The Pennsylvania State University

The Graduate School

Department of Plant Sciences

GENETIC DIVERSITY AND GENOME WIDE MAPPING OF STRESS INDUCED SECONDARY METABOLITES IN SORGHUM (*Sorghum bicolor* (L.) Moench)

A Dissertation in

Agronomy

by

Dinakaran Elango

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2018

The dissertation of Dinakaran Elango was reviewed and approved* by the following:

Surinder Chopra
Professor of Maize Genetics
Dissertation Co-Advisor

Gregory W. Roth
Professor of Agronomy
Dissertation Co-Advisor
Head of the Department of Plant Sciences

Yinong Yang
Professor of Plant Pathology

Lavanya Reddivari
Assistant Professor of Plant Sciences

Majid R. Foolad
Professor of Plant Genetics

*Signatures are on file in the Graduate School

ABSTRACT

Sorghum is the fifth most important cereal crop and an excellent model system to study the role of secondary metabolites in response to biotic and abiotic stresses. Sorghum produces unique and rare flavonoid phytoalexins in the poaceace family. Phytoalexins are low molecular weight, stress inducible, secondary metabolites with activity against multiple biotic and abiotic stressors. Sorghum phytoalexin class contains 3-deoxyanthocyanidins (3-DAs), which consists of luteolinidin and apigeninidin and their methyl and acyl derivatives. Sorghum 3-DAs play a key role in maintaining plant health by restricting anthracnose fungal pathogen proliferation and also imparting resistance against certain insect pests. The occurrence of these inducible 3-DAs are low in nature and genotypes differ greatly for their ability to synthesize of these compounds in response to biotic and abiotic stresses. So, it is imperative to evaluate the large genetic pool of sorghum accessions for these novel induced phytoalexins and profile them for further genetic improvement. It is known that anthracnose resistant sorghum genotypes synthesize more phytoalexins at a faster rate at the infection site than the susceptible genotypes. We have used two sorghum mapping populations with dense SNPs (Single Nucleotide Polymorphisms) to map genes controlling the accumulation of 3-DAs, epi-cuticular wax (EW), and anti-oxidant activities in sorghum. Mapping populations consist of Sorghum Association Panel (SAP; 377 accessions) and an International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) sorghum mini-core panel (242 accessions).

Colletotrichum sublineola is an aggressive hemi-biotrophic fungal pathogen that causes anthracnose leaf blight and stem rot in sorghum. In severe conditions, it cause substantial yield loss up to 45%. Often major gene resistance fails due to varying degree of virulent pathotypes across sorghum growing locations and availability of other host species for pathogen survival. One mechanism to control anthracnose leaf blight in sorghum is to identify genetic factors associated with anti-fungal production. It is known that *Colletotrichum sublineola* induce the production of anti-fungal compounds known as 3-deoxyanthocyanidins (3-DAs) up on the fungal ingress in sorghum. These compounds are initially colorless and move to the site of infection. After 36 hours, they turn into brick red color and kill the invading fungus and the cell. This mechanism prevents the fungal pathogen spread from one area to other, thus effectively preventing the disease spread. Our goal of this investigation was to identify the genetic factors associated with 3-DA bio-synthesis and regulation in sorghum. A genome wide association study (GWAS) was performed using the sorghum association panel (SAP), ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) mini-core (MNC) panel, and nested association mapping (NAM) panel. Community resources of 265,487 SNP (Single Nucleotide Polymorphism) markers for SAP and MNC panel, and 5108 SNP markers for NAM panel were used in this study. Three biological replicates of SAP and MNC leaf samples were collected at V_7 stage for fungus infection, whereas 21 days seedlings were used for NAM population screening for anthracnose disease. The quantitative measurement of 3-DAs was carried out using spectrophotometer (λ_{max} 480 nm) and confirmation of a sub-set by HPLC (High Performance Liquid Chromatography) with

known 3-DA standards. Significant variability for 3-DAs among the genotypes observed (0.02 (IS 1219) to 2.99 (IS 11619)) with a mean of 0.52 (abs/mg). GWAS identified 11 candidate genes (*p*-value ≤7.73E-06) for 3-DA bio-synthesis and regulation in sorghum which include Sb04g013160 (Guanylate-binding protein), Sb01g044000 (Sugar transporter), Sb02g034830 (BCS1 AAA type ATPase), Sb04g010280 (MYB-like DNA-binding domain), Sb09g004660 (Peroxidase) and others. Our results provide basis for anthracnose disease improvement through marker assisted selection and genomic selection for plant secondary metabolites that act as fungicides in sorghum.

Sorghum exhibits poor tolerance to cold and frost. A number of flavonoid secondary metabolites are induced during the plant–stress cross talk, and they play a major role in imparting stress tolerance to plants. Our objective of this study is to quantify the flavonoid induction before and after frost stress, and determine its anti-oxidant activity in the global sorghum panel and also identify the key candidate genes involved in the process through genome wide association mapping. GWAS identified 51 genes for DPPH (anti-oxidant activity) and 20 genes for total phenolic content (TP) before frost. Whereas, 19 genes were identified 3-deoxyanthocyanidins (3-DAs) and 6 genes for TPC under after frost stress. Most of the identified genes are involved in plant defense pathways for biotic and abiotic resistance. The probable candidates for after stress were peroxidase (sb01g041770), vesicle fusing ATPase (sb01g041930), leucine rich repeats (sb08g001430), flavonol reductase (sb01g025770), cytosine deaminase (sb02g041770), and anthocyanidin reductase (sb02g038520) for 3-DAs; and leucine rich repeat (Sb08g023030) and UDP-glucosyl

transferase (Sb06g021900) for TP. Anti-fungal gene (Sb06g014340) was the most probable candidate for DPPH before frost stress. The identified candidate genes can be used in breeding climate resilient sorghum.

Sorghum accumulates epi-cuticular wax (EW) or bloom in plant surfaces such as leaves, sheaths, and culm. EW reduces the non-transpiration water loss and protects the plant from severe drought stress and also imparts resistance against various insect pests. We present here results from the analysis of epi-cuticular wax (EW) content of 387 diverse sorghum accessions and its genome-wide association study (GWAS). EW content in sorghum leaves ranged from 0.1 mg cm^{-2} to 29.7 mg cm^{-2} with a mean value of 5.1 mg cm^{-2}. GWAS using 265,487 single nucleotide polymorphisms (SNPs) identified thirty-seven putative genes that were associated ($p < 9.89E-06$) with EW biosynthesis and transport in sorghum; out of these, Sobic.002G310400 (3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal), Sobic.004G154200 (Ankyrin repeats), and Sobic.003G004500 (bHLH-MYC and R2R3-MYB transcription factors N-terminal) have been shown to be involved in EW biosynthesis; Sobic.001G447300 (ABC transporter), Sobic.004G089400 (Lipid exporter ABCA1 and related proteins, ABC superfamily), Sobic.002G311200 (Multidrug resistance protein), Sobic.001G447200 (Inositol 1, 3, 4-trisphosphate 5/6-kinase), and Sobic.005G217500 (Cytochrome p450) are involved in EW regulation or transport in sorghum. GWAS results from this study demonstrate the potential for genetic manipulation of EW content in sorghum and other grain cereals for better adaptation to biotic and abiotic stresses.

Brown mid-rib (BMR) forage sorghum (*Sorghum bicolor* (L.) Moench) silage is a reasonable alternative to corn silage for areas with limited soil moisture. Traditional forage sorghum varieties are tall and prone to lodging with lower forage quality. Brachtyic dwarf BMR forage lines are shorter, lodging resistant and have higher forage quality. Newer, earlier hybrids have expanded the potential adaptation of forage sorghums to more northern areas. A two-year study was conducted during the 2014 and 2015 growing seasons using newly available brachytic dwarf BMR forage hybrids to determine the effects of different seeding rates and N (nitrogen) fertilization rates on forage dry matter (DM) yield and quality for two hybrids. The experimental design was split-split-plot with four replications. In each replication, main plots were two hybrids (AF7202 and AF7401), subplots were two seeding rates (198,000 and 296,400 seeds ha^{-1}), and sub-subplots were two N rates (123 and 168 kg ha^{-1}). DM yield and forage quality parameters were measured for each treatment. We observed significant varietal differences for all the parameters except neutral detergent fiber digestibility (NDFD) in 2015. The early maturity line, AF7202, had higher yields, higher starch and net energy for lactation (NE_L) levels than AF7401. The dwarf line, AF7401, had higher crude protein (CP) and NDFD than AF7202. AF7202 was more responsive to the higher N rate than AF7401. CP was increased with increased N for both varieties. Other forage quality traits were unaffected by N rates. Neither variety responded to an increase in seeding rate. This study showed that the earlier brachytic dwarf forage sorghums such as AF7202, managed with recommended seeding rates and possibly higher N rates have good potential for high forage yield and quality in central PA.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My PhD is the product of a wonderful cooperation with many people to whom I would like to express my gratitude and appreciation. Overall, my experience at Penn State was very enriching and fruitful. I would like to thank all the course teachers for their dedication and relentless hard work.

I acknowledge my advisor, Dr. Surinder Chopra, and my co-advisor, Dr. Gregory W Roth, for their guidance, support and understanding through my graduate carrier at Penn State. I was extremely fortunate to have a wide range of responsibilities in Chopra's lab while keeping a high degree of freedom. I owe Dr. Chopra and Dr. Roth the sincere respect as a great personality and good mentors.

My deep gratitude and appreciation also goes to my research committee, Drs. Yinong Yang, Majid Foolad, and Lavanya Reddivari, for their valuable discussion and constant encouragement. Their constructive criticism and helpful suggestions have been of great impact on shaping my research ideas and improving my scientific thinking.

I would like to thank present and past members of Chopra lab for their help and support. Kam for planting sorghum trials. Qin, Deb, Bin, Laura, Doug, and Cullen for helping me in field work and sample collection. Past and present members of Reddivari's lab for helping me in anti-oxidant analysis and sample collection for quantifying the induced level of phenolic compounds after frost stress. Special thanks to Drs. Iffa Gaffoor (fungal pathogen cultures and infection assay) and Weiya Xue (data analysis) for their suggestions throughout the course of my lab and field work.

My sincere thanks to GRIN germplasm network, Dr. Wilfred Vermerris, and Dr. Geoff Morris for providing us the sorghum germplasm resources and SNP genotypic data. I also would like to thank GAPIT forum for helping me in figuring out the association mapping models.

