Genetic Analysis of Inflorescence Architecture and Yield Components in Sorghum

January 2014

Khaing Pann Witt Hmon

Genetic Analysis of Inflorescence Architecture and Yield Components in Sorghum

A Dissertation Submitted to

the Graduate School of Life and Environmental Sciences,

the University of Tsukuba

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Agricultural Science

(Doctoral Program in Biosphere Resource Science and Technology)

Khaing Pann Witt Hmon

Table of contents

CHAPTER 2

Variation of inflorescence architecture associated with yield component traits in a sorghum germplasm

CHAPTER 3

QTLs underlying inflorescence architecture in sorghum detected by genome wide association analysis

List of Abbreviations

ANOVA	Analysis of variance
AwnL	Awn length
cDNA	complementary DNA
CIAT	Compact inflorescence architecture
CTAB	Hexadecyltrimethylammonium bromide
CulmL	Culm length
D’	Linkage disequilibrium coefficient
Df	Degree of freedom
DH	Days to heading
DNase	An enzyme to eliminate DNA
dNTP	Deoxynucleoside triphosphate
DTT	Dithriothreitol
EDTA	Ethylenediaminetetraacetate
EtBr	Ethidium bromide
FAO	Food and Agriculture Organization
GLM	General linear model

GNP	Grain number per panicle
GW	Grain weight (100 grain weight)
GWAS	Genome wide association mapping
GWP	Grain weight per panicle
IBPGR	The International Board for Plant Genetic Resources
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
L	Litre
LG	Linkage group
LIAT	Loose inflorescence architecture
MAS	Marker-assisted selection
MCMC	Markov chain Monte Carlo
mg	milligram(s)
ml	milliliter(s)
MLM	Mixed linear model
mm	millimetre(s)
MaxLBZ	Maximum length of primary branch zone
NecD	Neck diameter

NecL	Neck length
ng	Nanogram(s)
NIAS	National Institute of Aerobiological Sciences
NJ	Neighbor joining (bottom-up clustering method)
nM	Nanomoler
PanB	Panicle broadness
PanD	Panicle diameter
PanEx	Panicle exsersion
PAGE	Polyacrylamide gel ectrophoresis
PanH	Pant height
PanL	Panicle length
PanN	Panicle number
PanT	Panicle type
PanS	Panicle shape
PanW	Panicle width
Pend	Penducle length
QTL	Quantitative trait loci

Rac	Rachis length
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
rpm	Revolution per minute
SAS	Statistical analyses software
SDRS	Sorghum diversity research set
SEM	Structural equation modeling methodology
SNP	S***ingle-nucleotide polymorphism***
SPAGeDi	Spatial pattern analysis of genetic diversity
spp.	Specie(s)
SSR	Simple sequence Repeat
TASSEL	Trait analysis by association, evolution and linkage
TBE	Tris. HCl / Boric acid / EDTA
TE	Tris.HCl / EDTA
TotBr	Total number of branch
TotN	Number of total node
VG	Visual assessment by a single observation of a group of plants or parts of

	plants
VS	Visual assessment by observation of individual plants or parts of plants
WarpPLS	Script warp systems partial least squares
w/v	Weight/volume

List of Tables

List of Figures

CHAPTER 3

CHAPTER 1

General introduction and literature review

1. Introduction

To date, the increase in crop production and productivity has become a major challenge under the global explosion of human population in the 21st Century. Throughout the world, grass species are economically important both as staple grain food for humans and as feedstock for animals. Sorghum is more commonly grown in countries in the tropical and subtropical zones (FAO, 1999; Taylor and Dewar, 2001). It is a major dryland cereal crop in arid and semi-arid environments with low and unpredictable rainfall and also can adapt well to various soil types and toxicities (Singh *et al*., 2001; Vermerris *et al*., 2007). Sorghum is grown in the arid and semi-arid regions of the world as a cereal crop due to its high stress tolerance caused by numerous morphological and physiological characteristics. It is utilized for cereal grain, stalk, fiber, forage, fermented foods, beverages, sugars and building materials. The world largest producers of sorghum include USA, Mexico, India, China, Nigeria and Sudan. Over 61 million tons of sorghum was harvested from over 39 million ha grown worldwide in 2010 (http://www.agrostats.com/world-statistic/world-sorghum-production). The grasses (Family: Poaceae) constitute a large family containing about 10,000 species in the monocotyledonous plants (Kellogg, 2000). The grass species include some of the most important food crops in the world. The most remarkable example of grass species such as rice, maize and sorghum show remarkable diversity in morphological, physiological, genetic and ecological traits. They have been greatly improved in plant architecture and grain yield productivity. One of the important approaches is the improvement of plant architecture. It is considered as a valuable approach to increase grain yield, because crop plants with desirable architecture are able to produce much higher yields. Sorghum has its center of

diversity in Africa, where the crop and wild sorghums co-exist (Doggett, 1998). The origin and domestication of sorghum is estimated to have occurred around 3000 BC in Africa, in particular, Ethiopia and part of the Congo region, from where they migrated to Asia during the period of human migration (Kimber, 2000). The secondary centers of diversity include India, Sudan and Nigeria (Ayana, 1998). Some authors suggested that sorghum was domesticated in South China before 1045 BC and they introduced to North China as early as 850 BC (Kimber, 2000). Early introductions of this crop to the United States of America occurred in 1853 (Maunder 1999), when a sweet Chinese amber sorghum was introduced from France (Martin and Leonard, 1949) (**Figure 1.1**).

Cultivated sorghum is classified into five main races (Bicolor, Guinea, Caudatum, Durra and Kafir) (Barnaud *et al*., 2008a; Harlan and de Wet, 1972) (**Figure 1.2**). Among these races, Guineas is the oldest of the specialized races because of its relatively wide distribution and diversity (Harlan *et al*., 1976). According to Stemler *et al*., (1977) the race Caudatum was domesticated after Bicolor and Guinea. The race Kafir was derived from an early Bicolor race by de Wet (1978). The genetic variation of both cultivated and wild sorghum has been studied by many researchers to classify landrace groups and to investigate phylogenetic differentiation. However, all these divisions are mostly based on their morphological traits, especially panicle and grain characteristics (Harlan and de Wet, 1972). Among those traits, sorghum's panicle shape, type and grain color are the most important in sorghum identification.

Moreover, the knowledge of the genetic basis of the link between sorghum inflorescence architecture and yield related traits can complement breeder's efforts on genetic improvement of sorghum breeding. In recent years, the genetics of inflorescence architecture has been studied extensively from a molecular biological point of views, and many genes controlling the inflorescence architecture and the development of the floral organs has been

cloned and characterized (Weigel and Meyerowitz, 1994). Few studies have been carried out on the morphological diversity of sorghum panicles. In sorghum most yield related traits are polygenic, and inflorescence architecture also remains a complex traits controlled by multiple genes or polygenes (House 1985; Bello *et al*., 2001; Zou *et al*., 2011). Sorghum inflorescences (panicles) have a large diversity of types, ranging from very open and loose type to a very compact type. The size and shape of inflorescence organs generally show continuous variation in many plant species and are quantitative traits (Shore and Barrett 1990). Elucidation of the quantitative traits of the genetic base underlying variation of inflorescence architecture might allow us to understand how diverse the variation in inflorescence morphology has occurred.

This study first clarified the relation between geographic origin and the variation of inflorescence architecture to provide the information on the origins of the accessions using sorghum world-wide germplasm. Secondly is to study the relation between inflorescence architecture and yield potential at the intraspecific levels. Finally, QTLs associated with sorghum inflorescence architecture using the sorghum diversity research set (SDRS) of 107 landraces from worldwide sorghum germplasm to be identified by genome wide association analyses.

1.1. Sorghum morphology, taxonomy and classification

Sorghum species are both annual and perennial herbaceous plant (maturing approximately within 60 to 180 days). Sorghum has well developed root system and the stalk (stem) is strong, hard, and smooth divided by nodes. They germinate from grain forms and grow up to 75 – 250 cm, with various thickness (1 – 2.5 cm) and have either a dry, semi-dry or juicy marrow. Basal plant color is either red, reddish brown or purple. The leaves vary

from 8 to 20 in numbers, 50 – 100 mm in width and 0.5 – 1 m in length. Sorghum inflorescence (panicle) has different shape and size, usually 10 – 40 cm long. Inflorescence types vary from open to compact with a wide range of dimensions. Sorghum is essentially allogamous but often self fertilized. Sorghum seeds are round or oval -shaped with various colors such as white, creamy, yellow, pink, orange, brown and violet. Common name "Sorghum*"* is in the family Poaceae, subfamily Panicoideae and the tribe of Andropogoneae. *Sorghum bicolor* (L.) Moench. Commercially, it is known as grain sorghum, fodder sorghum, broom corns, and sweet sorghum. The most complete and defined classification of cultivated sorghum, *Sorghum bicolor* (Linn.). Moench was determined by Snowden (1936) in which hereafter all classifications have been modified according to the Snowden's system. Harlan and de Wet (1972) reported a simplified classification of sorghum which has proven to be of real practical utility for sorghum workers instead of wading through the 31 categories in the Snowden's key for cultivated sorghums. The genus *Sorghum* belongs to one of the 16 subtribes of the tribe Andropogoneae. The common scientific classification of sorghum is as follow:

Kingdom - Plantae- Plants

Subkingdom - Tracheobionta- Vascular plants

Superdivision - Spermatophyta- Seed plants

Division - Magnoliophyta - Flowering plants

Class - Liliopsida- Monocotyledons

Subclass - Commelinidae

Order - Cyperales

Family - Poaceae – Grass family

Genus - Sorghum

Species - *Sorghum bicolor*

Subspecies - *Sorghum bicolor* ssp. *arundinaceum* (Common wild Sorghum)

Subspecies - *Sorghum bicolor* ssp. *bicolor* (Grain Sorghum)

Subspecies - *Sorghum bicolor* ssp. *drummondii* (Sudan grass)

Species- *Sorghum almum* (Columbus grass)

Species- *Sorghum halepense* (Johnson grass)

Species - *Sorghum propinquum* (Columbus grass)

Harlan and de Wet (1972) partitioned the primary gene pool of *Sorghum bicolor* Moench, into the following races by using a much clearer and simpler system. The races are, for the most part could be identified easily by its spikelet morphology (inflorescence morphology).

Cultivated races: *S. bicolor ssp bicolor*

Basic races:

Race (1) bicolor (B),

Race (2) guinea (G),

Race (3) caudatum (C),

Race (4) kafir (K),

Race (5) durra (D).

I. **Intermediate races: (all combinations of basic races)**
    a. Race (6) guinea-bicolor (GB)
    b. Race (7) caudatum-bicolor (CB)
    c. Race (8) kafir-bicolor (KB)
    d. Race (9) durra-bicolor (DB)

e. Race (10) guinea-caudatum (GC)

f. Race (11) guinea-kafir (GK)

g. Race (12) guinea-durra (GD)

h. Race (13) kafir-caudatum (KC)

i. Race (14) durra-caudatum (DC)

j. Race (15) kafir-durra (KD)

II. Spontaneous races: *S. bicolor ssp arundinaceum*.

a. Race (1) arundinaceum

b. Race (2) aethiopicum

c. Race (3) virgatum

d. Race (4) verticilliflorum

e. Race (5) propinquum

f. Race (6) shattercane

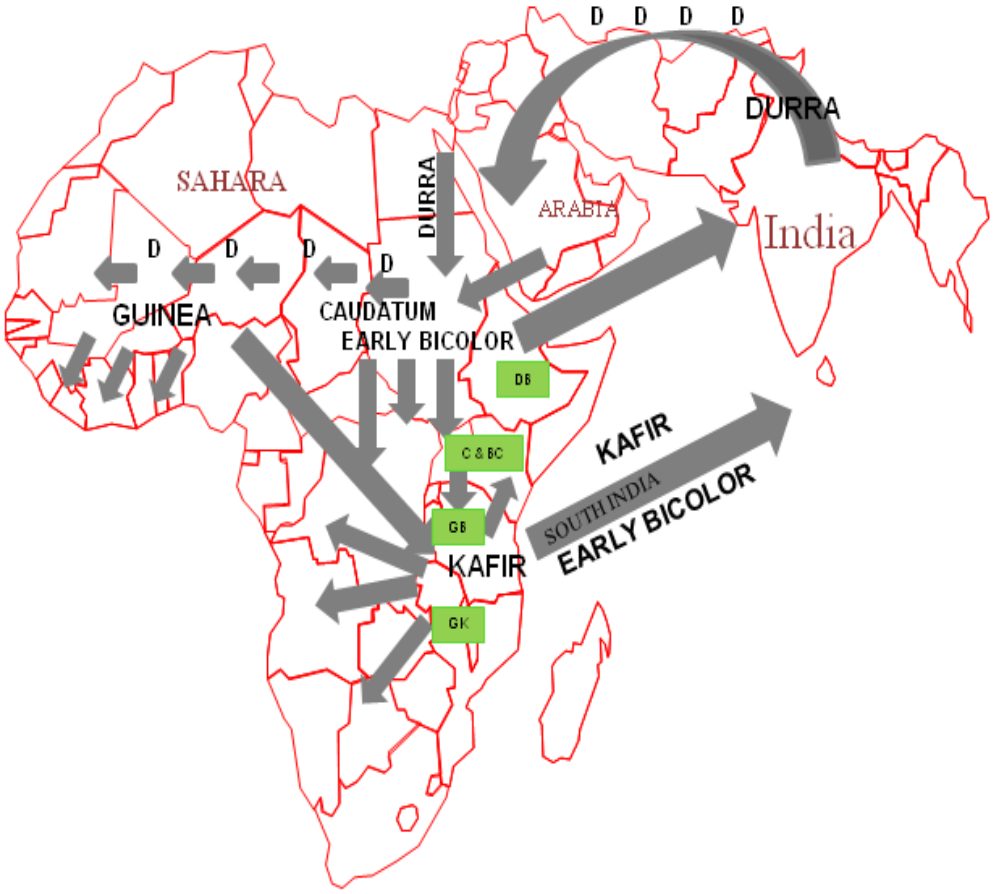
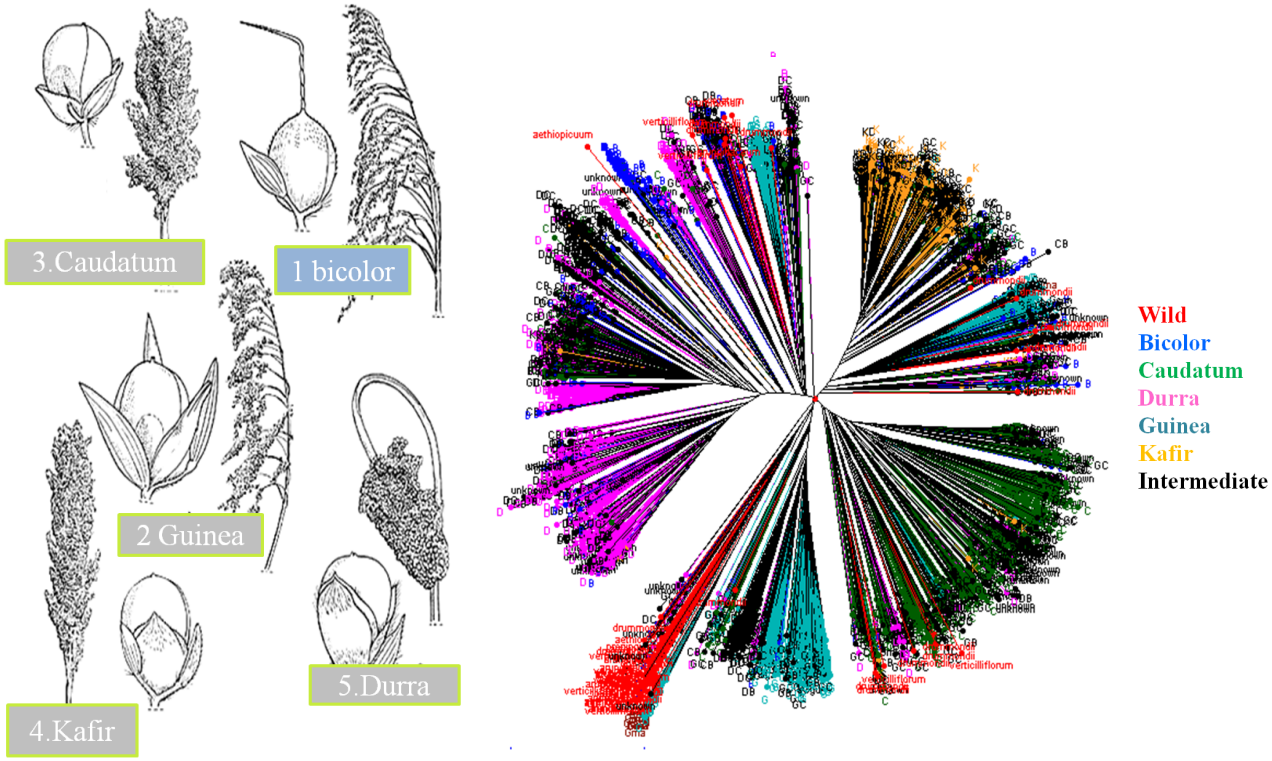


Figure 1.1. Area of origin, development for the domesticated races and migration routes.

(This figure is developed based on the reference: Kimber *et al.*, (2013))



(i) (ii)

Figure 1.2. (i) The different panicle types and grain characters of the 5 main sorghum races, Source: PROTA (Plant Resources of Tropical Africa), http://www.prota4u.org/search.asp.

(ii) Wild and cultivated races of sorghum by Harlan and de Wet, 1972.

(This figure is developed based on two references: Baloe *et al.,* 2006 and Geoffrey *et al*., 2012)

1.2. Sorghum inflorescence architecture

Sorghum inflorescence (panicle) has a large diversity of types, ranging from very open and loose types to a very compact head pattern types. In cereal crops, inflorescence (panicle) development and productivity are the principal factors for yield potential. There are three distinct phases of the inflorescence growth stage: vegetative, floral initiation (panicle development) and grain filling. In the inflorescence anatomy, the inflorescence meristem essentially bridges the gap between the two main categories of aerial meristem: the vegetative meristem produces leaves and stems, and the reproductive meristem produces only floral organs. Inflorescence (panicle) architecture is regulated the four kinds of reproductive meristems: rachis meristem, branch meristem, spikelet meristem and floral meristem. The change from vegetative apical meristem to reproductive meristem for panicle formation. The apical meristem of the primary branch produces secondary branches and spikelets. The spikelet meristem is then transformed to one or more inflorescence meristems. Panicle formation begins at about 4 leaves stage and reaches above the ground and begins to enlarge at about 6 leaves stage. Sorghum inflorescence is 5 – 60 cm long, 3 – 30 cm wide for both the open and compact types. The shape and color of the inflorescence (panicle) varies from cultivar to cultivar. Generally the basic structure of the mature panicle is a combination of the following components:

Peduncle: The panicle (head) emerges from the flag leaf sheath and is individually supported by a portion of stalk called peduncle. Peduncle is usually straight but in some cases, for example in broom corn, the head is curved.

Rachis: Rachis is the main axis, the length of the panicle starting from the whorl, bottom part of the branch to the top. Rachis can be straight, curved or drooping and its length varies, ranging from length of 2 cm to 52 cm.

Internodes on the rachis: Inflorescence branches remain intact at maturity with 1 to 5 nodes along the rachis. Each node can produce several branches.

The number of branches (whorl) on the rachis: Several branches (whorls) could be produced from the internodes of the rachis. Each lateral may also branch out repeatedly with each primary branch dividing into secondary branches and the secondary branch dividing into the tertiary branches. The final branches then carry the spikelets which can be single or consist of several paired spikelets.

Spikelet: Spikelets are borne on the branches formed off the main rachis primary branches compound terminates in racemes with 2 to 7 spikelet pairs. The lower one, sessile spikelet is bisexual and fruit bearing of about 3 – 9 mm long and 2 – 5 mm wide, elliptic to ovate in shape; calluses blunt; glumes coriaceous to membranous, glabrous, densely hirsute, or pubescent, keels usually winged; upper lemmas unawned or with a geniculate, twisted, awn 5 – 30 mm; anthers 2 – 2.8 mm. Pedicels is 1 – 2.6 mm in length. Pedicellate spikelet is 3 – 6 mm in length and is usually shorter than the sessile spikelet. It is staminate or sterile. Caryopses are exposed at maturity.

Grain: The number of grains per panicle is determined by the number of branch whorls, the number of primary branches per whorl, and the number of grains per primary branch. Each panicle contains about 800 to 3000 seeds which are usually partly covered by the glumes.

1.3. Mapping of QTLs for inflorescence architecture

Genetic information about sorghum inflorescence (panicle) characters is limited to date. Ayyangar and Ayyar (1938) and Ghawchawe *et al.,* (1996) reported that panicle

density and sterility are basically controlled by a single gene. Fanous *et al.,* (1971), Patel *et al.,* (1983), Kukadia *et al.,* (1983), Kumar and Singhania (1984) and Wenzel (1990) identified that panicle length and seed branch length are highly heritable traits. Kirby and Atkins (1968) identified that a major portion of the genetic variation for some panicle characteristic is due to the additive genetic effects. Pereira *et al.,* (1994) identified 10 linkage groups using several DNA markers, most likely corresponding to the haploid chromosome number of sorghum. Pereira and Lee (1995) identified 4 QTLs for plant height and three QTLs for flowering. Six QTLs of panicle related traits were identified by Pereira *et al.,* (1994), six QTLs for pre-flowering stress were detected using RIL population by Tuinstra *et al.,* (1996) and Crasta *et al.,* (1999), four major QTLs associated with the "stay green" trait by Xu (2002), Kebede *et al.,* (2001) were identified, three QTLs for stay green and Klein *et al.,* (2001) identified QTLs for foliar disease and grain mold resistance in a sorghum RIL population, Paterson *et al.,* (1998 and 2009) identified four QTLs for seed size and seed number. QTLs for other morphological traits of sorghum related to inflorescence architecture have not been identified.

1.4. Research aims and objectives

Systematic genomic analysis of sorghum panicle traits can lead to accelerate genetic improvement for the increment in yield capacity. Therefore this study is aimed to exploit the intraspecific variation of panicle traits across the sorghum core collection from around the world to evaluate not only a wide range of the phenotypic diversity but also its suitability for association analysis**.** The value of these traits, the relationship to the yield components and the preliminary effort for the association mapping may provide useful information for sorghum breeding. Therefore the main objectives of this study are:

(1) to evaluate the variation of panicle patterns across a world-wide collection of sorghum germplasm,

(2) to identify the main components related to sorghum inflorescence architecture,

(3) to clarify the relation between panicle component traits and yield related traits, and

(4) to identify QTL underlying inflorescence architecture by genome wide association analysis.

CHAPTER 2

Variation of inflorescence architecture associated with yield component traits in a sorghum germplasm

2.1. Introduction

The grasses (Family: Poaceae) constitute a large family containing about 10,000 species in the monocotyledonous plants (Kellogg, 2000). The most remarkable example of grass species such as rice, maize and sorghum show remarkable diversity in morphological, physiological, genetic and ecological traits. Sorghum is grown for cereal grain, stalk, fiber, forage, fermented foods, beverages, sugars and building materials. The world's largest producers of sorghum include USA, Mexico, India, China, Nigeria and Sudan. Over 61 million tons of sorghum was harvested from over 39 million ha grown worldwide in 2010 (http://www.agrostats.com/world-statistic/world-sorghum-production). Sorghum has its center of diversity in Africa, where the crop and wild sorghums co-exist (Doggett 1988). Nothing is known about when *Sorghum bicolor* was first brought into cultivation along with several West African crops, although it was domesticated some 7000 years ago. It reached to India not earlier than 1500 BC and China by 900 AD. The secondary centers of diversity include India, Sudan and Nigeria (Ayana and Bekele, 1998). These domesticated races have been associated with human migrations in Africa from where they migrated to Asia (Teshome *et al.*, 1997; Kimber, 2000; Kimber *et al.*, 2013). Early introductions of this crop to the United States occurred in 1853, when a sweet Chinese Amber sorghum was introduced from France (Martin and Leonard, 1949). Cultivated sorghum was first introduced to the Americas and Australia about 100 years ago (Kimber *et al.,* 2013). Sorghum is distributed in wild forms in Africa and other countries (Mann *et al.,* 1983). In sorghum, domestication was initiated by allelic changes at only two loci resulting from different selection pressures. The

essential step adopted in domestication was the harvest of the whole inflorescence, and the utilization of the grain for seed. The types in which panicle components, rachis, panicle branch spikelet node remained intact had a selective advantage for domestication. This recessive characteristic remained fixed to complete the process of domestication (Mann *et al.,* 1983). The domestication process continued for several thousand years. Several authors explain the center of origin of sorghum was in African countries (Mann *et al.,* 1983; Kimber*,* 2000; Kimber *et al.,* 2013), but Harlan and de Wet (1972) using archaeological, palaeobotanical, anthropological evidences as well as botanical evidence believed that the center of origin of sorghum extends from near Lake Chad in Africa. These areas represent the diversity and abundance of wild and weedy species as well as a presence of a primitive race of bicolor. Snowden (1936) reported sorghum to have separate centers of origin for different types. Cultivated sorghum is classified into five main races (Bicolor, Guinea, Caudatum, Durra and Kafir) (Harlan and de Wet, 1972; Barnaud *et al.,* 2008b; Dahlberg 2000; Kimber, 2000; Kimber *et al.,* 2013) and their divisions are mostly based on panicle and grain characteristics (Murray *et al*., 2009). Among these races *Guinea* is the oldest of the specialized races because of its relatively wide distribution and diversity (Harlan and de Wet, 1972). According to Stemler *et al.,* (1977) the race *Caudatum* is a later domesticated than *Bicolor* and *Guinea*. The race *Kafir* was derived from an early bicolor race by de Wet (1978) and Harlan *et al*., (1976). The genetic variation of both cultivated and wild sorghum has been studied by many researchers to classify landrace groups and to investigate phylogenetic differentiation. However, all these divisions were mostly based on their morphological traits, especially panicle and grain characteristics (Harlan and de Wet, 1972). Among those traits sorghum panicle type and its grain color are the most important traits in sorghum identification (House, 1985).

The agronomic performance of cereal crops is significantly influenced by the complexity of inflorescence (panicle) patterns. Sorghum germplasm can be identified according to their morphological traits (Kaitaniemi *et al*., 1999), especially the degree of expression of panicle and grain characteristics (Abdi *et al*., 2002; Harlan and de Wet, 1972). Sorghum inflorescences (panicle) have a large diversity of types, ranging from very open and loose types to a very compact head pattern. Inflorescence called panicle architecture is considered as a breeding target for plant architecture in crop species because the pattern of the panicle is an important character in sorghum to identify race classifications and species value. The structure and type of panicles are not only important agronomic factors in sorghum identification, but also thought to contribute to yield and grain quality as shown in rice and maize (Bommert *et al*., 2005; Ikeda *et al*., 2010; Yan *et al*., 2007; Doust *et al*., 2004, 2005; Hu *et al.,* 2003). These characteristics are very helpful for breeding and botanical purposes and still remain as a great interest for breeders (Doust *et al*., 2005; Bala *et al*., 1996; Doust and Kellogg, 2002; Futsuhara *et al*., 1979a, 1979b; Kellogg 2000; Zhu *et al*., 2010 and Tao *et al*., 1993). The knowledge of the genetic basis of sorghum inflorescence architecture and its component traits can enhance the process of improvement in sorghum breeding. Few studies have been carried out on the morphological diversity of sorghum panicles. In sorghum most yield related traits are polygenic and especially inflorescence architecture is probably controlled by the multiple genes or QTLs (House 1985; Bello *et al*., 2001; Zou *et al*., 2011). We expected that the systematic analysis of panicle traits could lead to improvement of grain yield in sorghum. Therefore in this study we have measured panicle dimensions, architecture and yield related traits to capture the intraspecific variation of panicle traits using (1) a large collection of 206 sorghum accessions from around the world and (2) the diversity research set (SDRS) of 107 landraces as core collection from 206 sorghum world-wide germplasm accessions. The value of these traits and the link between geographical information on the

origins of the accessions may be useful information for sorghum breeding. Therefore our objectives of this chapter were: (1) to identify the key components of variation in sorghum inflorescence architecture, (2) to clarify the association between panicle component traits and yield related traits and (3) to clarify the relationship between panicle pattern and their origin.

2.2. Materials and methods

2.2.1. Plant materials

In this study we used 206 accessions of sorghum which originated from 27 countries in Asia (East, Southeast, South and Southwest Asia) and Africa. These accessions were obtained from the germplasm collections at the National Institute of Aerobiological Science, Genebank, Japan. We categorized materials into 3 main different sub-populations based on geographical distributions are as follow: East Asia group (66 accessions from East Asia), other regions of Asia group (2 from Southeast Asia, 60 from South Asia, 2 from Southwest Asia) and African group (76 accessions from Africa) **(Table 2.1).** We grouped the accessions based on six panicle types; open type (33 accessions), intermediated type (40 accessions), semi-compact type (22 accessions), compact type (80 accessions), broom type (24 accessions) and mixed type (9 accessions). From the whole collection 206 sorghum accession, previously selected sorghum diversity research set (SDRS) of 107 landraces accessions (Shehzad *et al.,* 2009a) were set up to analysis based on inflorescence architecture. We categorized the plant materials into three different groups involving 25 accessions from East Asia, 2 accessions from Southeast Asia, 26 accessions from South Asia, 2 accessions from Southwest Asia and 52 accessions from Africa **(Table 2.10).**

2.2.2. Experimental methods

Field experiments were conducted in the year 2010 and 2011 sorghum main growing seasons in the experimental field at Agricultural and Forestry Center, University of Tsukuba, Japan. The sowing dates were 22 May in 2010 and 27 May in 2011. Dried seeds were sterilized in fungicide 1 week before sowing and seeds were sown manually by dibbling method directly into the field plots. Seeds are dibbled into a 2-3 cm deep and 1 cm apart in one hole. Four individual plants were grown out for each of the 206 accessions including 107 accessions, and plants from the different accessions were grown together in the same field in 60m x 1m plots. Weed control was done around plots in every week during crop season. Fertilizer and nutrients were used at 35 days after germination. Disease and insects were controlled with fungicide and insecticide. At anthesis stage all panicle were covered by pollination paper bags to prevent from out-crossing and several damages. The bags were removed after complete flowering when hardened grain stage. From all labeled plants mature panicles were harvested and dried naturally with net bags before traits measurement.

2.2.3. Trait measurements and methods

Mature panicles were harvested and dried before measurement of panicle traits. For 206 accessions all the data for 18 morphological components of panicle traits and yield related traits were recorded and data on 14 panicle traits were recorded for 107 core collection accessions according to the sorghum descriptors from IBPGR, ICRISAT and NIAS, Genebank detailed in (**Table 2.2)** and **(Table 2.11).** For each accession, panicle traits were measured using a single panicle from three separate plants and the results averaged. All measurements were made in metric units. At maturity stage main components of panicle traits

including rachis length (Rac), panicle length (PanL), penducle length (Pend), panicle shape (PanS), panicle type (PanT) and other yield component traits plant height (PanH) , culm length (Culm) were evaluated (**Figure 2.1 and Figure 2.6)**. Panicle diameter (PanD) and panicle width (PanW) were measured with calipers. Rachis length was measured as the distance from the bottom whorl to the topmost one. Peduncle length (Pend) was measured as the distance from flag leaf to the lowest primary branch zone. After harvesting, all labeled panicles were dried and cleaned before trait measurement. After cleaning total nodes (TotN), total number of primary branch (TotBr) and maximum length of primary branch (MxLBZ/MaxLBZ) were manually measured. At the basal part of the panicle, total primary branches (TotBr) were removed and counted individually. Total nodes (total number of the whole on the rachis) (TotN) was counted along main axis. For maximum length of branch zone (MxLBZ/MaxLBZ), three random branches were selected from the longest branch zone in the bottom third whole of panicle. The stage of panicle traits was also measured to support the visual assessment of panicle shape and type such as panicle exertion (PanEx) and panicle broadness (PanB). Next evaluations are actual number of grains per panicle (GNP) average of 3 panicles in the accession, grain weight per panicle (GWP) and 100 grain weight (GW) from each panicle. All grains were threshed and measured for GWP, GNP and GW. Additional data on ordinal grouping observations characters *i.e.*, neck diameter (NecD), neck length (NecL), awn presence/absent (Awn) and awn length (AwnL) were also recorded but data were not shown. All main panicle traits and panicle related traits were evaluated in three panicles per accession. For yield traits were measured in three labeled plants of each accession per row included plant high, culm length, penducle length and total panicle number per plant.

2.2.4. Statistical and general analyses

The overall data were divided into two clusters to calculate the association of panicle traits with other yield traits based on three different origins (East Asia, Other regions of Asia, Africa) and association of panicle traits with other yield related traits based on panicle types. All statistical analyses (ANOVA, correlation and PCA analysis) were performed using the JMP version 9.0 (SAS Institute, Inc, 2010). Analysis of variance (ANOVA), principle component analysis (PCA) were run on the total panicle traits and cluster analysis to perform compare the level of phenotypic variation for each trait across 3 different regions and six different panicle types. Correlation analysis was performed to elucidate the relationships among the investigated panicle traits. To understand the patterns of correlation among inflorescence architecture and their direct and indirect effects toward the yield, path analysis with structural equation modeling methodology (SEM) was carried out using WarpPLS software ver 3.0 (Kock, 2012).

2.3. Results

2.3.1. Frequency distributions of inflorescence architecture in sorghum worldwide germplasm.

We examined the variability of the 14 quantitative traits related to inflorescence architecture and yield of sorghum accessions from the three geographic regions over two years to identify the characters responsible for the majority of the variation in the inflorescence architecture. Phenotypic variations of these traits in 206 accessions are shown in (**Table 2.3**) **(Figure 2.2).** Sorghum accessions showed a broad range of variability of many traits measured for two years. The mean and range of the traits were shown widely distributed

and the ranges were broad in TotBr, Pend, PanD, PanN, GNP, GWP and GW traits varied significantly between two years. Analysis of variance ANOVA indicated that PanW, PanD, PanN, GNP, GWP and GW traits measured showed differences between two growing seasons. We examined the variability of these traits across three different regions. Phenotypic variations of these traits from East Asia, other Asian regions, and African region showed a broad range of variability of many traits measured. The mean and range of the traits were shown widely distributed and the ranges were broad in East Asian accessions such as Rac, PanH, CulmL, TotBr, GNP, GWP, GW and PanD (**Table 2.4**). In other regions of Asian accessions, most of the traits were also widely distributed such as Rac, PanL, MaxBLZ, PanH, TotBr, GNP, GWP, GW and PanD. The Culm, TotBr, GNP, GW and PanD traits were widely distributed in African accessions. Rac, MaxLBZ, TotN, GNP, GDP, PanW, PanD and PanN traits varied significantly between East Asian and African accessions. The African accessions had reproducibly shorter penducle length (Pend), more nodes (TotN) and longer rachis length (Rac). For the yield related traits grain weight (GW) and panicle number (PanN) differed significantly.

2.3.2. Phenotypic correlation among traits in sorghum worldwide germplasm.

We next considered whether there were significant correlations among measured traits. Trait pairs with significant correlations ($p<0.001$) are summarized in **(Table 2.5)**. Twelve of the fourteen traits showed at least one highly significant ($p<0.001$) correlation with another (the exceptions were Pend and GW). In terms of panicle architecture, most of the length based measurements showed a trend to be correlated. In particular panicle length positively correlated with most of the length based measurements (Rac, MaxLBZ, PanW, PanH, CulmL), and PanN and TotBr between multi years. Additionally the total number of branches

(TotBr) and number of panicle (PanN) also correlated with PanL and PanH. Importantly, grain yield related traits were also positively correlated with most of the length based measurements. In particular grain weight per panicle (GWP) was correlated with panicle width (PanW; 0.29) in 2010, panicle diameter (PanD; 0.27, 0.23) in 2010 and 2011; plant height (PanH; 0.23, 0.24) in 2010 and 2011; panicle length (PanL; 0.22) in 2010; culm length (CulmL; 0.22, 0.22) in 2010 and 2011. GWP was correlated with the non-length measurements, total branch on the rachis (TotBr; 0.21) in 2010; number of panicles (PanN; 0.22, 0.17) in 2010 and 2011. Grain weight per panicle (GWP) was no trait correlation with PanW, PanL, TotBr in 2011. Also grain number per panicle (GNP) was correlated with panicle width (PanW; 0.23) in 2010 but there was no correlation between PanW in 2011. As expected grain number per panicle (GNP) was highly correlated with grain weight per panicle (GWP; 0.60, 0.72) in 2010 and 2011. Penducle length was negatively correlated with yield related traits (GNP; -0.17) and (GWP; -0.16) in 2010.

2.3.3. Comparison between loose and compact inflorescence architecture associated with yield traits in sorghum worldwide germplasm.

In this study we further investigated the association of panicle traits with yield and yield components traits by two separate inflorescences architectures based on loose and compact panicle types and comparing their relationship. Comparison between loose and compact inflorescence architecture (by using only compact panicle type and open panicle type data) there are most of characteristics have shown to be related each other (**Table 2.6**). In the first growing season (Year 2010) data analysis revealed that the total of 82 phenotypic correlations were found to be significant among traits in compact and loose inflorescences architecture. Forty nine correlations were significant among traits in compact inflorescence

architecture (CIAT) (**Table 2.6**). Thirty three correlations were significant for only in loose inflorescence architecture (LIAT). The association between yield related traits and inflorescence architecture 12 traits were significantly correlated with yield related traits in compact inflorescence architecture (CIAT) and 9 traits were significantly correlated in loose inflorescence architecture (LIAT). Seven of eight total main panicle traits are highly significant correlation with each other except for TotN in LIAT such as Pend, Rac, PanL, MaxLBZ, TotBr, PanD and PanW. In both of loose and compact panicle types the correlation coefficient of PanL was significantly associated with most of the characters among accessions but the characteristic of total node (TotN) was not significantly correlated with other panicle traits in LIAT. Correlation coefficient table revealed that the different characteristics were associated with different inflorescence architectures such as PanL, MaxLBZ, TotN, TotBr, PanD and PanW were obviously significant with other characters in CIAT while PanL, TotBr and Rac traits showed a significant correlation with other characters in LIAT. The association of morphological traits and yield related trait such as PanH, CulmL, and GWP were more or less equally correlated with other traits between different inflorescence architecture. The traits most highly correlated with panicle trait pairs in CIAT are presented in Table 2.6 ($p<0.001$) and highly positive correlation of PanL with Rac, MaxLBZ, PanH, PanN and significant ($p<0.01$) and positive correlation of PanL with TotBr, PanS and PanW as well. The correlation of PanL was positive and significant ($p<0.05$) with CulmL, PanD in CIAT. The panicle trait pairs of loose inflorescence architecture are presented in Table 2.6 showed significant ($p<0.001$) and most highly positive correlation of PanL with Rac, TotBr and PanH. The correlation of PanL was positive and significant ($p<0.01$) with GWP, CulmL and PanW. The traits most highly correlated with yield in compact inflorescence architecture CIAT are significant ($p<0.01$) and positive correlation of grain yield with panicle width (0.38), plant high (0.28), culm length (0.29) and significant

($p<0.05$) and positive correlation of rachis length with grain yield (0.19), total branch with grain yield (0.16). Furthermore the most highly correlated with yield in the loose inflorescence architecture are significant ($p<0.01$) and positive correlation of panicle length with grain yield (0.35), total branch on the rachis with grain yield (0.36) and significant ($p<0.05$) and positive correlation of plant height with grain yield (0.31) and culm length with grain yield (0.29). Grain number per panicle (GNP) had positive and highly significant ($p<0.001$) correlation with grain yield (GWP). GNP had positive and highly significant ($p<0.001$) correlation with GWP (0.79) in both compact and loose inflorescence types, CIAT and LIAT. In the second growing season (Year 2011) data analysis revealed that the total of 69 phenotypic correlations was found to be significant among traits in compact and loose inflorescences architecture (**Table 2.7**). Twenty eight correlations were significant among traits in compact inflorescence architecture (CIAT). Forty one correlations were significant for only in loose inflorescence architecture (LIAT). The association between yield related traits and inflorescence architecture 10 traits were significantly correlated with yield related traits in compact inflorescence architecture (CIAT) and 6 traits were significantly correlated in loose inflorescence architecture (LIAT) (**Table 2.7**).

