were found in Wello (at Zobel, W-6), the highest elevation (1711m) of all the Wello collection sites. Farmers around that area grow compact headed durra type sorghum landraces. At Abuare and Aradom (W-4 and W-5), weedy sorghums were found in the improved sorghum variety 76T1#23 in private commercial farms.

There were two geographically distinct regions of collection in Tigray. The first extends from the Alamata plain adjacent to Wello (south Tigray) to Adenkel, 15Km away from Mekele town towards Tenben (Central Tigray). The second region extends from Hagere-selam to Humera bordering Sudan and Eritrea (western Tigray). In this area, both compact (mainly in T-7) and loose panicle wild and weedy sorghum types were found in

sorghum fields. However, unlike the Zobel (W-6) area the compact panicle weedy sorghums in those areas were morphologically distinct from the caudatum landraces.

The wild sorghum populations at the Awash National Park were found to be very interesting for this study because they were more than 30 Km away from places where the crop sorghum was grown. However, phenotypic measurements were not recorded as the seeds were shattered during arrival. In Beneshangu-Gumuz, six sites were sampled from the border of Amhara regional state to village 7 including the Pawe Agricultural Research Center close to the abandoned airport where Amsalu Ayana *et al.* (2001) reported.

In addition, 10 accessions of wild sorghum, which were originated in 7 countries and identified by sub-species and races, were acquired from ICRISAT.

3.1.2. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (l.) Moench] using SSR markers

Samples were collected from 19 *in situ* wild sorghum populations (AW, G-2, G-4, G-6, H-1, H-2, P-1, P-4, P-5, P-6, T-1, T-4, T-5, T-6, T-8, W-2, W-4, W-6, and W-7) for this study (Table 1, Fig. 1). The distance among populations was calculated from global positioning system (GPS) coordinates and ranged from 3.5 km (G-2 and G-4) to 828 km (H-1 and T-5).

In November 2009, DNA samples were also collected from 20 plants in each of eight cultivated sorghum populations that were planted near wild populations shown in Table 3.1, in order to examine evidence for historical crop-wild gene flow. Four cultivars were local landraces of durra, caudatum or durra-caudatum, and one, WC2, was an improved cultivar of caudatum ($76T_1$#23). Two of the durra cultivars, at WC3 and TC1, were known locally as "Degalit".

To compare wild sorghum populations with earlier accessions, 10 wild sorghum accessions originated from at least 6 African countries were acquired from ICRISAT (International Crops Research Institute for the Semi-Arid Tropics; Table 3.3). The goal of including these materials was to determine whether the unidentified Ethiopian wild sorghum collections clustered with any of the accessions according to the sub-species and races.

3.1.2.1 DNA extraction

DNA was extracted in two steps: *in situ* collection of leaf squashes from live plants using the Whatman FTA cards and Whatman chromatography paper in Ethiopia, and sample purification in Aronoff laboratory following the manufacturer's procedure and optimized for sorghum by Asfaw Adugna *et al*. (2011) at Ohio State University.

3.1.2.1.1 *In situ* squash collection

Samples were taken from the second or the third leaf, whichever was more succulent and free from disease lesions. Flag leaves of mature plants did not produce adequate squashes and thus were not used. A good time for sqash collection is until 10:30 in the morning. Whatman chromatography paper cards were prepared by folding 3cm × 10cm pieces of Whatman chromatography paper in half. Before collecting squashes, cards were labeled according to the codes of the region of collection, population number and individual number. The cards were placed in polytene bags and placed on a wooden board covered with polytene sheet. A piece of the leaf was placed in between the polyethylene bag and the appropriate section of the card with the back of the leaf facing towards the card. As suggested by the manufacturer (Long *et al*, 2008), leaf squashes were taken onto the cards by pounding with a pestle until the card was stained. Alternately, a shoe horn was used to rub the leaves until the squash was well soaked from the back of the card. Cards were allowed to air dry for a minimum of 1 hour under shade before storage. Materials (scissors, the polythene sheet covering the crushing board and shoe horn/pestle) were rinsed or wiped with 70% ethanol between each sample to avoid cross-contamination. Moreover, a new polythene bag was used for each extraction.

3.1.2.1. 2 Sample purification

This step was carried out in Aronoff Stanley J. laboratory at the Department of Evolution, Ecology and Organismal Biology, Ohio State University, Columbus, Ohio, USA. A single hole paper punch was used to remove a 6mm disc from the center of the dried sample area on the FTA® and/or Whatman card. The disc was transferred to a 1.5ml

microfuge tube. In order to ensure no residue was carried over to the next sample, the punch tip was triple rinsed in distilled water, wiped with a tissue paper, and punched onto a blank chromatography paper between samples. After adding 400µl of FTA® purification reagent (Whatman Inc., USA) to each tube, the tubes were vortexed and incubated for 4 minutes at room temperature. Used FTA® purification reagent was discarded using a pipette so that the disc remained in the tube. The FTA® purification reagent wash was repeated one more time. The disk was then rinsed with 400µl of a modified TE^{-1} (10 mM Tris, 0.1 mM EDTA) buffer in a similar manner. The TE^{-1} buffer rinse was repeated once. Using tweezers, the disk was transferred to a 0.5ml microfuge tube and 80µl of TE (10mM Tris, 1mM EDTA) was added. After centrifugation, the disk was incubated in the buffer for 5 minutes at 95°C, cooled on ice, and stored at -20°C. To check the quantity and quality of the DNA, samples were loaded with markers on 3% super fine resolution (SFR) agarose gels stained with ethidium bromide and viewed under UV transilluminator.

3.1.2.2. Polymerase chain reaction (PCR) and SSR fragment detection

Polymerase chain reactions (PCR) were run using 12 sorghum microsatellite loci that were previously mapped (Brown *et al*., 1996; Taramino *et al*., 1997; Bhattramakki *et al*., 2000; Li *et al*., 2009) and represented all of the 10 sorghum linkage groups (Table 4.5). These loci were selected based on high polymorphism and compatibility for multiplexing. Three loci (Sb6-342, Sb4-15, and Sb5-236) failed to amplify in many samples from 4 of the 19 wild populations (populations AW, H-2, T-1, and W-7) and hence were excluded from the final analysis (Table 4.1).

Polymerase chain reactions followed the QIAGEN® multi-master mix kit protocol for SSR multiplex, and forward primers were labeled with different fluorescent colors. PCR was carried out in 12µl total volume of reaction mix containing 1µM of each primer pair in a multiplex, 1µl of template DNA, 2.6µl of sterile ddH_2O, 6µl of QIAGEN® Multiplex PCR 2X Master mix. Polymerase chain reactions were run in a Master cycler (Eppendorf™) with an initial denaturation step of 15 min at 95°C, followed by 35 cycles of 30 sec at 94°C, 90 sec at 58°C, 60 sec at 72°C, 30 minutes at 60°C, and held at 4°C following QIAGEN® protocol for microsatellite multiplexes.

For fragment-size determinations, 2 µl of the PCR product was diluted with 14 µl of ddH_2O and then 2 µl of the diluted PCR product was added to 14 µl of 36:1 Hi-Di-Formamide: GenScan™/350 Rox™ size standard in a 96 well microtitter plate and was denatured at 95 °C for 5 minutes and cooled on ice for at least 5 minutes. Fragment analysis of alleles was done by ABI 3100 Genetic Analyzer (DNA sequencer) and sizes were read using the associated GeneMapper 3.7 software (Applied Biosystems Inc., CA, USA). To verify the repeatability of allele scoring, samples of a well-studied sorghum inbred line, BTx623, were included by obtaining seeds from the Department of Agronomy, Purdue University. To exclude the possible effects of imprecise DNA fragment sizes due to stuttering, large allele drop out, or null alleles on genotyping, the software Allelobin (Prasanth *et al.*, 2006) was used to classify observed SSR allele sizes into representative discrete allele sizes using a variation of the least-square minimization algorithm of Idury and Cardon (1997).

3.2 Estimation of mating system parameters in *in situ* wild/weedy sorghum using SSR markers

For this part of the study, populations G-2, H-1, P-6, T-6, T-8, W-4 and W-7 were used (Table 3.1). From each maternal plant 8 progeny plants (hereafter referred to as family) were sampled for DNA extraction. Hence, a total of 1120 progeny plants were included in the experiment. In other words, each population was represented by 20 families, which is equivalent to 160 progeny plants per population. Moreover, the 140 maternal plants of the sampled progenies were included. This set of 1260 plant samples was assumed to be adequate to represent all of the regions of collection. Sampling from the geographically diverse regions was aimed at assessing ecotypic variation on outcrossing rate and/ or for the possible inclusion of more subspecies and races prevailing in different regions.

3.3. Morphology and fitness of wild × crop F_1 hybrids of sorghum

3.3.1. Plant materials and crossing design

The list of the plant materials used in the study is presented in Table 3.4. Twenty-three wild sorghum lines and two released cultivated sorghum varieties were involved for initial crossing. Crossing was made in 2010 main rainy season at Melkassa Agricultural Research Center (39°21′E, 8°24′N, altitude=1550m, located in the Rift Valley) by hand emasculation and pollen shading. The wild sorghum lines served as female parents and the pollinators were the two improved sorghum varieties, 76T1#23 and WSV 387 (Melkam) as this was the natural pattern of gene flow intended to be investigated. The

crop varieties belong to race caudatum and are early maturing, recommended for production in lowlands of Ethiopia similar to the places of collection of most of the wild sorghum parents used in this study. Three of the wild parents were received from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India. Out of the ICRISAT wild parents, two belong to race *aethiopicuum* (originated from Sudan) and one belongs to race *virgatum* from USA. The remaining parents were Ethiopian collections in November 2008 from Hararghe and Tigray belonging to race *arundinaceum* and from Wello which had unknown grouping (most probably belonging to the *ssp drummondii*). The Ethiopian collections were grown in progeny rows (head-to-row). The seeds of all of the experimental parents were multiplied under the same management at the same field where crossing of hybrids was made. The harvested seeds were kept in the laboratory at room temperature until the beginning of the experiment.

3.3.2. Setup of the laboratory and field experiment

Germination test was carried out just after sowing the field experiment for those entries which had adequate amount of remnant seeds. Before the test the glumes were removed with fine sand paper to facilitate germination. Twenty-five seeds of the hybrids and their parents were germinated on 115mm petridishes lined with number one Whatman® filter paper moistened with distilled water in 3 replications. Each replication was represented by a compartment in the incubator. The petridishes were kept at 30±3°C constant temperature in the dark. Distilled water was added as needed.

Table 3.4 Hybrids and their parents involved in the experiment

Entry No.	Genotype	Ssp./Race	Source
	F1 hybrids		
1	H2-1 × WSV 387		
2	H2-16 × 76T1#23		
3	T1-1 × WSV 387		
4	W5-20 × 76T1#23		
5	IS18822 × WSV 387		
6	IS14485 × 76T1#23		
7	IS18804 × 76T1#23		
	Wild parents		
8	H2-1	*arundinaceum*	Hararghe collection
9	H2-16	*arundinaceum*	Hararghe collection
10	T1-1	*arundinaceum*	Tigray collection
11	W5-20	ssp. *Drummondii*?	Wello collection
12	IS18822	*aethiopicum*	ICRISAT/Sudan
13	IS14485	*aethiopicum*	ICRISAT/Sudan
14	IS18804	virgatum	ICRISAT/USA
	Cultivar parents		
15	76T1#23		MARC-released
16	WSV 387		MARC-released

The field experiment was carried out at Melkassa Agricultural Research Center (MARC) during the main rainy season of 2011 (from 29 June to 29 October 2011). The amount of rainfall during the season was 589.7mm. This amount was largely distributed in the first three months (June-July=131.2mm, August=208.8mm, and September=197.5mm). There was no rainfall in October. The soil at the experimental site is silty clay loam Andosol with pH=7.8. Out of a pool of crosses made in 2010, seven F1 hybrids were selected for this study based on the availability of adequate seed. Hence, a total of 16 entries including hybrids, their cultivated and wild parents were involved (Table 3.4). The design

was randomized complete block (RCBD) with 5 replications. The plot size was a single row of length 3m. The inter row spacing was 0.75m and the space between plants was 15cm. The space between blocks was 1.5m. Two seeds were dropped by hand per hill and latter the seedlings were thinned to a single plant per hill. The thinned seedlings were transplanted wherever there were some missing hills. In order to avoid transplanting shock, transplanting was done while it was raining. Because the wild parents and some of the hybrids tended to lodge, they were braced with stakes. All management practices were done as per the recommendation for sorghum in the lowlands where the wild sorghums are better adapted in Ethiopia. Accordingly, Diammonium phosphate (DAP) fertilizer was applied at the rate of 100kg/ha during planting. Moreover, urea was applied at the rate of 50kg/ha when the plants reached 20-30cm in height. To control sorghum shoot fly and stem borers, Karate 5% E.C. was appled at the arte of 320ml/ha. The field was kept free of weeds throughout the duration of the experiment.

3.4. Ecotypic variation for seed dormancy, longevity and germination requirements in wild/weedy sorghum

3.4.1 Effect of burial time and depth on seed longevity of different ecotypes of wild sorghum

This study was composed of two sets:

Set-1. Of the total 30 sites where seed samples were collected in 2008, 5 populations representing each region of collection were selected for this study based on the

availability of adequate seeds. These were Ghibe (G-1), Hararghe (H-2), Pawe (P-6), Tigray (T-8), and Wello (W-5) (refer to Table 3.1). Each population was represented by 5 individual plants. A total of 400 observations were originally buried in March 2009 at Melkassa Agricultural Research Center. It was a factorial experiment in randomized complete block design (RCBD) consisting of 5 individuals × 5 regions × 4 exhume dates × 4 replications.

Set-2: For the second set of this study, seed from 3 to 4 plants from each of the two wild sorghum populations collected from Pawe and Ghibe areas in November, 2009 was bulked. Moreover, one cultivated sorghum variety, WSV387 was included. It was arranged in a factorial experiment in RCBD with 3 replications. The 3 factors were 2 burial depths (10cm and 20cm), 3 Populations/genotypes (2 wild and 1 cultivar) and 3 exhume dates (6, 12, and 18 months).

For both sets of the experiment, double layered permeable nylon mesh bags of size 15cm × 35cm were prepared for seed burial to create the natural soil environment. The mesh bags were prepared in such a way that they had folds to hold Nylon strings for tying. Fine soil particles were observed to pass through the mesh bags. To avoid this, first, the dry soil was sieved with a 2mm sieve. The resulting soil was again sieved with a 1mm sieve and what was left on the sieve (with particle size $1<X<2$mm) was used for the experiment. All of the wild sorghum seeds had more than 2mm size. Therefore, during each excavation period, it was easier to recover seeds using a 2mm sieve, which would pass all of the soil in the bags but not the seeds. In many of earlier studies (e.g., Omami *et*

al., 1999; Narwal *et al*., 2008) a large number of seeds mixed with soils are placed in small bags, usually 10cm × 10cm. However, high seed density in the mesh bags may influence seed mortality because of elevation of pathogen levels as it is higher than the natural density in the soil bank (Van Mourik, *et al*., 2005). In this experiment a more ideal proportion of soil to seed mixture was used as suggested by Teo-Sherrell *et al*. (1996). Accordingly, 50 seeds from the intended samples were mixed with 500 gram of sieved soil in the mesh bag. The soil was collected at Melkassa where the species was not grown for more than two years. Paper labels sealed in polythene bags were included in each bag for identification of samples during exhume. In order to prevent loss of seeds from the bags and to avoid entrance of soils and other species weed seeds, the bags were tightly fastened and buried in horizontal position. The bags were buried on 14-15 March 2009 for set-1 in a 10cm depth with their colored strings extended to the outside of the soil for easy detection of the bags during exhumation (e.g., Van, 2005). The bags were buried 50cm apart from one another. There was no rainfall during the dry season except a few days about 3 months before the burial date of set-1. The land where the seeds were buried was previously used for doing other crop experiments and has been extensively disturbed. However, during the duration of the experiment it was kept undisturbed (no cultivation and weeding). Every six months (6, 12, 18 and 24), 20 bags (5 genotypes x 4 replications) from each region of collection (a total of 5 x 20 = 100) were exhumed for viability test for set-1. For set-2, 18 bags were exhumed every 6 months and the last samples were exhumed after 18 months. Exhumes were held on 16 September 2009, 26 February 2010, 5 October 2010, and 16 March 2011, for the first, second, third and fourth exhumes, in that order for set-1. For set-2, bags were buried on 26 April 2010 and the

first, second, and third exhumation dates were 25 September 2010, 16 March 2011, and 3 November 2011, respectively.

During each exhume, the seeds were sorted from the soil with 2mm sieve and graded based on their physical appearance and strength. During the first exhume in both sets, the seeds germinated in the field were found with their intact dead plumule and radicle tissues. Moreover, some of those seeds which were died in the field crushed easily when minimum force was applied between the thumb and index finger and the endosperm looks decomposed (had brown endosperm). However, after a year it was difficult to identify those seeds germinated in the field as the dried tissues decompose. The stiff seeds were washed and kept in germination boxes and soaked in 2% Clorox for 5 minutes in order to sterilize them and rinsed twice with distilled water. The seeds were then placed in new 11cm×11cm×3.5cm germination boxes on blotters wetted with distilled water. The germination boxes were placed in incubator adjusted at 30ºC temperature. The temperature was selected based on calibration. The germinated seeds were counted and removed starting from the third day for two weeks. Later, the non germinated seeds were evaluated for viability by dissecting and soaking them in 1% Tetrazolium chloride (2,3,5-triphenyltetrazolium chloride) solution for 6-8 hours at 35ºC in an incubator. Those seeds that fully or >2/3 stained (red) were considered viable. Therefore, the whole data in each set was classified in to three proportions: germinated in lab (germinable), dormant and dead. The proportion that was considered as dormant includes those seeds, which didn't germinate during lab germination test, but were found viable using TTC test. The dead group includes seeds those germinated and/or decayed in the field and those found intact

but confirmed dead after TTC test. All of the proportions of the viability measures were estimated based on the original 50 seeds. Data on mean monthly soil temperature at 10cm and 20cm depth and rainfall measurements were recorded at a meteorological weather station of the research center located about 200 meters from the experimental plot.

3.4.2. Response of wild sorghum ecotypes to differential moisture stress and temperature regimes

This experiment was carried out in September, 2011 at Melkassa Agricultural Research Center in the laboratory. Polyethylene glycol (PEG 6000) is the best chemical known for simulating drought stress in seed germination experiments. It was a 3 factor factorial experiment in a randomized block design (RCBD) with 4 replications. The factors were osmotic potential (OP) with 5 levels (0MPa (control), -0.3MPa, -0.6MPa, -0.9MPa, and -1.2MPa), 3 constant temperature levels (15°C, 23°C and 30°C) and 3 wild sorghum populations (H-1, P-5, and W-4). The above levels of OP were produced by mixing 0g (distilled water, control), 135.6g, 204.4g, 257.6g, and 302.5g at 15°C; similarly, 0g, 148.1g, 219.5g, 274.7g, 321.2g at 23°C; and 0g, 160.5g, 234.7g, 291.6g, and 339.7g of polyethylene glycol (PEG 6000) at 30°C in a liter of distilled water according to Michel and Kaufman (1973) using the formula:

$$\Psi s = (-1.18 \times 10^{-2})C - (1.18 \times 10^{-4})C + (2.67 \times 10^{-4})CT + (8.39 \times 10^{-7})C^2T$$

Where, Ψs =osmotic potential in Mega Pascals (MPa), C=PEG concentration, T=Temperature. The minus sign indicates that dissolved PEG reduces the water potential of a solution relative to the reference state of pure water (Tazi and Zeiger, 2003).

The wild sorghum genotypes used in the study were grown at Melkassa Agricultural Research Center during the 2010 main crop season under the same management and harvested in November. The seeds were kept in the laboratory at room temperature (25±5°C) until the beginning of the experiment. The genotypes were selected based on their representation to 3 different agroecologies and by their high rate of germination (≥88%). In order to avoid mechanical seed dormancy, the grains were deglumed before the experiment using fine sand paper. Twenty-five seeds from each of the 3 weedy sorghum populations were placed in a 115mm petridishes on top of Whatman® number one filter paper. Seeds were treated with ApronStar® to avoid decay of seeds by fungi. Five milliliter of each solution was applied to each petridish and the petridishes were sealed with a 2 inch wide Parafilm® to avoid evaporation. The petridishes were randomly placed in each of the two upper compartments of a WTC binder 78532 incubator (Tuttlingen/ Germany) in the dark. In order to avoid breakage of the Parafilm, one liter of distilled water was kept in a 5 liter beaker at the bottom compartment of the incubator to increase the humidity of the incubator.

3.5. Data recording and statistical analysis

3.5.1. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia

3.5.1.1. Estimation of phenotypic diversity

Readings of the coordinates and altitudes of the collection sites were recorded by a Global Positioning System (GPS), which was later overlaid on to the regional map of Ethiopia using ArcGIS version 9.3 (Fig. 1). *In situ* phenotypic measurements were made based on 20 plants from each of the 30 populations (Total 600). For analysis of qualitative traits, 516 samples with complete data were included. Measurements were made on 7 quantitative characters as per Appendix Table 1. In addition, hundred seed weight was measured after harvest. Because the seeds of the wild sorghum were fully covered by glumes, in order to record grain color and size, they were deglumed using fine sand paper. Seed size was estimated based on visual relative comparison among populations. For this, prior to measurement, 5 samples from each categorical class were prepared as standard with which all the samples were compared. Simple magnifying hand lens was used to observe the presence of hair on glumes. All of the other measurements were recorded by integrating the descriptors for sorghum (IPGRI/ICRISAT, 1993 and Bioversity International, 2010). GenStat 7th ed. (DE3) version 7.2 (VSN International Ltd., 2008) was used to calculate simple descriptive statistics. The resulting frequencies of qualitative phenotypic measurements were used to compute Shannon and Weaver Diversity Index (Hutcheson, 1970) as follows:

$$H_c = -\sum_{i=1}^{n} P_i \ln Pi$$

Where, for a given character *C*, H is the Shannon Weaver diversity index, *n* is the number of phenotypic classes in a given categorical character and *p* is the proportion of observation in the i^{th} class pi (1, 2,…n). In order to keep all values between 0 and 1, H was later standardized to H′ as the ratio of H and ln n, where n is the total phenotypic character classes.

Principal component analysis was performed using the correlation matrix and Pearson's coefficient of correlation between all pairs of quantitative characters was also performed and their significance tested using t test.

3.5.1.2 Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (l.) Moench] using SSR markers

3.5.1.1.1. SSR polymorphism and analysis of genetic diversity

Different measures of diversity were analyzed using two or more software programs in order to verify consistency of the results. For estimating observed heterozygosity, total expected heterozygosity, polymorphism information content (PIC) (Botstein *et al*., 1980), and allele frequency based genetic distance analysis, PowerMarker software V3.25 (Liu and Muse, 2005) was used and FSTAT software (Goudet, 2002). Nei's heterozygosity estimates (H_o, H_s, and H_t) were computed using FSTAT software (Goudet, 2002). The rarefaction method for allelic richness and private allelic richness trims unequal samples

to the same standardized sample size, a number less than or equal to the smallest sample size across populations (Szpiech *et al*., 2008). This method, implemented in HP-Rare 1.1 software (Kalinowski, 2005) was used to compute allelic richness (A_{taxon}^{S}) and private allelic richness ($\prod_{taxon}^{S}$) of a taxon. A taxon represents a single population, total populations of wild sorghum, cultivated sorghum or ICRISAT accessions. For the ease of presentation, the notations R_s and R_p are used instead of A_{taxon}^{S} and $\prod_{taxon}^{S}$ to represent allelic richness and private allelic richness, respectively. Allelic richness (R_s) was computed as:

$$R_s \triangleq \sum_{i=1}^{m} P_{i,taxon}^{S}$$

Where, $P_{i,taxon}^{S}$ is the probability that each allele (i=1, 2, 3, …m) is represented in a subsample of size S taken without replacement from the actual sample size N collected from the taxon. Private alleleic richness (Rp) was also computed using the following equation:

$$R_{p_{taxon}} \cong \sum_{i=1}^{m} U_{i,taxon}^{S}$$

Where,

$$U_{i,taxon}^{S} = P_{i,taxon}^{S} Q_{i,taxon'}^{S}$$

Where,

$$P_i^S = 1 - Q_{i,taxon}^{S}$$

Where,

$$Q^{S}_{i,taxon} = \frac{[N_{(taxon)S_0} - N_{i,taxon}]}{[N_{(taxon)_{S_0}}]}$$

$U^{S}_{i,taxon}$ is the probability that the ith allele will be found only in the subsample from the taxon under consideration and $Q^{S}_{i,taxon}$ is the probability that the ith allele is not contained in a sample of size S from a taxon. N is the number of genes sampled and N_i is the frequency that the ith allele was observed in the sample, and S_0 is the number of genes for which rarefaction will be computed. Significance of differences in gene diversity, allelic richness, and private allelic richness was compared among the three sorghum types (wild populations, cultivars and accessions) using Wilcoxon's signed rank test. Kruskall-Walis test was also used to compare differences in the diversity parameters (H_e, R_s and R_p) among populations in each of the three sorghum types.

Shared alleles distance matrix (Jin and Chakraborty, 1993) was used to construct UPGMA (unweighted pair-group method using arithmetic averages) clusters for the various populations using PowerMarker, and the resulting UPGMA trees were viewed using TreeView 1.6.6 (Page, 2001, available at, http://taxonomy.zoology.gla.ac.uk/rod/rod.html).

The shared allele distance (D_{SA}) of Jin and Chakraborty (1993) was computed using the following equation:

$$D_{SA} = \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{a_i} min\left(p_{ij}, q_{ij}\right)$$

Where p_{ij} and q_{ij} are the frequencies of the i^{th} allele at the jth locus; *m* is the number of loci examined; and a_i *is* the number of alleles at the j^{th} locus.

To study the distribution of the components of variance within and among the Ethiopian wild sorghum populations, analysis of molecular variance (AMOVA) was performed using Arlequin v 3.1 (Excoffier *et al*., 2005). Regions component was not included for AMOVA as some regions had more extended areas of collection than others.

3.5.1.1.2. Estimation of population structure and gene flow

F statistics (F_{ST}) and Slatkin's R_{ST}, were computed using FSTAT software v.2.9.3.2 (Goudet, 2002). Further, the pattern of population structure and detection of probable introgression was visualized using a Bayesian model based clustering method implemented in STRUCTURE software, Version 2.2 (Pritchard *et al*., 2000). For this, the admixture model with correlated allele frequencies was used. A burnin period of 500,000 was used followed by 1,000,000 Markov Chain Monte Carlo (MCMC) replications for data collection for K=1 to K=11. For each K value, five replicates were run. This procedure clusters individuals into populations and estimates the proportion of membership in each population for each individual (Falush *et al*., 2003). Moreover, the model does not assume a particular mutation process (Pritchard *et al*. 2000). The optimum number of clusters was predicted between K=1 and K=11 following the

simulation method of Evanno *et al.* (2005) using the web based software STRUCTURE HARVESTER v0.6.8. (Earl and von Holdt, 2011).

Two indirect estimates of gene flow based on the number of migrants per generation (N_m) were computed using Wright's (1951) fixation index (F_{ST}) and Slatkin's R_{ST} as:

$$N_m = \frac{(1 - F_{ST}\ or\ R_{ST})}{4F_{ST}\ or\ R_{ST}}$$

R_{ST} is similar to θ (standardized F_{ST}) defined by Weir and Cockerham (1984), but in the former case allele sizes are taken into account whereas in the latter analysis only identity or nonidentity of allelic states are entered (Slatkin, 1995). The geographic distance matrix of the sampling sites was calculated from geographical coordinates marked with the aid of GPS using the Geographic Distance Matrix Generator (GDMG) version 1.2.3 software of the American Museum of Natural History, Center for Biodiversity and Conservation (http://biodiversityinformatics.amnh.org/open_source/gdmg/index.php). The correlation between the resulting geographic distance matrix of the collection sites and Rousset's (1997) genetic distance matrix, which is most convenient for isolation-by-distance models (Templeton, 2006), was computed and the significance was tested by a Mantel test (Mantel 1967) using IBDWS v. 1.52 web-based program software (Bohonak, 2002) to examine the geographic pattern of gene flow. Rousset's (1997) genetic distance was computed as $\frac{F_{ST}}{1-F_{ST}}$. In order to estimate the level of divergence between wild and cultivated sorghum, F_{ST} was computed by grouping the whole data set into wild and

cultivated using PowerMarker (Liu and Muse, 2005). For this purpose, only those wild and cultivated sorghum populations that were collected in the same region (see Table 3.1), were included.

3.5.1.1.3 Detection of recent population bottlenecks and estimation of effective population size

This study aimed at investigating the occurrence of recent bottlenecks due to habitat change as it was assumed by Amsalu Ayana *et al*. (2001). For this study, only two populations, which were isolated from the crop and presumed to be more vulnerable to direct effect of environmental change, were included. These were Awash (AW) and Ghibe Bereha/Sateri (G-2) populations. Two methods of detecting population bottlenecks: heterozygosity excess and the mode-shift were applied. In this study, the program BOTTLENECK 1.2.02 (Cornuet and Luikart, 1996) was used to detect bottleneck in the wild sorghum populations. The program estimates the number of alleles per locus and from which generates the expected heterozygosity and the equilibrium expected heterozygosity (HeEq). The statistical tests that the program applies include the sign test, the standardized differences test, wilcoxon sign rank test and mode shift under the three mutation models (infinite allele model (IAM), two-phase model (TPM), and ste-wise mutation model (SMM)). However, for the standardized differences test to work, the number of markers should be at least 20. Because only 9 SSR markers were used, this test was not applied and the most appropriate tests were the Wilcoxon sign rank test and the mode shift test (Cornuet and Luikart, 1996). In the mode shift test, the allele frequency

distribution shifts from the normal L-shape in the presence of a bottleneck. Significance tests were based on 10000 iterations.

Effective population size (N_e) was estimated using two methods on the basis of the BOTTLENECK program output. The first one was the method of Hartl and Clark (1989), which is based on the expected heterozygosity (H_e) and mutation rate (μ) computed using the following equation:

$$N_e = \frac{H_e}{4\mu(1-H_e)}$$

Where, H_e is the expected heterozygosity and μ is the rate of microsatellite mutation. This equation assumes the IAM of microsatellite mutation. The second method was the method of Ohta and Kimura (1973) which assumes the SMM (Nei, 1987) as follows:

$$N_e = \frac{\left[\frac{1}{(1-He)}\right]^2 - 1}{8\mu}$$

Because the rate of mutation of microsatellites in sorghum is not available, that of maize was used, which belongs to the same tribe, andropogoneae (5.1×10^{-5}, Vigouroux *et al*. (2002)).

3.5.2. Estimation of mating system parameters in *in situ* wild/ weedy sorghum using SSR markers

Mating system parameters were estimated both at the population level and at the family level based on the mixed mating model of Brown and Allard (1970) using the most popular procedure of Ritland and Jain (1981) (revised in 1990 and 1996) multilocus mating system program (MLTR for windows revised version 3.3 (Ritland, 2008) accessible at "http://genetics.forestry.ubc.ca/ritland/programs.html".

3.5.2.1. Estimation of *in situ* outcrossing rate

For estimation of outcrossing rate at the population level and at the family level, the families within populations and the individuals within families respectively were used as the units of resampling. The numerical method used was the Newton-Raphson (NR) as recommended in the manual. Standard errors of the estimates were based on 500 bootstraps. The recommended initial values that allow iterations to start for each parameter were used as they were checked by default. Assumptions of the model are described in Ritland (2002). This model specifies that both selfing and outcrossing (mixed-mating) occur in a given population (Shaw and Allard, 1982; Ritland, 2002).

