Phylogenetics of the Genus *Sorghum*, Genetic Diversity and Nutritional Value of its Cultivated Species

Dickson Ng'uni
Faculty of Landscape Planning, Horticulture and Agricultural Science
Department of Plant Breeding and Biotechnology
Alnarp

Doctoral Thesis
Swedish University of Agricultural Sciences
Alnarp 2011

Acta Universitatis agriculturae Sueciae
2011:63

ISSN 1652-6880
ISBN 978-91-576-7607-8

Print: SLU Service/Repro, Alnarp 2011

Phylogenetics of the Genus *Sorghum*, Genetic Diversity and Nutritional Value of its Cultivated Species.

Abstract

Sorghum (*Sorghum bicolor* (L.) Moench) ranks fifth among cereal crops and second highest in production after maize in Africa particularly in the semi arid regions where it is a food security crop. This study assessed phylogenetic relationships of the species within the genus *Sorghum*, genetic diversity and the nutritional value of cultivated sorghum for its use for breeding and conservation.

The phylogenetic analyses based on sequence data from four chloroplast DNA (cpDNA) regions and the internal transcribed spacer (ITS) revealed that *S. laxiflorum* and *S. macrospermum* were more closely related to *Eu-sorghum* species than to any other section and that the former two species are best merged into one section. Assessment of 27 Zambian sorghum accessions and 14 accessions from Malawi, Tanzania and Zambia for genetic diversity based on microsatellite markers revealed a significant genetic variation within and largely among (>80%; $p < 0.001$) sorghum accessions. Bioassay for grain mineral contents of 27 farmer varieties of sorghum from southern Africa and 13 improved varieties showed that improved sorghum varieties were superior in macronutrients while farmer varieties showed superiority for grain Fe and Zn contents. Morphological characterization of 17 accessions from southern Africa revealed considerable variation among accessions and plant height, days to 50% flowering and inflorescence length were more important discriminating traits. The studies in this thesis provide insights into the extent and pattern of genetic relationships within the genus *Sorghum* and reveal significant genetic variation for nutritional value improvement.

Keywords: cpDNA, genetic diversity, ITS, macronutrients, micronutrients, phylogenetics, proteins, Sorghum, SSR

Author's address: (1) Dickson Ng'uni, Swedish University of Agricultural Sciences (SLU), Department of Plant Breeding and Biotechnology,
P.O. Box 101, SE-230 53 Alnarp, Sweden.
E-mail: Dickson.Nguni@slu.se; or (2) Mount Makulu Research Station, P/B 7, Chilanga, Zambia. *E-mail*: dickson.nguni@gmail.com

Dedication

To my family for the love, passion, unfailing and unconditional support provided throughout my study period. You stood by in all circumstances.

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Ng'uni, D., Geleta, M., Fatih, M. & Bryngelsson, T. (2010). Phylogenetic analysis of the genus *Sorghum* based on combined sequence data from cpDNA regions and ITS generated well-supported trees with two major lineages. Annals of Botany 105(3), 471-480.

II Ng'uni, D., Geleta, M & Bryngelsson, T. (2011). Genetic diversity in sorghum [*Sorghum bicolor* (L.) Moench] accessions of Zambia as revealed by simple sequence repeats (SSR). Hereditas 148, 52-62.

III Ng'uni, D., Geleta, M., Johansson, E., Fatih, M. & Bryngelsson, T. Characterization of the Southern African sorghum varieties for mineral contents: Prospects for breeding for mineral dense lines (African Journal of Food Science, in press).

IV Ng'uni, D., Geleta, M., Fatih, M. & Bryngelsson, T. Comparative genetic diversity and nutritional quality variation among Southern African sorghum [*Sorghum bicolor* (L.) Moench] (Submitted).

V Ng'uni, D., Geleta, M., Fatih, M. & Bryngelsson, T. Agromorphological traits variation in sorghum [*Sorghum bicolor* (L) Moench] germplasm from southern Africa (Manuscript).

Paper I is reproduced with the permission of the publishers.

The contribution of Dickson Ng'uni to the papers included in this thesis was as follows:

I Planned, requested germplasm, raised and managed seedlings in a greenhouse, extracted DNA, carried out all laboratory work, analysed data and wrote the manuscript with input from co-authors

II Planted seeds in a greenhouse, collected young leaf tissue for DNA extraction, planned and carried out all laboratory work, analyzed data and wrote the manuscript with input from co-authors

III Milled seed samples into flour, planned and carried out laboratory digestion and analysis of samples, analysed data and wrote the manuscript with input from co-authors

IV Planted seed for accessions in a greenhouse, collected young leaf tissue for DNA extraction, milled seed samples into flour, planned and carried out all laboratory work, analyzed data and wrote the manuscript with input from co-authors

V Planned, set up and carried out the experimental field trial, scored and analysed data and wrote the manuscript with input from co-authors

1 Introduction

1.1 Nomenclature and a historical perspective of the genus *Sorghum*

Sorghum belongs to the family Poaceae, subfamily Panicoideae, the tribe Andropogoneae, subtribe Sorghinae and genus Sorghum (Clayton & Renvoize, 1986). Garber (1950) and Celarier (1959) further segmented the genus into five subgenera: *Eu-sorghum* Stapf emend. Snowden, *Chaetosorghum*, *Heterosorghum*, *Parasorghum* and *Stiposorghum*. Genus *Sorghum* Moench is highly heterogeneous and together with the genus *Cleistachne* Bentham they form Sorghastrae (Garber, 1950), one of the 16 sub-tribes in the tribe Andropogoneae. The seventeenth century saw an increase in the number of references to sorghum with several authors describing the genus (Smith & Frederiksen, 2000).

Sorghum was first described by Linnaeus in 1753 and was referred to as Holcus. Later on in 1794, Moench distinguished the genus *Sorghum* from the genus *Holcus* (Clayton, 1961). Several other authors have discussed the systematics, origin and evolution of sorghum after Linnaeus (Snowden, 1936; de Wet & Harlan, 1971; Doggett, 1988).

The subgenus *Sorghum* (*Eu-sorghum*) includes annual cultivated sorghum from Africa and perennial taxa from S. Europe and Asia. Three species are recognized under the subgenus *Eu-sorghum* as *S. halepense* (L.) Pers. occurring in India, *S. propinquum* (Kunth) Hitchc found in Southeast Asia and *S. bicolor* (L.) Moench originated in Africa (de Wet, 1978).

1.2 Evolution and taxonomy of sorghum

Sorghum, maize (*Zea mays* L.), sugarcane (*Saccharum officinarum* L.) and millet (*Eleusine coracana* L. Gaerth) are grass species that diverged from rice about 50 million years ago (Doebley *et al.*, 1990). Sorghum and maize shared a common ancestor as recently as 20–24 million years ago (Gaut & Doebley, 1997). Sorghum is the closest relative of sugarcane which diverged from the latter around 5 million years ago (Al-Janbi *et al.*, 1997). Phylogenetic analysis of molecular data from nuclear rRNA sequences showed that *S. bicolor* and *S. officinarum* are very closely related and that these taxa recently shared a common ancestry (Hamby & Zimmer, 1988). Springer et al. (1989) reaffirmed the sequence similarity between the two through comparison of rDNA sequence which also revealed that they are closely related to *Z. mays* L. Recently, the chroloplast DNA of sorghum (cpDNA; Fig. 1) has been completely sequenced (Saski *et al.*, 2007). The phylogenetic relationships inferred from five intergenic spacer regions of chroloplast genome confirm the closeness among *Z. mays*, *S. officinarum* and *S. bicolor* (Saski *et al.*, 2007).

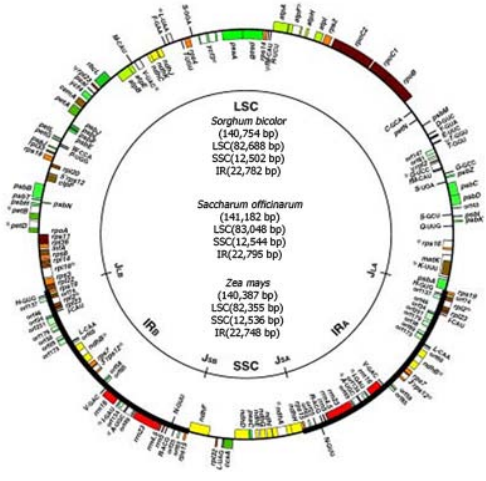

Figure 1. Gene map of *Sorghum bicolor* chloroplast genome. Modified from Saski et al. (2007).

The genus Sorghum has been variously described by several taxonomists (Garber, 1950; Lazarides et al., 1991; Spangler, 2003). The genus *Sorghum* has 25 recognized species that have been classified into five subgenera or sections namely *Eu-sorghum*, *Chaetosorghum*, *Heterosorghum*, *Parasorghum* and *Stiposorghum* (Garber, 1950). Section *Eu-sorghum* includes cultivated

sorghum *S. bicolor* (2n = 20) and its subspecies *drummondii* and *arundinaceum*, and the wild species, *S. propinquum* (Kunth) Hitch, (2n = 20), *S. halepense* (L.) Pers. (2n = 40) and *S. almum* Parodi (de Wet, 1978). Species of section *Eu-sorghum* have a natural range through Africa and southern Asia (de Wet, 1978; Duvall & Doebley, 1990). *Eu-sorghum* was earlier considered under two subsections, Arundinacea and Halepensia (Snowden, 1936). The subsection Arundinacea, commonly found in tropical Africa and India, consists of *S. bicolor* (L.) Moench, *S. arundinaceum* (Desv.) Stapf (2n = 20) and *S. drummondii* (Steud.) Millsp. *S. propinquum* (Kunth) Hitchcock, *S. halepense* (L.) Pers and *S. almum* Parodi form subsection Halepensia, and are found in the Mediterranean region and Southeast Asia. However, *Eu-sorghum* was considered to include cultivated sorghums and their closest wild relatives (De Wet and Huckay, 1967).

Based on the gene pool concept, the genus *Sorghum* has been classified into three gene pools. The primary gene pool includes *S. bicolor* subsp. *bicolor* cultivars and races, *S. bicolor* subsp. *arundinaceum*, *S. bicolor* subsp. *drummondii* and *S. propinquum* (Stenhouse et al., 1997). The secondary gene pool consists of *S. halepense*. The tertiary genepool of the genus *Sorghum* includes all wild sorghum belonging to subgenera *Chaetosorghum*, *Heterosorghum*, *Parasorghum* and *Stiposorghum*.

1.3 Biogeography of the genus *Sorghum* and domestication of its cultivated species

There are about 25 species in the genus *Sorghum*, two third of which are wild *Sorghum* species of sections *Chaetosorghum*, *Heterosorghum*, *Parasorghum* and *Stiposorghum*. Sections *Chaetosorghum* and *Heterosorghum,* each containing single species *S. macrospermum* E.D. Garber and *S. laxiflorum* F.M. Bailey, respectively occur in Australia and South Pacific. Section *Para-sorghum* consists of the five Australian species *S. grande* Lazarides, *S. leiocladum* (Hack.) C.E. Hubb., *S. matarankense* E.D. Garber & Snyder, *S. nitidum* (Vahl) Pers., *S. timorense* (Kunth) Buse, and the two southeastern African, India and Southeast Asian species *S. purpureo-sericeum* (Hochst. ex A. Rich.) Asch. & Schweinf. and *S. versicolor* Andersson. Section *Stiposorghum* has ten species that are endemic to Australia including *S. amplum* Lazarides, *S. angustum* S.T. Blake, *S. brachypodum* Lazarides, *S. bulbosum* Lazarides, *S. ecarinatum* Lazarides, *S. exstans* Lazarides, *S. interjectum* Lazarides, *S. intrans* F. Muell. ex Benth., *S. plumosum* (R. Br.) P. Beauv., and *S. stipoideum* (Ewart

& Jean White) C.A. Gardner & C.E. Hubb (Garber, 1950; Lazarides *et al.*, 1991).

There are two schools of thought that relate to the domestication of *Sorghum bicolor*. One school of thought suggests that the Mande people around the headwaters of the Niger River may have domesticated sorghum (Murdock, 1959). The other with backing of some archeological evidence postulates that sorghum was domesticated in the northeast Africa about 3000 BC (Doggett, 1965). There are also suggestions that cultivated sorghum was domesticated by selections from a wild progenitor, subspecies *verticilliflorum*, about 5000-7000 years ago (Purseglove, 1972). The northeastern quadrant of Africa seems to be strongly supported as the region where sorghum was domesticated given the existence of greatest variability of cultivated and wild sorghum in that region (de Wet & Harlan, 1971; Doggett, 1988) and the restriction of wild progenitors to the region (Zohary & Hopf, 2000). The diversity in sorghum was mainly created through the practice of disruptive selection and isolation, recombination in the extremely varied habitats and movement of people carrying one or more varieties of the species (Doggett, 1970). A balance of farmer selection for cultivated traits and natural selection for wild characteristics has generated improved sorghum types, wild types and intermediate types (Doggett, 1970).

Cultivated sorghum varieties were spread to other regions of Africa, India (approx. 1500–1000 BC), the Middle East (approx. 900–700 BC) through the movement of people along trade routes. It appears that sorghum moved into eastern Africa from Ethiopia around 200 AD or earlier. It was adopted and carried to the savannah countries of eastern and southern Africa by the Bantu people, who used the grain mainly to make beer. The Bantu people probably began their expansion from the region of southern Cameroon about the first century AD, moved along the southern border of the Congo forest belt and reached eastern Africa possibly before 500 AD. The present-day sorghums of central and southern Africa are closely related to those of the Tanzania and more distantly related to those of West Africa, as the equatorial forests were an effective barrier to this spread.