I would like to thank Mark Antle and Austin Kirt for helping us in planting sorghum forage trials. My sincere thanks to Scott Harckom, Agronomy farm manager and his team members for arranging all the necessary things in the field for smooth running of the experiments. Also, I want to thank green house manager, Scott DiLoreto for helping in getting the work done in controlled facilities.

Special thanks to two amazing undergrad students Ryan Fritz and John Frank for their help in NAM population screening for anthracnose disease in the green house facility. Thanks to Drs. HC Sharma, and Are Ashok Kumar for accommodating me at ICRISAT for a semester in 2015 to work on sorghum-aphid interaction project. I would also like to thank many wonderful people I met at Penn State. Thanks to the hard-working staff in the academic units and the international programs for their effort for helping us not worried about anything but research. Thanks for the department of plant science graduate assistantship and competitive research grant award.

No matter the details I can give, I will not be able to adequately express my gratitude and appreciation to my parents, Elango and Rajam, who endure the greatest pain in raising us. They spared no effort and time in supporting us with care, and guidance. They have been my greatest motivation for perusing knowledge. I am indebted to my wife Vimala for her

sacrifice, wonderful support, love, and care. I also want to thank my ever-loving sister, Malini for her support and love and my nephew, Bharathi and niece, Abirami. Also, I want to thank my brother in law, Neelakandan for his help. I dedicate this thesis work to my loving grandmother, Indiramani, and grandfather, Chellaiah who always blessing me from heaven.

Finally, my great thanks go to the funding agency, department of plant science, friends, and colleagues here at Penn State for their help and support to achieve my dreams.

Chapter 1

Introduction

1.1. General description of sorghum origin and botany

Sorghum is a C_4 cereal crop (food, feed, fodder and fuel) and well adapted to harsh climatic conditions. Currently, sorghum is the fifth most important cereal crop after rice, wheat, maize and barley and is the dietary staple of more than 500 million people in over 90 countries, primarily in the developing world (Reddy et al., 2012). Worldwide, it is grown on about 40 million ha. More than 35 percent of world sorghum production is going for food consumption (Awika and Rooney, 2004). The acreage under sorghum has been increased by 24% in the year 2015 in the USA alone.

Sorghum is an often-self-pollinated crop, in which cross-pollination does not exceed more than 25% under favorable conditions (House, 1985). Producing pure lines and maintaining purity is much easier in sorghum, whereas, making heterozygous cross combinations require skills such as emasculation (removal of anther) and identification of right flower stage. Sorghum is a short-day plant, the vegetative bud remains vegetative until the day length becomes short enough for the floral bud to develop and this point of conversion is called the critical photoperiod (Miller et al., 1968). Once the plant achieves critical photoperiod it starts to initiate flowering (House, 1985). Most of the tropically adapted sorghum lines are not suitable for temperate climatic conditions because they do not flower under temperate conditions due to long day length (more than 12.5 hrs).

Sorghum originated in Ethiopia, and most of the germplasm accessions came from tropics

of Africa and India. So, photo-period sensitivity is common in global germplasm accessions. With this in mind, Casa et al. (2008) developed a community resource, which consists of 377 sorghum accessions. They converted all these sorghum accessions for temperate climatic adaptations especially for USA (United States of America). This sorghum association mapping panel (SAP) has been characterized using 47 simple sequence repeat (SSR) markers (Casa et al., 2008). The sorghum association panel is freely available for sorghum genetic improvement through USDA's germplasm resource information network (GRIN).

The cultivated *Sorghum bicolor* has five distinct races namely bicolor, guinea, caudatum, kafir and durra and ten hybrid races (Fig. 1). All of the fifteen races have been easily identifiable based on their spikelet structure (Harlan and de Wet, 1972). Sorghum usually starts to heading after 45 days from planting, and panicle emergence will be followed by 10 to 15 days from heading. Early maturing lines start to flower around 50 to 55 days. Late maturing lines typically flower around 80 to 85 days. Day length determines the flowering phase in sorghum by initiating the critical photo-period required for flowering. Sorghum flowers from top to bottom. We need to keep this phenomena in mind, while performing emasculation or self-pollination, sometime this helps in identification of the right flower for emasculation or self-pollination. Sorghum is gaining importance in the food value chain as it is gluten free and have anti-inflammatory and anti-proliferative properties against various cancer cell lines (Shih et al., 2007; Devi et al., 2011; Massey et al., 2014; Yang et al., 2014). Sorghum is also known to synthesize varying degree of flavonoid compounds that are toxic to certain insect pests compared to their cereal counterparts in the poaceace

family (Ejike et al., 2013).

1.2. General flavonoid pathway and function

Flavonoids are plant secondary metabolites that share a basic C_6-C_3-C_6 structure, consisting of two aromatic rings joined by a three-carbon heterocyclic ring (Awika, 2011). They are present ubiquitously throughout the plant kingdom. Over 9,000 compounds of this group are known (Buer et al., 2010). Their biosynthesis pathway begins with the condensation of one p-coumaroyl-CoA molecule with three molecules of malonyl-CoA to yield chalcone, catalyzed by chalcone synthase (CHS). The next step is isomerization of chalcone to flavanone by chalcone isomerase (CHI). From this step onwards, the pathway branches to several different flavonoid classes, including aurones, dihydrochalcones, flavanonols (dihydro flavonols), isoflavones, flavones, flavonols, leucoanthocyanidins, anthocyanins and proanthocyanidins.

Flavonoids undergo further modifications, for example, methylation by methyltransferases and glycosylation by specific glycosyltransferases. These modifications often alter their solubility, reactivity, and stability (Mierziak et al., 2014). The majority of flavonoids are present in the form of glycosides under natural conditions. Because of their diverse chemical structure and variety resulting from the attached substituents, they have a number of important functions in plants. They participate in plant protection against biotic (herbivores, pathogens) and abiotic stresses (UV radiation, heat), and due to their anti-oxidative properties, they also maintain a redox state in cells (Mierziak et al., 2014). They play a role in providing color, fragrance and taste to the fruits, flowers, and seeds, which

makes them attractants for insects, birds or mammals, which aid in pollen or seed transmission (Grotewold, 2006).

1.3. Flavonoid Phytoalexins

Flavonoid phytoalexins have been discovered in gymnosperms and angiosperms. The first characterized phytoalexin, *pisatin*, was isolated and reported in 1960 from *Pisum sativum*. Since then, phytoalexins have been found in at least 75 host plants (Ejike et al., 2013). Phytoalexins are defined as low molecular weight, inducible, secondary metabolites with activity against multiple biotic attackers (Müller and Börger 1940). The name phytoalexin was first used by Müller and Börger (1940) who observed the above phenomenon in potato tubers infected by the oomycete *Phytophthora infestans*.

Among poaceace family crop members, sorghum produces novel phytoalexins. However, other cereal counterparts such as rice, maize, wheat and barley also synthesize various secondary phytoalexin compounds. In rice, 15 phytoalexins that include 14 diterpenoid phytoalexins and one flavonoid phytoalexin have been identified following treatment with elicitors such as chitin oligosaccharide (biotic) and in leaves infected with the blast fungus *Magnaporthe grisea* (biotic) or irradiated with UV light (abiotic) (Kodama et al., 1988; Kodama et al., 1992; Koga et al., 1995). Other known elicitors of phytoalexin production in rice are cerebrosides and xylanase protein from *Trichoderma viride*, and ethylene-inducing xylanase in rice cultured cells. These elicitors are known to bring about a variety of defense responses in rice (Kurusu et al., 2010; Okada et al., 2009). They are classified as oryzalexins, phytocassanes, momilactones and sakuranetin. Maize also produces

labdane-type diterpenoid phytoalexins termed kauralexins as well as acidic sesquiterpenoids known as zealexins (Ejike et al., 2013). At concentrations as low as 10 μg ml^{-1}, kauralexin B3 significantly inhibited the growth of the opportunistic necrotroph *R. microsporus* and the causal agent of anthracnose stalk rot, *Colletotrichum graminicola* (Schmelz et al., 2011). Table 1 describes the induction of different types of flavonoid metabolites due to biotic and abiotic stress signals in various crop plants (Dao et al., 2011). Often, phytoalexin synthesis requires the external stress elicitor (pathogen infection) and their production requires the expenditure of plant energy.

1.4. Sorghum Phytoalexins

Plant resist pathogens through pre-formed structures and compounds, inducible post-infection defense compounds (anti-fungal phytoalexins), PAMP- triggered immunity (pathogen associated molecular patterns) and effector triggered immunity. Sorghum synthesizes flavonoid class anti-fungal phytolaexins in response to *Colletotrichum sublineola* ingress (Nicholson et al. 1987). These anti-fungal phytolaexins in sorghum are known as 3-deoxyanthocyanidins (3- DAs). The 3-deoxyanthocyanidins may arise via colorless precursor flavan-4-ol (λmax 564nm) in the flavonoid bio-synthetic pathway (Styles and Ceska 1977; Shekon and Chopra 2009). The major 3-DA family compounds in sorghum are apigeninidin (λmax 480nm) and luteolinidin (λmax 500nm). The synthesis of 3-DAs occur at the attempted site of pathogen entry (Snyder and Nicholson 1990). Initially, they are colorless, 30 hours after post infection the intensity of the anti-fungal compound color increases and release their contents into the cytoplasm of the fungus, in the event of

that they produce additional red colored anti-fungal compounds in surrounding cells and killing the pathogen eventually (Fig. 2). These 3-DA family compounds play a major role in restricting the fungus hyphae growth and imparting the anthracnose disease resistance to sorghum (Ibraheem et al., 2010).

Anthracnose is the major fungal disease in sorghum that causes yield loss of more than 50% under favorable environmental conditions (Thakur & Mathur, 2000) (Fig 3). Anthracnose fungus can overwinter in soil and plant debris. This overwintered inoculum is sought to cause disease next year. *Colletotrichum sublineola*, a pathogen of sorghum anthracnose and *Colletotrichum graminicola*, a pathogen of maize anthracnose has been shown to be interchangeable between the two-plant species (Vernard and Villancourt, 2007; Wheeler et al., 1974). There is no major gene resistance reported for anthracnose. Often resistance in one location fails in another location. Thus, developing durable disease resistant sorghum cultivar is very important.