Mixed-mating model is based on three important assumptions (Clegg, 1980): 1) self-fertilization occurs at a rate of s and random mating (outcrossing) at a rate of t =1-s, 2) the gene frequencies among pollen are distributed identically over all maternal plants, and 3) maternal genotype doesn't determine the rate of outcrossing. The mating system parameters that were estimated from the progeny array data and the maternal genotype

data of each population and the combined analysis include a) the multilocus population outcrossing rate (t_m); (b) the single-locus population outcrossing rate (t_s); c) the single locus inbreeding coefficient of maternal parents (Fm). The significance of the multilocus outcrossing rates of the different populations was tested using Wilcoxon paired-match test using GenStat Software v.7.2.222. Pearson's correlation of the multilocus outcrossing rate of the progenies with that of plant height, population density, and gene diversity of the maternal plants was also calculated.

3.5.2.2. Estimation of inbreeding and biparental inbreeding

Three types of inbreeding coefficient: inbreeding coefficient of the maternal plants (Fm), inbreeding coefficient of the progenies (F_p), and equilibrium inbreeding coefficient expected from observed outcrossing (F_{eq}) were computed. The single locus inbreeding coefficient of maternal parents (F_m) was computed with MLTR program. Coefficient of inbreeding of the progenies (F_p) was computed using FSTAT (Goudet, 2002) by randomly taking one progeny individual from each family. The inbreeding coefficient expected at equilibrium (F_{eq}) was computed using the method of Fyfe and Bailey (1951) as $F_{eq} = (1-t_m)/(1+t_m)$, where t_m is the population level outcrossing rate.

The presence of biparental inbreeding, mating between related individuals was tested in several ways as follow: 1) the difference between t_m and t_s. The presence of biparental inbreeding is declared if this difference is positive because single-locus estimates include all apparent selfing due to biparental inbreeding, whereas multilocus estimates exclude much of the apparent selfing due to biparental inbreeding (Shaw and Allard, 1982), as an

observed outcross at any locus overrides the apparent selfing manifested at other loci (Ritland, 2002), 2) by comparing the coefficient of inbreeding of the progenies (F_p) and the inbreeding coefficient expected at equilibrium (F_{eq}) following Brown (1979). In the presence of biparental inbreeding F_{eq} will be less than F_p (Neel *et al*., 2001). 3) Evaluating the correlation of paternity (fraction of siblings that share the same father), 4) correlation of selfing among loci (r_s). With biparental inbreeding, this quantity is less than one (Ritland, 2002), 5) single vs. multilocus correlated paternity. If single is greater than multilocus, there is an effect of population substructure on the male similarity between outcrosses (Ritland, 2008). Methods 1, 3, 4, and 5 were performed by the MLTR program software.

3.5.3. Morphology and fitness of wild × crop F1 sorghum hybrids

Data were recorded at the three main life history stages following Burke *et al*. (1998). Germination percentage and days to emergence were considered to be juvenile survival traits. Days to flowering, plant height, number of productive tillers, flag leaf length and width, leaf number, head length and head width were regarded as adult (growth) characters. Likewise seed yield per plant (female fecundity) was regarded as fertility (reproductive) traits. Data on germination percentage was taken in the laboratory every day starting from the third day through day 10 and the germinated seeds were removed after count. Data on days to 50% seedling emergence and flowering were collected on plot basis. Plant height, flag leaf length and width, leaf number, head length, and head width were recorded at flowering on individual plant basis from six to ten plants those were randomly selected and permanently tagged in each plot throughout the field

experiment. Number of productive tillers was counted at maturity. Panicles were covered using water proof Lawson® paper bags when flowering was completed and started producing seeds. Number of seeds per plant (female fecundity), was estimated from grain yield and thousand grain weight, which was counted using an electronic grain counter (Wagtech International, Tachtam, Berkshire, U.K.). Grain yield per plant was determined in each plot on the tagged plants. For estimating grain yield per plant, all heads from the six-ten main plants and from all productive tillers were threshed together, cleaned carefully and weighed and the total weight was divided by the number of plants. Count data were log transformed and percentage data were arc sine transformed. To assess variation among the genotypes (parents and crosses) all measured variables were subjected to analysis of variance and mean separation by Duncan's Multiple Range Test (DMRT) using SPSS Release 17.0. Two methods were followed to estimate the level of fitness: mid-parent heterosis and direct comparison of relative fitness. Mid parent heterosis (MPH) of the hybrids was computed using the formula:

$$\text{MPH} = \frac{(F1 - MP)}{MP} * 100$$

where F1 is the mean value of wild × crop hybrid, MP= the mean values of the parents involved in the cross. Significance of mid-parent heterosis was tested using t test following the procedure of Wynne *et al*. (1970) as follows:

$$\text{t} = \frac{(F1ij - MPij)}{\text{S. E.}}$$

$$S.E. = \sqrt{\frac{3}{2r}MSE}$$

Where, $F1_{ij}$=the mean of the ijth cross, MP_{ij}= the average of the parents of the ijth cross, S.E. = standard error, MSE = Mean square of error, and r= number of replications. The computed t value was compared with the tabular t value at degree of freedom of error at 95% and 99% levels of probability.

Relative fitness of the hybrids and their parents was computed following the procedure of Song *et al*. (2004). Accordingly, the mean values of the three types for each trait were arranged in an ascending order with the highest mean value receiving relative fitness of 1.0. The remaining values were computed from that value as a starting point and the significance was tested using Duncan's Multiple Range Test.

3.5.4. Ecotypic variation in wild sorghum for dormancy, longevity and germination requirements

For all experiments, germination data were recorded every day starting from the second day until day 15. A seed was considered germinated if the radicle was at least 2mm long. All germinated seeds were removed after counting. The germination and TTC viability data were converted into percentages of the total seeds, 50 in each bag, for the seed burial experiment and 25 for the experiment involving temperature and osmotic potential. Total viability was computed as the percent of seeds germinated in the lab plus percent of

dormant seeds (TTC viable). General Linear Model (GLM) ANOVA with a Poisson distribution was computed on the arcsine transformed data using MINITAB software release 14. Replication was considered as random effect and genotype, region, depth, temperature, osmotic potential, and duration of burial were regarded as fixed effects. Analysis was done separately for each of the three data sets, set-1 and set-2 of the seed burial experiment, and for the experiment involving temperature and osmotic potential. For all sets, survival curves were plotted. Because most of the genotypes had zero viability at the 24th month, the data in set-1 recorded at the final exhumation date was excluded from ANOVA. Data were compared with means and standard deviations. Because the viability of buried seeds in soils follow exponential decline (e.g., Conn *et al*. 2006; Nielson *et al*., 2009), the rate of decline was computed from the exponential decay function as:

$$P_{(t)} = P_0 e^{-rt}$$

Where, $P_{(t)}$ = the number of seeds still viable after time t, P_0 = the initial number of seeds buried, r = the decay rate, and t = the time (duration) of burial (in years) in the soil. Moreover, the time taken for 50% of the seeds to lose their viability ($t_{0.5}$) was estimated from the above formula:

$$t_{0.5} = \frac{[\ln(0.5) - (\ln P_0)]}{-r}$$

4. RESULTS

4.1. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia

4.1.1. Geographical distribution and phenotypic diversity of wild/weedy sorghum

Unlike the cultivated sorghum, which grows in wide altitude ranges of as low as 400m and as high as 2500m, most of the wild sorghums were more abundant in lower altitudes. They were found in all the surveyed areas from hot and low rainfall lowlands of 600m at Humera to a relatively cooler and higher elevation (1925m) at Adenkel, both in Tigray. However, they were not found in Hararghe during the season except at two locations, Bedesa-Gelemso road side and at Bisidimo. They were more abundant in *Beneshangul-Gumuz* (*Metekel* zone), than any other region of the country where collection was made. Starting from a few kilometers from the first collection site P-1 (border of Amhara and Benishangul Gumuz regional states) they continuously colonized the farm and the uncultivated lands. No other species of wild sorghum with perennial nature was found in all of the surveyed areas. The only populations found in a wild habitat far from any crop were G-2, G-3, and AW. However, AW was not included in this study as the seeds were already shattered during arrival for collection.

In Ethiopia, wild sorghums exist as weeds in cultivated sorghum or other cereals (maize, tef and barley) or in naturally isolated and disturbed (e.g., near fences, irrigation ditches and abandoned farm lands) and relatively undisturbed (e.g., national parks and forests) habitats. They appear very vigorous wherever they occur as weeds in cultivated sorghum

and volunteers in maize fields, perhaps they are hybridized forms and those found in small cereals such as tef and barley appear very weak statured.

Quantitative traits

There was remarkable variation among the wild sorghum collections for quantitative morphological characters. Table 4.1 shows the mean, range (min. and max.) and the standard deviation of the eight quantitative characters measured *in situ*. The phenotypic characters of the wild sorghum populations extend from weak stemmed, grass like open panicle forms (typical wild) to those with vigorous stalks and panicles as compact as the cultivated race durra (stabilized hybrid forms). In general, the magnitude of the quantitative traits were in the range of 90cm (H-2) to 530cm (P-2) with an average of 277cm for plant height, 1.5cm to 10cm for leaf width, 18cm to 95cm for leaf length, 3 to 18 for number of leaves per plant, 0 to 24 for number of tillers per plant, 15 to 57cm for head length, 4cm in the compact weedy types to 43cm in those with very lax panicle typical wild types for head width. There was significant correlation ($p<0.001$) among all pairs of characters except for hundred seed weight and number of tillers per plant.

The first two principal components (PC1 and PC2) explained 73.89% of the total variation in the data set. Plant height had the largest contribution to the first principal component (PC1) (45%) followed by leaf number (44%) and leaf width (44%). The second component was mostly influenced by head width (65%) and hundred seed weight (51%). Based on this, in the principal component biplot (Fig. 4.1) the short stature, loose panicle and perhaps the 'typical wild' sorghum forms (T-1, T-2, T-3, H-2, W-7 and W-8)

took their position in one side of the PCA biplot. Five of the six Metekel zone populations were clustered together in the biplot. The putative hybrid populations of the weedy sorghum H-1, T-6, T-7, and W-6 also took one side of the biplot.

Table 4.1 Summary of the descriptive statistical values of the quantitative morphological characters measured on the 600 wild sorghum plants in the collection field

Character	Min.	Mean	Max.	SD
Head length (cm)	15	33.73	57	7.70
Head width (cm)	4	19.85	43	7.28
Leaf length (cm)	18	54.90	95	13.91
Leaf width (cm)	1.5	5.08	10	1.90
Leaf number	3	9.18	18	3.37
Plant height (cm)	90	277.10	530	93.40
Number of tillers	0	1.76	24	2.79
Hundred seed weight (gm)	0.4	1.41	3.8	0.63

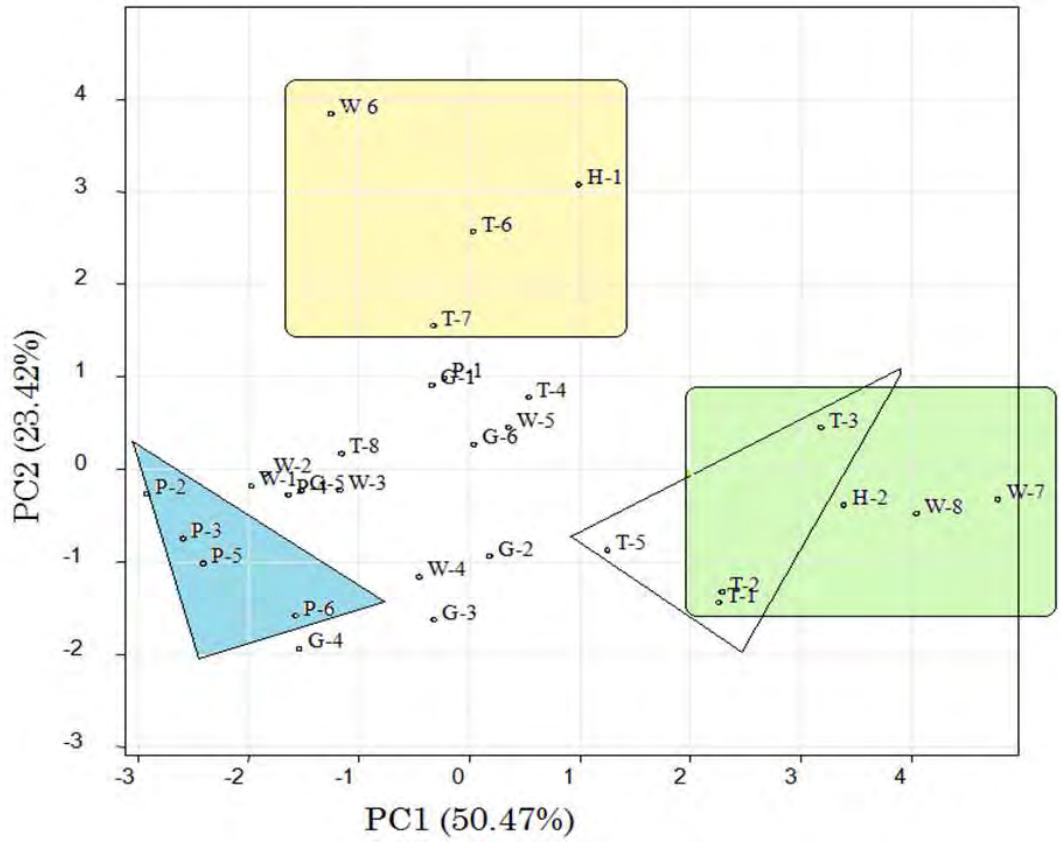


Fig. 4.1 Principal component biplot showing the population structure of wild sorghum for the quantitative phenotypic characteristics

Qualitative traits

Qualitative traits are known to be less affected by environmental variation and are more reliable phenotypic traits than quantitative traits. The frequency distribution of these traits in Ethiopian wild sorghum was such that about 90% had awns at maturity (Table 4.2). Moreover, 67.1% had high to very high shattering. Further, 52.7% had very loose to loose panicles with both erect and drooping primary branches. The weedy sorghum populations with compact panicles were found in Wello (Zobel) and in Tigray (Golel and

Hagereselam/Idris). The panicle architecture of some wild/weedy sorghum samples is presented in Fig. 4.2. For the grain associated traits, 65.7% had hairs on their glumes; 58.9% of the collections had grains fully covered; 41.1% had glumes longer than grain, and 75.9% had very small (as fine as finger millet seed) to medium grain size. The remaining proportion of them, mainly the stabilized hybrids, had large to very large seeds resembling the cultivated sorghum and was distributed in all of the 5 regions. As much as 81.4% of the populations had light to dark brown seed colors. The observed grain colors in the total collection were more than what was in the descriptors (Fig. 4.3), but they were rounded to the nearest color to fit the descriptors. White grain samples were observed only in Wello and Yellow grains were found in Tigray and Hararghe only. In Tigray all color grains except white were observed. In total, 74.6% of the collections had black (brown-black) glume color group.

The overall standardized Shannon-Weaver diversity index (H′) of the qualitative traits was in the range of 0.48 (for presence of awn at maturity) to 0.98 (for grain covering). Its pooled value over all characters was also in the range of 0.33 for Hararghe to 0.84 for Tigray with an average of 0.76 (Table 4.3). Moreover, the distribution of the qualitative traits in agro-climatic groups showed that warm semi-arid lowlands (SA2), which comprised of Tigray collections supported the highest Shannon-Weaver diversity index (H′) for 4 qualitative characters (Table 4.4).

Table 4.2 Frequency distribution of the various qualitative morphological characters by their region of origin and the total percentage contribution of each character class to the total variation (see appendix Table 2 for description of character codes)

		Awn		Shattering					Panicle compactness									Glume hair		Grain covering	
Region	Sample size	1	2	1	3	5	7	9	1	2	3	4	5	6	7	8	10	1	2	7	9
Ghibe	85	84	1	0	0	42	15	28	18	0	0	49	0	13	5	0	0	56	29	55	30
Hararghe	38	35	3	0	0	0	0	38	0	0	18	0	0	20	0	0	0	38	0	38	0
Pawe	110	109	1	0	19	22	60	9	0	0	0	31	16	40	23	0	0	66	44	77	33
Tigray	106	75	31	12	4	0	56	34	0	17	17	38	0	16	0	18	0	84	22	61	45
Wello	177	161	16	0	18	53	0	106	0	11	81	38	0	24	6	0	17	95	82	73	104
Total	516	464	52	12	41	117	131	215	18	28	116	156	16	113	34	18	17	339	177	304	212
Percent of the total		89.9	10.1	2.3	7.9	22.7	25.4	41.7	3.5	5.4	22.5	30.2	3.1	21.9	6.6	3.5	3.3	65.7	34.3	58.9	41.1

Table 4.2 continued

		Grain size					Grain color											Glume color		
Region	Sample size	1	2	3	4	5	1	2	4	5	6	7	8	9	10	11	13	2	4	6
Ghibe	85	23	12	36	13	1	0	0	0	3	8	0	10	59	5	0	0	0	28	57
Hararghe	38	18	0	1	6	13	0	0	0	4	1	12	1	20	0	0	0	1	1	36
Pawe	110	0	34	55	20	1	0	0	0	1	7	1	7	82	12	0	0	1	7	102
Tigray	106	2	43	20	19	22	0	2	2	3	1	23	25	46	2	1	1	18	30	58
Wello	177	61	46	40	22	8	2	0	0	3	17	0	18	113	20	2	2	2	43	132
Total	516	104	135	152	80	45	2	2	2	14	34	36	61	320	39	3	3	22	109	385
Percent of the total		20.2	26.2	29.5	15.5	8.7	0.4	0.4	0.4	2.7	6.6	7.0	11.8	62.0	7.6	0.6	0.6	4.3	21.1	74.6

Fig. 4.2 Diversity in panicle architecture in wild sorghum from a single population in Ghibe valley (upper) and Pawe/Mandura (lower)

Fig. 4.3 Diversity in Ethiopian wild sorghum seed colour and size. Color example: White (W4-2), Chalky white (W4-12), Yellow (T8-1), Grey (T6-13), Light red (T8-10), Red (W6-17), Light brown (G4-7), Dark brown (P2-9), Black (T2-3), Purple (T6-5). Size example: Very small (T2-3), Small (W4-12), Medium (G1-2), Large (T8-2), Very large (W6-17).

Table 4.3 Estimates of standardized Shannon-Weaver Diversity Index (H′) of wild sorghum collections for the qualitative morphological traits by their region of origin

Region*	N	GLC	GRCV	GLH	AW	GRC	GRS	PCS	SH	Mean
Ghibe (5)	85	0.91	0.94	0.92	0.09	0.63	0.83	0.79	0.93	0.76
Hararghe (2)	38	0.22	0.00	0.00	0.40	0.70	0.80	1.00	0.00	0.39
Metekel/Pawe (6)	110	0.26	0.88	0.97	0.07	0.50	0.76	0.96	0.84	0.66
Tigray (6)	106	0.90	0.98	0.74	0.87	0.65	0.86	0.96	0.77	0.84
Wello (10)	177	0.56	0.98	0.99	0.43	0.58	0.90	0.82	0.82	0.76
All	516	0.62	0.98	0.92	0.48	0.56	0.96	0.83	0.83	0.76

**Figures in parentheses show the number of populations represented in each region, each population consisted of 17-20 individual plants; N=Sample size, GLC=Glume color, GRCV=Grain covering, GLH=Presence of glume hair, AW=Presence of awn at maturity, GRC*=Grain color, GRS=Grain size, PCS=Panicle compactness and shape, SH=Shattering

Table 4.4 Estimates of standardized Shannon-Weaver Diversity Index (H′) of wild sorghum collections for the qualitative morphological traits by agroecologies

Agro-ecology*	N	GLC	GRCV	GLH	AW	GRC	GRS	PCS	SH
SA2 (4)	71	0.93	0.98	0.46	0.95	0.73	0.97	0.99	0.85
SM2 (5)	89	0.42	0.99	0.51	0.68	0.53	0.77	0.83	0.73
SM3 (4)	69	0.63	0.99	0.99	0.32	0.67	0.94	0.95	0.94
M2 (4)	74	0.13	0.72	0.92	0	0.58	0.81	0.84	0.95
M3 (4)	74	0.66	0.99	0.91	0.25	0.77	0.92	0.84	0.99
SH2 (6)	103	0.85	0.91	0.91	0.07	0.57	0.81	0.87	0.97
SH3 (2)	36	0.65	0.96	0.97	0.18	0.31	0.80	0.94	0.81

**Figures in parentheses show the number of populations represented in each climatic zone, each population consisted of 17-20 individual plants; N=Sample size, GLC=Glume color, GRCV=Grain covering, GLH=Presence of glume hair, AW=Presence of awn at maturity, GRC*=Grain color, GRS=Grain size, PCS=Panicle compactness and shape, SH=Shattering

4.1.2. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (l.) Moench] using SSR markers

4.1.2.1. SSR polymorphism

The observed microsatellite allele size ranges were more or less similar to that reported in earlier studies using the same loci (Table 4.5). The mean number of observed alleles and polymorphic information content (PIC) are presented in Appendix Table 3. The nine SSR loci produced a total of 185 alleles from 640 individual samples pooled from Ethiopian wild, cultivated, and ICRISAT accessions, of which 139 (75.13%) were rare (with frequency ≤0.05). The Ethiopian wild population had a total of 165 alleles, of which 117 (71%) were rare. The cultivated sorghum also had 97 alleles, of which 64 (66%) were rare. Moreover, the ICRISAT wild accessions contained 103 total alleles, of which 54 (52.4%) were rare. On average the number of alleles per the 9 SSR loci was 18.33, 10.78, and 12.67 for the wild populations, ICRISAT accessions, and cultivars, respectively. The polymorphism information content (PIC) in all populations ranged from 0.66 (SbKAFGK1) to 0.92 (Sb1-1) with a mean of 0.81. Moreover, availability of alleles in each locus (the proportion of loci without missing alleles) ranged from 0.96 (Sb6-57) to 1.0 (Sb5-206, Sb5-256, Sb4-121, and SBKAFGK1).

Table 4.5 Characteristics of the microsatellite loci used in the study.

Multiplex Set	Marker	Flanking sequences (5'-3') [b]	Repeat motif	LG[cd]	Size range (bp) Earlier studies[d]	This study
1	Sb5-206	F: HEX-8ATTCATCATCCTCATCCTCGTAGAA R: AAAAACCAACCCGACCCACTC	$(AC)_{13}/(AG)_{20}$	E	92-156	86-150
	Sb1-1	F: FAM-6TCCTGTTTGACAAGCGCTTATA R: AAACATCATACGAGCTCATCAATG	$(AG)_{16}$	H	241-300	236-308
	Sb6-34	F: HEX-8AACAGCAGTAATGCCACAC R: TGACTTGGTAGAGAACTTGTCTTC	$[(AC)/(CG)]_{15}$	I	168-208	181-213
	Sb5-256	F: FAM-6AATTTGCTTTTTGGTCCGTTT R: TAGGAAAGACAGTACTAGAGGTCA	$(AG)_{8}$	C	162-214	157-187
	Sb4-72	F: NED-TGCCACCACTCTGGAAAAGGCTA R: CTGAGGACTGCCCCAAATGTAGG	$(AG)_{16}$	B	182-350	172-238
2	Sb6-84	F: FAM-TAACGGACCACTAACAAATGATT R: TAACGGACCACTAACAAATGATT	$(AG)_{14}$	F	170-222	174-218
	Sb4-121	F: NED-FAM-GAAAAATCTCCGTCAATCCCAAAAATAA R: CGCTGAACAACGAAAGGAATAAGTG	$(AC)_{14}$	D	200-229	181-228
	Sb6-342[a]	F: HEX-6TGCTTGTGAGAGTGCCTCCCT R: GTGAACCTGCTGCTTTAGTCGATG	$(AC)_{25}$	A	250-320	252-350
3	Sb4-15[a]	F: HEX-GCTGCTAAGCCGTGCTGA R: TTATTTGGGTGAAGTAGAGGTGAACA	$(AG)_{16}$	E	119-135	106-142
	Sb5-236[a]	F: NED-6GCCAAGAGAAACACAAACAA R: AGCAATGTATTTAGGCAACACA	$(AG)_{20}$	G	158-222	154-218
	Sb6-57	F: HEX-FAM-ACAGGGCTTTAGGGAAATCG R: CCATCACCGTCGGCATCT	$(AG)_{18}$	C	283-320	284-314
4	SBKAFGK1	F: FAM-6GCTTTCGGCGAGCATCTTACAA R: GCGGTTGGATTCGCCATG	$(AAC)_{9}$	J	140-320	227-284

[a] markers excluded from final analysis; [b] (F= Forward primer, R= Reverse primer); [c]LG = Linkage Group

[d] Earlier results were compiled from Brown *et al*. (1996), Dean *et al*. (1999), Bisrat Ghebru *et al*. (2002), Agrama and Tuinstra (2003), Abu Assar *et al*. (2005)

4.1.2.2. Genetic diversity of wild populations, accessions, and cultivars

Measures of genetic diversity for wild sorghum populations, cultivars and ICRISAT accessions are presented in Table 4.6. The overall allelic richness over the nine SSR loci was 14.93 in the wild sorghum pool, which was significantly higher ($P \leq 0.001$) compared both to cultivars (Rs = 10.0) and accessions (Rs = 11.41). The wild sorghum populations were also characterized by significantly higher levels ($P \leq 0.05$) of private allelic richness (Rp = 3.79) than both the cultivars (Rp = 1.35) and the accessions (Rp = 1.59). Gene diversity was also relatively higher in the wild pool (He = 0.89) compared both to the cultivars (He = 0.68) and the accessions (0.79), even though the difference was only significant ($P \leq 0.01$) for the wild *vs* cultivar comparison. The level of diversity varied significantly among the wild populations (Table 4.6) in allelic richness (range: 1.87 in T-1 to 4.98 in G-2; $P < 0.001$) and in gene diversity (range: 0.17 in T-1 to 0.75 in G-1; $P < 0.001$) but not in private allelic richness (range: 0.02 in W-4 to 0.50 in P-2; $P > 0.05$). Observed heterozygosity (Ho) was >0.50 in population T-4, while populations H-2 and T-1 had the lowest observed heterozygosity (Ho = 0.01 and 0.04, respectively). The latter two populations, along with AW (Ho=0.11) and W-4 (Ho = 0.14), formed a separate cluster that was distinct from the cultivars in the UPGMA analysis (Fig. 4.4).

Table 4.6 Genetic diversity indices of the Ethiopian wild and cultivated sorghum populations, and ICRISAT wild sorghum accessions over all 9 loci (Rs= allelic richness, Rp= private allelic richness, Ho= average observed heterozygosity, He= gene diversity, F_{IS}= inbreeding coefficient).

	Population	Rs[a]	Rp[a]	He	Ho	Fis
Ethiopian wild (N=380)	AW	2.95	0.12	0.43	0.11	0.80
	G-2	4.87	0.32	0.75	0.25	0.71
	G-4	4.98	0.39	0.77	0.31	0.59
	G-6	4.01	0.26	0.56	0.17	0.68
	H-1	4.29	0.28	0.60	0.43	0.27
	H-2	2.63	0.17	0.29	0.04	0.89
	P-1	3.40	0.39	0.54	0.10	0.77
	P-4	4.12	0.50	0.69	0.13	0.77
	P-5	3.20	0.29	0.59	0.23	0.69
	P-6	3.19	0.28	0.57	0.16	0.78
	T-1	1.87	0.09	0.17	0.01	0.93
	T-4	2.94	0.18	0.47	0.20	0.64
	T-5	3.85	0.22	0.64	0.18	0.73
	T-6	4.81	0.34	0.65	0.54	0.13
	T-8	4.28	0.22	0.69	0.17	0.80
	W-2	2.89	0.05	0.39	0.19	0.51
	W-4	3.57	0.04	0.67	0.14	0.78
	W-6	1.99	0.14	0.20	0.10	0.53
	W-7	2.35	0.02	0.34	0.14	0.58
	Overall	14.93	3.76	0.83	0.19	0.66
Ethiopian Cultivated (N=160)	GC1	1.93	0.11	0.15	0.12	0.25
	GC2	3.24	0.17	0.50	0.25	0.55
	PC1	4.14	0.80	0.41	0.10	0.76
	PC2	4.43	0.77	0.52	0.24	0.55
	TC	2.82	0.56	0.31	0.01	0.95
	WC1	5.57	1.05	0.74	0.18	0.71
	WC2	2.88	0.53	0.33	0.05	0.80
	WC3	2.87	0.11	0.34	0.10	0.68
	Overall	9.95	1.35	0.68	0.13	0.65

Table 4.6 continued

	Population	Rs[a]	Rp[a]	He	Ho	Fis
ICRISAT (N=100)	South Africa-arundinaceum	2.01	0.00	0.22	0.02	0.90
	Swaziland-arundinaceum	3.08	0.44	0.57	0.08	0.84
	Angola-arundinaceum	3.36	0.76	0.55	0.18	0.70
	USA-virgatum	2.92	0.55	0.55	0.32	0.46
	Egypt-virgatum	1.76	0.32	0.27	0.07	0.79
	Sudan-aethiopicuum-1	3.62	1.15	0.63	0.35	0.45
	Sudan-aethiopicuum-2	2.29	0.96	0.27	0.09	0.61
	Sudan-verticilliflorum	2.71	0.36	0.44	0.06	0.72
	Sudan-drummondii	1.64	0.44	0.18	0.13	0.46
	Ethiopia-drummondii	2.02	0.38	0.32	0.14	0.62
	Overall	11.41	1.59	0.79	0.14	0.65

[a]To facilitate comparisons, allelic richness and private allelic richness were computed with corrected sample sizes among the three sorghum groups (wild, cultivar and accessions; N=92 for each) and among populations (N = 7 in wild, N = 15 in cultivar and N = 7 in accessions) in each of the groups.

4.1.2.3. Geographical and agroecological pattern of microsatellite diversity in wild sorghum populations

Allelic richness (R_s) and private allelic richness (Rp) were computed by grouping populations according to their region of collection and agroecologies. Allelic richness of the regions of origin ranged from 6.42 (Wello) to 9.94 (Tigray), but the difference was significant between Ghibe and Hararghe (p=0.008), Ghibe and Pawe (p=0.05) and Tigray and Wello (p=0.008). Similarly, private allelic richness (R_p) ranged from 0.3 (Wello) to 1.68 (Pawe). The variation in private allelic richness was significant only between Ghibe and Wello (p=0.02) and between Tigray and Wello (p=0.027). Overall, Tigray had the highest Rs whereas Pawe had the highest Rp. On the other hand, Wello had the lowest wild/ weedy sorghum diversity based on these two indices.

Variations in R_s and R_p were significant among the agroecologies. The range of allelic richness of the agroeologies was from 2.95 (A2) to 5.99 (SH2). Private alleleic richness was also ranged from 0.27 (A2) to 1.26 (SH2). In general, warm sub-humid lowlands (SH2), which comprised of Ghibe and Mandura (Metekel) populations supported the highest Rs and Rp, whereas Warm arid lowland plains (A2), which contains only Awash populations in the present study, supported the lowest diversity based on these two indices. Altitude had significant negative correlation with gene diversity (r=-0.529, p=0.02), average observed heterozygosity (r=-0.448, p=0.054) and with allelic richness (r=-0.526, p=0.021) of wild sorghum populations.

4.1.2.4. Relationship among wild populations

The 19 wild sorghum populations were distributed into 5 clusters (Fig. 4.4). The population that was collected from Zobel (W-6), which was morphologically closer to the cultivated sorghum is placed as single (I). The second clauster (II) consisted of the typically wild populations (AW, H-2, T-1 and W-7). Moreover, all of the four Metekel zone populations (P-1, P-4, P-5, and P-6) clustered together in cluster III. The fourth cluster (IV) consisted of populations from South Wello (W-2), Ghibe (G-6) and from West Tigray (T-4). The fifth cluster (V) is the biggest cluster consisted of three sub-clusters. In the first sub-cluster, the population from Abuare (W-4) was placed. The second sub-cluster consisted of the two Ghibe populations (G-2 and G-4) and the third one consisted of the three Tigray populations (T-5, T-6, and T-8) and the population from Hararghe (H-1).