Presently, *Sorghum bicolor* (L.) Moench is identified with several given names in many parts of the world such as great millet and guinea corn in West Africa, kafir corn in South Africa, durra in Sudan, mtama in eastern Africa, jowar in India and kaoliang in China (Purseglove, 1972). Based on the morphological features of the inflorescence, grain and glumes, *S. bicolor*

subsp. *bicolor* as classified today include five basic races (Fig. 2): *bicolor*, *guinea*, *caudatum*, *durra* and *kafir* as well as ten intermediate races (Harlan & de Wet, 1972). According to Harlan and de Wet (1972), race *caudatum* is dominant on the north-eastern African savanna, in Sudan and Chad, parts of Nigeria, Cameroun, Uganda and Kenya. Race *guinea* is mostly grown in high rainfall regions of West Africa and also Mozambique, Malawi and Swaziland as well as in the southwest of Ethiopia. Race *kafir* is grown south of the equator covering parts of southern Africa. Race *durra* spans northeastern Africa, Arabia and parts of Asia and race *bicolor* is mainly grown in the west of the Rift Valley (Harlan & de Wet, 1972).

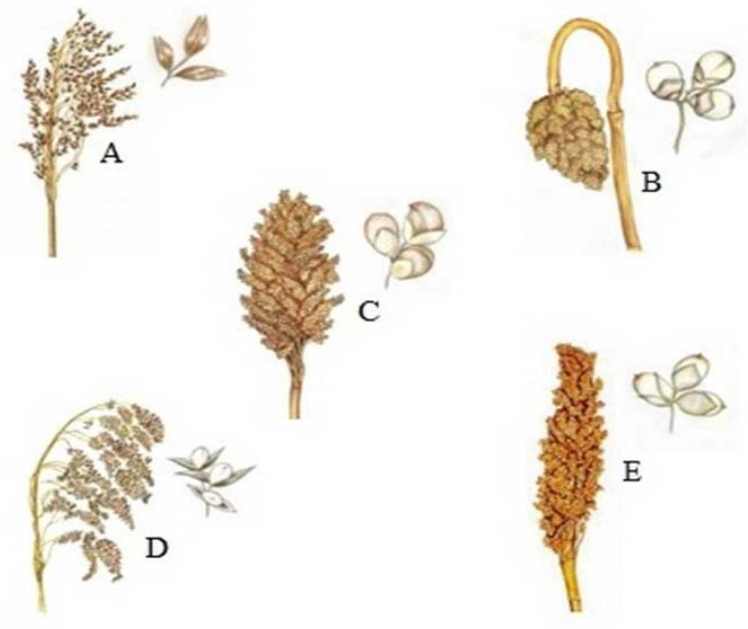

Figure 2. Sorghum races; A = *bicolor*, B = *caudatum*, C = *durra*, D = *guinea* and E = *kafir* (Harlan and de Wet, 1972).

1.4 Sorghum characteristics and adaptation

Sorghum is an annual grass species up to 5 m tall (Fig. 3), with one to several tillers that originate from the basal stem nodes supported by an extensive and deep root system. The plant has wide adaptation and can be grown between 40°N and 40°S across the equator at altitude of up to 2300 m (Doggett, 1988) and tolerates a wide range of soil conditions. Sorghum

performs better at temperature of 25 - 31°C but is susceptible to frost. It is mainly a rainfed crop of lowland, semi-arid areas of the tropics and sub-tropics (Craufurd *et al.*, 1999). Sorghum is adapted to drought conditions due to a number of morphological and physiological characteristics, including an extensive root system, waxy bloom on leaves that reduces water loss, and the ability to stop growth in periods of drought and resume it when the stress is relieved (Stenhouse *et al.*, 1997). An annual total of 400-800 mm of well distributed rainfall over the cropping season is adequate for the crop to reach maturity. The crop tolerates water logging and can also be grown in high rainfall areas. Naturally, sorghum is a short-day plant (SDP) but a wide genetic variation exists for its adaptation to the wide range of photoperiod and temperature of different environments(Craufurd *et al.*, 1999).

Figure 3. Sorghum plants at grain filling stage (A) lax inflorescence (B) Semi-compact to compact upright inflorescence and (C) semi-compact inflorescence with curved peduncle.

1.5 World sorghum production

Sorghum ranks fifth in the world after wheat, rice, maize and barley in terms of production, but in Africa the crop is the second highly produced cereal crop after maize (FAO, 2011). The crop plays a significant role in the provision of food security in Africa and Asia particularly in the semi arid tropics of the two continents. According to FAO 2008 ranking of countries by commodities, USA, Nigeria, India, Sudan and Argentina accounted for

the largest portion of world sorghum production (Fig. 4). Africa accounted for 42% of the world sorghum production (FAO, 2011), which is mainly grown for food at the small scale farmer level. In industrialized countries, sorghum is grown at commercial scale and largely for use as feed and fodder for livestock (Doggett, 1988) and sweet sorghum is used for ethanol production.

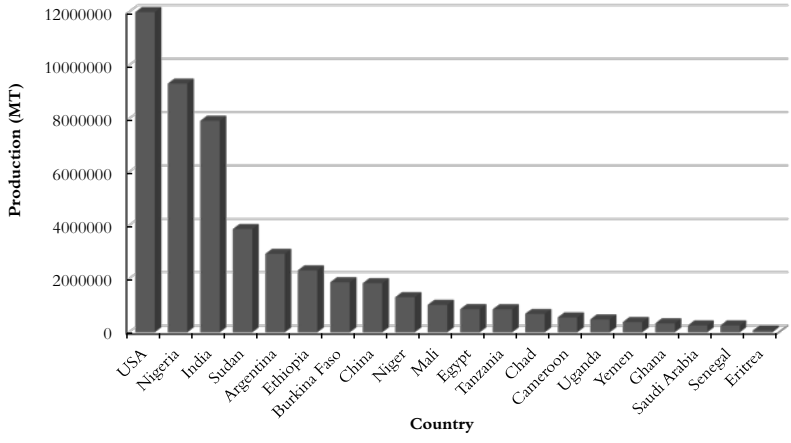

Figure 4. World leading sorghum producing countries, source: FAO (2011).

1.6 Uses and nutritive value of sorghum

The sorghum grain varies in colour from white, red and brown to pale yellow and deep purple-brown. The most common grain colours are white, bronze and brown. Grains are mainly spherical but vary in size and shape. The grains can be rounded and bluntly pointed, 4-8 mm in diameter (Purseglove, 1972). Grain size varies with variety and large grains with corneous endosperm are usually preferred for human consumption. Yellow endosperm with carotene and xanthophyll increases the nutritive value. Sorghum grain has a seed coat that may contain tannin in varying proportions depending on the variety (Dykes & Rooney, 2006).

Sorghum has a wide range of uses, such as food, beer brewing and livestock feed and fodder (Chakauya *et al.*, 2006). In addition, the crop is also used in the production of commercial alcohol, adhesives, waxes, construction materials and bio-ethanol from sweet sorghum varieties (Antonopoulou *et al.*, 2008). In the confectionary industry, the sorghum grain is used for baking of unleavened bread, biscuits and tortilla.

Starch is the main component of the sorghum grain and is the major form of carbohydrate in sorghum comprising amylopectin and amylase. Starch content of sorghum grain ranges from 56 to 73 g/100g with an average of 69.5 g/100g (Jambunathan & Subramanian, 1988). The sorghum grain protein content is within the range of 7 to 15% (Beta *et al.*, 1995; Nkongolo *et al.*, 2008; Wong *et al.*, 2010). Using the solubility-based classification (Jambunathan *et al.*, 1975), sorghum proteins have been divided into albumins, globulins, kafirins, cross-linked kafirins and glutelins. The kafirins, which are aqueous alcohol-soluble prolins account for about 50-70% of the proteins (Oria *et al.*, 1995; Duodu *et al.*, 2003; Vendemiatti *et al.*, 2008). Sorghum is also a good source of vitamin B complex and more than 20 mineral elements (BSTID-NRC, 1996) and is specifically rich in phosphorus, potassium, iron and zinc (Glew *et al.*, 1997; Anglani, 1998). For example, sorghum is a better source of zinc, an important micronutrient for pregnant women than corn and wheat (Hopkins *et al.*, 1998).

1.7 Pest and diseases of sorghum

A range of insect pests and diseases have been reported to cause economic loss in sorghum (House, 1985). Notable among the insect pests are sorghum aphids (*Melanaphis sacchari*), sorghum shoot fly (*Atherigona soccata* Rondani) and stalk borers [*Busseola fusca*; *Chilo partellus* (Swinhoe)]. Sorghum midge, *Contarinia sorghicola* Coquillett, has been reported to occasionally cause serious crop damage to developing grain (House, 1985; Reddy *et al.*, 2006). Sorghum diseases of economic importance include grain moulds, downy mildew (*Peronosclerospora sorghi*), charcoal rot (*Macrophomina phaseoline*) and Anthracnose (*Colletotrichum graminicola*) (Reddy *et al.*, 2006). The parasitic plant, striga (*Striga* spp), is one of the major problems in affecting sorghum production in Africa and Asia. Birds, especially *Quelea quelea*, can be a problem in sorghum and can cause serious economic losses. There are also post harvest grain losses in sorghum due to damage caused by storage pests, of which the principal ones are rice weevil (*Sitophilus oryzae*), flour beetle (*Tribolium castaneum*) and the grain moth (*Sitotroga cerealella*).

1.8 Genetic diversity, conservation and utilization of crops and their wild relatives

Globally, a total of 167, 890 sorghum accessions are reported held in different germplasm collection centres representing about 86% of the total 194,250 that has been documented in the Bioversity Germplasm Database (January 2006). Of the 167 890 accessions, 43, 104 sorghum accessions (25.7% of the documented) are held by the USDA-ARS-PGRCU of the USA.

ICRISAT, for which sorghum is one of its mandate crops, is one of the major repositories holding a total of 36,774 accessions (21.9%) from 91 countries (Reddy *et al.*, 2006). A preliminary survey indicated that the largest number of accessions (47, 963; 28.6%) was held by gene banks in Asia including China and India. An approximate total of 31, 200 sorghum accessions (16.1%) were held in Africa, and collectively gene banks in East Africa are holding larger collections of sorghum landraces than Southern and West Africa. Globally, a total of 1, 240 accessions of wild relatives of cultivated sorghum are maintained *ex-situ* in 19 centres. Of these accessions, the large proportion (37%) was being conserved at ICRISAT. This was followed by Australia – DPI and India-NBPGR that were maintaining 344 (28%) and 237 (19%) accessions respectively.

1.9 Traits of economic importance and sorghum breeding

Like in other domesticated crops, sorghum domestication involved artificial selection. The process resulted in gradual changes from the small-seeded, shattering open panicles to larger, non-shattering seeds and more compact panicles. Several characteristics of the sorghum plant were changed. These included the increase in the number of branches within the inflorescence; decreasing the internode length of the rachis; and an increase in seed size that significantly protruded from the glumes (House, 1985).

Sorghum is mainly grown for human consumption in Asia and Africa, but used for livestock feed in China, Australia and the Americas. In Africa and Asia, sorghum is mostly grown at subsistence scale. The crop's grain yield and quality are challenged by a broad spectrum of biotic and abiotic factors (Reddy Belem *et al.*, 2004). There has been a shift in the breeding goals for sorghum from wide adaptability to specific regional adaptation

(Reddy Belem *et al.*, 2004). Most of the agricultural research efforts were focused on the development of high yielding, resistant varieties, and adaptability to drought-prone environment in order to enhance both crop productivity and yield stability. Achievements in sorghum breeding in Africa have mainly been in the development and release of improved varieties based on high grain yield and resistance to diseases, insect pests and striga (Obilana, 2004).

ICRISAT working in collaboration with various national agricultural research institutes has adopted five different phases in its global sorghum breeding program (Reddy et al., 2004). The goals are phased as follows: (1) wide adaptability and high grain yield (1972–75); (2) wide adaptability and breeding for biotic and abiotic constraints (1976–79); (3) regional adaptations and resistance breeding (1980–84); (4) specific adaptations and resistance breeding (1985–89); and (5) trait-based breeding, sustainable productivity and upstream research (1990–2004).

1.10 CpDNA and ITS of nrDNA for phylogenetic inference

Chloroplast DNA (cpDNA), is the DNA located in the chloroplast organelles and is maternally-inherited in most angiosperms. The complete chloroplast genome sequence of sorghum has been published (Saski *et al.*, 2007), which facilitates the use of different regions for phylogenetic analyses of species in the genus *Sorghum, as* cpDNA is useful in providing information for inference of the evolutionary patterns and processes in plants (Raubeson & Jansen, 2005). The non-coding chloroplast regions are phylogenetically more informative than the coding regions at lower taxonomic levels because they are under less functional constraints and evolve more rapidly (Gielly & Taberlet, 1994).

The internal transcribed spacer (ITS) region of the 18S–5.8S–26S nuclear ribosomal DNA (nrDNA) has been commonly used for phylogenetic inference at the generic and infrageneric level in plants. Biparental inheritance, universality of primers, intragenomic uniformity and intergenomic variability merit the utility of ITS for phylogenetic reconstruction (Baldwin *et al.*, 1995). The two intergenic spacers, ITS1 and ITS2, evolve more rapidly than coding regions of nrDNA and have shown to be equally informative, being able to differentiate between closely related species (Baldwin, 1992) as exemplified in the genus *Sorghum* (Dillon *et al.*,

2001). The ITS has been used either solely (Sun *et al.*, 1994; Dillon *et al.*, 2001; Guo *et al.*, 2006) or in combination with regions of the cpDNA (Dillon *et al.*, 2004; Dillon *et al.*, 2007) for inferring phylogenetic relationships in the genus *Sorghum*.

1.11 Microsatellite and morphological markers for assessment of genetic diversity

Microsatellites, which are also called Simple Sequence Repeats (SSRs), which are mainly caused by polymerase slippage during DNA replication or slipped strand mispairing, represent tandem repeats with repeat motifs of 1-6 base pairs. The strengths of the markers include their high genomic abundance in eukaryotes, codominance inheritance, locus specificity, and multi-allelic character. SSRs are highly polymorphic in cultivated sorghum (*Sorghum bicolor* (L.) Moench) (Taramino *et al.*, 1997; Bhattramakki *et al.*, 2000; Kong *et al.*, 2000) except those located in relatively conserved coding regions (Schloss *et al.*, 2002). These markers have been widely used in the assessment of genetic diversity (Menz *et al.*, 2004; Manzelli *et al.*, 2007; Li *et al.*, 2010), population genetic structure and relatedness within or among sorghum landraces (Uptmoor *et al.*, 2003; Folkertsma *et al.*, 2005; Mutegi *et al.*, 2011) as well as in the construction of the framework architecture of a highly dense genetic map (Wu & Huang, 2006).