Different genotypes have different levels of phenolic content, especially colored sorghum grains shown to have high phenolic content and anti-oxidative activities compared to white grains (Dykes et al., 2014; Pfeiffer and Rooney, 2015; Taleon et al., 2014; Vietor et al., 2010). Sorghum genotypes also differ in their ability to synthesize 3-DAs (Ibraheem et al., 2010). Anthracnose disease resistant genotypes synthesize 3-DAs higher and faster compared to susceptible genotypes (Basavaraju et al., 2009; Ibraheem et al., 2010). The 3-DA synthesis in sorghum is dependent on the presence of functional copies of the *yellow seed1* (*y1*) gene (Ibraheem et al., 2010; Ibraheem et al., 2015). We wanted to screen the large collection of sorghum accessions for 3-DA variability and its association to fungal

toxicity and also to map the 3-DA regulators through genome-wide association study. Therefore, this method will enable us to select disease resistant genotypes and accelerate field breeding efforts to develop disease resistant cultivars.

So far, we know seven of the anti-fungal compounds (3-DAs), which includes luteolinidin, apigeninidin, and their methyl and acyl derivatives namely apigeninidin-5- glucoside, 7-O-methyl-apigeninidin, luteolinidin-5-glucoside, 5- methoxyluteolinidin and a newly identified novel pyrano-apigeninidin-4- vinyl phenol (Awika and Rooney, 2004; Bai et al., 2014). Besides 3-DAs anti-fungal activity, 3-DA offers plenty of health benefits for human well-being. The 3-DAs are known to have high anti-oxidant and phase II enzyme activities, and therefore can suppress the growth of various cancer cells (colon cancer, breast cancer) in humans (Awika and Rooney, 2004; Massey et al., 2014; Yang et al., 2009). Generally, 3-DAs are highly stable at high temperature and in acidic conditions (Awika and Rooney, 2004; Yang et al., 2014). Hence, 3-DAs can also be used as a natural food colorant. In Africa, people grow sorghum for extracting red color compounds from sheaths to use in cheese coloring. Recent report indicates that, this sorghum sheath extract possesses high 3-DAs and anti- oxidant activities (Kayode et al., 2011). The demand for natural and organic food colorants are increasing day by day. In this context, profiling and quantification of 3-DA variations in the large sorghum genetic pools will help to identify ectopic expressers of 3-DAs for improving plant and human health.

1.5. Epi-cuticular wax pathway and their role in plant defense

Cuticular waxes are the hydrophobic compounds on the surface of the plant that can be

removed by a brief immersion in an organic solvent such as chloroform or hexane. They are complex mixtures of primarily very long chain ($>C_{18}$) fatty acids, hydrocarbons, alcohols, aldehydes, ketones, esters, triterpenes, sterols, and flavonoids. The wax pathway starts from precursors of fatty acids that are likely derived from de novo synthesis in plastids. The de novo fatty acid biosynthesis is catalyzed by a series of enzymatic steps, collectively referred to as fatty acid synthase (FAS). The initiation of fatty acid synthesis is the condensation of malonyl-acyl carrier protein (ACP) with acetyl-CoA, followed by the sequential reduction of 3- ketoacyl-ACP, the dehydration of 3- hydroxyacyl-ACP, and the reduction of trans- Δ^2-enoyl-ACP. The fatty acyl primer remains esterified to the ACP cofactor and is further extended two carbons at a time by the donor, malonyl-ACP. For each two- carbon addition, there is a sequential round of condensation, reduction, dehydration, and second-reduction steps. NAD(P)H serve as reducing equivalents for the two reductases. The long chain products (C16, C18) are subsequently processed by one or more enzymes, including stearoyl-ACP desaturase, plastidial acyltransferases, and acyl-ACP thioesterases (hydrolases). Fatty acids are then utilized for glycerolipids, waxes, or cutin and suberin biosynthesis, depending on the tissue type and developmental stage.

The proportions of the wax major classes vary among plant species and also vary within species and tissues. Sorghum is known to accumulate varying classes of EW in stems and leaves at varying degrees. EW deposition is highly visible on the abaxial side of the leaf blade, culm, and peduncle of sorghum (Burow et al., 2009). EW production is beginning to appear at seedling stage in sorghum. However, the peak synthesis of EW occurs between pre-flowering to maturity stage. EW provides the outermost barrier between plants and

their environment. It has been shown to impart tolerance to a number of environmental stresses, including freezing (Thomas and Barber 1974), insects (Edwards 1982; Maiti et al. 1984), pathogens (El-Otmani et al. 1989), mechanical damage (Eglinton and Hamilton 1967), air pollutants and acid rain (Percy and Baker 1990), excessive ultraviolet radiation (Eglinton and Hamilton 1967; Reicosky and Hanover 1978), heat (Jefferson et al. 1989), and drought (Blum 1975; Bengston et al. 1978). However, reduced wax (sparse bloom) or no wax (bloomless) sorghum plants known to exhibit non-preference effect to greenbugs (Peiretti et al. 1980) and resistance sheath blight disease and susceptible to sorghum leaf blight upon *Exserohilum turcicum* infection (Jenks, 1992). The amount of wax coverage on leaf sheath, leaf blade, internode, panicle branch and glume determine the type of bloom genotypes. More wax coverage on the plant body is called profuse bloom (BTx 623, 296B), whereas light coverage of wax is called sparse bloom (IS 18551, KFS 2021) and no coverage is called bloomless.

Sorghum variants for bloom (EW) production were first reported by Ayyangar and Ponnaiya (1941) and were designated either bloomless (lacking visible epi- cuticular wax) or sparse bloom (reduced visible wax). Allelism tests indicated that each of these phenotypes was determined by individual homozygous recessive genes which were designated *bmbm* (bloomless) and *hh* (sparse bloom). Peterson et al. (1982) performed tests of allelism on four sparse bloom and five bloomless mutants in sorghum. Two different loci were shown to control the expression of bloomless, designated *bm*, and *bm2*, and at least three loci, designated *h*, *h2*, and *h3*, governed expression of the sparse bloom phenotype.

Madhusudhana and Patil (2012), identified a major QTL *Dw4*, which co- segregate for bloom production in sorghum while mapping for plant height in the cross between 296 B and IS 18551. (Burow et al., 2009) identified a locus called *BLMC* (BLooMCuticle) responsible for the production of profuse wax in sorghum plants from the mutant population derived from BTx 623 (profuse wax donor) and KFS 2021 (sparse wax). Recently a genomic region associated with bloom inhibition (bloomless production) have been identified by Mizuno et al. (2013) from the gamma irradiated mutant populations. They found that, inversion of Sb06g023280 gene is responsible for inhibition of wax secretion in sorghum. Further, gene annotation revealed that, Sb06g023280 is in fact a family of ABC transporters. Briefly, wax biosynthesis is controlled by single dominant and recessive genes, dominant for profuse and recessive for bloomless.

1.6. Sorghum Diversity Panels

There are five different global sorghum diversity panels available for the sorghum research community to carry out various research activities of interest.

1. SAP (Sorghum Association Panel)
2. ICRISAT mini-core (MNC)
3. NAM (Nested Association mapping)
4. CIRAD landraces
5. NIAB (National Institute for Agro-Biodiversity)

The SAP panel consists of 377 global accessions, which are converted into photoperiod

insensitive lines to US climatic adaptations to facilitate phenotypic characterization in temperate regions. This collection has accessions from all across the globe with historical importance in sorghum breeding (Casa et al., 2008). The ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) mini- core (MNC) collection consist of 242 accessions captured from 2247 sorghum core collections with phenotypic data measured for 11 qualitative traits and 10 quantitative traits (Upadhyaya et al., 2009). This mini-core collection is not converted to temperate adaptations. So, most of them won't set seed due to photoperiod sensitivity in temperate regions. Both SAP and MNC collections are diverse geographically and racially. The Nested Association Mapping (NAM) panel has been developed to capture advantages lying in both QTL linkage mapping and association mapping approach. There are two NAM panels available now. The first panel consist of ten founder lines (SC283, SC1103, Segaolane, Macia, SC35, Ajabsido, SC971, SC265, SC1345, and P898012) crossed to RTx 430. From each cross, 200 RIL's were derived to form a sorghum NAM panel with 2000 RIL's (Paterson ed., 2013). The second NAM panel was developed with ten different founder lines crossed to Tx 623. CIRAD population consist of 200 landraces that are representative of the race, the latitude of origin, response to day length, and production system and characterized them with restricted fragment length polymorphism (RFLP) probes (Deu et al., 2006). The final population NIAB from Japan has 3500 accessions, mostly came from subpopulations. So, they have a narrow genetic base.

1.7. Sorghum Genomics

The sorghum genome was assembled (Paterson et al., 2009) based on whole- genome

shotgun sequence with a detailed genetic map (Bowers et al. 2003) and a BAC-based physical map (Bowers et al. 2005). The breeding line BTx623 was used for genome sequencing. BTx623 is a homozygous advanced breeding line released by Texas A&M University, widely exploited in sorghum breeding as well as for genomics research. More than 98.46% sorghum genome is complete and accurate to less than one error per kb. The assembled sorghum genome size is 730 Mb, which consists of 7.5% DNA transposons, 55% retrotransposons and 1491 loci corresponding to alternate splicing events with a prediction of 34,496 genes out of which 27,640 genes were supported by homology and ab initio gene prediction methods (Paterson et al., 2009). Its small genome size (730 Mb) makes an attractive model system to study the functional genomics of grass families (Paterson et al., 2009).

1.8. Genome-Wide Association Study

Association mapping can be conducted by genotyping all individuals with tens of thousands of SNPs (single nucleotide polymorphism) instead of focusing on candidate genes or regions. Genome-wide association studies have been extensively conducted to dissect the genetic causes of complex human diseases for many years. Recent years GWAS gaining much attention in crop improvement due to the availability of large-scale SNP markers by the advent of low-cost sequencing methodologies. GWAS offers high allelic diversity and high recombination rate due to the deployment of many markers across the genome and also from the large population size. Successful GWAS studies in sorghum identified genomic regions associated with trait of interest for agroclimatic traits (Morris et al., 2013), adaptive traits (Lasky et al. 2015), grain quality (Sukumaran et al., 2012),

seed size (Zhang et al., 2015), grain polyphenol concentrations (Rhodes et al., 2014), Fe and Zn density, stalk rot resistance (Adeyanju et al., 2015), and aluminum tolerance. However, we need to pay critical attention to GWAS approach while working with diverse populations, they may have different maturity groups, photoperiod sensitivity, and flowering time. So, appropriate field design, blocking, the timing of record taking, data analysis, and interpretation of results all required for getting a good phenotypic data, good marker-trait associations and also to avoid false-positive associations for the traits of interest. Missing heritability is another problem reported in human GWAS results. The classic example of missing heritability is human adult height variation. Forty loci have been implicated in controlling adult height variation, but together they explain only 5% of phenotypic variation even though the estimated heritability of this trait is about 80%. Potential causes of this problem include rare allele frequency, epistasis, sample size, structure variants, and the interaction between genotype and environment. The reported level of LD (linkage disequilibrium) decay in sorghum is about 150 kb. The average level of recombination rate in sorghum is 1.4/kb with considerable variations across chromosomes (Morris et al., 2013). We may get to map single gene resolution level at telomeric regions, whereas a Mb level resolution in centromeric regions, due to the high recombination rate at telomeric region and low recombination to no recombination at centromeric regions (Morris et al., 2013).