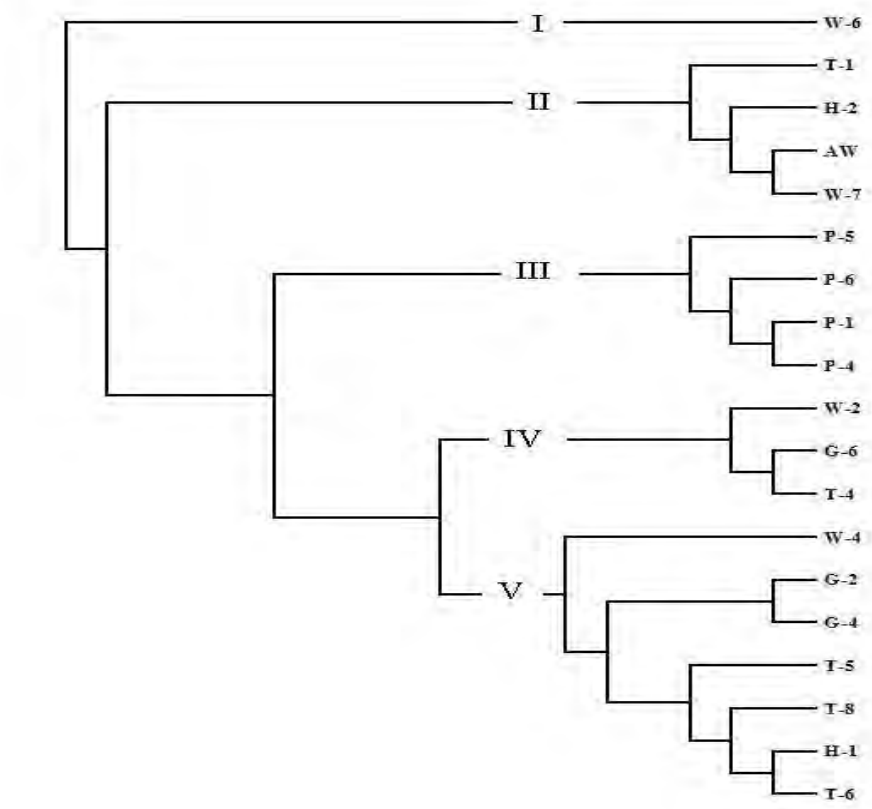


Fig. 4.4 Cladogram showing relationships among the wild sorghum populations collected from different geographical regions based on microsatellite data. See Table 3.1 and Figure 3.1 for collection site information.

4.1.2.5. Relationship of wild populations to ICRISAT accessions

The cladogram that shows the relationship among the Ethiopian *in situ* wild sorghum collections and the ICRISAT accessions (Fig. 4.5) is similar to the Fig. 4.4, but three of the 6 clusters joined ICRISAT accessions with the Ethiopian wild sorghum populations. Cluster I joined the Egypt and USA accessions belonging to race virgatum, which may suggest that Egypt could be the country of origin for the USA material. Cluster II joined the 4 Ethiopian wild populations (AW, H-2, T-1, and W-7) with *S. bicolor* ssp.

drummondii in the ICRISAT accession from Ethiopia. Below, additional evidence will be presented showing that these 4 wild populations are genetically distinct from the others. Cluster V joined the 2 Ghibe (G-1 and G-2), the 3 Tigray (T-5, T-6, T-8) and the Hararghe (H-1) populations with Sudan *aethiopicuum*-1. Cluster VI also joined W-4 with the rest of the ICRISAT accessions.

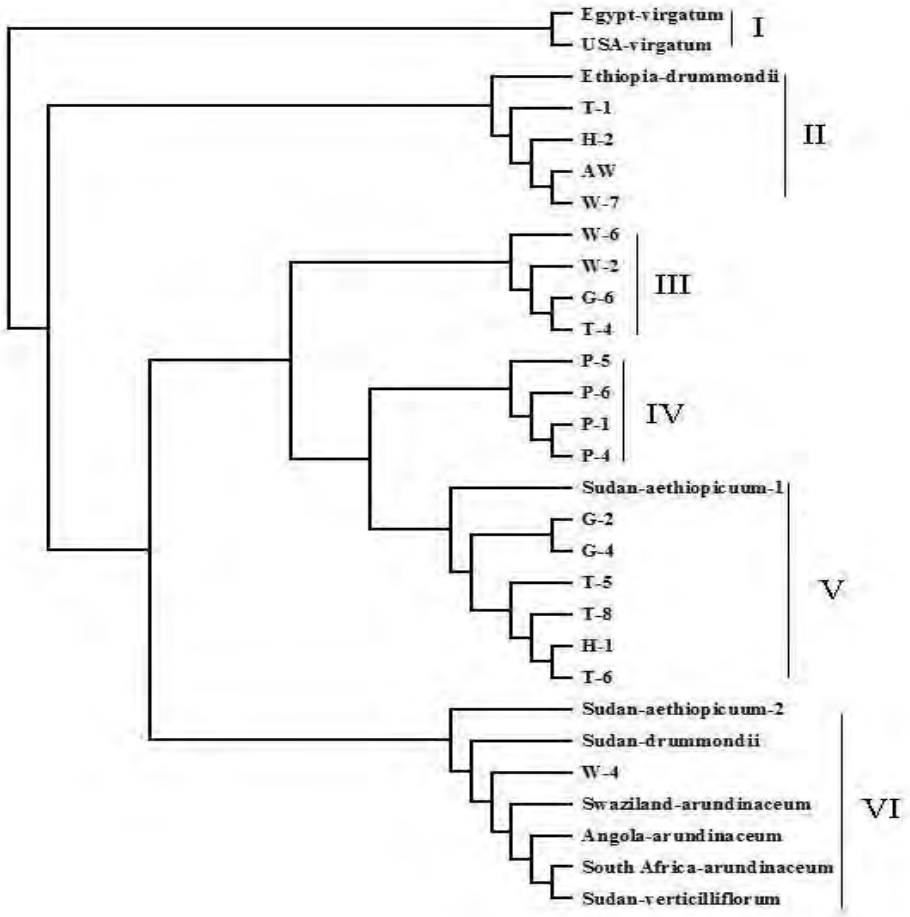


Fig. 4.5 Cladogram showing relationship of Ethiopian wild sorghum populations and ICRISAT accessions. See Table 3.1 and Figure 3.1 for collection site information.

4.1.2.6. Relationship of wild populations to cultivars

All of the cultivars had compact and semi-compact panicles typical of the durra and caudatum landraces, which are common throughout Ethiopia, and one, WC2, was an improved caudatum variety, $76T_1$#23. Two durra populations, WC3 and TC1, were collected in the same Kobo-Alamata plain located in adjacent districts (~35 km apart) and had the same local name, "Degalit". UPGMA relationships among the 8 cultivars (Fig. 4.6) showed four clusters, with the two "Degalit" landrace cultivars grouped together (cluster II). These two landrace cultivars together with an adjacent wild sorghum population collected from Zobel, W-6 also clustered together in a cladogram (Fig. 4.7). Evanno *et al*. (2005) simulation model predicted K=4 to be the optimum cluster (Fig. 4.8). Similar to Fig. 7, the two "Degalit" landrace cultivars and the adjacent wild sorghum population collected from Zobel clustered together at K=2 through K=4 in the STRUCTURE bar graphs (Fig. 4.9).

The cultivated sorghum landraces from Pawe and Ghibe clustered together (cluster-III, Fig. 4.6), as did the durra landraces from Wello and neighboring TC1 (Cluster-II). In Fig. 4.7, an improved variety, 76T1#23 (WC2) and intermixed weed (W-4) were in one cluster (Cluster-III). The landraces from Pawe and Ghibe also clustered together, as did the landraces from Wello (WC3) and neighboring TC1 (Fig. 4.6). In Fig.4.7, four of the 7 clusters joined wild and cultivated sorghum together. Very interesting is that an exotic, improved variety, 76T1#23 (WC2) remained single in Fig. 4.6. In Fig. 4.7, 76T1#23 and intermixed weed (W-2) with it were in one cluster. Clustering of Ghibe cultivar (GC2) with South Wello cultivar (WC1) could be due to long distance movement of seed.

During the interview with farmers, they responded that they came from South Wello and North Shewa areas as settlers.Overall, the results suggest that wild populations are closely related to the cultivars (Figs. 4.7 and 4.9), as did a Principal Component Analysis (Appendix Fig. 1). However, in spite of the fact that many wild populations grew in close proximity to the crop (Table 1; Tesso *et al*. 2008), evidence for gene flow between the cultivars and wild populations that were sampled was only moderate.

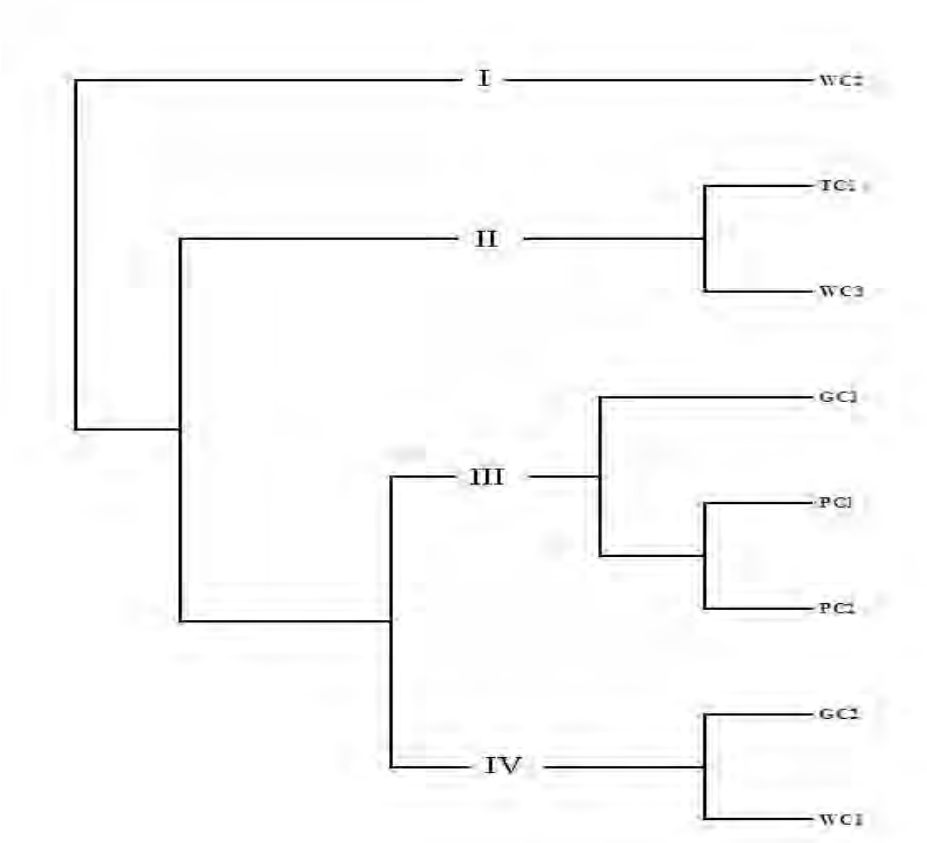


Fig. 4.6 Cladogram showing relationships among the 8 cultivated sorghum populations. See Table 3.1 and Figure 3.1 for collection site information.

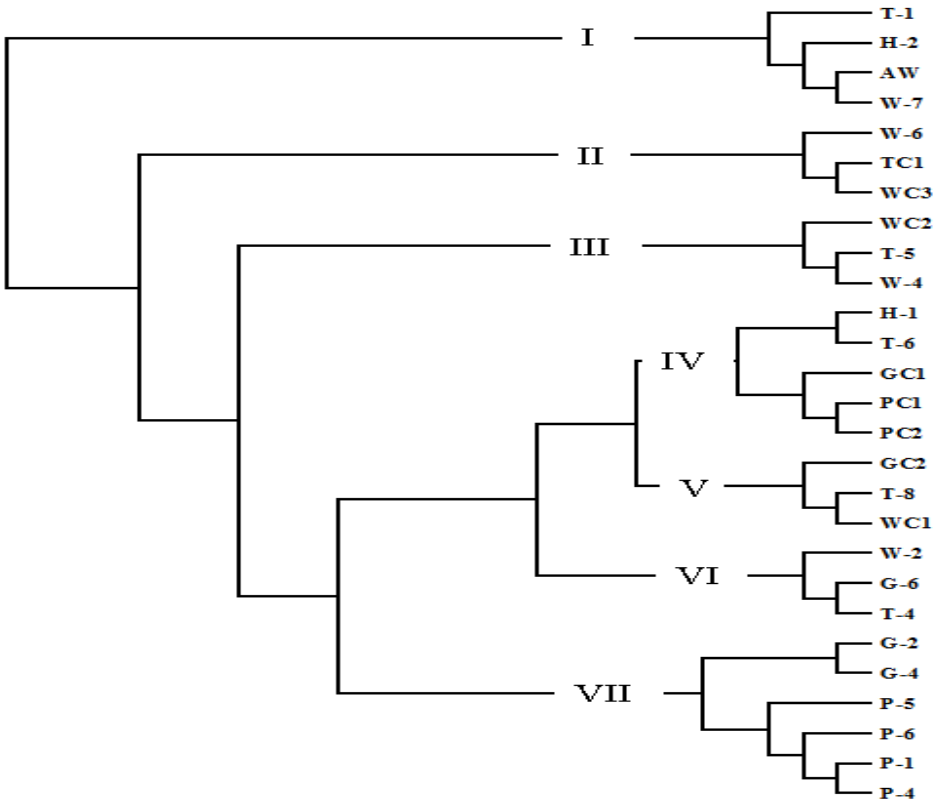


Fig 4.7 Cladogram showing relationships among the 19 wild and 8 cultivated sorghum collections. See Table 3.1 and Figure 3.1 for collection site information.

4.1.2.7. Genetic structure and gene flow in wild and cultivated sorghum populations

Wild populations displayed considerable genetic differentiation, some of which were associated with geographic proximity. For example, overall values of F_{ST} and R_{ST} for the 19 wild populations were 0.374 and 0.376, respectively (Appendix Table 3), which indicates very great differentiation. Likewise, analysis of molecular variance (AMOVA) showed that 41.2% of the variation occurred among wild populations and 58.8% found within populations (Table 4.7). Indirect estimates of gene flow using Wright's F_{ST} and Slatkin's R_{ST} were similar and showed that the average number of migrants per generation across wild populations was N_m= 0.43, with higher gene flow in some

adjacent wild populations. For instance, the largest number of migrants per generation was between G-2 and G-4 (N_m=3). Similarly, N_m= 2 between P-1 and P-4, and N_m= 1 between P-4 and P-6. In general, gene flow was evident in 18 population pairs (~10.53% of the total population combinations) (with N_m=1 to N_m=3) computed from pair wise F_{ST} (Appendix Table 5).

Table 4.7 Analysis of molecular variance (AMOVA) of the 19 wild sorghum populations as revealed by SSR markers

Source of variation	d.f.	Sum of squares	Variance components	Variation (%)	P value
Among populations	18	9099.1	1.20	41.2	<0.001
Within populations	7961	13629.6	1.71	58.8	<0.001
Total	7979	22728.7	2.91		

Four wild populations consistently clustered together in the UPGMA analysis (Figs. 4.4, 4.5 and 4.7) and in each level of K in the STRUCTURE analysis (Fig. 4.9). These were W-7 and T-1 in the North Wello-South Tigray region and AW and H-2 in the Awash-Haraghe region. Interestingly, these populations were grouped with the Ethiopian *drummondii* accession acquired from ICRISAT in the cladogram, and did not show affinity with the cultivars in either type of cluster analysis (Figs. 4.7 and 4.9). Thus, they can be considered to represent a distinct sub-group.

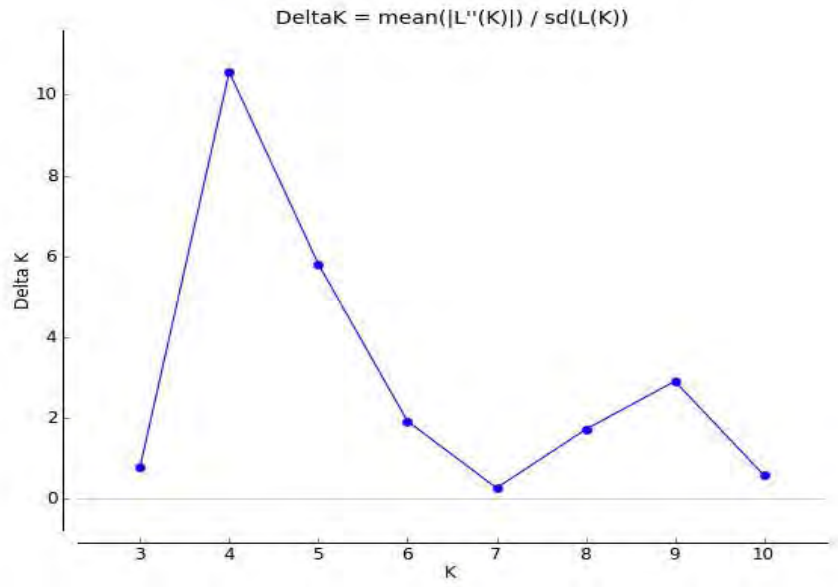


Fig. 4.8 Evanno *et al*. (2005) method predicted that the modal value (Delta K) at k=4 (the highest peak) to be the optimum k value

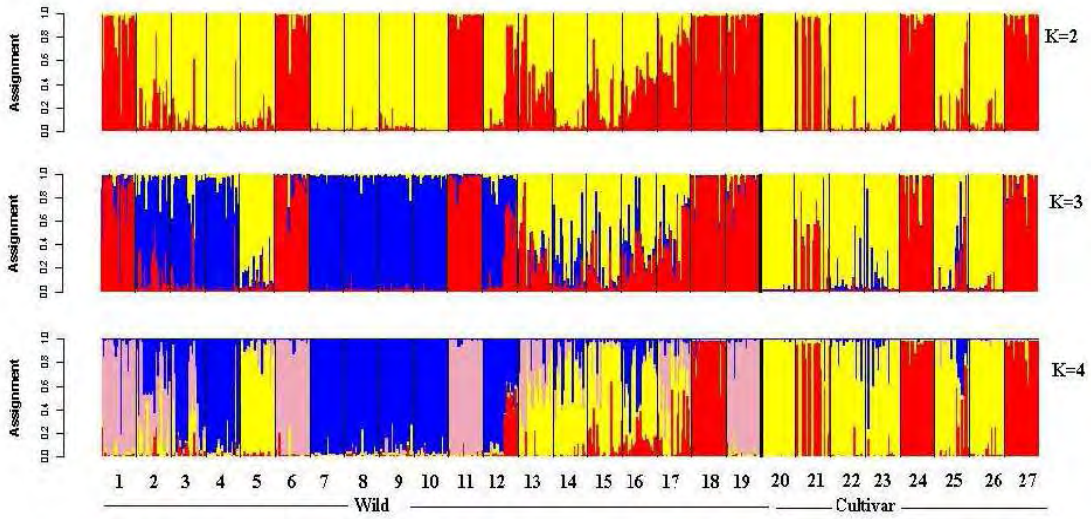


Fig. 4.9. STRUCTURE bar graph of 19 wild and 8 cultivated sorghum predefined populations (x-axis) with k=2 to k=4. Figures in the y-axis show coefficient of membership/ assignment. Each of the 540 plants is represented by a single vertical line of a color according to the estimated membership fraction averaged over all the 9 loci and each color group represents a cluster (K). Wild: 1=AW, 2=G-2, 3=G-4, 4=G-6, 5=H-1, 6=H-2, 7=P-1, 8=P-4, 9=P-5, 10=P-6, 11=T-1, 12=T-4, 13=T-5, 14=T-6, 15=T-8, 16=W-2, 17=W-4, 18=W-6, 19=W-7; Cultivated: 20=GC1, 21=GC2, 22=PC1, 23=PC2, 24= TC1 [Degalit cultivar], 25= WC1, 26= WC2 [improved cultivar 76T1#23], 27=WC3 [Degalit cultivar].

There was weak geographic structuring in the wild populations. Only some of the wild populations formed separate clusters according to their geographical regions (Fig. 4.4). For example, all of the four Pawe populations and two of the three Ghibe populations (G-2 and G-4) clustered according to their geographical origins. The three northwest Tigray populations (T-5, T-6 and T-8) clustered with each other but this group also included population H-1 (Hararghe) in the east. Similar patterns were observed in the STRUCTURE analysis in that the four Pawe populations clustered together at all levels of K as did G-2 and G-4. The remaining populations were clustered with cultivars in the UPGMA (Fig. 4.7) and Structure (Fig. 4.9) analyses.

The divergence among all pairs of geographical regions was significant ($P<0.05$), but it was relatively lower between Ghibe and Tigray and higher between Pawe (Metekel) and Hararghe (Table 4.8). Similarly, population differentiation among all pairs of the climatic zones (agroecologies) was significant ($p<0.05$). SH2 had the smallest divergence with M2 and SA2 (Table 4.9). In general, the wild sorghum growing agroecologies were more differentiated ($F_{ST}=0.179$) than the regions of collection ($F_{ST}=0.134$).

Mantel test of the correlation between Rousset's genetic distance and geographic distance matrices was not significant ($r = -0.1520$, $p=0.0682$). Even after transformation, the correlation between log (genetic distance) and log (geographic distance) matrices was not significant ($r = 0.1012$, $p=0.9106$).

Table 4.8 Pair wise F_{ST} showing wild sorghum population differentiation across the regions of collection.

	Hararghe	Ghibe	Pawe	Tigray
Ghibe	0.136*			
Pawe	0.196*	0.111*		
Tigray	0.109*	0.083*	0.120*	
Wello	0.126*	0.157*	0.179*	0.109*

*significant (P<0.05)

Table 4.9 Pair wise F_{ST} showing wild sorghum population differentiation across the climatic zones (agro-ecologies).

	SH2	M3	SH3	M2	SM2	SM3
M3	0.132*					
SH3	0.136*	0.252*				
M2	0.085*	0.213*	0.164*			
SM2	0.181*	0.250*	0.173*	0.211*		
SM3	0.222*	0.244*	0.272*	0.291*	0.290*	
SA2	0.084*	0.109*	0.158*	0.139*	0.163*	0.234*

*significant (P<0.05)

4.1.2.8. Detection of recent bottlenecks and estimation of effective population size

Using the heterozygosity excess and mode shift methods, there were no bottlenecks observed in the studied populations (Table 4.10). The effective population size of Ghibe Bereha (G-2) population was more in both the IAM (Ne=9951) and the SMM (Ne=37201) than a population from Awash National Park, which had IAM (Ne=5631) and SMM (Ne=16015) (Table 4.11).

Table 4.10 Sign, wilcoxon and mode shift tests for existence of bottleneck in wild sorghum populations from Awash National Park and Ghibe river valley (Ratios represent the number of microsatellite loci exhibiting heterozygosity deficiency vs. excess).

Population	Test	IAM		TPM		SMM		Mode shift
Awash park	S	0.405	5:4	0.342	5:4	0.036	7:2	None
	W	0.820		0.359		0.019		
Ghibe valley	S	0.211	2:7	0.538	4:5	0.109	6:3	None
	W	0.019		0.496		0.250		

S=Sign test; W= Wilcoxon test (two tail for H excess and deficiency)

Table 4.11 Number of samples (n), observed SSR alleles (ko), expected heterozygosity (He), average expected heterozygosity (Heq) computed from He, and effective population size (Ne) (based on IAM and SMM mutation models) of wild sorghum populations from Awash national park and Ghibe valley

	Awash							Ghibe valley						
Locus	n	ko	He	Heq	Ne (IAM)	Heq	Ne (SMM)	n	ko	He	Heq	Ne (IAM)	Heq	Ne (SMM)
Sb5-206	40	2	0.19	0.23	1440	0.26	4428	38	10	0.89	0.82	21739	0.85	106085
Sb1-1	40	5	0.66	0.59	6996	0.65	19894	38	7	0.73	0.71	12001	0.76	42199
Sb6-34	40	5	0.72	0.59	7025	0.65	19669	40	9	0.83	0.78	17481	0.82	77357
Sb5-256	40	4	0.35	0.5	4961	0.56	12776	40	4	0.7	0.5	4922	0.56	12718
Sb4-72	18	5	0.75	0.69	10809	0.72	31714	40	6	0.75	0.66	9307	0.71	29144
Sb6-84	40	3	0.38	0.39	3134	0.44	7760	40	6	0.65	0.66	9307	0.71	28746
Sb4-121	30	4	0.63	0.53	5572	0.59	14369	40	4	0.7	0.5	4961	0.56	12893
Sb6-342	40	6	0.53	0.66	9307	0.71	29144	40	4	0.47	0.5	4941	0.56	12834
SbKAFGK1	40	2	0.1	0.23	1431	0.25	4381	40	4	0.51	0.5	4902	0.56	12834
Mean	36	4	0.48	0.49	5630	0.54	16015	40	6	0.69	0.63	9951	0.68	37201

4.2. Estimation of mating system parameters in *in situ* wild/weedy sorghum using SSR markers

4.2.1. *In situ* outcrossing rate of wild sorghum in Ethiopia

The SSR primer pair, sb4-72 failed to amplify the DNA of most of the families in *Kobo* population. Hence, this population, amplified with the remaining four pair of primers was analyzed separately and excluded from the combined analysis. The number of observed alleles in the total families ranged from 16 (sb5-256) to 56 (sb1-1) showing that the SSR markers used were highly polymorphic. There was significant variation (p=0.016) in outcrossing rate among wild sorghum populations collected from different geographical regions. Multilocus outcrossing rate (t_m) ranged from 0.33 (*Humera/Donkey*) to 0.65 (*Abuare*). The average multilocus outcrossing rate (tm) of all populations excluding Kobo (which had 0.19) was 0.51. However, the value was in the extreme range of 0 to 100% in some families.

Multilocus family outcrossing rate in the progenies was found to have no correlation with plant height (r=-0.046, p=0.61), head length (r=-0.068, P=0.0.45) and width (r=0.026, p=0.775). Moreover, the correlation between population outcrossing rate and panicle compactness was negative, but not significant (r=-0.374, p=0.465). The population outcrossing rate had highly significant correlation with gene diversity (expected heterozygosity) of the maternal plants (r=0.714, p=0.01), but no correlation with altitude (r=0.013, p=0.978). The inbreeding coefficient of the maternal plants (F_M) had no correlation with plant density (r=0.52, p=0.291), and altitude (r=0.59, p=0.217).

Correlation of outcrossed paternity (r_P) had negative, but not significant correlation with plant density (r=-0.358, p=0.48).

4.2.2. Inbreeding and biparental inbreeding

Fig. 4.10 shows the three coefficients of inbreeding estimated from the population and family data. In the figure, inbreeding coefficient of the progenies (F_p) was greater than both inbreeding coefficient of the maternal plants (F_m) and equilibrium inbreeding coefficient expected from the observed outcrossing (F_{eq}) in all populations except in Mandura where it was less than F_m. Inbreeding coefficient of the maternal plants (Fm) was in the range 0.016±SE0.091 in Humera/Donkey to 0.662±SE0.063 in Mandura/Pawe with an average of 0.496±SE0.029. Similarly, F_p ranged from 0.436±SE0.036 in Humera/Donkey to 0.710±SE0.022 in Abuare with an average of 0.691±0.043.

Table 4.12 Estimates of the mating system parameters of the wild and weedy sorghum, figures in parentheses are standard errors

Estimate*	Gibe	Hararghe	Pawe	Humera/ Donkey	Hagereselam	Abuare	Kobo	Combined**
t_m	0.569(0.073)	0.479(0.071)	0.453(0.053)	0.331(0.067)	0.609(0.058)	0.654(0.054)	0.190(0.072)	0.505(0.027)
t_s	0.242(0.037)	0.231(0.042)	0.265(0.033)	0.230(0.055)	0.343(0.064)	0.376(0.046)	0.138(0.068)	0.205(0.014)
t_m-t_s	0.328(0.041)	0.248(0.037)	0.187(0.028)	0.101(0.023)	0.266(0.040)	0.278(0.032)	0.051(0.017)	0.300(0.016)
$r_{p(m)}$	0.732(0.127)	0.926(0.096)	0.138(0.040)	0.711(0.120)	0.545(0.115)	0.541(0.108)	0.205(0.230)	0.783(0.042)
$r_{p(s)}$	0.487(0.164)	0.724(0.167)	-0.069(0.038)	0.721(0.179)	0.238(0.084)	0.419(0.133)	-0.027(0.230)	0.512(0.058)
r_s	0.128(0.066)	1.100(0.183)	0.341(0.068)	0.408(0.145)	0.270(0.124)	0.239(0.084)	0.577 (0.177)	1.088(0.198)

*tm = multilocus outcrossing rate, ts = single locus outcrossing rate, $rp_{(m)}$=correlation of multilocus paternity, $r_{p(s)}$= correlation of single locus paternity, r_s= correlation of selfing among loci; ** Combined value doesn't include *Kobo* families

Table 4.13 Multilocus outcrossing (tm) rate of wild sorghum as affected by head length (HL), head width (HW), plant height (PH), and panicle compactness (PCOM)

Population*	HL	HW	PH	tm	PCOM**
Ghibe	36.9	18.1	335	0.569	1
Hararghe	24.8	12.5	183.5	0.479	6
Pawe	43.5	26.1	346	0.453	6
Humera/Donkey	30.8	10.5	207.5	0.331	6
Hagereselam	39.7	21.4	254	0.609	4
Abuare	35.8	26.7	256.5	0.654	6
Kobo	25.4	18.1	137.5	0.19	3
r	0.323	0.648	0.259	1	-0.374

*r= Pearson's correlation coefficient

**1= very lax panicle, 3=very loose, 4= loose, 6= semi-loose

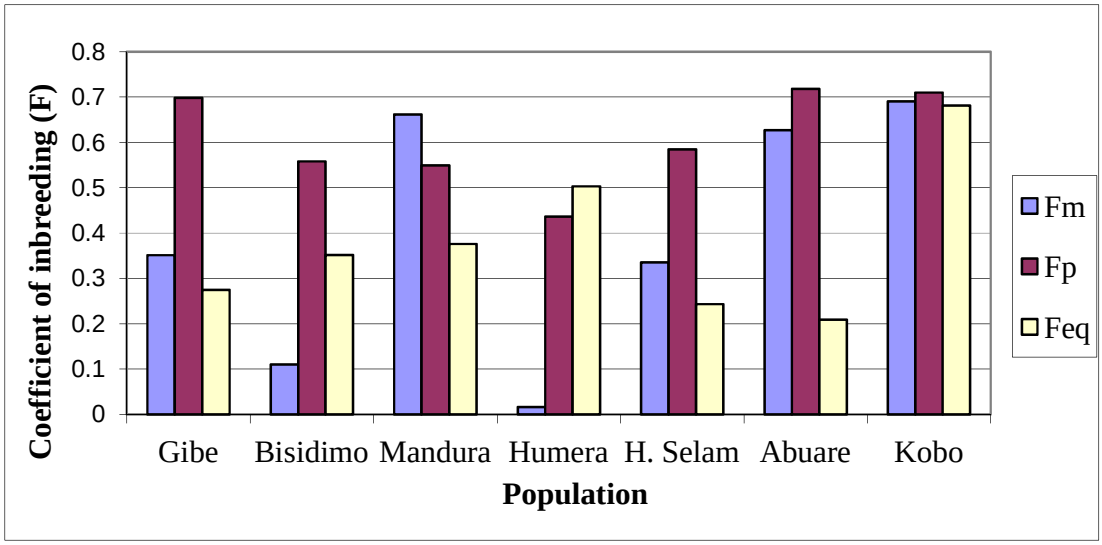


Fig. 4.10 The three coefficients of inbreeding estimated from the population and family data (see text for meaning of Fm, Fp, and Feq)

Biparental inbreeding was evident in all populations with variable magnitude as measured by the different procedures. First, the single locus outcrossing rate (t_s) was consistently smaller than that of the multilocus outcrossing rate (t_m) in all populations (Table 4.12). On average t_m-t_s = 0.24, which accounted for 48% of the total outcrossing rate in all families, but the extent was lower in *Humera* (0.101±SE0.023) and higher in *Ghibe* (0.328±SE0.041) families. Second, the coefficient of inbreeding of the progenies (F_p) was greater than F_{eq} in six of the seven populations, which reveals that there was more inbreeding in the progenies than expected from the observed outcrossing rate. Third, correlation of outcrossed paternity (r_p) was high in five of the seven populations indicating the contribution of few fathers for the pollen pool, which in turn was an indication of the presence of biparental inbreeding or non random mating. Fourth, the correlation of selfing among loci was less than one in all except in Bisidimo populations. Fifth, the single locus correlated paternity was less than the multilocus ones in all

populations except in Humera/Donkey. Evaluation of the coefficient of inbreeding, it was found that F_p was greater than both F_m and F_{eq}.