In most crops, analyses of morphological traits that inherit according to Mendelian genetic principles were the earliest methods for estimating genetic diversity (Doggett, 1988). The synthesis and categorization of morphological data into morphological and presumably genetic similarity groups is most useful when none is known about the population structure in a collection (Marshall & Brown, 1975). Phenotypic diversity index of morphological traits and/or multivariate analysis of qualitative and/or quantitative characters have been used to measure genetic relationships within sorghum e.g. (Ayana & Bekele, 1999; Abdi *et al.*, 2002; Geleta & Labuschagne, 2005).

2 Objectives

The purpose of this doctoral thesis was to study the phylogenetic relationships between the *Sorghum* species, the assessment of genetic diversity and nutritional value for its cultivated species for its conservation and sustainable utilization in breeding programmes. Specifically, the study sought to achieve the following objectives:

1. to resolve the phylogenetic relationships between species of the genus *Sorghum* based on cpDNA and ITS of nrDNA,
2. to estimate the genetic diversity and relationships among sorghum accessions from southern Africa,
3. to analyze the grain mineral contents of sorghum farmer and improved varieties from southern Africa and evaluate their potential for biofortification,
4. to assess the patterns of agromorphological variation among sorghum accessions from Malawi, Mozambique, Tanzania and Zambia.

3 Material and methods

3.1 Plant material

The germplasm used included wild *Sorghum* species and farmer and improved varieties of cultivated sorghum.

Forty eight accessions from 21 *Sorghum* species were used for the phylogenetic study (Paper I). These accessions included wild species belonging to sections *Stiposorghum* (20 accessions), *Heterosorghum* (2 accessions), *Para-sorghum* (8 accessions) and *Chaetosorghum* (1 accession). The rest of the materials used in this study were 17 accessions of *S. bicolor* (L.) subsp. *bicolor* and their close wild relatives of the section *Eu-sorghum*. Wild sorghum accessions were obtained from the Australian Tropical Crops & Forage Genetic Resource Centre (ATCFC), Biloela, Queensland, Australia while most of *S. bicolor* (L.) subsp. *bicolor* accessions were obtained from the Zambian National Plant Genetic Resources Centre.

Forty sorghum accessions including 27 accessions of farmer varieties and 13 accessions of improved varieties were used for macro and micronutrient analysis (Paper III). Farmer varieties were obtained from national gene banks of Malawi (7 accessions), Tanzania (5 accessions) and Zambia (15 accessions). Thirteen accessions of improved varieties were obtained from the sorghum and millet improvement program of Zambia.

Twenty seven accessions of *S. bicolor* (L.) subsp. *bicolor* obtained from the national gene bank of Zambia were used in the SSR based genetic diversity analyses of Zambian sorghum (paper II). These accessions were collected across regions inhabited by four ethnic groups of Zambia namely Chikunda, Soli, Tonga and Lozi. The other genetic diversity study using SSR markers

(Paper IV) involved fourteen sorghum accessions originally collected from Malawi (6 accessions), Tanzania (4 accessions) and Zambia (4 accessions) and obtained from the national gene banks of the respective countries.

In the assessment of agromorphological variation in *S. bicolor* (L.) subsp. *bicolor* from southern Africa, data was collected on 17 accessions (Paper V). The national gene banks of Malawi, Mozambique, Tanzania and Zambia provided five accessions each, of which one Mozambican and two Zambian failed to emerge in the field.

3.2 Methods

3.2.1 DNA extraction

Sorghum seedlings were raised in a greenhouse and fresh leaf tissues were sampled for DNA extraction at two weeks of age. Twelve plants per accession were used for DNA extraction. Individually sampled fresh leaf tissue was placed in eppendorf tubes, frozen in liquid nitrogen, freeze dried and ground into powder. DNA was extracted using a modified CTAB method (Bekele *et al.*, 2007).

3.2.2 Sequencing

Primers for amplification and sequencing of the *trnS-trnfM*, *trnY-psbM and trnT-trnD* intergenic spacers of cpDNA were designed during this study while the cpDNA *trnT-trnL* intergenic spacer was amplified and sequenced using the universal primers designed by Taberlet et al. (1991). Universal primers ITS4 and ITS5 (White *et al.*, 1990) were used for amplification and sequencing of the ITS-1, 5S and ITS-2 of nrDNA. All the necessary steps from amplification and cleaning of the amplified products to sequencing and subsequent processing of the sequences were described in detail in paper I.

3.2.3 SSR-PCR

The SSR primer screening for amplification, optimization of PCR conditions and detection of polymorphism lead to the selection of ten SSR primer pairs for genetic diversity analyses of both Zambian and SADC sorghum germplasm. The selected primers, their linkage groups and the amplification conditions were provided in Papers II and IV. Confirmed amplified PCR products were separated on readymade high resolution polyacrylamide gels. Allelic data for each locus was recorded as fragment size in comparison with

a standard 50 bp DNA ladder and also as binary data coded as 1 or 0 for the presence or absence for each allele.

3.2.4 Determination of macro- and micronutrients

About 50 g of grains of each accession was milled to flour using a laboratory mill (Yellow line, A10, IKA-Werke, Staufen, Germany). Following milling, samples were freeze dried to constant dry weight over a period of four days. About 0.5 g of each flour sample was digested as described in Hussain *et al.* (2010). The digested samples were analyzed for mineral contents at the ICP laboratory, Department of Ecology, Lund University, Sweden using Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES; Perkin-Elmer, OPTIMA 3000 DV). The atomic spectrometry standards from Perkin-Elmer, SPEX, AccuStandard and Merck were used during the analysis. The ICP-AES instrument was calibrated using a mixed multi-component standard at three contents within the factor of 50.

3.2.5 Protein determination

Sorghum grain samples were milled and freeze dried to constant weight prior to total nitrogen analysis. The samples were weighed (2-5 grams) using 5 x 9 mm tin capsules. Capsules containing samples were rolled into pellets. An aliquot was burned in an elemental analyzer (Nitrogen analyzer, NA 1500 series 2; Micromass, Carlo Erba Instruments, Rodano (Milan), Italy) at 1020°C and interfaced with an isotope ratio mass spectrometer (Optima; Micromass) leading to the release of CO_2, H_2O and N_2. CO_2 and H_2O are removed by passing the gasses over special absorbent columns. The nitrogen content was estimated in a column containing a thermal conductivity detector. Acetanilide (C_8H_9NO; C = 71.09%, H = 6.71%, N = 10.36%, O = 11.84%) was used as the standard reference material in this assay. A protein factor of 6.25, equivalent to 0.16 g nitrogen per gram of protein, was used to estimate protein content in sorghum as recommended by Merrill and Watt (1973).

3.2.6 Total starch determination

Approximately 50 mg of dry flour sample was weighed in duplicates and placed in a glass centrifuge tube (16 x 120 mm; 17 ml capacity). Starch content was determined using the total starch analysis protocol developed by AA/AMG; Megazyme International, Wicklow, Ireland. The enzymatic (α-amylase/amyloglucosidase) digestion (hydrolysis) of starch produces glucose which was spectrophotometrically quantified at 510 nm.

3.2.7 Agromorphological characterization

Sorghum accessions were sown in 2 row plots of 3 m long and 0.75 m wide and intra row spacing of 0.2 m, with two replications in a randomized complete block design at Mount Makulu Central Research Station in Zambia. Both qualitative (10) and quantitative (5) traits (Table 1) were recorded to estimate the levels of variation among the sorghum accessions. Measurements and observations were taken based on 10 randomly selected and tagged plants from each accession. The tagged plants were used for collection of all the data. The character states associated with the vegetative growth stage, panicle and grain were measured, observed and recorded based on sorghum descriptor of the International Board for Plant Genetic Resources (IBPGR & ICRISAT, 1993).

3.3 Data analysis

3.3.1 Phylogenetic analysis

The DNA sequence data was obtained from the University of Oslo, Norway (http://www.bio.uio.no/ABI-lab/). The quality of the sequences was visually inspected using Sequence Scanner version 1.0 (Applied Biosystems). Multiple sequence alignment was performed using ClustalX version 2.1.10 (Larkin *et al.*, 2007). The sequences were edited using BioEdit version 7.0.9 (Hall, 1999) and PAUP* 4.0 Beta 10 was used for phylogenetic analyses. The phylogenetic analyses of sequence data from the ITS and intergenic spacers of cpDNA were carried out both separately and combined. In both cases, indel positions were treated as missing data. *Zea mays* L. (GenBank accession number U04796) was used as an out-group species.

3.3.2 Genetic diversity analysis

Allelic size data for each SSR locus was used to estimate percentage of polymorphic loci, Shannon's information index (I), Nei's gene diversity (h), observed and expected heterozygosities using POPGENE version 1.31 (Yeh & Boyle, 1997). Genetic variation within and among accessions as well as among different groups of sorghum accessions was estimated through analysis of molecular variance (AMOVA) using the Arlequin 3.0 (Excoffier et al., 2005). Cluster analysis based on Unweighted Pair Group Method with Arithmetic Average (UPGMA) method within sequential agglomerative hierarchical nested (SAHN) and principal co-ordinate analysis were performed based on Nei's distance matrix using NTSYSpc (Rohlf 1998). To estimate of robustness of obtained trees, bootstrap values were obtained

through 1000 bootstrap resampling procedure using FreeTree – Freeware program (Pavlicek et al., 1999).

3.3.3 Micronutrient, macronutrients, protein and total starch analysis

Data for macronutrients, micronutrients, starch and protein contents were subjected to analysis of variance (ANOVA). Tukey's test was carried out for pairwise comparisons of means nutrient values of accessions using Minitab version 16.

3.3.4 Analysis of agromorphological data

Range, mean, standard error of mean and coefficient of variation were computed for five quantitative characters plant height, days to flowering, thousand seed weight, inflorescence length and width. Subsequently, these data were subjected to analysis of variance and Tukey's test. The qualitative data was analyzed for the frequency of a particular character state using InfoStat (Di Rienzo et al., InfoStat version 2010). A combined dataset of qualitative and quantitative characters was standardized and subjected to cluster analysis whereas the standardized quantitative dataset was subjected to principal component analysis. Minitab version 16 was used for ANOVA and multivariate analyses.

4 Summary of results and discussion

4.1 Phylogenetic relationships between *Sorghum* species (Paper I)

Sorghum species exhibited sequence length differences for the sequenced regions of the cpDNA and the ITS of nrDNA (Table 1). For example, the length of *trnY-psbM* intergenic spacer ranged from 1028 bp (*S. drummondii*) to 1053 bp (*S. exstans*). The sequences from this spacer of eight accessions of *S. bicolor* revealed 2–3 bp differences between them. The size of the *psbZ-trnG* spacer ranged between 286 bp (section *Eu-sorghum*) and 291 bp (*S. intrans*). The size difference between the two sections is attributed to the 5-bp indels within the spacer. Interestingly, this indels also exist in *S. laxiflorum* and *S. macrospermum*. Sequence length differences were also revealed in the *trn*T-*trn*L spacer, ranging from 684 bp (*S. arundinaceum*) to 693 bp (*S. leiocladum* and *S. laxiflorum*). Low sequence length differences of 2 bp in the trnT-*trn*L spacer were observed among the S. bicolor accessions. The *trn*T-trnL spacer from *S. bicolor*-12, -13 and -14 showed differences from the rest of the *S. bicolor* accessions arising from transitions and transversions at eight positions. The sequences obtained from the trnY-trnD spacer were between 318 bp (S. *amplum*, *S. angustum*) and 329 bp (*S. exstans*). ITS sequences had small length differences of 528-534 bp across the Sorghum species.

Of the cp*DNA* regions used in this study, the highest number of parsimony informative characters was obtained from *trnY-psbM* which was closely followed by *trnT-trnL* and *trnY-trnD* (Table 1). The *trnY-psbM*, *trnT-trnL* and *trnY-trnD* intergenic spacers were useful in the inference of phylogenetics at low taxonomic level. Similarly, *trn*T-*trn*L and *psb*M-*trn*D were identified as suitable for low taxonomic level phylogenetic studies (Shaw *et al.*, 2005).

Table 1. *Sequence characteristics and tree statistics of the cpDNA and ITS regions from maximum parsimony (MP) analysis.*

	cpDNA regions						
	*psb*Z-*trn*G	*trn*Y-*trn*D	*trn*Y-*psb*M	*trn*T-*trn*L	ITS	Combined cpDNA regions	Combined cpDNA and ITS
LAS	286-291	318-329	1028-1053	684-693	528-534	2316-2366	2844-3111
PICs[a]	8(3%)	12(4%)	32(4%)	19(3%)	69(13%)	71(3%)	140(4%)
TL	16	48	101	57	190	536	743
CI	0.9375	0.8958	0.6931	0.8947	0.8737	0.6250	0.6743
HI	0.0625	0.1048	0.31	0.1053	0.1263	0.3750	0.3257
RI	0.9846	0.9734	0.93	0.9757	0.9764	0.8463	0.8938
RC	0.9231	0.8720	0.6489	0.8730	0.8531	0.5252	0.6027

[a] Inclusive of the outgroup; LAS=Length of aligned sequences: PICs=Parsimony informative characters (number & percent); TL=Tree length; CI=Consistency index; HI=Homoplasy index; RI=Retention index; RC=Rescaling consistency index

Two lineages, A and J, are resolved by separate parsimony analysis of sequence data of four regions of cpDNA (Fig. 5a) and the combined sequence data from the cpDNA regions and the ITS (Fig. 5b). Lineage A contains all the *Eu-sorghum* species in clade B which include *S. bicolor* and their immediate wild relatives, *S. almum*, *S. halepense*, *S. drummondii* and *S. arundinaceum* with 100% bootstrap support. These results are consistent with earlier findings based on combined cpDNA and ITS sequences (Dillon *et al.*, 2004; Price *et al.*, 2005; Dillon *et al.*, 2007). Lineage J consists of all Australian wild *Sorghum* species except *S. laxiflorum* and *S. macrospermum* with equally high bootstrap support (Fig. 5). *S. laxiflorum* and *S. macrospermum* show closer relationship with the *Eu-sorghum* species with 100% bootstrap support than with other Australian wild *Sorghum* species. Earlier studies based on DNA sequence data also reported similar results (Sun *et al.*, 1994; Dillon *et al.*, 2007) which has led to the argument against placement of *S. laxiflorum* and *S. macrospermum* in separate taxonomical sections. In fact earlier taxonomic studies have also shown that *S. macrospermum* was more closely related to *Eu-sorghum* species than to any other *Sorghum* species (Garber, 1950; Celarier, 1958).