1.9. Rationale and Research objectives

Sorghum is the climate resilient versatile crop grown on marginal soils. However, sorghum

is highly susceptible to foliar diseases. Anthracnose is a major fungal foliar disease in sorghum, which causes yield reduction up to 50%. Anthracnose pathogen in sorghum and maize are known to cross-infect each other if they are grown on the close proximities. The recent introduction of sorghum to newer areas in the USA and elsewhere may increase the disease pressure. Also, sorghum is getting introduced to high altitude regions such as Pennsylvania for bio-energy and silage purpose. So, identifying resistant genes and other mechanisms to breed better biotic and abiotic tolerant sorghum cultivars are very important for wider adaptation. In this context, our aim is to screen and quantify the stress induced flavonoid metabolites in the large collection of global sorghum accessions and map the genomic regions associated with stress induced flavonoid metabolites in sorghum.

Objectives of this research study:

1. Assessing the genetic variability for 3-DAs present in the large set of sorghum accessions and map the genes responsible for 3-DA biosynthesis.

2. Identification of SNPs associated with frost induced secondary metabolites in sorghum.

3. Assessing the epi-cuticular wax variability present in the large set of sorghum accessions and map the wax biosynthesis genes.

4. Understand the role of seeding and N fertilization rate effect on brachytic BMR forage sorghums in central PA environment.

Literature Cited

Adeyanju, A., Little, C., Yu, J., and Tesso, T. (2015). Genome-Wide Association Study on Resistance to Stalk Rot Diseases in Grain Sorghum. *G3 (Bethesda)* **5**, 1165-75.

Awika, J. M. (2011). Sorghum Flavonoids: Unusual Compounds with Promising Implications for Health. *In* "Advances in Cereal Science: Implications to Food Processing and Health Promotion", pp. 171-200.

Awika, J. M., and Rooney, L. W. (2004). Sorghum phytochemicals and their potential impact on human health. *Phytochemistry* **65**, 1199-221.

Ayyangar, G. N., and B. W. X. Ponnaiya. (1941). The occurrence and inheritance of a bloomless sorghum. *Curr Sci* **10**:408-409.

Bai, Y., Findlay, B., Maldonado, A. F., Schieber, A., Vederas, J. C., and Ganzle, M. G. (2014). Novel pyrano and vinylphenol adducts of deoxyanthocyanidins in sorghum sourdough. *J Agric Food Chem* **62**, 11536-46.

Basavaraju, P., Shetty, N. P., Shetty, H. S., de Neergaard, E., and Jorgensen, H. J. (2009). Infection biology and defence responses in sorghum against Colletotrichum sublineolum. *J Appl Microbiol* **107**, 404-15.

Bengston, C., S. Larsson, and C. Liljenberg. (1978). Effects of water stress on cuticular transpiration rate and amount and composition of epicuticular wax in seedlings of six oat varieties. *Physiol Plant* **44**:319-324.

Blum, A. (1975). Effect of the Bm gene on epicuticular wax deposition and the spectral characteristics of sorghum leaves. *SABRAO J* **7**,45-52.

Bowers, J.E., Colette Abbey, Sharon Anderson, Charlene Chang, Xavier Draye, Alison H. Hoppe, Russell Jessup, Cornelia Lemke, Jennifer Lennington, Zhikang Li, Yann-rong Lin, Sin-chieh Liu, Lijun Luo, Barry S. Marler, Reiguang Ming, Sharon E. Mitchell, Dou Qiang, Kim Reischmann, Stefan R. Schulze, D. Neil Skinner, Yue-wen Wang, Stephen Kresovich, Keith F. Schertz and Andrew H. Paterson. (2003). A High-Density Genetic Recombination Map of Sequence-Tagged Sites for Sorghum, as a Framework for Comparative Structural and Evolutionary Genomics of Tropical Grains and Grasses. *Genetics* **165**, 367-386.

Bowers, J.E., Miguel A. Arias, Rochelle Asher, Jennifer A. Avise, Robert T. Ball, Gene A. Brewer, Ryan W. Buss, Amy H. Chen, Thomas M. Edwards, James C. Estill, Heather E. Exum, Valorie H. Goff, Kristen L. Herrick, Cassie L. James Steele, Santhosh Karunakaran, Gmerice K. Lafayette, Cornelia Lemke, Barry S. Marler, Shelley L. Masters, Joana M. McMillan, Lisa K. Nelson, Graham A. Newsome, Chike C. Nwakanma, Rosana N. Odeh, Cynthia A. Phelps, Elizabeth A. Rarick, Carl J. Rogers, Sean P. Ryan, Keimun A. Slaughter, Carol A. Soderlund, Haibao Tang, Rod A. Wing, Andrew H. Paterson. (2005). Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. *Proceedings of the National Academy of Sciences* **102**, 13206-13211.

Buer, C. S., Imin, N., and Djordjevic, M. A. (2010). Flavonoids: new roles for old molecules. *J Integr Plant Biol* **52**, 98-111.

Burow, G. B., Franks, C. D., Acosta-Martinez, V., and Xin, Z. (2009). Molecular mapping and characterization of BLMC, a locus for profuse wax (bloom) and enhanced

cuticular features of Sorghum (Sorghum bicolor (L.) Moench.). *Theor Appl Genet* **118**, 423-31.

Edwards, P. B. (1982). Do waxes on juvenile Eucalyptus leaves provide protection from grazing insects? *Aust. J. Ecol.* **7**, 347-352.

Eglinton, G., and R. J. Hamilton. (1967). Leaf epicuticular waxes. *Science* **156**, 1322-1335.

El-Otmani, M., M. L. Arpaia, C. W. Coggins, Jr., J. E. Pehrson, Jr., and N. V. O'Connell. (1989). Developmental changes in 'Valencia' orange fruit epicuticular wax in relation to fruit position on the tree. *Sci. Hort.* **41**, 69-81.

Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., Franks, C. D., and Kresovich, S. (2008). Community Resources and Strategies for Association Mapping in Sorghum. *Crop Science* **48**.

Dao, T. T., Linthorst, H. J., and Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochem Rev* **10**, 397-412.

Deu, M., Rattunde, F., and Chantereau, J. (2006). A global view of genetic diversity in cultivated sorghums using a core collection. *Genome* **49**, 168-80.

Devi, P. S., Kumar, M. S., and Das, S. M. (2011). Evaluation of antiproliferative activity of red sorghum bran anthocyanin on a human breast cancer cell line (mcf-7). *Int J Breast Cancer* **2011**, 891481.

Dykes, L., Hoffmann, L., Portillo-Rodriguez, O., Rooney, W. L., and Rooney, L. W. (2014). Prediction of total phenols, condensed tannins, and 3-deoxyanthocyanidins in sorghum grain using near-infrared (NIR) spectroscopy. *Journal of Cereal Science* **60**, 138-142.

Ejike, C. E. C. C., Gong, M., and Udenigwe, C. C. (2013). Phytoalexins from the Poaceae: Biosynthesis, function and prospects in food preservation. *Food Research International* **52**, 167-177.

Grotewold, E. (2006). The Science of Flavonoids. *Springer*, 1-269.

Harlan, J. R., and de Wet, J.M.J. (1972). A Simplified Classification of Cultivated Sorghum. *Crop Science* **12**, 172-176.

House, L. R. (1985). A guide to sorghum breeding. *ICRISAT* **2**, 1-205.

Ibraheem, F., Gaffoor, I., and Chopra, S. (2010). Flavonoid phytoalexin-dependent resistance to anthracnose leaf blight requires a functional yellow seed1 in Sorghum bicolor. *Genetics* **184**, 915-26.

Ibraheem, F., Gaffoor, I., Tan, Q., Shyu, C. R., and Chopra, S. (2015). A sorghum MYB transcription factor induces 3-deoxyanthocyanidins and enhances resistance against leaf blights in maize. *Molecules* **20**, 2388-404.

Jefferson, P. G., D. A. Johnson, and K. H. Asay. (1989). Epi cuticular wax production, water status and leaf temperature in Triticeae range grasses of contrasting visible glaucous ness. *Can. J. Plant Sci.* **69**, 513-519.

Jenks, M. A., P.J. Rich, P.J. Peters, J.D. Axtell, and E.N. Ashworth (1992). Epicuticular wax morphology of bloomless (bm) mutants in Sorghum bicolor. *Int J Plant Sci* **153**, 311-319.

Jesse R. Lasky, H. D. U., Punna Ramu, Santosh Deshpande, C. Tom Hash, Jason Bonnette, Thomas E. Juenger, Katie Hyma, Charlotte Acharya, Sharon E. Mitchell, Edward S. Buckler, Zachary Brenton, Stephen Kresovich, Geoffrey P. Morris (2015).

Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances* **1**, 1-13.

Kayode, A. P., Nout, M. J., Linnemann, A. R., Hounhouigan, J. D., Berghofer, E., and Siebenhandl-Ehn, S. (2011). Uncommonly high levels of 3-deoxyanthocyanidins and antioxidant capacity in the leaf sheaths of dye sorghum. *J Agric Food Chem* **59**, 1178-84.

Kodama, O., Miyakawa, J., Akatsuka, T., and Kiyosawa, S. (1992). Sakuranetin, a flavanone phytoalexin from ultraviolet-irradiated rice leaves. *Phytochemistry* **31**, 3807–3809.

Kodama, O., Suzuki, T., Miyakawa, J., & Akatsuka, T. (1988). Ultraviolet-induced accumulation of phytoalexins in rice leaves. *Agricultural and Biological Chemistry* **52**,
2469–2473.

Koga, J., Shimura, M., Oshima, K., Ogawa, N., Yamauchi, T., & Ogasawara, N. (1995). Phytocassanes A, B, C and D, novel diterpene phytoalexins from rice, Oryza sativa L. *Tetrahedron* **51**, 7907–7918.