2.3.4. Path analysis among panicle traits in sorghum worldwide germplasm.

Given the observed correlations between the various length measurements and grain yield traits, we carried out path coefficient analysis to understand the interaction among all measured traits. Moreover, panicle main components and their direct and indirect effect on dependent variable panicle length were analyzed by path coefficient analysis. Path analysis (**Table 2.8 and Figure 2.3)** showed yield was directly influenced by panicle length (PanL; 0.255), panicle width (PanW; 0.343), total branches (TotBr; 0.240), and panicle diameter

(PanD; 0.281). As expected grain number per panicle (GNP) had the highest positive direct influence on grain yield (0.768). Yield was also indirectly affected by MaxBL via PanL (0.303), Rac via PanL (0.707) and TotN via Rac (0.409). In terms of relationships between panicle traits, direct relationships were identified for PanL with Rac (0.548), MaxLBZ (0.303), and PanW (0.106).

2.3.5. Principal component analysis of 206 accessions based on 6 different panicle types

To clarify the feature of accessions with the different panicle types and panicle traits, we used PCA analysis across 206 sorghum accessions. The central purpose of PCA is to reduce the dimensionality of a data set consisting of a large number of correlated variables, while retaining as much as possible of the variation present in the data set (Jolliffe 1986). This is achieved by identifying uncorrelated linear combinations of traits, the principal components (PCs), which are derived from the components of the eigenvectors of the phenotypic covariance or correlation matrix. The PC scores are calculated for each experimental unit by applying a characteristic linear combination of traits as indicated by the respective eigenvector. A scatter diagram of all accessions were made by their PC1 and PC4 score (**Figure. 2.3**). Along the PC1 axis, we considered whether the data separated based on any of the panicle types **(Figure. 2.3 A, B and Table 2.9)**. We observed that PC1 appears to separate the open type, some of intermediated type and broom type accessions from the others, but the remaining types were not separated by PC2, PC3 and PC4. Eigenvector analysis **(Table 2.9)** suggested that rachis length (Rac), panicle length (PanL), total branch number (TotBr), panicle diameter (PanD), maximum length of primary branch (MxLBZ) and panicle width (PanW) are largely responsible for this separation and are criteria characteristics of inflorescence architecture. Together this suggested that the panicle

measurements used in this study can at least partially capture the diversity of sorghum panicle architecture. For yield related traits PC1 the loadings of PanH, CulmL and GWP were substantial. For PC2 yield component trait such as PanN traits was substantial. The loadings of Rac, TotN, GNP and GWP were highest in the PC3. For PC4 TotN, PanW, PanH and CulmL traits were substantial. The first and second PC was a good measure of the morphological characters and yield traits. The first and second PC was both associated with the plant height, culm length, grain number and grain weight. Although eigenvalue of the third PC were small except for the GW trait contribution, this PC seemed to express 100 grain weight. The eigenvalues and contributions to the total variance of the first three PCs of the twelve panicle traits are shown in (**Table 2.9).** For the two growing seasons data analysis of PCA results revealed that the contribution of the PCA explained the total variation about (23.15 %, 12.43 %, 10.74 % and 9.41%) in 2010 (20.27%, 12.76%, 11.29% and 10.41%) in 2011, respectively. Between multi years 23.15% and 20.27% for the first two components for total accession among three groups. The highest contribution and positive value on PC1 with traits are PanL, GNP, GWP and TotBr as substantial (i.e., -30 > loadings > 0.30). The highest contributed trait on PC2 is Rac, PanL, PanD, PanN, MxLBZ and PanW. For PC3 the loadings of Rac and TotN trait were substantial. The loading of PanW, TotN, PanH and CulmL were highest in the PC4. Factor loading indicated that PC1 was positively correlated with PanL and TotBr. The trait variation in PC2 indicted that panicle size belonging to panicle diameter and negatively contributed with length traits such as Rac and PanL. The third PC was indicated that node number belonging to penicle elongation traits and negatively contributed with yield related traits. Factor loading indicated that PC4 was positively correlated with PanW and branching trait such as TotN.

The correlation matrix among 14 quantitative traits was examined using principal component analysis based on regional their origins (PCA) (**Table 2.9**) (**Figure 2.5-A,B,C**). In the PCA PC1, 2 and 3 showed no significant separation of 206 sorghum accessions. The eigenvalues and contributions to the total variance of the first three PCs of the fourteen panicle traits are shown in (**Table 2.9).** Along the PC1 axis, some of South, South East and Southwest Asian accessions were distributed and slightly overlapped with African accessions on the upper left side, while many East Asian accessions were scattered along the PC1. The PC1 and PC2 scores in other regions of Asia accessions showed but there was no distinct difference among regions. Examining PC1, 2 and 3 showed no significant separation based on regional origin, however, PC4 appeared to separate East Asian accessions from the others. The PC1 and PC4 scores few of other regions of Asia accessions are also slightly overlap with East Asia accessions. Most of East Asia accessions indicating that elongation traits tended to cluster (**Table 2.9**) and (**Figure. 2.4 A, B, C**).

2.3.6. Geographic distribution of different panicle types in 206 accessions and 107 accessions.

In this study we used the inflorescence traits revealed six panicle types and they varied from open/ loose panicle to compact elliptic panicle indicating a high diversity of sorghum accessions among three different origins. In the 206 accessions that highest percentage (Africa-46.3%, East Asia-21.3% and other regions of Asia-32.5%) of the accessions had compact type, open type (Africa-42.4%, East Asia-42.4% and other regions of Asia-15.2%) and intermediated type (Africa-37.5%, East Asia-35% and other regions of Asia-27.5%) when compare between the center of diversity area and edge of diversity areas **(Table 2.12 and Figure 2.7-A).** The percentage of other types of panicle such as broom type

(Africa-16.7%, East Asia-54.2% and other regions of Asia-29.2%) and mixed type (Africa-14.3%, East Asia-14.3% and other regions of Asia-71.4%) also are much different among different origins. The highest percentage 46.3% of compact panicle types are from Africa, 42.4% open type panicle are from East Asia and Africa, 54.2% broom type panicle are from East Asia and 71.4% mixed panicle types are from the other regions of Asia when compared between center of diversity and edge of diversity. Secondly we used previously selected sorghum diversity research set (SDRS) of 107 landraces and we categorized the accessions into three different groups (Africa-52.94%, East Asia-20.59% and other regions of Asia-26.47%). The inflorescence traits revealed six panicle type classes and they varied from open/ loose panicle to compact elliptic panicle indicating a high diversity of sorghum accessions among three different origins. Many types of panicle exist in Africa and Asia. Semi compact type (Africa-23.08%, East Asia-23.08% and other regions of Asia-53.85%), open type (Africa-76.47%, East Asia-11.76% and other regions of Asia-11.76%), (Africa-25%, East Asia-43.75% and other regions of Asia-31.25%) of the accessions had broom type and intermediated type (Africa-54.17%, East Asia-25% and other regions of Asia-20.83%) when compare between center of diversity area and edge of diversity areas **(Table 2.12 and Figure 2.7-B)**. Many types of panicle exist in both diversity areas, Africa and Asia. The highest percentage (Africa-52.17%) of the accessions had compact type, open type (Africa-76.47%), broom type East Asia-43.75% and other regions of Asia-53.85 with semi compact type when compare between center of diversity area and edge of diversity areas.

2.3.7. Phenotypic variations of panicle traits and other yield traits of sorghum diversity research set.

We examined the variability of the 11 quantitative traits related to inflorescence architecture and yield of sorghum to identify the characters responsible for the majority of the variation in the inflorescence architecture based on 107 core collection accessions. Phenotypic variations of these traits in 107 accessions are shown in (**Table 2.13 and Figure 2.8**). Eleven quantitative traits showed approximately normal distributions. The mean and range of the 11 traits showed a wide range of variation and the ranges were broad in 2010 growing season such as Pend, PanL, PanW, PanD,PanN and MaxLBZ (**Table 2.13**). In 2011 growing season most of traits were also widely distributed such as TotBr, Rac, PanL, PanW, PanN, MaxLBZ and GWN. The frequency distribution of the 8 traits didn't show obvious differences among 2 different growing seasons except for TotBr, PanD and GNP. Among these traits PanL, PanD, PanN and MaXLBz traits showed larger mean value in both of growing seasons. Analysis of variance indicated that the panicle traits, PanL PanD, PanN and MaxLBZ differed significantly at the 1% level among accessions and different seasons.

2.3.8. Phenotypic correlation of panicle traits and other yield traits of sorghum diversity research set

The correlation coefficients (*r*) were calculated for all traits. Testing the correlation of 11 quantitative traits with each other was shown in (**Table 2.14**). Focusing on the trait pairs correlation between 2 different growing seasons are presented in Table S4 showed significant ($p<0.001$) (**Table 2,14**). Ten traits showed high significant ($p<0.001$) correlation with another (the exceptions was PanN). In terms of panicle architecture, most of the length based measurements were correlated. In particular panicle length positively correlated with most of

the length based measurements (Rac, MaxLBZ, PanD), and yield components (GNP and GWP). Additionally the total number of branches (TotBr) and number of total nod (TotN) also correlated with PanL. Importantly, grain yield related traits were also positively correlated with most of the length based measurements. In particular grain weight per panicle (GWP) was correlated with panicle width (PanW; 0.49) in FY 2010 and (PanW; 0.53) in FY 2011, panicle diameter (PanD; 0.44) in FY 2011 and the non-length measurements, total branch number (TotBr; 0.38) in FY 2010, (TotBr; 0.46) in FY 2011. Also grain number per panicle (GNP) was correlated with the total branch number (TotBr; 0.39), panicle width (PanW; 0.43) and panicle length (PanL; 0.32). As expected grain number per panicle (GNP) was highly correlated with grain weight per panicle (GWP; 0.79), (GWP;0.67) in both seasons. Peduncle length was negatively correlated with yield related traits (GNP; -0.32 and GWP; -0.31).

2.3.9. Comparison of inflorescence type based on criteria characteristics

The six different panicle types such as open type, intermediated type, semi-compact type, compact, broom type and mixed types were screened to identify their external feature. The figure showed using the 8 characters of inflorescence architecture such as Pend, Rac, PanL, PanD, MaxLBZ, TotN, TotBr and PanW **(Figure 2.9)**. The mean values of these 8 traits in each panicle type were used to generate the Figure. Broom type and open panicle type are showed higher mean value of length based traits such as PanL. Morever total number branch (TotBr) showed obvious differences across all panicle types but where this trait did not show obvious differences between semi compact type and mixed panicle type.

2.3.10. Principal component analysis of 107 accessions based on 6 different panicle types

We considered whether the data separated based on any of the panicle types (**Table 2.15**). The highest contribution and positive value on PC1 with traits are Pend, PanW, GNP and GWP as substantial (i.e., -30 > loadings > 0.30) in 2010 and Pend, TotBr, PanD, PanW, PanEx and GWP in 2011. The loading of PanL, TotN, MaxLBZ, PanS and PanT in 2011were highest in the PC2 in 2010 and PanL, TotN, MaxLBZ, PanD were highest in the PC2 in 2011. For PC3 the loadings of Pand, Rac, PanS, PanT, PanD, PanEx and PanN were substantial. The loading of TotN, MaxLBZ, PanS, and PanT in Year-2010, Rac,TotN, MaxLBZ and GNP in 2011 were highest in the PC4. We found that PC1 appears to separate the open type and broom type accessions from the others, but the remaining types were not separated by PC2, PC3 and PC4. Eigenvector analysis (**Table 2.14**) suggested that panicle length (PanL), rachic length (Rac), panicle diameter (PanD), panicle width (PanW) and primary branch number (MaXLBZ), total grain number (GNP) and total grain weight (GWP) are largely responsible for this separation and are key characteristics of inflorescence architecture. Together this suggests that the panicle measurements used in this study can at least partially capture the diversity of sorghum panicle architecture.

2.4. Discussion

2.4.1. Variation of inflorescence architecture associated with yield component traits in 206 sorghum world wide germplasm.

Variation in sorghum inflorescence architecture is a not only the result of differences in panicle elongation but also different in the branching and panicle diameter (Brown *et al.,* 2006; Witt Hmon *et al.,* 2013). These characteristics are very helpful for breeding and

botanical purposes because these panicle characters can help to separate a varities and races within a sorghum species but still remain of great interest to breeders (Bala *et al*., 1996; Doust *et al*., 2002; Doust *et al*., 2005; Futsuhara *et al*., 1979a, 1979b; Kellogg 2000; Zhu *et al*., 2010). This study was performed to understand the diversity in inflorescence architecture of sorghum germplasm and to choose the plant materials for further study based on inflorescence architecture. Firstly, the distribution of several components of sorghum inflorescence architecture influenced yield components. We examined the correlations between the measured traits and we observed that many of the length based elongation measurements showed a positive correlation. In particular panicle width (PanW), plant height (PanH), rachis length (Rac), culm length (CulmL), maximum length of the primary branch zone (MaxLBZ) was positively correlated with panicle length (PanL) **(Table 2.5)** Interestingly panicle length also positively correlated with the number of panicles per plant (PanN) and the total number of branches (TotBr). This relationship between the patterns of branching traits and panicle characteristics are in agreement with another previous study (Vollbrecht *et al*., 2005) reported that inflorescence architecture comprises the stereotypical number and arrangement of floral branches in grasses including the domesticated cereals. In terms of the relationship between yield and panicle architecture we found that the width of the panicle (PanW), panicle diameter (PanD) and plant height (PanH) correlated with the grain yield per panicle (GWP). In addition the number of panicles per plant (PanN) was correlated with panicle width (PanW). Our results agreed with previous studies by Maman *et al*., (2004) and Saeed and Francis, (1983) who reported that yield per plant and head length was highly correlated. Path analysis indicated that panicle length had a positive effect on yield related traits. Moreover PanW, TotBr, PanD and GNP were also shown to strongly affect yield. We can classify the panicle types by using these traits **(Figure. 2.9).** This result revealed that the selection for panicle length, rachis length, total branch number, panicle

diameter and panicle width may improve grain yield. Comparisons between loose and compact inflorescence architecture there are the distribution of different characteristics have shown to be different associations (**Table 2.6 and Table 2.7**). This is because of the nature of the different panicle morphology among different races may be different. Among (12) quantitative traits many variations were observed for PanL, Rac, TotN, Rac, PenD, PanW and GNP to extract unique characteristics of the panicle structure of this sorghum germplasm. For loose inflorescence architecture PanL, TotBr and PanW were significantly associated with other characters. This result suggests these characters influenced on other characters among loose/semi loose panicle type. Negative correlation of TotBr with Pend trait can be explained that loose panicle accessions has the long peduncle with less number of total branch. Number of total branch along the rachis was significantly associated with PanL and other traits. The trait associations of PanL and TotBr are remarkable key character of loose inflorescence architecture. The association of various traits relating to grain yield and panicle traits was a little weak among loose panicle type accessions when compared with compact panicle pattern. The association of MaxLBZ, TotN and PanD can be clearly revealed that the maximum length of primary branch length decreased, the density of grain increased with thick head and also the number of total node (whole) is increase. These associations were indicated the elliptic, oval and short cylinder panicle type because these types are with short primary branch length, increase number of whole along rachis with dense panicle head and high grain density. The trait association of TotN and PanW can be express that the remarkable variation of compact panicle characters. Among compact panicle types there are panicle length, maximum length of primary branch, number of total branch, panicle diameter most frequently have shown to be related with other characters. This suggested that the other characters affected on PanL, MaxLBZ and TotN in compact inflorescence architecture. These trait associations showed remarkable variations in compact inflorescence architecture. The

correlation among traits showed the great importance for success in the selection units to be led to breeding programs. Correlation coefficients and path analysis indicated that the panicle length, total number of branches, rachis length and panicle width had a positive direct effect on grain yield. We found that the variation of sorghum inflorescence architecture is dependent not only on the panicle length, but also the total branch number, total node number, maximum length of primary branch, rachis length, panicle diameter and panicle width. There are major panicle determinants which are strongly associated with grain yield in breeding programs. According to these results TotBr, TotN, MaxLBZ, PanW, PanL and Rac traits should be consider as new preliminary information of sorghum inflorescence architecture to emphasize yield improvement. Moreover, among the panicle trait combination the emphasizing of trait selection is lacking especially the elongation trait TotN and branching trait MaxLBZ. These traits should be considered for trait selection of morphological diversity in modern sorghum plant breeding. Analysis of the panicle diversity using 206 accessions from germplasm collection we used 18 measured traits to attempt to capture the separation of varieties based on panicle type, the first component of the PCA was able to partially separate the broom panicle and open panicle types from the others **(Figure 2.4)**. The first PC was a good measure of the elongation traits of inflorescence architecture. The second PC component for attributions was associated with the diameter. Although eigenvalue of the third PC were small, this PC seemed to node number along main axis and their exsertion pattern related with rachis length and yield components. This suggested that the measurements were at least partially captured the diversity of panicle architectures. The traits associated with this separation were mainly the length based traits suggesting that they were key determinants for describing panicle diversity. Additional traits such as TotBr and TotN are also important for further study on the structure of the panicle, Our results agreed with the previous suggestion by Ikeda *et al.,* (2010) that TotBr and TotN traits were important for

future advanced methods that automatically captured the architecture using image analysis is useful in the future. We set out to measure the intraspecific variation of panicle traits across a large collection of 206 sorghum accessions to identify the relationship between these traits and yield components. To do this we used a set of measurements to approximate panicle architecture, including length and branching measurements. We showed that not only many of length based measurements but also elongation based measurements are correlated with yield related traits, which may suggest a common genetic regulation, and that several traits are likely to influence yield.

2.4.2. Geographical patterns of inflorescence variation

We examined in panicle diversity using 206 accessions assessed from a germplasm collection that covers most of the diversity of geographic origins and panicle types known in sorghum. We found that the patterns of observed panicle traits reflected the distribution of different origins. We used 18 traits to capture the diversity in inflorescence architecture. Among the sorghum accessions from Africa, the center of diversity of sorghum, represented open type, intermediate type, compact type, semi compact type and mixed type. Africa, the center of diversity area, represents highest variation when compare with the other diversity area. The most of the variability among main panicle traits has different clusters between other region of Asian accessions and East Asian accessions and African accession. These results can agree with the previous study (Shehzad *et. al*., 2009a) who reported that sorghum bicolor does not have a characteristic geographical distribution between any region of Asia and Africa because it was introduced from Africa to East Asia through South Asia. Our results are an agreement with the previous report by Brown *et al.,* (2011) that among five main races, S. *bicolor* can grow everywhere in Africa and Asia without any separation of

geographic origin or adaptation level. Inflorescence architecture in most of African accessions was more or less compact panicle type and open panicle type if compared with other regions of Asian accessions and East Asian accessions. In sorghum five main races (Bicolor, Guinea, Caudatum, Durra and Kafir) can be identified based on their critical characteristics. According to their divisions identification system, our results can agree with the previous report by Brown *et al.,* (2011) and Casa *et al.,* (2008) who reported that sorghum races would be classified based on panicle and spikelet in genotype-based classification by using structure analysis except bicolor race. Our results shown in this chapter the distribution of the broom panicle type was less frequent among African accessions. The distribution of the open panicle type was less frequent among the other diversity regions, (Southeast Asia, Southwest Asia and South Asia). The semi compact, intermediate, broom and mixed panicle types were found more or less nearly equally among the accessions from three different regions. In this chapter total 206 accessions generally formed groups classified based on inflorescence architecture. Analysis of variance (ANOVA) could clearly explain significant differences among accessions, characters and regions suggesting that this sorghum population was highly variable for almost all traits measured. ANOVA could identify significant differences between East Asian and other Asian accessions for half of these traits indicated with the African accessions tending to have shorter penducle length (Pend), more nodes (TotN) and longer rachis length (Rac). Despite this, principle component analysis was only able to partially separate of the East Asian varieties from the African and other regions of Asian accessions. This suggested there is lower diversity in the East Asian varieties which agrees with an African origin of sorghum and that early domestication in Africa and Western Asia lead to a higher diversity (Harlan *et al*., 1976; House, 1985; Shehzad *et al*., 2009a). According to our results the open panicle types may have originated as an adaptive trait to allow quick drying of panicles in high humidity environments and minimize damage by

fungal diseases. The guinea race provide less/open panicle type. Harlan *et al*., (1976) reported that the guineas to be the oldest of the specialized races because of its relatively wide distribution and diversity. Our result indicated that the open panicle type was not only widely distributed in the center of diversity areas, Africa, but also more or less frequently distributed in the edge of diversity areas, East Asia. According to Stemler *et al.,* (1977), caudatum is a later domesticated race than bicolor and guinea. Our result indicated the compact panicle type was widely distributed in the center of diversity, Africa. The guinea race had compact panicle types. The compact panicle type with curve panducle and predominantly white seeds of race durra can be adapted to low-rainfall environments with a low risk of grain mold (Mann *et al.,* 1983). It is an important type in India and may have been domesticated there (Harlan *et al*., 1976). Until recently, the durra is grown in the Islamic and Hindu areas of India and Pakistan as an important panicle type, South Asia. The result on the diversity of compact and semi-compact panicle type agreed with these earlier reports because the semi-compact, compact and intermediate panicle types were found nearly equally among the accessions from different regions, other region of Asia and East Asia. The previous reports determined that broom type sorghum has a different story among sorghum panicle types, as are the sorgos such as amber cane. Our result agreed with these previous reports with the distribution of the broom panicle type was less frequent among African accessions. In Southeast Asia and Indonesia, the sorghum is different as well. *S. propinquum* is found in South China through Thailand, Cambodia, Malaya, and Myanmar to the Philippines. These sorghum types were characterized by very large, loose, open panicles and might also have a history different from those of the African based races (Harlan and de Wet 1972; Doggett 1988). These previous reports can agree with our results because the distribution of the mixed panicle type was very high frequent among other regions of Asian accessions (South Asia, Southeast Asia and Southwest Asia). However, our results in this chapter agreed with previous evidences in the

complex species S. *bicolor* included all cultivated sorghum as well as semi-wild plants mostly associated with them as weeds. According to the results in this chapter we could observed that sorghum panicle types distributed throughout the African and Asian regions most likely vary according to different climate conditions, temperature, humidity and rainfall patterns. The diversity in shape and compactness is likely to indicate selection for varieties that can survive in different local environments, and is largely independent of grouping by continent. Moreover, we analyzed the variability in inflorescence architecture in a wide range of sorghum germplasm. In addition, the total number of nodes (TotN) was largely responsible for separation among different origins by eigenvector analysis (PCA analysis). As a result of this trait would consider for preferences in trait selection among the complex sorghum inflorescence traits. It is worth noting that we were unable to separate African from the Other Asian varieties by the PCA analysis. However, cluster and scatter plot analysis identified that the pattern of distribution of the inflorescence (panicle) traits reflected the distribution of different origins and it can exhibits a great range of phenotypic variation in inflorescence architecture. By applying the cluster methods, selection of parental lines with desirable panicle traits would be achieved. This information generated from this study would allow selecting for appropriate sorghum materials for further breeding program.

2.4.3. Variation of inflorescence architecture associated with yield component traits in sorghum diversity research set.

The correlation coefficient contingency table revealed a very intense character association such as PanL, PanD, PanW, MaXLBZ and those panicle traits were significantly associated with most of the characters. Examining the correlations between the measured traits we observed that many of the length based elongation measurements showed positive

correlation. In terms of the relationship between yield and panicle architecture we found that the width of the panicle (PanW), panicle length (PanL), total number of branch (TotBr) and peduncle length (Pend) were correlated with the grain yield (GWP). Our results agreed with previous studies by (Maman *et al*., 2004) and (Saeed and Francis 1983) who reported that yield per plant and head length was highly correlated. Interestingly panicle diameter also positively correlated with the grain number (GNP) and grain weight (GWP). Our results for grain yield are in agreement with another previous study by (Maman *et al*., 2004). GWP and GWP were also significantly correlated with PanD and PanL even high coefficients of genotypic correlation between PanL and other traits such as, total number of branch per panicle have been reported by many authors (Brown *et al,* 2006; Srinivas *et al.,* 2009). When we explained the separation of varieties based on panicle type, the first component of the PCA was able to partially separate the broom panicle and open panicle types from the others. This suggests that the measurements used can at least partially capture the diversity of panicle architectures. The traits responsible for this separation were mainly the length based traits; rachis length (Rac), panicle length (PanL), panicle diameter (PanD), panicle width (PanW) and primary branch length (MaXLBZ/MxLBZ), suggesting they are key determinants for describing panicle diversity. Additional traits such as Pend and TotN are also important in further describing the structure of the panicle, and advanced methods that automatically capture the architecture using image analysis will be useful in the future (Ikeda *et al.* 2010).

Table 2.1. List and origin of 206 sorghum (*Sorghum bicolor* L. Moench) accessions from NIAS, Genebank.

No.	Geography/Group	Cultivar Name	Origin
1	East Asia	OOTOYO-MURA ZAIRAI	Japan
2	East Asia	TAKAKIMI	Japan
3	East Asia	IKEDACHO MATSUO ZAIRAI	Japan
4	East Asia	KOUCHI OUKAWA ZAIRAI	Japan
5	East Asia	DANGOMOROKOSHI	Japan
6	East Asia	TOKIBI	Japan
7	East Asia	HIMEKI ZAIRAI	Japan
8	East Asia	KIKUCHI ZAIRAI	Japan
9	East Asia	AKAHO	Japan
10	East Asia	KANAGAWAZAIRAI	Japan
11	East Asia	72-10-10-5	Japan
12	East Asia	TAKAKIBI	Japan
13	East Asia	KOUCHI MONOBE ZAIranAI	Japan
14	East Asia	COL/NAGASAKI/1994/MAFF/114	Japan
15	East Asia	KIBI OGASAWARAZAIranAI	Japan
16	East Asia	YATABU	Japan
17	East Asia	CHAKIBI	Japan
18	East Asia	KIBI	Japan
19	East Asia	KIBI	Japan
20	East Asia	TOUHOKUZAIranAI 35	Japan
21	East Asia	CHIBAKURO	Japan
22	East Asia	ZAIranAISHU 51-12	Japan
23	East Asia	KIKUCHI ZAIranAI KUMA 101	Japan
24	East Asia	TOKUSHIMA ZAIranAI	Japan
25	East Asia	AKAKIBI	Japan
26	East Asia	76-7-31	Japan
27	East Asia	MOROKOSHI	Japan
28	East Asia	72-10-8-2	Japan
29	East Asia	73-10-25-9	Japan
30	East Asia	HANGETSUTOSUI	Korea
31	East Asia	KOUSHUU ZAIRAISHU	Korea
32	East Asia	CHAL WAXY SORGHUM	Korea
33	East Asia	KOUBOUSHI	Korea
34	East Asia	MOCTAC LOCAL	Korea
35	East Asia	SENKINHAKU	Korea
36	East Asia	CHOONCHAN LOCAL	Korea
37	East Asia	KOKKAKU SOHANSHIN	Korea
38	East Asia	KOKKAKU 2	Korea
39	East Asia	KOKKAKU SOUSHINHAN	Korea
40	East Asia	72-8-13	Taiwan
41	East Asia	AI HUI	China
42	East Asia	NUO GAO LIANG	China
43	East Asia	ER BAI SHE YAN	China

Table 2.1. continued

No.	Geography/Group	Cultivar Name	Origin
44	East Asia	LIAOZA 1	China
45	East Asia	BIG WHITE HULL	China
46	East Asia	XIONG YUE 334	China
47	East Asia	DA ZHAI LIN HUANG JIAO GAO LIANG	China
48	East Asia	KAI 64	China
49	East Asia	TAISAIranINOUKOUKOURYAN	China
50	East Asia	HU 4	China
51	East Asia	CHI 7321	China
52	East Asia	HU 22	China
53	East Asia	TA HUNG KAK	China
54	East Asia	GAO GAN NUO GAO LIANG	China
55	East Asia	NUO XIAO GAO LIANG	China
56	East Asia	AI GAO LIANG	China
57	East Asia	HUNGPANTSE	China
58	East Asia	TENKOURYAN	China
59	East Asia	DAIranYUKO(KATSUZAN)	China
60	East Asia	WHITE BIG BELLIED	China
61	East Asia	JI 7384	China
62	East Asia	XIONG YUE 253	China
63	East Asia	BIG YELLOW AMBRELLA 250	China
64	East Asia	WHITE SORGHUM	China
65	East Asia	GOLDEN LIGHT SORGHUM	China
66	East Asia	KORYANKAI 64	China
67	Southeast Asia	BATTANBAN	Cambodia
68	Southeast Asia	AS 5781 HUAN SA PHAUNG AH LPYSU	Myanmar
69	South Asia	Y. E. (I. P.) INT. TYPE	India
70	South Asia	KALJANPUR	India
71	South Asia	SC NO.0217 CI1197	India
72	South Asia	GOOSENECK	India
73	South Asia	MARIANGARIJORA MUDDAHIHAL	India
74	South Asia	DHOOTI ANEHULA	India
75	South Asia	RABI YANGAR JORA MITHUGADUR	India
76	South Asia	AS 4136 MASAKA LUWEMEA	India
77	South Asia	M.35-1 DODDA MAGADI	India
78	South Asia	JHANJHARALA	India
79	South Asia	SA 9798 Y. E. KAFIran	India
80	South Asia	SWEET JOWAR SELECTED	India
81	South Asia	PI 248293	India
82	South Asia	JALWARA	India
83	South Asia	M 35-1 DODDA MAGADI	India
84	South Asia	GOKUWASE INDOSHU	India
85	South Asia	IS 8722	India

Table 2.1. continued

No.	Geography/Group	Cultivar Name	Origin
86	South Asia	EC 21434 SB 77	India
87	South Asia	SA 9020 A	India
88	South Asia	MN 405	India
89	South Asia	119475 COL/PAK/1989/IBPGR/2386(2)	Pakistan
90	South Asia	119513 COL/PAK/1991/IBPGR/2724(2)	Pakistan
91	South Asia	119485 COL/PAK/1989/IBPGR/2420(1)	Pakistan
92	South Asia	119487 COL/PAK/1989/IBPGR/2427(5)	Pakistan
93	South Asia	119488 COL/PAK/1989/IBPGR/2439(1)	Pakistan
94	South Asia	119489 COL/PAK/1989/IBPGR/2444(1)	Pakistan
95	South Asia	119494 COL/PAK/1989/IBPGR/2550(1)	Pakistan
96	South Asia	119496 COL/PAK/1989/IBPGR/2553(4)	Pakistan
97	South Asia	COL/PAK/1989/IBPGR/2411(1)	Pakistan
98	South Asia	COL/PAK/1989/IBPGR/2416(2)	Pakistan
99	South Asia	COL/PAK/1989/IBPGR/2592(7)	Pakistan
100	South Asia	87-9-21-3-1	Pakistan
101	South Asia	87-9-21-3-2	Pakistan
102	South Asia	COL/PAK/1989/IBPGR/2390(2)	Pakistan
103	South Asia	COL/PAK/1989/IBPGR/2391(1)	Pakistan
104	South Asia	COL/PAK/1989/IBPGR/2410(1)	Pakistan
105	South Asia	COL/PAK/1989/IBPGR/2549(1)	Pakistan
106	South Asia	COL/PAK/1991/IBPGR/2748(7)	Pakistan
107	South Asia	COL/PAK/1989/IBPGR/2590(2)	Pakistan
108	South Asia	RED TYRI	Pakistan
109	South Asia	COL/PAK/1989/IBPGR/2272(1)	Pakistan
110	South Asia	COL/PAK/1989/IBPGR/2369(1)	Pakistan
111	South Asia	COL/PAK/1989/IBPGR/2371(4)	Pakistan
112	South Asia	COL/PAK/1989/IBPGR/2574(1)	Pakistan
113	South Asia	COL/PAK/1989/IBPGR/2581(1)	Pakistan
114	South Asia	COL/PAK/1989/IBPGR/2587(2)	Pakistan
115	South Asia	COL/PAK/1989/IBPGR/2543(2)	Pakistan

Table 2.1. continued

No.	Geography/Group	Cultivar Name	Origin
116	South Asia	COL/PAK/1989/IBPGR/2550(2)	Pakistan
117	South Asia	COL/PAK/1989/IBPGR/2593(10)	Pakistan
118	South Asia	COL/PAK/1989/IBPGR/2345(2)	Pakistan
119	South Asia	COL/PAK/1989/IBPGR/2580(2)	Pakistan
120	South Asia	COL/PAK/1989/IBPGR/2598(1)	Pakistan
121	South Asia	87-9-20-2-1	Pakistan
122	South Asia	87-9-20-4-1	Pakistan
123	South Asia	87-9-21-3-3	Pakistan
124	South Asia	COL/PAK/1989/IBPGR/2416(2)	Pakistan
125	South Asia	COL/PAK/1989/IBPGR/2592(7)	Pakistan
126	South Asia	ALLAKH	Bangladesh
127	South Asia	EC 18868	Nepal
128	South Asia	JUNELO	Nepal
129	Southwest Asia	PI 229486 VULGARE	Iran
130	Southwest Asia	HAZERA 6014	Israel
131	Africa	E 9	Chad
132	Africa	PI 282834	Chad
133	Africa	E 17	Congo
134	Africa	MAKHOTLONG I	Lesotho
135	Africa	TENANT WHITE	Lesotho
136	Africa	NYAKASOBA BEST	Lesotho
137	Africa	MAKHOTLONG II	Lesotho
138	Africa	AIT BRAHIM	Morocco
139	Africa	CODY	Morocco
140	Africa	KOURNIANIA	Morocco
141	Africa	PHATSAI	Morocco
142	Africa	SCHROCK	Morocco
143	Africa	ESHOME	South Africa
144	Africa	E 232 INGWARUMA PEARLY	South Africa
145	Africa	AW 70/12 DL/59/1532	South Africa
146	Africa	E 233 BARNARD RED	South Africa

Table 2.1. continued

No.	Geography/Group	Cultivar Name	Origin
147	Africa	RED KAFIR	South Africa
148	Africa	S.BASUTORUM DL/60/97	South Africa
149	Africa	EAR FROM PIETESBURG DL/60/107	South Africa
150	Africa	RADAR	South Africa
151	Africa	SWAZI RED	South Africa
152	Africa	WHITE KAFIran	South Africa
153	Africa	S. SACCHARATUM DL/59/1544 SUGARDIP	South Africa
154	Africa	PINK KAFIran	South Africa
155	Africa	WAIN DL/60/760	South Africa
156	Africa	1824 NTULI RED EX SWAZIsraelAND L/60/13	South Africa
157	Africa	E 238 BIranD PROOF	South Africa
158	Africa	SOGALANE	South Africa
159	Africa	MILO PET. 139/51 EX TANGANYIKA	Central Africa
160	Africa	117 SBI 20	Central Africa
161	Africa	HEGARI MALOWAR	Sudan
162	Africa	143 DINDERAWI 1	Sudan
163	Africa	REDBINE 655	Sudan
164	Africa	E 1089	Sudan
165	Africa	LAMBAS	Sudan
166	Africa	DINDERAWI 1	Sudan
167	Africa	240 WAD UMM BENEIN	Sudan
168	Africa	MUGBASH WHITE	Sudan
169	Africa	B-112	Sudan
170	Africa	E 1091	Sudan
171	Africa	109 TONJI	Sudan
172	Africa	E 1094	Sudan
173	Africa	ACCA KODRI 30	Sudan
174	Africa	ZA113 DAWA PAS PARA	Nigeria
175	Africa	AS 4547 JARDIRA	Nigeria
176	Africa	MN 1277 MUHEYAR	Nigeria
177	Africa	KA 24	Nigeria
178	Africa	PI 221543 Q 2/3/61	Nigeria

Table 2.1. continued

No.	Geography/Group	Cultivar Name	Origin
179	Africa	BC 59 MERE	Nigeria
180	Africa	AS 4548 TI2 YANDER	Nigeria
181	Africa	PI 221615	Nigeria
182	Africa	MN 401	Algeria
183	Africa	S. VULGARE 72-726-7	Uganda
184	Africa	S. VULGARE 72-728-1	Uganda
185	Africa	E 276 FRAMIDA	Uganda
186	Africa	UGANDA L1	Uganda
187	Africa	E 75	Uganda
188	Africa	E 83	Uganda
189	Africa	S. VULGARE 71-728-3	Uganda
190	Africa	E 67	Uganda
191	Africa	MORABA 74	Ethiopia
192	Africa	THIBA RED	Ethiopia
193	Africa	SC112	Ethiopia
194	Africa	GIZA 3/59	Ethiopia
195	Africa	PI 329762	Ethiopia
196	Africa	AKLMOI WHITE	Kenya
197	Africa	E 959	Kenya
198	Africa	PI 152748 C	Kenya
199	Africa	WAD YABOO 132/53	Zimbabwe
200	Africa	CAPE COLO 28/53	Zimbabwe
201	Africa	TSETA LOCAL NATURE TYPE 27/51	Zimbabwe
202	Africa	CAPE CALO 28/53 EX	Zimbabwe
203	Africa	WADYABGO	Zimbabwe
204	Africa	AS 5885 MALA MATUBA	Zimbabwe
205	Africa	AS 4637 NHORONGO NENPI	Tanzania
206	Africa	E 37	Tanzania

Table 2.2. List of traits and sorghum descriptors analyzed.