4.3. Morphology and fitness of wild × crop F_1 hybrids of sorghum

4.3.1. Performance of F_1 hybrids and their parents for juvenile survival, adult (Growth), and fertility traits

Analysis of variance has shown significant differences for all the measured adult survival, growth and fertility traits (Table 4.14). Laboratory seed germination and field emergence of crop-wild hybrids were compared with that in the parents as juvenile survival traits. Early emergence was regarded as better fitness. There were significant differences among parents and hybrids for laboratory germination and field emergence.

Mean comparisons of the various traits are presented in Table 4.15. Three of the 5 hybrids compared had better mean germination than their wild parents. In one case the hybrid had no significant germination than both of its parents. However, hybrids showed relatively later emergence than their wild parents.

Because the parents were extremely different for some characters, most of the hybrids assumed intermediate magnitude for most of the growth traits. Four of the seven hybrids had earlier flowering than their cultivar parents. The remaining three hybrids had flowering times not significantly different from both of their parents. All hybrids except W5-20 × 76T1#23 had heights significantly taller than their cultivar parents. Moreover, three were significantly taller than both parents.

All hybrids except W5-20 × 76T1#23 had significantly lower number of tillers per plant than their wild parents and significantly higher tillering than their cultivated parents. Five of the seven hybrids had significantly longer flag leaves than their wild parents, three of which had also significantly shorter flag leaf length than their crop parents. Two hybrids (W5-20 × 76T1#23 and IS18804 × 76T1#23) had leaf length not significantly different from their cultivar parent. All of the hybrids had significantly wider leaves than their wild parents but had significantly narrower leaves than their cultivar parents. All of the hybrids had significantly more number of leaves than their wild parents and less number of leaves than their cultivar parents except IS18804 × 76T1#23, which had leaf traits at par with its wild parent (Table 4.15). The average number of seeds per plant was in the range of 2249 (76T1#23) to 12363 (H2-16 × WSV 387). Virtually, all of the hybrids had higher fecundity than both of their parents, but it was relatively lower in W5-20 × 76T1#23. Generally, cultivar parents had the lowest fecundity than wild parents and hybrids.

Table 4.14 Mean squares from analysis of variance (ANOVA) of F1 hybrids and parents for different characters measured at Melkassa in 2011

Source	DF*	GERM	DTE	DTF	PH	HL	HW	LL	LW	TILL	LN	NSPP
Replication	4 (3)	10.29	3.67**	135.72*	3815.95*	34.31	17.40	11.24	0.37	8.33*	2.24	1919410
Entry	15(13)	383.65**	2.40**	206.49**	10575.5**	203.64**	220.46**	108.32**	15.10**	47.61**	6.32**	45219200**
Residual	60(39)	15	0.91	30.15	1305.71	25.19	20.75	8.49	0.22	2.65	0.23	3409020
Total	79(55)	101.88	1.33	68.97	3192.90	59.53	58.50	27.58	3.05	11.47	1.49	11272200

*Figures in parentheses show the DF for GERM; GERM= germination percentage; DTE= days to emergence; DTF= days to flowering; PH= plant height; HL= head length; HW= head width; LL= leaf length; LW= leaf width; TILL= number of tillers per plant; LN= leaf number; NSPP= number of seeds per panicle

Table 4.15 Mean of various life history traits measured on wild × crop F1 sorghum hybrids and their parents at Melkassa in 2011

Genotype	Strain	GERM*	DTE	DTF	PH	TILL	HL	HW	LL	LW	LN	NSPP	SSW
H2-1	wild parent	94^{c}	11^{bc}	57^{ab}	182.9^{bcde}	9.2^{cd}	28.9^{bc}	23.3^{defg}	39.4^{bcd}	2.9^{a}	4.7^{ab}	5365de	0.0096^{a}
H2-16	wild parent	93^{c}	10.8^{bc}	54^{a}	193.1^{bcdef}	9^{cd}	32^{bcdef}	25.4^{efg}	40.8^{bcde}	2.8^{a}	4.5^{a}	5766^{e}	0.0091^{a}
T1-1	wild parent	84^{b}	10.8^{bc}	67^{c}	198.9^{cdefg}	12.6^{d}	28.4^{bc}	23^{defg}	42.9^{defg}	3.7^{b}	4.4^{a}	4685^{cde}	0.0087^{a}
W5-20	wild parent	92^{c}	9.8^{abc}	52^{a}	172.5^{abcd}	4.3^{b}	26.8^{b}	14^{abc}	39.2^{bcd}	4.9^{cd}	5^{abc}	4228^{cde}	0.0148^{bc}
IS18822	wild parent	68	11.4^{c}	55^{ab}	225.2^{defgh}	8.6^{cd}	42.5^{gh}	29.3^{g}	33.9^{a}	3.8^{b}	4.8^{ab}	3847^{cd}	0.0157^{c}
IS14485	wild parent	86^{b}	10^{abc}	53^{a}	227.9^{defgh}	6.9^{c}	36.9^{efgh}	25.9^{fg}	38.3^{bc}	4.6^{c}	4.9^{abc}	4112^{cde}	0.017^{c}
IS18804	wild parent	75^{a}	10.6^{bc}	60^{abc}	246.9^{fghi}	9^{cd}	37.8^{fgh}	26.9^{fg}	37.6^{ab}	3.5^{ab}	5.2^{abcd}	3384^{bc}	0.0112^{ab}
76T1#23	cultivar parent	99^{d}	8.8^{a}	63^{bc}	126.4^{a}	1.2^{a}	19.8^{a}	7.9^{a}	48.6^{h}	8.6^{h}	7.3^{f}	2249^{a}	0.0182^{c}
WSV 387	cultivar parent	98^{d}	11^{bc}	77^{d}	143.8^{ab}	1.3^{a}	30.4^{bcde}	10.8^{a}	53.2^{i}	8.7^{h}	8.9^{g}	2553^{ab}	0.0246^{d}
H2-1 × WSV 387	F1 hybrid	93^{c}	10^{abc}	53^{a}	264.6^{hi}	6.9^{c}	33^{bcdef}	19.1^{cde}	44^{efg}	5^{cd}	5.4^{bcd}	9651^{fg}	0.016^{c}
H2-16 × 76T1#23	F1 hybrid	NA	10^{abc}	57^{ab}	234^{efgh}	6.4^{c}	34.6^{cdef}	21.4^{def}	43.3^{defg}	5.2^{cde}	5.4^{bcd}	12363^{g}	0.0159^{c}
T1-1 × WSV 387	F1 hybrid	100^{d}	10.2^{abc}	59^{abc}	252.7^{ghi}	7^{c}	39^{fgh}	24.2^{efg}	42.4^{cdef}	5.8^{ef}	5.6^{bcde}	10756^{fg}	0.014^{bc}
W5-20 × 76T1#23	F1 hybrid	99^{d}	9.5^{ab}	58^{abc}	163.1^{abc}	3.1^{b}	26.6^{b}	11.4^{ab}	46.8^{gh}	6.2^{fg}	5.8^{cde}	4696^{cde}	0.0181^{c}
IS18822 × WSV 387	F1 hybrid	99^{d}	11^{bc}	54^{a}	296^{i}	4.3^{b}	44^{h}	28.8^{g}	43.1^{defg}	5.6^{def}	5.5^{bcd}	4300^{cde}	0.0234^{d}
IS14485 × 76T1#23	F1 hybrid	NA	9.6^{ab}	54^{a}	218.9^{defgh}	4.2^{b}	29.4^{bcd}	17.1^{bcd}	42.3^{cdef}	6.5^{g}	6.2^{de}	5132de	0.0225^{d}
IS18804 × 76T1#23	F1 hybrid	98^{d}	10.2^{abc}	57^{ab}	239.8^{fgh}	6.5^{c}	36.5^{defg}	21.8^{def}	46^{fgh}	5.3^{de}	5.6^{bcde}	8924^{f}	0.0175^{c}

*NA= Not Available (seed was not adequate), Figures preceded by similar letter are not significantly different ($p<0.05$); GERM= germination percentage; DTE= days to emergence; DTF= days to flowering; PH= plant height; HL= head length; HW= head width; LL= leaf length; LW= leaf width; TILL= number of tillers per plant; LN= leaf number; NSPP= number of seeds per panicle; SSW= single seed weight

4.3.2. Mid-parent heterosis of wild × crop crosses

Mid-parent heterosis values are presented in Table 4.16. Three of the five hybrids, which had data, had significant mid-parent heterosis for germination. There was no hybrid that showed significant mid-parent heterosis for days to seedling emergence and number of tillers per plant. Most hybrids exhibited earlier flowering (10-20%) than their mid parents, but three of the seven hybrids had significant mid-parent heterosis. All of the hybrids except W5-20 × 76T1#23 had significantly high mid-parent heterosis for plant height. Three hybrids showed significant mid-parent heterosis for head length and 2 others had significant mid-parent heterosis for head width. Most of the hybrids had head shapes intermediate between both of their parents (Fig. 4.11). Four hybrids had significant mid-parent heterosis for leaf number, 4 hybrids for leaf width and one for leaf length. Four hybrids exhibited significantly higher mid-parent heterosis for number of seeds per panicle. There was no cross that showed consistently superior mid-parent heterosis for all the characters. However, two crosses, T1-1 × WSV 387 and IS18822 × WSV 387 exhibited significant mid-parent heterosis for most of the traits. On the contrary, one cross (W5-20 × 76T1#23) had not significantly different from the mid-parent for all the characters except for leaf width.

Table 4.16 Mid parent heterosis (%) for different life history traits measured in wild × crop hybrids

Entry	Cross	GERM	DTE	DTF	PH	TILL	HL	HW	LN	LL	LW	NSPP
1	H2-1 × WSV 387	-3.1	-9.1	-21.4**	62.0**	31.4	11.3	12.0	-20.6**	-17.3	-13.8**	143.8**
2	H2-16 × 76T1#23	-	2.0	-3.4	46.5**	25.5	33.6**	28.5	-8.5	-10.9	-8.8	208.5**
3	T1-1 × WSV 387	9.9**	-6.4	-18.1**	47.5**	0.7	18.4**	43.2*	-15.8**	-20.3**	-6.5	197.2**
4	W5-20 × 76T1#23	3.7	2.2	0.9	9.1	12.7	14.2	4.1	-5.7	-3.7	-8.1*	45.0
5	IS18822 × WSV 387	19.3**	-1.8	-18.3**	60.4**	-13.1	20.7	43.6**	-19.7**	-19.0	-10.4*	34.4
6	IS14485 × 76T1#23	-	2.1	-6.6	23.6*	3.7	3.7	1.2	1.6	-13.0	-1.5	61.4
7	IS18804 × 76T1#23	12.6**	5.2	-7.0	28.5*	27.5	26.7*	25.3	-10.4*	-5.3	-12.4**	216.8**

* Significantly different ($p<0.05$), ** significantly different ($p<0.01$); GERM= germination percentage; DTE= days to emergence; DTF= days to flowering; PH= plant height; HL= head length; HW= head width; LL= leaf length; LW= leaf width; TILL= number of tillers per plant; LN= leaf number; NSPP= number of seeds per panicle

4.3.3. Relative fitness

Hybrids showed superior fitness for plant height and number of seeds per panicle (Table 4.17). Cultivar parents also had higher fitness for leaf characters (flag leaf length and width, and leaf number). Fitness of the cultivars was also higher for germination, but the difference with hybrids was not significant. Wild parents had the lowest fitness for these characters. However, their fitness was higher for number of tillers per plant. Hybrids had better reproductive fitness than both of their parents. However, as far as weediness is concerned the hybrids should be compared with their wild parents (Snow *et al.*, 1998). Hence, hybrids had better juvenile and reproductive fitness than their wild parents, but differences for growth fitness were not significant. Hybrids also had the highest composite fitness (0.85).

Table 4.17 Juvenile survival, growth and fertility fitness of wild-crop sorghum hybrids relative to their parents

Genotype	Order*	Type	Juvenile fitness		Growth fitness						Reproductive fitness	composite fitness
			GERM	DTE	DTF	PH	TILL	LL	LW	LN	NSPP	
WSV 387	1	cultivar	1	0.9	0.55	0.54	0.14	1	1	1	0.26	
76T1#23	2	cultivar	NA	1	0.83	0.54	0.13	1	1	1	0.18	
WSV 387	3	cultivar	0.98	1	0.69	0.57	0.1	1	1	1	0.24	
76T1#23	4	cultivar	1	1	0.79	0.73	0.28	1	1	1	0.48	
WSV 387	5	cultivar	0.99	1	0.57	0.48	0.15	1	1	1	0.59	
76T1#23	6	cultivar	NA	1	0.81	0.55	0.17	1	1	1	0.44	
76T1#23	7	cultivar	1	1	0.89	0.51	0.13	1	1	1	0.25	
Mean			0.99	0.99	0.73	0.56	0.16	1	1	1	0.35	0.63
H2-1 × WSV 387	1	F1	0.95	1	1	1	0.75	0.83	0.57	0.61	1	
H2-16 × 76T1#23	2	F1	NA	0.86	0.94	1	0.71	0.89	0.6	0.74	1	
T1-1 × WSV 387	3	F1	1	0.93	1	1	0.55	0.8	0.67	0.63	1	
W5-20 × 76T1#23	4	F1	1	0.92	0.88	0.94	0.72	0.96	0.72	0.79	1	
IS18822 × WSV 387	5	F1	1	1	1	1	0.5	0.81	0.64	0.62	1	
IS14485 × 76T1#23	6	F1	NA	0.91	0.98	0.96	0.61	0.87	0.75	0.85	1	
IS18804 × 76T1#23	7	F1	0.99	0.84	1	0.97	0.72	0.95	0.62	0.77	1	
Mean			0.99	0.92	0.97	0.98	0.65	0.87	0.65	0.72	1	0.85
H2-1	1	wild	0.96	0.9	0.92	0.69	1	0.74	0.33	0.53	0.55	
H2-16	2	wild	NA	0.77	1	0.82	1	0.84	0.32	0.62	0.45	
T1-1	3	wild	0.84	0.98	0.86	0.79	1	0.81	0.42	0.49	0.43	
W5-20	4	wild	0.93	0.89	1	1	1	0.81	0.57	0.68	0.9	
IS18822	5	wild	0.69	0.96	0.98	0.76	1	0.64	0.44	0.54	0.89	
IS14485	6	wild	NA	0.86	1	1	1	0.79	0.53	0.67	0.8	
IS18804	7	wild	0.76	0.8	0.95	1	1	0.77	0.41	0.71	0.38	
Mean			0.84	0.88	0.96	0.87	1	0.77	0.43	0.61	0.63	0.69

* Related groups have similar numbering according to their order

Fig. 4.11 Differences in panicle morphology between hybrids (in the middle of each slide) and their crop (right) and wild (left) parents. A) IS14485 × 76T1#23, B) IS18804 × 76T1#23, C) W5-20 × 76T1#23

4.4. Ecotypic variation for seed dormancy, longevity and germination requirements in wild/weedy sorghum

4.4.1. Fate of the seeds of different ecotypes of wild sorghum after 18 and 24 months of burial in the soil

General Linear Model (GLM) ANOVA showed no significant differences among genotypes and among the regions of collection, but it was significant ($p<0.01$) for the duration of burial (Table 4.18). The significance of the second order interaction ($p<0.01$), Genotype × Duration of burial × Region indicates that genotypes (ecotypes) collected from different geographical regions showed differential response on seed viability to the different periods of burial. Some of them showed drastic decline in viability, but others showed relatively slower, but ultimately, most of the seeds were depleted within the 24 months of burial.

The data of all populations in set-1 pooled together, 71% of all the seeds were viable of which 49.7% were germinated in the laboratory, 21% were dormant and 24.14% were germinated and/or dead in the field during the first 6 months of burial (Table 4.19). The remaining 4.8% of the seeds were found intact but were confirmed as died using 1% TTC test. Hence, a large proportion of the seeds were found to be viable after 6 months of burial in the soil. After 12 months of burial, however, the viability sharply declined to 9.06% of which 2.32% were dormant. After 18 months of burial in the soil, only 1.82% of the seeds remained viable of which 1.34% were still dormant. Furthermore, after 24 months of burial, 1.24% remained viable, of which 0.74% were still dormant. Fig. 4.12 shows the survival curve of wild sorghum seeds during the 24 months period of burial in

the soil. On the basis of the total viability of seeds in the entire population, ecotypic variation in the wild sorghum populations was not evident during the first exhume, but there were differences among ecotypes for dormancy (the proportion of seeds that confirmed viable by TTC test, but did not germinate). The highest dormancy was recorded by Ghibe (39.1%) and Pawe (38.5%) populations, but the lowest (4.2%) was by Hararghe population. After a year of burial, Ghibe population showed better total viability than the rest of the populations included in the experiment followed by Pawe. During the third exhume, Pawe and Ghibe populations still showed better total viability than the remaining populations. After 24 months of burial, Ghibe populations had an average viability of 4.8 %, more than 50% of which (2.6%) were still dormant. The other populations had <1% viability.

Table 4.18 General Linear Model (GLM) ANOVA showing the effect of duration of burial on viability of seeds from different wild sorghum populations collected in 5 geographical regions*

Source	DF	SS	MS	F	Prob
Rep	3	0.06904	0.02301	0.86	0.461
Genotype (G)	4	0.0956	0.0239	0.9	0.466
Region (R)	4	0.20486	0.05121	1.92	0.108
Duration of burial (D)	2	40.72649	20.36325	764.33	0.000
G×R	16	0.67858	0.04241	1.59	0.072
G×D	8	0.10915	0.01364	0.51	0.847
R×D	8	0.7411	0.09264	3.48	0.001
G×R×D	32	1.50215	0.04694	1.76	0.01
Error	222	5.91454	0.02664		
Total	299	50.04149			

*values are based on Arcsine transformed viability data

Table 4.19 Survival of buried seeds from five populations (ecotypes) of wild sorghum after 6, 12, 18, and 24 months, based on pooled data from 20 mesh bags per population

Population	Average percent of viable seeds	Percent germinated seeds in lab (non-dormant)	Percent TTC viable seeds (dormant)	Percent intact but dead seeds (TTC non-viable)	Percent germinated/ decomposed in field
Excavated after 6 months of burial					
Ghibe	67.1	28	39.1	10.2	22.7
Hararghe	75.2	71	4.2	2.5	22.3
Pawe	66	27.5	38.5	2.5	31.5
Tigray	73.6	60.2	13.4	2.1	24.3
Wello	73.4	61.7	11.7	6.7	19.9
Average	71.06	49.7	21.38	4.8	24.14
SD	4.19	20.44	16.28	3.56	4.40
Exhumed after 12 months of burial					
Ghibe	22.5	17.8	4.7	8.1	69.4
Hararghe	4.8	2.6	2.2	3.4	91.8
Pawe	9.2	7.7	1.5	5.4	85.4
Tigray	4.3	3.1	1.2	0.5	95.2
Wello	4.5	2.5	2.0	1.0	94.5
Average	9.06	6.74	2.32	3.68	87.26
SD	7.78	6.55	1.39	3.16	10.71
Exhumed after 18 months of burial					
Ghibe	3.4	1	2.4	3	96.6
Hararghe	0	0	0	0	100
Pawe	4.3	1.4	2.9	3.2	95.7
Tigray	1.3	0	1.3	1.5	98.7
Wello	0.1	0	0.1	1	99.9
Average	1.82	0.48	1.34	1.74	98.18
SD	1.95	0.67	1.31	1.36	1.95
Exhumed after 24 months of burial					
Ghibe	4.8	2.2	2.6	12.8	82.4
Hararghe	0.2	0	0.2	0.9	98.9
Pawe	0.7	0.2	0.5	6.1	93.2
Tigray	0.2	0.1	0.1	0.7	99.1
Wello	0.3	0	0.3	0.5	99.2
Average	1.24	0.5	0.74	4.2	94.56
SD	2	0.95	1.05	5.35	7.26

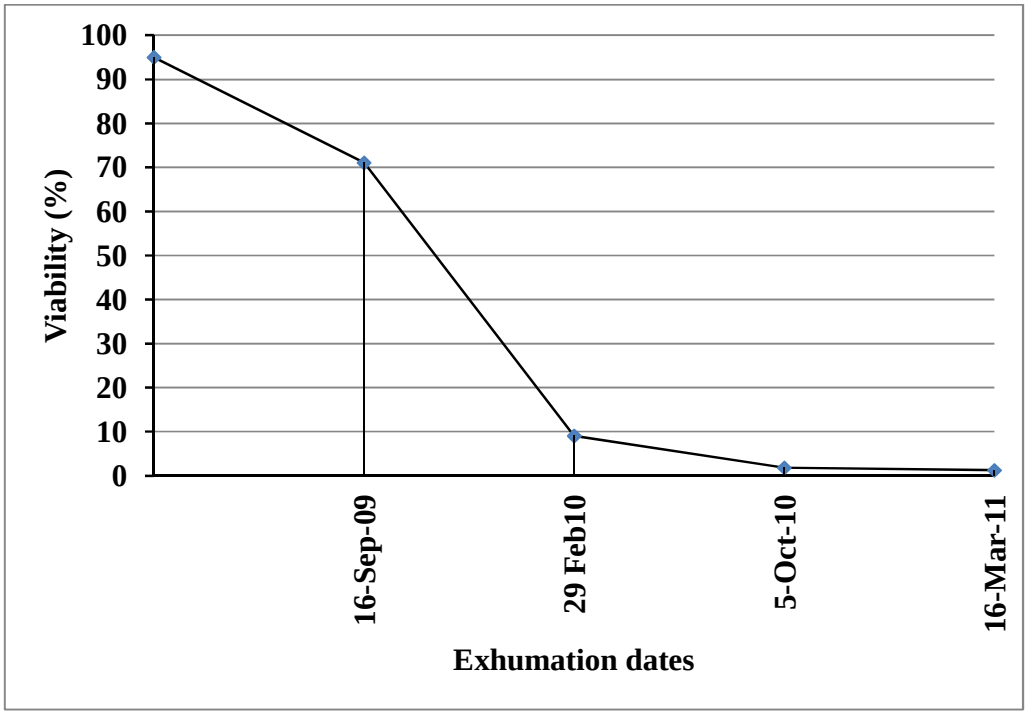


Fig. 4.12 Survival of 25 wild sorghum population seeds after 24 months of burial in soil

The differences among all factors, except period of burial-by-depth interaction were highly significant for total viability (survival). However, because the second-order interaction (D × t × G) was highly significant, all of the three factors were not independent of each other. Hence, different genotypes had different viability after some period of burial at different burial depths. Differences for viability associated with dormancy were highly significant only for genotype factor (Table 4.20). In general, crop seeds lost their viability within the first 6 months regardless of the depth of burial (Fig. 4.13). However, wild sorghum retained their viability up to 31% and 33% in Ghibe and Pawe populations, respectively after 18 months of burial at 20cm soil depth (also Table 4.21). At a depth of 10cm the rate of decline in viability was sharp in Pawe populations

compared to the Ghibe populations. At 20cm depth, however, both showed more or less similar pattern of decline.

Table 4.20 General Linear Model (GLM) ANOVA showing the effect of burial time and depth on longevity of two wild sorghum ecotypes and a cultivar seed (Rep=random, others=fixed) for total viable and dormant (TTC stained) seeds

Source	DF	MS (Total viable)	MS (TTC viable/ dormant)
Replication	2	0.016	0.001
Depth of burial (D)	1	0.235**	0.017
Duration of burial (t)	2	0.403**	0.027
Genotype (G)	2	1.388**	0.293**
D × t	2	0.002	0.01
D × G	2	0.126**	0.013
t × G	4	0.130**	0.012
D × t × G	4	0.074**	0.028
Error	34	0.017	0.015

** significant ($p<0.01$), Arcsine transformed percentage data were used for analysis

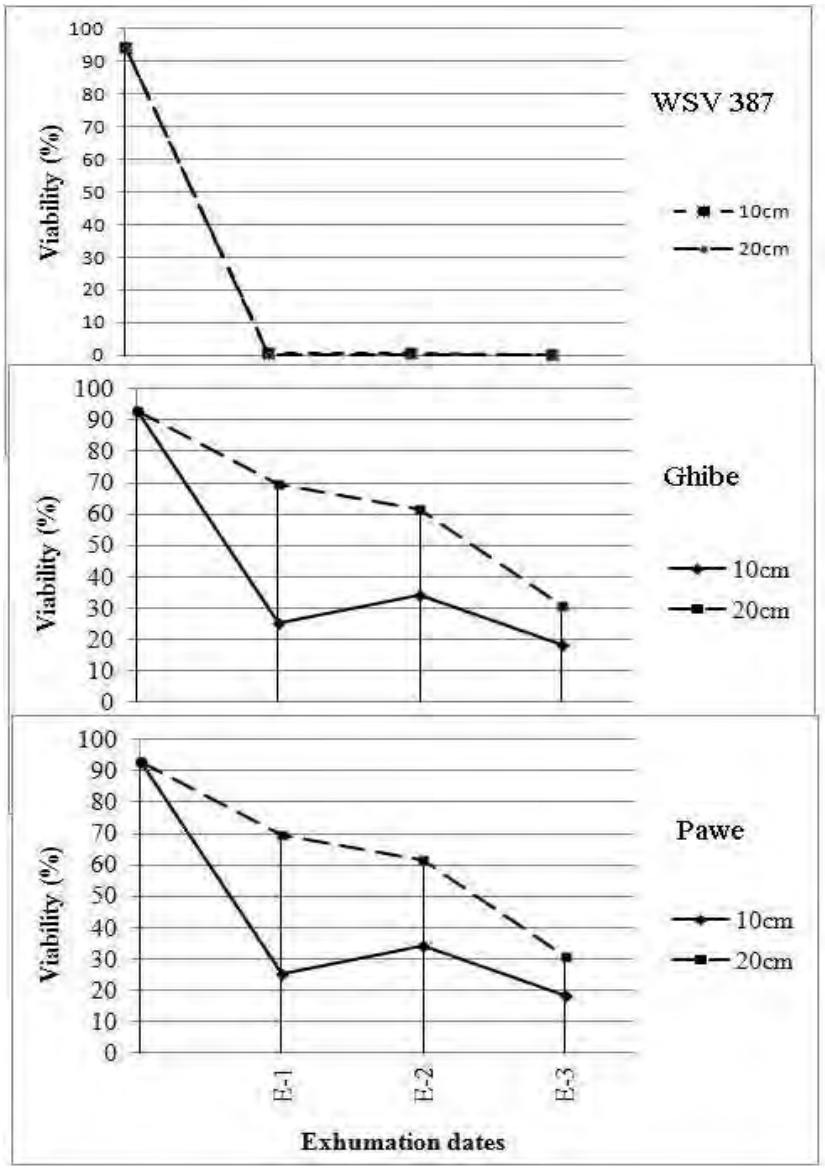


Fig. 4.13 Effect of burial time and depth on viability of wild sc Pawe ed collected from Ghibe and Pawe and sorghum variety WSV 387 (E-1= 25 Sept. 2010, E-2= 16 Mar 2011, E-3= 03 Nov 2011)

Table 4.21 Percent of viable seeds, the rate of decline and the time it takes for 50% of wild and cultivated sorghum seeds to lose viability after 0.5, 1 and 1.5 years of burial in the soil

		10cm depth			20cm depth		
population	Duration of burial, t (years)	Average viable seeds (%)	Rate of decline (r)	1/2t	Average viable seeds (%)	Rate of decline (r)	1/2t
	0.5	72.7	0.64	1.085	57.3	1.11	0.623
	1	55.3	0.59	1.171	67.3	0.40	1.753
Ghibe	1.5	4.0	2.15	0.323	32.7	0.75	0.929
	0.5	25.3	2.75	0.252	69.3	0.73	0.946
	1	34.0	1.08	0.643	61.3	0.49	1.418
Pawe	1.5	18.0	1.14	0.606	30.7	0.79	0.880
	0.5	0.7	10.02	0.069	0.0	0.60	0
	1	0.7	5.01	0.138	0.0	0.60	0
WSV387	1.5	0.0	0.60	0.087	0.0	0.60	0

*1/2t is the time at which 50% of the seeds lose viability at time, t

4.4.2. Response of wild sorghum ecotypes for varying moisture and temperature regimes

General Linear Model ANOVA showed that differences among all factors and their interactions were significant ($p<0.01$) (Table 4.22). The three genotypes revealed differential response to the varying levels of temperature and moisture regimes for seed germination. At all temperature levels, germination was greater or equal to 80% in all the genotypes under no osmotic potential (control) condition. In each genotype there was a trend of increase in germination with an increase in temperature in each level of osmotic potential. Accordingly, at the lowest temperature (15°C), all of the three genotypes showed no germination less than Ψs=-0.6 MPa of osmotic potential (Fig. 4.14). At 23°C all of them showed germination at additional Ψs= -0.9 MPa, and at 30°C, all of them showed germination at all levels of osmotic potential. In general, there was a trend that decrease in osmotic potential (increase in moisture stress) decreases seed germination in

all genotypes, but the extent was different with temperature and genotype. There was no significant difference between 0 MPa and 0.3 MPa, and even better germination was observed at 0.3 MPa than the control in some genotypes. At the highest temperature (30°C) W-4 population showed the highest germination when there was mild osmotic potential (Ψs= -0.3 MPa) (Fig. 4.14). This germination was comparable to the highest germination in the control treatment levels at all other temperature levels for the same population. Similarly, at Ψs=0.6 MPa, this population showed better germination than the other two populations.