Kurusu, T., Hamada, J., Nokajima, H., Kitagawa, Y., Kiyoduka, M., Takahashi, A., Hanamata, S., Ohno, R., Hayashi, T., Okada, K., Koga, J., Hirochika, H., Yamane, H., and Kuchitsu, K. (2010). Regulation of Microbe-Associated Molecular Pattern-Induced Hypersensitive Cell Death, Phytoalexin Production, and Defense Gene Expression by Calcineurin B-Like Protein-Interacting Protein Kinases, OsCIPK14/15, in Rice Cultured Cells. *Plant Physiology* **153**, 678-692.

Madhusudhana, R., and Patil, J. V. (2012). A major QTL for plant height is linked with bloom locus in sorghum [Sorghum bicolor (L.) Moench]. *Euphytica* **191**, 259-268.

Massey, A. R., Reddivari, L., and Vanamala, J. (2014). The Dermal Layer of Sweet Sorghum (Sorghum bicolor) Stalk, a Byproduct of Biofuel Production and Source of Unique 3-Deoxyanthocyanidins, Has More Antiproliferative and Proapoptotic Activity than the Pith in p53 Variants of HCT116 and Colon Cancer Stem Cells. *J Agric Food Chem* **62**, 3150-3159.

Mierziak, J., Kostyn, K., and Kulma, A. (2014). Flavonoids as important molecules of plant interactions with the environment. *Molecules* **19**, 16240-65.

Miller F.R, D. K. B., and H. J. Cruzado (1968). Effect of Tropical Photoperiods on the Growth of Sorghum When 1 Grown in 12 Monthly Plantings. *Crop Science* **8**, 499-509.

Mizuno, H., Kawahigashi, H., Ogata, J., Minami, H., Kanamori, H., Nakagawa, H., and Matsumoto, T. (2013). Genomic inversion caused by gamma irradiation contributes to downregulation of a WBC11 homolog in bloomless sorghum. *Theor Appl Genet* **126**, 1513-20.

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., and Kresovich, S. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc Natl Acad Sci U S A* **110**, 453-8.

Müller, K. O., & Börger, H. (1940). Experimentelle untersuchungen über die phytophthora-resistenz der kartoffel, arbeiten aus der biologischen reichsanstalt für land-und forstwirtschaft. *Berlin-Dahlem* **23**, 189–231.

Nicholson, R. L., S. S. Kollipara, J. R. Vincent, P. C. Lyons and G. Cadena-Gomez. (1987). Phytoalexin synthesis by the sorghum mesocotyl in response to infection by pathogenic and nonpathogenic fungi. Proc. Natl. Acad. Sci. 84, 5520–5524.

Okada, A., Okada, K., Miyamoto, K., Koga, J., Shibuya, N., Nojiri, H., and Yamane, H. (2009). OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice. *J Biol Chem* **284**, 26510-8.

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, R., Ware, 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.

Percy, K. E., and E. A. Baker. (1990). Effects of simulated acid rain on epicuticular wax production, morphology, chemical composition and on cuticular membrane thick ness in two clones of Sitka spruce [Picea sitchensis (Bong.) Carr.]. New Phytol. 116, 79-87.

Peterson, G. C., S. Krittika, and D. E. Weibel. (1982). In heritance of some bloomless and sparse-bloom mutants in sorghum. Crop Sci. 22, 63-67.

Pfeiffer, B. K., and Rooney, W. L. (2015). Sunlight Induces Black Color and Increases Flavonoid Levels in the Grain of Sorghum Line Tx3362. *Crop Science* **55**.

Reddy, B. V. S., Reddy S, Sadananda A.R, Elango Dinakaran, Kumar A.A, Deshpande S.P, Srinivasa Rao, Sharma H.C, Rajan Sharma, Krishnamurthy L, Patil J.V. (2012). Postrainy season sorghum: Constraints and breeding approaches. *SAT eJournal* **10**.

Reicosky, D. A., and J. W. Hanover. (1978). Physiological effects of surface waxes. I. Light reflectance for glaucous and non-glaucous Picea pungens. *Plant Physiol.* **62**, 101- 104.

Rhodes, D. H., Hoffmann, L., Jr., Rooney, W. L., Ramu, P., Morris, G. P., and Kresovich, S. (2014). Genome-wide association study of grain polyphenol concentrations in global sorghum [Sorghum bicolor (L.) Moench] germplasm. *J Agric Food Chem* **62**, 10916-27.

Schmelz, E. A., Kaplan, F., Huffaker, A., Dafoe, N. J., Vaughan, M. M., Ni, X., Rocca, J. R., Alborn, H. T., and Teal, P. E. (2011). Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proc Natl Acad Sci U S A* **108**, 5455-60.

Sekhon, R. S., and S. Chopra. (2009). Progressive loss of DNA methylation releases epigenetic gene silencing from a tandemly repeated maize Myb gene. *Genetics* **181**, 81–91.

Shih, C.H., Elaine Wong, Lawrence C. M. Chiu, Ivan K. Chu, and Clive Lo (2007). Quantitative analysis of anticancer 3-deoxyanthocyanidins in infected sorghum seedlings. *Journal of Agricultural and Food Chemistry* **55**, 254-259.

Snyder, B. A., and Nicholson, R.L. (1990). Synthesis of Phytoalexins in Sorghum as a Site-Specific Response to Fungal Ingress. *Science* **248**, 1637-1639.

Styles, E. D., and O. Ceska. (1977). Genetic control of flavonoid synthesis in maize. Can. J. Genet. and Cytol. 19, 289–302.

Sukumaran, S., Xiang, W., Bean, S. R., Pedersen, J. F., Kresovich, S., Tuinstra, M. R., Tesso, T. T., Hamblin, M. T., and Yu, J. (2012). Association Mapping for Grain Quality in a Diverse Sorghum Collection. *The Plant Genome Journal* **5**.

Taleon, V., Dykes, L., Rooney, W. L., and Rooney, L. W. (2014). Environmental effect on flavonoid concentrations and profiles of red and lemon-yellow sorghum grains. *Journal of Food Composition and Analysis* **34**, 178-185.

Thakur, R.P. and K. Mathur, 2000. Anthracnose, in Compendium Sorghum Diseases, R.A. Frederiksen and G.N. Odovody, Editors. *APS Press: St. Paul*, MN. p. 10-12.

Thomas, D. A., and H. N. Barber. (1974). Studies on leaf characteristics of a cline of Eucalyptus urnigera from Mount Wellington, Tasmania. II. Reflection, transmission and ab sorption of radiation. Aust. J. Bot. 22, 701-707.

Upadhyaya, H. D., Pundir, R. P. S., Dwivedi, S. L., Gowda, C. L. L., Reddy, V. G., and Singh, S. (2009). Developing a Mini Core Collection of Sorghum for Diversified Utilization of Germplasm. *Crop Science* **49**.

Venard, C. and L. Vaillancourt, Penetration and colonization of unwounded maize tissues by the maize anthracnose pathogen Colletotrichum graminicola and the related nonpathogen C. sublineolum. *Mycologia*, 2007. 99(3): p. 368-377.

Vietor, D. M., Rhodes, G. A., and Rooney, W. L. (2010). Relationship of phenotypic variation in sorghum to nutritive value of crop residues. *Field Crops Research* **118**, 243-250.

Wheeler, H., D.J. Politis, and C.G. Poneleit., 1974. Pathogenicity, host range, and distribution of Colletotrichum graminicola on corn. *Phytopathology*, 64(3): p. 293-296.

Yang, J., Martinson, T. E., and Liu, R. H. (2009). Phytochemical profiles and antioxidant activities of wine grapes. *Food Chemistry* **116**, 332-339.

Yang, L., Dykes, L., and Awika, J. M. (2014). Thermal stability of 3-deoxyanthocyanidin pigments. *Food Chem* **160**, 246-54.

Zhang, D., Li, J., Compton, R. O., Robertson, J., Goff, V. H., Epps, E., Kong, W., Kim, C., and Paterson, A. H. (2015). Comparative Genetics of Seed Size Traits in Divergent Cereal Lineages Represented by Sorghum (Panicoidae) and Rice (Oryzoidae). *G3 (Bethesda)* **5**, 1117-28.

Figure Legends

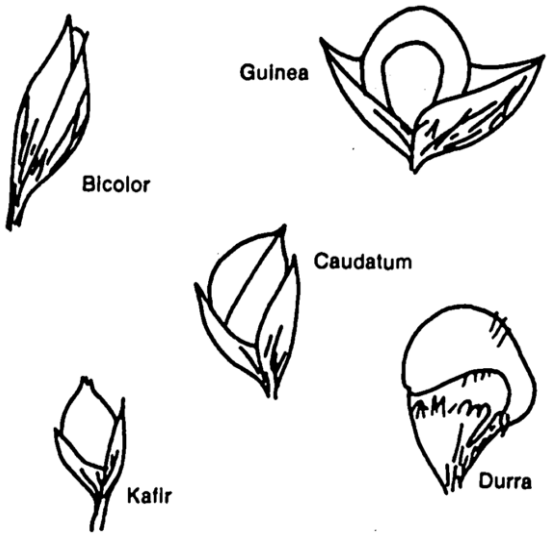

Figure 1. Spikelet types of basic sorghum races (Adapted from House, 1985).

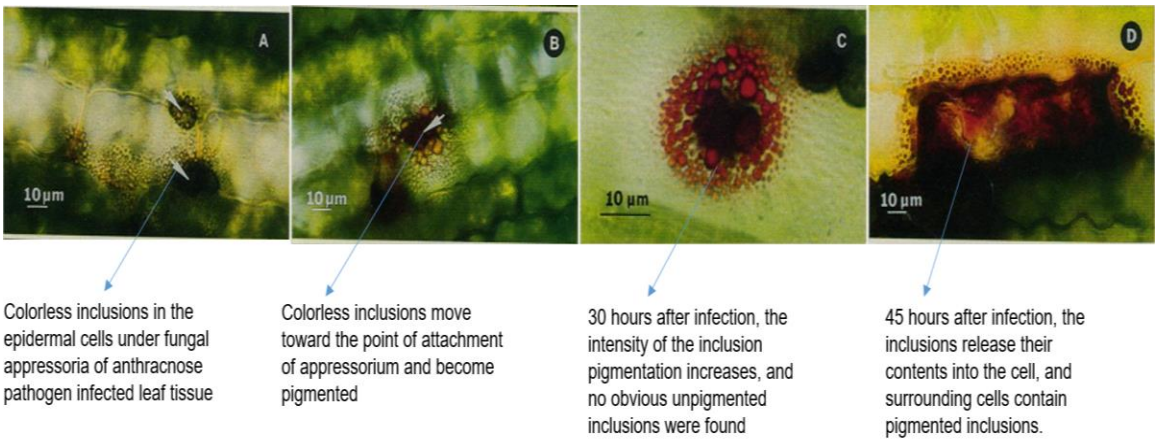

Figure 2. Sorghum anti-fungal phytoalexins (3-DAs) induction upon anthracnose fungus ingress (Adapted from Synder and Nicholson, 1990).