No.	Traits/ character	Trait abbreviation	Evaluation method	Type of assessment	Rank	Descriptor referred
Elongation/length based traits						
1	Plant total height	PanH (cm)	From ground to the tip of panicle at 50% flowering (Length of main stalk)	MS	maturity	IBPGR/ICRISAT
2	Culm length	CulmL (cm)	Length from the ground to the neck node of	MS	maturity	NIAS
3	Penducle length	Pend (cm)	Neck length emerged from flag leaf sheath to the base of panicle	MS	maturity	IBPGR/ICRISAT
4	Rachis length	Rac (cm)	Length from the neck node to the base of top branching zone	MS	maturity	NIAS
5	Panicle length	PanL(cm)	Length from the base of the head to the tip	MS	maturity	NIAS
6	Maximum length of primary branch	MaxLBZ/ MxLBZ(cm)	Length of branches (middle third of panicle)	MS	maturity	IBPGR/ICRISAT
Branching traits						
7	Number of node	TotN (number)	Total node on the rachis	MS	maturity	IBPGR/ICRISAT
8	Total branch	TotB (number)	Total primary branch on the rachis	MS	maturity	IBPGR/ICRISAT
Dimentional traits						
9	Panicle diameter	PanD (mm)	natural position at the widest part of head	MS	maturity	IBPGR/ICRISAT
10	Panicle width	PanW (mm)	natural position at the thickness of head	MS	maturity	IBPGR/ICRISAT
11	Panicle shape	PanS (shape)	Inflorescence shape at maturity	VG	maturity	IBPGR/ICRISAT
12	Panicle type	PanT (type)	Inflorescence compactness at maturity		maturity	NIAS

Type of assessment of characteristics indicated in: MG : Measurement by a single observation of a group of plants or parts of plants, MS : Measurement of a number of individual plants or parts of plants, VG : Visual assessment by a single observation of a group of plants or parts of plants, VS : Visual assessment by observation of individual plants or parts of plants Descriptors Referred: National Institute of Aerobiological Sciences (NIAS), The International Board for Plant Genetic Resources (IBPGR), International Crops Research Institute for the Semi-Arid Tropics (ICRIST) (1993).

Table 2.2. (continued)

No.	Traits/ character	Trait abbreviation	Evaluation method	Type of assessment	Rank	Descriptor referred
Yield related traits						
13	Panicle number	PanN (number)	Number of mature panicles per plant	MG	Harvesting	NIAS
14	Grain number per panicle	GNP (number)	Number of grain per panicle on main stem	MG	after threshing	IBPGR/ICRISAT
15	Grain weight per panicle	GWP (g)	Weight of cleaned grains per panicle on main stem	MG	after threshing	NIAS
16	100 grain weight	GW (g)	Weight of 100 grain	MG	after threshing	ICRISAT
Miscellaneous						
17	Panicle exsertion	PanEx (stage)	position of peduncle emergence between flag leaf and inflorescence based	MS	maturity	IBPGR/ICRISAT
18	Panicle broader	PanB (stage)	Position of on panicle broader at maturity	VG	maturity	NIAS

Type of assessment of characteristics indicated in: MG : Measurement by a single observation of a group of plants or parts of plants, MS : Measurement of a number of individual plants or parts of plants, VG : Visual assessment by a single observation of a group of plants or parts of plants, VS : Visual assessment by observation of individual plants or parts of plants Descriptors Referred: National Institute of Aerobiological Sciences (NIAS), The International Board for Plant Genetic Resources (IBPGR), International Crops Research Institute for the Semi-Arid Tropics (ICRIST) (1993).

Table 2.3. Analysis of variance (ANOVA) and value of 14 quantitative traits in 206 sorghum accessions.

Traits	TotBr	Pend	Rac	PanL	PanW	TotN	PanD	MxLBZ	PanH	CulmL	PanN	GNP	GWP	GW
Year-1 (FY 2010)														
Mean	45.93	11.05	16.99	20.44	0.18	7.78	3.46	9.26	151.6	131.8	3.24	1175.97	22.19	2.23
(±SD)	(±15.6)	(±6.7)	(±6.7)	(±7.8)	(±0.2)	(±2.5)	(±2.4)	(±7.4)	(±57.5)	(±55.6)	(±1.4)	(±1018.5)	(±20.5)	(±2.9)
Year -2 (FY 2011)														
Mean	16.79	9.91	16.79	21.98	0.23	6.96	3.82	10.29	181.12	159.18	1.89	930.26	13.99	1.56
(±SD)	(±6.4)	(±5.6)	(±6.4)	(±7.5)	(±0.1)	(±2.2)	(±1.6)	(±8)	(±60.3)	(±58.3)	(±1.1)	(±646.7)	(±10.6)	(±0.7)
ANOVA														
F Ratio	3.998	9.1265	0.5273	0.3045	39.9684	2.2282	35.9739	1.3293	0.4581	0.4369	37.5922	40.4284	80.7396	311.389
P-Value	0.0456*	0.0025**	0.4678	0.5811	<.0001***	0.1355	<.0001***	0.2489	0.4985	0.5086	<.0001***	<.0001***	<.0001***	<.0001***

* , * *,* ** signifi cantly different at P< 0.05 and P < 0.01, P < 0.001 respectively; *Legend for quantitative trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant height, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle,GW:100 grain weight, PanN: panicle number per plant. *Note: Qualitative traits*; PanS: panicle shape, PanT: panicle type, PanEx: panicle exsertionand PanB: panicle broader were also recorded.

Table 2.4. Analysis of variance (ANOVA) and value of 14 quantitative traits across 3 different regions.

Origins	East Asia		Other region of Asia		Africa		ANOVA	
	M (±SD)						p-value	p-value
Trait	2010	M (±SD) 2011	M (±SD) 2010	M (±SD) 2011	M (±SD) 2010	M (±SD) 2011	(2010)	(2011)
Pend	12.7 (±7)	11.7 (±6.5)	12.6 (±6)	10.9 (±4.8)	8.4 (±6)	7.5 (±4)	<.0001***	<.0001***
Rac	17.1 (±7)	17.4 (±5.8)	14.7 (±6)	15.5 (±6.8)	18.7 (±6)	18.6 (±5.2)	0.0019**	0.0093**
PanL	20.7 (±7)	23.6 (±7.4)	19.4 (±10)	19.1 (±8.2)	21 (±6)	23.1 (±6.3)	0.4596	0.0008**
PanH	149.1 (±56)	196.1 (±58.4)	153.5 (±53)	175.7 (±57.5)	152.2 (±63)	170.4 (±60.2)	0.9030	0.0220*
CulmL	128.5 (±54)	173.9 (±56)	134.3 (±50)	159.7 (±54.5)	132.6 (±62)	143.8 (±58.4)	0.8932	0.0069**
MxLBZ	9.8 (±4)	10.3 (±4.2)	9.2 (±10)	8.6 (±9.4)	8.8 (±7)	9.3 (±4.5)	0.8351	0.3433
TotN	6.9 (±2)	6.4 (±2.6)	8 (±2)	6.8 (±2.2)	8.3 (±3)	7.1 (±1.9)	0.0034**	0.1520
TotBr	42.1 (±14)	47.4 (±15.4)	47.7 (±17)	55.2 (±25.7)	47.2 (±16)	47.4 (±11.3)	0.0782	0.0152*
PanW	0.4 (±0.2)	2 (±1)	0.5 (±0.2)	2.3 (±1.1)	0.5 (±0.2)	2.6 (±0.7)	0.0497	0.0040**
PanD	2.3 (±1.4)	4.2 (±2.2)	2.9 (±1.6)	3.6 (±1.3)	2.9 (±1.2)	3.8 (±1)	0.0341	0.1069
PanN	3.8 (±1)	1.4 (±0.6)	2.9 (±1.6)	1.8 (±3.7)	4.5 (±4.1)	1.5 (±1.5)	0.3236	0.4781
GNP	1120 (±911.8)	1048.8 (±585.9)	1085.8 (±1090.5)	810.6 (±721.6)	1448.8 (±1855.4)	974.6 (±664.4)	0.4969	0.1073
GWP	19.8 (±18.8)	16.8 (±11.7)	21.8 (±25.1)	11.9 (±11)	24.5 (±17.3)	13.3 (±8.9)	0.3915	0.1900
GW	2.2 (±3.5)	1.6 (±0.6)	2 (±1.2)	1.6 (±0.8)	2.5 (±3.4)	1.5 (±0.6)	0.5020	0.4770

* , * *,* ** signifi cantly different at P< 0.05 and P < 0.01, P < 0.001 respectively; *Legend for quantitative trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant height, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle,GW:100 grain weight, PanN: panicle number per plant. *Note: Qualitative traits*; PanS: panicle shape, PanT: panicle type, PanEx: panicle exsertion were also recorded.

Table 2.5. Correlation coefficient among quantitative traits of 206 sorghum accessions.

Traits	*Pend*	*Rac*	*PanL*	*PanH*	*CulmL*	*MxLBZ*	*TotN*	*TotBr*	*GNP*	*GWP*	*GW*	*PanD*	*PanN*	*PanW*
Pend	1	-0.08	-0.09	0.08	0.13	-0.13	0.11	-0.07	-0.02	0.01	0.00	0.11	0.11	-0.03
Rac	-0.20**	1	0.44***	0.03	-0.06	0.23***	0.31***	0.01	0.12	0.12	0.00	0.04	0.16*	0.00
PanL	-0.09	0.52***	1	0.24***	0.12	0.33***	-0.18	-0.10	0.10	0.11	-0.01	0.23***	0.02	0.01
PanH	0.01	0.17*	0.36***	1	0.94***	0.09	-0.01	0.30***	0.17*	0.24***	0.15	0.38***	0.22	-0.02
CulmL	0.01	0.1	0.23***	0. 99***	1	0.02	-0.01	0.32***	0.12	0.22**	0.19	0.34***	0.25***	-0.06
MxLBZ	-0.11	0.05	0.46***	0.15*	0.09	1	-0.12	0.00	0.09	0.05	-0.06	0.15*	-0.03	0.13
TotN	0.09	0.29***	-0.01	-0.01	0.01	-0.17*	1	0.18*	0.04	0.10	0.10	-0.10	0.02	-0.02
TotBr	-0.20**	0.25***	0.33***	0.30***	0.26***	0.16*	0.12	1	0.02	0.08	0.06	0.25***	0.01	0.32***
GNP	-0.17*	0.08	0.14	0.18**	0.19**	0.07	-0.06	0.14	1	0.72***	-0.16	0.17*	0.00	0.05
GWP	-0.16*	0.12	0.22**	0.23***	0.22**	0.06	-0.05	0.21**	0.60***	1	0.30***	0.23***	0.17*	0.00
GW	-0.11	-0.01	-0.01	-0.05	-0.05	-0.08	-0.01	0.01	-0.09	0.19**	1	0.06	0.18*	-0.04
PanD	-0.08	-0.03	0.08	0.13	0.14*	0.04	-0.01	0.17*	0.21**	0.27***	0.03	1	0.11	0.32***
PanN	-0.09	0.14*	0.24***	0.23***	0.21**	0.07	-0.05	0.15*	0.1	0.18*	0.06	0.22**	1	-0.21*
PanW	-0.18*	0.17*	0.31***	0.09	0.06	0.11	-0.08	0.13	0.23***	0.29***	0.04	0.31***	0.27***	1

*P<0.05, ** P<0.01, ***P<0.001, respectively. (Upper diagonal= The correlation coefficient among traits for 2011 growing season, Lower diagonal = The correlation coefficient among traits for 2010 growing season). *Legend for quantitative trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MxLBZ (MaxLBZ): maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant height, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanN: panicle number per plant.

Table 2.6. Correlation coefficient among traits in loose inflorescence architecture and compact inflorescence architecture in year 2010.

Traits	Pend	Rac	PanL	PanH	CulmL	MxLBZ	TotN	TotBr	GNP	GWP	GW	PanD	PanN	PanW
Pend	1	-0.22**	0.05	-0.02	-0.02	-0.06	0.06	-0.14*	-0.15*	-0.15*	-0.11	-0.15*	-0.09	-0.22**
Rac	-0.16	1	0.54***	0.07	0.01	-0.06	0.34**	0.19*	0.11	0.19*	0.01	0.01	0.13	0.24**
PanL	-0.20	0.47**	1	0.30***	0.18*	0.56***	-0.13	0.25**	0.12	0.20**	-0.03	0.14*	0.32***	0.35**
PanH	0.09	0.42**	0.51***	1	0.99***	0.26**	-0.04	0.29**	0.25**	0.22**	-0.07	0.17*	0.39***	0.06
CulmL	0.13	0.38**	0.38***	0.98***	1	0.19*	-0.02	0.27**	0.25**	0.21**	-0.07	0.17*	0.37***	0.03
MxLBZ	-0.24*	0.23 *	0.25*	-0.07	-0.12	1	-0.39**	0.13	0.07	0.08	-0.08	0.13	0.20**	0.21**
TotN	0.13	0.21*	-0.16	0.05	0.08	0.15	1	0.14*	-0.03	0.01	0.01	-0.03	-0.04	-0.08
TotBr	-0.37**	0.46**	0.57***	0.30*	0.23*	0.28*	0.11	1	0.11	0.16*	-0.03	0.20**	0.14*	0.12
GNP	-0.11	0.02	0.24*	0.31*	0.29*	0.11	-0.12	0.23*	1	0.79***	-0.11	0.27**	0.09	0.21**
GWP	-0.17	-0.05	0.35**	0.28*	0.24*	0.06	-0.15	0.36**	0.89**	1	0.18*	0.30***	0.22**	0.26**
GW	-0.11	-0.06	0.12	0.04	0.02	-0.08	-0.03	0.21*	-0.01	0.33*	1	0.03	0.08	0.02
PanD	0.1	-0.04	0.01	0.04	0.04	-0.08	0.03	0.09	0.14	0.19	0.01	1	0.37**	0.29**
PanN	-0.09	0.16	0.15	-0.03	-0.05	-0.08	-0.06	0.16	0.13	0.12	0.03	0.03	1	0.22**
PanW	0.09	0.04	0.36***	0.22*	0.17	-0.09	-0.03	0.1	0.35**	0.38**	0.08	0.37**	0.51***	1

*P<0.05, ** P<0.01, ***P<0.001, respectively. (Upper diagonal= The correlation coefficient among traits in *compact inflorescence architecture*, Lower diagonal = The correlation coefficient among traits in *loose inflorescence architecture*). Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MxLBZ (MaxLBZ): maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant height, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanN: panicle number per plant.

Table 2.7. Correlation coefficient among traits in loose inflorescence architecture and compact inflorescence architecture in year 2011.

Traits	Pend	Rac	PanL	PanH	CulmL	MxLBZ	TotN	TotBr	GNP	GWP	GW	PanD	PanN	PanW
Pend	1	-0.19*	-0.19*	0.10	0.15	-0.02	0.08	-0.03	-0.01	-0.12	-0.15	0.06	0.04	-0.04
Rac	-0.02	1	0.62***	0.12	0.01	0.20**	0.30**	0.02	0.16	0.18*	0.06	-0.06	0.03	0.01
PanL	0.22*	0.19	1	0.12	-0.04	0.65***	-0.07	-0.12	0.10	0.12	0.03	0.15	-0.05	0.01
PanH	0.19	-0.10	0.47***	1	0.96***	0.15	0.09	0.35***	0.13	0.26**	0.25**	0.14	0.24**	-0.15
CulmL	0.23*	-0.12	0.41***	0.94***	1	0.05	0.07	0.38***	0.11	0.22**	0.22**	0.13	0.22**	-0.18*
MxLBZ	-0.19	0.62***	-0.16	-0.29*	-0.33**	1	-0.24**	-0.19*	0.11	0.04	0.10	0.28**	-0.11	0.02
TotN	-0.10	0.46***	-0.51***	-0.23*	-0.25*	0.33**	1	0.15	-0.02	0.05	-0.05	-0.18*	0.05	-0.14
TotBr	-0.09	-0.09	-0.41***	0.17	0.24*	0.06	0.40***	1	0.03	0.12	0.09	0.41***	0.05	0.33***
GNP	0.33**	0.34**	-0.02	0.05	0.03	0.46***	0.17	0.30*	1	0.87***	-0.05	0.14	-0.07	0.09
GWP	0.18	0.21*	-0.11	0.15	0.17	0.30*	0.24*	0.14	0.66***	1	0.34***	0.19*	0.11	0.06
GW	-0.15	-0.02	-0.10	0.10	0.17	-0.08	0.19	0.00	-0.42***	0.24*	1	0.08	0.17	-0.10
PanD	0.31**	0.11	0.19	0.55***	0.54***	-0.07	-0.01	0.14	0.25*	0.44***	0.13	1	-0.03	0.40***
PanN	0.28*	0.36***	0.09	0.27*	0.30**	0.12	0.13	-0.09	0.21*	0.33**	0.21*	0.22*	1	-0.22*
PanW	0.11	-0.14	-0.10	-0.04	-0.02	-0.07	0.11	0.47***	0.01	-0.07	-0.05	0.01	-0.21*	1

*P<0.05, ** P<0.01, ***P<0.001, respectively. (Upper diagonal= The correlation coefficient among traits in *compact inflorescence architecture*, Lower diagonal = The correlation coefficient among traits in *loose inflorescence architecture*). Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MxLBZ (MaxLBZ): maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant height, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanN: panicle number per plant.

Table 2.8. Path coefficient analysis showed direct and indirect effects of panicle characters on grain yield.

Cause & effect	Correlation	*P*_value	Cause & effect	Correlation	*P*_value	Cause & effect	Correlation	*P*_value	Cause & effect	Correlation	*P*_value
PanL vs. yield	0.23	<0.01	PanW vs yield	0.30	<0.001	TotBr vs. yield	0.21	<0.01	GNP vs. yield	0.60	<0.001
Direct effect	0.26	<0.001	Direct effect	0.34	<0.001	Direct effect	0.24	<0.001	Direct effect	0.77	<0.001
indirect via TotN	-0.13	0.16	Indirect via PanL	0.08	0.110	Indirect via PanL	0.08	0.16	Indirect via PanL	-0.01	0.460
indirect via MaxLBZ	0.30	<0.001	Indirect via TotN	-0.09	0.310	Indirect via TotN	0.10	0.15	Indirect via TotN	0.19	0.070
Indirect via Rac	0.55	<0.001	Indirect via MaxLBZ	0.01	0.380	Indirect via MaxLBZ	0.07	0.27	Indirect via MaxLBZ	0.16	0.480
Indirect via TotBr	0.10	0.06	Indirect via Rac	0.23	0.030	Indirect via Rac	0.22	0.06	Indirect via Rac	0.13	0.38
Indirect via PanW	0.11	0.03	Indirect via TotBr	0.15	0.110	Indirect via PanW	0.12	0.1	Indirect via PanW	0.09	0.100
Indirect via PanD	-0.03	0.41	Indirect via PanD	0.26	<0.001	Indirect via PanD	0.20	0.003	Indirect via PanD	0.16	<0.01
Indirect via GNP	0.04	0.22	Indirect via GNP	0.13	0.140	Indirect via GNP	0.1	0.18	Indirect via TotBr	0.10	0.200
Total effect	0.22		Total effect	0.29		Total effect	0.21		Total effect	0.60	

Legend for trait: Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, GNP: grain number per panicle.

Table 2.8. (Continued)

Cause & effect	Correlation	*P*_value	Cause & effect	Correlation	*P*_value	Cause & effect	Correlation	*P*_value	Cause & effect	Correlation	*P*_value
TotN vs. yield	-0.09		MaxLBZ vs yield	0.02		Rac vs. yield	0.04		PanD vs yield	0.28	<0.001
Direct effect	-0.07	0.24	Direct effect	0.12	0.098	Direct effect	0.14	0.038	Direct effect	0.28	<0.001
indirect via PanL	0.01	0.49	indirect via PanL	0.30	0.02	indirect via PanL	0.71	<0.001	indirect via PanL	-0.01	0.42
indirect via MaxLBZ	-0.29	0.27	indirect via TotN	-0.24	0.29	indirect via TotN	0.39	<0.001	indirect via Rac	0.03	0.36
Indirect via Rac	0.41	<0.001	Indirect via Rac	0.07	0.45	indirect via MaxLBZ	0.41	0.049	indirect via TotN	0.03	0.46
Indirect via TotBr	0.14	0.08	Indirect via TotBr	0.09	0.24	Indirect via TotBr	0.10	0.07	indirect via MaxLBZ	0.08	0.44
Indirect via PanW	-0.09	0.28	Indirect via PanW	-0.07	0.39	Indirect via PanW	0.02	0.31	Indirect via TotBr	0.08	0.07
Indirect via PanD	-0.02	0.43	Indirect via PanD	0.04	0.47	Indirect via PanD	-0.06	0.45	Indirect via PanW	0.40	<0.001
Indirect via GNP	-0.16	0.07	Indirect via GNP	0.09	0.46	Indirect via GNP	-0.01	0.41	Indirect via GNP	0.13	0.06
Total effect	-0.09		Total effect	0.02		Total effect	0.03		Total effect	0.28	

Legend for trait: Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, GNP: grain number per panicle.

Table 2.9. Eigenvectors for the inflorescence architecture in the principal component (PC) analysis of the 206 sorghum accessions.

(A)Year-1 (2010)

Principle component (Year-1)	PC1	PC2	PC3	PC4
Eigenvalue	3.24	1.74	1.50	1.32
Percent %	23.15	12.43	10.74	9.41
Cumulative (%)	23.15	35.57	46.31	55.72
Traits				
TotBr	**0.30**[a]	0.23	0.12	0.22
Pend	-0.11	-0.29	0.20	-0.20
Rac	0.20	**0.42**[a]	**0.30**[a]	0.27
PanL	**0.35**[a]	**0.35**[a]	0.13	-0.03
PanW	0.29	**-0.31**[a]	0.01	**0.41**[a]
TotN	0.02	-0.09	**0.34**[a]	**0.51**[a]
PanD	0.28	**-0.30**[a]	-0.02	0.25
MxLBZ	0.18	**0.30**[a]	-0.04	-0.26
PanH	**0.45**[a]	-0.17	0.24	**-0.33**[a]
CulmL	**0.42**[a]	-0.24	0.23	**-0.33**[a]
PanN	-0.01	**0.39**[a]	0.16	-0.14
GNP	**0.31**[a]	-0.04	**-0.49**[a]	-0.03
GWP	**0.30**[a]	0.05	**-0.54**[a]	0.09
GW	-0.05	0.20	-0.25	0.16

B) Year-2 (2011)

Principle component (Year-2)	PC1	PC2	PC3	PC4
Eigenvalue	2.85	1.79	1.57	1.46
Percent %	20.27	12.76	11.29	10.41
Cumulative (%)	20.27	33.04	44.32	54.72
Traits				
TotBr	0.24	-0.17	**-0.30**[a]	**0.31**[a]
Pend	-0.05	-0.26	0.12	-0.026
Rac	0.13	**0.43**[a]	0.24	-0.11
PanL	0.21	**0.46**[a]	-0.06	**-0.37**[a]
PanW	0.07	0.13	**-0.48**[a]	**0.40**[a]
TotN	0.05	-0.09	**0.33**[a]	0.28
PanD	**0.36**[a]	0.05	**-0.30**[a]	0.06
MxLBZ	0.13	**0.44**[a]	-0.20	-0.17
PanH	**0.49**[a]	-0.20	-0.11	-0.22
CulmL	**0.47**[a]	**-0.30**[a]	-0.09	-0.20
PanN	0.21	-0.13	0.25	**0.43**[a]
GNP	0.26	0.28	**0.30**[a]	**0.42**[a]
GWP	**0.35**[a]	0.14	**0.38**[a]	**0.36**[a]
GW	0.16	-0.21	0.22	-0.07

([a]PC loadings larger than 0.30 and smaller than -0.30 were regarded as substantial shown in bold)

Legend for quantitative trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant high, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanN: panicle number per plant. *Note: Qualitative traits*; PanS: panicle shape, PanT: panicle type, PanEx: panicle exsertion and PanB: panicle broader were also recorded.

Table 2.10 List and origin of sorghum diversity research set (SDRS) of 107 landraces (*Sorghum bicolor* (L.) Moench) accessions from NIAS, Genebank.

No.	Serial No.	Stock No.	Cultivar Name	Origin
1	61	589	E 9	Chad
2	83	48433	OOTOYO-MURA ZAIRAI	Japan
3	84	48442	HANGETSUTOSUI	Korea
4	87	48445	KOUSHUU ZAIRAISHU	Korea
5	88	48446	CHAL WAXY SORGHUM	Korea
6	98	48458	AI HUI	China
7	5	48491	Y. E. (I. P.) INT. TYPE	India
8	142	48544	AIT BRAHIM	Morocco
9	143	48545	CODY	Morocco
10	144	48546	KOURNIANIA	Morocco
11	145	48548	PHATSAI	Morocco
12	147	48550	SCHROCK	Morocco
13	178	48692	ESHOME	S.Africa
14	246	119475	COL/PAK/1989/IBPGR/2386(2)	Pakistan
15	194	48757	ZA113 DAWA PAS PARA	Nigeria
16	202	48881	PI 229486 VULGARE	Iran
17	286	119448	TAKAKIMI	Japan
18	274	119513	COL/PAK/1991/IBPGR/2724(2)	Pakistan
19	27	403	HEGARI MALOWAR	Sudan
20	47	515	E 232 INGWARUMA PEARLY	S.Africa
21	49	519	AW 70/12 DL/59/1532	S.Africa
22	56	534	E 233 BARNARD RED	S.Africa
23	81	48419	IKEDACHO MATSUO ZAIRAI	Japan
24	129	48519	KALJANPUR	India
25	139	48531	EC 18868	Nepal
26	140	48532	JUNELO	Nepal
27	141	48543	MN 401	Algeria
28	156	48567	143 DINDERAWI 1	Sudan
29	186	48727	RED KAFIR	S.Africa
30	201	48779	PI 282834	Chad
31	204	49005	PI 220636 Q 2/3/56	Afghanistan
32	15	59655	SC NO.0217 CI1197	India
33	207	54763	KOUCHI OUKAWA ZAIRAI	Japan
34	228	76744	MAKHOTLONG I	Lesotho
35	231	91317	NUO GAO LIANG	China
36	240	91326	ER BAI SHE YAN	China
37	284	119430	DANGOMOROKOSHI	Japan
38	290	119461	TOKIBI	Japan
39	254	119485	COL/PAK/1989/IBPGR/2420(1)	Pakistan

Table 2.10. (Continued)

No.	Serial No.	Stock No.	Cultivar Name	Origin
40	256	119487	COL/PAK/1989/IBPGR/2427(5)	Pakistan
41	257	119488	COL/PAK/1989/IBPGR/2439(1)	Pakistan
42	258	119489	COL/PAK/1989/IBPGR/2444(1)	Pakistan
43	261	119494	COL/PAK/1989/IBPGR/2550(1)	Pakistan
44	262	119496	COL/PAK/1989/IBPGR/2553(4)	Pakistan
45	16	119481	COL/PAK/1989/IBPGR/2411(1)	Pakistan
46	69	45423	HIMEKI ZAIRAI	Japan
47	72	45428	KIKUCHI ZAIRAI	Japan
48	122	48512	GOOSENECK	India
49	17	119484	COL/PAK/1989/IBPGR/2416(2)	Pakistan
50	272	119509	COL/PAK/1989/IBPGR/2592(7)	Pakistan
51	170	48617	S. VULGARE 72-726-7	Uganda
52	171	48619	S. VULGARE 72-728-1	Uganda
53	2	235	KOUBOUSHI	Korea
54	4	251	REDBINE 655	Sudan
55	6	260	MORABA 74	Ethiopia
56	7	290	THIBA RED	Ethiopia
57	8	291	E 276 FRAMIDA	Uganda
58	11	294	E 1089 MARIANGARIJORA	Sudan
59	12	297	MUDDAHIHAL	India
60	15	311	AKAHO	Japan
61	20	377	BATTANBAN	Cambodia
62	21	381	AS 4547 JARDIRA	Nigeria
63	28	45432	KANAGAWAZAIRAI	Japan
64	24	400	DHOOTI ANEHULA RABI YANGAR JORA	India
65	25	401	MITHUGADUR	India
66	33	480	HAZERA 6014	Israel
67	34	490	AKLMOI WHITE	Kenya
68	35	491	LAMBAS	Sudan
69	36	492	DINDERAWI 1	Sudan
70	38	494	240 WAD UMM BENEIN	Sudan
71	42	498	MUGBASH WHITE	Sudan
72	51	521	S.BASUTORUM DL/60/97 EAR FROM PIETESBURG	South Africa
73	52	522	DL/60/107	South Africa
74	57	544	WAD YABOO 132/53	Zimbabwe
75	58	545	CAPE COLO 28/53	Zimbabwe

Table 2.10. (Continued)

No.	Serial No.	Stock No.	Cultivar Name	Origin
76	63	635	MN 1277 MUHEYAR	Nigeria
77	66	651	PI 220636 Q2/3/56	Afghanistan
78	67	42156	LIAOZA 1	China
79	73	45437	MOCTAC LOCAL	Korea
80	76	45451	B-112	Sudan
81	91	48449	SENKINHAKU	Korea
82	101	48466	AS 5781 HUAN SA PHAUNG AH LPYSU	Myanmar
83	109	48476	AS 4136 MASAKA LUWEMEA	India
84	162	48608	SC112	Ethiopia
85	166	48612	GIZA 3/59	Ethiopia
86	168	48615	UGANDA L1	Uganda
87	173	48630	AS 4637 NHORONGO NENPI	Tanzania
88	174	48631	E 37	Tanzania
89	188	48738	TSETA LOCAL NATURE TYPE 27/51	Zimbabwe
90	192	48755	E 17	Congo
91	196	48759	KA 24	Nigeria
92	209	54766	CHOONCHAN LOCAL	Korea
93	213	76728	BIG WHITE HULL	China
94	222	76738	XIONG YUE 334	China
95	226	76742	TENANT WHITE	Lesotho
96	227	76743	NYAKASOBA BEST	Lesotho
97	-	-	72-8-13	Taiwan
98	-	-	72-10-10-5	Japan
99	-	-	87-9-21-3-1	Pakistan
100	-	-	87-9-21-3-2	Pakistan
101	37	262	E 1091	Sudan
102	41	48598	109 TONJI	Sudan
103	47	54652	PI 329762	Ethiopia
104	48	512	E 959	Kenya
105	49	81250	PI 152748 C	Kenya
106	58	48752	MILO PET. 139/51 EX TANGANYIKA	Central Africa
107	138	48530	ALLAKH	Bangladesh

Table 2.11. List of traits and sorghum descriptors analyzed for sorghum diversity research set 107 accessions.

No.	Traits/ character	Trait abbreviation	Evaluation method	Type of assessment	Rank	Descriptor referred
Elongation/length based traits						
1	Penducle length	Pend (cm)	Neck length emerged from flag leaf sheath to the base of panicle	MS	maturity	IBPGR/ICRISAT
2	Rachis length	Rac (cm)	Length from the neck node to the base of top branching zone	MS	maturity	NIAS
3	Panicle length	PanL(cm)	Length from the base of the head to the tip	MS	maturity	NIAS
4	Maximum length of primary branch	MaxLBZ/ MxLBZ(cm)	Length of branches (middle third of panicle)	MS	maturity	IBPGR/ICRISAT
Branching traits						
5	Number of node	TotN (number)	Total node on the rachis	MS	maturity	IBPGR/ICRISAT
6	Total branch	TotB (number)	Total primary branch on the rachis	MS	maturity	IBPGR/ICRISAT
Dimensional traits						
7	Panicle diameter	PanD (mm)	natural position at the widest part of head	MS	maturity	IBPGR/ICRISAT
8	Panicle width	PanW (mm)	natural position at the thickness of head	MS	maturity	IBPGR/ICRISAT
9	Panicle shape	PanS (shape)	Inflorescence shape at maturity	VG	maturity	IBPGR/ICRISAT
10	Panicle type	PanT (type)	Inflorescence compactness at maturity		maturity	NIAS

Type of assessment of characteristics indicated in: MG : Measurement by a single observation of a group of plants or parts of plants, MS : Measurement of a number of individual plants or parts of plants, VG : Visual assessment by a single observation of a group of plants or parts of plants, VS : Visual assessment by observation of individual plants or parts of plants. Descriptors Refered: NIAS, IBPGR and ICRIST (1993).

Table 2.11. (Continued)

No.	Traits/ character	Trait abbreviation	Evaluation method	Type of assessment	Rank	Descriptor referred
Yield related traits						
11	Panicle number	PanN (number)	Number of mature panicles per plant	MG	Harvesting	NIAS
12	Grain number per panicle	GNP (number)	Number of grain per panicle on main stem	MG	after threshing	IBPGR/ICRISAT
13	Grain weight per panicle	GWP (g)	Weight of cleaned grains per panicle on main stem	MG	after threshing	NIAS
Miscellaneous						
14	Panicle exsertion	PanEx (stage)	position of peduncle emergence between flag leaf and inflorescence based	MS	maturity	IBPGR/ICRISAT

Type of assessment of characteristics indicated in: MG : Measurement by a single observation of a group of plants or parts of plants, MS : Measurement of a number of individual plants or parts of plants, VG : Visual assessment by a single observation of a group of plants or parts of plants, VS : Visual assessment by observation of individual plants or parts of plants. Descriptors Refered: NIAS, IBPGR and ICRIST (1993).

Table 2.12. Distribution different sorghum panicle types of 206 accessions and 107 SDRS accessions in three main regions.

(i) N=206

Regions	No. of accessions	% of population	Diversity of panicle type					
			Open	Inter	Semi	Compt	Broom	Mix
Africa	76	36.89%	14	15	5	37	4	1
East Asia	66	32.04%	14	14	7	17	13	1
Other regions of Asia	64	31.07%	5	11	10	26	7	7
Total	206	100%	33	40	22	80	24	9

(ii) N=107

Regions	No. of accessions	% of population	Diversity of panicle type					
			Open	Inter	Semi	Compt	Broom	Mix
Africa	52	52.94%	13	13	3	18	4	1
East Asia	25	20.59%	2	6	3	7	7	0
Other regions of Asia	30	26.47%	2	5	7	9	5	2
Total	107	100%	33	40	22	80	24	9

Legend for panicle types: Open: open/ loose panicle type, Broom: broom panicle type, Compt: compact panicle type, Inter: intermediate panicle type, Semi: semi-compact panicle type, Mix: mixed panicle type.

Table 2.13. Analysis of variance and value of traits in sorghum diversity research set (SDRS) 107 landraces accessions.

Traits	TotBr	Pend	Rac	PanL	PanW	TotN	PanD	MxLBZ	PanN	GNP	GWP
Year-1(2010)											
Mean	47.06	11.48	17.24	21.53	2.97	8.19	4.39	8.66	2.67	1168.05	22.03
(±SD)	(± 16)	(± 7.1)	(± 6.91)	(± 8.4)	(± 1.3)	(± 2.3)	(± 2.7)	(± 7.8)	(± 1.3)	(± 963.9)	(± 18.1)
Year -2 (2011)											
Mean	51.64	9.27	16.22	22.08	2.71	6.89	4.08	10.84	1.41	920.75	12.13
(±SD)	(± 19.9)	(±4.1)	(±6.6)	(±7.9)	(±0.9)	(±2.2)	(±1.3)	(±8.5)	(±1.1)	(± 510.9)	(± 7.7)
ANOVA											
F Ratio	4.93	0.16	29.97	0.33	277.39	0.13	51.99	0.78	3.75	36.86	65.97
P-Value	0.0264*	0.6856	<.0001***	0.5634	<.0001***	0.7226	<.0001***	0.3764	0.0529*	<.0001***	<.0001***

* , * *,* ** significantly different at P< 0.05 and P < 0.01, P < 0.001 respectively; *Legend for quantitative trait*: TotBr: total branch per panicle, Pend: peduncle length, Rac: rachis length, PanL: panicle length, PanW: panicle width, TotN: total node along rachis, PanD: panicle diameter, PanN: panicle number per plant, GNP: grain number per panicle, MaxLBZ: maximum length of primary branch, GWP: total grain weight per panicle. *Note: Qualitative traits*; PanS: panicle shape, PanT: panicle type and PanEx: panicle exsertion were also recorded.

Table 2.14. Correlation coefficient among traits in sorghum diversity research set.

Traits	*TotBr*	*Pend*	*Rac*	*PanL*	*PanW*	*TotN*	*PanD*	*PanN*	*GNP*	*MXL*	*GWP*
TotBr	1	n.s	n.s	-0.25**	0.46***	0.19*	n.s	n.s	n.s	n.s	n.s
Pend	-0.22**	1	n.s	n.s	0.19*	n.s	0.26**	n.s	n.s	n.s	-0.25**
Rac	n.s	n.s	1	0.34***	n.s	0.34***	n.s	n.s	n.s	n.s	n.s
PanL	0.29**	n.s	0.24**	1	n.s	n.s	0.34***	n.s	n.s	0.53***	n.s
PanW	0.34***	-0.20*	n.s	0.21*	1	n.s	n.s	0.34***	n.s	n.s	0.53***
TotN	n.s	n.s	0.32***	-0.35***	n.s	1	n.s	n.s	n.s	-0.44***	n.s
PanD	n.s	n.s	n.s	0.26**	0.57***	n.s	1	0.23**	0.38***	0.23**	0.44***
PanN	n.s	n.s	n.s	n.s	n.s	n.s	n.s	1	n.s	n.s	n.s
GNP	0.39***	-0.32***	n.s	0.32***	0.43***	n.s	0.24**	n.s	1	n.s	0.67***
MXL	n.s	n.s	n.s	0.52***	0.21*	-0.43***	0.21*	n.s	0.22*	1	n.s
GWP	0.38***	-0.31**	n.s	0.31**	0.49***	n.s	0.26**	n.s	0.79***	n.s	1

n.s = non significant, P<0.05, ** P<0.01, ***P<0.001, respectively. (Upper diagonal= The correlation coefficient among traits for 2011 growing season, Lower diagonal = The correlation coefficient among traits for 2010 growing season). *Legend for quantitative trait*: TotBr: total branch per panicle, Pend: peduncle length, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, PanD: panicle diameter, PanW:Panicle width, PanN: panicle number per plant, GNP: grain number per panicle, MaxLBZ(MXL): maximum length of primary branch, GWP: total grain weight per panicle.

Table 2.15. Proportions (%) of variance components for the principle components of sorghum diversity research set.

([a]PC loadings larger than 0.30 and smaller than -0.30 were regarded as substantial shown in bold)

Principal component	1st	2nd	3rd	4th
Yearr.2010				
Eigenvalue	3.52	2.13	1.53	1.35
Percentage contribution	25.14	15.19	10.96	9.64
Cumulative %	25.14	40.33	51.29	60.92
ChiSquare	542.22	383.19	301.42	251.19
Traits loadings				
Pend	**-0.30**[a]	0.29	**0.37**[a]	-0.13
Rac	0.08	-0.29	0.29	0.10
PanL	0.27	**0.33**[a]	0.09	-0.15
TotN	-0.14	**-0.36**[a]	0.24	**0.43**[a]
TotBr	0.29	-0.01	0.17	0.09
MxLBZ	0.23	**0.32**[a]	-0.07	**-0.47**[a]
PanD	0.24	0.20	**0.34**[a]	0.13
PanW	**0.34**[a]	0.14	0.26	0.24
PanS	-0.16	**0.45**[a]	-0.09	**0.47**[a]
PanT	-0.20	**0.42**[a]	-0.18	**0.40**[a]
PanEx	-0.27	0.21	**0.49**[a]	-0.12
PanN	-0.10	-0.06	**0.46**[a]	-0.13
GNP	**0.42**[a]	0.02	0.01	0.12
GWP	**0.42**[a]	0.02	0.04	0.19

Principal component	1st	2nd	3rd	4th
Yearr.2011				
Eigenvalue	2.79	2.24	1.85	1.63
Percentage contribution	19.93	15.98	13.21	11.63
Cumulative %	19.93	35.91	49.12	60.75
ChiSquare	507.08	413.06	334.43	264.92
Traits loadings				
Pend	**0.35**[a]	-0.09	-0.08	0.03
Rac	-0.06	0.08	**-0.33**[a]	**0.50**[a]
PanL	-0.07	**0.51**[a]	-0.24	0.06
TotN	0.08	**-0.30**[a]	0.00	**0.49**[a]
TotBr	**0.31**[a]	-0.25	0.06	-0.17
MxLBZ	-0.07	**0.48**[a]	-0.01	**-0.34**[a]
PanD	**0.42**[a]	**0.33**[a]	-0.12	-0.08
PanW	**0.41**[a]	0.06	0.09	-0.25
PanS	-0.01	0.17	**0.59**[a]	0.22
PanT	0.00	0.11	**0.61**[a]	0.11
PanEx	**0.43**[a]	-0.26	-0.03	-0.15
PanN	0.10	0.03	-0.28	0.12
GNP	0.29	0.25	0.05	**0.33**[a]
GWP	**0.37**[a]	0.23	0.05	0.29

Legend for traits: TotBr: total branch per panicle, Pend: peduncle length, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, PanD: panicle diameter, PanW: panicle width, PanN: panicle number per plant, GNP: grain number per panicle, MaxLBZ: maximum length of primary branch zone, GWP: total grain weight per panicle, PanS: panicle shape, PanT: panicle type, and PanEx: panicle exsertion.