Within the *Eu-sorghum* section, clade D excludes *S. arundinaceum* from the rest of the species, but a subgroup comprising *S. halepense*-1, *S. drummondii, S. almum*, and *S. bicolor*-1, -2, -5 and -13 is formed as clade F with 99% bootstrap support (Fig. 5). The strongly supported (94 %) clade E consists of three accessions of *S. bicolor* (-3, -11 and -14). The *S. bicolor*

accessions in this clade originated from southern Africa, one from Zimbabwe (*S. bicolor*-3) and the other two from Zambia. *S. bicolor*-2, an accession from Yemen, seems to be distantly related to *S. bicolor* accessions from southern Africa but has a stronger association (clade H) with *S. almum* with strong bootstrap support. Earlier on, an allozyme variation study also revealed that *S. halepense* grouped closely with *S. bicolor* (Morden *et al.*, 1990), suggesting that the latter was most likely one of the parental species of *S. halepense*.

Stiposorghum and *Para-sorghum* form clade J with 100% bootstrap support. The internal nodes of this particular clade, however, in most cases had moderate level of bootstrap support. Most of the *Para-sorghum* and all of the *Stiposorghum* species form clade K with moderate bootstrap support. Clade M consists of *S. brachypodum* and *S. exstans* with 95% bootstrap support. *S. intrans* and *S. stipoideum*-1 form clade N, and *S. amplum* and *S. ecarinatum* form clade O but with only moderate bootstrap support (78 %).

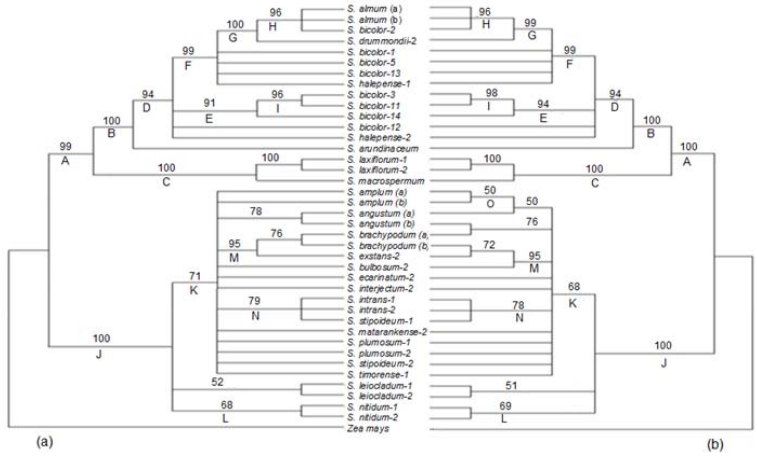

Figure 5. Maximum-parsimony 50% majority rule consensus trees generated using PAUP (1000 bootstrap replicates with 100 random additions; MaxTrees 100) (a) Sequence data from the four regions of cpDNA (b) Combined sequence data from four regions of cpDNA and ITS of the nrDNA of 21 Sorghum species. *Zea mays* was used as an out-group species. Indels are treated as missing data. Clades are indicated by letters below the branch. Bootstrap values of 50% and higher are indicated above the branches.

4.2 Genetic diversity in *Sorghum bicolor* (L.) Moench accessions (Papers II & IV)

The amplification of ten microsatellite loci in 324 individuals from 27 sorghum accessions from Zambia revealed that all loci were polymorphic with a total of 44 alleles. The number of alleles per locus ranged from 2 (*sb6-36*) to 9 (*sb5-236*) and averaged 4.4 alleles. Similar analysis of sorghum accessions from three SADC countries (Malawi, Tanzania and Zambia) revealed that all the ten SSR loci used were polymorphic generating 47 alleles in 14 accessions and the number of alleles per locus ranged from 3 (sbAGB03, Xcup05 and Xcup50) to 8 (Xcup02). One of the reasons for variation in number of alleles obtained per locus could be due to differences in number of mutations per locus and that mutation rates tend to be locus specific (Estoup *et al.*, 2002). Similar SSR polymorphic levels and ranges of fragment size of alleles were reported from earlier studies (Brown *et al.*, 1996; Ghebru *et al.*, 2002; Uptmoor *et al.*, 2003; Shehzad *et al.*, 2009). The observed heterozygosity (*H*oL) at each locus over 27 Zambian accessions ranged from 0.01 (*sb1-10*) to 0.09 (*Xtxp285*) with an average of 0.04 per locus. Three loci (*sb5-236*, *sb6-36* and *Xtxp285*) had *H*oL that was higher than the average.

The average number of alleles per locus, percent polymorphic loci, observed and expected heterozygosity and genetic variation measures estimated by Shannon's diversity index (I) and Nei's gene diversity (h) for the sorghum accessions in Zambia and at regional level are given in Table 2. Analysis of Zambian sorghum accessions revealed that accessions from agroecological region I exhibited moderately higher genetic variation (I = 1.30 and h = 0.64) than accessions from agroecological region II. Nei's gene diversity measure on these accessions showed clear differences between ethnic groups (Table 2). The highest Nei's gene diversity (h = 0.60) was found in the sorghum accessions collected from the area mainly inhabited by Lozi tribe. Sorghum accessions from the Tonga tribe dominated areas followed closely with Nei's gene diversity of 0.58. In this group, sorghum accessions from Sinazongwe exhibited higher Nei's gene diversity (h = 0.49) and Shannon diversity index (I = 0.89) than accessions from Kazungula (h = 0.34; I = 0.51) and Gwembe (h = 0.10; I = 0.14).

Table 2. *Number of alleles, mean percent polymorphic loci (%PL), mean observed heterozygosity (Ho) and gene diversity (He), mean Shannon diversity index (I) and Nei's gene diversity (h) for sorghum accessions (n) from Zambia and southern Africa partitioned as Malawi, Tanzania and Zambian.*

Origin	Group	Category	n	na	%PL	*Ho*	*He*	I	h
Zambia	Region	Region I	16	6.4	100	0.04	0.64	1.30	0.64
		Region II	11	5.3	100	0.03	0.63	1.26	0.63
	Ethnic group	Soli	7	3.4	90	0.03	0.49	0.89	0.49
		Chikunda	3	2.1	70	0.06	0.33	0.50	0.32
		Lozi	6	4.2	100	0.03	0.61	1.11	0.60
		Tonga	11	4.8	100	0.03	0.58	1.12	0.58
	District	Chongwe	7	3.4	90	0.03	0.49	0.89	0.49
		Luangwa	3	2.1	70	0.06	0.33	0.50	0.32
		Sesheke	4	2.8	90	0.02	0.45	0.74	0.44
		Kazungula	3	1.9	70	0.00	0.35	0.51	0.34
		Sinazongwe	6	3.7	100	0.06	0.49	0.89	0.49
		Shang'ombo	2	2.3	80	0.04	0.35	0.56	0.34
		Gwembe	2	1.2	20	0.00	0.10	0.14	0.10
Southern Africa	Country	Malawi	6	2.7	80	0.01	0.41	0.71	0.41
		Tanzania	4	2.2	70	0.00	0.25	0.42	0.25
		Zambia	4	2.3	90	0.00	0.42	0.64	0.42

4.2.1 Genetic variation within accessions

Analysis of data generated from sorghum accessions from three SADC countries revealed the highest Shannon diversity index (I = 0.71) in the Malawian accessions (Table 2). Zambian sorghum accessions closely followed with I = 0.64 and h = 0.42. Tanzanian sorghum exhibited the lowest Shannon diversity index and Nei's gene diversity (I = 0.42; h = 0.25). The amount of heterozygosity across loci, which is synonymous with allelic variation, indicates the amount of genetic variability which has a bearing on the survival of a species and allows organisms to adapt to changing environments provided that some loci have adaptive values.

Observed heterozygosity ranged from 0 to 0.06 with an average of 0.01 for the sorghum accessions from SADC countries, which was lower than the average observed heterozygosity (Ho = 0.04) obtained in the study involving 27 Zambian sorghum accessions and 10 SSR loci. The low level of observed heterozygosity is more attributable to the predominantly inbreeding nature of sorghum than samples sizes used. In fact, when a bottleneck occurs in a population, allelic diversity is reduced faster than is heterozygosity (Nei et al., 1975), which is a result of loss of rare alleles from

the population contributing little to the overall heterozygosity (Muraya et al., 2010). Sorghum is a self pollinating crop, although a wide range of outcrossing rates of 7–30% or higher have been reported (Dje *et al.*, 2004; Barnaud *et al.*, 2008). Outcrossing rates in sorghum are influenced by factors such as morphology of the inflorescence (Dje *et al.*, 2004), floral traits (Barnaud *et al.*, 2008) and environment (Abdel-Ghani *et al.*, 2004). Loose panicles such as of race guinea favour outcrossing, whereas the architecture of compact panicles e.g. of durra race impedes outcrossing (Dje *et al.*, 2004). Cleistogamy (flowers remain enclosed) in sorghum due to very long glumes prevent pollen movement and thus strongly promoting selfing (Barnaud *et al.*, 2008).

The predominantly selfing nature of the species explains the observed lower genetic variation within than among accessions in this study as revealed by AMOVA (Table 3). Similar results were reported from a study involving sorghum accession originated from Somalia (Manzelli et al., 2007). Breeding systems of plant species are reported to have a significant impact on population variability with self pollinating species being the least diverse and exhibiting higher between population than within population variation (Nybom & Bartish, 2000). In fact, according to Hamrick and Godt (1996), the breeding system is one of the strongest predictors of within population genetic diversity. Low levels of genetic variation among self pollinated plants is attributed to limited movement of genes via pollen, which also results in greater differentiation among populations (Hamrick, 1983).

4.2.2 Genetic variation among accessions

Analysis of molecular variance (AMOVA) of SSR data revealed a significant genetic variation both among and within accessions studied both for Zambian (82% and 18%, respectively) and SADC accessions (83% and 15%, respectively; $P < 0.001$; Table 3). Observed genetic variation in the sorghum accessions both at national (Zambia) and regional (SADC) levels was higher among the accessions than within accessions. When the 27 Zambian accessions were grouped based on agroecological regions, AMOVA revealed a significant genetic variation between the groups (12%; $P < 0.001$; Table 3). Furthermore, AMOVA on sorghum accessions grouped based on the four ethnic groups (Soli, Chikunda, Lozi and Tonga) associated with the collection sites revealed a significant genetic variation among groups (23%; $P < 0.001$). AMOVA of SSR data from the regional sorghum accessions grouped based on country of origin revealed a highly significant genetic

variation among countries (42%; $P < 0.001$). Similarly, AMOVA on regional sorghum accessions grouped based on altitude of collection sites revealed a significant genetic variation among groups (12%; $P < 0.05$). Earlier genetic diversity studies involving microsatellites on cultivated sorghum also revealed a higher genetic diversity among than within accessions. For instance, Ghebru *et al.* (2002) observed significant genetic variation among 28 Eritrean accessions for all measured variance components in which differences among accessions accounted for 50.4% of the variation while within accession diversity accounted for 49.6%. Similarly, Dje *et al.* (2000) reported overall gene diversity (H_T) of 0.9 in 25 accessions from the world germplasm collection, with differentiation among accessions (G_{ST} = 67%) accounting for two thirds of the diversity. On the contrary, the study involving nine Somali sorghum accessions using five SSR loci, Manzelli *et al.* (2007) observed that most of the genetic diversity (*Hs*) resided within accessions relative to the genetic differentiation between accessions (G_{ST}), demonstrating that the accessions are not under artificial selection processes and/or there is a continuous gene flow in form of seeds among accessions. A number of factors such as agronomic, economic and cultural in the traditional farming systems have been reported to impact on levels of genetic diversity in sorghum (Chakauya *et al.*, 2006; Mutegi *et al.*, 2011). Following the plant domestication stage, artificial selection has been identified as one of the factors contributing to the reduced genetic diversity of crops (Gepts, 2004). In most cases, traditional farmers maintain more than one distinct local variety selected for particular characteristics of interest to them and specific use. These landraces are perpetuated as farmer varieties from generation to generation. The driving forces behind the practice of maintenance of two or more sorghum landraces per household are twofold. Local farmers select landraces that could cope with local environmental factors such as duration of rainy season. Therefore, early maturing local varieties are usually planted by most households to provide food early in the season and thus ensuring attainment of household food security.