Figure 3. Anthracnose disease severity in the sorghum genotype, H 112 at the Penn state agronomy farm, Rock springs, PA.

Figure 4. Epi-cuticular wax variability in the profuse bloom (296B) and sparse bloom (IS 18551) producing sorghum genotypes (Adapted from Madhusudhana and Patil, 2012).

Table 1. Induction of flavonoid compounds in different crops due to biotic and abiotic stresses (Adapated from Dao et al. 2011).

Host	Pathogen/stresses	Metabolites	References
Petroselinum crispum (Parsley)	UV	Flavonoids	Schmelzer et al. (1988), Schulze et al. (1989)
Phaseolus vulgaris cells (French bean)	*Colletotrichum lindemuthianu*		Ryder et al. (1984)
Arabidopsis cells	UV-B and UV-A/blue light		Christie and Jenkins (1996)
Arabidopsis thaliana	Low temperature	Anthocyanins	Leyva et al. (1995)
	UV-B, UV-A, and blue Light		Fuglevand et al. (1996), Hartmann et al. (1998), Wade et al. (2001)
	High-intensity lights	Anthocyanins	Feinbaum and Ausubel (1988)
	SA, ethylene, methyl jasmonate		Schenk et al. (2000)
	Pseudomonas syringae	Phenolic compounds	Soylu (2006)
Petunia hybrida	UV	Anthocyanin	Koes et al. (1989)
	Low temperature		Shvarts et al. (1997)
*Petroselinum hortense*cells	UV		Kreuzaler et al. (1983)
Pinus sylvestris (Scots pine)	UV-B	Phenolic compounds, flavonoids, catechin	Schnitzler et al. (1996)
Picea abies (Norway spruce)	*Ceratocystis polonica*		Nagy et al. (2004)
	Ophiostoma polonicum and wounding	Catechin	Brignolas et al. (1995)
Secale cereale	UV		Haussuehl et al. (1996)
Hordeum vulgare (Barley)	*Blumeria graminis*		Christensen et al. (1998)
	Erysiphe graminis		
	UV		
Medicago truncatula (Alfalfa)	*Glomus versiforme*	Isoflavonoid	Harrison and Dixon (1993)

Antirrhinum majus (Snapdragon)	*Erwinia chrysanthemi*		Junghans et al. (1993)
	Rhizobium meliloti		
	$CuCl_2$		
	Wounding		
	Phoma medicaginis		
	Colletotrichum lindemuthianum		Dalkin et al. (1990)
	UV		Lipphardt et al. (1988)
			Staiger et al. (1989)
Glycine max (Soybean)	*Pseudomonas syringae* pv *glycinea*		Dhawale et al. (1989)
	Phytophthora megasperma f. sp.*Glycinea*		
Picea glauca (White Spruce)	Wounding, JA, MeJ		Richard et al. (2000)
Daucus carota (Carrot cell)	UV, *Pythium aphanidermatum*	Anthocyanin	Gläßgen et al. (1998)
Brassica rapa (Turnip)	UV	Anthocyanin	Zhou et al. (2007)
Sorghum bicolor (Sorghum mesocotyl, juvenile sorghum tissues)	*Colletotrichum graminicola*	3-Deoxyanthocyanidins (apigeninidin; luteolinidin)	Lue et al. (1989), Nicholson et al. (1987)
	Helminthosporium maydis		

Chapter 2

Genome wide mapping of anti-fungal phytoalexins in sorghum (*Sorghum bicolor* (L.) Moench)

Abstract

Colletotrichum sublineola is an aggressive hemi-biotrophic fungal pathogen that causes anthracnose leaf blight and stem rot in sorghum. In severe conditions, it cause substantial yield loss up to 45%. Often major gene resistance fails due to varying degree of virulent pathotypes across sorghum growing locations and availability of other host species for pathogen survival. One mechanism to control anthracnose leaf blight in sorghum is to identify genetic factors associated with anti-fungal production. It is known that *Colletotrichum sublineola* induce the production of anti-fungal compounds known as 3-deoxyanthocyanidins (3-DAs) up on the fungal ingress in sorghum. These compounds are initially colorless and move to the site of infection. After 36 hours, they turn into brick red color and kill the invading fungus and the cell. This mechanism prevents the fungal pathogen spread from one area to other, thus effectively preventing the disease spread. Our goal of this investigation was to identify the genetic factors associated with 3-DA bio-synthesis and regulation in sorghum. A genome wide association study (GWAS) was performed using the sorghum association panel (SAP), ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) mini-core (MNC) panel, and nested association mapping (NAM) panel. Community resources of 265,487 SNP (Single Nucleotide Polymorphism) markers for SAP and MNC panel, and 5108 SNP markers for NAM panel were used in this study. Three biological replicates of SAP and MNC leaf samples were collected at V_7 stage for fungus infection, whereas 21 days seedlings were used for NAM population screening for anthracnose disease. The quantitative measurement of 3-DAs was carried out using spectrophotometer (λ_{max} 480 nm) and

confirmation of a sub-set by HPLC (High Performance Liquid Chromatography) with known 3-DA standards. Significant variability for 3-DAs among the genotypes observed (0.02 (IS 1219) to 2.99 (IS 11619)) with a mean of 0.52 (abs/mg). GWAS identified 11 candidate genes (*p*-value ≤7.73E-06) for 3-DA bio-synthesis and regulation in sorghum which include Sb04g013160 (Guanylate-binding protein), Sb01g044000 (Sugar transporter), Sb02g034830 (BCS1 AAA type ATPase), Sb04g010280 (MYB-like DNA-binding domain), Sb09g004660 (Peroxidase) and others. Our results provide basis for anthracnose disease improvement through marker assisted selection and genomic selection for plant secondary metabolites that act as fungicides in sorghum.

Keywords: 3-deoxyanthocyanidins, Anthracnose leaf blight, *Colletotrichum sublineola*, GWAS, Phytoalexins, SNPs.

Abbreviations

3-DAs	- 3-deoxyanthocyanidins
ALB	- Anthracnose Leaf Blight
CMLM	- Compressed Mixed Linear Model
DPI	- Days Post Infection
GAPIT	- Genomic Association and Prediction Integrated Tool
GBP	- Guanylate Binding Protein
GV	- Genotypic Variation
GWAS	- Genome Wide Association Study
HPLC	- High Performance Liquid Chromatography
HR	- Hypersensitive Response
ICRISAT	- International Crops for the Semi-Arid Tropics
LD	- Linkage disequilibrium
LRR	- Leucine Rich Repeats
MNC	- ICRISAT Mini-core Collection
MLM	- Mixed Linear Model
NAM	- Nested Association Mapping
PCA	- Principal Component Analysis
PCD	- Programmed Cell Death
PV	- Phenotypic Variation
Q-Q plot	- Quantile-Quantile plot
RILs	- Recombinant Inbred Lines
RLK	- Receptor like Kinases
SAP	- Sorghum Association Panel
SNPs	- Single Nucleotide Polymorphisms
SUPER	- Settlement of MLM under Progressively Exclusive Relationship

Introduction

Sorghum is the fifth most important cereal crop and a dietary staple for more than 500 million people living in semi-arid tropics (Reddy et al., 2012). Anthracnose is a major fungal disease in sorghum that causes yield loss up to 50% under favorable environmental conditions (Thakur and Mathur, 2000). Anthracnose fungus can over winter in soil and plant debris. This overwintered inoculum is sought to cause disease next year. *Colletotrichum sublineola*, a pathogen of sorghum anthracnose and *Colletotrichum graminicola*, a pathogen of maize anthracnose has been shown to be interchangeable between the two-plant species (Vernard and Villancourt, 2007; Wheeler et al. 1974). There is no major genic resistance reported for anthracnose in sorghum. Often resistant lines in one location fail in another location, because of the aggressiveness of different strains of fungus. Hence, the discovery and validation of putative genes or genic regions conferring resistance to anthracnose is of paramount importance.

Plants combat pathogens through various mechanisms including: pre-formed structures (cuticle, waxy outer layers) and compounds (hormones), post inducible compounds (stress inducible secondary metabolites), PAMP- triggered immunity (pathogen associated molecular patterns) and effector triggered immunity (ETI). Sorghum is known to synthesis flavonoid class anti-fungal compounds known as 3-deoxyanthocyanidins (3-DAs) upon anthracnose fungus invasion (Nicholson et al. 1987; Snyder and Nicholson, 1990). These 3-DAs are known to exhibit fungi-toxicity effect towards anthracnose pathogen, *Colletotrichum sublineola* (Snyder and Nicholson, 1990; Nielsen et al., 2004; Ibraheem et al., 2010; Ibraheem et al., 2012). Initially 3-DAs are synthesized in inclusions (vesicles)

within the cell under attack and later on they move in the cell to the attempted site of fungal penetration and release their contents into the cytoplasm (Snyder and Nicholson, 1990; Nielsen et al., 2004). The compounds released at the primary site of infection kill the penetrating fungal hypha as well as the cell that produce the compounds (Nielsen et al., 2004). The major 3-DA compounds in sorghum are luteolinidin (λ_{max} 500nm) and apigeninidin (λ_{max} 480nm), and their methyl and acyl derivatives (Awika and Rooney, 2004).

Although the 3-DAs have been implicated in resistance against anthracnose, their genetic and bio-chemical pathway information is incomplete. Previous research in our laboratory has shown that, sorghum lines that lack a functional *y1* (yellow seed1) gene are defective in their synthesis of 3-DAs (Ibraheem et al., 2010). Secondly, several near-isogenic lines with functional and non-functional *y1* gene were used to demonstrate that anthracnose resistance requires a functional *y1* gene (Ibraheem et al., 2010). Based on the above-mentioned background and preliminary results from our previous work, following two questions remain unanswered. Question 1. What's the genetics of 3-DA dependent sorghum-anthracnose resistance? Question 2. What are all the other putative genes involved in 3-DA biosynthesis other than *y1* in a whole genome level? To address the above questions, genome wide association mapping approach has been adopted using three community sorghum mapping panels such as sorghum association panel (SAP), ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) mini-core (MNC), and nested association mapping (NAM) panel to map the putative genes responsible for 3-DA biosynthesis in sorghum. Association mapping can be conducted by genotyping all

individuals with tens of thousands of single nucleotide polymorphisms (SNPs) instead of focusing on candidate genes or regions. Genome-wide association studies (GWAS) have been extensively conducted to dissect the genetic causes of complex human diseases for many years. In recent years GWAS gaining much attention in crop improvement due to the availability of large-scale SNP markers by the advent of low-cost sequencing methodologies.