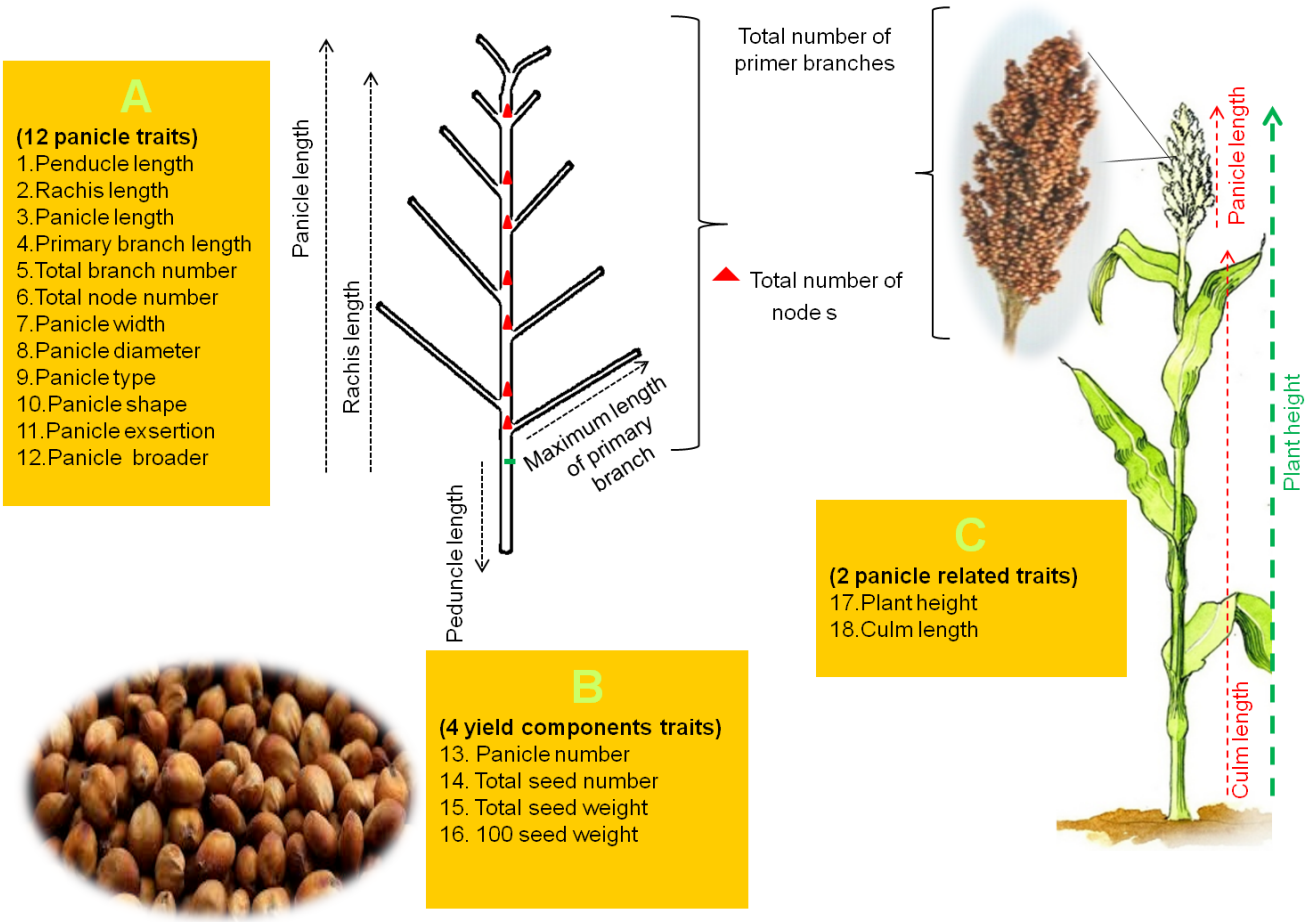


Figure 2.1. Scheme of sorghum inflorescence (panicle) traits investigated.

Legend for Figure: A= Panicle traits- Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanS: Panicle shape, PanN: panicle number per plant, PanT: Panicla type, PanEx: Panicle exsertion, PanBr: panicle broader; B= Panicle component traits/Yield related traits- PanN: panicle number, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight; C= Plant traits- PanH: plant high, CulmL: culm length.

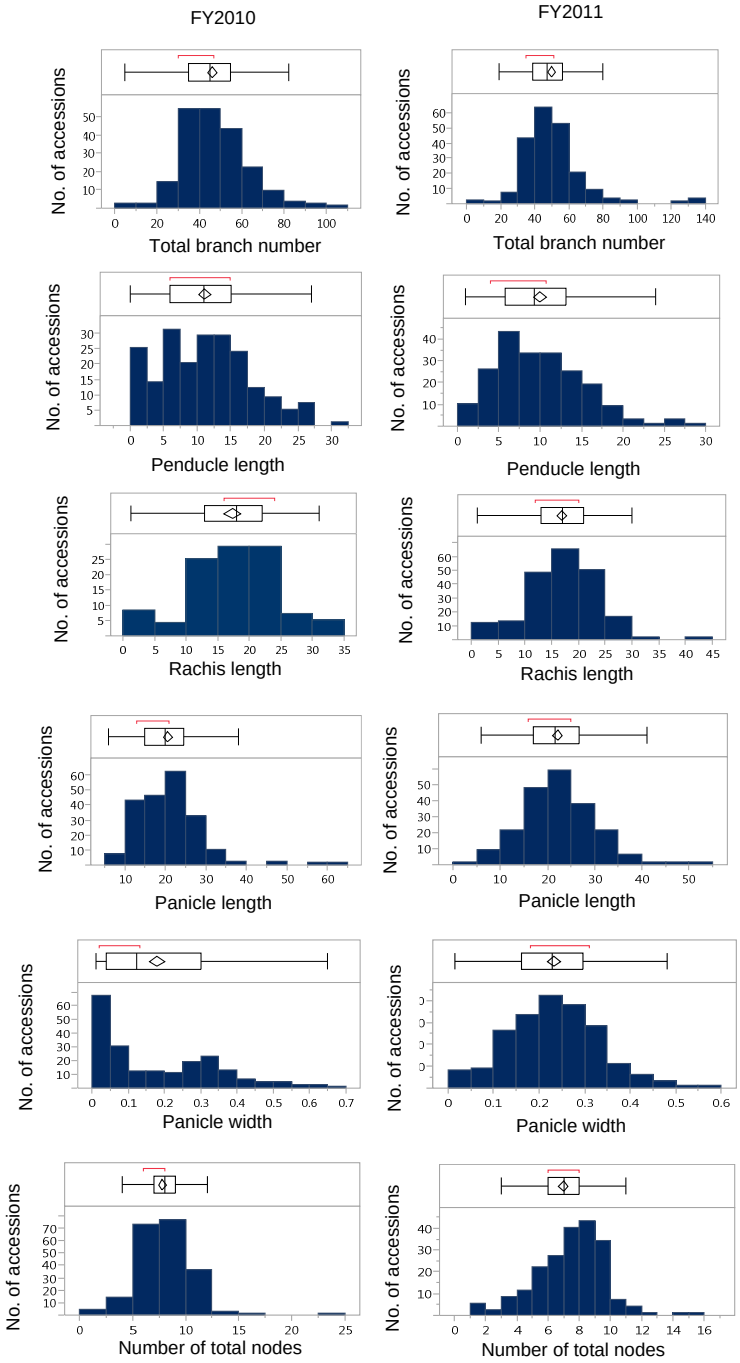


Figure 2.2. Frequency distribution of inflorescence architecture in two different growing seasons (N=206).

Legend for figure: Year-1: FY 2010,Year-2: FY 2011*; Legend for trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant high, CulmL: culm length, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanS: panicle shape, PanT: panicle type, PanEx: panicle exsertion and PanB: panicle broader.

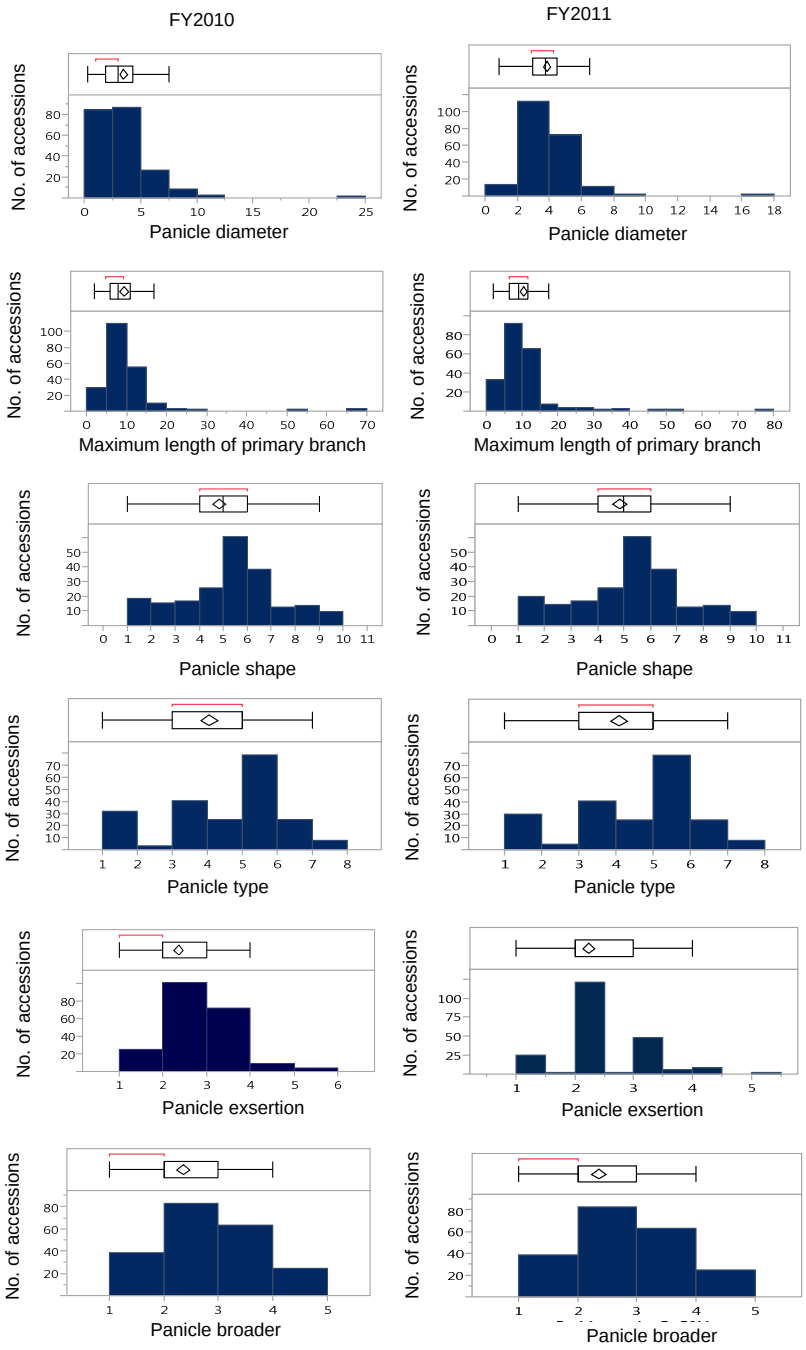


Figure 2.2. Frequency distribution of inflorescence architecture in two different growing seasons (N=206).

Legend for figure: Year-1: FY 2010,Year-2: FY 2011*; Legend for trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant high, CulmL: culm length, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanS: panicle shape, PanT: panicle type, PanEx: panicle exsertion and PanB: panicle broader.

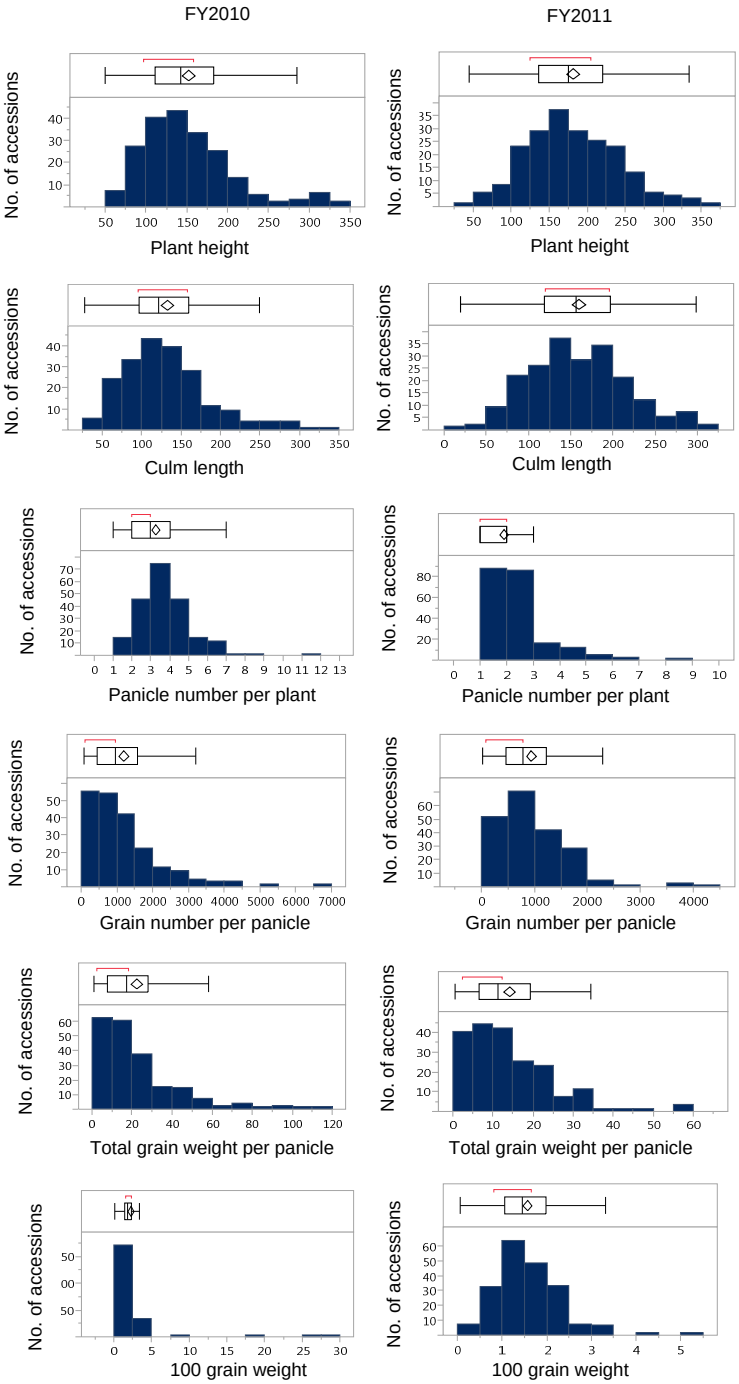


Figure 2.2. Frequency distribution of inflorescence architecture in two different growing seasons (N=206).

Legend for figure: Year-1: FY 2010,Year-2: FY 2011*; Legend for trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant high, CulmL: culm length, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanS: panicle shape, PanT: panicle type, PanEx: panicle exsertion and PanB: panicle broader.

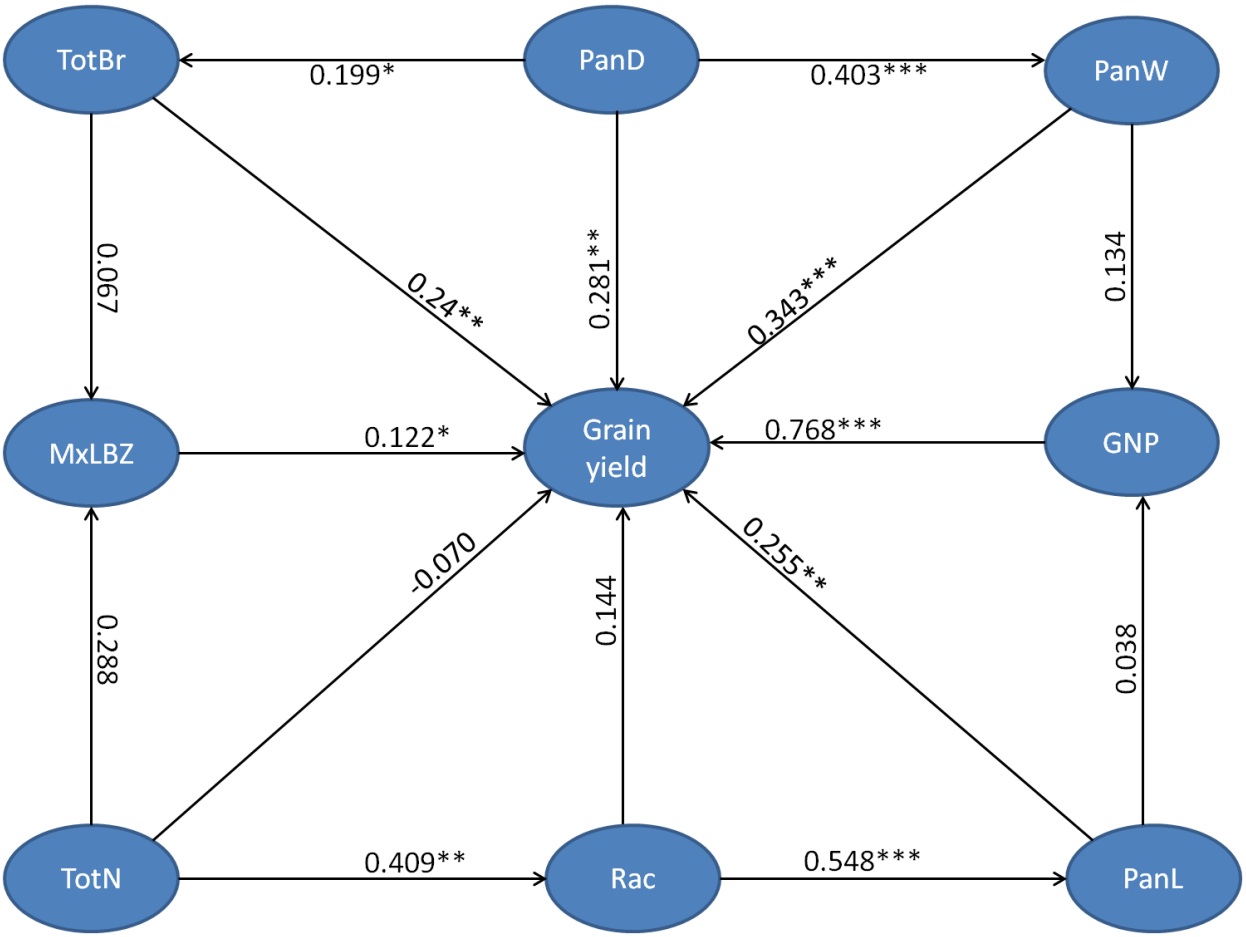
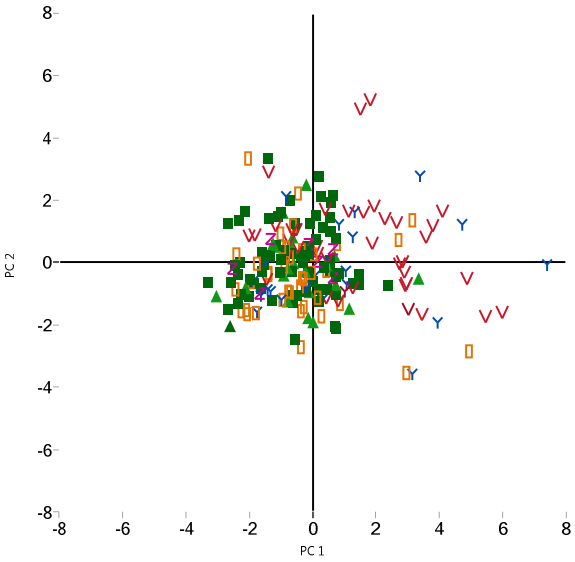


Figure 2.3. Phenotypic path diagram showing the influence (direct and indirect effect) of panicle characters on grain yield.

Legend for figure: *P<0.05, ** P<0.01, ***P<0.001, respectively. *Legend for trait*: PanL: panicle length, Rac: rachis length, TotN: total node along rachis, PanW: panicle width, PanD: panicle diameter, TotBr: total branch per panicle, GNP: grain number per panicle.

(A)Year-1 (2010)



(B) Year-2 (2011)

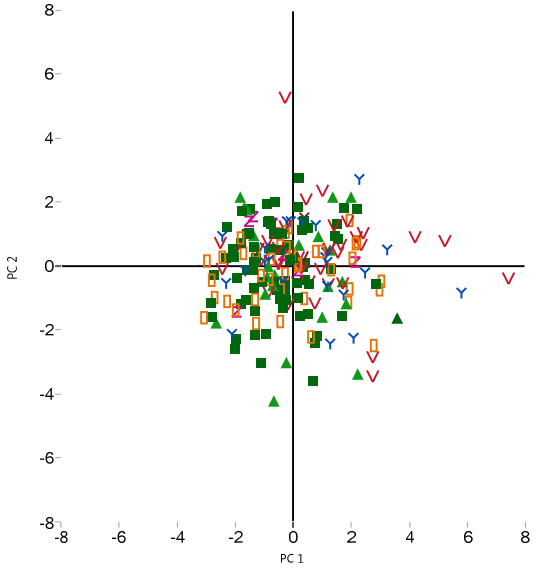


Figure 2.4. Scatter plot of the four principal components for 206 accessions across three different origins summarized from inflorescence traits by the principal component analysis based on correlation.

Legend for figure: Plot of PC1 vs PC2; The coordinates are divided based on the opposite to the clock counter-wise; coordinate I: top right; coordinate II: top left; coordinate III: bottom left; coordinate IV: bottom right. Arrows indicate eigenvectors for the traits. *Legend for accessions*: (v)open type, (▯) intermediate type, (♦) semi compact type, (■) compact type, (Y) Broom type, (Z) mix type.

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant high, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanN: panicle number per plant.

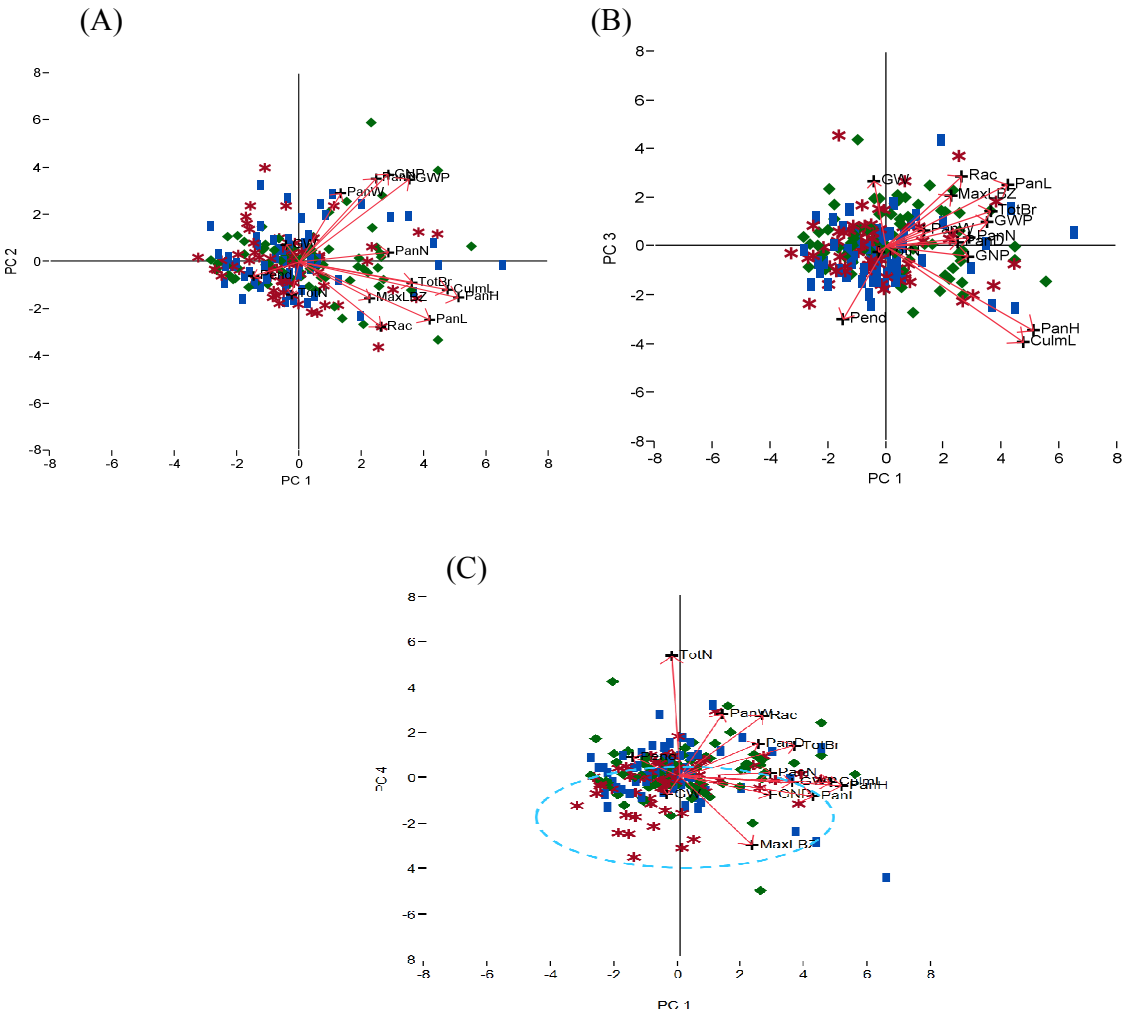
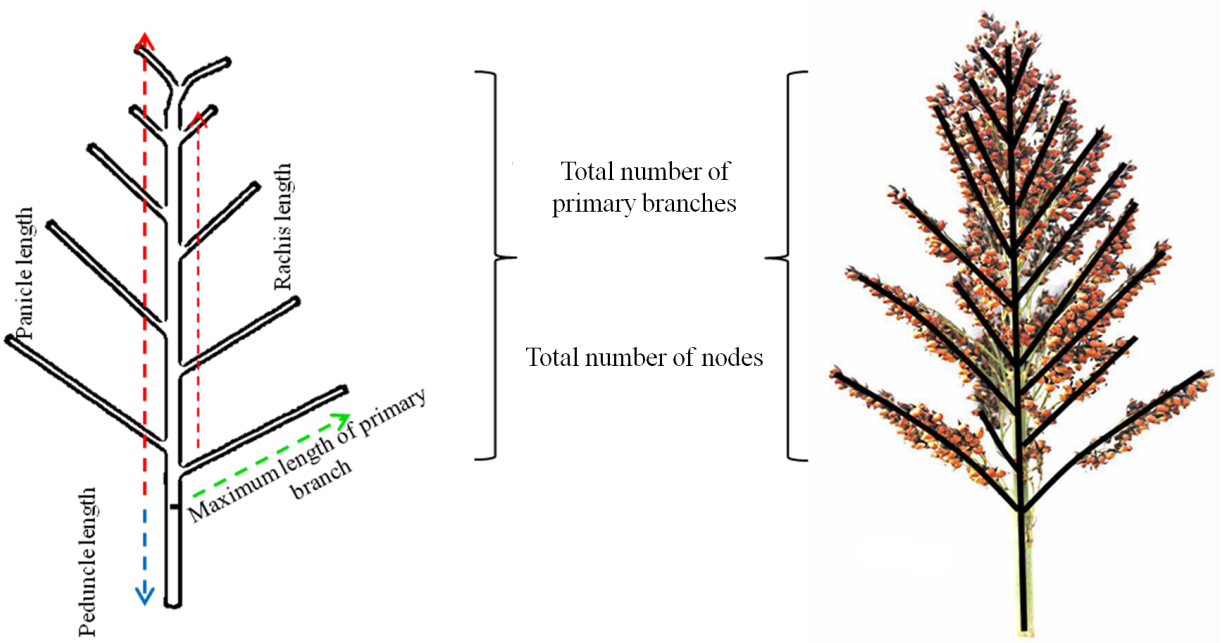


Figure 2.5. Scatter plot of the four principal components for 206 accessions across three different origins summarized from inflorescence traits by the principal component analysis based on correlation.

Legend for figures: Plot of PC1 vs PC2, B) Plot of PC1 vs PC3, C) Plot of PC1 vs PC4. The coordinates are divided based on the opposite to the clock counter-wise; coordinate I: top right; coordinate II: top left; coordinate III: bottom left; coordinate IV: bottom right. Arrows indicate eigenvectors for the traits. *Legend for accessions*: (*) East Asia, (♦) Africa, (l) Other regions of Asia; Arrows indicate eigenvectors for the traits; *Legend for trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanH: plant high, CulmL: culm length, GNP: grain number per panicle, GWP: total grain weight per panicle, GW:100 grain weight, PanN: panicle number per plant.

(i)



(ii)

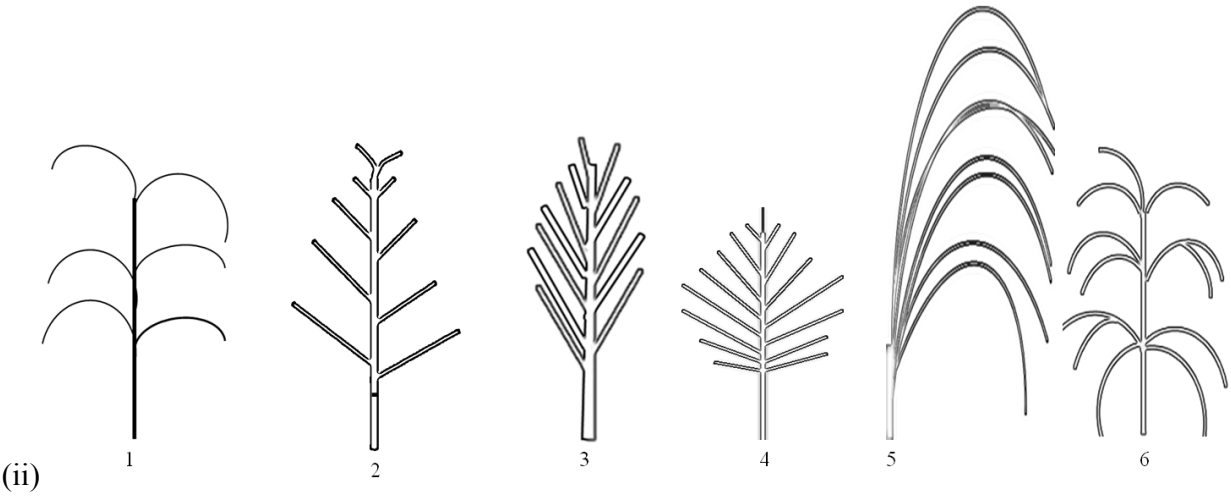


Figure 2.6. (i) Scheme of sorghum inflorescence (panicle) traits investigated for sorghum diversity research set and (ii) images and schematic diagrams of phenotype characterization of different panicle types analyzed.

Legend for Figure: 1- open / loose panicle type, 2- Intermediate panicle type, 3- Semi-compact panicle type, 4-compact panicle type, 5- broom panicle type, 6- Mixed panicle type.

(A)

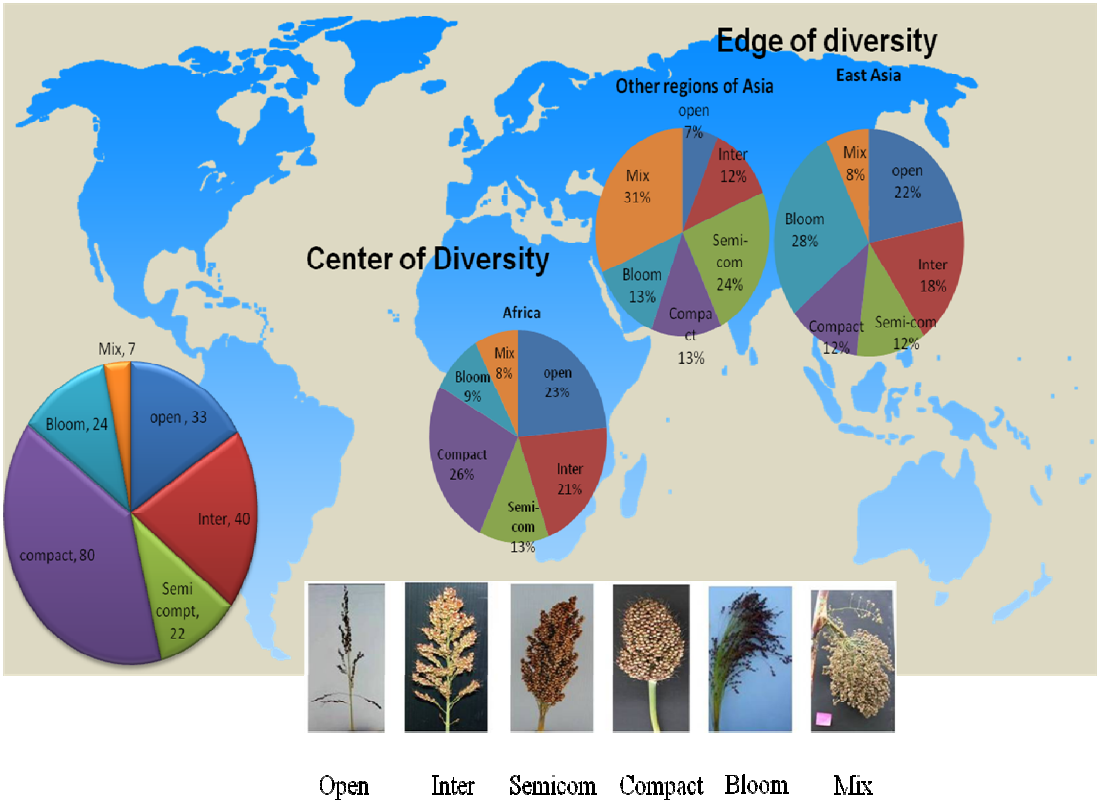
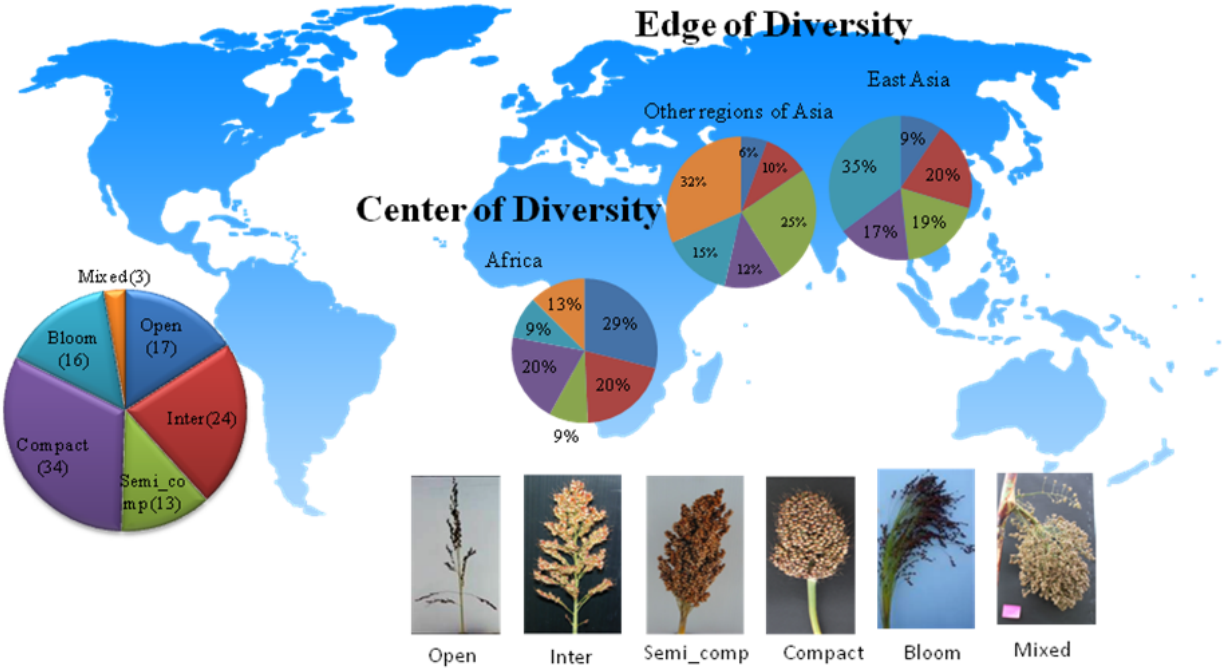


Figure 2.7. (A) Distribution different sorghum panicle types of 206 landraces accessions based on their geographic origins.

The large pie chart summarizes the distribution of accession in the 206 sorghum core collection, and the small pie charts on the world map correspond to the country-specific distribution of sorghum accessions. The color within each small pie charts are reflective of the percentage of accessions in each origin .

Legend for pie chart: Open: Open panicle type, Inter: Intermediate panicle type, Semi-comp: Semi compact panicle type, Compact: Compact panicle type, Broom: Broom panicle type, Mixed: Mixed panicle type.

(B)



(B) Distribution different sorghum panicle types of 107 landraces accessions based on their geographic origins. The large pie chart summarizes the distribution of accession in the 107 sorghum core collection, and the small pie charts on the world map correspond to the country-specific distribution of sorghum accessions. The color within each small pie chart are reflective of the percentage of accessions in each origin.

Legend for pie chart: Open: Open panicle type, Inter: Intermediate panicle type, Semi-comp: Semi compact panicle type, Compact: Compact panicle type, Broom: Broom panicle type, Mixed: Mixed panicle type.