Table 3. *AMOVA based on SSR data for sorghum accessions from Zambia, Malawi and Tanzania: (A) without grouping the accessions, (B) grouping by geographical location, (C) grouping the accessions according to two altitude levels and (D) grouping the accessions based on the ethnic group of collection sites.*

Origin of germplasm	Groups	Source of variation	df	Variance	%variation
Zambia	(A) ungrouped	AA	26	Va = 2.26	82.44***
		WA	621	Vb = 0.48	17.56***
	(B) regions	AG	1	Va = 0.36	12.45***
		AAWG	25	Vb = 2.07	71.08***
		WA	621	Vc = 0.48	16.47***
	(C) altitudes	AG	1	Va = -0.04	-1.39ns
		AAWG	25	Vb = 2.28	83.71***
		WA	621	Vc = 0.48	17.68***
	(D) ethnicity	AG	3	Va = 0.67	22.91***
		AAWG	23	Vb = 1.75	60.60***
		WA	621	Vc = 0.48	16.49***
Malawi, Tanzania, Zambia	(A) ungrouped	AA	13	Va = 2.75	83.36***
		WA	126	Vb = 0.50	15.13***
	(B) countries	AG	2	Va = 1.58	41.90***
		AAWG	11	Vb = 1.64	43.52***
		WA	126	Vc = 0.50	13.25***
	(C) altitudes	AG	1	Va = 0.41	11.77*
		AAWG	12	Vb = 2.53	72.49***
		WA	126	Vc = 0.50	14.30***

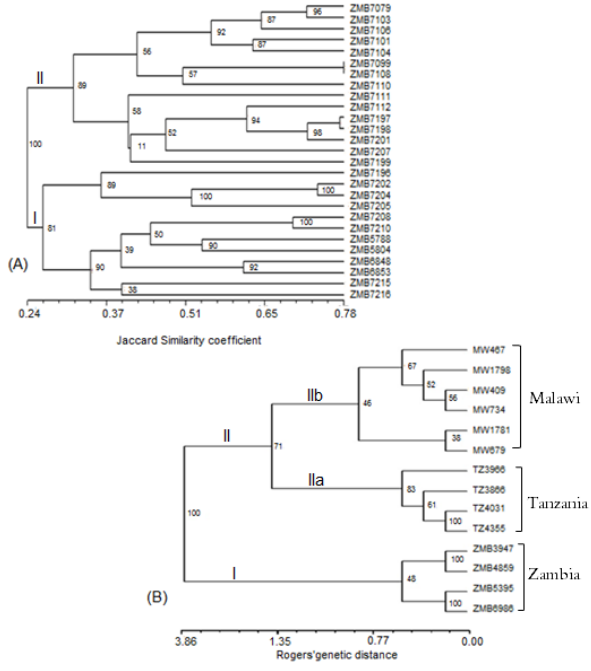

Figure 6. UPGMA cluster analysis based dendrograms generated based on SSR data (A) for Zambian sorghum accessions and (B) for sorghum accessions from Malawi, Tanzania and Zambia. Bootstrap values from 1000 resamplings are in between two branches.

Cluster analysis of 27 accessions from Zambia based on the Jaccard similarity coefficient (Fig. 6A) revealed two major clusters of sorghum accessions, I and II, with substantial bootstrap support. Internal grouping of accessions mainly put sorghum accessions in similarity clusters according to location of collection sites. Cluster I consists of 12 accessions from Sinazongwe, Gwembe, Shangombo and Sesheke. Cluster II was composed of 15 sorghum accessions mainly from Chongwe and Luangwa. Other accessions in the same cluster were from Sesheke, Kazungula and Sinazongwe (Fig. 6A). In most cases, sorghum accessions collected from the same location showed higher genetic similarity than accessions collected from different sites. ZMB7099 and ZMB7108 collected from Chongwe were revealed to be the most similar pair of accessions. Similarly, ZMB7197 and ZMB7198 from Sesheke, ZMB7202 and ZMB7204 from Kazungula, ZMB7208 and ZMB7210 from Sinazongwe and ZMB6848 and ZMB6853 from Gwembe are highly similar (Fig. 6A). However, there seem to be less similarity between ZMB7196 from Sesheke and the three other accessions, ZMB7197, ZMB7198 and ZMB7199, collected from the same location. Similarly, ZMB7201

from Kazungula showed clustering pattern that indicated that it is genetically distant from ZMB7202 and ZMB7204, also of the same site.

Cluster analysis of microsatellite data from sorghum accessions from SADC region involving Malawi, Tanzania and Zambia also revealed a dendrogram with two major clusters (clusters I and II) that were strongly bootstrap supported (Figure 6B). Cluster I consisted of only accessions from Zambia. The Zambian accessions were genetically less similar to the accessions in cluster II from Tanzania (sub cluster IIa) and Malawi accessions (sub cluster IIb). Ghebru *et al.* (2002) also reported both distinct and mixed clusters among 60 sorghum accessions according to geographic origin. However, Dje *et al.* (2000) reported that accessions belonging to the same race or geographic origin were widely scattered when a matrix plot of individual sorghum accessions based on R_{ST} distance matrix was used. A recent study based on SSR data reported a clustering pattern of sorghum germplasm that was according to geographic origin (Geleta *et al.*, 2006). Similarly, genetic distance data from polymorphic loci in the present study clustered SADC accessions according to country of origin (Figure 6B). It is also evident that sorghum accessions from Zambia were genetically distant from accessions obtained from Malawi and Tanzania. The close grouping of accessions from Malawi and Tanzania may suggest the presence of higher levels of gene flow between sorghum populations of the two countries than between these countries and Zambia. Gene flow encompasses several mechanisms of gene exchange among populations, including movement of gametes, seed, individuals or groups of individuals from one place to another (Slatkin, 1987). Seed exchange practices between communities could be the main factor for the observed similarities among sorghum accessions originating from different geographical regions.

4.3 Characterization of the SADC sorghum varieties for micronutrients, (Papers III & IV)

Highly significant differences ($p < 0.0001$) were revealed among the sorghum accessions in terms of their grain mineral, protein and starch contents (Table 4). Improved sorghum varieties exhibited higher profiles for grain macronutrient contents than farmer varieties. Among the improved varieties, ELT-1-17, Macia, MMSH-1040, MMSH-1257, MMSH-1324, MMSH-1365 and MMSH-740 ranked high in grain K, Mg and P contents whereas Kuyuma and Sima, which are widely grown improved sorghum varieties in

Zambia, had the lowest grain K, Mg and P contents. Farmer varieties were superior to improved sorghum varieties in grain Zn, Cu and Mn contents. The highest grain Fe content was obtained from ZMB5788 (8.03 mg/100 g) followed by MW734 (6.33 mg/100 g). The highest grain Zn content (5.51 mg/100 g) was obtained from TZ4031 which was followed by TZ3966, MW734 and ZMB7111 with 4.51 mg/100 g, 3.89 mg/100 g and 3.89 mg/100 g, respectively. Grain-Fe and Zn contents of 4.6 mg/100 g and 3.7 mg/100 g, respectively, were reported in an earlier study involving 29 sorghum accessions from the ICRISAT core collection (Ashok Kumar *et al.*, 2009). Grain Fe and Zn contents obtained in the present study were comparable with those reported in 76 Benin sorghum accessions (Kayode *et al.*, 2006b; Kayode *et al.*, 2006a). This study revealed that accessions MW734, TZ3966 and TZ4031 have potential for use in a micronutrient enrichment programme as they exhibited higher grain Fe and Zn contents than the suggested minimum of 5 mg/100 g and 3.7 mg/100 g respectively (Ashok Kumar *et al.*, 2009).

A significant correlation was found between grain colour and mineral contents for Ca, K, and S (Table 4). Sorghum varieties with brown grain exhibited significantly higher grain content for Ca (17.9 mg/100 g) and K (446.7 mg/100 g) content than varieties with white grains (Table 4). However, white sorghum varieties exhibited significantly higher grain S content (159.2 mg/100 g) than varieties with brown grains. No significant differences were observed between sorghum accessions with brown and white grains for grain Fe and Zn supporting the report of Kayode *et al.* (2006b). The results imply that whereas high grain macronutrients of Ca and K showed close association with brown grained varieties, selection for high Fe or Zn grain content may not automatically translate into selection for a desired grain colour. However, further studies involving increased number of accessions of different seed colour should be conducted to establish the suggested association.

Analysis of variance for grain mineral contents of sorghum farmer varieties grouped according to country of origin, demonstrated that farmer varieties showed significant differences in grain mineral content for Ca, K, P, S, Zn and Fe (Table 4). Tanzanian sorghum accessions showed significantly higher Ca, K, P, Zn and Fe grain contents than Malawian and Zambian accessions. The study demonstrated a significant variation in grain mineral content among the sorghum farmer varieties. However, grain mineral contents are influenced by genotype, environment and probably by

genotype x environment interactions (House, 1999; Zhang *et al.*, 2010). For example, genetic as well as environmental factors have been shown to significantly affect Fe and Zn levels in maize and wheat (Graham *et al.*, 1999; Bänzinger & Long, 2000). Therefore, part of the significant variation among farmer varieties in relation to grain mineral contents obtained in this study is due to genotypic variation and suggests an existing genetic potential in the studied genetic material for the improvement of sorghum varieties in macro and micronutrients.

Table 4. *Mean comparison of grain mineral content (mg/100 g) in sorghum accessions based on variety type, grain colour (brown and white) and country of origin. Only farmer varieties were used for mean mineral content comparisons based on country of origin.*

Group	Ca	Mg	K	P	S	Zn	Cu	Fe	Mn
All 40 accessions	15.4	163.9	436.9	343.8	149.4	2.97	0.5	4.1	2.1
(P < 0.0001)									
Level of genetic improvement									
Improved variety (n=13)	15.0a	179**b**	525**b**	387**b**	164**b**	2.7**a**	0.4**a**	3.9a	1.5**a**
Farmer variety (n=27)	15.6a	157**a**	395**a**	323**a**	142**a**	3.1**b**	0.5**b**	4.2a	2.3**b**
Grain colour									
Brown (n= 19)	17.9**b**	161.7a	446.7**b**	343.2a	138.5**a**	3.1a	0.5a	4.2a	2.2a
White (n=21)	13.1**a**	165.9a	428.1**a**	344.3a	159.2**b**	2.8a	0.5a	4.1a	2.0a
Origin of farmer varieties									
Malawi (n=7)	11**a**	153.9a	350.9**a**	294.8a	158.0**b**	2.8a	0.4a	3.9a	2.1a
Tanzania (n=5)	20**c**	163.1a	453.2**c**	360.1**b**	134.8a	3.8**b**	0.5a	5,0**b**	2.1a
Zambia(n=15)	16**b**	155.6a	395.6**b**	323.5**a**	137.3**a**	2.9**a**	0.5a	4.0**a**	2.5a

Means with the same letters within a column and in each category do not differ significantly (Tukey Test at $p \leq 0.05$).

Pearson correlation of grain protein, starch, macro and micro nutrients and one thousand seed weight revealed a significant positive correlation between grain Fe and Zn, Zn and P, K and Mg, K and P, K and S, Mg and P, Mg and S, S and P (Table 5). Grain Ca, K, Mg, P, S and Zn showed negative correlations with thousand seed weight, r = -0.44, -0.56, -0.58, -0.52, -0.29 and -0.32, respectively. The significant positive correlations between grain-Fe and Zn also have been previously reported (Reddy Belum *et al.*, 2005; Kayode *et al.*, 2006b). The observed positive correlation between grain Fe and Zn content implies that there are possibilities of combining the selection

for both minerals in a single agronomic background. Protein exhibited a positive correlation with Fe, Zn, K, Mg, P and S. Total starch showed negative correlation with all the minerals studied except grain Na and thousand seed weight. The negative correlation between thousand seed weight and grain micronutrients such as Fe and Zn could be attributed to the dilution effect caused by enhanced grain starch content on mineral contents in larger seeds (Bänzinger & Long, 2000). In order to realize the desired impact of micronutrient-dense improved cultivars in human nutrition, the micronutrients must be delivered in sorghum varieties that also meet the farmer-preferred traits such as early maturity, grain size and colour.

Table 5. *Estimates of correlation coefficients between mineral elements and thousand seed weight.*

	TSW	Protein	Starch	Fe	Zn	Ca	K	Mg	Na	P
Protein	-0.34**									
Starch	0.52***	-0.41***								
Fe	-0.20ns	0.23*	-0.28*							
Zn	-0.32**	0.44***	-0.18 ns	0.46***						
Ca	-0.44***	-0.13ns	-0.19ns	0.12 ns	0.11 ns					
K	-0.56***	0.23*	-0.39***	0.12 ns	0.10 ns	0.39***				
Mg	-0.58***	0.51***	-0.45***	0.11 ns	0.17 ns	0.14 ns	0.67***			
Na	0.34**	-0.10 ns	0.39***	-0.10 ns	-0.18 ns	-0.06 ns	-0.06 ns	0.03 ns		
P	-0.52***	0.38***	-0.40***	0.13 ns	0.26*	0.15 ns	0.77***	0.88***	-0.02 ns	
S	-0.29**	0.49***	-0.38***	0.01 ns	0.10 ns	-0.15 ns	0.42***	0.66***	0.12 ns	0.53***

*, **, *** = Significant correlation at P < 0.05, 0.01, 0.001 (2 tailed); ns = non-significant

4.4 Agromorphological trait variation in SADC sorghum accessions

4.4.1 Qualitative characters

A total of 17 sorghum accessions were used in the analyses. Of these accessions, 88% had non-juicy stalks and had white leaf midrib colour. Only two accessions, ZMB6731 and ZMB7016, exhibited juicy stalks and dull green leaf midrib colour. Wide variation among accessions was observed for shape and compactness of the inflorescence (Fig. 7). Twelve percent of the accessions (MW1798, ZMB7198), exhibited compact inflorescence. On the other extreme of the character states, two Mozambique accessions (MZ1537, MZ1553) had lax inflorescence, which was typical of wild sorghum. In between was TZ4148 with semi compact elliptic inflorescence. Other accessions exhibited inflorescence with semi loose drooping primary

branches (TZ3616, TZ3943, ZMB6731), half bloom corn (MZ1542, TZ3835, TZ3864), very loose erect primary branches (ZMB7016), very loose drooping primary branches (MW409, MZ1574), loose erect primary branches (MW1819) and loose drooping primary branches (MW1788, MW619). A wide distribution of character classes for inflorescence compactness and shape indicates the existence of different races and/or intermediate races. The domination of compact elliptic inflorescence types was reported among 45 sorghum accessions from the eastern highlands of Ethiopia (Geleta & Labuschagne, 2005). This perhaps explains the dominance of certain sorghum races in particular sub-regions of Africa. Kafir is the dominant race in southern Africa and occurs as a pure race and as intermediate hybrids mainly from crosses with race guinea (de Alencar Figueiredo *et al.*, 2008). The race durra, characterized by a compact oval inflorescence, is a dominant race in the northeast Africa (Stemler *et al.*, 1977; Doggett, 1988; Ayana & Bekele, 1999).

In the case of grain covering, 58% of the accessions had panicles with fully covered grains. TZ3943, MW1798, TZ4148 and ZMB7016 had 50% or more exposed grains. The rest of the accessions, representing 18%, exhibited three quarters of the grains covered. Sorghum accessions that exhibited full grain covering also showed little or no grain shattering, a trait associated with grain loss in the field following physiological maturity. This group includes all Mozambican accessions, three Malawian accessions, MW1788, MM409 and MW619 and three Tanzanian accessions, TZ3616, TZ3835, TZ3864. The dominant character states for grain covering were either full grain coverage or that the glumes protruded over the length of the grain. On the contrary, Geleta & Labuschagne (2005) reported that the distribution of 25% grain covering character was dominant in Ethiopian sorghum accessions. The pattern of the shattering character followed a similar pattern as that of grain covering.