GWAS offers high allelic diversity and high recombination rate due to the deployment of many markers across the genome and also from the large population size. Successful GWAS studies in sorghum identified genomic regions associated with trait of interest for agroclimatic traits (Morris et al., 2013), adaptive traits (Lasky et al. 2015), grain quality (Sukumaran et al., 2012), seed size (Zhang et al., 2015), grain polyphenol concentrations (Rhodes et al., 2015), Fe and Zn density, stalk rot resistance (Adeyanju et al., 2015), and aluminum tolerance. However, we need to pay critical attention to GWAS approach while working with diverse populations, they may have different maturity groups, photoperiod sensitivity, and flowering time. So, appropriate field design, blocking, data analysis, and interpretation of results all required for getting a good phenotypic data, good marker-trait associations and also to avoid false-positive associations for the traits of interest. Therefore, a GWAS study was undertaken to validate the role of 3-DAs in anthracnose resistance and to identify genes that play a role in 3-DA induction during sorghum-anthracnose resistance response.

Materials and Methods

Plant materials

We used three community mapping populations in our experiment such as SAP (377), MNC (242), and NAM RIL family of SC 265 x RTx 430 (220) along with 3 near-isogenic lines (*Y1-rr3, y1-ww2* and *y1-ww4*) that differ in their ability to produce 3-DAs (Chopra et al., 2002; Ibraheem et al., 2010). The SAP panel consists of 377 global accessions, which are converted to US climatic adaptations to facilitate phenotypic characterization in temperate regions, this collection assembled with accessions from all across the globe with historical importance in sorghum breeding (Casa et al., 2008). This whole collection was characterized with 47 simple sequence repeat (SSR) markers. The ICRISAT mini-core collection consists of 242 accessions captured from 2247 sorghum core collections with phenotypic data measured for 11 qualitative traits and 10 quantitative traits (Upadhyaya et al., 2009). This mini-core collection is not converted to temperate adaptations. So, most of them do not flower due to photoperiod sensitivity in temperate regions. Both mapping panels are diverse geographically and racially. Because of the poor seed supply, we did not perform a disease resistance assay in these study population. NAM panel consists of ten RILs with the common male parent, RTx 430. We have used only the SC 265 × RTx 430 RIL family for our disease assay, which consists of 220 lines.

Seedling disease screening assay

We have used a set of near-isogenic lines which differ in their ability to synthesize 3-DAs (*Y1-rr3, y1-ww2* and *y1-ww4*) along with NAM parental lines, and a NAM RIL family, SC 265 × RTx 430 (220 accessions) for seedling disease screening assay. This experiment was

carried out in controlled conditions in a greenhouse. Twenty-one day old sorghum seedlings were inoculated using *Colleotrichum sublineola* (strain CgSl1) with a concentration of 10^6 spores ml^{-1}. The night and day cycles were maintained inside the greenhouse as 14 hrs of dark and 10 hrs of light, respectively. The humidity and temperature of the chamber were controlled by using a mist humidifier (SU-4010G, Sunpentown, China). Day time humidity was kept at 75% with 30^{o}C temperature, whereas night time humidity of 90-95% with 25^{o}C was maintained for 11 days to allow optimal infection conditions. At the end of 11th day, sampling was done for 3-DA quantification and the disease rating was done using the scale of 1-5, where 1 = no visible lesions; 2 = < 5% leaf area with lesions; 3 = 6-20% leaf area with lesions; 4 = 21-40% leaf area with lesions; and 5 = > 40% leaf area with lesions (Pande et al., 1994). Then, the samples were freeze dried and ground to fine powder to extract 3-DAs using acidified-methanol solvent (99.9% methanol with 0.1% Hcl).

Detached leaf assay

We have used field grown V_7 stage leaf samples in triplicates for infection and quantification (vegetative stage 7 – approximately 50 days old) of fungal induced 3-DAs from SAP and ICRISAT mini-core panels. Disease infection was performed using anthracnose fungal spores of *Colletotrichum sublineola* (strain# CgSl1) at the concentration of 10^6 spores per ml. Infected leaflets were placed in a water-agar medium under constant light for 3 to 4 days. Samples were collected three or four days after infection to quantify the fungal-induced 3-DAs using spectrophotometer (Cytation 5,

BioTek, USA) and the sub-set of these samples were validated using high-performance liquid chromatography (Prominence UFLC, Shimadzu, Japan).

Spectrophotometry and HPLC

To detect the presence of 3-deoxyanthocyanidins, 11 dpi (days post infection) samples were incubated in 99.9% methanol with 0.1% HCl for 48 hrs at -20^{o} C. The clear extracts were collected after centrifugation at 14,000 rpm for 15 min. Ether was added to remove the chlorophyll content in the extract. Total anthocyanidin content was quantified spectrophotometrically at 480 nm using a Cytation 5 spectrophotometer (BioTek, USA), and the concentration was expressed as micrograms using the luteolinidin and apigeninidin standards (ChromaDex, Inc, USA).

For HPLC analysis of compounds, *C. sublineolum* inoculated field grown V_7 plants and greenhouse grown 21 days old seedlings lesions were dissected and placed in 1 ml of 99.9% HPLC-grade methanol with 0.1% Hcl, and compounds were allowed to leach at -20^{o} C for 48 hrs in the dark. Extracts were collected by centrifugation at 14,000 rpm for 14 min and filtered through 0.45-mm acrodisc LC 13-mm syringe filters (Gelman laboratory, Ann Arbor, MI), and evaporated down to 100 ml. To separate compounds, two reversed phase C18 columns connected in tandem (Supelco, Bellefonte, PA) were used, and a 50-ml sample was applied to the column fitted on a prominence HPLC separator with a PDA detector system. Spectral measurements were taken over a wavelength range of 230–550 nm, which is known to detect all flavonoid compounds (Grotewold et al., 1998). Detection of 3-deoxyanthocyanidins was carried out at 480 nm as described previously (Snyder and

Nicholson 1990; Lo and Nicholson 1998). Pure luteolinidin and apigeninidin were dissolved in HPLC-grade methanol and used as standards. Compounds were separated on a 1 × 150-mm beta basic C18 column with a solvent gradient of water containing 0.15% formic acid (solvent A) and methanol (solvent B) with a flow rate of 0.05 ml min^{-1} delivered using a pre-injection split. The program of solvent composition was established as the following: 0–5 min (99% solvent A; 1% solvent B) followed by a linear gradient to 50% solvent A, 50% solvent B at 12 min, another linear gradient to 100% of solvent B at 20 min, with a hold at 100% solvent B until 25 min.

Association analyses

GWAS was performed using models implemented in the R-package GAPIT (Genome Association and Prediction Integration Tool) (Lipka et al., 2012) with 265,487 SNP markers (Morris et al., 2013; Lasky et al., 2015) for SAP and MNC, whereas 5108 SNP markers for the NAM panel cross, SC 265 × RTx 430. Both phenotypic and marker data sets were fed into the GAPIT package to determine marker-trait associations. Association scans were performed using the CMLM (compressed mixed linear model) in combination with SUPER (Settlement of MLM under Progressively Exclusive Relationship) model to determine marker-trait associations. The compressed mixed linear model (CMLM) involves genetic marker–based kinship matrix modeling of random effects used jointly with population structure estimated by principal components analysis (PCA) to model fixed effects (Zhang et al., 2010). SUPER model (Wang et al., 2014) extracts a small subset of SNPs from CMLM and uses them in FaST-LMM method. This method increases statistical power and is computationally advantageous. The compression level and optimal number

of principal components that adequately explain population structure were previously determined by the GAPIT (Lipka et al., 2012). Before presenting GWAS results, the model fit was compared by examining the quantile-quantile (Q-Q) plots (Fig. 3 and 8), and the CMLM with SUPER was selected as the model with superior fit. As several studies have suggested varying number of linkage disequilibrium (LD) blocks in sorghum, LD range of 2–5 kb was used for the analysis (Morris et al., 2013). Genes potentially linked to any significantly associated SNP were identified by scanning version 2.1 of the *S. bicolor* genome (Goodstein et al., 2012). Gene function was determined using the panther classification system (Mi et al., 2013), the European bioinformatic institute's PFAM identification (Finn et al., 2014), and gramene biomart (www.gramene.org), and phytozome v.10 (www.phytozome.net).

Statistical Analysis

The total measurements included two years (2014 and 2015) field × a year (2017) green house for NAM lines × a treatment (quantification of fungal induced 3-DAs) × 722 lines × three biological replicates for a total of 4332 assays. *P*-values of < 0.05 were considered statistically significant. Statistical analysis was conducted with R, a statistical computing language and environment (R Core Team 2015). Data obtained were expressed as mean ± standard deviation.

Results

Quantitative variation of sorghum anti-fungal compounds in lines carrying *y1* alleles

In sorghum it was reported that 3-deoxyanthocyanidin (3-DA) production requires functional allele of *yellow seed1* (*y1*) gene (Ibraheem et al., 2010) and the *y1* gene was transformed to control anthracnose leaf blight (ALB) in maize (Ibraheem et al., 2015). Thus, we have quantified the anthracnose fungal (*Colletotrichum sublineola*) induced 3-DAs in lines carrying functional (*Y1-rr3*) and non-functional (*y1-ww2, y1-ww4*) *y1* alleles along with two standard sorghum lines (H 112 (Moderately resistant to ALB) and IS 19153 (ALB resistant check). We have observed early induction of 3-DAs in the lines carrying the functional *y1* gene (*Y1-rr3*) and they were revealed fewer disease symptoms on leaves compared to the lines that do not carry functional *y1* alleles (*y1-ww2* and *y1-ww4*) (Fig. 5 and Table 2). HPLC analysis revealed that the ALB resistant lines substantially produced 3-DAs after fungal ingress, they shown to accumulate both 3-DA types such as luteolinidin and apigeninidin (Fig. 5 and Table 2). Whereas, the moderately resistant line H 112 accumulated only apigeninidin in trace amounts (Table 2). On the other hand, the susceptible lines that do not carry functional *y1* alleles (*y1-ww2* and *y1-ww4*) did not accumulate any of these 3-DA types and they were shown to have more disease symptoms on leaves after fungal infection (Fig. 5 and Table 2). Our results supported the *y1* dependent 3-DA induction and resistant response to anthracnose imparted via 3-DAs in sorghum (Ibraheem et al., 2010). We only know about the *y1* gene regulated 3-DA biosynthesis in sorghum, we want to understand the other genes that might play a vital role in 3-DA bio-synthesis and transport. Hence, screening of large set of sorghum accessions were carried out to quantify the anthracnose fungal induced 3-DA variations and to map the genes responsible for its bio-synthesis and regulation in sorghum.