Table 4.22 Mean square from General Linear Model (GLM) ANOVA showing the effect of individual and combinations of temperature and osmotic potential on seed germination of wild sorghum seed

Source	DF	MS
Rep	3	0.015
Genotype (G)	2	0.992**
Temperature (T)	2	0.381**
Osmotic Potential (OP)	4	10.052**
G*T	4	0.048**
G*OP	8	0.201**
T*OP	8	0.132**
G*T*OP	16	0.035**
Error	132	0.019
Total	179	

**, significant ($p<0.01$)

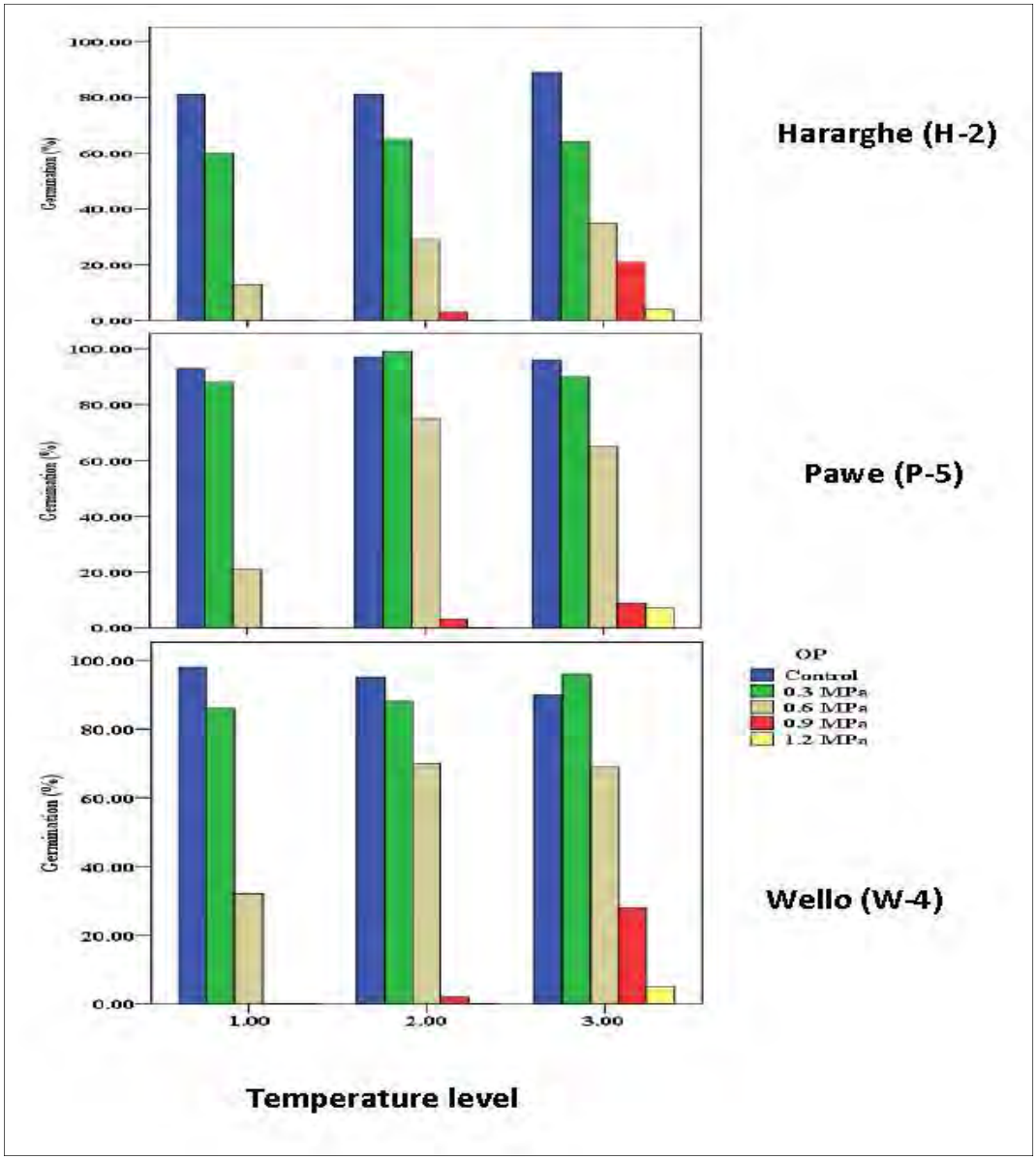


Fig. 4.14 Response of wild sorghum populations from Hararghe (H-2), Pawe (P-5) and Wello (W-4) for differential osmotic potential at different levels of temperature (1=15°C, 2=23°C, and 3=30°C)

5. DISCUSSION

5.1. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (L.) Moench] in Ethiopia

5.1.1. Distribution and *in situ* phenotypic diversity of wild sorghum in Ethiopia

Knowledge of the distribution of wild and weedy relatives of crop species is essential for their conservation and risk assessment of transgenic crops (Chandler and Dunwell, 2008). This is due primarily to two reasons. First, gene flow from transgenic crops to their wild relatives may aggravate weediness and also reduces genetic diversity in the wild. Second, related species may provide a genetic bridge to plants more remotely related to the transgenic crop and hence a comprehensive survey is required to identify any plants in the taxonomic group that may be noxious weeds (Gealy *et al.*, 2003; Chandler and Dunwell, 2008). In Ethiopia, wild sorghums exist in sympatric and allopatric ranges to the cultivated sorghum in a wide range of geographical regions. In addition to the survey covered by this study, early herbarium specimens from even a lower altitude of 180m at the bank of Wabishebele River in the Somali regional state are preserved in the national herbarium. Herbarium specimens are also available for Welega in Oromiya and Konso and Omo areas in the Southern Ethiopia regional state.

In most sorghum growing areas of Ethiopia farmers practice mixed cropping system and the farming system is predominantly mixed crop-livestock production. In the Ethiopian settings particularly in South Wello and North Shewa, sorghum is cultivated in close association with oil crops (*Guizotia abyssinica* and *Sesamum indicum*) mainly because of their combined uses in the cultural feeding system of local people and the presence of

multipurpose sorghum land races in the area has played a significant role for oil crop species diversity and their *in situ* conservation (Mulatu Geleta, 2001). It is common to most areas that the farmers notice the weedy sorghums only after flowering, and when they do so, they cut them and feed to their livestock around homesteads. In Kenya farmers consume the intermediate wild sorghum forms (crop-wild hybrids) (Mutegi *et al*., 2010), but there was no similar evidence in the studied areas of Ethiopia.

In areas where crop-livestock farming system is practiced it is hard to find wild sorghum in open fields such as abandoned farm lands. However, those plants grown nearby fences and in areas inaccessible to livestock have the chance to escape grazing and thus can mature and shatter their seeds. Due to this disturbance, wild sorghums that are isolated from the cultivated sorghum are found in patches. This patch occupancy and relative abundance vary from place to place due to different farmers' practices. Also, the proportion of the weedy sorghum to the cultivated sorghum in a given farmer's field was found to be variable with location. Some farmers remove the wild sorghums during flowering, but others do not notice until maturity. For instance, in a single farmer field at Gibe (G-5), there were 39 weedy sorghum plants, but there were only 20 cultivated sorghum plants in a quadrant (4m × 4m) of land after the first weeding (the removed plants were laying under). In many places farmers responded that wild sorghums preferentially grow in soils with high organic matter (manure) content. However, we analyzed soil samples from 6 sites for organic matter and found no correlation between genetic diversity and soil organic matter content. Probably it is not because the wild

sorghums prefer high organic matter; instead they are dispersed by cow manure as they are good fodder sources for livestock.

Unlike the reported high diversity in the cultivated sorghum landraces in Hararghe (e.g. Amsalu Ayana and Endashaw Bekele, 1998; Amsalu Ayana *et al*., 2000; Nemera Geleta *et al*., 2006), it was difficult to find a single head of wild and weedy sorghum during the season when survey was made. Similarly, Tesfaye Tesso *et al*. (2008) reported low frequency of wild sorghum occurrence in Hararghe during their survey. Farmers in that area predominantly grow compact headed durra type sorghum landraces; hence, the loose panicle wild sorghum plants are easily identified and removed by farmers during flowering and fed to livestock (also, Tesfaye Tesso *et al*., 2008). Perhaps it is due to such kind of farmers' practice that there were no stabilized hybrid forms with compact panicle those mimic the durras as they were in Zobel (Wello). Another possible reason is that due to the absence of land that remained uncultivated and the use of even marginal lands for the cultivation of *Chat* (*Catha edulis*) wild sorghums have no place to grow. Hence, wild sorghums were not found in Hararghe following the major routes, except at two locations, Bisidimo and Gelemso-Bedesa road side during the season. In the former location, farmers witnessed that they had never seen the weedy sorghums those were grown during the season intermixed with the cultivated sorghum variety, which in turn was also different from the farmers' durra landrace. It was said to have appeared in the area recently with seed aid. The wild sorghum near the road sides past Bedesa town to Gelemso were found in tef field, which were similar in their morphology to those found in some parts of Wello (Kobo-Alamata plain) and central Shoa. Like in the Bisidimo

area, farmers in Bedesa area said that the wild sorghums were recently escaped from the nearby tree seedling nursery managed by the district agricultural bureau and thus they might not be native to the area.

Earlier, Stemler *et al*. (1977) reported that the places where wild sorghum grows in Ethiopia are limited and relatively unpopulated. The current observation also confirmed that they mainly occupy patches wherever they are not intermixed with cultivated sorghum. However, they are widely grown in the Ethiopian lowlands (<1600m) and to some extent in the intermediate altitudes (1600-1900m). Of course, the wild and weedy sorghums were more abundant in the Benishangul-Gumuz region (Metekel zone) than in any other sites where the survey was made. This can be because of the early colonization of the species in that area as it is one of the areas situated west of the Ethiopian highlands where sorghum was believed to be first domesticated (Stemler *et al*, 1977). Moreover, the high rainfall and hot climate might have favored the expansion of the wild and weedy sorghums in that area though they are supposed to thrive dry climates as well. Still, the other reason can be the relatively larger area of land that remained uncultivated in that region and the wide natural forest cover. Compared to the Wello and Tigray areas the Benishangul-Gumuz area adopted agriculture in recent historical times. The farmers in that region very rarely rear livestock (e.g., Ahrens, 1996) as a result of which the wild sorghums were freely expanding.

In general, high diversity was observed in wild/weedy sorghum for the measured phenotypic characters. Some characters like white, yellow and purple seed colors are rare

and found only in specific sampling sites. Principal component analysis of the quantitative traits also revealed some level of geographic structuring in wild sorghum populations originated from diverse regions of Ethiopia. In the principal component biplot (Fig. 4.1) five of the six Pawe populations clustered together and the Tigray populations (T-1, T-2, T-3, and T-5) also clustered together. Moreover, the three South Wello populations (W-1, W-2, and W-3) clustered together. In some cases the clustering seems to have racial background. For instance, T-1, T-2, T-3, H-2, W-7 and W-8 took their position in one side of the PCA biplot. These populations are morphologically similar to one another in that they have short stature, loose panicle, weak culms and perhaps the 'typical wild' sorghum forms. They also share similar habitat, found either in tef or in barley fields. Morphologically, they resemble the already identified race *arundinaceum* accessions from ICRISAT. The well stabilized hybrid forms of the weedy sorghum H-1, T-6, T-7, and W-6 also took one side of the PC biplot.

The number of populations sampled for phenotypic data analysis was not equal for all geographical regions as they primarily differ in the area coverage and abundance of wild sorghum. They were 6 for each of Ghibe and Pawe, 2 for Hararghe, and 8 for each of Tigray and Wello. Despite this inconsistency in the number of samples, correlation between phenotypic diversity and sample size was not significant ($r=0.672$, $p=0.214$).

5.1.1.1. Implications for gene flow and conservation

One of the limitations of using morphological based genetic diversity estimates is that these characters are limited in number and are influenced by the environment (van

Beuningen and Busch, 1997). The measurements of genetic diversity using different techniques would have better meaning if they are carried out *in situ*. Although it is common to use gene bank accessions for determining population genetic studies in sorghum, inferences based on these conserved *ex situ* collections may have little relevance for current ecological and evolutionary assessments because the accessions were often collected years ago and may no longer represent sorghum populations at the site of collection (Asfaw Adugna *et al*., 2011). The present *in situ* morphological based study showed that wild and weedy sorghums are highly diverse in Ethiopia. This variability could be important source for future breeding programs provided that the bridge between pre-breeding and breeding is well established. In the surveyed areas due to extreme cultural and social values associated with the crop, farmers grow sorghum after sorghum, which aggravates the buildup of the weedy sorghums and helps them to become noxious. Nonetheless, most of this weed seeds could have been depleted if rotation is practiced due to loss of viability after a year (See section 5.4).

During domestication from wild sorghum, humans selected characters desirable for their own use like large seed, reduced glume covering and low shattering. These are domestication traits. For instance, high shattering (67% in the present experiment, Table 4.2) is the most conspicuous adaptation that separates wild from cultivated cereals (Harlan *et al*., 1973), the more stabilized forms being with minimum shattering. Due to disruptive selection of humans on one direction and nature in another direction the diversity has increased within the species (Dogget and Majisu, 1968), which was further facilitated by outcrossing (hybridization) with other species of sorghum. Hybridization

between the crop and the wild/weedy sorghum is not a recent phenomenon and has an age since the domestication of the sorghum crop. However, gene flow through hybridization between crops and their wild relatives has received attention in recent times due to the potential risks associated with the introduction of transgenic crops. To date no transgenic sorghum is under commercial production. However, nutritionally enhanced transgenic sorghum is being developed for Africa by the Africa Biofortified Sorghum (ABS) project (Zhao, 2007). The present study has shown that gene flow is likely to have occurred between Ethiopian cultivated and wild sorghums as the two congeners coexist, have overlapping flowering windows and there is no apparent genetic barrier for hybridization. Weedy sorghums in cultivated sorghum fields that show well stabilized morphological domestication traits due to repeated pollen mediated gene flow and introgression could be good evidences for this. This may call for effective conservation (*ex situ*), transgenic mitigation and containment measures to be in place if transgenic sorghum is to be deployed in Ethiopia.

5.1.2. Population genetic structure, gene flow and genetic diversity in wild and weedy sorghum [*Sorghum bicolor* (l.) Moench] using SSR markers

5.1.2.1. Comparisons among wild populations, accessions, and cultivars

This study revealed significantly higher levels of diversity in wild population compared to cultivars and accessions in terms of overall allelic richness, private allelic richness and gene diversity. In similar surveys, previous investigators also reported greater genetic diversity in wild accessions than in sorghum cultivars (e.g., Aldrich *et al*.1992; Casa *et al*. 2005; Mutegi *et al*., 2011), consistent with the expectation that cultivars have

undergone genetic bottlenecks during domestication and further breeding. However, in a recent study at a local scale in Cameroun (Barnaud *et al*., 2009), in Mali and Guinea, Sagnard *et al*. (2011) and in Kenya (Mutegi *et al*., 2012) reported similar levels of genetic diversity between cultivated and wild sorghum, which they attributed to crop-to-wild gene flow. The authors also reported higher levels of diversity in a pool *ex situ* wild accession compared both to extant wild and cultivar sorghum, which is opposite to what was observed in this study. The significantly lower genetic diversity observed in the 10 ICRISAT accessions originating from 6 African countries suggests that extant wild populations from Ethiopia are more diverse than these *ex situ* collections. This result also corresponds well with the suggestion that Ethiopia is a center of origin and diversity for sorghum.

5.1.2.2. Patterns of genetic structure in wild sorghum populations

Although much smaller than reported earlier in cultivated sorghum (E.g. Djé *et al*., 2000; Ghebru *et al*., 2002), the fairly high and significant F_{ST} value (0.37, P <0.001 based on 1000 randomizations) observed in wild sorghum in this study shows high differentiation among the various populations collected from geographically different locations, in addition to being highly diverse. In comparison, Amsalu Ayana *et al*. (2001) reported that 25% of the variation in RAPD markers was found among 10 *in situ* populations in central, south western and western Ethiopia. In the current study, four widely separated populations (W-4, T-1, AW, and H-2) were closely related to each other and to the ICRISAT accession of subspecies *drummondii* from Ethiopia (Fig. 4.5). Three of these populations were collected in tef fields, close to cultivated sorghum, while the Awash

(AW) population was found in a national park and was isolated from any cultivated fields. Subspecies *drummondii* is often described as a stabilized crop-wild hybrid lineage (e.g. Doggett, 1988; Harlan and De Wet, 1972), and the data of this study suggest that these populations form a fairly distinct group. The four wild populations shared morphological characteristics such as short stature (91-230 cm), weak stalks, profuse tillering, dense racemes, the peduncle well *exserted*, narrow leaves (2-5cm wide), the grain fully covered, and awned glumes. Another shared feature was that three of the 12 SSR primers those were originaly screened failed to amplify for most of these samples (see Methods).

Two of the three Ghibe populations clustered with race *aethiopicuum*-1 from Sudan. One population from Wello with compact panicles clustered with Sudan *drummondii* (Fig. 4.2). Neighboring wild populations often clustered together, especially in Pawe, while others showed greater affinity with distant populations (Figs. 4.1). The latter could be related to long-distance seed movement, perhaps as contaminants of crop seed. For example, during interview farmers at site H-1 (Bisidimo) said that wild sorghum was a recent introduction to their area and its appearance coincided with food grain sorghum aid. Moreover, the rare yellow seed color was observed in these two regions only (see section 4.1.1). Therefore, clustering of the H-1 population from eastern Ethiopia with Tigray populations near Sudan may occur because grain that was given as relief to farmers was perhaps purchased in Tigray and the wild sorghum seeds might have moved with it. Alternatively, these two areas might have received sorghum grain as food aid from the same source (donor). Moreover, relatedness could be due to common ancestry.

In Kenya, Mutegi *et al*. (2011) reported consistently strong spatial genetic structure in a study of wild sorghum plants collected from throughout the country. These authors attributed correlations between genetic and geographic distances to patterns of seed movement within regions. Interestingly, sorghum landraces showed similar pattern of spatial structure from these regions (Mutegi *et al*., 2011). In the present study, there was no significant pattern of gene flow by isolation-by-distance presumably because of other barriers including the steep mountainous topography in some of the collections sites.

The pattern of genetic diversity and population structure among the regions of origin of populations was less than among the climatic zones, which may indicate that the latter is more informative for population genetic studies in wild sorghum in Ethiopia. The negative correlation between gene diversity, allelic richness, and average observed heterozygosity with altitude may indicate high genetic diversity in the more lowlands where the wild sorghums adapt better than in the higher altitudes. The negative correlation between gene diversity and altitude was in agreement with Baek *et al*. (2003) who after examining the microsatellite diversity of 306 wild barley (*Hordeum spontaneum*) individuals from 16 populations with altitude ranges of -120-1650masl in Jordan, found higher gene diversity (H_e) in the lower altitudes than in the higher altitudes.

5.1.2.3. Evidence for crop-wild gene flow

Different methods confirmed the existence of historical gene flow with different magnitude between the crop and the wild sorghum. This evidence was found by pooling together the SSR data of individuals from both taxa and subjecting to F_{ST}, Structure, and

PCoA analyses. The divergence between cultivated and wild sorghum was moderate (F_{ST}=0.105) compared to the inter population divergence within each taxon, indicating the possible gene flow and introgression. Unlike a recent study by Mutegi *et al*. (2011), the results of this study were not consistent with the hypothesis that repeated gene flow has blurred distinctions between wild and cultivated sorghum. As noted above, four Ethiopian wild populations that clustered with an accession of ssp. *drummondii* were very distinct from the landraces in this study. Likewise, some populations were also differentiated from the crop. However, similar to Mutegi *et al*. (2011) STRUCTURE analyses at K=2 (Fig. 4.9) did not differentiate between wild and cultivated plants. Moreover, a neighbor joining cluster dendrogram based on pair-wise simple matching distance among the 540 individuals of cultivated and wild populations showed some level of admixture (Appendix Fig. 2). These patterns could reflect extensive hybridization and/or common ancestry of the cultivars from their wild progenitors in Ethiopia.

Differences between the findings of this study and those of Mutegi *et al*. (2011) could be attributed to several factors, including sampling designs, differences in SSR markers, and variation in farmers' practices of weeding out the wild or hybrid sorghum before it is able to set seed. In Hararghe, for example, it was observed that farmers frequently remove "off-types" and weeds during the growing season to feed livestock, making it difficult for researchers to find surviving wild plants with flowering panicles. Tesfaye Tesso *et al*. (2008) also reported that wild sorghum was less common in Hararghe than in other regions of Ethiopia where it occurred intermixed with the crop.

In agreement with the present results, Casa *et al*. (2005) surveyed many African accessions and reported that the degree of differentiation between wild *vs*. cultivated was moderate (F_{ST} = 0.13). In this study crop-wild differentiation was also moderate (F_{ST}=0.105). Similarly, Cui *et al*. (1995) reported moderate differentiation among wild and cultivated sorghum. However, these findings are not inconsistent with the expectation that a limited amount of gene flow has occurred, as many previous investigators have speculated (e.g., Harlan and De Wet, 1972; Doggett, 1988; Aldrich *et al*., 1992). Analyses of progeny segregation, allozymes, and RFLPs revealed crop-specific alleles in wild *S. bicolor* when it co-occurred with the crop in Africa, suggesting that intraspecific hybridization and introgression are common (Aldrich and Doebley, 1992). On the other hand, the result was against the result of Mutegi *et al*. (2012) who analyzed wild and cultivated sorghum at a local scale in Kenya and found great differentiation between the two congeners (F_{ST}=0.27).

A broad survey involving more cultivar samples and crop-specific molecular markers might provide stronger evidence for crop-wild gene flow in Ethiopia, but data from the current study do not suggest that such gene flow has been extensive. This is rather surprising because cultivated and wild sorghum frequently co-occur and have overlapping flowering times (e.g., Tesfaye Tesso *et al*., 2008; Mutegi *et al*., 2010). In addition, F_1 crop-wild sorghum hybrids do not show fitness penalties (Muraya *et al*., 2012) and may even exhibit heterosis. As noted earlier, outcrossing rates of wild sorghum ranged from 0.33-0.65 in a survey of six populations in Ethiopia and from 0.09-0.70 in 12 populations in Kenya (Muraya *et al*., 2011b). Populations with relatively high

outcrossing rates should be more likely to hybridize with the crop, depending on the likelihood of pollen flow and the likelihood of being weeded out by farmers.

5.1.2.4. Detection of recent bottlenecks and estimation of effective population size

Detection of bottlenecks and estimation of effective population size are of paramount importance for the conservation of plant genetic resources, especially if they are indigenous to a specific country. Because the immediate progenitors of cultivated sorghum are indigenous to Ethiopia, studying their population status is timely for providing advice to conservationists. In this study, both the heterozygosity excess (mainly the TPM, which is known to be more reliable) and mode shift methods did not detect any recent bottlenecks in the studied populations, which were previously assumed to have threatened (Amsalu ayana *et al*., 2001). The effective size of the two populations also showed no risk of population reduction in wild sorghum in these regions of Ethiopia. The methods followed do not require prior information on historical population size (Cornuet and Luikart, 1996). Therefore, the allopatric wild sorghum populations of Ethiopia can survive by occupying patches in road sides, fences, abandoned farm lands, forests, etc. and this shows that their wild characteristics of adaptation have been adequate to survive from extinction at least until recently. However, this doesn't guarantee the survival of these species for the future and *ex situ* conservation is suggested to maintain the population diversity.

5.1.2.5. Implications for germplasm conservation and use

We found that considerable genetic diversity remains in extant wild populations, some of which (the isolated types) were thought to be threatened by human overpopulation and expansion of modern agriculture (Amsalu Ayana *et al*., 2001). The existence of such diversity in the weedy types will continue as it is the result of adaptation due to co-evolution with the crop. The weedy sorghums are intimately associated with the crop by a system of mimicry (Barrett, 1983), whereby they resemble the crop before flowering and, as a result of misidentification, escape eradication. In addition, most of them shatter their seeds before harvest of the crop. However, the result of the present study has shown that genetic swamping from the crop has not occurred, although some crop-to-wild gene flow is likely. This implies that transgenes could also spread to wild populations. The significantly higher level of private allelic richness observed in wild sorghum populations relative to cultivars may be beneficial to sorghum breeding as it may be linked to unique traits. Further conservation effort may be necessary to safeguard this unique diversity.

As with many crop species, rigorous comparisons between *in situ* and *ex situ* populations are generally needed to fully assess genetic resources for germplasm preservation (Heywood and Iriondo, 2003; Maxted, 2003). Sampling genetic diversity and storing or propagating the collected material outside the natural environments using *ex situ* conservation methods is useful to safeguard endangered species (Heywood and Iriondo, 2003). For instance, after reintroduction of 20 cultivated sorghum accessions managed in gene bank for an average of 17.5 years to their original places in Wello, Ethiopia, Beyene Seboka and van Hintum (2006) found that 5 of them (those collected 22 years ago) were

lost or not remembered by the farmers. *Ex situ* conservation saved them from extinction. Similar study was conducted in North Eastern Ethiopia (Wello) to investigate the possible genetic erosion that occurred in cultivated sorghum for 30 years by making new collections in 2003 and comparing with those similar types collected in 1973 and conserved in gene bank (Godwin *et al*., 2008). Based on the preliminary results, these authors reported the loss of some important landraces. Moreover, Rice (2004) found that the diversity of older Jala maize gene bank populations (*ex situ* conserved) was less and more differentiated from recent Jala (*in situ* conserved) due to gene bank collection and regeneration methods. However, it has limitations such as poor genetic or demographic management almost inevitably resulting in genetic erosion and spontaneous hybridization (Volis *et al*., 2009) in addition to the costly operation (Maxted, 2003). On the other hand, *in situ* conservation allows evolutionary change to continue in the component species and populations (Heywood and Iriondo, 2003) and provides better representation of genetic diversity (Volis *et al*., 2009). In this study, unlike the Ethiopian collections, which were sampled *in situ*, the reduced gene diversity and allelic richness in the ICRISAT accessions was perhaps because they have been under stringent selection pressure during maintenance. Sampling accessions from seed banks that have been increased to maintain the germplasm was suggested to be one reason for low variation in sorghum accessions by Morden *et al*. (1989). Maintaining the high outcrossing rate and other life history traits may help sustain the high diversity (Barnaud *et al*., 2008).

5.2. Estimation of mating system parameters in *in situ* wild/weedy sorghum [*Sorghum bicolor (L.) Moench]* using SSR markers

5.2.1. Variation in outcrossing rate among wild sorghum populations from different geographical regions

The high variation in outcrossing rate (0.33-0.65) among wild sorghum populations collected from different geographical regions in Ethiopia can be attributed to variation in geoclimatic factors. These factors that might have affected outcrossing rate may include temperature, humidity (and moisture), and wind (speed and direction). It is known that the sorghum growing environments of Ethiopia are diverse in the geoclimatic variables. The wild sorghum populations included in the present experiment were collected at elevations ranging from 631-1709m. This wide range of elevations is accompanied by modification of climatic and edaphic factors that might have favored or disfavored outcrossing rate and inbreeding of populations from different geographical regions. However, the correlation between altitude and outcrossing rate *per se* was not significant in this study. Moreover, against the results of Dje *et al*. (2004) all of the measured phenotypic traits (plant height, panicle length, width, and compactness) had no significant correlation with multilocus outcrossing rate. Probably the proximity to cultivated sorghum might have affected the mating systems because all of the wild sorghum populations were collected from cultivated sorghum fields as weeds except Ghibe populations, which were found in isolated habitat. This proximity to the cultivated sorghum might be associated with the panicle size and shape of the cultivars, and the relative density of the crop and the cultivar in the mixture as it may also affect pollen competition. In sorghum, self pollen (within the wild or within the cultivated) is reported

to have more seed siring efficiency than pollen from other taxa (Muraya *et al*., 2011b). Differences in farmers' cultivation practices were also suggested to affect mating systems (Barnaud *et al*., 2008).

The relatively lower outcrossing rate in Kobo families might be due to the reduced number of SSR loci used for outcrossing rate is a function of the number of loci used. On the other hand, the highest outcrossing rate in Abuare families was probably the result of better synchronization of flowering and better anther extrusion ability of the intermixed improved variety developed through selection. Although crop specific SSR loci were not used, the highest outcrossing rate that was observed in this weedy sorghum associated with the improved variety is an indication of the high potential of gene flow from improved varieties to the wild sorghums, which in turn is an indication of the high risk of potential transgene flow from the future GM sorghum to the non-GM cultivated sorghum and wild and weedy sorghum.

The presence of significant correlation between gene diversity and outcrossing rate indicates the role of mating systems in shaping the genetic diversity of wild sorghum. The negative correlation between the correlation of outcrossed paternity (r_P) and plant density (r=-0.605, p=0.203) though not significant, shows that when maternal plant density decreases, the probability that any two progenies from the same mother share the same father (being full sibs) increases due to low concentration of pollen admixture received via wind as it was observed in Bombacaceous trees (Murawski *et al*., 1990). The high and variable outcrossing rate in the present experiment was in agreement with

the results of Muraya *et al*. (2011b). It also verifies that outcrossing rate was more than it was reported in cultivated sorghum in some of the earlier studies (Ellstrand and Foster, 1983; Dje *et al*., 2004). This can be expected because one of the evolutionary consequences of domestication in self compatible crop plants is the evolution towards inbreeding (Jain, 1976).

5.2.2. Biparental inbreeding

Biparental inbreeding has significant consequences for the evolution of mating systems, yet is difficult to estimate in natural populations (Griffin and Christopher, 2003) and its frequency is unknown (Kelly and Willis, 2002). However, recently, some procedures have been proposed to estimate biparental inbreeding based on genetic data (e.g., Brown, 1979; Ritland, 1889; Ritland, 2002; Kelly and Willis, 2002). In this experiment, three of the five methods applied revealed biparental inbreeding in the Ethiopian wild sorghum populations.

The Ghibe population was found isolated from the crop sorghum on either sides of the main road from Addis Ababa to Jimma crossing a natural forest. It occurs at high density (135 plants per quadrant, $16m^2$ area) and in a continuous hedge for more than 3Km distance. However, no correlation was observed between multilocus outcrossing rate and plant density (r=0.04) in the current set of populations, which was in agreement with the results of Neel *et al*. (2001) in *Eriogonum ovalifolium var. vineum*. Moreover, the plants in this population had highly laxed panicles, which contribute to high outcrossing (eg. Dje *et al*., 2004). The high biparental inbreeding in families derived from this population

was perhaps due to the recurrent shattering and emergence of seeds in similar localities (limited seed dispersal). However, the biparental inbreeding observed in the rest of the families, whose parents were found intermixed as weeds with the cultivated sorghum or found in close proximity might be due to farmer practices and/or in some cases the reduced synchronization of flowering between the crop and the wild sorghum.

Each weedy sorghum plant is surrounded by a mixture of cultivated sorghum plants and another weedy plants, most of which found in close proximity could be closely related. Hence it may receive pollen by wind from these related plants and/or possibly from unrelated individuals located a bit farther. If there is reduced synchronization of flowering between the cultivated and the weeds, and due to better pollen competition, the greatest chance of mating will be between related individuals, which are in close proximity. Although Tesfaye Tesso *et al*. (2008) reported the overlap in flowering between the cultivated and the wild sorghum in most regions where they made the survey, the wild sorghum seed for this experiment was collected from those main plants which matured earlier than the cultivated sorghum in all except at *Abuare* where the cultivars were also early maturing. The interview with farmers also supports the view that wild sorghum flowers earlier than the cultivated. This was further confirmed by a seed maintenance activity held at Melkassa in the main season, 2010 where two samples from each of the 30 wild sorghum populations were grown with the 8 cultivated sorghum landrace populations collected from the same place where the 8 wild sorghum populations were collected. All of the wild sorghum collections flowered earlier than the cultivated ones. However, this doesn't rule out the overlap in flowering between the

cultivated sorghum and the late coming tillers of the wild sorghum as wild sorghum has extended flowering time. In plants like wild and weedy sorghums whose seeds are dispersed over short distances from the maternal plants, biparental inbreeding is not unexpected.

5.2.3. Implications for crop-wild gene flow and conservation

The high outcrossing observed in wild sorghum in Ethiopia can have implications for crop-wild gene flow as there are no barriers for hybridization to occur between the two congeners. Moreover, it has implications on the spread of crop genes into the wild sorghum pool. However, it is not only outcrossing potential that is required for gene flow but also other factors like the presence of sexually compatible recipient plants in the vicinity (e.g., Wilkinson *et al*., 2003; Jenczewski *et al*., 2003) and further the ability of introgression in the recipient populations (Slatkin, 1987). Recent case studies also showed that crop-wild gene flow is likely in sorghum (Tesfaye Tesso *et al*., 2008; Muraya *et al*., 2010, 2011b; Mutegi *et al*., 2011).

The wild/weedy sorghums have their own mechanisms to survive in the environment such as mimicry inside crop fields in which case the farmers are not able to identify the weedy types. During flowering cross pollination takes place to both directions and can bring about differential consequences, but then contributes to the perpetuation of the wild forms. On the other hand, sometimes the farmers in Ethiopia may deliberately leave the wild sorghums in and around their field to undergo interpollination with the cultivated sorghum which could lead to beneficial characteristics being attained by the cultivated

landraces (Awgchew Teshome *et al*., 1999). Therefore, this kind of variation in farmers' practices in different parts of the country might have contributed to the variation in outcrossing rate of the wild sorghum populations collected from diverse environments. Therefore, maintaining life history traits such as high outcrossing is important for preserving genetic diversity in natural populations.