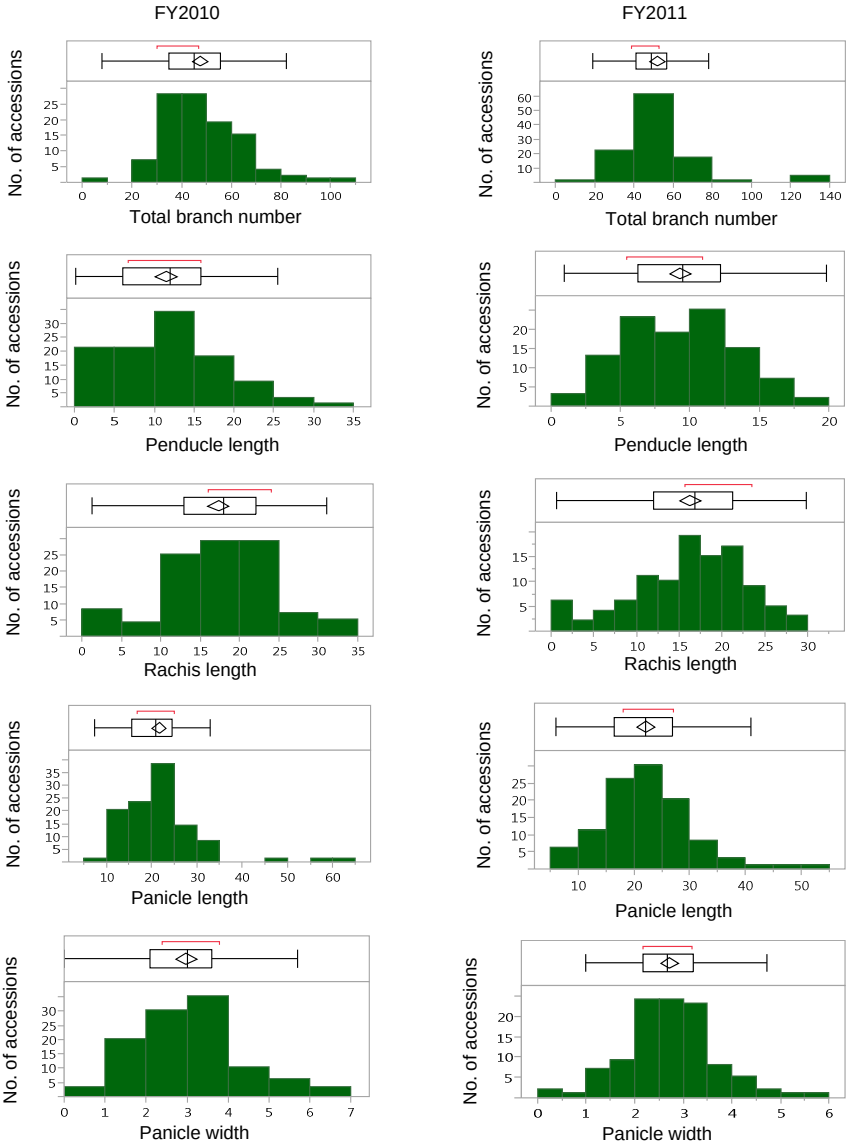


Figure 2.8. Frequency distribution of inflorescence architecture in two different growing seasons (N=107).

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanS: Panicle shape, PanT: Panicle type, PanEx: Panicle exsertion, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle.

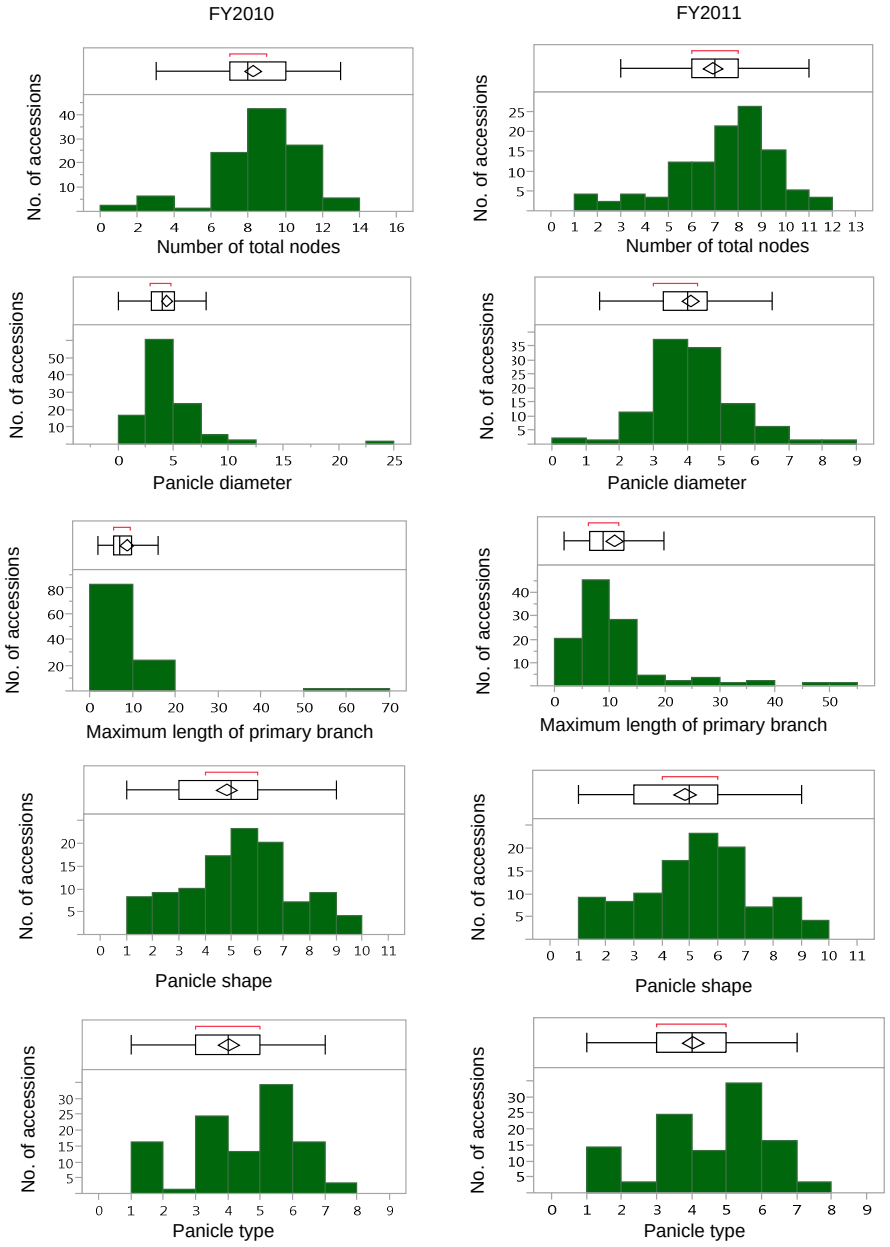


Figure 2.8. Frequency distribution of inflorescence architecture in two different growing seasons (N=107).

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanS: Panicle shape, PanT: Panicle type, PanEx: Panicle exsertion, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle.

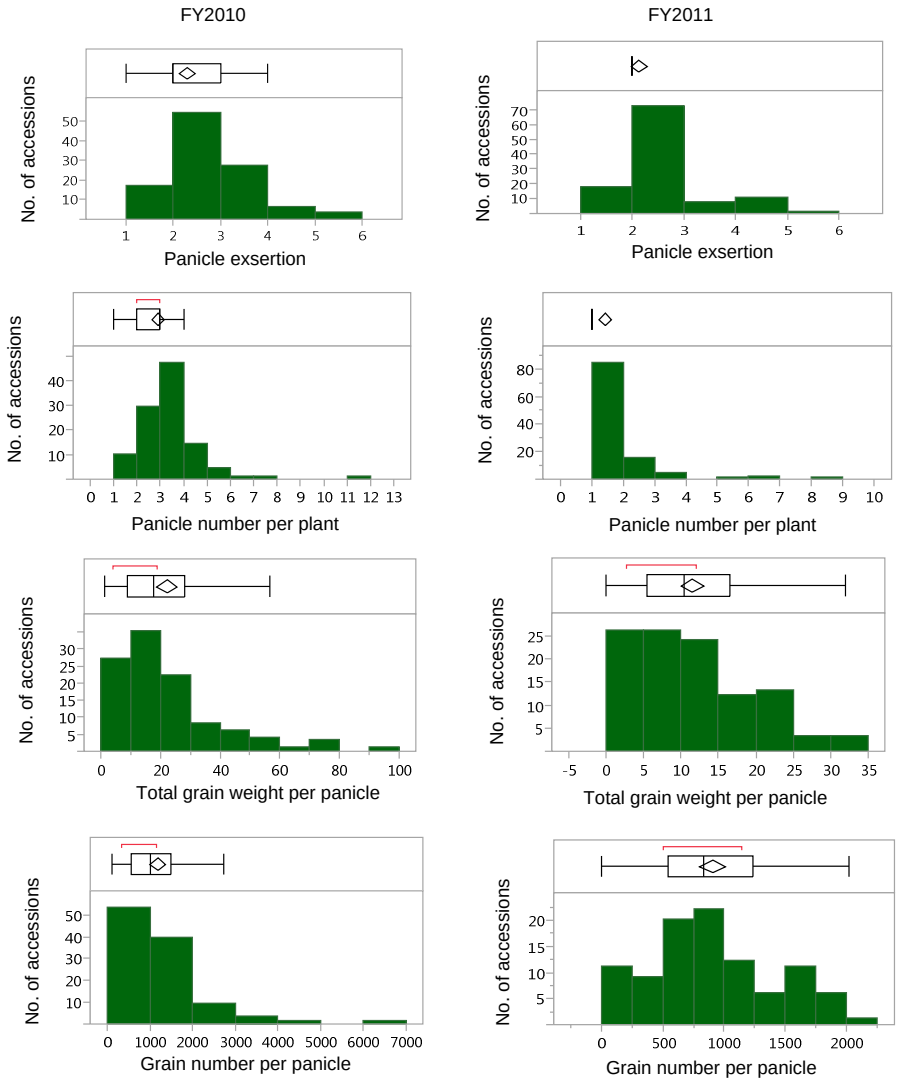


Figure 2.8. Frequency distribution of inflorescence architecture in two different growing seasons (N=107).

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanS: Panicle shape, PanT: Panicle type, PanEx: Panicle exsertion, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle.

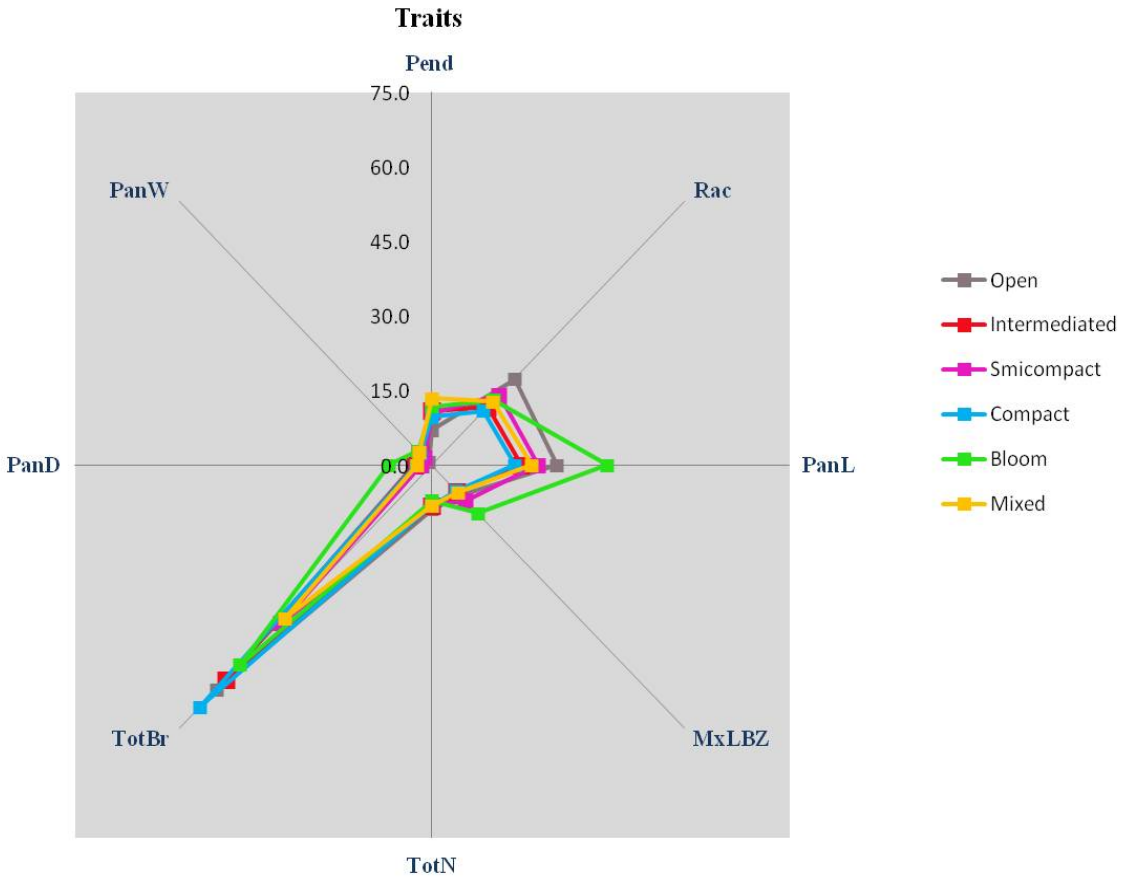
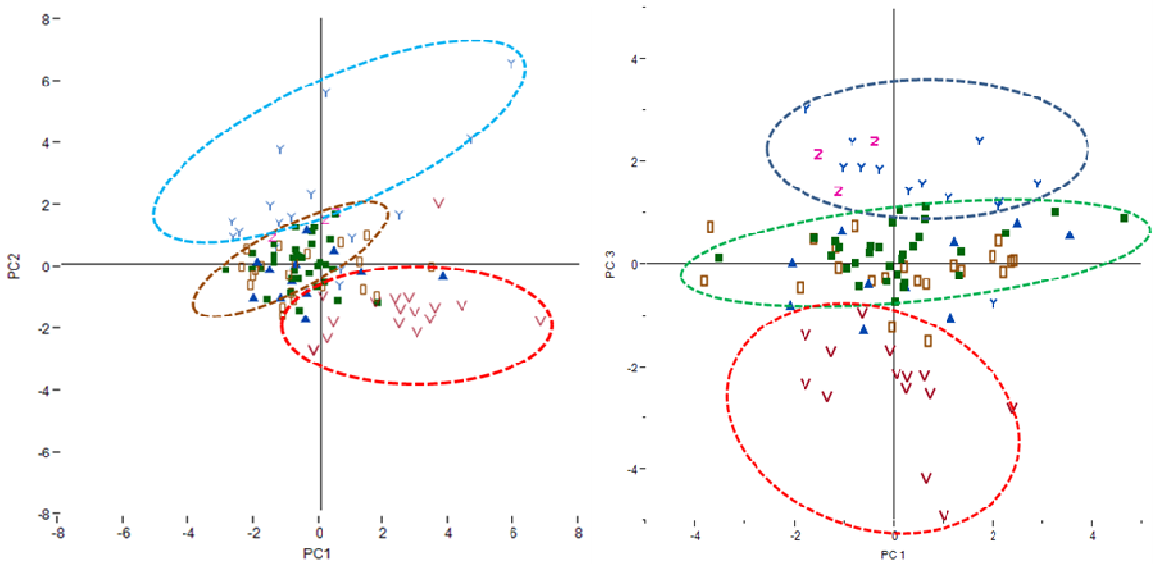


Figure 2.9.Comparison of 6 different panicle types based on criteria characteristics.

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width.



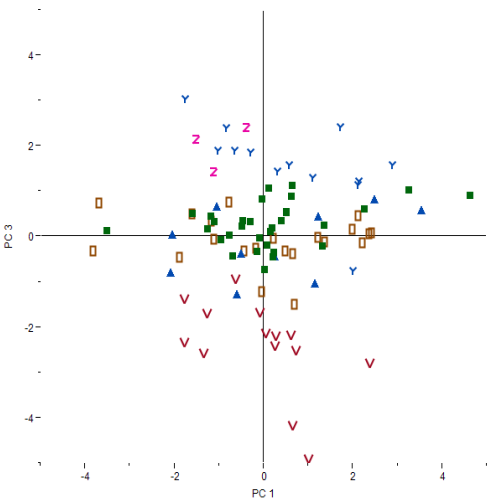
(A) Year- 2010 (B) Year- 2011

Figure 2.10. Scatter plot of the four principal components for 107 landrace accessions across six different panicle types summarized from inflorescence traits by the principal component analysis based on correlation (Labeled by panicle types). (**A**) Plot of PC1 vs PC2 [Yr. 2010], (**B**) PC1 vs PC2 [Yr.2011];

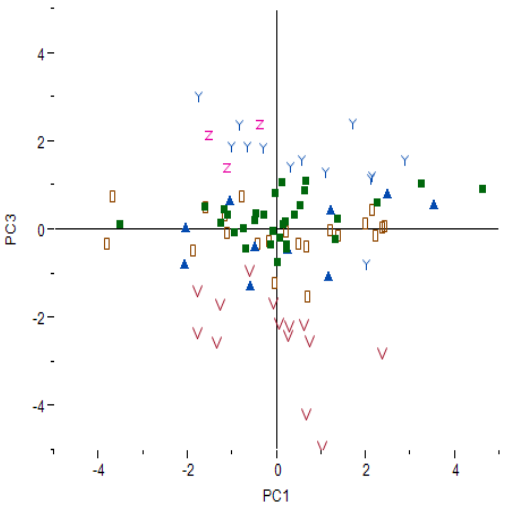
Legend for accessions: (**V**) open type, (▯) intermediate type, (▲) Semi-compact type, (■) compact type, (**Y**) Broom type, (**Z**) mix type.

Figure 2.10. (Continued)

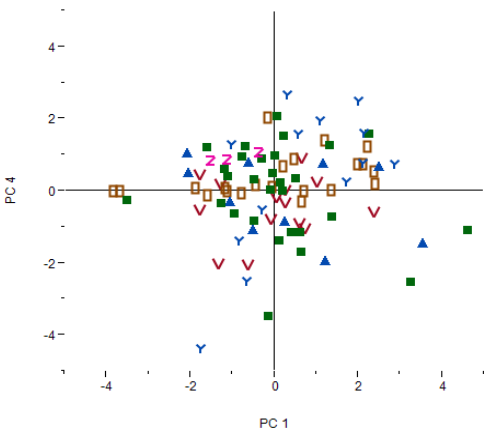
(C)Year- 2010



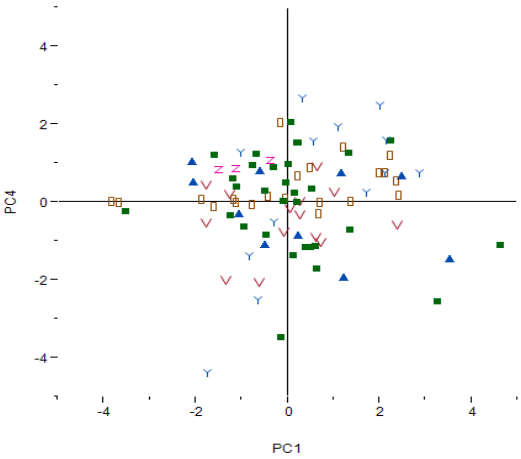
(D)Year- 2011



(E)Year-2010



(F)Year-2011



C) Plot of PC1 vs PC3 [Year. 2010], D) PC1 vs PC3 [Year. 2011], E) Plot of PC1 vs PC4 [Year. 2010], PC1 vs PC4 [Yr.2011].

Legend for accessions: (**V**) open type, (▯) intermediate type, (▲) Semi-compact type, (■) compact type, (**Y**) Broom type, (**Z**) mix type.

Chapter 3

QTLs underlying inflorescence architecture in sorghum detected by association analysis

3.1. Introduction

The grass inflorescence is one of the most important staple grain food resource for humanity and it provided more than 70% of human food. Among the grass species maize and rice are two of the leading model systems for genome research, and inflorescence mutants from these species have been used to characterize a number of genes involved in the control of grass inflorescence architecture (Bommert *et al.,* 2005). Some of these genes are appeared to affect quantitative variation in inflorescence traits (Upadyayula *et al*., 2005). Sorghum panicles also known as the inflorescence show a remarkable diversity in morphological, physiological, genetic and ecological traits. It has a diverse set of morphologies and complex morphological characters. The pattern of sorghum panicle is an important character in sorghum for identifying race divisions and species value (Abdi *et al*., 2002; Harlan and de Wet, 1972; Murray *et al*., 2009). Variation in sorghum inflorescence architecture is not only a result due to differences in the panicle elongation but also differences in the branching and panicle diameter (Brown *et al.,* 2006; Witt Hmon *et al.,* 2013). The knowledge of the genetic basis of sorghum inflorescence architecture and its component traits can enhance the process of genetic improvement in sorghum breeding but still remain of great interest to breeders (Bala *et al*., 1996; Doust and Kellogg, 2002; Doust *et al*., 2005; Futsuhara *et al*., 1979a, 1979b; Kellogg 2000; Zhu *et al*., 2010). Breeders have greatly improved inflorescence architecture, potential energy and grain yield productivity. Panicle morphology directly affects grain yield, therefore knowledge of the genetic basis of sorghum inflorescence architecture and its components can complement the breeder's efforts to improve sorghum. Sorghum inflorescence architecture is not only important

factor for sorghum identification which contribute to both yield and quality of sorghum but also the important determinant of rice panicle and maize tassel because of its close associations with grain yield and grain quality as they include several commercially important traits (Bommert *et al*., 2005; Ikeda *et al*., 2010; Yan *et al*., 2007; Tesso *et al*., 2011). Of these, sorghum inflorescences have become a new model system for functional genomics from agronomic, developmental, and evolutionary viewpoints as well as the other important model cereal crops, maize and rice, however only a few morphological characters of inflorescence architecture have been mapped as major genes using genetic linkage maps (Harlan and de Wet, 1972). Mapping approach to know the genetic basis of identifying genes and QTLs underlying sorghum inflorescence architecture has been undertaken in the same way as QTLs analysis in rice and maize (Colasanti *et al*., 1998; Ikeda *et al*., 2005). Genome wide association studies (GWAS) has been widely applied to identify the causal genes association with agronomical traits in cereal crops. The use of GWAS in sorghum is a newly developed and linkage map construction. Inflorescence pattern of sorghum is a complex trait which involves many genes. In sorghum most yield traits and yield related traits are polygenic, but inflorescences architecture probably remains the most polygenic and complex trait (House 1985; Bello *et al*., 2001; Zou *et al*., 2011). Quantitative traits are the most valuable traits for crop improvement and QTLs analysis is a useful tool because of their common feature of natural variation in a population. Sorghum panicle is also an excellent model for the study of quantitative variation in high order inflorescence pattern in the grass family because it is more highly branched than the inflorescences pattern of other cereal crops such as rice or maize. Otherwise, sorghum genome is more closely related to many major cereal crops with more complex genome and higher levels of gene duplication than rice. Small genome of sorghum is an attractive model for advancing the

study of structure, function and evolution of cereal crop genomes. High resolution linkage maps with many genetic markers is useful for genome-wide association analysis (GWAS) based on linkage disequilibrium (LD) and for mapping of (QTL) underlying agronomic traits due to their wide advantages such as cost-effective and other useful processes. Alternatively association mapping known as LD mapping depends on existing natural variation in crop populations of plants to overcome the constraints inherent to linkage mapping. The genome-wide association analysis (GWAS) is one of the important strategies of LD mapping because the efficiency of association mapping depends on the degree of LD between the functional genetic polymorphisms and genotyped markers across all chromosomes. More than 40000 accessions of sorghum germplasm collections have been used to generate a core collection in mapping the important trait loci. Previous studies have been identified for its plant morphology, environmental stress tolerance, disease resistance and other agromophological traits. Genome wide association analysis studies have been carried out to clarify the genetic bases of agronomic traits in sorghum (Bouchet, 2012; Brown *et al.,* 2006 and 2008, Casa *et al.,* 2005, 2008; Sherzad *et al.,* 2009b) but compared with maize and rice, the genome wide association of sorghum inflorescence architecture has not been extensively studied. Thus this study attempted to detect QTLs for the inflorescence architecture, panicle traits and yield component traits by genome wide association mapping (GWAS). The systematic genomic analysis of sorghum panicle traits may lead to improve yield in breeding programs. Intraspecific variation of panicle traits across the 107 sorghum diversity research set (SDRS) as core collection from around the world to characterize not only a wide range of genetic but also the phenotypic diversity and its suitability for association analysis. This chapter was undertaken to better understand the genetic basis of sorghum inflorescence architecture and its association with yield-related traits. We used 98

simple sequence repeat (SSRs) markers mapped on 10 sorghum chromosomes and the sorghum diversity research set (SDRS) of 107 landraces from world-wide sorghum germplasm. Fourteen sorghum panicle and panicle component traits were analyzed to confirm the effectiveness of a core collection to identify QTLs. Revealing the genetic basis of sorghum inflorescence architecture has been one of the major scientific challenges for sorghum improvement. The value of these traits, the relationship to the yield components and the preliminary effort for the association mapping analysis may be useful information to sorghum breeding. Therefore the objectives of this study were to identify the chromosomal regions underlying sorghum inflorescence architecture.

3.2. Materials and methods

3.2.1. Plant materials, trait measurements and methods

In this chapter we used previously selected sorghum diversity research set (SDRS) of 107 landraces accessions as core collection from Asia and Africa (Shehzad *et al.,* 2009a). It is important to use genetic analysis on establishing core collection at molecular level. These core collection accessions were effectively utilized in previous mapping research and several loci have been identified to be associated with morphological traits. We categorized the plant materials into three different groups involving 25 East Asian accessions, other Asian accessions group (2 from Southeast Asian, 26 from South Asian, 2 from Southwest Asian) and 52 African accessions group **(Table 2.10).** Phenotypic data was recorded for 14 panicle traits according to the sorghum descriptors from IBPGR, ICRISAT (IBPGR and ICRISAT, 1993) and NIAS, Genebank **(Table 2.11).** The basic structure of the mature panicle is a head supported by a stem

(peduncle) which is usually straight but in some cases curved. The main axis is the rachis which runs the length of the panicle from the bottom to the top of branch. Several branches (whorls) arise from the internodes of the rachis. Each lateral may also branch repeatedly with each primary branch dividing into secondary branches and tertiary branches coming from secondary branches. The final branches then carry the spikelets. At maturity stage main component of panicle traits including rachis length (Rac), panicle length (PanL), peduncle length (Pend), panicle shape (PanS), panicle type (PanT) were evaluated. Panicle diameter (PanD) was measured with digital vernier caliper. Rachis length was measured as the distance from the bottom whorl to the topmost one. Peduncle length (Pend) was measured the distance from flag leaf to the lowest primary branch zone. After harvesting, all panicles were dried and cleaned before trait measurement. After cleaning number of total nodes (TotN), total number of primary branch (TotBr) and maximum length of primary branch (MaxLBZ/MxLBZ) were manually measured. At the basal part of the panicle, total primary branches (TotBr) were removed and counted individually. Number of total nodes (total number of the whole on the rachis) (TotN) was counted along main axis. For maximum length of branch zone (MaxLBZ), three branches were randomly chosen from the longest branch zone in the bottom third whole of panicle were counted. The actual number of grains per panicle (GNP) was averaged over 3 panicles with grain weight per panicle (GWP) from each panicle. All grains were threshed and measured for GWP and GNP. Additional data on ordinal grouping observations characters i.e. plant hight (PanH), panicle broadness (PanB), neck length (NecL), panicle width (PanW), awn presence (Awn) and awn length (AwnL) were also recorded but data were not shown. The details of the traits measured in this study and the stage of measurement are explained and listed in **(Table 2.11).** All main panicle traits and panicle related traits were evaluated for three panicles per each accession

of each replication. Totally six panicles per each accession of two replications were measured. Yield traits were measured in three plants of each accession per row included a total panicle number per plant.

3.2.2. Experimental design and field procedure

The study was conducted in the experimental field at the Agricultural and Forestry Center, University of Tsukuba, Japan in FY 2010 (Year-1) and FY 2011(Year-2). A field design was used with (60m x 1m) of 107 accessions by two replications, each accession with 4 individual plants. Dried seeds were prepared with fungicide 1week before sowing time then the seeds were sown manually by dibbling method directly into the field plots. At anthesis stage all panicles were covered by paper bags to prevent from out crossing. The bags were removed at the maturity grain stage.

3.2.3. Genomic DNA isolation, PCR amplification and gel electrophoresis

The SDRS was developed using 98 SSR markers by Shehzad *et al.,* (2009b). The SSRs were screened from published linkage maps of sorghum as revealed by Bhattramaki *et al.,* (2000), Kong *et al.,* (2000), Taramino *et al.,* (1997). In this study the previous genotypic data of Shehezad *et al.,* (2009a) was used. The list of total sorghum microsatellite markers with chromosome location, sequence information, size range and other information are given in (**Table 3.1**). DNA samples were extracted by using CTAB method from the leaves of 40 days old plants as described by Murry and Thompson (1980) with some modification. Ninety eight microsatellite markers were chosen for analysis. The extraction buffer was composed of 2%

CTAB, 50 mM Tris-HCL (pH 8.0), 10 mM EDTA, 0.7 M NaCL, 0.1% Proteinase K, 2% insoluble and 2% 2-mercaptoethanol. Chloroform-isoamyl alcohol (24:1 v/v) extraction was perform to remove the cellular debris and proteins. The DNA was precipitated by adding 2-propanol, and the precipitate was rinsed with 70% and then 95.5% ethanol. The final precipitate was dissolved in 50 µl 1/10 TE solution and stored at 4˚C. The DNA concentration was measured by NanoDrop ND-1000 (Thermo scientific) spectrometer and diluted to a working concentration of 5ng/ µl.

PCR amplification of the sorghum SSRs were performed in 10 µl reaction mixture containing 10 ng DNA template, 10X PCR buffer (Mg^{2+} concentration: 20 nM), 2mM dNTPs, 25 ng of each primer and 0.02 U of Taq polymerase (Toyobo Co., LTD., Japan) enzyme using Applied Biosystem 9700 and 2700 thermal cyclers. Annealing temperature was determined for all primers by using Eppendorf Master Cycler ep gradient S. PCR was conducted with a profile as denaturation at 94 ˚C for 5 min, followed by 35 cycles at 94˚C for 1 min, SSR 55˚C/ SSR 61˚C/ SSR 67˚C for 1 min, 72˚C for 2 min with a final extension at 72˚C for 5 minutes and then cooling at 4˚C. PCR products were fractionated through 10% (w/v) polyacrylamide gel (10 cm in size) with constant supply of 200 V powers, 500 mA current for 65 min to 120 min depending upon the size of PCR product. 10xTBE buffer was used in making the gel while 1xTBE Buffer was subjected to the tank and the gel was stained in ethidium bromide solution. The gel was revealed by using Kodak Digital Science EDAS 290 ver. 3.6 with Kodak ID Image analysis software ver. 3.5. Different bands of the same SSR primers were grouped according to their respective sizes by comparing with 50 bp DNA size marker ladder and genotyping was done visually according to the format of different softwares used.

3.2.4. Statistical analyses and association mapping

The overall data were divided into two clusters to calculate the association of panicle traits with other yield traits based on three different origins (East Asian, Other region of Asia, Africa) and the association of panicle traits with other yield traits based on panicle pattern. The population structure among the 107 accessions using the genotype data of 98 SSR markers was estimated with the statistical package STRUCTURE ver. 2.2 (Pritchard *et al.,* 2000). The detail of structure analysis is described in the next section. To obtain *P* values representing the significance association of markers with traits of LD we used the statistical software TASSEL (Trait Analysis by Association, Evolution and Linkage) ver. 2.0.1 (Bradbury *et al.,* 2007), with a general linear model (GLM) and a mixed linear model (MLM), were used to obtained the *P*-values, which represented the significance of LD and the fraction of total variation R^2 value revealed by marker effect.

3.2.5. Population structure and kinship matrix

The program STRUCTURE, version 2.2.3 (Pritchard *et al.,* 2000), was used to analyze population structure and assign individuals to sub-populations. The STRUCTURE program was run 10 times for each number of sub-population (*J)*, ranging from 1-9 by using Bayesian clustering analyses with the admixture model. Markov chain Monte Carlo (MCMC) sampling was repeated 1 x 10^5 times after 1 x 10^4 cycles of a burn-in period (i.e., j=2 to 8). The optimal number of populations was determined on the basis of estimated logarithmic posterior probability of the Bayesian clustering. The final sub-populations were determined on the basis of 1) likelihood plot of models, 2) stability of grouping patterns across 10 runs, 3) cluster analysis (NJ tree), and 4) principal component analysis (PCA). The analysis was repeated three times for

each number of *J*. On the basis of this information, we chose $j = 3$ as the optimal grouping. The posterior probability of $J = 3$ was the largest among other values of *J* (Table 3.2). Thus, we chose $J = 3$ and obtained estimates for the proportion of accession *i*'s genome that originated from population *j*, q_{ij}. A Q matrix, whose (*i,j*)-th element was represented as q_{ij}, was incorporated into the association mapping models in which the effect of population structure was considered. The combined display of the color-coded sub-population memberships from STRUCTURE with other analyses are shown Kinship (K) was calculated with SPAGeDi 1.3 (Loiselle *et al.,* 1995; Hardy and Vekemans, 2002). A kinship matrix (K) was calculated as the allele-sharing rates of the 98 SSR markers as suggested by Zhao *et al.,* (2007) and used in the models that included a *K* effect.

3.2.6. Linkage disequilibrium (LD) plot

Simple sequence repeat (SSRs) were checked for the distribution of alleles among populations (Figure. 3.1). LD between SSR markers were estimated by *D'* and *r* (*D'* represents the standardized disequilibrium coefficient and *r* is the correlation between alleles at two loci (Farnir *et al.,* 2000; Elhan *et al.,* 2009) for all possible combinations of alleles, and weighting them according to allele's frequency.

3.2.7. Model comparison and association analysis

To identify QTLs significantly associated with panicle traits and its component traits and to assess the effect of population structure on association mapping of these traits, we compared two different models, *i.e*., a general linear model (GLM) and a mixed linear model (MLM) by using TASSEL software. The *P*-values obtained from all models were converted into -Log10 (*P*).

We compared different association models (two models from a general linear model- GLM, and two models from a mixed linear model- MLM) to evaluate the possibility of false positive among in these models where were observed *P*-values were plotted against expected *P*- values as demonstrated (**Figure. 3.7**). As for GLM, we approached two different models (1) the Naive model, which there is no control of population structure and kinship, and (2) the Q model, which is based on population structure (Yu *et al.,* 2006). In MLM, we approached two different models such as (1) a model based on kinship (K) and (2) a model unified both population structure and kinship (Q+K). Among 4 different models naive model showed the highest deviation from y=x line then other models such as K, Q and Q+K models**.** K model was better than naive when compare of naive but the results obtained from naive and K model detected the largest number of markers associated with different panicle traits among all single QTL models. The method of both naive and K models might detect a larger number of false positives than others. The Q and Q+K models showed comparable results because they gave the lowest deviation from the y=x line as indicating these two methods might have the smallest possibility of detecting false positive among all models. Among all possible models of two different single QTL approaches for association analysis, GLM and MLM, the MLM approach was shown to be superior to more conventional linear models (Yu *et al.,* 2006). After comparisons across different models we selected Q+K model as the best fit model to determine the association of SSRs markers with each trait for sorghum panicle QTL. The selected models were then used to test marker-trait associations between 98 SSRs and 14 sorghum inflorescence architecture traits.

3.2.8. Marker localization and homology to known genes

The significant loci lined with panicle architecture were physically localized by BLAST in http://www.phytozome.net/sorghum, http://www.plantgdb.org/SbGDB/,or http://www.gramene.org/. Markers previously identified as linked to known genes were localized to the genome-based sequence information provided in Map Viewer at the NCBI website (http://www.ncbi.nlm.nih.gov/mapview/) and sorghum genome database in http://www.phytozome.net/sorghum. Protein sequences of genes were also used to search by BlastP, and the homologous sorghum genes were identified in http://www.plantgdb.org/SbGDB/.

3.3. Results

3.3.1. Linkage disequilibrium (LD) plot

A short to medium range of pair wise LD statistic was observed for total germplasm. The pairwise LD triangle plot between polymorphic marker sites in a hypothetical genome fragment. The pairwise LD values of polymorphic sites were plotted on both the X and Y axis, upper diagonal show r^2 values and the corresponding p-values from rapid 1000 shuffle permutation test shown in below diagonal (**Figure.3.1**). Each cell represents the relationship between two markers with the color codes indicating the significance of LD. Maximum number of SSR markers with highly significant LD ($P<0.0001$) were situated on linkage groups A and B (marker index 1-41). On the other hand, a short range of LD between markers closely locating on the chromosomes was not obvious.

3.3.2. Population structure (Q) and inflorescence architecture

The population structure was inferred with Bayesian clustering analyses with the admixture models in which the number of populations (*J*) ranged from 2 to 9. Markov chain Monte Carlo (MCMC) sampling was repeated 1 x 10^5 times after 1 x 10^4 cycles of a burn-in period. The optimal number of populations was determined on the basis of estimated logarithmic posterior probability of the Bayesian clustering. The analysis was repeated three times for each number of *J*. The posterior probability of $J = 3$ was the largest among other values of *J* (**Table. 3.2**). Thus, we chose $J = 3$ and obtained estimates for the proportion of accession *i*'s genome that originated from population *j,* q_{ij}. Three sub-groups (sub-populations) were detected across the sorghum diversity research set (SDRS) of 107 landraces (Group I- *J*=1, Group II- *J*=2, Group III- *J*=3) that contained 35, 35 and 37 accessions and these groups comprised of six different panicle types (**Figure 3.2)**. We chosen the other inflorescence traits such as PanL, PanD and yield related traits such GWP and GNP to examine the association of sub-group membership. Plotting the STRUCTURE results on the panicle length and panicle diameter showed significant difference among three sub-populations (**Figure 3.3 and Figure 3.4)** but yield related traits were not clearly distinguish a specific major group among the sub-populations (**Figure 3.5 and Figure 3.6)**.

3.3.3. Comparison between GLM and MLM

We used the statistical software TASSEL (Trait Analysis by Association, Evolution and Linkage) ver.2.0.1 (Bradbury *et al.,* 2007) to obtained *P* values representing the significance of LD. We compared different association models (two models from a general linear model - GLM, and two models from a mixed linear model- MLM). To evaluate the possibility of the false

positives in association models, we plotted observed *P*-value against expected *P*-values as described by Stich *et al.*, (2008) and Shehzad *et al.*, (2009b) (**Figure. 3.7**). As for GLM, we approached two different models (1) the naive model, which there is no control of population structure and kinship, and (2) the Q model, which is based on population structure (Yu *et al.,* 2006). The association analysis by using the GLM model without population structure and kindship detected a large number of associations between genotypes and phenotypes. This model had no control for heterogeneity of genetic background (i.e., population structure and familial relatedness among accessions) and thought to be affected largely by false positive. In MLM, we approached two different models such as (1) a model based on kinship (K) and (2) a model unified both population structure and kinship (Q+K). Among 4 different models naive model showed the highest deviation from y=x line then other models such as K, Q and Q+K models. K model was better than naive when compare of naive but the results obtained from naive and K model detected the largest number of markers associated with different panicle traits among all single QTL models. The results obtained from naïve and K models detected large number of markers associated with different panicle traits. The method of both naive and K models might detect a larger number of false positives than others. The Q and Q+K models showed comparable results because they gave the lowest deviation from the y=x line indicating these two methods might have the smallest possibility of false positive among all models. After taking consensus among models, a total of 15 loci were identified by any two models that have strong association with 14 panicle traits. After comparisons across different models K and Q models are not so bad but we selected Q+K model as the best fit model to determine sorghum panicle QTL.