The distribution of grain colour among sorghum accessions was skewed with accessions with white grain (47%) dominating. Other accessions were yellow (12%), red (12%) brown (18%) and buff or yellow-brown (12%). All Mozambican accessions in the study had white grains. Two of the Malawian accessions had white grains and the other three showed yellow, brown and buff coloured grains. Ayana & Bekele (1998) reported dominance of both white and brown grained sorghum but that landraces with other grain colours were possible depending on the region within the country. Grain colour in sorghum is an important character as it relates to end use of a

particular variety. Human selection was reported as another important factor influencing the domination of certain landraces with particular grain colour (Ayana & Bekele, 1998).

Figure 7. Genetic diversity in sorghum accessions based on inflorescence shape and exsertion. 1: bloom type; 2: semi compact with recurved pudencle; 3: upright semi compact; 4: upright narrow; 5: semi compact; 6: upright compact.

4.4.2 Quantitative characters

ANOVA of the quantitative characters revealed significant variation among sorghum accessions for all the traits analyzed ($p \leq 0.001$; Table 6). Plant height ranged from 172.5 to 407 cm with an average of 307.4 cm. Sorghum accessions that exhibited dwarf characteristic included MW1798, MW409, ZMB7016 and ZMB7198 and accessions with tall plants were MW1819, MW619, MZ1537, MZ1553 and TZ3835. Sorghum accessions exhibited a moderate range of 88-125 days to flowering. Tukey comparison of means of days to flowering, grouped sorghum accessions into three flowering groups namely early, medium and late flowering. Early flowering plants within accessions were observed in MW409, ZMB7198, TZ3943, ZMB7016 and MW619. Mozambican accessions, MZ1542, MZ1574, MZ1537 and Tanzanian accessions TZ3864, TZ3616, TZ3875 comparatively exhibited the longest vegetative growth stage. One Mozambican sorghum accession, MZ1574, exhibited the longest inflorescence (49 cm) and a Zambian accession, ZMB7198 had the shortest inflorescence (14.6 cm). Thousand seed weight, which is indicative

of seed size, ranged from 3 - 29 g with a mean of 18 g. MW619 had the highest thousand seed weight of all the accessions used.

Table 6. *Sorghum accessions used in the study, location of collection sites and average plant height, days to flower (DTF), thousand seed weight (TSW), inflorescence length (Infl. length), inflorescence width (Infl. width) data and analyzed ranges, means, standard errors of mean and coefficients of variation.*

Accession	Country	Nearest town	Plant height (cm)	DTF (days)	TSW (g)	Infl. Length (cm)	Infl. Width (cm)
MW1788	Malawi	Msangu	332.3cd	109.2b	28b	44.3b	14.1a
MW1798	Malawi	Chitala	172.5h	91.2d	22d	21.1g	13.5ab
MW1819	Malawi	Mwalawanjuchi	405.0a	105.0bc	26c	36.0d	9.7f
MW409	Malawi	Therere	184.5gh	88.6d	27b	26.3f	13.0b
MW619	Malawi	Thyolo	357.5abcd	102.7c	29a	36.4d	11.0cd
MZ1537	Mozambique	Mogovolas	402.3a	119.4a	10i	41.5c	10.9cd
MZ1542	Mozambique	Murrupula	310.5de	124.6a	7k	35.6d	11.3cd
MZ1553	Mozambique	Malema	389.8ab	106.6bc	21e	36.7d	13.6ab
MZ1574	Mozambique	Alua	371.8abc	121.1a	11h	48.6a	9.8ef
TZ3616	Tanzania	Mtwara	323.0cd	122.1a	9j	37.6d	11.9c
TZ3835	Tanzania	Newala	396.3a	121.3a	3l	37.4d	10.4def
TZ3864	Tanzania	Nachingea	339.0bcd	124.9a	9ij	37.5d	11.2cd
TZ3943	Tanzania	Mugumu	255.8ef	89.1d	26c	17.1h	10.7de
TZ4148	Tanzania	Biharamulo	303.8de	106.7bc	16g	20.1g	8.2g
ZMB6731	Zambia	Lufwanyama	235.3fg	107.1bc	18f	19.9g	10.9cd
ZMB7016	Zambia	Chama	224.8fgh	89.4d	20e	30.4e	7.3g
ZMB7198	Zambia	Sesheke	216.3fgh	89.0d	21e	14.7h	7.3g
Range			172-405	88-125	3-29	14.6-49	7.2-14.1
Mean			307.4***	106.9***	18***	31.8***	10.9***
SE Mean			13.2	2.32	1.4	1.71	0.35
CV			25.1	12.6	45.8	31.4	18.6

Means with the same letters within a column (trait) do not differ significantly (Tukey Test at $p \leq 0.05$; ***Significant at $p \leq 0.001$

Cluster and principal component analyses generated a dendrogram and a matrix plot, respectively, that separated Zambian sorghum accessions from other accessions used in the study (Fig. 8 & 9). The dendrogram had two main clusters, I and II (Fig. 8). Cluster I consisted of sorghum accessions from Malawi (MW1798), Tanzania (TZ4148, TZ3943) and all Zambian accessions (ZMB7198, ZM7016, ZMB6731; Fig.8). Cluster II consisted of a larger group which included Tanzanian accessions (TZ3616, TZ3864, TZ3835), Malawi (MW409, MW619, MW1819, MW1788) and all Mozambican (MZ1542,

MZ1574, MZ1537, MZ1553) accessions. It is interesting to note that all the Zambian and Mozambican accessions were completely placed under different clusters. Dissimilarity exhibited between Mozambican and Zambian accessions were mainly attributed to differences in plant height, days to flowering and inflorescence length.

The two clusters further generated four sub clusters A, B, C and D (Fig. 8). Two Zambian accessions (ZMB7016, ZMB6731) formed a similarity sub-cluster A. In sub-cluster B, two Tanzanian accessions (TZ3943, TZ4148) grouped with ZMB7198 and MW1798. In cluster II, MW409 was less similar to the rest of the accessions (Fig. 8). Seed exchange among farming communities especially along the common borders of the three countries may be one of the major factors contributing to the observed clustering pattern. Seed exchange among the local farmers was also reported as an important source of sorghum seeds for planting as the case for cucurbits in Zambia (Gwanama & Nichterlein, 1995). The possible reason for the dissimilarity between Mozambican and Zambian accessions could be that since Mozambican accessions were collected from the Nampula region, Malawi seems to provide a physical barrier to direct gene flow through seed exchange between farming communities of Mozambique in that region and Zambia. If there is traditional seed exchange between communities of the two countries, most probably the sorghum varieties from either Zambia or Mozambique are grown and subjected to farmer variety selection in Malawi before farming communities in either countries gain access to them.

Principal component analysis of the quantitative data revealed that the first two principal components with eigenvalues > 1.0 and cumulatively accounted for 81.4% of the total variation (Table 7). The cumulative proportion of the variation reached 94.9% in the first three PC axes. The first PC axes has variance of 2.9 and accounted for 57.5% of the total variation while the second and third principal components account for 24% and 13.5% of the total variance, respectively. The variation in PC1 was positively associated with plant height, days to flowering, inflorescence length and width (Table 7). The second axis was positively and moderately associated days to flowering. A third principal component was moderately associated with 100 seed weight and inflorescence length. Taking eigenvalues > 1 as significant and component loading greater than $\pm$ 0.30 as meaningful (Hair *et al.*, 1998) and that a high coefficient for a trait indicates the relatedness of that trait to the respective PC (Sneath & Sokal, 1973), only the first two components were significant in this study. The high degree of

variation in the first two PC axes indicated a high degree of variation for these characters.

Table 7. *Proportion and cumulative variances and Eigen-vectors on five principal components (PC) based on quantitative traits.*

	PC1	PC2	PC3	PC4	PC5
Eigenvalue	2.873	1.198	0.674	0.180	0.074
Proportion	0.575	0.240	0.135	0.036	0.015
Cumulative	0.575	0.814	0.949	0.985	1
PH	0.502	-0.060	0.540	-0.617	0.268
DTF	0.559	0.094	-0.262	-0.119	-0.772
100SW	-0.413	-0.487	0.548	-0.043	-0.538
Infl. length	0.506	-0.296	0.292	0.747	0.116
Infl. width	0.092	-0.814	-0.504	-0.214	0.170

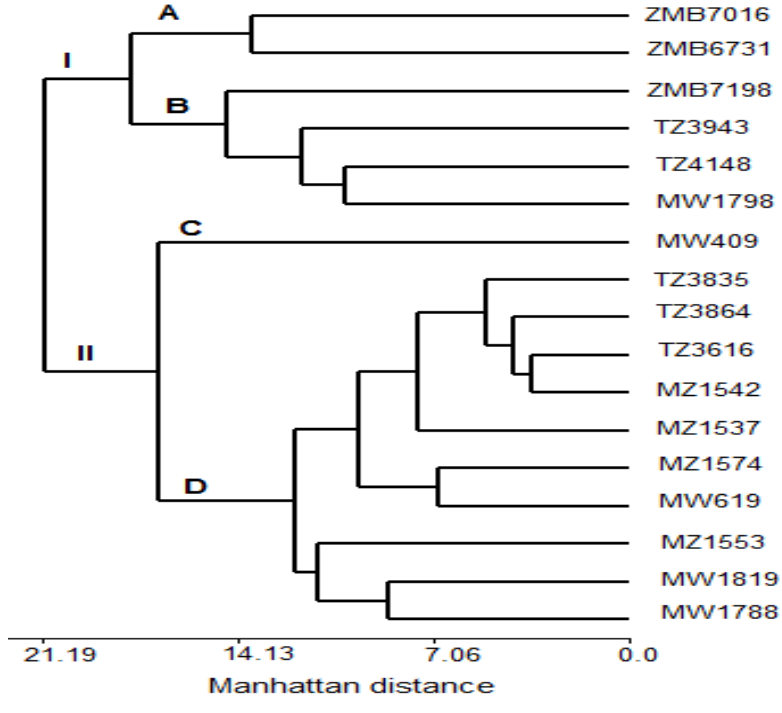

Figure 8. Manhattan distance and average linkage clustering of sorghum accessions from four countries in southern Africa.

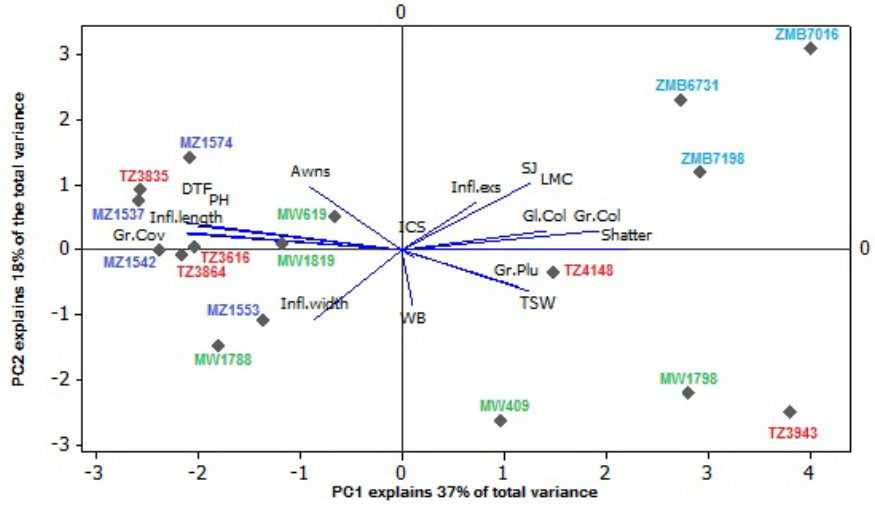

Figure 9. A two-dimension plot of sorghum traits and accessions from Malawi (green), Mozambique (purple), Tanzania (red) and Zambia (blue). The traits used are days to flower (DTF), plant height (PH), inflorescence length (Infl. length), grain covering (Gr. Cov), inflorescence width (Infl. width), wax bloom (WB), inflorescence compactness and shape (ICS), grain plumpness (Gr. Plu), hundred seed weight (100SW), shattering (Shatter), glume colour (Gl. Col), grain colour (Gr. Col), leaf midrib colour (LMC), stalk juiciness (SJ) and inflorescence exsertion (Infl. exs).

The plot pattern of ZMB6731, ZMB7016 and ZMB7198 in the two-dimension plot (Fig. 9), is explained by similarity of these accessions for grain shattering, inflorescence exsertion, glume colour, grain colour and stalk juiciness. Typically, the observed variation among sorghum accessions was mainly as a result of the high loading for plant height, days to flower and inflorescence length. In addition to the three characters identified in this study, Bucheyeki *et al.* (2009) also reported high loading of grain number per panicle, weight of five panicles and inflorescence width. The differences in number of characters with high loading could be attributed to differences in sample size used in the studies.

5 Conclusions

1. The close phylogenetic relationship between *S. macrospermum* and *S. laxiflorum* suggest their merger into one taxonomic section.

2. The *S. macrospermum* and *S. laxiflorum* showed more close relationship with *Eu-sorghum* species than to any other Australian wild *Sorghum* species and hence provide opportunities of their utilization for improvement of cultivated sorghum varieties.

3. *S. almum* is more closely related to *S. bicolor* than to *S. halepense*, one of its known parents. As the chloroplast DNA is maternally inherited, *S. bicolor* is the most probable maternal parent of *S. almum*.

4. The *S. bicolor* accessions from southern Africa form a distinct, well-supported clade separated from the accession originally from Yemen, indicating the geographical divergence of cultivated sorghum due to separation of the two sorghum gene pools probably in the early stage of the domestication of the crop.

5. A significant genetic diversity exists in the sorghum accessions from Malawi, Mozambique, Tanzania and Zambia. Sorghum accessions that exhibited high gene diversity and those that represent different clusters/sub-clusters should be prioritized for core collection irrespective of their geographic origin and would be useful for breeding and conservation.