5.3. Morphology and fitness of wild × crop F1 sorghum hybrids

Because genetic transformation of sorghum has been underway (e.g., Jeoung *et al*., 2004; Zhao *et al*., 2000; Zhu *et al*., 1998; Liang and Gao, 2001; Zhao *et al*., 2003) and is expected to be released in the near future (Zhao, 2007; Muraya *et al*., 2011b), it is timely to undertake risk assessment studies case by case in each country (Ejeta and Grenier, 2005). Furthermore, although wild/weedy types of sorghum populations could be sources of genes for crop improvement, they have direct negative consequences on sorghum production. Hence, in order to provide better weed control methods, it is important to perform studies to learn about crop traits affecting competitiveness and reproduction that will make their way into weed populations.

Studying the fitness/breeding value of crop wild/weed hybrids is crucial to predict the future fate of the rare hybrids on the environment as it determines the effectiveness of the initial gene flow, especially in cases where alien transgenes might be included in the hybrids and cause biosafety concerns (Snow *et al*., 2003; Groot *et al*., 2003; Song *et al*., 2004). Ejeta and Grenier (2005) stressed the critical need of determining the fitness of

crop-wild sorghum hybrids in order to evaluate the potential genetic introgression into weedy types.

The present experiment was dealt with comparing crop-wild sorghum hybrids and their parents for various fitness related traits: juvenile survival (germination and emergence), adult growth (leaf, tillering, height, flowering and panicle traits) and reproductive/fertility (female fecundity). These traits are believed to have significant contribution for competition and reproductive potentials in crop-wild hybrids and thus are useful to estimate fitness. Earlier during crossing there was no any form of barrier for hybridization observed between the wild and the cultivated sorghum. Usually, the low fertility of the early generation hybrids has been considered to be the major obstackle to introgression (e.g., Weis, 2005). F1 fitnesses from crosses between wild and cultivated rice were generally high, but hybrids from certain crosses showed reduced fertility (Ellstrand *et al*., 1999). However, that is commonly the case in interspecific crosses. Arriola and Ellstrand (1996, 1997) detected interspecific hybridization between the tetraploid *S. halepense* and the diploid sorghum crop (*S. bicolor*) using allozyme progeny analysis and found that the resulting hybrids showed no fitness differences from the weed under field conditions. However, intraspecific gene flow occurs more readily than interspecific gene flow, since the cross compatibility is high and improved offspring fitness (heterosis) is common (Jørgensen and Wilkinson, 2005). To this end, cultivated sorghum and its wild relatives belong to the same species (*S. bicolor* (L.) Moench) and can freely hybridize (de Wet, 1978). Due to this reason, such penalties in fertility were not observed in the present study involving wild-crop sorghum hybrids. Moreover, all of

the hybrids did not show any fitness costs with regard to the measured phenological and morphological traits. In some cases even the hybrids exhibited significant mid-parent heterosis for some of these traits.

In agreement with the present study, some empirical fitness assessment studies have been carried out in crop-wild hybrids in different species with and without transgenes and the results were quite variable. Some field experiments showed that different fitness parameters exhibited by the non-modified hybrids could be as high as and even higher than those of the weedy parent (Klinger and Ellstrand, 1994; Arriola and Ellstrand, 1997; Hauser *et al*., 1998; Hauser *et al*., 2003). Di *et al*. (2009) also reported intermediate composite fitness and lowest reproductive fitness of crop wild hybrids formed between transgenic oilseed rape (*Brassica napus* L.) and wild brown mustard [*B. juncea* (L.) Czern et Coss.]. Arnold and Hodges (1995, cited in Weis (2005) reviewed 44 studies on species that hybridize naturally, and compared fitness components of hybrids to parental taxa and found that in over half the cases, hybrids were as fit as or fitter than the parental taxa.

Survivorship and reproduction of early-generation hybrids in a field environment determines gene flow between crops and weedy relatives (Raybould and Cooper, 2005). The better fitness of the hybrids at the juvenile and fertility (fecundity) stages than their wild parents in the present study may indicate that crop genes in the hybrids are working in two different directions to recombine the parental characters. First, unlike the wild parents dormancy could be removed in the hybrids and second, unlike the crop parents,

fecundity could be enhanced. This recombination can be beneficial to the survival and spread of transgenes.

Some levels of mid-parent heterosis were observed at all stages in different wild × crop cross combinations, but were not evident on emergence and tillering. Heterosis was negative for days to flowering, which is considered to be advantageous in crop improvement. This can also help the hybrids as an adaptation to shatter their seeds earlier during the season. With the exception of a single hybrid, all have overlapping flowering time with their wild parents. This was in agreement with the results of Snow *et al*. (1998) and will help the hybrids to backcross with their wild parents and may ensure introgression. Further, while plants in the cultivated sorghum plots flowered in synchrony, the tillers in wild plants continued flowering for more than 15 days. This behavior of asynchronous and wider range of flowering in the wild sorghum is in agreement with Jenczewski *et al*. (2003) and is what opens overlapping widows with the cultivated sorghum for possible natural outcrossing.

Heterosis was also negative in leaf traits (number, length and width) because the domesticated parents have values approaching to the positive maximum for these traits. But it was positive for germination, plant height, head length and width, and number of seeds per plant (fecundity). The number of seeds per panicle was lower in the hybrids and wild parents than the cultivars. But superiority of most hybrids in fecundity over their cultivar parents was attributed to the large number of productive tillers. Fecundity itself was considered to substitute fitness, and all other things being equal, plants producing

more seeds will make a greater proportional increase to the next generation (Hails and Raymond, 2004). Similar heterosis pattern in crop-wild crosses was also observed in sorghum (Muraya *et al*., 2012) and in other crop species (Hauser *et al*., 2004; Guadagnuolo *et al*., 2006). The observed heterosis in the present crosses can be degraded in subsequent generations. Nevertheless, based on two races of ssp. verticilliflorum (i.e., verticilliflorum and arundinaceum), Muraya *et al*. (2012) studied the fitness of wild × cultivated sorghum in F1, F2, and F3 generations in Kenya and Mali and found no consistent differences among the three generations. However, they didn't observe fitness differences in seed yield (fecundity) of all the generations.

Such substantial degree of fitness in the wild × crop sorghum F1 hybrids in the present experiment without the addition of alien transgene(s) into the cultivated parents may have implications on introgression and spread of crop genes to the wild relatives. However, the addition of transgenes may alter the fitness and competitiveness cost to either direction depending on the behavior of the inserted gene and the environment. Usually, for weed control purposes, transgenes that have fitness costs to the weedy species or tightly linked to a disadvantageous gene are more preferable for transfer (Andow and Zwahlen, 2006). So far it has been suggested that hybridization between wild and cultivated parents may result in novel combination of genes, which may enhance fitness relative to wild plants (Hauser *et al*., 1998; Kiær *et al*., 2007) and the results of the present study also support this idea with evidences based on morphological observations.

5.4. Ecotypic variation for seed dormancy, longevity and germination requirements in wild/weedy sorghum

5.4.1. Effect of burial time and depth for seed germination of wild sorghum ecotypes and the crop

The result of the present study on the fate of wild sorghum seed after 18 and 24 months of burial was intermediate between the results of Burnside *et al*. (1996) and Teo-Sherrell and Mortensen (2000). Burnside *et al*. (1996) found little viability in weedy sorghum seed after 5 years of burial, but they also observed 1% germinated after 17 years of burial in eastern Nebraska. However, Teo-Sherrell and Mortensen (2000) reported more than 80% mortality in weedy sorghum seed within three months after burial (between November and March) in Nebraska because of winter cold. However, such kind of winter environment of the temperate regions is not available in most parts of Sub-Sahara Africa including Ethiopia. The observation that an increase in depth of burial increases longevity was in agreement with the previous reports in crimson and red clovers (Rampton and Ching, 1970), in jointed goat grass (Donald and Zimdahl, 1987), in wild oat (Miller and Nalewaja (1990), and in *Amaranthus retroflexus* seeds (Omani *et al*., 1999), to mention a few.

The fact that differences for viability associated with dormancy were highly significant only for genotype factor may indicate that dormancy is mainly a characteristic of the plant (genotype). It could be largely influenced by the physiological and genetic makeup within the seed. In general, crop seeds lost their viability within the first 6 months regardless of the depth of burial. The difference in longevity between the crop and the

wild seeds could be attributed to two reasons. First, it may be related to differences in their initial dormancy. The wild sorghum seeds from Ghibe and from Pawe had 35% and 45% dormancy, respectively before burial whereas the cultivated sorghum had near complete germination (94%) (data not shown). This much dormancy in the wild was observed after five months of storage in the laboratory at room temperature (after ripening). It is well known that there is a relationship between seed dormancy and plant longevity (e.g., Jurado and Flores, 2005). The second reason could be that wild sorghum seeds are covered by lemma, palea, and glumes, which may impose mechanical dormancy for considerable period of time. This behavior of wild sorghum seeds was also observed in earlier studies and degluming was found to improve germination (Sastry *et al*., 2006). Similar observation was also reported in wild rice (Gu *et al*., 2003).

Soil temperature at Melkassa seemed to have less effect to bring about variation in dormancy among the different exumation periods as the trend was more or less similar during the duration of burial (Appendix Fig. 3). However, the rainfall distribution might have affected. Seed bags in set-2 of this experiment were buried in April, 2010. After the burial, there was high rainfall; even the highest rainfall during the duration of burial was in July-August of the same year (Appendix Fig. 4). That could be the main reason for the depletion of crop seeds within the first 6 months of burial.

5.4.2. Response of wild sorghum ecotypes for varying moisture and temperature regimes

In this study germination of wild sorghum seed decreased with a decrease in osmotic potential in all genotypes and at all temperature levels. Smith and Hoveland (1986) also found 44% reduction in germination of cultivated sorghum seed when osmotic potential decreased from 0 to -1.0 MPa. In agreement with this study, Smith *et al*. (1989) observed no germination at 15 °C and at an OP of -1.2MPa in cultivated sorghum and in pearl millet. Similarly, Lippai *et al*. (1996) studied the response of Horehound seed at different temperature and osmotic potentials and found increased germination for each degree increase in temperature but decreased with decreases in water potential and ceased at -1.5 MPa. The superior germination observed in W-4 population over the other two populations at the highest temperature (30°C) with mild moisture stress (Ψs=-0.6 MPa) might be associated with the fact that this wild population is adapted to the dry lowlands characterized by high temperature and drought stress conditions. The observed better germination in some of the test ecotypes at -0.3MPa of osmotic potential than the control may indicate that some level of moisture stress is useful to enhance wild sorghum seed germination. Perhaps this is among the reasons why wild sorghum is predominantly distributed in the lower altitude, moisture stressed areas of Ethiopia than in the highlands.

6. CONCLUSION AND RECOMMENDATIONS

The exact number of wild sorghum accessions held *ex situ* in the institute of biodiversity (IBC) as of 2011 is not known. The total germplasm of sorghum (12000-15000) has been preserved under *S. bicolor* with no classification into cultivated and wild (Mrs.

Weynshet, personal communication). Moreover, the herbarium specimens kept at the national herbarium are not adequate to represent the wild/weedy populations at the country scale. Therefore, much needs to be done to collect, characterize and conserve these genetic resources, which can be useful for future crop improvement to introgress genes to meet the emerging needs. To this end, the contribution of the present study is great. The collected samples used in this study are preserved at the national sorghum research coordinating center (Melkassa Agricultural Research Center) and seeds will be delivered to IBC with their complete passport data.

The present study was aimed at studying the relationship between cultivated and wild sorghum (including the extent of gene flow). Hence, survey was made in the major sorghum growing regions only. However, wild sorghums also grow near irrigation ditches, river banks and areas where sorghum is not grown. Therefore, future studies should focus on the geographical distribution of the various sub-species and races of wild sorghums in areas of the country where this study didn't cover.

The study has shown that wild and weedy sorghum of Ethiopia possesses higher genetic diversity than the cultivated, which may indicate that the latter have undergone bottlenecks during domestication. This high diversity and the presence of high private allelic richness may be useful for future crop improvement. Moreover, there was no signature of recent bottlenecks observed in the studied isolated populations. Therefore, efforts should be made to safeguard such high diversity. Wild sorghum populations are more differentiated within each region than between regions, which may indicate the

importance of long distance seed movement among regions. The divergence between the crop and the wild populations was smaller than the divergence within each taxon, which may indicate gene flow between the two congeners. The pattern of variation among climatic regions is generally more informative than among regions of collection for the study of wild sorghum genetic structure in Ethiopia. More diversity could be observed in extant populations than in gene bank conserved accessions, which may show the advantage of *in situ* conservation over *ex situ* conservation. Further fine tuned study should be carried out involving more cultivar samples and crop-specific molecular markers to predict the consequences of gene flow between the two congeners.

Wild sorghum in different geographical regions of Ethiopia exhibited variation in outrcossing rate, but in general the rates were higher than has been reported in cultivated sorghum in other countries. Several factors could affect the rate of outcrossing, but altitude, plant height and panicle traits were not found to have significant correlation. Differences in farmers' practices of selective removal vs. maintaining them in different parts of the country might have contributed to the variation. Because of the observed high outcrossing, crop genes (including transgenes) transferred to the wild pool can easily spread. Depending on the fitness advantage conferred by such genes/traits, this may cause genetic erosion in the wild genetic resources, which calls for effective conservation measures to be put in place.

The study has also shown that wild × crop F_1 hybrids could reproduce and had better fitness than their wild parents. In some cases heterosis was also observed for some traits.

The direction of the fitness could be determined by the behavior of the inserted gene, but in the mean time the results of this study showed that hybrids containing crop genes (including transgenes) could have fitness advantage and could be weedier than their weedy parents. For a more robust conclusion regarding the effect of the environment on the fitness of wild × crop hybrids, however, multilocation trials should be carried out.

If a genetically engineered gene escapes to the neighboring non-genetically engineered crop or wild relatives belonging to the same species via pollen, the seed that will be formed and shattered into the soil bank will have two outcomes. A proportion of that seed may germinate and grow into a new generation plant whereas the remaining will remain dormant and persist in the soil for some period of time. Therefore, studying the longevity of wild/weed seeds in the soil bank is an important component of the biosafety risk assessment procedures. In the present study that consisted of five populations (ecotypes) representing different sorghum growing regions, on average 1.24% of the seeds remained viable at 10cm depth after 24 months of burial. This proportion cannot be considered as negligible so far as biosafety is concerned. Therefore, dispersed seeds carrying transgenes could persist in the soil for considerable period of time, which may call for devising effective risk management strategies if transgenic sorghum is to be deployed in Ethiopia. However, from weed control point of view, because most of the seeds die within the first year, it is possible to manage the weedy sorghum through crop rotation. Moreover, wild sorghum seeds can germinate at reduced osmotic potential (high moisture stress) if it is associated with high temperature.

7. REFERENCES

Abebe Menkir, Goldsbrough, P. and Gebisa Ejeta. (1997). RAPD Based Assessment of Genetic Diversity in Cultivated Races of Sorghum. *Crop Science* **37**:564-569.

Abu-Assar, A.H., Uptmoor, R., Abdelmula, A.A., Salih, M., Ordon, F., Friedt, W. (2005) Genetic in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). *Crop Science* **45**:1636–1644.

Ahrens, J.D. (1996). Benishangul-Gumuz, Situation Report, 10/96. Emergencies Unit for Ethiopia (UNDP-EUE) Development Programme. (http://www.africa.upenn.edu/EUE/eue_ahrens1096.html).

Agrama, H.A., Tuinstra, M.R. (2003). Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *African Journal of Biotechnology* **2**: 334-340.

Aldrich, P.R., and Doebley, J. (1992) Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild *Sorghum bicolor*. *Theoretical and Applied Genetics* **85**: 293–302.

Aldrich, P.R., Doebley, J., Schertz, K.F., Stec, A. (1992). Patterns of allozyme variation in cultivated and wild *Sorghum bicolor*. *Theoretical and Applied Genetics* **85**:451–60.

Alemu Tirfessa (2009). Analysis of genetic diversity among Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] germplasm collections. MSc Thesis, Hawassa University, Awassa, Ethiopia.

Ali, M.L., Rajewski, J.F., Baenziger, P.S., Gill, K.S., Eskridge, K.M., Dweikat, I. (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Molecular Breeding* **21**: 497–509.

Allard, R. W. (1965). Genetic systems associated with colonizing ability in predominantly self-pollinated species. In: Baker, H.G. and Stebbins, G.L. (eds.), *The Genetics of Colonizing Species*. Proceedings of the First International Union of Biological Sciences Symposium on General Biology. Academic Press: New York, pp. 49.75.

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

Amsalu Ayana, Bryngelsson, T., Endashaw Bekele (2000). Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). *Genetic Resources and Crop Evolution* **47**:471–481

Amsalu Ayana, Endashaw Bekele and Bryngelsson, T. (2001). 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.

Anderson, E. (1949). Introgressive hybridization. New York: John Wiley & Sons.

Andow, D.A. and Zwahlen, C. (2006). Assessing environmental risks of transgenic plants. *Ecology Letters* **9**: 196–214.

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

Arriola, P.E. and Ellstrand, N.C. (1997). Fitness of interspecific hybrids in the genus *Sorghum*: persistence of crop genes in wild populations. *Ecological Applications* **7**:512–518.

Arriola, P.E., Ellstrand, N.C. (1996). Crop-to-weed gene flow in the genus *Sorghum* (Poaceae): spontaneous interspecific hybridization between johnsongrass, *Sorghum halepense*, and crop sorghum, *S. bicolor*. *American Journal of Botany* **83**: 1153–1160.

Arriola, P.E. and Ellstrand, N.C. (2002). Gene flow and hybrid fitness in the *Sorghum bicolor –Sorghum halepense* complex. In: Proceedings of the Ecological and Agronomic Consequences of Gene Flow from Transgenic Crops to Wild Relatives Meeting, March 5 - 6, The Ohio State University, Columbus, Ohio.

Asfaw Adugna, Snow, A.A. and Sweeney, P. (2011). Optimization of a high throughput, cost effective and all-stage DNA extraction protocol for sorghum. *Journal of Agricultural Science and Technology, USA* **5**: 243-250.

Awgchew Teshome, 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 SouthWelo, Ethiopia. *Euphytica* **97**: 255–263

Awgchew Teshome, Fahrig, L., Torrance, J.K., Lambert, J.D., Arnason, Z.J. and Baum, B.R. (1999). Maintenance of sorghum (*Sorghum bicolor*, Poaceae) landrace diversity by farmers' selection in Ethiopia. *Economic Botany* **53**: 79-88.

Baek, H.J., Beharav, A. and Nevo E. (2003). Ecological-genomic diversity of microsatellites in wild barley, Hordeum spontaneum, populations in Jordan. *Theoretical and Applied Genetics* **106**:397–410

Baker, H.G. (1974). The evolution of weeds. *Annual Review of Ecology and Systematics* **5**: 1–24.

Balloux, F. and Goudet, J. (2002). Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology* **11**: 771–783.

Balloux, F. and Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**: 155-165.

Barnaud, A., Deu, M., Garine, E., Chantereau, J., Bolteu, J., Ouin, E., KoÏda, E.O, Mckey, D. and Joly, H. I. (2009). A weed – crop complex in sorghum: the dynamics of Genetic diversity in a traditional farming system. *American Journal of Botany* **96**: 1869–1879.

Barnaud, A., Deu, M., Garine, E., McKey, D. and Joly, H.I. (2007). Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. *Theoretical and Applied Genetics* **114**: 237–248.

Barnaud, A., Trigueros, G., McKey, D. and Joly, H. (2008). High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity* **101**: 445–452.

Barrett, S.C.H. (1983). Crop mimicry in weeds. *Economic Botany* **37**: 255-82

Bartsch, D., Lehnen, M., Clegg, J., Pohl-Orf, M., Schuphan, I. and Ellstrand, N.C. (1999). Impact of gene flow from cultivated beet on genetic diversity of wild sea beet populations. *Molecular Ecology* **8**:1733–1741.

Battraw, M. and Hall, T.C. (1991). Stable transformation of *Sorghum bicolor* protoplasts with chimeric *neomycin phosphotransferase* II and *b-glucuronidase* genes. *Theoretical and Applied Genetics* **82**: 161-168.

Bennett, S.J. (1999). Ecotypic variation between and within two populations of Trifolium tomentosum (woolly clover) from Syria and Western Australia: its success as a colonising species. *Australian Journal of Agricultural Research* **50**: 1443-1450.

Benvenuti, S. (2003). Soil Texture Involvement in Germination and Emergence of Buried Weed Seeds. *Agronomy Journal* **95**:191–198.

Beyene Seboka and van Hintum, T. (2006). The dynamics of on-farm management of sorghum in Ethiopia: Implication for the conservation and improvement of plant genetic resources. *Genetic Resources and Crop Evolution* **53**: 1385-1403.

Bhattramakki, D., Dong J., Chhabra A.K. and Hart, G.E. (2000). An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* **43**: 988–1002.

Bioversity International. (2010). Key access and utilization descriptors for sorghum genetic resources. (www.bioversityinternational.org).

Bisrat Ghebru, Schmidt, R.J. and Bennetzen, J.L. (2000). Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. *Theoretical and Applied Genetics* **105**: 229-236.

Bohonak, A.J. (2002). IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* **93**: 153-154.

Borrell, A.K., Hammer, G.L. and Douglas, A.C.L. (2000). Does maintaining green leaf area in sorghum improve yield under drought? I. Leaf growth and senescence. *Crop Science* **40**: 11026-103.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**: 314-331.

Brhane Gebrekidan. (1982). Utilization of germplasm in sorghum improvement. In: House, L.R., Mughogho, L.K. and Peaxoxk J.M. (eds.), Sorghum in the eighties, Vol. 1. Procedings of the international symposium on Sorghum, 2-7 Nov 1981, ICRISAT, Patancheru, A.P., India., pp. 335-345.

Brown, A.H.D. (1979). Enzyme polymorphisms in plant populations. *Theoretical Population Biology* **15**: 1-42.

Brown, A.H.D. and Allard, R.W. (1970). The estimation of Wright's fixation index from genotypic frequencies. *Genetica* **41**: 399-406.

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]. *Theorical and Applied Genetics* **93**: 190-198.

Buckley, Y.M., Ree M., Paynter Q. and Lonsdale M. (2004). Modelling integrated weed management of an invasive shrub in tropical Australia. *Journal of Applied Ecology* **41**: 547–560.

Burke, J.M., Carney, S.E. and Arnold, M.L. (1998). Hybrid fitness in the Louisiana irises: analysis of parental and F1 performance. *Evolution* **52**: 37-43.

Burnside, O.C., Wilson, R.G., Weisberg, S. and Hubbard, K.G. (1996). Seed Longevity of 41 Weed Species Buried 17 Years in Eastern and Western Nebraska. *Weed Science* **44**: 74-86.

Casa, A.M., Mitchell, S.E., Hamblin, M.T., Sun, H., Bowers, J.E., Paterson, A.H., Aquadro, C.F. and Kresovich, S. (2005). Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. *Theoretical and Applied Genetics* **111**: 23–30.

Casas, A.M., Kononowicz, A.K., Zehr, U.B., Tomes, D.T., Axtell, J.D., Butter, L.G., Bressan, R.A. and Hasegawa, P.M. (1993). Transgenic sorghum plants via microprojectile bombardment. *Proceedings of the National Academy of Sciences, USA* **90**: 11212–11216.

Celarier, R. P. (1959). Cytotaxonomy of the Andropogoneae. III. Subtribe Sorgheae genus *Sorghum. Cytologia* **23**: 395-418.

Central Statistical Agency (CSA) (2003-2011). Agricultural sample survey, report on area and production of major crops (private peasant holdings, Meher season), September-December, Addis Ababa, Ethiopia.

Chandler, S. and Dunwell, J.M. (2008). Gene flow, risk assessment and the environmental release of transgenic plants. *Critical Reviews in Plant Sciences* **27**: 25–49.

Clayton, W.D. and Renvoize, S.A. (1986). Genera Graminum grasses of the world. Kew Bulletin Addition Series XIII. London: Royal Botanic Gardens, Kew, pp. 338–345.

Clegg, M.T. (1980). Measuring plant mating systems. *Bioscience* **30**: 814-818.

Cohen, D. (1966). Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology* **12**:119–129.

Conn, J.S., Beattie, K.L. and Blanchard, A. (2006). Seed viability and dormancy of 17 weed species after 19.7 years of burial in Alaska. *Weed Science* **54**: 464–470.

Cornuet, J.M. and Luikart, G. (1996). Description and power analysis for two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.

Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M. and Wayne, K. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**: 290–295.

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

Dahlberg, J.A. (2000). Classification and characterization of sorghum. In: Smith, C.W. and Frederiksen, R.A. (eds.) *Sorghum: origin, history, technology, and production*. John Wiley & Sons, Inc., New York. pp. 99-130.

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

Darwin, C. (1859). On the origin of species by means of natural selection. Murray, London.

De Wet, J.M.J. (1978). Systematics and evolution of sorghum sect. Sorghum (Gramineae). *American Journal of Botany* **65**: 477-484.

De Wet, J.M.J. and Harlan, J.R.. (1971). The origin and domestication of *Sorghum bicolor*. *Economic Botany* **25**: 128-135.

De Wet, J.M.J. and Harlan, J.R. (1975). Weeds and domesticates: Evolution in the man-made habitat. *Economic Botany* **29**: 99–107.

De Wet, J.M.J., Harlan, J.R. and Price, E. G. (1970) Origin of variability in the spontanea complex of *Sorghum bicolor*. *American Journal of Botany* **57**: 704-707.

De Wet, J.M.J. and Huckabay, J.P. (1967). The origin of *Sorghum bicolor*. II. Distribution and Domestication. *Evolution* **21**: 787-802.

Dean, R.E., Dahlberg, J.A., Hopkins, M.S., Mitchell, S.E. and Kresovich, S. (1999). Genetic redundancy and diversity among 'orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. *Crop Science* **39**: 1215-1222.

Defelice, M.S. (2006). Shattercane, *Sorghum bicolor* (l.) Moench ssp. *Drummondii* (Nees ex steud.) de Wet ex Davidse—black sheep of the family. *Weed Technology* **20**:1076-1083.

Deu, M., Sagnard, F., Chantereau, J., Calatayud, C., He´rault, D., Mariac, C., Pham, J.L., Vigouroux, Y., Kapran, I., Traore, P.S., Mamadou, A., Gerard, B., Ndjeunga, J. and Bezancon, G. (2008). Niger-wide assessment of *in situ* sorghum genetic

diversity with microsatellite markers. *Theoretical and Applied Genetics* **116**: 903–913.

Dewoody, J., Nason, J.D. and Hipkins, V.D. (2006). Mitigating scoring errors in microsatellite data from wild Populations. *Molecular Ecology Notes* **6**: 951–957.

Di, K., Stewart, C.N., Wei, W., Shen, B., Tanga, Z. and Ma, K. (2009). Fitness and maternal effects in hybrids formed between transgenic oilseed rape (*Brassica napus* L.) and wild brown mustard [*B. juncea* (L.) Czern et Coss.] in the field. *Pest Management Science* **65**: 753–760.

Dillon, S.L., Lawrence, P.K. and Henry, R.J. (2005) The new use of *Sorghum bicolor*-derived SSR markers to evaluate genetic diversity in 17 Australian Sorghum species. *Plant Genetic Resources* **3**: 19-28.

Dillon, S.L., Shapter, F.M., Henry, R.J., Cordeiro, G., Izquierdo, L. and Lee, S. (2007). Domestication to crop mprovement: genetic resources for sorghum and saccharum (Andropogoneae). *Annals of Botany* **100**: 975 – 989.

Djé , Y., Heuertz, M., Ater, M., Lefebyre, C. and Vekemans, X. (2004). *In situ* estimation of outcrossing rate in sorghum landraces using microsatellite markers. *Euphytica* **138**: 205-212.

Djé , Y., Heuertz, M., Lefebvre, C. and Vekemans, X. (2000). Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. *Theoretical and Applied Genetics* **100**: 918-925.

Dje`, Y., Forcioli, D., Ater, M., Lefe`bvre, C. and Vekemans, X. (1999). Assessing population genetic structure of sorghum landraces from North-western Morocco

using allozyme and microsatellite markers. *Theoretical and Applied Genetics* **99**: 157-163.

Doggett, H. (1965). The development of cultivated sorghums. In: Hutchinson, J. (ed.), *Crop plant evolution*. Cambridge Univ. Press, London.

Doggett, H. (1976). Sorghum. In: Simmonds, N.W. (ed.), *Crop Evolution*, Longman, pp. 112–117.

Doggett, H. (1988). Sorghum. 2nd edn. Longman, London.

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

Donald, W.W and Zimdahl, R.L. (1987). Persistence, germinability and distribution of jointed goat grass (Aegilops cylindrical) seed in soil. *Weed Science* **35**: 149-154.

Donohue, K. (2002). Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology* **83**: 1006–1016.

Dunbabin, M.T. and Cocks, P.S. (1999). Ecotypic variation for seed dormancy contributes to the success of capeweed (*Arctotheca calendula*) in Western Australia. *Australian Journal of Agricultural Research* **50**:1451-1458.

Earl, D.A. and von Holdt, B.M. (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. DOI: 10.1007/s12686-011-9548-7, Version: v0.6.8 Oct 2011

Edwards, K.J., Barker, J.H.A., Daly, A., Jones, C. and Karp, A. (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques* **20**: 758-759.

Ellstrand, N.C and Foster, F. (1983). Impact of population structure on the apparent outcrossing rate of grain sorghum (*Sorghum bicolor*). *Theoretical and Applied Genetics* **66**: 323–327.

Ellstrand, N.C, Prentice, H.C. and Hancock, J.F. (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecological and Systematics* **30**: 539-563.

Ellstrand, N.C. (2003). Dangerous Liaisons? When Cultivated Plants Mate with Their Wild Relatives. Johns Hopkins University Press, Baltimore, MD.

Ellstrand, N.C. and Elam, D.R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual Reviews of Ecology and Systematics* **24**: 217–242.

Enjalbert, J. and David, L. (2000). Inferring Recent Outcrossing Rates Using Multilocus Individual Heterozygosity: Application to Evolving Wheat Populations. *Genetics* **156**: 1973–1982.

Escalante, A., Coello, M.G., Eguiarte, L.E. and Piñero, D. (1994). Genetic structure and mating systems in wild and cultivated populations of Phaseolus coccineus and P. vulgaris (Fabaceae). *American Journal of Botany* **81**: 1096-1103.

Estoup, A., Jarne, P. and Cornuet, J. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**:1591–1604.

Evanno, S., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

Excoffier, L., Laval, G. and Schneider, S. (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47-50.

Falush, D., Stephens, M. and Pritchard, J.K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Felber, F., Kozlowski, G., Arrigo, N., and Guadagnuolo, R. (2007). Genetic and ecological consequences of transgene flow to the wild flora. In: Fiechter, A. and Sautter, C. (Eds.) *Green Gene Technology. Advances in Biochemical Engineering/Biotechnology*, Vol. 107. Springer, Berlin and Heidelberg, Germany pp. 173–205.

Folkertsma, R.T., Frederick, H., Rattunde, W., Chandra, S., Raju, G.S. and Hash, C.T. (2005). The pattern of genetic diversity of guinea-race *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. *Theoretical and Applied Genetics* **111**: 399-409.

Fraser, D.F. (2000). Species at the edge: the case for listing of “peripheral” species. In: Darling, L.M. (ed.). At Risk - Proceedings of a Conference on the Biology and Management of Species and Habitats at Risk. Vol. 1, 15–19 February 1999, University College of the Cariboo, Kamloops, Ministry of Environment, Lands, and Parks, Victoria, B.C. pp. 49–53.

Frew Mekbib. (2007). Infra-specific folk taxonomy in sorghum (*Sorghum bicolor* (L.) Moench) in Ethiopia: folk nomenclature, classification, and criteria. *Journal of Ethnobiology and Ethnomedicine* **3**:38. doi:10.1186/1746-4269-3-38

Fyfe, J.L. and Bailey, N.T.J. (1951). Plant breeding studies in leguminous forage crops. 1. Natural cross breeding in winter beans. *Journal of Agricultural Science* **41**: 371–378.

Garber, E.D. (1950). Cytotaxonomic studies in the genus *Sorghum. University of California Publication in Botany* **23**: 283-361.