3.3.4. QTL detection by association analysis

A total of 14 panicle traits were evaluated on the (SDRS) of 107 landraces population. The selected model, Q+K, was used to test marker trait associations between 98 SSRs and 14 panicle traits because model comparisons showed that the Q+K model suitably controlled the false positive rate and gave appropriate associations with 14 panicle traits as the best model. The data analysis revealed that total 44 QTLs for 14 panicle traits and panicle-related traits were detected at $-\text{Log}_{10}$ *P*-value ≥ 1.3 as the threshold value, and this threshold level corresponds to the $P<0.05$ level of significance (**Figure 3.8, Figure. 3.9 and Table 3.3**). The QTL results were presented based on the average trait values of panicle traits over two different growing seasons (Year 2010 and Year 2011). QTLs were not identified from each growing season with different QTLs data but the same QTLs were presented in both years and their chromosomal location is shown in (**Figure 3.9 and Table 3.3**). Fifteen loci on 9 chromosomes were found to be significantly related to the patterns of observed panicle traits. Among these loci, five QTLs were responsible for length-based traits, and three QTLs were responsible for dimentional traits. Five QTLs were responsible for panicle feature such as panicle type and shape, one QTL responsible for yield related trait such as grain weight. Two QTLs were responsible for branch-based traits, such as the total node number and branch number with $-\text{Log}_{10}$ (*P*) values ranging from 1.3 to 7.6 as the threshold value. Several genomic regions affected multiple traits, including one region that affected PanL and MxLBZ (MaxLBZ). QTLs for different traits tended to be found in the same region on chromosome 4 and on chromosome 9. Additionally, QTLs were involved on Chr-2, Chr-5, Chr-6 and Chr-10 as a novel QTLs that underlying rachis length, total node number, panicle diameter and panicle type were identified. Another single locus, *Xtxp10* (Chr-9), was found to be strongly associated with two of the main panicle traits, panicle length and maximum

length of the primary branch (PanL, MaxLBZ). Similarly, a single locus *Xtxp12* was highly associated with the PanL trait and was association with MaxLBZ on Chr-4. The panicle length (PanL) was associated with the maximum number of SSRs on four different chromosomes (Chr-1, Chr-2, Chr-4and Chr-9). *P*-values greater than 5.0 were treated as major QTLs. Our results suggested that the sorghum linkage group is heavily populated with loci that were responsible for the inheritance of panicle traits of 107 accessions, particularly the panicle elongation and branching traits. Similar results have been reported in other studies (Klein *et al.,* 2001; Shehzad *et al.,* 2009b; Srinivas *et al.,* 2009). In this study, the association analysis using 98 SSRs in the regions of Chr-5 and Chr-10 detected new associations for panicle and panicle-related traits. In the first growing season (Year 2010) data analysis revealed that 36 QTLs for 14 panicle traits were detected in the second growing season and (Year 2011) data analysis revealed that 21 QTLs were detected at $-\text{Log}_{10}$ *P*-value ≥ 1.3 as the threshold value, and this threshold level corresponds to the $P<0.05$ level of significance. Among these QTLs, locus *Xtxp212* on Chr-2 was highly associated with the panicle traits and yield related trait. One of the panicle traits, the panicle diameter (PanD), had a strong association with the single SSR markers *Xtxp25* on chromosome 2. One of the main panicle traits, the maximum length of primary branch (MaxLBZ), was associated with SSRs marker loci on six different chromosomes.

3.3.5. Physical co-localization of QTLs

To validate our results, we physically localized our markers and compared their positions with known genes. One QTL (*SbAGB03*) identified in this study was physically localized on Chr-2 at 58,128,106 bp and was homologous to a protein-coding gene SB02g024110 with a molecular function of binding DNA or protein. Similarly *Xtxp8* (LG B),

located on Chr-2 at 64,824,875 bp was found to be within the sequence of the gene Sb02g029730, which plays an important role in ATP binding and protein tyrosine kinase activity. Further experimentation is needed to establish whether either of these loci is related to drought tolerance traits in sorghum.

3.4. Discussion

Phenotypic data analysis result revealed that the selection for panicle length, total branch number and panicle diameter had a strong impact on grain yield because we have shown that many of these length-based measurements are correlated, which may suggest a common genetic regulation, and that several traits are likely to influence yield. Our result revealed that the PanL, PanD, PanW, TotN and TotBr are important component traits for the variation in sorghum inflorescence architecture and also that these traits contribute directly and indirectly to yield improvement. Genome wide association mapping (GWAS) is a powerful tool fine mapping of quantitative traits and is dependent on the structure of linkage disequilibrium of alleles at different loci (Flint- Garcia *et al.,* 2005). Association analysis is strongly affected by both false positive (addition of same subpopulation in population structure) as well as false negatives (statistical power in detecting QTLs). In this study we have used different models for GWAS to control both false positives (spurious association) and false negatives (increase statistical power of the models). Some of the significant markers showed same level of association in all models, while in some cases same markers identified with different level of significance by different models. The success of GWAS depends upon the possibility of detecting LD between marker alleles and alleles affecting the expression of phenotypic traits (Stich *et al.,* 2005). In this study, we found a wide-range LD, which ranged over chromosomes, whereas a short-range of LD

between markers closely locating on the same chromosome was not obvious. A wide range of LD might be caused by population structure, and might be responsible for a large number of false positives when the association mapping models did not take into account the population structure (i.e., in the naive and K models). A short range of LD is also caused mainly by physical linkage on the chromosome. Low LD in a short range may indicate that marker density in this study is not enough for detecting QTLs in a genome-wide manner. Thus, many QTLs might be missed because of the low density of markers used in this study, although some markers still captured the signal of QTL even though in this density. The naïve and K models, which did not control the effects caused by population structure, detected a large number of significant associations between markers and panicle traits. These models showed large discrepancy of observed *P*-values from the expected *P*-values, indicating these models were affected by a large number of spurious associations in comparison with the other models. When the population structure was taken into account in other models (i.e., Q and Q+K models), a number of significant associations is much less than in naïve and K models. These models showed the smaller discrepancy from the uniform distribution of *P*-values. Plotting the STRUCTURE results on the inflorescence/panicle types revealed association of sub-group membership (broom, open and compact) probability. These sub-populations differed in three groups of inflorescence/panicle types from six different panicle types, which were associated with 14 inflorescence traits. Cultivated sorghum is classified into five main races (Bicolor, Guinea, Caudatum, Durra and Kafir), and their identification of species values and divisions are primarily based on panicle and grain characteristics. According to their division identification system, broom corn generally falls into bicolor type, bicolor and guinea races have open panicles, kafir and durra races have compact panicles, and caudatum panicle types can vary (Harlan and de Wet,

1972). In previous successful reports by Brown *et al.,* (2011) and by Casa *et al.* (2008), the phenotype-based racial classification of sorghum, which was based on panicle and spikelet characteristics, is controlled by a limited number of genomic regions; however, in genotype-based classification using STRUCTURE analysis, sorghum races can be classified, except for bicolor because it is based on random markers that are distributed across the genome, which can capture the genomic variation among sorghum races. In this study, an SDRS of 107 genotypes generally formed a sub-population by STRUCTURE classification, which was based on the inflorescence architecture, but did not clearly distinguish a specific major group. Thus, our results indicate that sorghum panicle types that are distributed throughout the African and Asian regions most likely vary according to different adaptation levels, temperature, humidity and rainfall patterns. The diversity in shape and compactness are likely to indicate the selection of varieties that can survive in different local environments. However, our result is in agreement with the previous report by Brown *et al.,* (2011), which found that, among the five main races, sorghum bicolor can grow everywhere in Africa and Asia but does not form a clear, separate sub-population.

In total, 15 loci were detected using 98 SSRs mapped markers, which were distributed among 9 linkage groups. Our results showed the locations of QTLs for 14 panicle traits as four types: length-based traits (Rac, PanL, MaxLBZ), branching-based traits (TotN, TotBr), size and dimentional traits (PanD, PanW, PanTand PanS) and panicle-related yield traits (GWP and GNP). Our present research examined characteristics of panicle architecture. Among these characteristics, panicle length (PanL) has been investigated by many previous reports; however, the panicle length (PanL) consists of the rachis length (Rac) and the top most primary branches on the rachis. In this study, we analysed PanL by separating these different parts. Interestingly,

although many QTLs for PanL (10 QTLs in 2010, 4 QTLs in 2011) were detected on six chromosomes, our results did not reveal whether PanL itself and Rac were controlled by different QTLs. No QTLs of the latter were detected together. Our result indicated that these traits appeared to be under separate genetic control (from length based traits) in sorghum inflorescence architecture because there was no overlap between QTLs that were detected for these traits. In elongation, the characteristics of rachis length (Rac) and branches along the rachis, such as MaxLBZ, also appeared to be distinct processes. The association between QTLs for PanL and MaXLBz on Chr-4 and Chr-9 (LG- F) accounts for the strong correlation between these traits. In this case, the allele that was associated with greater panicle length was associated with increased primary branch length. Shehzad *et al.* (2009b) reported that the same location of two QTLs of panicle length (PanL) traits (*SbAGF06* on Chr-1 and *Xtxp7* on Chr-2) matched with the same chromosome location and the same QTL of the length-based trait PanL, which was identified in the study. Many QTLs for different traits tended to be found in the same region on chromosome 4 (LG-D) and on chromosome 9 (LG-F). Earlier studies on inflorescence architecture in sorghum, rice and maize suggested that branching characteristics had the most important role in panicle characteristics. QTLs for these traits were located in the same position**.** Our results of GWAS analysis demonstrated that PanL and MaxLBZ might be under the same genetic control. In the first growing season data (Year 2010) revealed five positively high associations, *Xtxp25*, *Xtxp297*, *Xtxp50*, *Xtxp211 and Xtxp84*, were detected on the same chromosome, approximately 10 bp apart. QTL analysis of panicle elongation traits showed major QTL was located on Chr-2. These regions overlapped completely with clustered QTLs. The results suggested that the sorghum linkage groups are heavily populated with loci that are responsible for the inheritance of panicle elongation and dimensional traits of SDRS genotypes.

The highly significant association of the panicle diameter on Chr-2 (LG-B) has not been previously reported in linkage mapping studies. Additionally, QTLs for that trait in this location have not been detected by the previous analysis. The genomic region on Chr-4 for QTLs showed an interaction with not only the panicle traits but also yield-related traits, such as the total grain weight. Panicle dimensional traits, such as the panicle diameter and width (PanD, PanW), are important traits for panicle improvement (i.e., dense panicles leads to low grain quality due to a disorder in panicle dimension and size). It is worth noting that we were unable to identify novel QTLs controlling PanD and PanW on three chromosomes (Chr-1, Chr-2 and Chr-3). Additionally, QTL was involved on Chr-1 for PanW, Chr-2 for TotN and PanD, Chr- 6 for Rac, Chr-10 for PanT as new QTLs associated with inflorescence arcchitecture. In each growing season the effects and relative positions of PanL, PanD, TotBr and Rac were in accordance with the QTL distribution of several inter-related other panicle and yield component traits, such as MaxLBZ, TotN, GWP and PanN traits. Several genomic regions affecting multiple traits, including one region affecting PanL, Rac, MaxLBZ and PanN and another region that influenced PanD, TotBr and GWP. Nine QTLs for the number of total branches (TotBr) were found on six chromosomes consisting of three on Chr-1, two on Chr-2, and others on Chr-3, Chr-4, Chr-7 and Chr-8. Among these QTLs, the QTL regions on Chr-1 with *Xtxp43* and *Xtxp40* on Chr-7 matched with similar positions detected by using RIL population as previously reported Srinivas *et al.,* (2009). Similarly, the two QTLs were mapped for TotBr in the study at similar position as QTL for the branch length, which was located on Chr-1 (100cM) and on Chr-3 by Brown *et al.* (2006). These results demonstrate that QTLs for the branch length are common in this study and in earlier studies on the branch trait in sorghum. The QTL for branch length (MaxLBZ) on Chr-3 was matched with a similar position as the QTL for branch length as

reported by Brown *et al.* (2006). GWAS analysis of inflorescence architecture can identify novel loci on these locations potentially with the key traits for sorghum inflorescence architecture. This location for these traits has not been detected by the previous studies. These traits can be used for selection and classification purposes by simplify procedures of evaluations. For both the panicle diameter QTLs on Chr-2, the allele was associated with a greater panicle diameter is associated with increased node number along longer panicle length similar to the loose panicle type. Other panicle component traits, such as MaxLBZ and PanL, were found on chromosome 9 and chromosome 4 (LG-F and LG-D) similar to the most significant QTL in this study. In the Poaceae family, sorghum has been much less studied for genes affecting inflorescence architecture than other cereal crops. Brown *et al.* (2006) have mapped two QTLs for panicle primary branch number and a single QTL for secondary branching number while characterising the inflorescence architecture in sorghum. Srinivas *et al.* (2009) mapped five QTLs that were detected for branching patterns. Other QTLs for Rac, PanD, PanW, TotN, MaxLBZ, PanS and PanT in the study were not related to any of the QTL for panicle branching patterns that were identified in previous reports. Therefore, it is likely that these QTLs are new loci that regulate the panicle and its component traits in sorghum which involved in the inflorescence architecture. These results will serve as preliminary findings of QTLs for the genetic basis of the inflorescence architecture and component traits, and the further evaluation of the germplasm for these traits is in progress. However, this study will provide other possibilities for more detailed studies, such as increasing the population size and saturating the target genomic regions by adding large-scale molecular markers with powerful molecular techniques (*i.e.*, next generation sequencing, genome-wide SNP discovery, whole genome re-sequencing and map-based cloning of the genes controlling the inflorescence architecture) would be required in future studies.

Table 3.1. List of 98 sorrghum SSR primers; (Bhattramakki *et al.,* 2000, Kong *et al.,* 2000 and Taramino *et al.,* 1997).

Index	Locus	LG	No. alleles	Type(s) of SSR(s) amplified and number of repeats	Sequence of forward primer	Sequence of reverse primer	Size range (bp)	Ann. temp(ºC)
1	Xtxp316	A	7	$(AGA)_{12}$	CCAGCTTCACTTACGAGGAGATG	ATGCCCGTTTTCTAATTCTTCTACT	340-480	55
2	Xtxp248	A	8	$(AG)_5(GA)_{28}$	GGGTGTCCAATGTTGTCTGC	GGCCGTTACTGTCCCTTACTCA	190-250	50
3	Xtxp340	A	5	$(TAC)_{15}$	AGAACTGTGCATGTATTCGTCA	AGAAACTCCAATTATCATCCATCA	180-280	55
4	Xtxp319	A	5	$(TC)_{17}$	TAGACATCTGAATTAAGGAGC	CATGCCCCTGAAAGAGA	130-165	55
5	Xtxp61	A	6	$(GA)_{13}$	GATGCCCATGCCTTGC	CCCACTAAACTAAAGCGGACA	175-210	55
6	Xtxp284	A	4	$(AAG)_{19}$	CCAGATTGGCTGATGCATACACACT	AAGGGTAATTTATGCACTCCAAGGTAGGAC	200-245	60
7	Xtxp229	A	3	$(GT)_8$	TGCCCAAGAGGATAAAAGGT	AGCGACGGCACATCAAT	165-170	55
8	Xtxp279	A	5	$(CTT)_{10}+(CTT)_3+(CTT)_6$	ATTCTGACTTAACCCACCCCTAAA	AGCTCATCAATGTCCCAAACC	270-290	55
9	Xtxp75	A	6	$(TG)_{10}$	CGATGCCTCGAAAAAAAAACG	CCGATCAGAGCGTGGCAGG	140-170	50
10	Xtxp58	A	5	$(AG)_{13}+(GA)_{16}$	CAAAGTGCCCGGTTAAGACCT	TTCCCTTGCTGTTGCTTGTG	150-175	55
11	Xtxp335	A	5	$(GT)_{12}$	TATTTCCTCTTGAAAGAATCAGGG	TATTCATCGAGCAAAAGGCA	160-245	60
12	Xtxp37	A	5	$(TC)_{23}$	AAC CTA AGA GGC CTA TTT AAC C	ACG GCG ACT ATG TAA CTC ATA G	170-190	55
13	Xtxp32	A	7	$(AG)_{16}$	AGA AAT TCA CCA TGC TGC AG	ACC TCA CAG GCC ATG TCG	130-160	60
14	Xtxp88	A	7	$(AG)_{31}$	CGTGAATCAGCGAGTGTTGG	TGCGTAATGTTCCTGCTC	150-190	53
15	Xtxp149	A	3	$(CT)_{10}$	AGCCTTGCATGATGTTCC	GCTATGCTTGGTGTGGG	160-165	60
16	Xtxp43	A	7	$(CT)_{28}$	AGT CAC AGC ACA CTG CTT GTC	AAT TTA CCT GGC GCT CTG C	190-230	60
17	Xtxp302	A	6	$(TGT)_8$	TAGGTTCTGGACCACTTTTCTTTTTGTGTT	GAATCAACTATGTGCTTGCATTGTGCT	180-240	55
18	SbAGF06	A	10	$(AG)_{35}$	GTTAAACGACCAATCACCC	TAGAGGTGTCACTGATGAGC	145-220	54
19	SbAGB02	A	4	$(AG)_{35}$	CTCTGATATGTCGTTGTGCT	ATAGAGAGGATAGCTTATAGCTCA	145-170	54
20	Xtxp197	B	3	$(AC)_{10}$	GCGTCAATTAATCCAAACAGCCTC	GAGTTCCTATTCCCGTTCATGGTGAT	150-160	60
21	Xtxp96	B	5	$(GA)_{24}$	GCTGATGTCATGTTCCCTCAC	CATTCGTGGACTCTGTCGG	195-235	52
22	Xtxp25	B	9	$(CT)_{12}$	CCA TTG AGC TTC TGC TAT CTC	CAT TTG TCA CCA CTA GAA CCC	120-190	55
23	Xtxp297	B	8	$(AAG)_{24}$	GACCCATATGTGGTTTAGTCGCAAAG	GCACAATCTTCGCCTAAATCAACAAT	170-400	55
24	Xtxp50	B	3	$(CT)_{13}(CA)_9$	TGATGTTGTTACCCTTCTGG	AGCCTATGTATGTGTTCGTCC	300-310	55
25	Xtxp211	B	6	$(CT)_{23}$	TCAACGGCCAATGATTTCTAAC	AGGTTGCGAATAAAAGGTAATGTG	180-240	55
26	Xtxp84	B	3	$(AG)_9$	CCGATCAGCACACCAG	GTACTAGGTCCAATCCAGC	210-240	50
27	Xtxp4	B	6	$(GA)_{23}$	AAT ACT AGG TGT CAG GGC TGT G	ATG TAA CCG CAA CAA CCA AG	145-180	55
28	Xtxp201	B	3	$(GA)_{36}$	GCGTTTATGGAAGCAAAAT	CTCATAAGGCAGGACCAAC	225-265	60
29	Xtxp19	B	3	$(AG)_5+(AG)_{10}$	CTT TCA ATC GGT TCC AGA C	CTT CCA CCT CCG TAC TC	270-300	55
30	Xtxp13	B	4	$(TG)_{13}$	TCT TTC CCA AGG AGC CTA G	GAA GTT ATG CCA GAC ATG CTG	120-175	55
31	Xtxp298	B	6	$(AGA)_{23}$	GCATGTGTCAGATGATCTGGTGA	GCTGTTAGCTTCTTCTAATCGTCGGT	175-210	55
32	Xtxp1	B	8	$(AG)_{34}$	TTG GCT TTT GTG GAG CTG	ACC CAG CAG CAC TAC ACT AC	170-220	55
33	Xtxp56	B	4	$(GA)_{39}$	TGTCTTCGTAGTTGCGTGTTG	CCGAAGGAGTGCTTTGGAC	310-450	60
34	Xtxp286	B	5	$(GCA)_4ACA(GCA)_5A(CAA)_5+(AAC)_9$	AGCAGCAGCAGCAACAG	GCGTGGTCTTTGTGGTTC	190-220	55
35	Xtxp348	B	4	$(TAA)_{37}$	CGACATCAGCGTTGTCTTTCTA	GCTTACGAATAGGGCAAAAGAACT	280-325	60

Table 3.1. (Continued)

Index	Locus	LG	No. alleles	Type(s) of SSR(s) amplified and number of repeats	Sequence of forward primer	Sequence of reverse primer	Size range (bp)	Ann. temp(°C)
36	Xtxp315	B	5	$(TAT)_{22}(CAT)_{18}CGT(CAT)_4$	AACCTACCCTAGCCCGCGCAACTG	CAACATGTCCGCAAGATTTGATGTGAC	230-350	55
37	Xtxp100(Kaf)	B	4	$(CT)_{19}$	CCGGCCGGCCAACCAACCAC	TGCCCCAACGCTCACGCTCCC	115-135	55
38	Xtxp7	B	5	$(CT)_{14}$	ACA TCT ACT ACC CTC TCA CC	ACA CAT CGA GAC CAG TTG	220-245	50
39	Xtxp207	B	4	$(CT)_{14}$	ACACATCTACTACCCTCTCACCCT	TGATAGACTTGTGAGCAGCTCC	170-190	55
40	Xtxp296	B	3	$(CA)_{18}$	CAGAAATAACATATAATGATGGGGTGAA	ATGCTGTTATGATTTAGAGCCTGTAGAGTT	150-170	55
41	Xtxp8	B	9	$(TG)_{31}$	ATA TGG AAG GAA GAA GCC GG	AAC ACA ACA TGC ACG CAT G	110-165	60
42	Xtxp69	C	6	$(TC)_{12}$	ACACGCATGGTTTGACTG	TTGATAATCTGACGCAACTG	180-250	50
43	Xtxp285	C	5	$(CTT)_{11}CTC(CTT)_{16}$	ATTTGATTCTTCTTGCTTTGCCTTGT	TTGTCATTTCCCCCTTCTTTCTTTT	205-260	60
44	**Xtxp38 (lg)**	C	4	$(AG)_{17}$	ACA AAC CGC GAC GAA GTA AC	ACA AGG CAA AGC ACA AAG C	430-500	60
45	Xtxp59	C	2	$(GGA)_5$	GAAATCCACGATAGGGTAAGG	GACCCAGAATAGAAGAGAGG	195-200	60
46	Cba	C	3	$(TA)_{18}$	AAAGCTCGGCGTTAGAAATA	CGCTTAACAACTCCTACCATC	195-230	60
47	Xtxp336	C	3	$(CGG)_4+(GAG)_6$	CAGCGAGCACCGACGAC	CCACCCAACCTGACCCTTCT	160-170	55
48	Xtxp31	C	7	$(CT)_{25}$	TGC GAG GCT GCC CTA CTA G	TGG ACG TAC CTA TTG GTG C	190-260	60
49	Xtxp205	C	3	$(AG)_{12}$	CCTGCCGTGTCTTCC	TATATGCATGCCGTAGATTT	190-200	60
50	Xtxp33	C	8	$(TC)_{20}$ $C(TG)_5+(CT)_9CC(TG)_7$	GAG CTA CAC AGG GTT CAA C	CCT AGC TAT TCC TTG GTT G	165-230	55
51	Xtxp228	C	3	$(TC)_{12}$	ACAGGTTGGCGATGTTTCTCT	TTCTTTTTCGAATTCATTCCTTTT	230-250	60
52	Xtxp266	C	2	(GT)8	GTTGTCTAGTATAGCAAGGTGGG	ATAATAGTAGATGCGTGTCAAAAGAA	190-195	55
53	Xtxp12	D	7	$(CT)_{22}$	AGA TCT GGC GGC AAC G	AGT CAC CCA TCG ATC ATC	155-190	55
54	Xtxp24	D	7	$(TC)_{21}$	TTG TGT AGT CCA TCC GAT GC	TTC TAA GCC CAC CGA AGT TG	120-160	60
55	Xtxp60	D	2	$(GT)_4GC(GT)_5$	GCTAGCTGACGCACGTCTCTG	TGCAACCGAGCGGTGACTA	220-225	60
56	**Xtxp212**	D	3	$(GT)_{10}$	TTTCCCCTCTTTCTTGTGTC	CTCGGCGTCGTCGTA	135-145	60
57	Xtxp51	D	2	$(TG)_{11}$	TCTCGGACTCAAGAGCAGAGG	GGACAGCAGCGGCTTCAG	225-230	60
58	Xtxp27	D	5	$(AG)_{37}$	AAC CTT GCC CTA TCC ACC TC	TAT GAT GAA TCA AGG GAG AGG	290-320	45
59	Xtxp21	D	5	$(AG)_{18}$	GAG CTG CCA TAG ATT TGG TCG	ACC TCG TCC CAC CTT TGT TG	160-195	60
60	Xtxp40	E	3	(GGA)7	CAG CAA CTT GCA CTT GTC	GGG AGC AAT TTG GCA CTA G	135-145	55
61	Xtxp36	E	2	$(GGA)_7GTA(T)_7+(A)_7$	ATG GGA CGG AAA TGC AGG AG	TTA TGC CTG CCA GCA ACT TG	180-185	60
62	Xtxp159	E	5	$(CT)_{21}$	ACCCAAAGCCCAAATCAG	GGGGGAGAAACGGTGAG	160-180	55
63	Xtxp312	E	6	$(CAA)_{26}$	CAGGAAAATACGATCCGTGCCAAGT	GTGAACTATTCGGAAGAAGTTTGGAGGAAA	90-185	60
64	Xtxp278	E	2	$(TTG)_{12}$	GGGTTTCAACTCTAGCCTACCGAACTTCCT	ATGCCTCATCATGGTTCGTTTTGCTT	240-250	60
65	**Xtxp92**	E	2	$(GAA)_5$	ACTTGCAGGTTAATTTCGTCC	GGCGAGCTTGCGGTAG	155-165	60
66	Xtxp295	E	4	$(TC)_{19}$	AAATCATGCATCCATGTTCGTCTTC	CTCCCGCTACAAGAGTACATTCATAGCTTA	145-160	60
67	SbAGE03	E	6	$(AG)_{34}GA(CA)_4$	AGCTCTCAGCCTTTCACAAT	GGAAGAAAGGAATGACTTGA	85-140	54
68	Xtxp10	F	3	$(CT)_{14}$	ATA CTA TCA AGA GGG GAG C	AGT ACT AGC CAC ACG TCA C	125-140	50
69	Xtxp67	F	4	$(GA)_{28}$	CCTGACGCTCGTGGCTACC	TCCACACAAGATTCAGGCTCC	145-170	60
70	Xtxp287	F	4	$(AAC)_{21}$	GCAAGCGAGCTGACTTATGTAACGAGA	CAAAGTGCTACTAAACCTATGCAGGGTGAA	330-350	60

Table 3.1. (Continued)

Index	Locus	LG	No. alleles	Type(s) of SSR(s) amplified and number of repeats	Sequence of forward primer	Sequence of reverse primer	Size range (bp)	Ann. temp(ºC)
71	Xtxp258	F	5	$(AAC)_{19}$	CACCAAGTGTCGCGAACTGAA	GCTTAGTGTGAGCGCTGACCAG	180-225	60
72	SbAGB03	F	9	$(AG)_{41}$	GTGTGTGTAGCTTCTTGGG	ACGTAGGAGTAGTTTCTAGGATT	90-145	54
73	Xtxp217	G	4	$(GA)_{23}$	GGCCTCGACTACGGAGTT	TCGGCATATTGATTTGGTTT	150-170	60
74	Xtxp20	G	5	$(AG)_{21}$	TCT CAA GGT TTG ATG GTT GG	ACC CAT TAT TGA CCG TTG AG	180-240	60
75	Xtxp270	G	6	$(GAA)_{12}(GAAA)_6 + (GAA)_{21} + (GTA)_5 + (GTA)_3 + (GTA)_3$	AGCAAGAAGAAGGCAAGAAGAAGG	GCGAAATTATTTTGAAATGGAGTTGA	280-320	60
76	Xtxp331	G	9	$(GAT)_{32}$	AACGGTTATTAGAGAGGGAGA	AGTATAATAACATTTTGACACCCA	170-280	60
77	PepC	G	3	$(AT)_{10}$	TGGGAAGCAGCTCAGG	AGGGTGGTGATGTAGGGA	200-220	60
78	**Xtxp273(Pbbf)**	H	4	$(TTG)_{20}$	GTACCCATTTAAATTGTTTGCAGTAG	CAGAGGAGGAGGAAGAGAAGG	180-220	60
79	Xtxp47	H	2	$(GT)_8(GC)_5 + (GT)_6$	CAATGGCTTGCACATGTCCTA	GGTGCGAGCTAGTTAAGTGGG	280-290	60
80	Xtxp294	H	3	(TG)10(GT)4	GCTGGGGCTCGAGGGTTTTCATT	AGCTTCCCAAGGACAACTAGCAAGGACA	280-290	50
81	Xtxp354	H	5	$(GA)_{21} + (AAG)_3$	TGGGCAGGGTATCTAACTGA	GCCTTTTTCTGAGCCTTGA	130-170	60
82	Xtxp18	H	7	(AG)21	ACT GTC TAG AAC AAG CTG CG	TTG CTC TAG CTA GGC ATT TC	220-270	55
83	Xtxp250	H	5	(AAG)17AAT(AAG)4AAA$(ACA)_9$	GCACATCCTCTAAAACTACTTAGT	GAACAGGACGATGTGATAGAT	270-300	50
84	Xtxp321	H	7	$(GT)_4+(AT)_6+(CT)_{21}$	TAACCCAAGCCTGAGCATAAGA	CCCATTCACACATGAGACGAG	180-220	55
85	Xtxp105	H	4	$(TG)_5 + (CT)_6GTCT(GT)_7$	TGGTATGGGACTGGACGG	TGTTGACGAAGCAACTCCAAT	290-325	60
86	SbAGA01	H	6	$(AG)_{33}$	CGAACCATGATAAATGACTG	ATCCGTTTCACAAAAAAAGT	95-120	54
87	Xtxp145	I	5	$(AG)_{22}$	GTTCCTCCTGCCATTACT	CTTCCGCACATCCAC	200-230	60
88	Xtxp274	I	6	$(TTC)_{19}$	GAAATTACAATGCTACCCCTAAAAGT	ACTCTACTCCTTCCGTCCACAT	280-320	60
89	Xtxp104	I	3	$(GGC)_6+(GT)_7$	TAACCTATGCGGATAAAACAG	GAATCGCTGCCAAATAAA	175-185	50
90	Xtxp97	I	3	$(CA)_8+(GCC)_6$	CAAATAAACGGTGCACACTCA	GTATGATTGGAGACGAGACGG	120-130	60
91	Xtxp95	I	5	$(GA)_{18}(GC)_4$	TCTCCGTTTGCCCGCCAG	CACCGTACCGCCTCCCGAATC	75-95	53
92	Xtxp65	J	3	$(ACC)_4+(CCA)_3CG(CT)_8$	CACGTCGTCACCAACCAA	GTTAAACGAAAGGGAAATGGC	120-130	60
93	Xtxp303	J	4	$(GT)_{13}$	AATGAGGAAAATATGAAACAAGTACCAA	AATAACAAGCGCAACTATATGAACAATAAA	140-160	60
94	Xtxp15	J	4	$(TC)_{16}$	CAC AAA CAC TAG TGC CTT ATC	CAT AGA CAC CTA GGC CAT C	190-220	55
95	Xtxp14	J	6	$(GA)_{15}$	GTA ATA GTC ATG ACC GAG G	TAA TAG ACG AGT GAA AGC CC	125-150	50
96	Xtxp23	J	4	$(CT)_{19}$	AAT CAA CAA GAG CGG GAA AG	TTG AGA TTC GCT CCA CTC C	175-195	60
97	Kaf2	J	3	$(CAA)_9$	TCGGCGAGCATCTTACA	TACGTAGGCGGTTGGATT	260-275	60
98	SbKAFGK1	J	4	$(ACA)_9$	AGCATCTTACAACAACCAAT	CTAGTGCACTGAGTGATGAC	120-140	54

Locus with bold face has homology with genes. LG; stands for linkage group (Bhattramakki *et al.* 2000)

Table 3.2. Summary statistics of 98 SSR markers estimated for 107 sorghum accessions (sorghum diversity research set).

A; Based on three populations (i.e. *J*=3) inferred from structure analysis.

Population (*J*)	No. of accessions	Panicle type (number)	No. of alleles	Gene diversity (He)	Allele richness (*Rt*)	F_{is}
J=1	35	O(2), I(8), S(6), C(12), B(7)	413	0.60	2.74	0.99
J=2	35	O(11), I(8), S(5), C(8), B(2), M(1)	405	0.57	2.65	0.99
J=3	37	O(4), I(8), S(2),C(14), B(7), M(2)	398	0.58	2.69	0.98
Total	107	107	470	0.64	4.6	0.99

Legend for panicle types: O: open type, I: intermediate type, S: semi-compact type, C: compact type, B: broom type, M: mixed type.

B; Based on the geographic distribution of accessions in five regions

Regions	No. of accessions	No. of alleles	Gene diversity (*He*)	F_{is}
East Asia	25	370	0.56	0.99
Southeast Asia	2	146	0.71	0.99
South Asia	26	415	0.62	0.98
Southwest Asia	2	138	0.64	0.99
Africa	52	444	0.62	0.99

Table 3.3.Genome wide association analysis for panicle traits and yield related traits of 107 sorghum accessions.

No.	Trait	Marker Index	Chromosome	Marker	[Year.2010] -Log 10 (*P*-value)	[Year.2011] -Log 10 (*P*-value)
1	Pend	M15	1	*Xtxp316*	3.31	0.71
	Pend	M27	2	*Xtxp50*	2.41	1.13
	Pend	M28	2	*Xtxp84*	6.23	0.06
	Pend	M64	7	*SbAGE03*	0.43	2.48
	Rac	M28	2	*Xtxp84*	8.69	1.02
	Rac	M68	9	*Xtxp10*	4.29	1.12
	Rac	M90	6	***Xtxp95***	2.19	1.55
3	PanL	M05	1	***SbAGF06***	3.59	1.30
	PanL	M09	1	*Xtxp75*	2.28	0.59
	PanL	M26	2	***Xtxp7***	2.72	1.82
	PanL	M33	2	*Xtxp96*	2.00	0.48
	PanL	M36	2	*Xtxp211*	2.66	0.44
	PanL	M45	3	*Xtxp228*	2.32	0.12
	PanL	M50	3	*Xtxp336*	3.00	0.20
	PanL	M56	4	***Xtxp12***	2.62	1.61
	PanL	M59	4	*Xtxp27*	2.07	0.47
	PanL	M68	9	***Xtxp10***	5.06	2.54
	PanL	M84	8	*Xtxp321*	1.19	2.31
4	TotN	M11	1	Xtxp229	2.00	0.25
	TotN	M17	1	*Xtxp340*	6.02	0.51
	TotN	M41	2	***Xtxp315***	2.62	2.48
	TotN	M56	4	*Xtxp12*	3.07	0.8
5	TotBr	M11	1	*Xtxp229*	2.33	0.55
	TotBr	M18	1	*Xtxp37*	1.07	3.78
	TotBr	M19	1	*Xtxp43*	0.29	2.96
	TotBr	M25	2	*Xtxp4*	2.55	0.10
	TotBr	M41	2	*Xtxp315*	0.29	6.45
	TotBr	M47	3	*Xtxp31*	2.04	0.14
	TotBr	M66	7	*Xtx40*	0.39	3.61
	TotBr	M82	8	***SbAGA01***	1.47	3.19
	TotBr	M55	4	*Xtxp212*	3.20	0.23

Table 3.3. (Continued)

No.	Trait	Marker Index	Chromosome	Marker	[Year.2010] -Log 10 (*P*-value)	[Year.2011] -Log 10 (*P*-value)
6	MaxLBZ	M07	1	*Xtxp32*	2.03	0.23
	MaxLBZ	M27	2	*Xtxp50*	4.88	0.87
	MaxLBZ	M39	2	*Xtxp297*	4.23	0.78
	MaxLBZ	M56	4	***Xtxp12***	2.00	1.95
	MaxLBZ	M89	6	*Xtxp274*	3.22	0.34
	MaxLBZ	M84	8	*Xtxp321*	2.04	0.36
	MaxLBZ	M68	9	***Xtxp10***	7.63	4.94
7	PanD	M05	1	*SbAGF06*	0.15	3.02
	PanD	M17	1	*Xtxp340*	2.11	0.57
	PanD	M29	2	***Xtxp8***	2.17	1.58
	PanD	M32	2	*Xtxp25*	10.37	0.89
	PanD	M52	3	*Xtxp266*	2.46	0.03
	PanD	M56	4	*Xtxp12*	2.23	0.63
	PanD	M96	5	*Xtxp14*	0.25	2.49
	PanD	M90	6	*Xtxp95*	2.04	0.50
	PanW	M14	1	***Xtp302***	1.74	1.49
	PanW	M045	3	***Xtxp228***	1.80	5.06
	PanW	M095	5	*SbKAFGK1*	0.07	2.19
9	PanS	M015	1	***Xtxp316***	1.68	1.65
	PanS	M098	5	***Xtxp23***	1.47	1.48
10	PanT	M14	1	***Xtxp302***	2.39	1.99
	PanT	M89	6	***Xtxp274***	1.50	1.50
	PanT	M73	10	***PepC***	1.33	1.51
11	PanEx	M16	1	*Xtxp319*	2.35	0.49
	PanEx	M82	8	*SbAGA01*	1.21	2.68
12	PanN	M05	1	*SbAGF06*	2.48	0.06
	PanN	M29	2	*Xtxp8*	3.10	0.27
	PanN	M34	2	*Xtxp100(Kaf)*	0.72	3.65
13	GNP	M55	4	*Xtxp212*	4.11	0.57
	GNP	M56	4	*Xtxp12*	4.11	0.57
	GNP	M69	9	*Xtxp67*	2.25	1.02
14	GWP	M22	2	*Xtxp201*	2.01	0.10
	GWP	M55	4	***Xtxp212***	2.15	1.89

MODELING ASSOCIATION: The Mixed Linear Model (Q+K), $-\text{Log}_{10}$ (*P)* values = 1.3 as threshold value (threshold level corresponds to a test at the 0.05 level of significant, respectively), *Legend for table:* The makers in bold and underline are QTLs identified in both years. *Legend for trait:* Pend: peduncle length, Rac: rachis length, PanL: panicle length, MaxLBZ: maximum length of primary branch zone, TotN: total node along rachis, TotBr: total branch per panicle, PanD: panicle diameter, PanW: panicle width, PanS: panicle shape, PanT: panicle type, PanEx: panicle exsersion, GNP: grain number per panicle, GWP: total grain weight per panicle, PanN: panicle number per plant.