6. A significant variation among the sorghum varieties from southern Africa was evident for total protein, starch, macro- and micronutrients. Superior sorghum accessions have a potential of contributing the best

starting genetic material for breeding improved varieties for desirable levels of micronutrient.

7. The significant positive correlation between protein and micronutrients exhibited by sorghum accessions suggest that the possibility of simultaneous increased content of two or more nutrients delivered in a single popular cultivar.

6 Recommendations and future prospects

1. Success in breeding is dependent on the availability of sufficient genetic variation in the original genetic material. The phylogenetic study has revealed that *S. laxiflorum* and *S. macrospermum,* wild *Sorghum* species from the tertiary genepool, are closely related with *S. bicolor*. The close relationships exhibited by cultivated sorghum with *S. macrospermum* and *S. laxiflorum* provide useful information for possible exploitation of these wild genetic resources for useful traits such as pest and disease resistance that affect production of cultivated sorghum. In order to tap into this potential genetic resource, it is recommended that further studies on crossability and hybridization of these species with *S. bicolor* are undertaken.

2. This study assessed the level and pattern of genetic diversity in cultivated sorghum from four countries in southern Africa, generating information of importance for conservation and breeding. A significant variation between sorghum accessions and their genetic differentiation based on agroecological regions and ethnicity has implication for its conservation and sustainable utilization. Given that financial resources for plant genetic resources are limited at both national and regional levels, when prioritizing accessions for conservation, there is need to conserve as broad the diversity as possible from all agroecological regions taking into account unique genetic variants. Consideration should be given to representative sorghum accessions of different clusters and sub-clusters, agroecological regions and ethnic background of the genetic resources.

3. A significant variation among the sorghum varieties was evident for both macro- and micronutrients. Identification of sorghum accessions that

have a potential of attaining desirable levels of macro and micronutrients from southern African countries is promising. Improved sorghum varieties ELT-1-17, MMSH-1040, MMSH-1257 and MMSH-1324 were more superior for grain macronutrient contents than farmer varieties and ranked highly for grain K, Mg, S and P contents. Farmer varieties, ZMB5788, MW734 and TZ4255 showed superiority for grain Fe content while TZ4031, TZ3966 and ZMB7111 were high in grain Zn content. Superior sorghum accessions identified in this study should further be studied both under the same and different environmental conditions in multiple locations. This is necessary not only for the evaluation of the heritability of the traits and the maximum potential of the accessions but also to determine suitable environmental conditions under which these desirable levels can be attained taking into account farmers' preferred varieties for delivery of, for example, high grain-Fe, Zn and protein contents.

References

Abdel-Ghani, A.H., Parzies, H.K., Omary, A. & Geiger, H.H. (2004). Estimating the outcrossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. *Theoretical and Applied Genetics* 109, 588-595.

Abdi, A., Bekele, E., Asfaw, Z. & Teshome, A. (2002). Patterns of morphological variation of sorghum [*Sorghum bicolor* (L.) Moench] landraces in qualitative characters in North Shewa and South Welo, Ethiopia. *Hereditas* 137 161-172.

Al-Janbi, S.M., Honeycutt, R.J., Peterson, C. & Sobral, B.W.S. (1997). Phylogenetics analysis of organellar DNA sequences in the Andropogoneae: Saccharum. *Theoretical and Applied Genetics* 88, 933-944.

Anglani, C. (1998). Sorghum for human food: a review. *Plant Foods for Human Nutrition* 52, 85-89.

Antonopoulou, G., Gavala, H.N., Skiadas, I.V., Angelopoulos, K. & Lyberatos, G. (2008). Biofuels generation from sweet sorghum. Fermentative hydrogen production and anaerobic digestion of the remaining biomass. *Bioresource Technology* 99, 110-119.

Ashok Kumar, A., Reddy Belum, V.S., Ramaiah, B., Sanjana Reddy, P., Sahrawat, K.L. & Upadhyaya, H.D. (2009). Genetic variability and plant character association of grain Fe and Zn in selected core collection of sorghum germplasm and breeding lines. *Journal of SAT Agricultural Research* 7.

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

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

Baldwin, B.G. (1992). Phylogenetic utility of the internal transcribed spacers of the nuclear ribosomal DNA in plants. An example from the Compositae. *Molecular Phylogenetics and Evolution* 1, 3-16.

Baldwin, B.G., Sanderson, M.J., Porter, J.J., Wojciechowski, M.F., Campbell, C.S. & Donoghue, M.J. (1995). The ITS region of the nuclear ribosomal DNA. A valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 85, 247-277.

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

Bekele, E., Geleta, M., Dagne, K., Jones, A.L., Barnes, I., Bradman, N. & Thomas, M.G. (2007). Molecular phylogeny of genus *Guizotia* (Asteraceae) using DNA sequences derived from ITS. *Genetic Resources and Crop Evolution* 54, 1419-1427.

Beta, T., Rooney, L.W. & Waniska, R.D. (1995). Malting characteristics of sorghum cultivars. *Cereal Chemistry* 72, 533-538.

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

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

BSTID-NRC (1996). *Lost crops of Africa*. Washington DC: Academic Press.

Bucheyeki, T.L., Gwanama, C., Mgonja, M., Chisi, M., Folkertsma, R. & Mutegi, R. (2009). Genetic variability characterisation of Tanzania sorghum landraces based on simple sequence repeats (SSRs) molecular and morphological markers. *African Crop Science Journal* 17, 71-86.

Bänzinger, M. & Long, J. (2000). The potential for increasing the iron and zinc density of maize through plant breeding. *Food and Nutrition Bulletin* 21, 397-400.

Celarier, R.P. (1958). Cytotaxonomy of the Andropogoneae. HI. Subtribe Sorgheae, Genus *Sorghum*. *Cytologia [Tokyo]* 23, 395-417.

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

Chakauya, E., Tongoona, P., Matiburi, E.A. & Grum, M. (2006). Genetic diversity assessment of sorghum landraces in Zimbabwe using microsatellites and indigenous local names. *International Journal of Botany* 2, 2006.

Clayton, W.D. (1961). Proposal to conserve the generic name *Sorghum* Moench (Gramineae) versus *Sorghum* Adans (Gramineae). *Taxon* 10, 242.

Clayton, W.D. & Renvoize, S.A. (1986). Genera Graminum Grasses of the World. *Kew Bulletin Additional Series, Royal Botanic Gardens, Kew, London*, 338-345.

Craufurd, P.Q., F, M., Bidinger, F.R., S.Z, M., Chantereau, J., Omanga, P.A., Qi, A., Roberts, E.H., Ellis, R.H., Summerfield, R.J. & Hammer, G.L. (1999). Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. *Theoretical and Applied Genetics* 99, 900-911.

de Alencar Figueiredo, L.F., Calatayud, C., Dupuits, C., Billot, C., Rami, J.-F., Brunel, D., Perrier, X., Courtois, B., Deu, M. & Glaszmann, J.-C. (2008). Phylogeographic evidence of crop neodiversity in sorghum. *Genetics* 179, 997-1008.

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. & Harlan, J.R. (1971). The origin and domestication of *Sorghum bicolor* (L) Moench. *Economic Botany* 25, 128-135.

Di Rienzo, J.A., Casanoves, F., Balzarini, M.G., Gonzalez, L., Tablada, M. & Robledo, C.W. (InfoStat version 2010). *InfoStat Group*. College of Agricultural Sciences, Nacional University of Cordoba, Argentina.

Dillon, S.L., Lawrence, P.K. & Henry, R.J. (2001). The use of ribosomal ITS to determine phylogenetic relationships within Sorghum. *Plant Systematics and Evolution* 230, 97-110.

Dillon, S.L., Lawrence, P.K., Henry, R.J. & Price, H.J. (2007). Sorghum resolved as a distinct genus based on combined ITS1, ndhF and Adh1 analyses. *Plant Systematics and Evolution* 268, 29-43.

Dillon, S.L., Lawrence, P.K., Henry, R.J., Ross, L., Price, H.J. & Johnston, J.S. (2004). *Sorghum laxiflorum* and *S. macrospermum*, the Australian native species most closely related to the cultivated *S. bicolor* based on ITS1 and ndhF sequence analysis of 25 Sorghum species. *Plant Systematics and Evolution* 249, 233-246.

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

Dje, Y., Heuertz, M., Lefebvre, C. & 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.

Doebley, J., Durbin, M., Golenberg, E.M., Clegg, M.T. & Ma, D.P. (1990). Evolutionary analysis of the large subunit of carboxylase (rbcL) nucleotide sequence among the grasses (Gramineae). *Evolution* 44, 1097-1108.

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

Doggett, H. (1970). *Sorghum*. 2nd. ed: Wiley-blackwell. ISBN 05 82463459.

Doggett, H. (1988). *Sorghum, 2nd edition*. London: Longman Scientific and technical.

Duodu, K.G., Taylor, J.R.N., Beltron, P.S. & Hamaker, B.R. (2003). Factors affecting sorghum protein digestibility. *Journal of Cereal Science* 38, 117-131.

Duvall, M.R. & Doebley, J.F. (1990). Restriction-site variation in the chloroplast genome of Sorghum (Poaceae). *Systematic Botany* 15, 472-480.

Dykes, L. & Rooney, L.W. (2006). Sorghum and millet phenols and antioxidants. *Journal of Cereal Science* 44, 236-251.

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

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

FAO (2011). FAOSTAT. http://faostat.fao.org.

Folkertsma, R.T., Rattunde, H.F.W., Chandra, S., Raju, G.S. & 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.

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

Gaut, B.S. & Doebley, J.F. (1997). DNA sequence evidence for the segmental allotetraploid origin of maize. *Proceedings of the National Academy of Sciences of the United States of America* 94, 6809-6814.

Geleta, N. & Labuschagne, M.T. (2005). Qualitative traits variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from eastern highlands of Ethiopia. *Biodiversity and Conservation* 14, 3055-3064.

Geleta, N., Labuschagne, M.T. & 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.

Gepts, P. (2004). Crop domestication as a long-term selection experiment, Volume 24, Part 2: Long-term selection: crops, animals, bacteria. In: Jannick, J. (Ed.) *Plant Breeding Reviews*. New York: Wiley.

Ghebru, B., Schmidt, R.J. & Bennetzen, J.L. (2002). Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. *Theoretical and Applied Genetics* 105, 229-236.

Gielly, L. & Taberlet, P. (1994). The use of chloroplast DNA to resolve plant phylogenies-noncoding versus RBCL sequences. *Molecular Biology and Evolution* 11, 769-777.

Glew, R.H., Vanderjagt, D.J., Lockett, C., Grivetti, L.E., Smith, G.C., Pastuszyn, A. & Millson, M. (1997). Amino acid, fatty acid, and mineral composition of 24 indigenous plants of Burkina Faso. *Journal of Food Composition and Analysis* 10, 205-217.

Graham, R., Senadhira, D., Beebe, S., Iglesias, C. & Monasterio, I. (1999). Breeding for micronutrient density in edible portions of staple food crops: Conventional approaches. *Field Crops Research* 60, 57-80.

Guo, Q., Huang, K., Yu, Y., Huang, Z. & Wu, Z. (2006). Phylogenetic relationships of Sorghum and related species inferred from sequence analysis of the nrDNA ITS region. *Agricultural Sciences in China* 5, 250-256.

Gwanama, C. & Nichterlein, K. (1995). Importance of cucurbits to small scale farmers in Zambia. *Zambia Journal of Agricultural Sciences* 5, 5-9.

Hair, J.F.J., Anderson, R.E., Tatham, R.L. & Black, W.C. (1998). *Multivariate Data Analysis*. 5th ed. London: Prentice-Hall international, Inc.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis programme of Windows 95/98/NT. *Nucleic Acids Symposium Series* 41, 95-98.

Hamby, R.K. & Zimmer, E.A. (1988). Ribosomal-RNA sequences for inferring phylogeny within the grass family (Poaceae). *Plant Systematics and Evolution* 160, 29-37.

Hamrick, J.L. (1983). The distribution of genetic variation within and among natural plant populations. In: Schoewald-Cox, C.M*., et al*. (Eds.) *Genetics and Conservation*. Menlopark: Benjamin Cummings.

Hamrick, J.L. & Godt, M.J.W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London, Series B* 351, 1291-1298.

Harlan, J.R. & de Wet, J.M.J. (1972). A simplified classification of cultivated sorghum. *Crop Science* 12, 172-176.

Hopkins, B.G., Whitney, D.A., Lamond, R.E. & Jolley, V.D. (1998). Phytosiderophore release by sorghum, wheat, and corn under zinc deficiency. *Journal of Plant Nutrition* 21, 2623-2637.

House, L.R. (1985). *A guide to sorghum breeding*. Second. ed. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

House, W. (1999). Trace element bioavailability as exemplified by iron and zinc. *Field Crops Research* 60, 115-141.

Hussain, A., Larsson, H., Kuktaite, R. & Johansson, E. (2010). Mineral composition of organically grown wheat genotypes: contribution to daily minerals intake. *International Journal of Environmental Research and Public Health* 7, 3442-3456.

IBPGR & ICRISAT (1993). Descriptors for sorghum [*Sorghum bicolor* (L.) Moench]. In: International Board for Plant Genetic Resources Rome Italy & International Crops Research Institute for the Semi-Arid Tropics Patancheru India (Eds.).

Jambunathan, R., Mertz, E.T. & Axtell, J.D. (1975). Fractionation of soluble proteins of high-lysine and normal sorghum grain. *Cereal Chemistry* 52, 119-121.

Jambunathan, R. & Subramanian, V. Grain quality and utilization of sorghum and pearl millet. In: *Proceedings of Biotechnology in tropical crop improvement. Proceedings of the International Biotechnology Workshop*, Patancheru, India, 12-15 January 1987 1988. pp. 133-139: Patancheru, ICRISAT.

Kayode, A.P.P., Linnemann, A.R., Hounhouigan, J.D., Nout, M.J.R. & Van Boekel, M. (2006a). Genetic and environmental impact on iron, zinc, and phytate in food sorghum grown in Benin. *Journal of Agricultural and Food Chemistry* 54, 256-262.