Gealy, D.R., Mitten, D.H. and Rutger, J.N. (2003). Gene flow between red rice (*Oryza sativa*) and herbicide-resistant rice (*O. sativa*): implications for weed management. *Weed Technology* **17**: 627–645.

Gebisa Ejeta and Grenier, C. (2005). Sorghum and its weedy hybrids. In: J. Gressel (ed), Crop ferality and volunteerism , 123-135, Taylor and Francis, Boca Raton, Florida, USA.

Gepts, P. and Papa, R. (2003). Possible effects of (trans)gene flow from crops on the genetic diversity from landraces and wild relatives. *Environmental Biosafety Research* **2**: 89–103.

Gillespie, J.H. (1998). Population Genetics: A Concise Guide. 1998 The Johns Hopkins University Press, Baltimore and London.

Godwin, I.D., Jordan, D.R. and Hailemichael Shewayrga (2008). Genetic erosion and changes in distribution of sorghum [*Sorghum bicolor* L.(Moench)] landraces in north-eastern Ethiopia. *Plant Genetic Resources* **6** 1: 1-10.

Goudet, J. (2002). FSTAT, version 2.932. Institute of Ecology, Biology Building, UNIL, Switzerland, Lausanne

Grenier, C., Deu, M., Kresovich, S., Bramel-Cox, P.J. and Hamon, P. (2000). Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum

collection using random vs nonrandom sampling procedures. B. Using molecular markers. *Theoretical and Applied Genetics* **101**: 197–202.

Griffin, C.A.M. and Eckert, C.G. (2003). Experimental analysis of biparental inbreeding in a self-fertilizing plant. *Evolution* **57**: 1513-1519.

Groot, M.H.M., van de, Wiel C.C.M., van Tienderen, P.H. and den Nijs, H.C.M. (2003). Hybridization and introgression between crops and wild relatives: Current knowledge and research priorities in lieu of impending introductions of GM crops. COGEM research 2003-02, University of Amsterdam & Plant Research International, Amsterdam & Wageningen.

Gu, X-Y., Chen, Z-X. and Foley, M.E. (2003). Inheritance of Seed Dormancy in Weedy Rice. *Crop Science* **43**:835–843.

Guadagnuolo, R., Clegg, J. and Ellstrand, N.C. (2006). Relative fitness of transgenic vs. non-transgenic maize × teosinte hybrids: a field evaluation. *Ecological Applications* **16**: 1967-1974.

Hagio, T., Blowers, A.D. and Earle, E.D. (1991). Stable transformation of sorghum cell cultures after bombardment with DNA-coated microprojectiles. *Plant Cell Reproduction* **10**: 260–264.

Hails, R.S. (2000). Genetically modified plants – the debate continues. *Trends in Ecology and Evolution* **15**:14–18

Hails, R.S. and Raymond, B. (2004). Insect-resistant transgenic plants and their environmental impact. In: den Nijs, H.C.M., Bartsch, D. and Sweet, J. (eds.) *Introgression from genetically modified plants into wild relatives*. CABI Publishing, Wallingford, Oxfordshire, UK. pp.279-295.

Halliburton, R. (2004). Introduction to population genetics. Pearson Education, Inc., Upper Saddle River.

Hamrick, J.L.J. and Godt, W. (1990). Allozyme diversity in plants. In: H. D. Brown, M. T. Clegg, A. L. Kahler and B. S. Weir (eds), Plant population genetics, breeding, and genetic resources. Sinauer Associates Inc. Publishers, Sunderland. pp. 43-63

Harding, J. and Tucker, C.L. (1964). Quantitative studies on mating systems. I. Evidence for the nonrandomness of outcrossing in *Phaseolus lunatus*. *Heredity* **19**: 369-381.

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

Harlan, J. and De Wet, J.M.J. (1972). A simplified classifi-cation of cultivated sorghum. *Crop Science* **12**: 172-176.

Harlan, J.R. and De Wet, J.M.J. (1974). Sympatric evolution in sorghum. *Genetics* **78**: 473-474.

Harlan, J.R., De Wet, J.M.J. and Price, E.G. (1973). Comparative evolution of cereals. *Evolution* **27**: 311-325.

Hartl, D. (1980). Principles of Population Genetics. Sinauer Assoc., Sunderland, Massachu-setts. 488 pp.

Hartl, D.L. and Clark, A.G. (1989). *Principles of Population Genetics*, 2nd edn. Sinauer & Associates, Sunderland, Massachusetts.

Hauser, T.P, Bjorn, G.K., Magnussen, L. and Shim, S.I. (2004). Hybrids between Cultivated and Wild Carrots: a Life History. In: den Nijs, H.C.M., Bartsch, D. and

Sweet, J. (eds.) *Introgression from genetically modified plants into wild relatives.* CABI Publishing, Wallingford, Oxfordshire, UK. pp. 41-51.

Hauser, T.P., Damgaard, C. and Jørgensen, R.B. (2003). Frequency dependent fitness of hybrids between oilseed rape (*Brassica napus*) and weedy *B. rapa* (*Brassicaceae*). *American Journal of Botany* **90**: 571-578.

Hauser, T.P., Shaw, R.G. and Østergård, H. (1998). Fitness of F-1 hybrids between weedy *Brassica rapa* and oilseed rape (*B. napus*). *Heredity* **81**: 429-435.

Haygood, R., Ives, R.A. and Andow, D.A. (2003). Consequences of recurrent gene flow from crops to wild relatives. *Proceedings of the Royal Society of London - Series B: Biological Sciences* **270**: 1879-1886.

Hermanutz, L.A. and Weaver, S.E. (1996). Agroecotypes or phenotypic plasticity? Comparison of agrestal and ruderal populations of the weed *Solanum ptycanthum*. *Oecologia* **105:** 271–280.

Heywood, V.H. and Iriondo, J.M. (2003). Plant conservation: old problems, new perspectives. *Biological Conservation* **113**: 321–335.

Holm, L.G., Plucknett, D.L., Pancho, J.V. and Herberger, J.P. (1977). The World's Worst Weeds: Distribution and Biology. Honolulu: Univ. Press Hawaii.

House, L.R. (1985). A Guide to Sorghum Breeding. 2nd Ed. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Andhra Pradesh 502 324, India

Huang, J.K., Rozelle, S., Pray, C. and Wang, Q.F. (2002). Plant Biotechnology in China. *Science* **295**: 674-677.

Hutchenson, K. (1970). A test for comparing diversities based on the Shannon formula. *Journal of Theoretical Biology* **29**: 151–154.

IPGRI and ICRISAT (1993) Descriptors for sorghum [*Sorghum bicolor* (L.) Moench]. International Board of Plant Genetic Resources, Rome, Italy. ICRISAT, Patancheru, India.

Idury, R.M. and Cardon, L.R. (1997). A simple method for automated allele binning in microsatellite markers. *Genome Research* **11**: 1104–1109.

Jain, S. K. (1976). The evolution of inbreeding in plants. *Annual Reviews in Ecology and Systematics* **7**: 469-495.

James, C. (2011). Global Status of Commercialized Biotech/GM Crops: International Service for the Acquisition of Agri-biotech Applications (ISAAA). Brief 43 (www.isaaa.org/resources/publications/briefs/43/).

Jarvis, D.I. and Hodgkin, T. (1999). Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Molecular Ecology* **8**: 5159-5173.

Jenczewski, E., Ronfort, J. and Chèvre, A.M.. (2003). Crop-to-wild gene flow, introgression and possible fitness effects of transgenes. *Environmental Biosafety Research* **2**: 9–24.

Jeoung, J.M., Muthukrishnan, S., Trick, H.N. and Liang, G.H. (1999). Agrobacterium-mediated transformation of sorghum. ASA abst. pp. 157, Salt Lake City, UT, USA.

Jeoung, J.M., Krishnaveni, S., Jayaraj, J., Haying, Y., Trick, H.N., Muthukrishna and S., Liang, G.H. (2004). Agrobacterium-mediated transformation of grain sorghum.

In: Seetharama, N. and Godwin, I. (eds.) *Sorghum tissue culture and transformation*. Science Publishers, pp 57–64.

Jin, L. and Chakraborty, R. (1994). Estimation of genetic distance and coefficient of gene diversity from single-probe multilocus DNA fingerprinting data. *Molecular Biology and Evolution* **11**: 120-127.

Jordan, D.R., Tao, Y.Z., Godwin, I.D., Henzel, R.G., Cooper, M. and McIntyre, C.L. (1998). Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. *Euphytica* **102**: 1-7.

Jørgensen, R.B. and Wilkinson, M.J. (2005). Rare hybrids and methods for their detection. In: Poppy, G.M. and Wilkinson, M.J. (eds.) *Gene flow from GM plants*. Blackwell Publishing Ltd., Oxford, UK.

Jurado, E. and Flores, J. (2005). Is seed dormancy under environmental control or bound to plant traits? *Journal of Vegetation Science* **16**: 559-564.

Kalinowski, S. (2005). HP-RARE 10: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology* **5**: 187–189.

Kanemasu, E.T., Bark, D.L. and Chin Choy, E. (1975). Effect of soil temperature on sorghum emergence. *Plant and Soil* **43**: 411-417.

Kelly, J.K. and Willis, J.H. (2002). A manipulative experiment to estimate biparental inbreeding in monkey flowers. *International Journal of Plant Science* **163**: 575-579.

Kiær, L.P., Philipp, M., Jorgensen, R.B. and Hauser, T.P. (2007). Genealogy, morphology and fitness of spontaneous hybrids between wild and cultivated chicory (*Cichorium intybus*). *Heredity* **99**: 112–120.

Klinger, T. and Ellstrand, N.C. (1994). Engineered genes in wild populations: fitness of weed–crop hybrids of *Raphanus sativus*. *Ecological Society of America* **4**: 117–120.

Ladizinsky, G. (1985). Founder effect in crop-plant evolution. *Economic Botany* **39**: 191-199.

Laidlaw, H.K.C. and Godwin, I.D. (2008). Sorghum. In: Kole, C. and Hall, T.C. (eds.). *Compendium of Transgenic Crop Plants: Transgenic Cereals and Forage Grasses*. Blackwell Publishing, Oxford, UK. pp 157-176.

Langridge, P. and Chalmers, K. (2004). The principle: Identification and application of molecular markers. In: Lörz, H. and Wenzel, G. (eds.) *Molecular marker systems in plant breeding and crop improvement*. Biotechnology in Agriculture and Forestry, Vol 55, Springer-Verlag Berlin Heidelberg, pp 3-22

Lazarides, M., Hacker, J.B. and Andrew, M.H. (1991). Taxonomy, cytology and ecology of indigenous Australian sorghums (Sorghum Moench: Andropogoneae: Poaceae). *Australian Systematic Botany* **4**: 591–635.

Li, B. and Foley, M.E. (1997). Genetic and molecular control of seed dormancy. *Trends in Plant Science* **2**: 384-389.

Li, M., Yuyama N., Luo L., Hirata M. and Cai H. (2009). In silico mapping of 1758 new SSR markers developed from public genomic sequences for sorghum. *Molecular Breeding* **24**: 41–47.

Liang, H. and Gao, Z.S. (2001). Phylogenetic analysis and transformation of sorghum [*Sorghum bicolor* (L.) Moench]. *Recent Research in Developmental and Plant Biology* **1**: 17-33.

Linder, C.R. and Schmitt, J. (1994). Assessing the risks of transgene escape through time and crop-wild hybrid persistence. *Molecular Ecology* **3**: 23–30.

Lippai, A., Smith, P.A., Price, T.V., Weiss, J. and Lloyd, C.J. (1996). Effects of Temperature and Water Potential on Germination of Horehound (Marrubium vulgare) Seeds from Two Australian Localities. *Weed Science* **44:** 91-99.

Liu, K. and Muse, S.V. (2005). PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics* **21**: 2128-2129.

Long, T., Karle, R., Igoe, F. and Smith, M. (2008). Collection, Transport, Purification and Storage of PCR-ready Plant DNA. Available online at: www.whatman.com., accessed, 26 January 2009.

Loveless, M.D. and Hamrick, J.L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Reviews on Ecology and Systematics* **15**: 65-95.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209-220.

Manzelli, M., Pileri, L., Lacerenza, N., Benedettelli, S. and Vecchio, V. (2007). Genetic diversity assessment in Somali sorghum (*Sorghum bicolor* (L.) Moench) accessions using microsatellite markers. *Biodiversity and Conservation* **16**:1715–1730.

Marks, M.K. and Nwachuku, A.C. (1986). Seed-bank characteristics in a group of tropical weeds. *Weed Research* **26**: 151-157.

Martinez-Ghersa, M.A., Satorre, E.H. and Ghersa, C.M. (1997). Effect of soil water content and temperature on dormancy breaking and germination of three weeds. *Weed Research* **45**: 791-797.

Maxted, N. (2003). Conserving the genetic resources of crop wild relatives in European Protected Areas. *Biological Conservation* **113**: 411–417.

Mellon, M. and Rissler, J. (2004). Gone to Seed: Transgenic Contaminants in the Traditional Seed Supply.UCS Publications, Cambridge, MA, pp. 40–42.

Menz, M.A., Klein, R.R., Unruh, N.C., Rooney, W.L., Klein, P.E. and Mullet, J.E. (2004). Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. *Crop Science* **44**: 1236–1244.

Messeguer, J., Marfà, V., Català, M.M., Guiderdoni, E. and Melé, E. (2004). A field study of pollen-mediated gene flow from Mediterranean GM rice to conventional rice and the red rice weed. *Molecular Breeding* **13:** 103–112.

Michel, B.E. and Kaufmann, M.R. (1973). The Osmotic Potential of Polyethylene Glycol 6000. *Plant Physiology* **51**: 914-916.

Miller, S.D. and Nalewaja, J.D. (1990). Influence of burial depth on wild oats (*Avena fatua*) seed longevity. *Weed Technology* **4**: 514–517.

MOARD (Ministry of Agriculture and Rural Development). (1998). Agro-ecological Zones of Ethiopia. Addis Ababa, Ethiopia.

MOARD (Ministry of Agriculture and Rural Development). (2005). Major Agro-ecological Zones of Ethiopia. Forestry, Landuse and Soil Conservation Department, Addis Ababa, Ethiopia.

Monaghan, N. (1979) The biology of Johnson grass *(Sorghum halepense)*. *Weed Research* **19**: 261-267.

Morden, C.W., Doebley, J.F. and Schertz, K.F. (1989). Allozyme variation in old world races of *Sorghum bicolor* (Poaceae). *American Journal of Botany* **76**: 247-255.

Mulatu Geleta. (2001). Ethnobotanical study of edible oil crops as a companion of Sorghum bicolor L. Moench and biochemical genetic analysis of in situ and ex situ conserved Guizotia abyssinica (L.F.) Cass. Germplasm from north Shoa and south Wello. MSc Thesis, Addis Ababa University.

Murawski, D.A., Hamrick, J.L., Hubbell, S.P. and Foster, R.B. (1990). Mating systems of two Bombacaceous trees of a neotropical moist forest. *Oecologia* **82**: 501–506.

Muraya, M.M., Hartwig, H.H., Mutegi, E., Kanyenji, B.M., Sagnard, F., de Villiers, S., Kiambi, D., Heiko, K. and Parzies, H.K. (2010). Geographical patterns of phenotypic diversity and structure of Kenyan wild sorghum populations (Sorghum spp.) as an aid to germplasm collection and conservation strategy. *Plant Genetic Resources* **8**:217–224.

Muraya, M.M., Geiger, H.H., Sagnard, F., Toure, L., Traore, P.C.S., Togola, S., de Villiers, S. and Parzies, H.K. (2012). Adaptive values of wild × cultivated sorghum (*Sorghum bicolor* (L.) Moench) hybrids in generations F1, F2, and F3. *Genetic Resources and Crop Evolution* **59**: 83–93.

Muraya, M.M., Hartwig, H.G., de Villiers, S., Sagnard, F., Kanyenji, B.M., Kiambi, D. and Parzies, H.K. (2011a). Investigation of pollen competition between wild and cultivated sorghums (*Sorghum bicolor* (L.) Moench) using simple sequence repeats markers. *Euphytica* **178**: 393–401.

Muraya, M.M., Mutegi, E., Geiger, H.H., de Villiers, S.M., Sagnard, F., Kanyenji, B.M., Kiambi, D. and Parzies, H.K. (2011b). Wild sorghum from different eco-geographic regions of Kenya display a mixed mating system. *Theoretical and Applied Genetics*. **122**: 1631-1639.

Mutegi, E., Sagnard, F., Muraya, M., Kanyenji, B., Rono, B., Mwongera, C., Marangu, C., Kamau, J., Parzies, H., de Villiers, S., Semagn, K., Traore´, P.S. and Labuschagne, M. (2010). Ecogeographical distribution of wild weedy and cultivated Sorghum bicolor (L.) Moench in Kenya: implications for conservation and crop-to-wild gene flow. *Genetic Resources and Crop Evolution* **57**:243–253.

Mutegi, E., Sagnard, F., Labuschagne, M., Herselman, L., Semagn, K., Deu, M., de Villiers, S., Kanyenji, B.M., Mwongera, C.N., Traore, P.C.S. and Kiambi, D. (2012). Local scale patterns of gene flow and genetic diversity in a crop–wild–weedy complex of sorghum (*Sorghum bicolor* (L.) Moench) under traditional agricultural field conditions in Kenya. *Conservation and Genetics.* DOI 10.1007/s10592-012-0353-y.

Mutegi, E., Sagnard, F., Muraya, M., Kanyenji, B., Rono, B., Mwongera, C., Marangu, C., Kamau, J., Parzies, H., de Villiers, S., Semagn, K., Traore´, P.S. and Labuschagne. M (2010). Ecogeographical distribution of wild weedy and cultivated *Sorghum bicolor* (L.) Moench in Kenya: implications for conservation and crop-to-wild gene flow. *Genetic Resources and Crop Evolution* **57**: 243–253.

Mutegi, E., Sagnard, F., Semagn, K., Deu, M., Muraya, M., Kanyenji, B., de Villiers, S., Kiambi, D., Herselman, L. and Labuschagne, M. (2011). Genetic structure and relationships within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. *Theoretical and Applied Genetics* **122**: 989–1004.

Narwal, S., Sindel, B.M. and Jessop R.S. (2008). Dormancy and longevity of annual ryegrass (Lolium rigidum) as affected by soil type, depth, rainfall, and duration of burial. *Plant and Soil* **310**: 225–234.

Neel, M.C., Ross-Ibarra, J. and Ellstrand, N.C. (2001). Implications of mating patterns for conservation of the endangered plant Eriogonum ovalifolium var. vineum (Polygonaceae). *American Journal of Botany* **88**: 1214-1222.

Nei, M. (1987). Molecular evolutionary genetics. Columbia University Press, New York.

Nemera Geleta, Labuschagne, M.T. and Viljoen, C.D. (2006). Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. *Biodiversity and Conservation* **15**: 3251-3265.

Nielson, R.L., McPherson, M.A., O'Donovan, J.T., Harker, K.N., Yang, R., and Hall, L.M.. (2009). Seed-Mediated Gene Flow in Wheat: Seed Bank Longevity in Western Canada. *Weed Science* **57**:124-132.

Nkongolo, K.K. and Nsapato, L. (2003). Genetic diversity in *Sorghum bicolor* (L.) Moench accessions from different ecogeographical regions in Malawi assessed with RAPDs. *Genetic Resources and Crop Evolution* **50**: 149–156.

Noldin, J.A., Chandler, J.M. and McCauley, G.N. (2006). Seed longevity of red rice ecotypes buried in soil. *Planta Daninha, Viçosa-MG* **24**: 611-620.

Ohta, T. and Kimura, M. (1973). The model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genetic Research Cambridge* **22**: 201–204.

O'Kennedy, M.M., Grootboom, A. and Shewry, P.R. (2006). Harnessing sorghum and millet biotechnology for food and health. *Journal of Cereal Science* **44**: 224–235.

Omami E., Haigh, A.M., Medd, R.W. and Nicol, H.I. (1999). Changes in germinability, dormancy and viability of *Amaranthus retroflexus* as affected by depth and duration of burial. *Weed Research* **39**: 345-354.

Page, R.D.M. (2001). TreeView 1.6.6. (http://taxonomy.zoology.gla.ac.uk/rod/rod.html)

Paterson, A.H., Bowers, J.E. and Feltus, F.A. (2008). Genomics of sorghum, a semi-arid cereal and emerging model for tropical grass genomics. In: Paul H. Moore Ray Ming (Eds) Genomics of tropical crop plants. Springer Science+Business Media, LLC, 233 Spring Street, New York, NY10013, USA. pp. 469-482.

Pedersen, J.F., Toy, J.J. and Johnson, B. (1998). Natural outcrossing of *Sorghum* and Sudangrass in the Central Great Plains. *Crop Science* **38**: 937 – 939.

Polowick, P.L., Quandt, J. and Mahon, J.D. (2000). The ability of pea transformation technology to transfer genes into peas adapted to western Canadian growing conditions. *Plant Science* **153**: 161–170.

Prasanth, V., Chandra, S., Jayashree, B. and Hoisington, D. (2006). AlleloBin-a program for allele binning of microsatellite markers based on the algorithm of Idury and Cardon (1997). International Crops Research Institute for the Semi, Arid Tropics (ICRISAT), Patancheru, AP, India

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Ramsay, G. (2005). Pollen dispersal vectored by wind or insects. In: Poppy, G.M., Wilkinson, M.J. (eds.) *Gene Flow from GM Plants*. Blackwell Publishing Ltd, Oxford.

Rampton, H.H. and Ching, T.M. (1970). Persistence of Crop Seeds in Soil. *Agronomy Journal* **62**: 272-277.

Raybould, A. and Cooper, I. (2005). Tiered tests to assess the environmental risk of fitness changes in hybrids between transgenic crops and wild relatives: the example of virus resistant *Brassica napus. Environmental Biosafety Research* **4**: 127–140.

Reed, J.D., Ramundo, B.A., Claflin, L.E. and Tuinstra, M.R. (2002). Analysis of resistance to ergot in sorghum and potential alternate hosts. *Crop Science* **42**: 1135–1138.

Rice, E.B. (2004). Conservation and change: a comparison of *in situ* and *ex situ* conservation of Jala maize germplasm in Mexico. PhD Dissertation, Faculty of the Graduate School of Cornell University .

Rich, P.J., Grenier, C. and Ejeta, G. (2004). Striga resistance in wild relative of sorghum. *Crop Science* **44**:

Ritland, K. (1983). Estimation of mating systems. In: Tanksley, S.D. and Orton T.J. (eds.), *Isozymes in Plant Genetics and Breeding*- Part A. Elsevier Science Publishers, Amsterdam. pp. 289-302.

Ritland, K. (1989). Correlated matings in the partial selfer, *Mimulus guttatus*. *Evolution* **43**: 848-859.

Ritland, K. (2002). Extensions of models for the estimation of mating systems using n independent loci. *Heredity* **88**: 221–228.

Ritland, K. (2008). Multilocus mating system program MLTR Version 3.2. http://genetics.forestry.ubc.ca/ritland/programs, accessed on 20 July 2009

Ritland, K. and Jain, S.K. (1981). A model for the estimation of outcrossing rate and gene frequencies using n independent loci. *Heredity* **47**: 35-52.

Ritter, K.B., McIntyre, C.L., Godwin, I.D., Jordan, D.R. and Chapman, S.C. (2007). An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* ssp. Bicolor (L.) Moench, using AFLP markers. *Euphytica* **157**: 161-176.

Rittner, D. and McCabe, T.L. (2004). Encyclopedia of biology. Facts on File Inc., New York.

Rousset, F. and Raymond, M. (1997). Statistical analyses of population genetic data: new tools, old concepts. *Trends in Ecology and Evolution* **12**: 313–317.

Sagnard, F., Deu, M., Dembe´le´, D., Leblois, R., Toure´, L., Diakite´, M., Calatayud, C., Vaksmann, M., Bouchet, S., Malle´, Y., Togola, S. and Traore, P.C.S. (2011). Genetic diversity, structure, gene flow and evolutionary relationships within the Sorghum bicolor wild–weedy–crop complex in a western African region. *Theoretical and Applied Genetics* **123**:1231-1246.

Sastry, DVSSR, Upadhyaya, H.D. and Gowda, C.L.L. (2006). Degluming, Storage and Testing Periods Influence Seed Germination of Wild Sorghum Germplasm. *Journal of SAT Agricultural Research* **2** (ejournal.icrisat.org)

Schlotterer, C. (2004). The evolution of molecular markers-just a matter of fashion? *Genetics* **5**:63-69.

Schmidt, M.R. and Bothma, G.C. (2006). Risk assessment for transgenic sorghum in Africa: Crop-to-crop gene flow in *Sorghum bicolor* (L.) Moench. *Crop Science* **46**: 790-798.

Shaw, D.V. and Allard, R.W. (1982). Estimation of Outcrossing Rates in Douglas-fir Using Isozyme Markers. *Theoretical and Applied Genetics* **62**: 113-120.

Shaw, D.V., Kahler, A.L. and Allard, R.W. (1981). A multilocus estimator of mating system parameters in plant populations. *Proceedings of the National Academy of Sciences USA* **78**: 1298-1302.

Shehzad, T., Okuizumi, H., Kawase, M. and Okuno, K. (2009). Development of SSR-based sorghum (*Sorghum bicolor* (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. *Genetic Resources and Crop Evolution* **56**: 809–827.

Simmonds, N.W. (1995). Evolution of crop plants. Longman Scientific and Technical, New York.

Simpson, R.L., Leck, M.A. and Parker, V.T. (1989). Seed banks: general concepts and methodological issues. Pages 3–8 in M. A. Leck, V. T. Parker, and R. L. Simpson, eds. Ecology of Soil Seed Banks. San Diego, CA: Academic.

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

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**: 393-430.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science* **236**: 787-792.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457-462.

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

Slatkin, M. and Maddison, W.P. (1989). A Cladistic Measure of Gene Flow Inferred from the Phylogenies of Alleles. *Genetics* **123**: 603-613.

Small, E. (1984) Hybridization in the domesticated-weed-wild complex. In: Grant WF (ed) Plant biosystematics, Academic Press, Toronto, pp. 195-210

Smith, B.D. (2001). Documenting plant domestication: the consilience of biological and archaeological approaches. *Proceedings of the National Academy of Sciences, USA* **98**: 1324–1326.

Smith, C.W. and Frederiksen, R.A. (2000). Sorghum: origin, history, technology, and production. John Wiley and Sons, Hoboken.

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

Smith, R.L. and Hoveland, C.S. (1986). Effects of temperature and water stress on seed germination and seeding depth on emergence of pearl millet and sorghum. *Agronomy abstracts*, ASA, Madison, WI. p.94.

Smith, R.L., Hoveland, C.S. and Hanna, W.W. (1989). Water stress and temperature in relation to seed germination of pearl millet and sorghum. *Agronomy Journal* **81**: 303-305.

Snow, A.A., Andow, D.A., Gepts, P., Hallerman, E.M., Power, A., Tiedje, J.M. and Wolfenbarger, L.L. (2005). Genetically engineered organisms and the

environment: Current status and recommendations. *Ecological Applications* **15**: 377-404.

Snow, A.A., Moran-Palma, P., Rieseberg, L.H., Wszelaki, A. and Seiler, G.J. (1998). Fecundity, phenology, and seed dormancy of F1 wild-crop hybrids in sunflower (*Helianthus annuus*, *Asteraceae*). *American Journal of Botany* **85**: 794–801.

Snow, A.A., Pilson, D., Rieseberg, L.H., Paulsen, M., Pleskac, N., Reagon, M.R., Wolf, D.E. and Selbo, S.M. (2003). A *Bt* transgene reduces herbivory and enhances fecundity in wild sunflowers. *Ecological Applications* **13**: 279-286.

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

Song, Z.P., Lu, B.R., Wang, B. and Chen, J.K. (2004). Fitness Estimation through Performance Comparison of F1 Hybrids with their Parental Species *Oryza rufupogon* and *O. sativa*. *Annals of Botany* **93**: 311-316.

Stemler, A.B.L., Harlan, J.R. and De Wet, J.M.J. (1977). The sorghums of Ethiopia. *Economic Botany* **31**: 446 – 460.

Szpiech, Z.A., Jakobsson, M. and Rosenberg, N.A. (2008). ADZE: A rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **24**: 2498–2504.

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

Taramino, G., Tarchini, R., Ferrario, S., Lee, M. and Pe, M.E. (1997). Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theoretical and Applied Genetics* **95**: 66–72.

Tazi, L. and Zeiger, E. (2003). Plant physiology. 3rd ed. Sunderland, Sinauer Associates.

Templeton, A.R. (2006). Population genetics and microevolutionary theory. John Wiley & Sons, Inc., Hoboken, New Jersey. pp. 716.

Teo-Sherrell, C.P.A. and Mortensen, D.A. (2000). Fates of Buried Sorghum bicolor ssp. drummondii Seed. *Weed Science* **48**: 549-554.

Teo-Sherrell, C.P.A., Mortensen D.A. and Keaton M.E. (1996). Fates of weed seeds in soil: a seeded core method of study. *Journal of Applied Ecology* **33**: 1107-1113.

Tesfaye Tesso, Kapran, I., Grenier, C., Snow, A.A., Sweeney, P., Pedersen, J., Marx, D., Bothma, G. and Gebisa Ejeta (2008). The potential for crop-to-wild gene flow in sorghum in Ethiopia and Niger: a geographic survey. *Crop Science* **48**: 1425–1431.

Uptmoor, R., Wenzel, W., Friedt, W., Donaldson, G., Ayisi, K. and Ordon, F. (2003). Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from Southern Africa by RAPDs, AFLPs and SSRs. *Theoretical and Applied Genetics* **106**: 1316–1325.

USDA ARS. (2007). National Genetic Resources Program. Germplasm Resources Information Network (GRIN). National GermplasmResources Laboratory. Beltsville, MD, available at http://www.ars-grin.gov/cgi-bin/npgs/html/tax_search.pl (accessed, 23 March 2011).

Van Beuningen, L.T. and Busch, R.H. (1997). Genetic diversity among North American spring wheat cultivars: III. Cluster analysis based on quantitative morphological traits. *Crop Science* **37**: 981–988.

Van Deynze, A. (2011). Gene Flow in Agriculture. In: The Science of Gene flow in Agriculture and its Role in Co-existence, Proceedings, Washington, DC, September 7-8, pp. 1-2.

Van Mourik, T.A., Stomph, T.J. and Murdoch, A.J. (2005). Why high seed densities within buried mesh bags may overestimate depletion rates of soil seed banks. *Journal of Applied Ecology* **42**: 299–305.

Van, T.K. (2005). Seed Longevity of *Melaleuca quinquenervia*: A Burial Experiment in South Florida. *Journal of Aquatic Plant Management* **43**: 39-42.

Vaughton, G.V. and Ramsey, M. (2001). Variation in summer dormancy in the lilioid geophyte *Burchardia umbeliata* (Colchicaeae). *American Journal of Botany* **88**:1223–1229.

Vigouroux, Y., Jaqueth, J.S., Matsuoka, Y., Smith, O.S., Beavis, W.D., Stephen, J., Smith, C. and Doebley, J. (2002). Rate and Pattern of Mutation at Microsatellite Loci in Maize. *Molecular Biology and Evolution* **19**:1251–1260.

Volis, S., Blecher, M. and Sapir, Y. (2009). Complex *ex situ* - *in situ* approach for conservation of endangered plant species and its application to Iris atrofusca of the Northern Negev. *BioRisk* **3**: 137–160

VSN International Ltd. (2007). GenStat discovery edition 3. Hempstead: VSN International Ltd.

Wang, M.L., Zhu, C., Barkley, N.A., Chen, Z., Erpelding, J.E., Murray, S.C., Tuinstra, M.R., Tesfaye Tesso, Pederson, G.A. and Yu, J. (2009). Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. *Theoretical and Applied Genetics* **120**: 13–23.