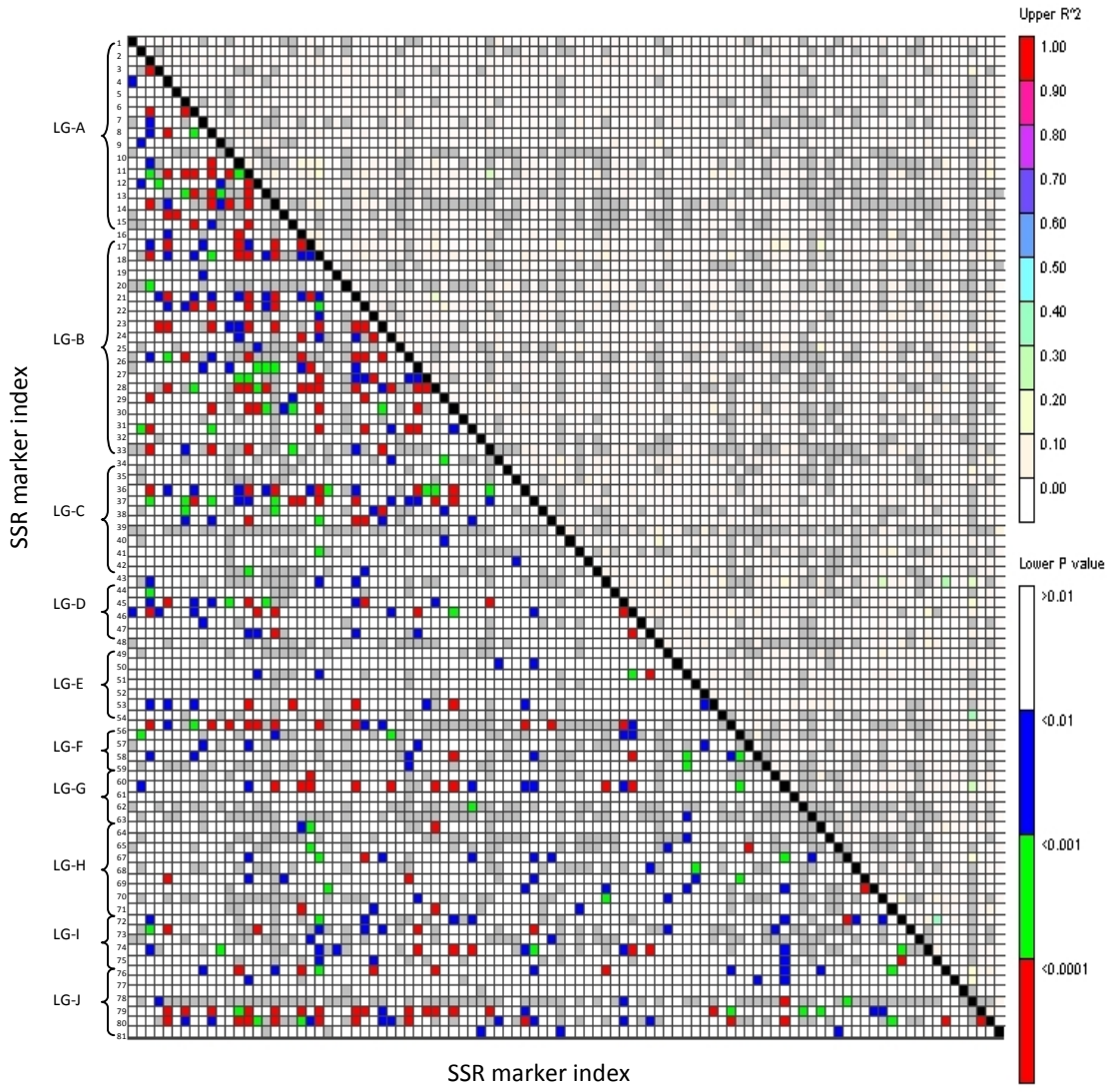
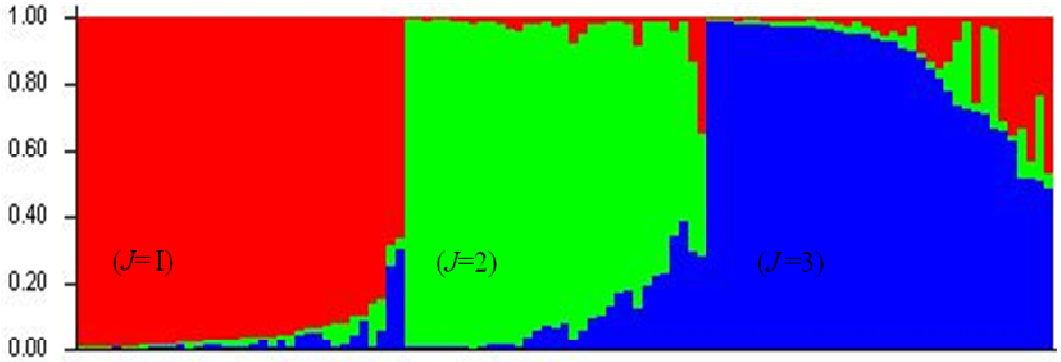


Figure 3.1. LD plot generated by SSR markers.

Legend for figure: Each cell represents the comparison of two pairs of marker sites with the color codes for the presence of significant LD. A colored bar code for the significant threshold levels.)

(i)



(ii)

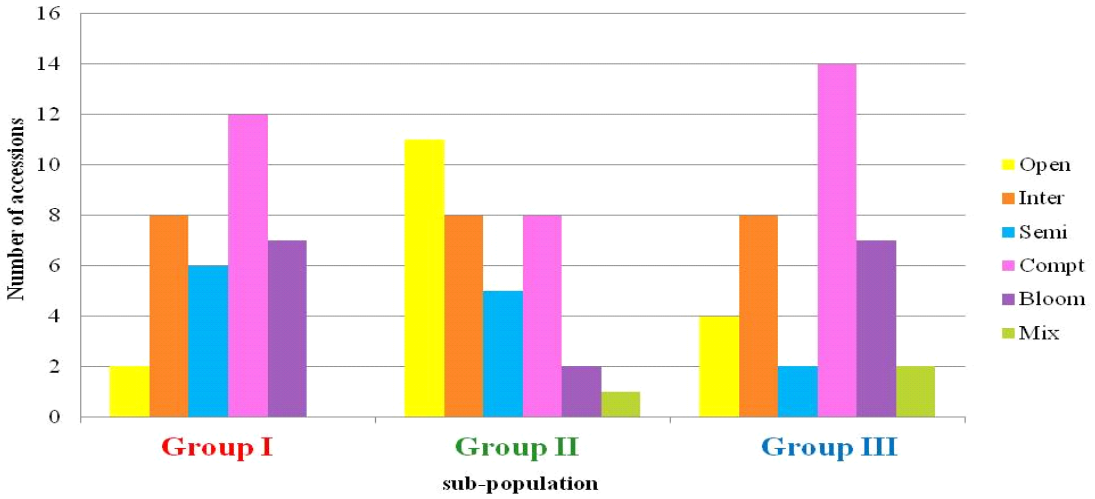


Figure 3.2. (i) Structure of the sorghum diversity research set divided into three sub-populations (*j*=3). (ii) Distribution of different panicle types among three sub-populations.

Legend for figure (i): Each color represents a subpopulation based on STRUCTURE results. *Legend for figure* (ii): Number of accessions involved in specific panicle type (y) within each sub-population (x).

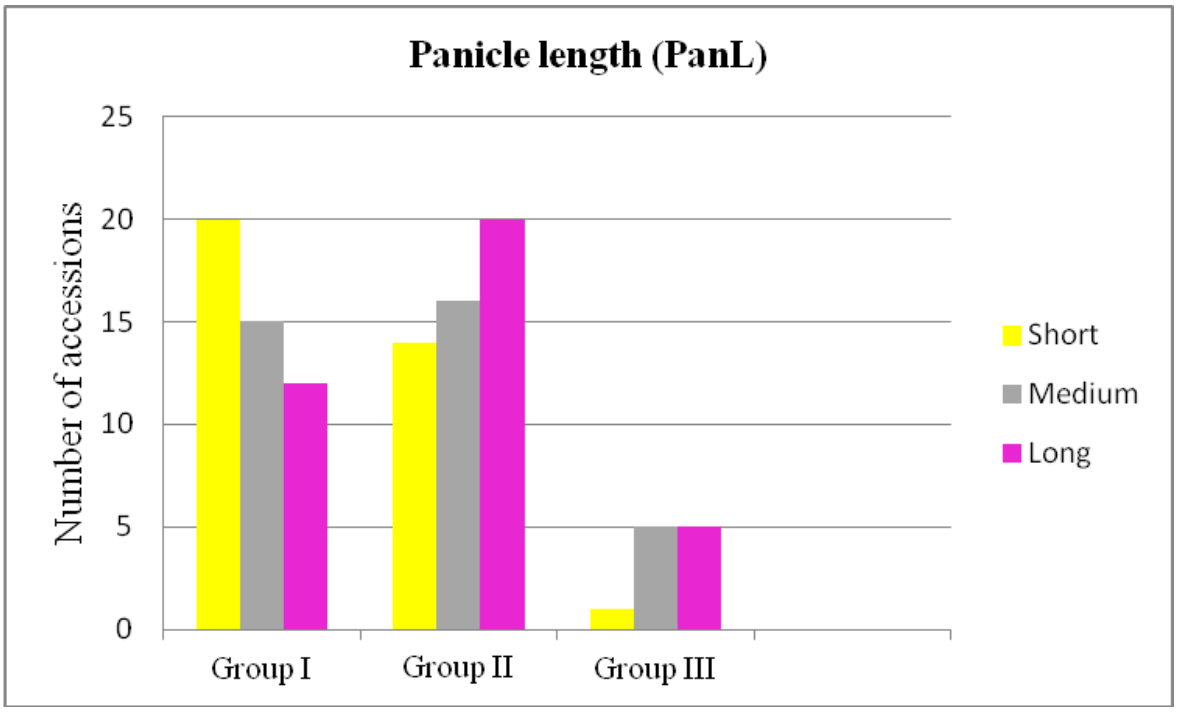


Figure 3.3. Distribution of accessions based on different range of panicle length among three sub-populations. [This figure is developed based on Fig.3.2 (i) Structure of the SDRS divided into three sub-populations]

Legend for figure (ii): Number of accessions involved in different range of panicle length (y) within each sub-population (x). Short: short panicle length (<21 cm), Medium: medium panicle length (<41cm), Long: long panicle length (<63cm).

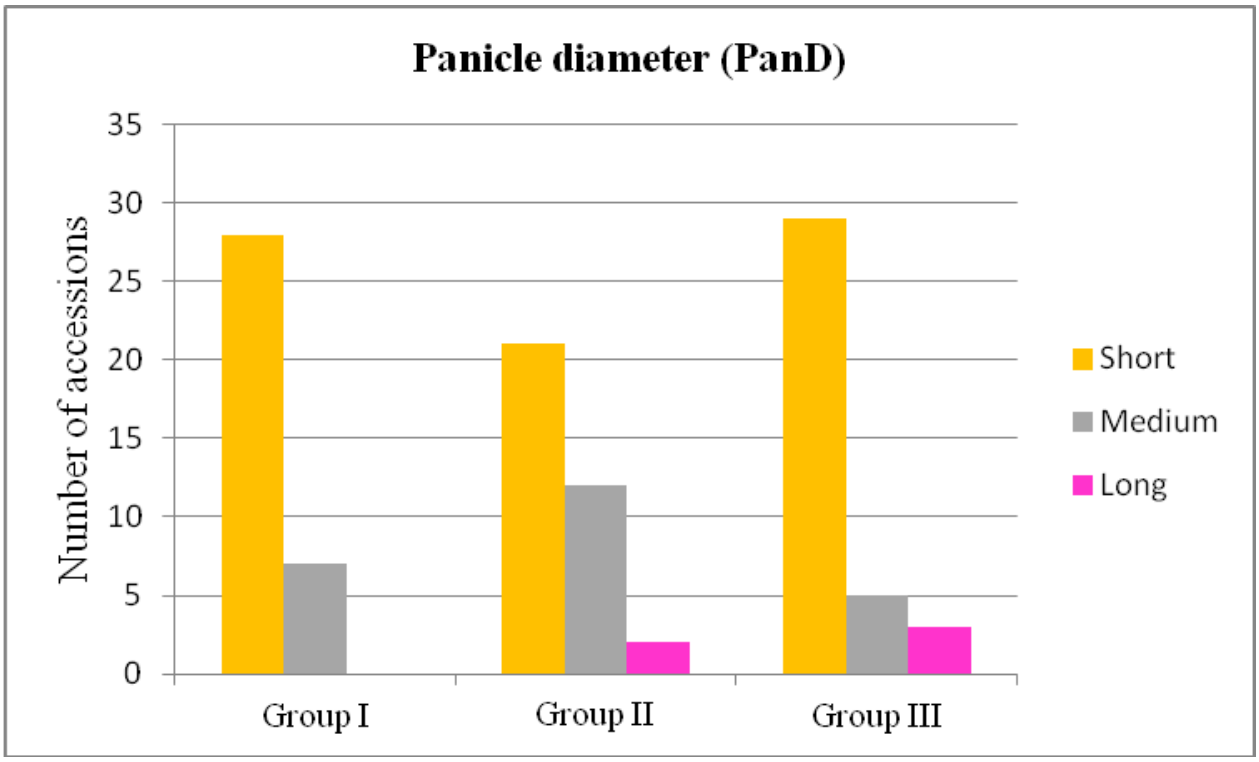


Figure 3.4. Distribution of accessions based on different range of panicle diameter among three subpopulations. [This figure is developed based on Fig.3.2 (i) Structure of the SDRS divided into three sub-populations]

Legend for figure: Number of accessions involved in different range of panicle diameter *(y)* within each sub-population *(x)*. Short: short panicle diameter (<4cm), Medium: medium panicle diameter (<8cm), Long: long panicle diameter (<12cm).

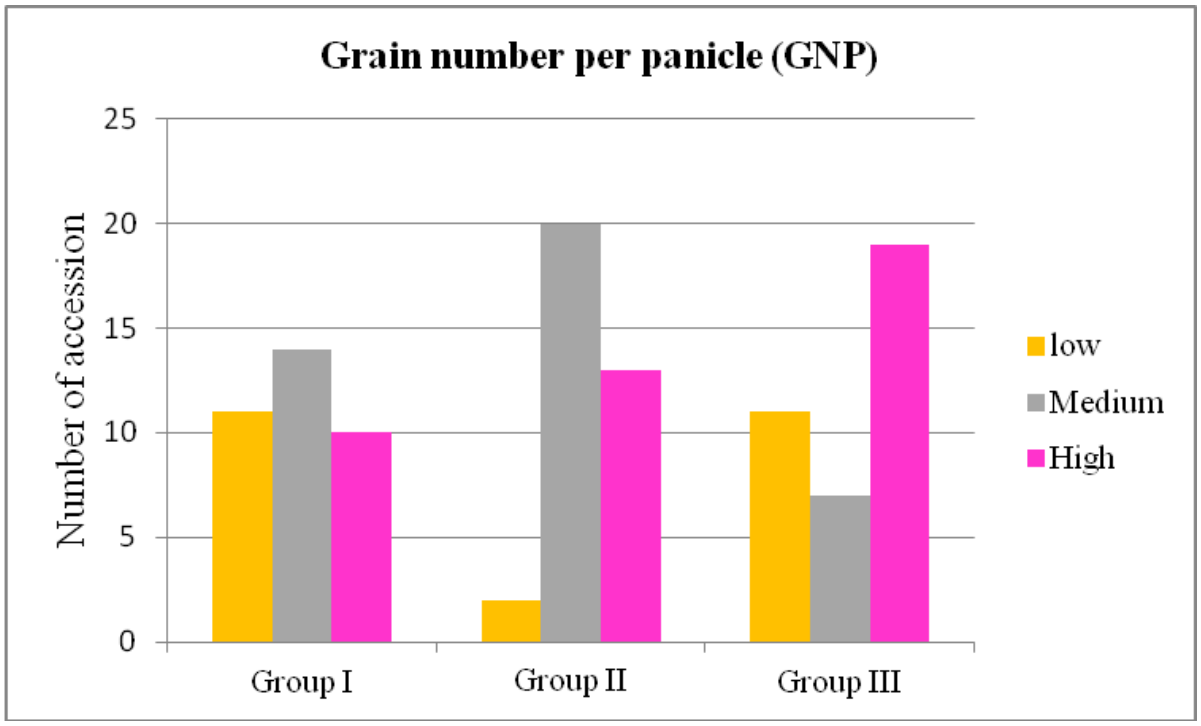


Figure 3.5. Distribution of accessions based on different range of grain number among three subpopulations. [This figure is developed based on Fig.3.2 (i) Structure of the SDRS divided into three sub-populations]

Legend for figure: Number of accessions involved in different range of grain number *(y)* within each sub-population *(x)*. Low: Low grain number (<1500), Medium: medium grain number (<3000), High: high grain number (<4500).

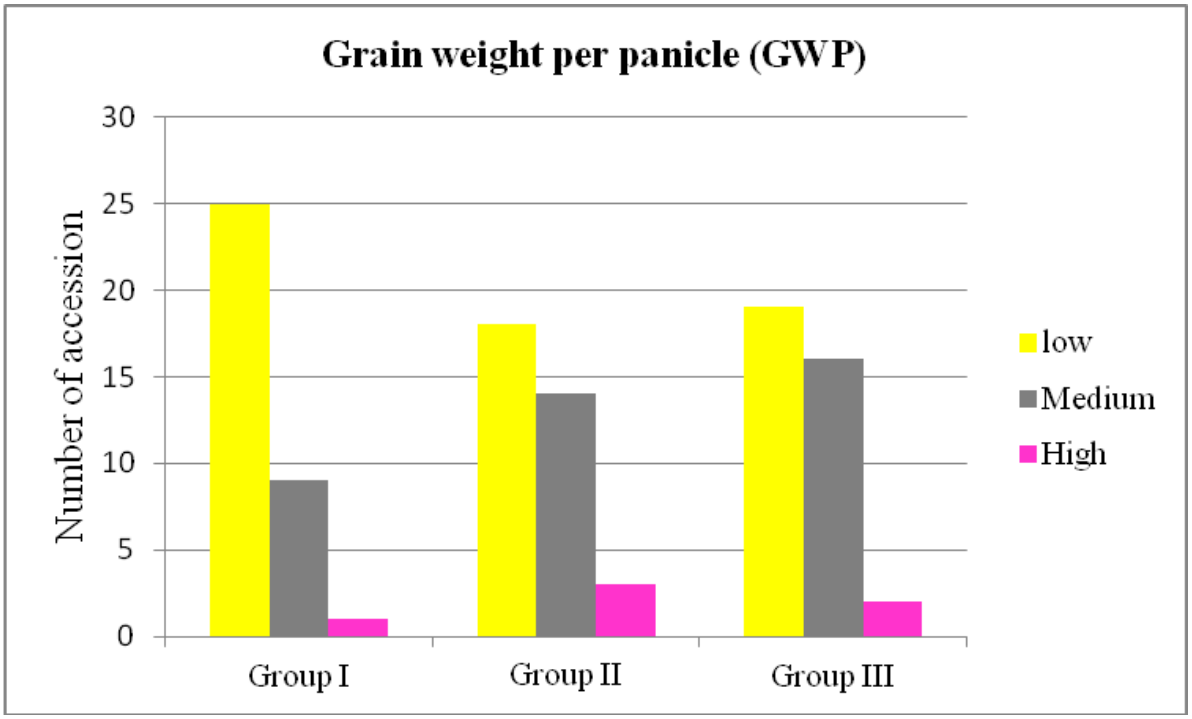
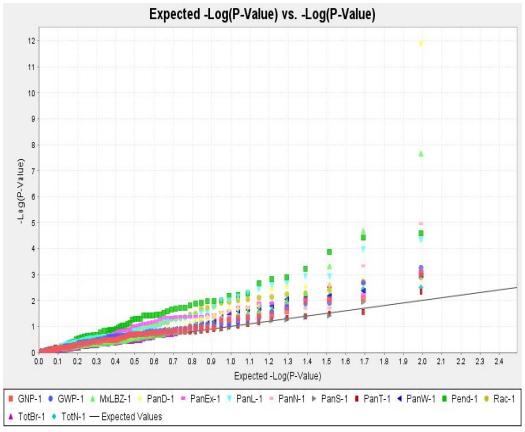


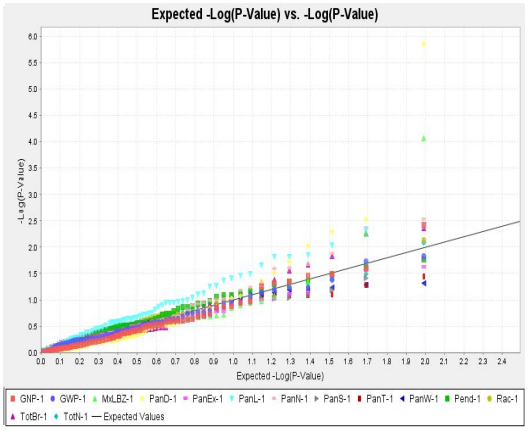
Figure 3.6. Distribution of accessions based on different range of grain weight among three subpopulations. [This figure is developed based on Fig.3.2 (i) Structure of the SDRS divided into three sub-populations]

Legend for figure: Number of accessions involved in different range of grain weight *(y)* within each sub-population *(x)*. Low: Low grain weight (<38g), Medium: medium grain weight (<76 g), High: high grain weight (<114 g).

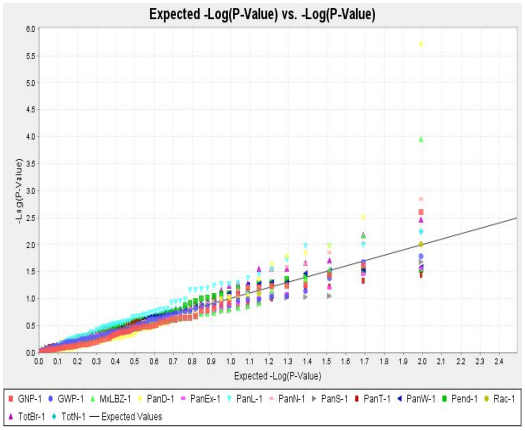
(i) Naïve



(ii) K



(iii)Q



(iv) Q+K

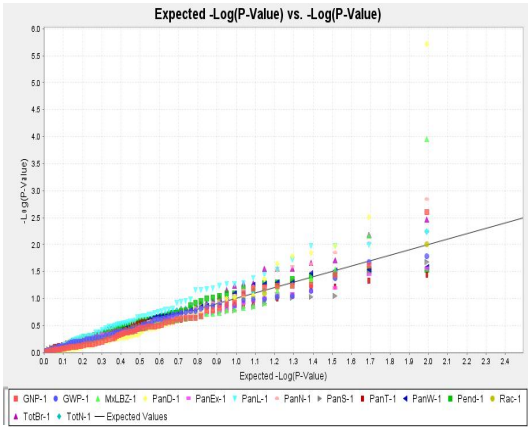


Figure 3.7. Quantile-quantile plots of the inflorescence architecture and yield related traits with 98 SSRs markers.

Legend for figure: The Q-Q plots showed the variation of observed *P*-value (x) against the expected *P*-values (y) using GLM and MLM models and control of type I error by the selected models. *Legend for trait*: Pend: peduncle length, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, TotBr: total branch per panicle, MaxLBZ: maximum length of primary branch, PanD: panicle diameter, PanW:Panicle width, PanS: panicle shape, PanT: panicle type, PanEx: Panicle exsersion, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle.

(i)

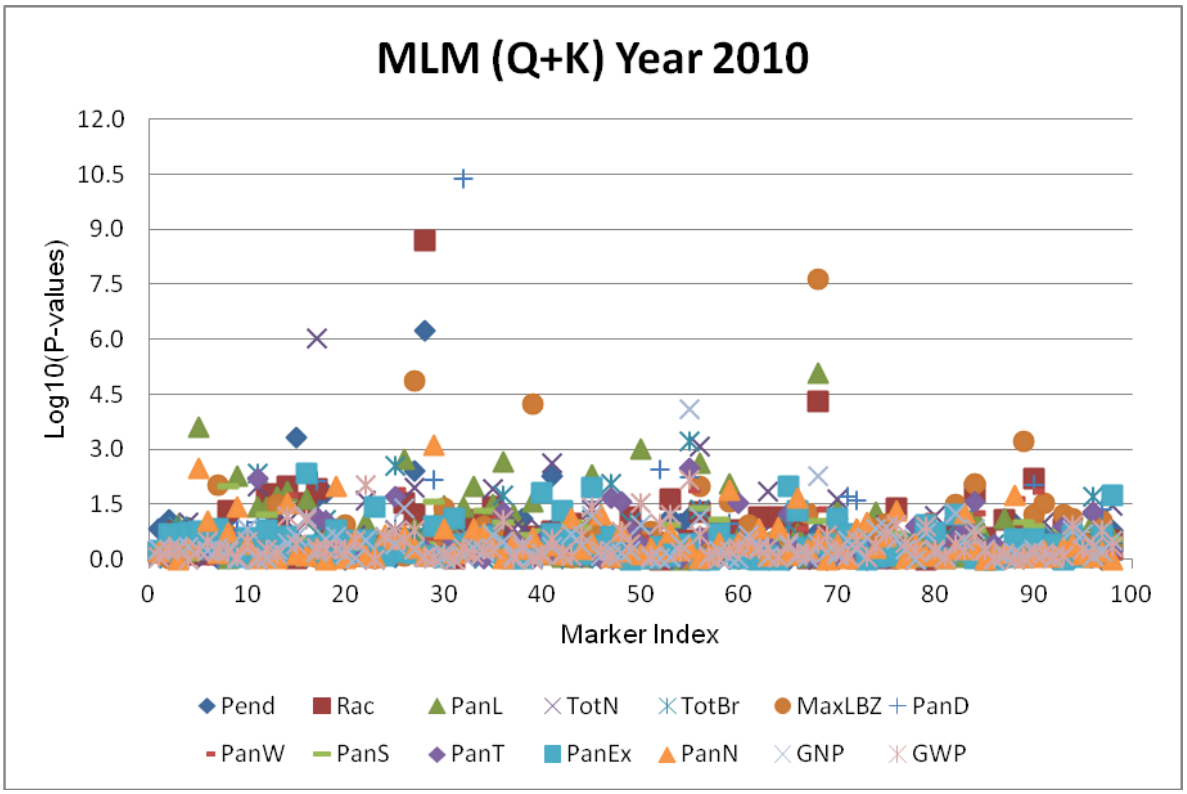


Figure 3.8. (i) Association analysis of 98 SSRs markers and 14 panicle traits by using Q+K model for 107 sorghum accessions in Year 2010 growing season.

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, TotBr: total branch per panicle, MaxLBZ: maximum length of primary branch, PanD: panicle diameter, PanW: panicle width, PanS: panicle shape, PanT: panicle type, PanEx: Panicle exsertion, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle.

(ii)

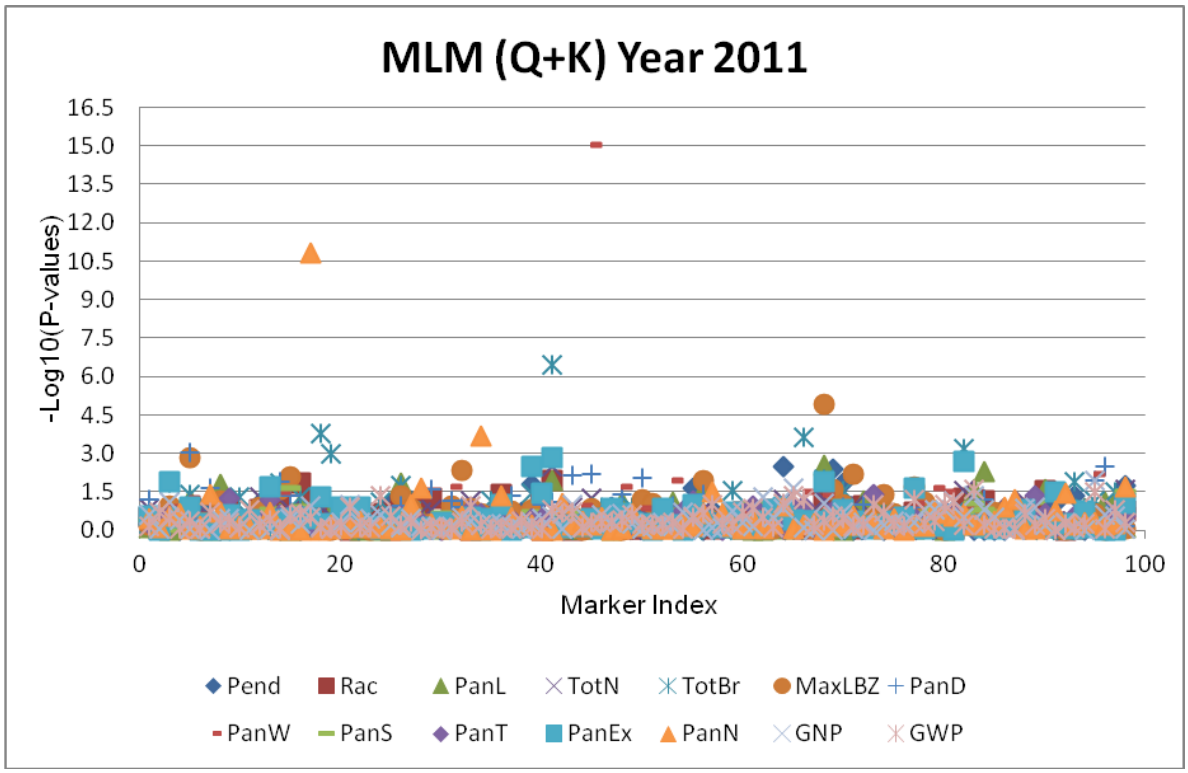


Figure 3.8.(cont) (ii) Association analysis of 98 SSRs markers and 14 panicle traits by using Q+K model for 107 sorghum accessions in Year 2011 growing season.

Legend for trait: Pend: peduncle length, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, TotBr: total branch per panicle, MaxLBZ: maximum length of primary branch, PanD: panicle diameter, PanW: panicle width, PanS: panicle shape, PanT: panicle type, PanEx: Panicle exsersion, PanN: panicle number per plant, GNP: grain number per panicle, GWP: total grain weight per panicle.

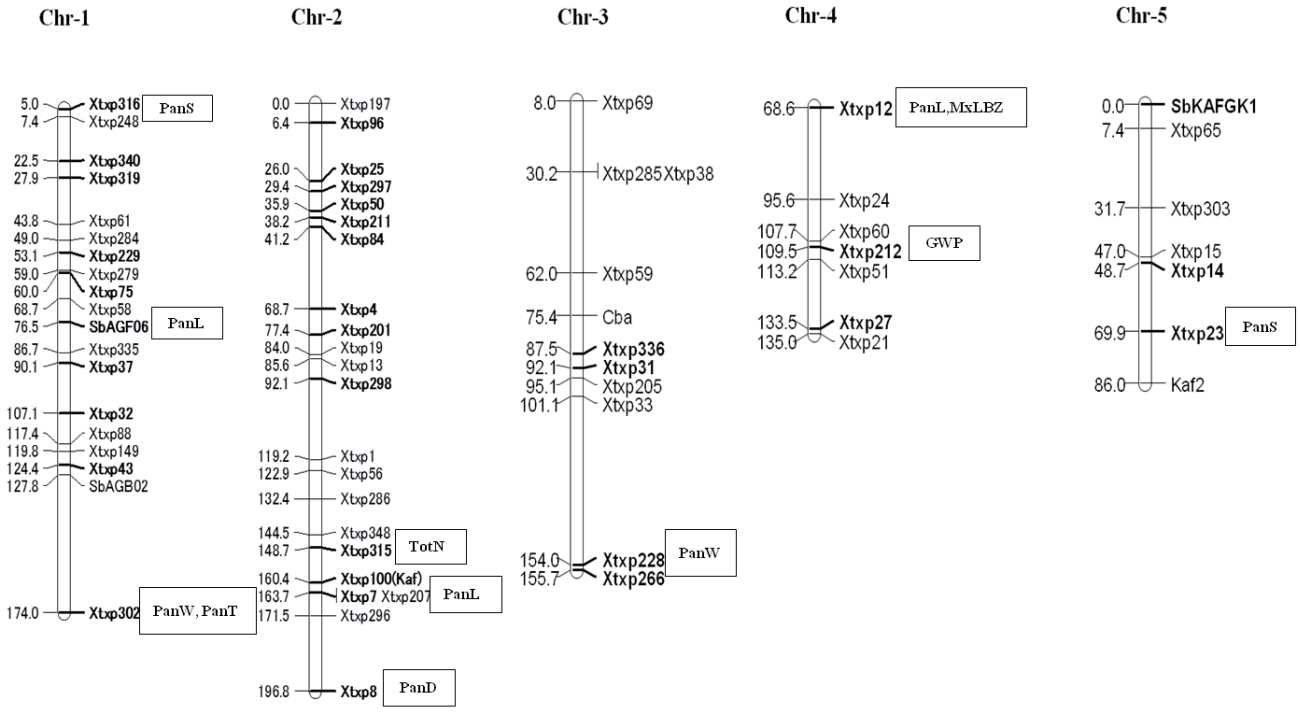


Figure 3.9. QTLs for sorghum inflorescence traits and location of SSRs on ten sorghum chromosomes are shown in order as describe in (Bhattramakki *et al.*, 2000; Kong *et al.,* 2000; Taramino *et al.,* 1997) and markers with bold face shows significant association with traits as resolved by Q + K model.

Legends for QTLs: 15 QTLs identified for 10 panicle traits and yield related traits in both years are enclosed within rectangle. *Legend for traits:* TotBr: total branch per panicle, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, PanD: panicle diameter, PanW: panicle width, MaxLBZ: maximum length of primary branch, GWP: total grain weight per panicle, PanS: panicle shape, PanT: panicle type.

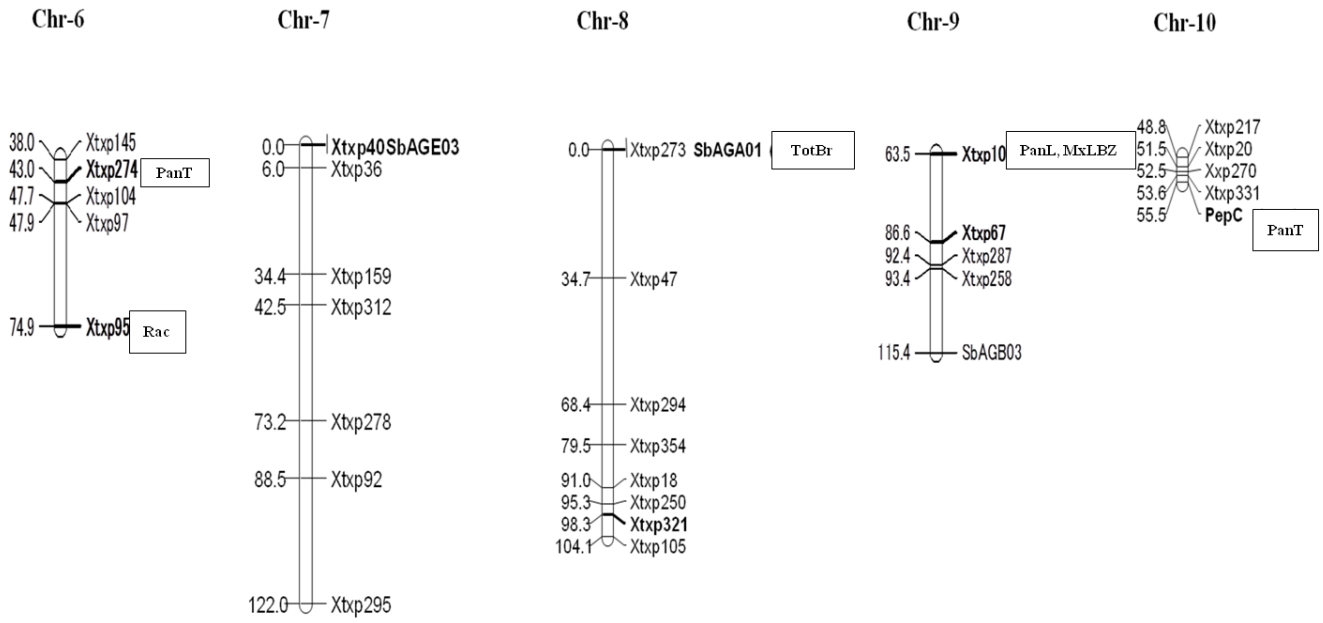


Figure 3.9 (cont). QTLs for sorghum inflorescence traits and location of SSRs on ten sorghum chromosomes are shown in order as describe in (Bhattramakki *et al.*, 2000; Kong *et al.,* 2000; Taramino *et al.,* 1997) and markers with bold face shows significant association with traits as resolved by Q + K model.

Legends for QTLs: 15 QTLs identified for 10 panicle traits and yield related traits in both years are enclosed within rectangle. *Legend for traits:* TotBr: total branch per panicle, Rac: rachis length, PanL: panicle length, TotN: total node along rachis, PanD: panicle diameter, PanW: panicle width, MaxLBZ: maximum length of primary branch, GWP: total grain weight per panicle, PanS: panicle shape, PanT: panicle type.

Chapter 4

General discussion

Sorghum (Family: Poaceae) is an important monocotyledonous food crop with a remarkable diversity in morphological, physiological, genetic and ecological traits. The size and shape of inflorescence organs generally show continuous variation in many plant species and are quantitative traits (Shore and Barrett 1990). Elucidation of the quantitative traits of the genetic base underlying the architecture in inflorescence architecture might allows us to understand how the diverse variation in inflorescence morphology is genetically controlled. Various genomic tools for sorghum are becoming available which will help research efforts on the improvement of this crop. Genetic maps based on molecular markers have several advantages over classical maps (Subudhi and Nguyen, 2000). The genome-wide association approach is one of the important strategies for LD mapping because the power of association studies depends on the degree of LD between the functional genetic polymorphisms and genotyped markers across all chromosomes. In sorghum, several linkage maps have been developed (Subudhi and Nguyen, 2000). More than 40,000 accessions of sorghum germplasm have been used to generate population for mapping of important trait loci. Pereira *et al.,* (1994) developed a sorghum linkage map with 10 complete linkage groups using maize and sorghum probes. Subudhi and Nguyen (2000) aligned the 10 linkage groups of sorghum using information generated from an RIL population with sorghum and maize probes. One of the most complete sorghum genetic maps was published by Menz *et al.,* (2004), who constructed a 1713 cM high-density map using 2454 AFLPs, 203 cDNAs and genomic clones from various grass species such as rice, barley, oat, and maize and 136 SSRs previously mapped in sorghum. In genome wide association studies, several agronomic traits have been studied in cereal crop germplasm (Bouchet, 2012, Brown *et al.,* 2006 and 2008, Casa

et al., 2005, 2008, Shehzad *et al.,* 2009b) but compared with that of maize and rice, successful genome wide association of traits affecting sorghum inflorescence architecture have not been studied extensively. The critical point of determining the genetic basis of sorghum inflorescence architecture has been one of the major scientific challenges to the process of sorghum crop improvement. The study detailed in this thesis attempts to clarify the link between inflorescence architecture and yield potential, and their morphological relationship at the intraspecific levels using 206 sorghum worldwide germplasm. The germplasm were obtained from germplasm collections in the National Institute of Aerobiological Science, Genebank, Japan. Secondly, QTLs underlying and gene influencing the intraspecific variation in the sorghum diversity research set (SDRS) of 107 landraces from worldwide sorghum germplasm were identified by using the genome wide association technique. Results of the present study are discussed under the following sub-headings:

(1) The key components of variation in sorghum inflorescence architecture,

(2) Influence of panicle characters on yield components

(3) Diversity of inflorescence architecture of world-wide sorghum germplasm,

(4) Identification of QTLs controlling inflorescence architecture in sorghum.