Kayode, A.P.P., Linnemann, A.R., Hounhouigan, J.D., Nout, M.J.R. & Van Boekel, M. (2006b). Genetic and environmental impact on iron, zinc, and phytate in food sorghum grown in Benin. *Journal of Agricultural and Food Chemistry* 54(1), 256-262.

Kong, L., Dong, J. & Hart, G.E. (2000). Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). *Theoretical and Applied Genetics* 101, 438-448.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948.

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

Li, R., Zhang, H., Zhou, X., Guan, Y., Yao, F., Song, G., Wang, J. & Zhang, C. (2010). Genetic diversity in Chinese sorghum landraces revealed by chloroplast simple sequence repeats. *Genetic Resources and Crop Evolution* 57, 1-15.

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

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

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

Merrill, A.L. & Watt, B.K. (1973). *Energy value of foods: Basis and derivation*. Washington D.C: ARS, United States Department of Agriculture, Handbook No. 74.

Morden, C.W., Doebley, J. & Schertz, K.F. (1990). Allozyme variation among the spontaneous species of *Sorghum* section *sorghum* (Poaceae). *Theoretical and Applied Genetics* 80, 296-304.

Muraya, M., Sagnard, F. & Parzies, H.K. (2010). Investigation of recent population bottlenecks in Kenyan wild sorghum populations (*Sorghum bicolor* (L.) Moench ssp. verticilliflorum (Steud.) De Wet) based on microsatellite diversity and genetic disequilibria. *Genetic Resources and Crop Evolution* 57, 995-1005.

Murdock, G.P. (1959). Staple subsistence crops of Africa. *Geographical Review* 50, 521-540.

Mutegi, E., Sagnard, F., Semagn, K., Deu, M., Muraya, M., Kanyenji, B., de Villiers, S., Kiambi, D., Herselman, L. & 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.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* 29, 1-10.

Nkongolo, K.K., Chinthu, K.K.L., Malusi, M. & Vokhiwa, Z. (2008). Participatory variety selection and characterization of Sorghum (*Sorghum bicolor* (L.) Moench) elite accessions from Malawian gene pool using farmer and breeder knowledge. *African Journal of Agricultural Research* 3, 273-283.

Nybom, H. & Bartish, I.V. (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics* 3, 93-114.

Obilana, T. (2004). Sorghum genetic enhancement- Research process, dissemination and impacts. In: Bantilan, M.C.S*., et al.* (Eds.). p. 320. Patancheru 502324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics.

Oria, M.P., Hamaker, B.R. & Shull, J.M. (1995). Resistance of Sorghum α -, β-, and γ - Kafirins to pepsin digestion. *Journal of Agricultural and Food Chemistry* 43, 2148-2153.

Pavlicek, A., Hrda, S. & Flegr, J. (1999). Free Tree-Freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of the genus *Frenkelia*. *Folia Biologica Prague* 45, 97-99.

Price, H.J., Dillon, S.L., Hodnett, G., Rooney, W.L., Ross, L. & Johnston, J.S. (2005). Genome evolution in the genus *Sorghum* (Poaceae). *Annals of Botany* 95, 219-227.

Purseglove, J.W. (1972). *Tropical crops; monocotyledons; Tropics*. London: Longman. ISBN 0582466482.

Raubeson, L.A. & Jansen, R.K. (2005). Chloroplast genomes of plants. In: *Plant diversity and evolution: genotypic and phenotypic variation in higher plants*. pp. 45-68. Wallingford UK: CABI Publishing. ISBN 0-85199-904-2.

Reddy Belem, V.S., Ramesh, S. & Sanjana Reddy, P. (2004). Sorghum Breeding Research at ICRISAT - Goals, Strategies, Methods and Accomplishments. *International Sorghum and Millets Newsletter* 45, 5-12.

Reddy Belum, V.S., Ramesh, S. & Longvah, T. (2005). Prospects of breeding for micronutrients and carotene-dense sorghums. *Journal of SAT Agricultural Research* 1, 1-4.

Reddy, B.V.S., Prakasha, R., Deb, U.K., Stenhouse, J.W., Ramaiah, B. & Ortiz, R. (2004). Global sorghum genetic enhancement processes at ICRISAT. In: Bantilan, M.C.S*., et al*. (Eds.) *Sorghum genetic enhancement: Research process, dissemination and impacts*. pp. 65-102. Patancheru 502 324 Andhra Pradesh India: International Crops Research Institute for the Semi-Arid Tropics.

Reddy, V.G., Upadhyaya, H.D. & Gowda, C.L.L. (2006). Current Status of Sorghum Genetic Resources at ICRISAT: Their Sharing and Impacts. *SAT eJournal* 2(1).

Rohlf , F.J. (1998). *NTSYS-pc 2.1. Numerical taxonomy and multivariate analysis system. Exeter Software:*. New York, USA.

Saski, C., Lee, S.B., Fjellheim, S., Guda, C., Jansen, R.K., Luo, H., Tomkins, J., Rognli, O.A., Daniell, H. & Clarke, J.L. (2007). Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. *Theoretical and Applied Genetics* 115, 571-590.

Schloss, S.J., Mitchell, S.E., White, G.M., Kukatla, R., Bowers, J.E., Paterson, A.H. & Kresovich, S. (2002). Characterisation of RFLP probes sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. *Theoretical and Applied Genetics* 105, 912-920.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W.S., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E. & Small, R.L. (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92, 142-166.

Shehzad, T., Okuizumi, H., Kawase, M. & 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.

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

Smith, C.W. & Frederiksen, R.A. (2000). *Sorghum: Origin, History, technology and Production*: John Wiley & Sons, Inc, USA. ISBN ISBN 0-471-24237-3.

Sneath, P.H.A. & Sokal, R.R. (1973). *Numerical taxonomy*. San Francisco: W.H. Freeman and Company.

Snowden, J.D. (1936). Cultivated race of sorghum. London: Adlard and Sons. 274 pp.

Spangler, R.E. (2003). Taxonomy of Sarga, Sorghum and Vacoparis (Poaceae: Andropogoneae). *Australian Systematic Botany* 16, 279-299.

Springer, P.S., Zimmer, E.A. & Bennetzen, T. (1989). Genomic organization of the ribosomal DNA of sorghum and its close relatives. *Theoretical and Applied Genetics* 77, 844-850.

Stemler, A.B.L., Harlan, J.R. & Dewet, J.M.J. (1977). Sorghums of Ethiopia. *Economic Botany* 31, 446-460.

Stenhouse, J.W., Prasada Rao, K.E., Gopal Reddy, V. & Appa Rao, S. (1997). Sorghum. In: Fuccillo, D., *et al.* (Eds.) *Biodiversity in Trust, Conservation and Use of Plant Genetic Resources in CGIAR Centres*. pp. 292-308. Cambridge: Cambridge University Press, United Kingdom.

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

Taberlet, P.L., Gielly, L., Pautou, G. & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17, 1105-1109.

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

Uptmoor, R., Wenzel, W., Friedt, W., Donaldson, G., Ayisi, K. & 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.

Vendemiatti, A., Rodrigues Ferreira, R., Humberto Gomes, L., Oliveira Medici, L. & Antunes Azevedo, R. (2008). Nutritional quality of sorghum seeds: Storage proteins and amino acids. *Food Biotechnology* 22, 377 - 397.

White, T.J., Bruns, T., Lee, S. & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: M. Innis, D.G., J. Sninsky and T. White (Ed.) *PCR Protocols: A Guide to Methods and Applications*. Academic Press, San Diego, California. pp. 315-322.

Wong, J.H., Marx, D.B., Wilson, J.D., Buchanan, B.B., Lemaux, P.G. & Pedersen, J.F. (2010). Principal component analysis and biochemical characterization of protein and starch reveal primary targets for improving sorghum grain. *Plant Science* 179, 598-611.

Wu, Y.Q. & Huang, Y.H. (2006). An SSR genetic map of *Sorghum bicolor* (L.) Moench and its comparison to a published genetic map. *Genome* 50, 84-89.

Yeh, F.C. & Boyle, T.J.B. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 129, 157.

Zhang, Y., Song, Q., Yan, J., Tang, J., Zhao, R., Zhang, Y., He, Z., Zou, C. & Ortiz-Monasterio, I. (2010). Mineral element concentrations in grains of Chinese wheat cultivar. *Euphytica* 174, 303-313.

Zohary, D. & Hopf, M. (2000). *Domestication of Plants in the Old World*. 3rd. ed. Oxford: University Press. ISBN 0198503571.

Acknowledgements

This may be seen as the tail end of the assignment but it marks the beginning of enormous challenges that lie ahead. I had the honour to steer this boat all the way to completion and finally the shore was in sight. Four years of my study life at SLU-Alnarp have changed my perception of life around humanity and the environment. The achievements made in this PhD study were not conceived and executed in a vacuum. I wish to attribute the successes that have been scored and the ability for timely completion of the study to the enabling environment at the University and in particular the support I received from my supervising staff, co-workers and colleagues within the SLU, Department of Plant Breeding and Biotechnology and Zambia Agriculture Research Institute (ZARI). Indeed, this work would not have been possible if it were not for the support, encouragement, advice and love I received from many other personalities of different persuasions, some of whom I may not mention in this section due to limitation of space.

I express my sincere gratitude to my main supervisor Professor Tomas Bryngelsson for his patience, tolerance, encouragement, support and seeing me through rather challenging times. Through him, my work environment at SLU–Alnarp was made fertile ground for all the research work that has been carried out. I have been enormously fortunate to have a supervisor who has always had the time, patience and was accommodating for any queries I had for him during the course of the programme. I am also thankful to him for untiringly checking different versions of the manuscripts prior to submission for publication without getting bored of mistakes in them.

I would like to thank Dr. Mulatu Geleta for the patience and untiring supervisory role he played during this programme. He understood my problems very well and most of the times he went miles off his busy

schedules in the endeavour to share his knowledge and experiences. I was always amazed by the amount of knowledge, his way of thinking, calm nature, and switching gears from one research area to another during the lab meetings. His enthusiasm always motivated, encouraged and compelled me to generate the best out of my research.

May I also thank Dr. Moneim Fatih for the key responsibility and facilitator role he played during this programme that enabled completion of what initially looked like an impossible mission. The timely provision of the required logistical support during the programme is highly appreciated. I wish to thank successive management teams of SADC Plant Genetic Resources Centre for ensuring that this programme became a reality even though it was implemented in the very last phase of the Nordic countries supported SADC Plant Genetic Resources Project. To the Board of SPGRC and the membership of the SADC PGR network, who I thank sincerely, the concerted effort on your part for this programme to take off was not in vain. Yes, its implementation in its early stages of the life of the SPGRC project could perhaps have had reaping effects. In this respect, the chairperson of the NPGRC-Zambia, Dr. S.W. Muliokela, who was chairperson for the SPGRC board and his team then, should be commended for the concerted effort towards realization of this programme.

I am indebted to The Swedish International Development Agency (Sida) for providing the financial support for my PhD study and assurance of provision of financial support towards the study programme beyond the SADC Plant Genetic Resources Project life at the end of 2009. The staff of SPGRC particularly Dr. P. Munyenyembe, T. Lupupa, B. Kapange and L. Qhobela should be commended for assisting in many ways during the period of my study.

I am indebted to the Australian Tropical Crops and Forage Genetic Resource Centre, Biloela, Queensland, Australia and in particular Dr Sally Dillon whose personal effort enabled timely access to sorghum accessions for a study. I wish to also thank the national genebanks of Malawi, Mozambique, Tanzania and Zambia and the sorghum and Millet Improvement Programme in Zambia for the provision of sorghum accessions.

I would like to express my special gratitude to all members of the Department of Plant Breeding at SLU-Alnarp for making my stay in Sweden a happy and comfortable one. Ann-Charlotte Strömdahl has not only helped me with the molecular laboratory work but also along with Jonas Hansson took their commitment to assist access health service as

needed. Pia Ohlsson was instrumental in the laboratory session on the molecular marker techniques when she was at Nordgen. Thanks to Ann-Sofie Fält, Helen Lindgren and Annelie Åhlman for the assistance provided in the laboratory particularly ordering of reagents and gels in the absence of Ann-Charlotte. I am thankful to Drs. Carlos Henry Loaisiga, Phuong Nguyen, Åsa Grimberg, Emily Too, Ida Lager, Toan Duc Pham and to Therese Bengtsson, Birjan Usubaliev, Sergey Hegay, Claid Mujaju, Helle Turesson, Lena Mabande, Isabel Herrera, Firuz Odilbekov, Fredrik Reslow, Maksat Amanov, Bakhromiddin Khusenov, Makhbubdzhon Rahmatov, Maria Luisa Prieto-Linde, Svetlana Leonova, Thuy Doan, Sonja Loots, Tiny Motlhaodi, Mbaki Muzila, Mohammed Elsafi, Abilio Afonso and Busisiwe Nsibande for the unfailing friendship.

It is also my great pleasure to thank Professor Eva Johansson, Professor Emeritus Waheeb Heenen, Dr. Erland Liljeroth, Dr. Anders Carlsson, Dr. Helena Persson, Dr Ramune Kuktaite, Dr. Anna Holefors, Dr Li-Hua Zhu, Dr Per Hofvander, Professor Sten Stymne, Professor Hilde Nybom, Dr Salla Marttila and many others for your encouragement and valuable contribution to this achievement.

Let me thank the Government of the Republic of Zambia for providing me with the full paid study leave to undertake the PhD programme. Allow me to also thank the management under the Directorate of Dr. Richard Kamona and staff of Zambia Agriculture Research Institute at Mount Makulu Research Centre and in particular staff of the National Plant Genetic Resources Centre for coming to my assistance with the field work. Mr. Kapalu Sakayula should be commended for his dedication and assistance provided with the field work.

I would like to express my deepest thanks and regards to my mother Nelie Kumwenda, mother in law, my brothers and sisters for your love, support, encouragement and reassurance during good and bad times of these years.

Finally, I would like to express my deepest love to my wife Mweemba and my children Joseph and Nchawaka. Thank you for your support, understanding and sharing happiness. You have been resilient enough all these years. Yes, it has not been that easy for you. God bless you.