Weir, S. and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

Weis, A.E. (2005). Assessing the ecological fitness of recipients. In: Poppy, G.M. and Wilkinson, M.J. (eds.), *Gene Flow from GM Plants*. Blackwell Publishing Ltd, Oxford, UK. pp. 143-168.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). DNA Fingerprinting in Plants: Principles, Methods, and Applications. 2nd Ed, CRC Press, Taylor & Francis Group, Boca Raton.

Wilkinson, M.J., Elliott, L.J., Allainguillaume, J., Shaw, M.W., Norris, C., Welters, R., Alexander, M., Sweet, J. and Mason, D.C. (2003). Hybridization between *Brassica napus* and *B. rapa* on a national scale in then United Kingdom. *Science* **302**: 457–459.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics* **15**:323-354.

Wright, S. (1969). Evolution and the genetics of populations, vol. 2, The Theoty of Gene Frequencies. The University of Chicago Press, Chicago, IL.

Wynne, J.C., Emery, D.A. and Rice, P.W. (1970). Combining Ability Estimates in *Arachis hypogaea* L.II. Field Performance of F1 Hybrids. *Crop Science* **10**: 713-515.

Yang, W., de Oliveria, A.C., Godwin, I., Schertz, K.F., Bennetzen, J.L. (1996). Comparison of DNA marker technologies in characterizing plant genome diversity: variability in Chinese sorghums. *Crop Science* **36**: 1669-1676.

Zhan, Q.W., Zhang, T.Z., Wang, B.H. and Li, J.Q. (2008). Diversity comparison and phylogenetic relationships of S. bicolor and S. sudanense as revealed by SSR markers. *Plant Science* **174**: 9–16.

Zhao, Z. (2007). The Africa biofortified sorghum project –applying biotechnology to develop nutritionally improved sorghum for Africa. In: Xu, Z. *et al* (eds) *Biotechnology and sustainable agriculture 2006 and beyond*, Springer, pp 273–277

Zhao, Z.Y., Cai, T., Tagliani, L., Miller, M., Wang, N., Pang, H., Rudert, M., Schroeder, S., Hondred, D., Seltzer, J. and Pierce, D. (2000). Agrobacterium mediated sorghum transformation. *Plant Molecular Biology* **44**: 789-798.

Zhao, Z.Y., Glassman, K., Sewalt, V., Wang, N., Miller, M., Chang, S., Thompson, T., Catron, S., Wu, E., Bidney, D., Kedebe, Y. and Jung, R. (2003). Nutritionally improved transgenic sorghum. In: Vasil, I.K. (ed.) *Plant biotechnology 2002 and beyond*, Kluwer Academic Publishers, Orlando, Florida, pp. 413–416

Zhu, H., Muthukrishnan, S., Krishnaveni, S., Wilde, G., Jeoung, J.M. and Liang, G.H. (1998). Biolistic transformation of sorghum using a rice chitinase gene. *Journal of Genetics and Breeding* **52**: 243-252.

Zongo, J., Gouyon, P.H. and Sandmeier, M. (1993). Genetic variability among sorghum accessions from the Sahelian agroecological region of Burkina Faso. *Biodiversity and Conservation* **2**: 627-636.

8. APPENDICES

Appendix Table 1. Description of measurements for quantitative traits

Character	Description
Plant height	The distance from the soil surface to the tip of the head during flowering measured in cm
Flag leaf length	The length of the flag leaf from the base to the tip measured in cm
Flag leaf width	Measured in cm at the widest part of the flag leaf
Leaf number	The total number of leaves
Number of tillers per panicle	The number of tillers
Head length	The length of the head measured in cm from the base to the tip
Head width	The width of the head measured in cm at the widest part
Hundred grain weight	Weight in grams of hundred grains

Appendix Table 2. Codes of the measured qualitative phenotypic characters and their description in wild sorghum (IBPGRI/ICRISAT, 1993; Bioversity International, 2010)

Trait	Code	Description	Code	Description
Awn	1	present	2	absent
Glume hair	1	present	2	absent
Glume color	2	sienna	6	black
	4	red		
Grain colour	1	white	8	light brown
	2	chalky white	9	brown
	4	grey	10	black
	5	light red	11	Purple
	6	red	13	reddish brown
	7	yellow		
Grain covering	4	grain fully covered	5	glumes longer than grain
Grain shattering	1	very low	7	high
	3	low	9	very high
	5	intermediate		
Grain size	1	very small	4	large
	2	small	5	very large
	3	medium		
Panicle compactness and shape	1	very lax panicle	6	semi-loose erect primary branches
	2	very loose erect primary branches	7	semi-loose drooping primary branches
	3	very loose drooping primary branches	8	Semi-compact elliptic
	4	loose erect primary branches	10	Compact elliptic
	5	loose drooping primary branches		

Appendix Table 3. Indices of diversity of wild and cultivated sorghum for the 9 microsatellite loci. NA = Number of alleles, PIC = polymorphism information content, Rs = allelic richness, Ho= observed heterozygosity, Hs=the mean expected heterozygosity in the sub populations, Ht=expected heterozygosity in the total population, D_{st}=Nei's (1978) unbiased genetic distance, G_{ST}= analogous to Wright (1951) F_{ST}, and Slatkin (1995) R_{ST}.

SSR Locus	NA	PIC	Rs	Ho	Hs	Ht	F_{ST}	D_{ST}	G_{ST}	R_{ST}
Cultivated										
Sb5-206	14.0	0.75	4.29	0.16	0.49	0.78	0.41	0.29	0.37	0.44
Sb1-1	27.0	0.85	5.77	0.15	0.53	0.87	0.42	0.34	0.39	0.77
Sb6-34	8.0	0.50	2.99	0.11	0.31	0.54	0.47	0.24	0.44	0.02
Sb5-256	3.0	0.54	2.24	0.32	0.38	0.61	0.41	0.23	0.38	0.45
Sb4-72	9.0	0.61	3.41	0.13	0.41	0.66	0.42	0.26	0.39	0.19
Sb6-84	12.0	0.72	3.88	0.13	0.45	0.75	0.44	0.30	0.40	0.71
Sb4-121	8.0	0.62	3.42	0.05	0.41	0.68	0.42	0.26	0.39	0.38
Sb6-57	8.0	0.48	2.58	0.12	0.38	0.57	0.36	0.19	0.33	0.12
SbKAFGK1	8.0	0.63	2.79	0.07	0.38	0.69	0.49	0.32	0.46	0.50
Mean	10.78	0.63	3.48	0.13	0.41	0.68	0.43	0.27	0.40	0.40
Ethiopian Wild										
Sb5-206	27.0	0.92	4.48	0.21	0.63	0.92	0.33	0.29	0.32	0.29
Sb1-1	32.0	0.92	3.85	0.17	0.53	0.93	0.44	0.40	0.43	0.68
Sb6-34	15.0	0.85	3.97	0.33	0.62	0.87	0.30	0.25	0.29	0.25
Sb5-256	11.0	0.71	2.94	0.07	0.48	0.76	0.38	0.28	0.36	0.22
Sb4-72	20.0	0.84	4.00	0.16	0.60	0.87	0.32	0.27	0.31	0.31
Sb6-84	21.0	0.82	3.39	0.20	0.52	0.83	0.38	0.31	0.37	0.39
Sb4-121	14.0	0.85	3.44	0.14	0.57	0.87	0.35	0.30	0.34	0.35
Sb6-57	15.0	0.73	2.85	0.14	0.42	0.76	0.46	0.34	0.45	0.39
SbKAFGK1	10.0	0.61	2.43	0.23	0.37	0.65	0.44	0.28	0.43	0.50
Mean	18.33	0.81	3.48	0.18	0.53	0.83	0.38	0.30	0.36	0.38
ICRISAT wild										
Sb5-206	17.0	0.82	3.02	0.26	0.43	0.84	0.51	0.40	0.48	0.33
Sb1-1	20.0	0.89	3.32	0.16	0.52	0.90	0.45	0.38	0.42	0.62
Sb6-34	8.0	0.77	2.13	0.09	0.33	0.80	0.62	0.48	0.59	0.81
Sb5-256	4.0	0.50	1.88	0.12	0.32	0.56	0.46	0.24	0.43	0.40
Sb4-72	16.0	0.85	2.93	0.13	0.44	0.87	0.53	0.43	0.50	0.66
Sb6-84	11.0	0.83	2.21	0.16	0.38	0.85	0.58	0.47	0.55	0.68
Sb4-121	9.0	0.78	2.46	0.14	0.45	0.81	0.47	0.36	0.44	0.45
Sb6-57	10.0	0.78	2.72	0.10	0.43	0.81	0.50	0.38	0.47	0.19
SbKAFGK1	8.0	0.58	2.18	0.12	0.30	0.61	0.53	0.31	0.51	0.65
Mean	11.44	0.76	2.54	0.14	0.40	0.78	0.52	0.38	0.49	0.53

Appendix Table 4. Pair wise F_{ST} (Cockerham and Weir, 1984) for the wild sorghum popula tions

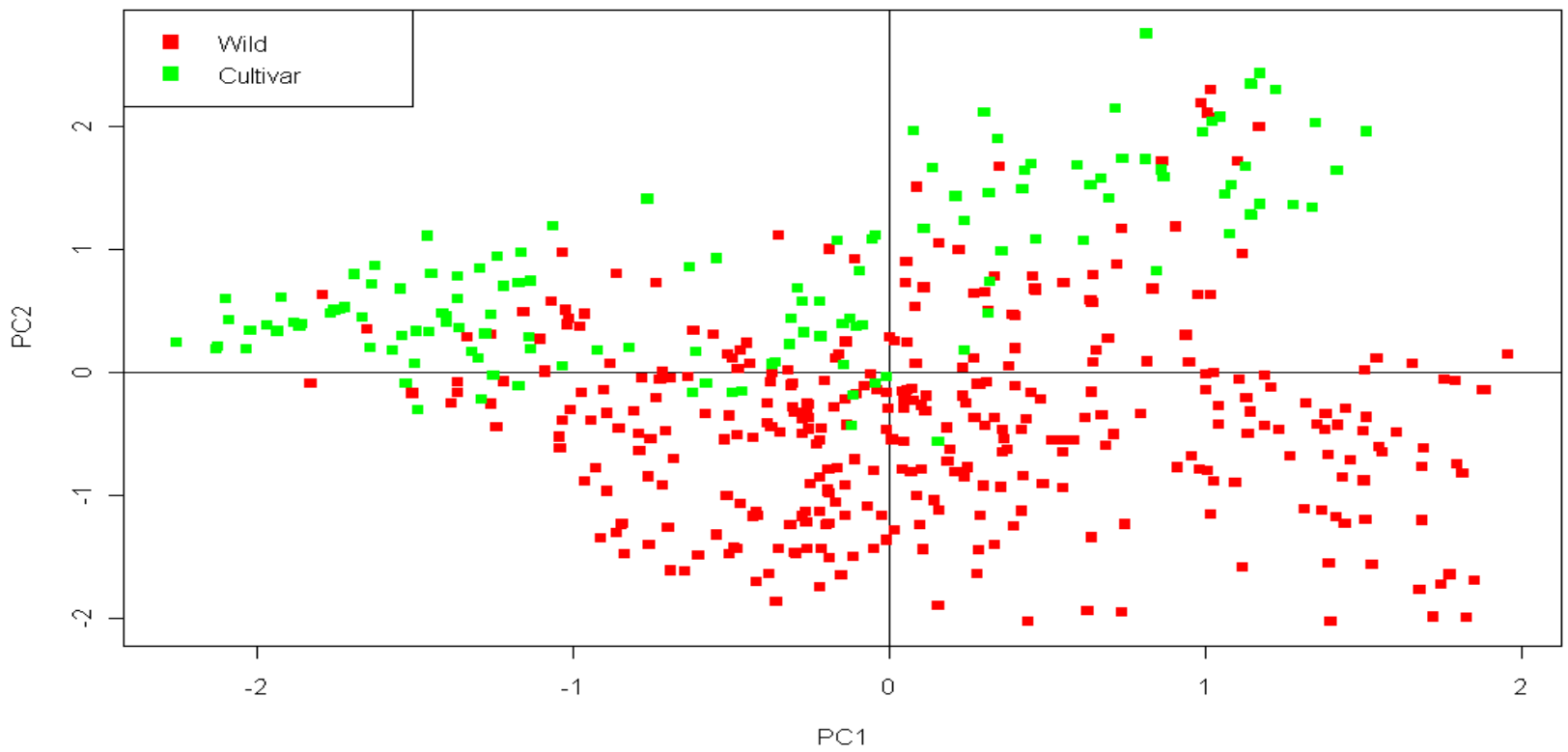
	AW	G-1	G-2	G-3	H-1	H-2	P-1	P-2	P-3	P-4	T-1	T-2	T-3	T-4	T-5	W-1	W-2	W-3
G-1	0.337																	
G-2	0.318	0.081																
G-3	0.412	0.282	0.145															
H-1	0.422	0.224	0.202	0.227														
H-2	0.399	0.405	0.408	0.507	0.524													
P-1	0.455	0.261	0.250	0.327	0.312	0.570												
P-2	0.376	0.160	0.161	0.267	0.257	0.482	0.128											
P-3	0.400	0.187	0.215	0.354	0.292	0.510	0.333	0.210										
P-4	0.452	0.234	0.253	0.344	0.284	0.560	0.299	0.179	0.271									
T-1	0.546	0.465	0.469	0.582	0.585	0.674	0.620	0.508	0.529	0.591								
T-2	0.468	0.348	0.288	0.293	0.337	0.571	0.341	0.310	0.407	0.368	0.634							
T-3	0.322	0.224	0.206	0.318	0.250	0.407	0.363	0.270	0.313	0.342	0.493	0.381						
T-4	0.368	0.215	0.185	0.215	0.102	0.472	0.311	0.227	0.262	0.287	0.508	0.299	0.248					
T-5	0.335	0.161	0.129	0.237	0.128	0.454	0.249	0.189	0.245	0.261	0.536	0.275	0.174	0.149				
W-1	0.531	0.369	0.344	0.362	0.333	0.626	0.455	0.334	0.453	0.400	0.678	0.415	0.368	0.362	0.313			
W-2	0.335	0.207	0.176	0.307	0.285	0.398	0.281	0.196	0.278	0.316	0.508	0.294	0.209	0.243	0.182	0.327		
W-3	0.635	0.503	0.474	0.567	0.522	0.698	0.582	0.527	0.543	0.566	0.794	0.528	0.538	0.515	0.470	0.560	0.461	
W-4	0.392	0.404	0.391	0.458	0.485	0.504	0.515	0.418	0.482	0.498	0.608	0.470	0.324	0.436	0.402	0.477	0.330	0.679

Appendix Table 5. Wild sorghum population pairs with some level of gene flow (Nm≥1)

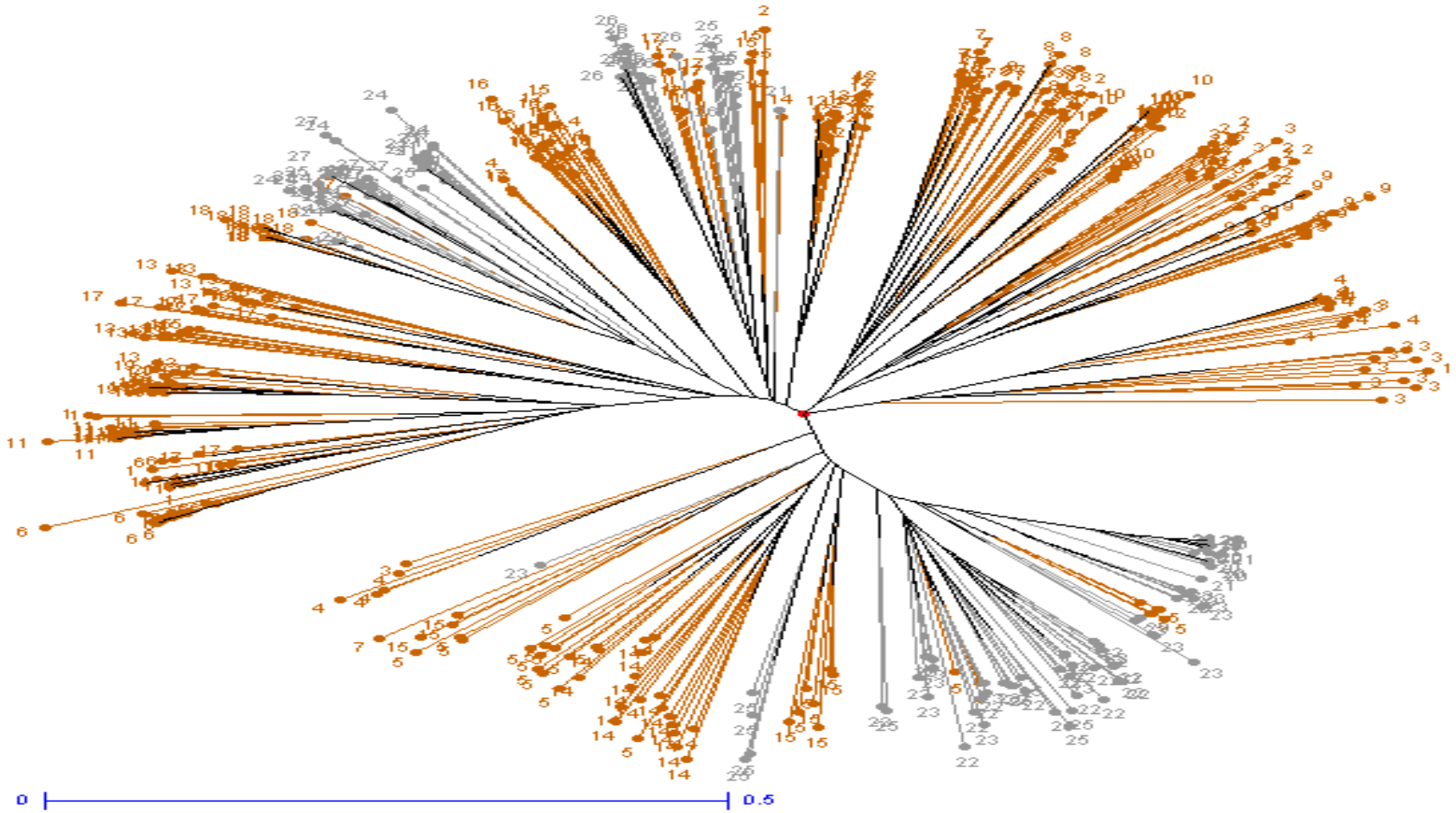
POP1	POP2	FST	Nm
G-1	G-2	0.081	3
G-1	P-2	0.160	1
G-1	P-3	0.187	1
G-1	T-5	0.161	1
G-2	G-3	0.145	1
G-2	P-2	0.161	1
G-2	T-4	0.185	1
G-2	T-5	0.129	2
G-2	W-2	0.176	1
H-1	T-4	0.102	2
H-1	T-5	0.128	2
P-2	P-1	0.128	2
P-2	P-4	0.179	1
P-2	T-5	0.189	1
P-2	W-2	0.196	1
T-3	T-5	0.174	1
T-4	T-5	0.149	1
T-5	W-2	0.182	1

Appendix Table 6. Rousset (1997) genetic distance (upper diagonal) and the geographic distance of the collection sites (lower diagonal) matrices used to estimate gene flow by isolation by distance

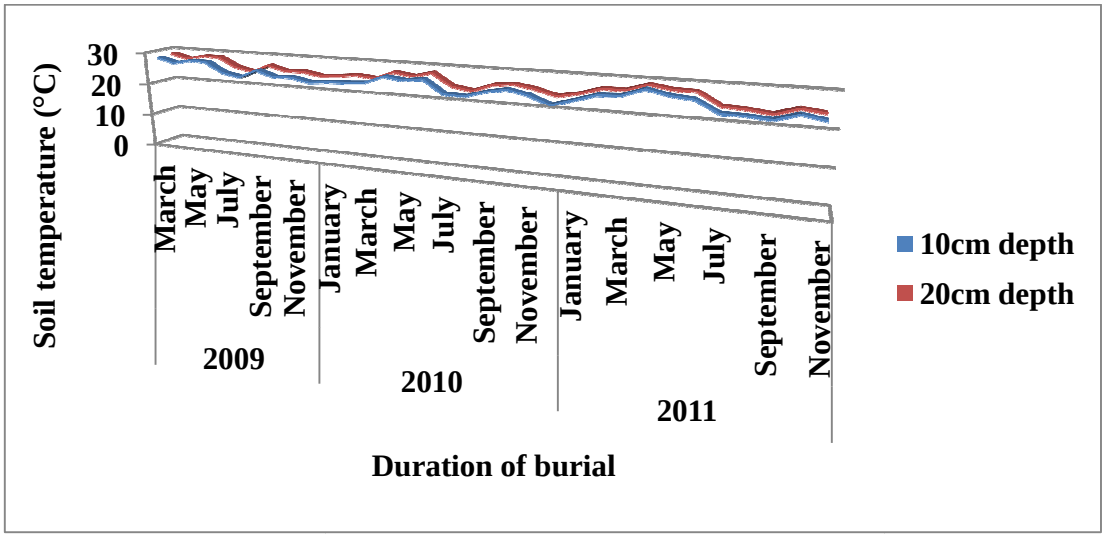
	AW	G-2	G-4	G-6	H-1	H-2	P-1	P-4	P-5	P-6	T-1	T-4	T-5	T-6	T-8	W-2	W-4	W-6	W-7
AW	0	0.509	0.465	0.700	0.730	0.663	0.834	0.603	0.666	0.823	1.204	0.878	0.475	0.582	0.503	1.132	0.504	1.737	0.645
G-2	290.7	0	0.088	0.392	0.289	0.681	0.353	0.190	0.230	0.305	0.870	0.533	0.288	0.274	0.192	0.585	0.260	1.010	0.678
G-4	288.7	3.5	0	0.169	0.253	0.689	0.334	0.192	0.273	0.339	0.884	0.405	0.260	0.227	0.147	0.524	0.213	0.899	0.641
G-6	281.2	11.0	8.0	0	0.293	1.027	0.485	0.364	0.548	0.524	1.392	0.414	0.466	0.274	0.310	0.568	0.443	1.309	0.844
H-1	236.1	525.2	523.4	516.0	0	1.099	0.452	0.346	0.411	0.397	1.408	0.509	0.333	0.113	0.147	0.499	0.399	1.090	0.941
H-2	67.4	354.1	352.3	344.9	171.1	0	1.327	0.931	1.042	1.274	2.063	1.332	0.687	0.894	0.832	1.675	0.661	2.313	1.016
P-1	460.4	337.4	334.7	333.0	662.6	522.4	0	0.147	0.500	0.426	1.630	0.518	0.569	0.451	0.331	0.834	0.391	1.391	1.060
P-4	480.8	362.7	360.0	358.4	679.2	541.9	25.4	0	0.265	0.218	1.033	0.450	0.369	0.294	0.233	0.501	0.244	1.116	0.719
P-5	486.1	374.5	371.7	369.9	681.5	546.5	37.5	13.5	0	0.372	1.125	0.687	0.455	0.355	0.325	0.829	0.386	1.188	0.930
P-6	467.2	345.8	343.1	341.4	668.0	528.8	8.5	16.9	29.3	0	1.444	0.581	0.519	0.402	0.353	0.667	0.461	1.302	0.991
T-1	393.3	523.2	519.7	512.9	459.7	414.2	377.0	376.8	370.1	376.7	0	1.730	0.972	1.034	1.154	2.107	1.033	3.847	1.551
T-4	637.6	689.0	685.5	680.5	722.5	668.5	414.6	398.5	385.8	409.1	263.0	0	0.616	0.426	0.380	0.708	0.417	1.118	0.885
T-5	702.1	683.3	680.2	676.8	828.8	745.3	361.2	338.5	325.0	353.6	380.2	166.4	0	0.330	0.211	0.581	0.264	1.165	0.479
T-6	695.8	679.0	675.9	672.5	821.8	738.8	358.2	335.8	322.3	350.7	372.9	160.2	7.3	0	0.175	0.568	0.322	1.062	0.772
T-8	666.1	659.8	656.6	652.9	788.2	707.8	346.4	325.2	311.7	339.3	338.1	131.6	42.3	35.0	0	0.455	0.223	0.887	0.671
W-2	190.2	376.2	373.0	365.3	298.2	213.8	380.7	393.5	393.8	384.8	203.4	454.7	538.7	531.9	499.7	0	0.486	1.275	0.913
W-4	353.8	491.3	487.9	480.9	425.7	375.0	368.8	371.2	365.8	369.4	39.5	298.9	407.7	400.5	366.1	163.9	0	0.854	0.491
W-6	363.5	506.8	503.4	496.3	426.4	382.3	383.7	385.7	380.0	384.2	33.4	296.4	410.9	403.7	369.1	173.3	16.4	0	2.116
W-7	359.8	494.7	491.3	484.3	432.0	381.3	367.5	369.5	363.9	368.0	33.6	292.5	401.7	394.5	360.1	170.0	6.5	16.2	0



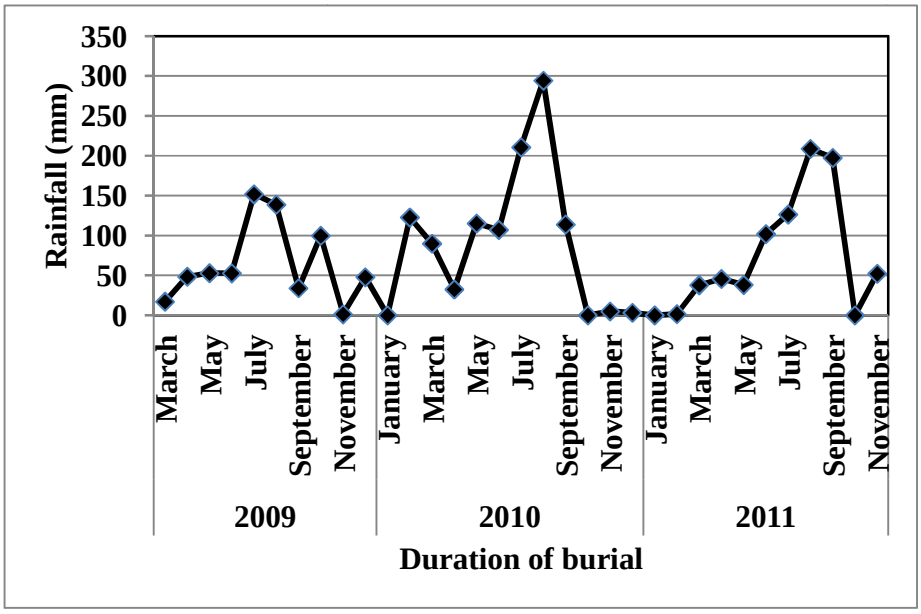
Appendix Fig.1. Principal component analysis showing the the clustering pattern of individual wild (red square) and cultivated (green square) sorghum plants



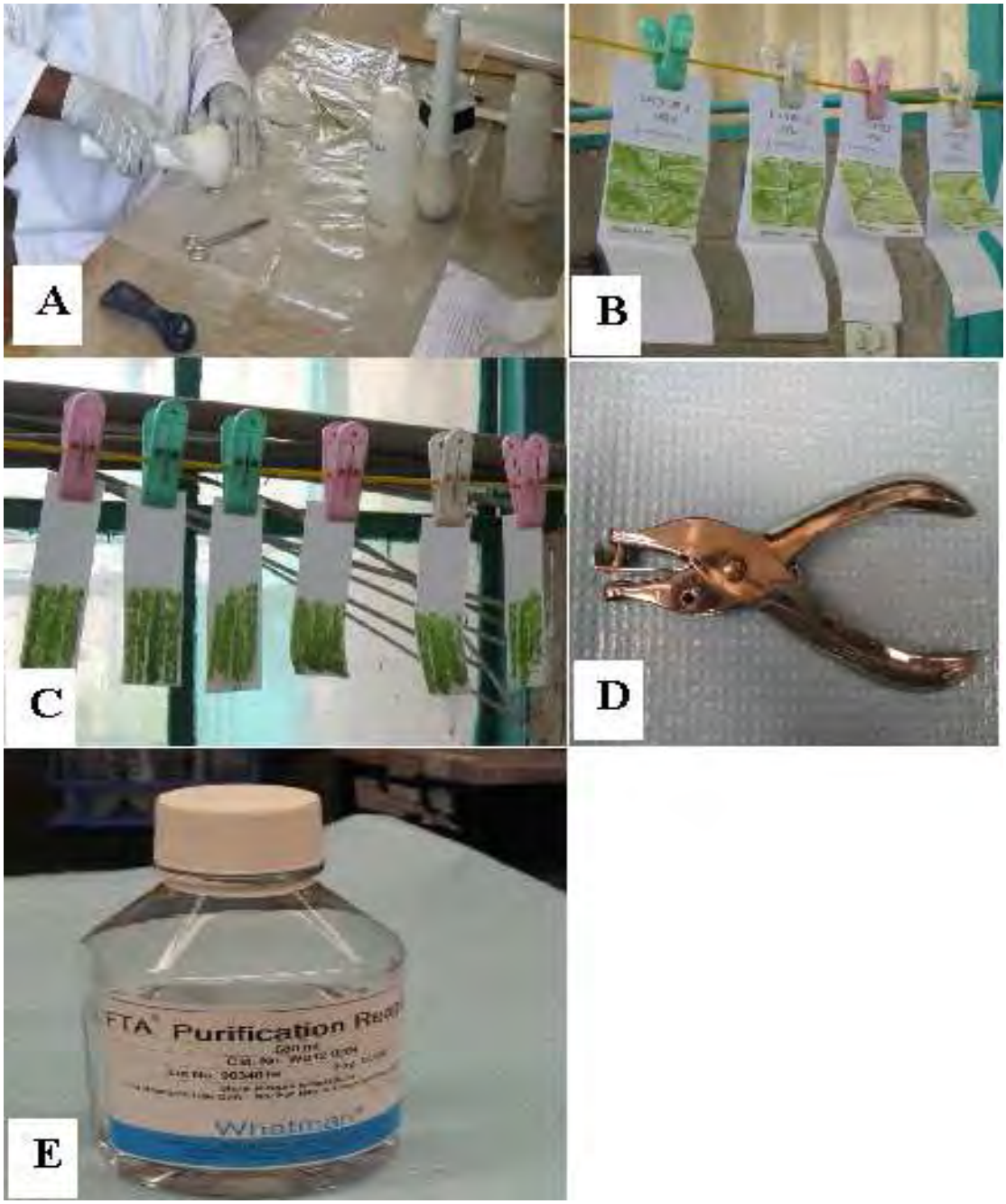
Appendix Fig. 2. A neighbor joining cluster dendrogram based on pair-wise simple matching distance among the 540 individuals of cultivated (gray) and wild (brown). Numbers represent source populations; population ID codes are, Wild: 1=AW, 2=G-1, 3=G-2, 4=G-3, 5=H-1, 6=H-2, 7=P-1, 8=P-2, 9=P-3, 10=P-4, 11=T-1, 12=T-2, 13=T-3, 14=T-4, 15=T-5, 16=W-1, 17=W-2, 18=W-3, 19=W-4; Cultivated: 20=GC1, 21=GC2, 22=PC1, 23=PC2, 24= TC1 [Degalit cultivar], 25= WC1, 26= WC2 [improved cultivar 76T1#23], 27=WC3 [Degalit cultivar]



Appendix Fig. 3. Monthly average soil temperature at Melkassa for the duration of the seed burial experiment



Appendix Fig 4. Monthly rainfall distribution at Melkassa for the duration of the seed burial and fitness experiments



Appendix Fig.5. A) Collection of squashes from sorghum leaf, B) Air drying FTA plant saver card and C) Whatman chromatography paper, D) Single hole paper punch, used to puch out a disc from cards, and E) FTA purification reagent

Declaration

I, the undersigned, declare that this Dissertation is based on my original work and that it has not been presented for a degree in any other university. All sources of materials have been duly acknowledged.

Asfaw Adugna Signature: _____

This Dissertation has been submitted for examination with my approval as supervisor of the Dissertation.

Prof. Endashaw Bekele Signature: _____