4.1. Key components of variation in sorghum inflorescence architecture

In this study, we found that the variation in the inflorescence architecture of sorghum accessions was not only dependent on the panicle length, but also on the total number of branches, the maximum length of primary branches, rachis length, panicle diameter and panicle width. We observed that the variation in sorghum inflorescence is not only the result of differences in panicle elongation and branching characters but also of difference in size and

diameters. Comparisons between loose and compact inflorescence architecture shown different trait associations. This is because different architectures based on different panicle assimilation. The nature of different panicle morphology among different races may be different. However, to understand better the relations between the compact type and open type components, further research has to be performed in other mapping populations and with larger more diverse collections that show considerable variation for both open type and compact type inflorescence architecture components. Among 12 quantitative traits many variations were observed for PanL, Rac, TotN, Rac, PenD, PanW and GNP as unique characteristics of the panicle structure of this sorghum population. The strong relationship between panicle length and other inflorescence characters suggested that these traits could be used as a performance indicator for other characters of the inflorescence architecture. Moreover, among the panicle trait combination, the emphasis of trait selection is still lacking in particular the elongation trait such as TotN and branching trait such as MaxLBZ. Thus, there are major panicle determinants that strongly associated with grain yield which should be considered in breeding programs. These results will serve as a starting point for further evaluation of sorghum germplasm via quantitative trait loci analysis and may be useful for improving yield, based on careful consideration of trait selection and inflorescence morphology.

4.2. Influence of panicle characters on yield components

The yield in cereal is determined by yield components, panicle number, grain number and grain weight. This study aimed to understand the diversity of sorghum germplasm and to help choose informative plant materials based on inflorescence architecture for further study. We detected significant correlations between different components of inflorescence architecture, and

some may be useful for selecting lines within this type of germplasm to improve yield capacity because we can predict seed yield from the complex structure of their sorghum inflorescence types. In terms of the relationship between yield and panicle architecture, we found that the total number of branch (TotBr) and panicle diameter (PanD) have strong impacts on grain yield and are responsible for separation of panicle types. These traits seem to be a good estimate of seed numbers and seed weights in sorghum. However, in general, there is a lot less variation in the open type inflorescence architecture compared to compact type inflorescence architecture with yield components traits. In this study, the result revealed that the selection for panicle length, rachis length, total branch number, panicle diameter and panicle width may improve grain yield. Results showed that the yield component traits are characteristics that would be convenient to measure and inflorescence architecture could be used easily as a selection method for reproductive characters in breeding programs.

4.3. Diversity of inflorescence architecture of world-wide sorghum germplasm

The panicles and grains of the *Sorghum* species vary widely in shape and size and represent a means for racial classification. Sorghum diversification into the five major races and thousands of different genotypes was the result of movement of peoples carrying the species throughout the continent. Africa especially the Ethiopian region is the center of origin of sorghum (Mann *et al.,* 1983) contains many snowdenian species and also several varieties of the durra type. Between the wild and cultivated species, human selection for cultivated characters (mainly non-shattering heads, large seeds and panicle, easy thresh ability, and suitable height and maturity) and natural selection for domesticated and non-domesticated character resulted in

divergence of sorghum populations. Cultivated races are also found in different regions of Africa according to their biological traits and also their histories of distribution. The complex species S. *bicolor* included all cultivated sorghum as well as semi-wild plants mostly associated with them as weeds. Bicolor is not only widely distributed in Africa but was apparently cultivated in Asia (de Wet and Price 1976). We examined panicle diversity using 206 accessions chosen from a world-wide germplasm collection that covers most of the diverse geographic origins in particular the range of variation panicle types of sorghum. Center of origin and early domestication in sorghum led to higher diversity in Africa and Western Asia. The diversity in shape and compactness is likely to indicate selection for varieties to survive in different local environments and is largely independent of geographic distribution from Asia and Africa. We can observe that sorghum panicle types distributed throughout the African and Asian regions has varied according to their adaptation to the difference in environments; temperature, humidity and rainfall patterns. Moreover we can investigate the variability in inflorescence architecture in a wide range of sorghum populations. African accessions tend to have panicle types with good morphological characters, selective traits to prevent adverse environmental effect. The most selected traits of the panicle from these origins were Rac, Pend and TotN and thus we need to know more genetic information for these traits. Cluster and scatter plot analyses identified that the pattern of distribution of the inflorescence (panicle) traits reflected the distribution of different origins and it exhibited a great range of phenotypic diversity based on inflorescence architecture. This is necessary not only for the evaluation of the variation 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. The information

generated from this study allows us to select the appropriate plant materials of sorghum among cultivars for further breeding programs.

4.4. Identification of QTLs controlling inflorescence architecture in sorghum

Currently, genomic resources are becoming available for sorghum breeding across the world. Molecular techniques can be used to analyze genetic distance; the linkage between genes, sequences and populations. Relatedness, identity, geneflow, linkage disequilibria can also be measured by molecular techniques. Recently a number of studies on genomic architecture of sorghum have been undertaken (Hamblin *et al.,* 2004, 2005 and 2007). Detection of loci involved in variation of agronomic traits is an important issue leading to marker assisted selection. In this study, we tried to achieve the possible use of sorghum diversity research sets as core collection in genome wide association mapping analysis. The core collection developed in our study has diverse collection of landraces selected from all parts of Africa and Asia without any improved variety. For the purpose we performed association mapping of sorghum core collection to identify QTLs responsible for sorghum inflorescence architecture. The knowledge of the genetic basic of sorghum inflorescence architecture and its component traits can enhance the process of genetic improvement in sorghum breeding. This study was undertaken to better understand the genetic basis of sorghum inflorescence architecture and its association with yield-related traits. A significant difference between these accessions was observed among 14 measured traits. In the several components of sorghum inflorescence architecture, we found that this trait variation is not only dependent on the panicle length but also on the total branch number, maximum length of primary branch, rachis length, panicle diameter and width. The genome-wide association analysis using 98 SSRs separated the panel into three sub-populations. They

differed in three major groups according to their inflorescence/panicle types (broom, open and compact) and detected 107 sorghum genotypes. Using different models of association analysis, 44 loci on 10 chromosomes were found as significantly related to the patterns of panicle traits in the two growing seasons. 15 loci on 9 chromosomes were found to be significantly related to the patterns of observed panicle traits over two different growing seasons. Among these characters, the mapping of QTLs associated with panicle length (PanL) had been reported in many previous studies. In this study, we analysed PanL and its related traits by separating these different parts. The traits for panicle length (PanL) consisting of the rachis length (Rac) and the top most primary branches on the rachis has not been identified through QTL mapping. Interestingly, our results did not detect whether PanL itself and Rac are controlled by different QTLs and QTLs for the rachis length. Our result indicated that these traits appeared to be under separate genetic control (from length based traits) in sorghum inflorescence architecture because there is no overlap between QTLs that were detected for these traits. Earlier studies on inflorescence architecture in sorghum, rice and maize have suggested that branching characteristics had the most important role in panicle characteristics. The rachis length (Rac) and branches along the rachis, such as MaxLBZ, appeared distinct processes and appeared to be controlled be different QTLs. More than one QTLs for different traits were found in the same region on chromosome 4 (linkage group D). Bortiri *et al*., (2006) reported that the gene *ramosa2* (ra^2) may be critical to the early steps of grass inflorescence architecture. When the authors examined (ra^2) mutants they observed increased branching, in particular short branches were replaced by long indeterminate ones. The gene also appears to be conserved in sorghum (*Sorghum bicolor*), rice (*Oryza sativa*) and barley (*Hordeum vulgare*) and shows a similar expression pattern, suggesting it is likely to play the same role in inflorescence architecture of these other species. Similarly, QTL studies

undertaken by Brown *et al.,* (2006) in sorghum and by Upadyayula *et al*., (2005) in maize show that allelic variation in genes such of the *ramosa* gene is a major determinant of morphological variation in inflorescence branch length within a species. In this study GWAS analysis demonstrated that PanL and MaxLBZ might be under the same genetic control. QTLs for these traits were located in the same position. In relation to yield, Gerik *et al.,* (2004) reported that panicle elongation was a good estimator of seed number in sorghum, while Brown (1980) reported that although numbers of spikelet per spike did not significantly correlate with yield, high seed yield comes from plants with larger heads. Therefore panicle (head) dimension and size has the potential to be used as a predictor for several inflorescence characters and also seed sets. Among inflorescence traits, panicle diameter can be use easily by breeders as a selection method in breeding programmes. The highly significant association of the panicle diameter on Chr-2 (LG-B) has not been previously reported in mapping studies. In addition, QTLs for this trait in this location have not been reported by the previous analysis. The genomic region on Chr-4 for QTLs showed an interaction not only with the panicle traits but also the yield-related traits, such as the total grain weight. Panicle dimensional traits; the panicle diameter and width (PanD, PanW), are important traits for panicle improvement for example, dense panicles, a disorder in panicle dimension and size, leads to low grain quality. It is worth noting that we were unable to identify novel QTLs for PanD and PanW, despite finding existing associations on three chromosomes (Chr-1, Chr-2 and Chr-3). However the QTLs involved on Chr-1 for PanW, Chr-2 for TotN and PanD, Chr- 6 for Rac, Chr-10 for PanT were identified as the new QTLs underlying inflorescence architecture. Most commercial crop verities contain dwarfing genes and it can give high yield through an improvement in the harvest index. The semi dwarfing gene is one of the most important genes deployed in modern crop breeding. In sorghum, the gene *dwarf3*

(*dw3*) of sorghum (Multani *et al.,* 2003) was used for lodging resistance and higher yields in this crop decades before the wheat or rice Green Revolutions were conceptualized (Quinby and Karper 1954). Dwarfing genes have been found to be useful for crop improvement but it is labor-intensive approach, costly and time-consuming. Recently other modifications of plant architecture might also be possible offers new possibilities for improved crop performance in the field (i.e., inflorescence architecture). GWAS on inflorescence architecture in this study identified novel loci associated with the key traits for sorghum inflorescence architecture. Genes underlying QTL influencing panicle and it related traits were benefit to not only the association of yield and inflorescence genetic basis of sorghum core collection accessions but also to understand the variation of different panicle types and yield components traits. This result suggested that the variation in inflorescence architecture in sorghum can be broadened, and there is room for advancement in genetic improvement of yield capacity of sorghum.

4.5. Conclusions

We have achieved our objectives and goals for performing this research. We have identified the patterns of diversity in sorghum worldwide accessions based on their panicle types and inflorescence traits. Moreover, our plant materials ,core collection, are landraces from worldwide germplasm collection and it was further utilized in association mapping of inflorescence traits and several loci controlling qualitative and quantitative traits were identified.

The future aspects of this research are (i) Construction of high linkage map of sorghum by using F2 populations obtained from parents of diverse origins, different panicle types and genetic background i.e., Africa and Asia accessions, compact and open panicle types. (ii) Identification of QTLs responsible for inflorescence architecture can serve as possible gene targets not only for

improving yield production but also for more emphasis combine grain yield with forage and stover yield.

SUMMARY

Throughout the world, grass species (Poaceae family) are economically important as both staple grain foods for humans and feedstock for animals. Sorghum (*Sorghum bicolor* L. Moench) is the fifth most important C_4 cereal crop globally (FAO, 1995, 1999) and can survive the harsh climatic conditions of arid and tropical environments. Unlike rice and other staple crops, which are widely used for food and industrial purposes, sorghum has thus far remained a traditional food crop of subsistence farmers (Rai *et al*., 1999; Singh *et al*., 2001). Cultivated sorghum is classified into five main races (*Bicolor, Guinea, Caudatum, Durra and Kafir*) (Barnaud *et al*., 2008a; Harlan *et al*., 1976). Sorghum types can be identified according to their morphological traits (Kaitaniemi *et al*., 1999). The identification of racial divisions and species values are primarily based on panicle and grain characters (Abdi *et al*., 2002; Harlan *et al*., 1976; Murray *et al*., 2009). Sorghum panicles, which are called inflorescences, show remarkable diversity in morphological, physiological, genetic and ecological traits. The agronomic performance of cereal crops is significantly influenced by the complexity of inflorescence/panicle patterns. The inflorescence architecture is an important agronomic factor and major determinant which contributes to both the yield and quality of sorghum. It is also an important determinant of rice and maize tassels because of its close associations with grain yield and grain quality, which include several commercially important traits (Bommert *et al*., 2005; Ikeda *et al*., 2010; Yan *et al*., 2007). Breeders have greatly improved plant architecture, potential energy and grain yield productivity. Panicle morphology can directly affect grain yield; therefore, the knowledge of the genetic basis of sorghum inflorescence architecture and its components can complement the breeder's efforts to improve sorghum. Of these characteristics, sorghum inflorescences have become a new model system for functional genomics from agronomic, developmental, and

evolutionary viewpoints similar to other important model cereal crops, such as maize and rice; however, a few morphological characteristics of inflorescence architecture have been mapped similar to major effect genes across a range of genetic linkage maps (Colasanti *et al*. 1998; Harlan *et al*. 1972; Ikeda *et al*. 2005). We expected that the systematic genomic analyses of sorghum panicle traits could lead to improved breeding programs and yields. Therefore we set out to measure the variation of a comprehensive set of sorghum inflorescence architecture traits based on a large collection of 206 geographically diverse sorghum accessions and the sorghum diversity research set (SDRS) of 107 landraces from worldwide sorghum germplasm.

This study first clarified the link between geographic origin and the variation of inflorescence architecture to compile useful information on the origins of the accessions from the sorghum world-wide germplasm collection. We found that the patterns of observed panicle traits only partially reflected the distribution of different origins. The diversity in shape and compactness are likely to indicate selection of varieties that can survive in different local environments. Secondly, the distribution of several components of sorghum inflorescence architecture influenced yield components. In several components of sorghum inflorescence architecture, we found that the significance and high correlations between pairs of traits revealed that these variations are dependent not only on the panicle length but also on the total branch number, maximum length of the primary branch, rachis length, and panicle diameter and width. These are major panicle determinants which are strongly associated with grain yield which should be considered in breeding programs to emphasize yield improvement. Moreover, among the panicle trait combination the emphasizing of trait selection is still lacking especially the elongation trait (TotN), dimensional trait (Pan D) and branching trait (MaxLBZ). These results

can be used as preliminary findings for QTL studies to find genetic markers for panicle traits in sorghum with the aim to improve yield of this crop.

Next, we investigated 98 simple sequence repeat (SSRs) maps in the (SDRS) of 107 landraces from the worldwide sorghum germplasm. A significant difference between accessions was observed for 14 measured traits. Molecular markers divided the germplasm into three sub-populations with different groups of inflorescence/panicle types (broom, open and compact). The SDRS was composed of six different panicle types that were associated with 14 inflorescence traits. Using different models of association analysis, 15 loci on 9 chromosomes were found to be significantly related to the patterns of observed panicle traits. Among these loci, five QTLs were responsible for length-based traits, and three QTLs were responsible for dimensional traits. Five QTLs were responsible for panicle features such as panicle type and shape, one QTL was responsible for yield related traits such as grain weight. Two QTLs were responsible for branch-based traits, such as the total node number and branch number with $-\text{Log}_{10}$ (P) values ranging from 1.3 to 7.6 as the threshold value. Our result revealed that several genomic regions affected multiple traits, including one region that affected PanL and MxLBZ (MaxLBZ). QTLs for different traits tended to be found in the same region on chromosome 4 and on chromosome 9. Additionally, QTLs on Chr-2, Chr-5, Chr-6 and Chr-10 (a novel QTL) were identified that control rachis length, total node number, panicle diameter and panicle type. Sorghum has been much less studied in the grass species, for genes affecting inflorescence architecture than other cereal crops. In this study the results of QTLs for Rac, PanD, PanW, TotN, MaxLBZ, PanS and PanT did not relate to any of the QTL for panicle branching pattern identified by previous reports. Therefore, it is likely that these QTLs are novel loci regulating the panicle and its component traits resulting in their involvement in sorghum inflorescence architecture. These results will

serve as a foundation for QTL work into sorghum inflorescence architecture, and a further evaluation of the germplasm for these traits is in progress. Moreover, this study helps pave the way for more detailed studies with increased population sizes and saturation of the target genomic regions by adding large scale molecular markers with powerful molecular techniques (*i.e.* next generation sequencing, genome-wide SNP discovery, whole genome re-sequencing and map-based cloning of the gene underlying the QTLs). In conclusion, these findings can lead to the emergence of a new era of sorghum genomics, and help bridge the knowledge gap between genotype and phenotype in sorghum inflorescence architecture.

ACKNOWLEDGMENTS

I would like to thank all the people who contributed in some way to the work described in this thesis. First and foremost, I would like to express my deepest sense of gratitude and appreciation to my senior advisor and supervisor, Professor Dr. Kazutoshi Okuno for accepting me into his group and thanks for all I have learned from him and for his continuous help and support in all stages of this thesis. I would also like to thanks him for being an open person to ideas, and for encouraging and helping me to shape my interest and ideas. His guidance helped me in all the time of research and writing of research manuscripts.

After that I am deeply thankful to my supervisor, Professor Dr. Ryo Ohsawa, Laboratory of Plant Genetics and Breeding Science, Biosphere Resource Science and Technology, Tsukuba University, for his invaluable advice, patient, motivation, enthusiasm and immense knowledge and inspiring guidance throughout the course of study.

My cordial thanks are forwarded to two members of supervisory committee, Prof. Dr. Sachio Maruyama (University of Tsukuba), Dr. Takeshi Hayashi (National Agricultural Researcher Center) and Dr. Tariq Shehzad (University of Tsukuba) for their valuable comments and suggestions in improving and uplifting the standard of my thesis.

I would like to acknowledge and convey my sincere thanks to Dr**.** Alistair Forrest**,** RIKEN, and Center for Life Science Technology for his great support, kindness and deep affection to me. I am also grateful to Professor Dr. Pi-Chao Wang, Cell Tissue Engineering Laboratory, Gene Research Center, Tsukuba University for her kind support.

I would like to express my truthful thanks to all my Senior Lab members Dr. Takashi Hara and Dr.Yasuko Yoshida for their guidance, support whenever I am in doubts and deeply

appreciate their valuable assistance in proofreading of this study. I sincerely appreciate all their help without this study would have been difficult to carry out.

I would like to acknowledge all my lab members, Laboratory of Plant Genetics & Breeding Science and my friends Foo Chuan Eng, Rita Andini and Ahmad Kadapi for their support and encouragement during this study.

I am highly indebted to my parents, especially my mother, Professor Dr. Khin San Yi, Botany Department and Microbiology Department, Yadanabon University, Myanmar, for her great encouragement and patience throughout the long term study in Japan.

Last but not least, I would like to acknowledge the Japanese Government (Monbukagakushore -MEXT: Ministry of Education, Culture, Sports, Science and Techonology, Government of Japan) for giving me this golden opportunity to study in Tsukuba University in Japan, a very prestigious experience. In addition, this study has been accomplished without the support, interest and help of many persons.

KHAING PANN WITT HMON

REFERENCES

Abdi A., Bekele E., Asfaw Z. and 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.

Ayana A. and Bekele E. (1989) Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor (*L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution* 46: 273–284.

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

Ayyangar G.N.R. and Ayyar M.A.S. (1938) Linkage between a panicle factor and the pearly chlalkly mesocarp factors in sorghum. *Proceedings of the Indian Academy of Science* 8: 100–107.

Bala R.S., Biswas P.K. and Ratnavathi C.V. (1996) Advances in value addition of Kharif sorghum. *Crop Improvement* 23:169–177.

Balole T.V. and Legwaila G.M. (2006) *Sorghum bicolor* (L.) Moench. [Internet] Record from PROTA4U. Brink, M. & Belay, G. (ed.) PROTA (Plant Resources of Tropical Africa)que tropicale), Wageningen, Netherlands. <http://www.prota4u.org/search.asp

Barnaud A., Monique D., Eric G., Jacques C., Justin B., Esaei O.K., Doyle M.K. and Helene I.J. (2008a) Weed–crop complex in sorghum: The dynamics of genetic diversity in a traditional farming system. *American Journal of Botany* 96:1869–1879.

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

Bello D., Kadams A.M. and Simon S.Y. (2001) Correlation and path coefficient analysis of grain yield and its components in sorghum (*Sorghum bicolor* (L.) Moench). *Nigerian Journal of Tropical Agriculture* 3: 4–9.

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

Bommert P., Satoh N.N., Jackson D. and Hirano H.Y. (2005) Genetics and evolution of inflorescence and flower development in Grasses. *Plant and Cell Physiology* 46: 69–78.

Bortiri E., Chuck G., Vollbrecht E., Rocheford T., Martienssen R. and Hake S. (2006) Ramosa2 encodes a lateral organ boundary domain protein that determines the fate of stem cells in branch meristems of maize. *Plant Cell* 18:574–585.

Bouchet S. (2012) Genetic structure, linkage disequilibrium and signature of selection in Sorghum: Lessons from physically anchored DArT markers. *PLoS ONE* 7: 33470.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y. and Buckler E.S. (2007) TASSEL Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635.

Brown K.R. (1980) Recent grass seed production studies in Canterbury. In: Lancashire J.A. (ed.) Herbage seed production. *Research and Practice Series no. 1*, Pamerston North, New Zealand: New Zealand Grassland Association, pp. 12–14.

Brown P.J., Klein P.E., Bortiri E., Acharya C.B., Rooney W.L. and Kresovich S. (2006) Inheritance of inflorescence architecture in sorghum. *Theoretical and Applied Genetics* 10:931–942.

Brown P.J., Rooney W.L., Franks C. and Kresovich S. (2008) Efficient mapping of plant height quantitative trait loci in sorghum association population with introgressed dwarfing genes. *Genetics* 180: 629–637.

Brown P.J., Myles S. and Kresovich S. (2011) Genetic support for phenotype-based racial classification in sorghum. *Crop Science* 51:224–230.

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

Casa A.M., Mitchell S.E., Hamblin M.T., Sun H., Bowers P.J., Paterson A.H., Aquadro C.F. and Krosovich S. (2008) Community resources and strategies for association mapping in sorghum. *Crop Science* 48:30–40.

Colasanti J., Yuan Z. and Sundaresan V. (1998) The indeterminate gene encodes a zinc finger protein and regulates a leaf-generated signal required for the transition to flowering in maize. *Cell* 93:593–603.

Crasta O.R., Xu W.W., Rosenow D.T., Mullet J. and Nguyen H.T. (1999) Mapping of post-flowering drought resistance traits in grain sorghum: Association between QTLs influencing premature senescence and maturity. *Molecular and General Genetics* 262:579–588.

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

de Wet J.M.C. (1978) Systematic and evaluation of sorghum sect. sorghum (Gramineae) *American Journal of Botany* 65: 477–484.

Doggett H. (1988) Sorghum. 2nd edition, Longman Group UK Ltd., Essex, Scientific and technical, London: Wiley, New York, pp. 75–97.

Doust A.N. and Kellogg E.A. (2002) Inflorescence diversification in the panicoid "bristle grass" clade (Paniceae, Poaceae): evidence from molecular phylogenies and developmental morphology. *American Journal of Botany* 89**:** 1203–1222.

Doust A.N., Devos K.M., Gadberry M.D., Gale M.D. and Kellogg E.A. (2004) Genetic control of branching in foxtail millet. *Proceedings of the National Academy of Sciences of the United States of America* 101:9045–9050.

Doust A.N., Devos K.M., Gadberry M.D., Gale M.D. and Kellogg E.A. (2005) The Genetic basis for inflorescences variation between foxtail and green millet (Poaceae). *Genetics* 169: 1659–1672.

Elhan S.E., Jianming Y. and Edward S.B. (2009) Applications of linkage disequilibrium and association mapping in maize. In: Kriz A.L., Larkins B.A. (eds.) *Molecular Genetic Approaches to Maize Improvement: Biotechnology in Agriculture and Forestry*, Vol.63, Springer-Verlag Berlin Heidelberg, pp. 175–176.

FAO (1995) Sorghum, In: Lupien J.R. (eds.) Sorghum and millets in human nutrition, Food and Agriculture Organization of the United Nations, Statistics Division, pp.3–12.

FAO (1999) Sorghum; Post harvest operation, In: Mejia D. Food Security Department (eds.) Food and Agriculture Organization of the United Nations, Statistics Division.

Fanous M.A., Weibel D.E. and Morrison R.D. (1971) Quantitative inheritance of some head and seed characteristics in sorghum (*Sorghum bicolor* (L.) Moench). *Crop Science* 11:787–789.

Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Mni M., Nezer C., Simln P., Vanmanshoven P., Wagenaar D. and Georges M. (2000) Extensive genome-wide linkage disequilibrium in cattle. *Genome Research* 10:220–227.

Flint-Garcia S.A., Anne-Ce´ L.T., Yu J., Gael P., Susan M. R., Sharon E.M., Doebley J., Kresovich S. and Buckler E.S. (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant Journal* 44:1054–1064.

Futsuhara Y., Kondo S. and Kitano H. (1979a) Genetically studies on dense and lax panicles in rice. I. Character expression and mode of inheritance of lax panicle rice. *Japanese Journal of Breeding* 29: 151–158.

Futsuhara Y., Kondo S. and Kitano H. (1979b) Character expression and mode of inheritance of dense panicle rice. *Japanese Journal of Breeding* 29: 239–247.

Morrisa G. P., Ramub P., Deshpandeb S.P, Hashc C.T, Shahb T., Upadhyayab H.D., Riera-Lizarazub O., Brown P.J., Acharyae C.B., Mitchelle S.E., Harrimane J., Glaubitze J.C., Bucklere E.S. and Kresovicha S. (2012) Population genomic and genome-wide association

studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences of the United States of America.* 110:453–8.

Gerik T.J., Rosenthal W.D., Vanderlip R.L.. and Wade L.J. (2004) Simulating seed number in grain sorghum from increases in plant dry weight. *Agronomy Journal*, 96:1222–1230.

Ghawchawe B.G., Bhale N.L. and Shekar V.P. (1996) Linkage group in sorghum. *Indian Journal of Genetic*. 26:317–326.

Hamblin M.T, Mitchell S.E., White G.M., Gallego J., Kukatla R., Wing R.A., Paterson A.H. and Kresovich S. (2004) Comparative population genetics of the Panicoid grasses: Sequence polymorphism, Linkage disequilibrium and selection in a diverse sample. *Genetics* 167:471–483.

Hamblin M.T., Salas F.M.G., Casa A.M., Mitchell S.E., Paterson A.H. and Kresovich S. (2005) Equilibrium processes cannot explain high levels of short- and medium-range linkage disequilibrium in the domesticated grass *Sorghum bicolor*. *Genetics* 171: 1247–1256.

Hamblin M.T., Salas F.M.G., Tuinstra M.R., Rooney W.L. and Kresovich S. (2007) Sequence variation at candidate loci in the starch metabolism pathway in sorghum: Prospects for linkage disequilibrium mapping. *Crop Science* 47:S125–S134.

Hardy O.J. and Vekemans X. (2002) SPAGeDi: A versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2:618–620.

Harlan J.R. and de Wet J.M.J. (1972) A simplified classification of cultivated sorghum. *Crop Science* 12:172–176.

Harlan J.R., de Wet J.M.J. and Stemler A.B.L., (1976) Plant domestication and indigenous African Agriculture. In: Harlan J.R., de Wet J.M.J. and Stemler A.B.L. (eds.) *Originas of African Plant Domestication*, Mouton Press, The Hague, pp. 3–19.

House L.R. (1985) A guide to sorghum breeding. Manual (online book). International Crops Research Institute for the Semi-Arid Tropics. Second Edition, India, pp.27–49.

Hu F.Y., Tao D.Y., Sacks E., Fu B.Y., Xu P., Li J., Yang Y., McNally K., Khush G.S. and Paterson A.H. (2003) Convergent evolution of perenniality in rice and sorghum. *Proceedings of the National Academy of Sciences of the United States of America*. 100:4050–4054.

IBPGR and ICRISAT (1993) Descriptors for sorghum [*Sorghum bicolor* (L.)Moench]. IBPGR, ICRISAT, Rome, Patancheru, pp. 12–29.

Ikeda M., Hirose Y., Takashi T., Shibata Y., Yamamura T., Komura T., Doi K., Ashikari M., Matsuoka M. and Kitano H. (2010) Analysis of rice panicle traits and detection of QTLs using an image analyzing method. *Breeding Science* 60: 55–64.

Kaitaniemi P., Room P.M. and Hanan J.S. (1999) Architecture and morphogenesis of grain sorghum, *Sorghum bicolor* (L.) Moench. *Field Crop Research* 61: 51-60(10).

Kebede H., Subudhi P.K., Rosenow D.T. and Nguyen H.T. (2001) Quantitative trait loci influencing drought tolerance in grain sorghum (*Sorghum bicolor* (L.) Moench). *Theoretical and Applied Genetics* 103:266–276.

Kellogg E.A. (2000) Molecular and morphological evolution in Andropogoneae. *Grass: Systematics and evolution.* In: Everett J.E. (ed.) CSIRO, Melbourne. Australia, pp. 148–229.

Kimber C.T. (2000) Origins of domesticated sorghum and its early diffusion to India and China. In: Smith *et al.*, (ed.) Sorghum: Origin, history, technology, and production. John Wiley & Sons, Inc., New York, pp. 3–98.

Kimber C.T., Dahlberg J.A. and Kresovich S. (2013) Genomics of the Saccharinae: The gene pool of *Sorghum bicolor* and its improvement, In: Paterson A.H. (ed.), *Plant Genetics and Genomics*, Crops and Models 11, pp. 2– 23.

Kirby J.S. and Atkins R.E. (1968) Heterotic response for vegetative and mature plant characters in grain sorghum, *Sorghum bicolor (*L). Moench. *Crop Science* 24: 649–654.

Klein R.R., Rodriguez-Herrera R., Schulueter J.A., Klein P.E., Yu Z.H. and Rooney W.L. (2001) Identification of genomic regions that affect grain-mouldincidence and other traits of agronomic performance in sorghum. *Theoretical and Applied Genetics* 102:307–319.

Kock N. (2012) *WarpPLS 3.0 User Manual***.** Laredo, Texas: ScriptWarp Systems.

Kong L., Dong J. and 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.

Kukadia M.U., Desai K.B., Desai M.S., Patel R.H. and Rajak R.V. (1983) Estimates of heritability and other related genetic parameters in sorghum. *Sorghum Newsletter* 26: 31–32.

Kumar S. and Singhania D.L. (1984) Genetic advance and heritability estimates for grain yield and its components. *Sorghum Newsletter* 27:15–16.

Loiselle B.A., Sork V.L., Nason J. and Graham C. (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany 82:1420–1425.

Maman N., Mason S.C., Lyon D.J. and Dhungana P. (2004) Yield components of pearl millet and grain sorghum across environments in the central great plains. *Crop Science* 44:2138–2148.

Mann J.A., Kimber C.T. and Miller F.R. (1983) The origin and early cultivation of sorghums in Africa. *Texas Agricultural Experiment Station Bulletin*: 1454.

Martin J.H. and Leonard W.H. (1949) Crops of the grass family. Sorghum. In: Richard P. (eds.) *Principles of Field Crop Production,* McMillan, New York. USA, pp.348–379.

Maunder A.B. (1999) History of cultivar development in the United States: From "memoirs of A.B. Maunder-sorghum breeder". In*:* Smith *et al*., (eds.) Sorghum: Origin, history, technology, and production. John Wiley & Sons, Inc., New York, pp. 191–223.

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

Multani D.S., Briggs S.P., Chamberlin M.A., Blakeslee J.J. and Murphy A.S. (2003) Loss of an MDR transporter in compact stalks of maize *br*2 and sorghum *dw*3 mutants. *Science* 302: 81–84.

Murry M.G. and Thompson W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8: 4321–4326.

Murray S.C., Rooney W.L., Hamblin M.T., Mitchell S.E. and Kresovich S. (2009) Sweet sorghum genetic diversity and association mapping for brix and height. *The Plant Genome* 2: 48–62.

Patel R.H., Desai K.B. and Desai S.P. (1983) Heritability estimates in grain sorghum. *Sorghum Newsletter* 26:89–90.

Paterson A.H., Lander E.S., Hewitt J.D., Peterson S., Lincoln S.E. and Tanksley S.D. (1998) Resolution of quantitative traits into mendelian factors using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721–736.

Paterson A.H., Bowers J.E., Bruggmann R., Dubchak I., Grimwood J., Gundlach H., Haberer G., Hellsten U., Mitros T., Poliakov A., Schmutz J., Spannagl M., Tang H., Wang X., Wicker T., Bharti A.K., Chapman J., Feltus F.A., Gowik U., Grigoriev I.V., Lyons E., Maher C.A., Martis M., Narechania A., Otillar R.P., Penning B.W., Salamov A.A., Wang Y., Zhang L., Carpita N.C., Freeling M., Gingle A.R., Hash C.T., Keller B., Klein P., Kresovich S., McCann M.C., Ming R., Peterson D.G., Mehboob-ur-Rahman, W.D., Westhoff P., Mayer K.F., Messing J. and Rokhsar D.S. (2009) The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457:551–6.

Pereira M.G., Lee M., Bramel-Cox P., Woodman W., Doebley J. and Whitkus R. (1994) Construction of an RFLP map using sorghum and maize DNA probes. *Genome* 37:590–594.

Pereira M.G. and Lee M. (1995) Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical and Applied Genetics* 90:380–388.

Pritchard J.K., Stephens M., Rosenberg N.A. and Donnelly P. (2000) Association mapping in structured populations. *The American Society of Human Genetics* 67:170–181.

Quinby J.R. and Karper R.E. (1954) Inheritance of height in Sorghum. *Agronomy Journal* 46:211–216.

SAS Institute Inc (2010) JMP statistical and graphics guide, version 9. SAS Institute Inc, Cary.

Saeed M. and Francis C.A. (1983) Yield stability in relation to maturity in grain sorghum. *Crop Science* 23: 683–687.

Shehzad T., Okuizumi H., Kawase M. and Okuno K. (2009a) 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.

Shehzad T., Iwata H. and Okuno K. (2009b) Genome-wide association mapping of quantitative traits in sorghum (*Sorghum bicolor* (L.) Moench) by using multiple models. *Breeding Science* 59: 217–227.

Shore J.S. and Barrett S.C.H. (1990) Quantitative genetics of floral characters in homostylous *Turnera ulmifolia* var. *anylstifolia* Wild. (Turneraceae). *Heredity* 64: 105–112.

Singh V., Oosterom E.J.V., Jordan D.R., Hunt C.H. and Hammer G.L. (2001) Genetic variability and control of nodal root angle in sorghum. *Crop Science* 51: 2011–2020.

Snowden J.D. (1936) The cultivated races of Sorghum. London: Allard and Son, pp. 274–279.

Srinivas G., Satish K., Madhusudhana R., Nagaraja R.R., Murali M.S. and Seetharama N. (2009) Identifcation of quantitative trait loci for agronomically important traits and their

association with genic-microsatellite markers in sorghum. *Theoretical and Applied Genetics* 118:1439–1454.

Stemler A.B.L., Harlan J.R. and de wet J.M.J. (1977). Sorghums of Ethiopia. *Economic Botany* 31: 446–460.

Stich B., Alberecht E., Frisch M.M., Maurer H.P., Heckenberger M. and Reif J.C. (2005) Linkage disequilibrium in Europen elite maize germplasm investigated with SSRs. *Theoretical and Applied Genetics* 111:723–730.

Stich B., Mohring J.M., Piepho H.P., Heckenberger M., Buckler E.S. and Melchinger A.E. (2008) Comparison of mixed-model approaches for association mapping. *Genetics* 178:1745–1754.

Subudhi P.K. and Nguyen H.T. (2000) New horizons in biotechnology. In: Smith et al. (ed.) Sorghum: Origin, history, technology, and production. John Wiley & Sons, Inc., New York, pp. 349–397.

Tao Y., Manners J.M., Ludlow M.M. and Henzell R.G. (1993) DNA polymorphisms in grain sorghum (*Sorghum bicolor* (L.) Moench). *Theoretical and Applied Genetics* 86:679–688.

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

Taylor J.R.N. and Dewar J. (2001) Developments in sorghum food technologies. *Advance in Food and Nutrition Research* 43: 217–264.

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

Tesso T., Tirfessa A. and Mohammed H. (2011) Association between morphological traits and yield components in the durra sorghums of Ethiopia. *Hereditas* 148: 98–109.

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

Upadyayula N., da Silva H.S., Bohn M.O. and Rocheford T.R. (2005) Genetic and QTL analysis of maize tassel and ear inflorescence architecture. *Theoretical and Applied Genetics* 112: 592–606.

Vermerris W., Saballos A., Ejeta G., Mosier N.S., Ladisch M.R. and Carpita N.C. (2007) Molecular breeding to enhance ethanol production from corn and sorghum stover. *Crop Science* 47: 142–S153.

Vollbrecht E., Patricia S.S., Lindee G., Edward S., Buckler I.V. and Robert M. (2005) Architecture of floral branch systems in maize and related grasses. *Nature* 436: 1119–1126.

Weigel D. and Meyerowitz E.M. (1994) The ABCs of floral homeotic genes. *Cell* 78:203–209.

Wenzel W.G. (1990) Inheritance and interrelationship of quantitative traits in sorghum. *Sorghum Newsletter* 31:1–3.

Witt Hmon K.P., Shehzad T. and Okuno K. (2013) Variation in inflorescence architecture associated with yield components in a sorghum germplasm. *Plant Genetic Resources: Characterization and Utilization* 11: 1–8.

Xu Y. (2002) Global view of QTL: Rice as a model. In: Kang M.S. (ed.) Quantitative genetics, genomics and plant breeding. CAB International, Cambridge, pp. 109–134.

Yan C.J., Zhou J.H.S., Yan F.C. and Yeboah M. (2007) Identification and characterization of a major QTL responsible for erect panicle trait in japonica rice **(***Oryza sativa* L.). *Theoretical and Applied Genetics* 115: 1093–1100.

Yu J.G., Pressoir W.H., Briggs I., Vroh B.M., Yamasaki J.F., Doebley M.D., McMullen B.S., Gaut D.M., Nielsen J.B., Holland S. and Kresovich E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetic* 38:203–208.

Zhao K., Aranzana M.J., Kim S., Lister C., Shindo C., Tang C., Toomajian C., Zheng H., Dean C., Marjoram P. and Nordborg M. (2007) An Arabidopsis example of association mapping in structured samples. *PLoS Genetic* 3:1–e4.

Zhu K., Tang D., Yan C., Chi Z., Yu H., Chen J., Liang I., Gu M. and Cheng Z. (2010) ERECT PANICLE2 encodes novel protein that regulates panicle erectness in Indica rice. *Genetics* 184: 343–350.

Zou G., Yan S., Zhai G., Zhang Z., Zou J. and Tao Y. (2011) Genetic variability and correlation of stalk yield-related traits and sugar concentration of stalk juice in a sweet sorghum

(*Sorghum bicolor* (L.) Moench) population. *Australian Journal of Crop Science* 5: 1232–1238.