3-DA concentrations in global sorghum accessions by race, seed color and geographic origin

Phenotypic variation for anti-fungal 3-DA concentrations were determined using a diverse association panel of 502 accessions which consist of SAP (Casa et al., 2008) and ICRISAT mini-core (Upadhyaya et al., 2009). The mean total 3-DA concentration was 0.52 (abs/mg) with a range between 0.02 and 2.99 (abs/mg). More than 200 genotypes accumulated 3-DAs above the mean level of 0.52 (abs/mg). The standard deviation between the genotypes was 0.38 (abs/mg) with a variance of 0.14 (abs/mg). Next, we have quantified the level of 3-DAs accumulation between races and intermediate races of sorghum. Durra-bicolor (0.72 abs/mg), Kafir-durra (0.62 abs/mg) and Caudatum (0.56 abs/mg) had the highest mean concentration of 3-DAs (Table 3). Moderate level of 3-DAs was observed in Kafir-bicolor (0.50 abs/mg), Kafir-caudatum (0.49 abs/mg), Durra-caudatum (0.48 abs/mg), Guinea (0.47 abs/mg), Durra (0.47 abs/mg), Guinea-caudatum (0.47 abs/mg), and Kafir (0.45 abs/mg). Whereas, races like Guinea-durra (0.36 abs/mg) and Guinea-kafir (0.32 abs/mg) had the lowest mean level of 3-DAs (Table 3). The genotypes from Australia (0.64 abs/mg), Africa (0.54 abs/mg) and America (0.54 abs/mg) had the highest mean level of 3-DAs (Table 2). 56% of the accessions used in this study came from African continent followed by Asia (18%) and America (17%). The genotypes IS 11619 (2.99 abs/mg), IS 9108 (2.92 abs/mg), and IS 11919 (2.20 abs/mg) had the highest 3-DA levels; whereas, the genotypes IS 1219 (0.02 abs/mg), PI 533869 (0.07 abs/mg), and PI 656093 (0.07 abs/mg) had the lowest level of 3-DAs (Table 4). Strikingly most of the highest 3-DA producing lines are colored (Reddish brown) as well as most of the lowest 3-DA producing lines are non-colored (White). However, we did not observe any significant mean difference between

colored and non-colored sorghum accessions with respect to 3-DA accumulation after fungal infection (Table 3). We have selected a handful of high and low 3-DA producing lines after *C. sublineola* infection for HPLC sub-set validation. HPLC chromatograms revealed the high 3-DA producing line SC 213 (PI 576391) accumulated both 3-DA types such as luteolinidin (R_T 18.87 min) and apigeninidin (R_T 19.37 min). Whereas the low 3-DA producing line SC 135 (PI 534148) did not accumulate either luteolinidin or apigeninidin (Fig. 6).

3-DA variation in NAM parental lines and in a SC 265 × RTx 430 RIL family

Ten NAM parental lines along with the common parent RTx 430 and the SC 265 × RTx 430 RIL family (220 accessions) were used to evaluate the 3-DA induction upon fungal infection under greenhouse condition. We have identified high (SC 1345 and Seagolone), medium (SC 35, SC 265, SC 283, P 898012) and low (RTx 430, Ajabsido, SC 971, SC 1103) 3-DA producers of NAM parental lines. The NAM RIL cross, SC 265 × RTx 430 exhibited significant variations for 3-DA induction upon fungal infection (Table 5 and fig. 6). The line PR1314_2772 exhibited high level of 3-DA induction (65.7 apigeninidin g^{-1}), and the line PR1314_2804 shown low level of 3-DA induction (1.9 apigeninidin g^{-1}) with an overall mean of 13.3 apigeninidin per g sorghum leaf tissue (Table 5). Thus, this NAM RIL panel was also used to map the fungal induced 3-DA genes in sorghum. The Q-Q plot exhibited significant deviation from the expected line (Fig. 11).

Genome wide association studies

To investigate the genetic factors associated with 3-DA bio-synthesis and regulation in sorghum, we have conducted a genome wide association study using 265,487 SNP markers (Morris et al., 2013) spread across the sorghum genome that account for 27,412 annotated genes in sorghum (https://plants.ensembl.org/Sorghum_bicolor) with 502 global sorghum germplasm accessions which are diverse racially and geographically (Casa et al., 2008; Upadhyaya et al., 2009) along with 5108 SNP markers for 220 accessions of SC 265 x RTx 430 RIL family. GWAS was performed using the GAPIT software (Lipka et al., 2012), which runs in the background of R program (R core team, 2015). The Q-Q plot of *P*-values showed significant association between the trait of interest and SNPs (Fig. 7 and 11). We have observed moderate level of heritability (48.6%) for 3-DAs in sorghum (Fig. 8). The CMLM (Zhang et al., 2010) with SUPER (Wang et al., 2014) in GAPIT program revealed 11 significant marker-trait associations (*p*-value ≤7.73E-06) in SAP and MNC panel with distinct association peaks on chromosomes 3, 4, 6 and 9 (Fig. 9). The peak on chromosome 3 was between 22 and 22.02 Mb and co-localized with a gene Sb04g013160, guanylate binding family protein (Table 5), involved in imparting immunity against microbial and viral pathogens (Kim et al., 2016). The peak on chromosome 1 was between 67 and 67.12 Mb, and co-localized with a gene Sb01g044000, sugar transporter, which plays a major role in imparting resistance against bacterial pathogens by allocating sugar metabolites within the host effectively to manage the availability of sugars in extra cellular surfaces for microbes (Yamada et al., 2016). The peak on chromosome 1 was between 36 and 36.12 Mb, and co-localized with a gene Sb01g004510, acyl-activating enzyme (AMP-binding enzyme), which involved in the regulation and synthesis of major secondary metabolites such as lignins, lignans, fatty acids, flavonoids and glucosinolates (Shockey and Browse,

2011). It was estimated to be involved in 4% of all enzymatic reactions, either as a co-factor or linked to particular substrates (Begley et al., 2011). The other probable candidates are listed on table 4, which includes Sb04g010280 (MYB-like DNA binding domain), Sb09g004660 (Peroxidase/luteolin triglucuronide degradation), Sb09g018180 (Glycosyl hydrolase family 1), Sb10g010580 (Putative expansin-B14 precursor), Sb07g005820 (SHR5-receptor-like kinase/ LRR), Sb01g040200 (ASC1-like protein/ LAG1), and Sb07g006550 (Rab-GTPase-TBC domain). Whereas, NAM RILs did not find any significant marker-trait associations (Fig. 12). However, the candidate gene annotation for the top SNPs in the GWAS results of SC 265 × RTx 430 RIL family revealed some common candidates across panels (Table 7). The common candidate identified across the panels are, BCS1 AAA-type ATPase (Sb03g027200), helix-loop-helix DNA-binding domain (Sb03g008290), nonspecific lipid-transfer protein (Sb01g017700), and serine/threonine protein kinase (Sb03g008160).

Discussion

GWAS identified 11 candidate genes for 3-DAs in sorghum. Among the candidates identified, guanylate-binding protein (Sb04g013160) is the most probable candidate gene (Table 5). 3-DAs known to accumulate rapidly around the site of fungal ingress and eventually kill the invading fungus as well as the host cell (Synder and Nicholson, 1990). It is synthesized in endoplasmic reticulum and transported to plasma membrane then to epidermis via golgi body associated vesicles (Synder and Nicholson, 1990). Guanylate binding protein (GBP) family is a family of GTPases. GBP known to involved in imparting resistant against viral and bacterial pathogens. It is also known to involved in signal

transduction especially vesicle mediated trafficking of metabolites (Vestal and Jeyaratnam, 2011). GBP might be the probable candidate for vesicle mediated 3-DA secretion and regulation in sorghum. Another probable candidate is sugar transporter (Sb01g044000). Sugar transporters are known to involved in allocation of sugar efflux between host and the invading microbes, it alters the sugar availability to the external cell membranes for microbes, thus impart resistance to the host for various pathogens (Chen et al., 2010; Morkunas and Ratajczak, 2014). We have identified sugar transporter (Sb01g044000) as the second most probable hit in our GWAS study. The third candidate gene was Sb02g034830, a BCS1 AAA-type ATPase. It is a known mitochondrial chaperone. Followed by Sb01g004510 gene was identified as a candidate for 3-DA bio-synthesis and secretion. Sb01g004510 gene encodes for acyl-activating enzyme complex (AMP-binding enzyme), which involved in the regulation and synthesis of major secondary metabolites such as lignins, lignans, fatty acids, flavonoids and glucosinolates (Shockey and Browse, 2011). Four percent of all enzymatic reactions are involved with acyl-activating enzymes either as a co-factor or linked to particular substrates (Begley et al., 2011).

P1 MYB in maize is the known phenylpropanoid pathway regulator (Grotewold et al., 1994; Hichri et al., 2011). Several Mybs have been shown to be involved in a variety of biological functions like biotic and abiotic stress (Segarra et al., 2009; Lippold et al., 2009), hormone responses (Urao et al., 1993), plant defense reactions (Yang and Klessig 1996; Liu et al., 2008), regulation of circadian clock (Schaffer et al., 1998) and telomeric DNA-binding protein (Yu et al., 2000). We have identified a Myb-like DNA-binding domain gene (Sb04g010280) as our significant candidate for 3-DA biosynthesis in

sorghum. Phenylpropanoid pathway begins with phenylalanine and goes into various steps of condensation to produce naringenin (colorless), a rate limiting step in the pathway. From naringenin, the pathway branches into produce phlobaphenes, 3-deoxyflavonoids, and anthocyanins. Phlobaphenes are red flavonoid pigments derived by polymerization of flavan-4-ols (Grotewold et al., 1994), which is under the control of R2R3-MYB protein encoded by *pericarp color1* (*p1*). In sorghum, phlobaphene accumulation in the pericarp, glume, and leaf is under the control of *y1* (*yellow seed1*), a gene that encodes an R2R3-MYB able to regulate the expression of the CHS (chalcone synthase), CHI (chalcone synthase isomerase), DFR (dihydroxyflavonol reductase), and F3'H (flavanone 3' hydroxylase) required for the synthesis of flavan-4-ols, which is a precursor for phlobaphenes and 3-deoxyanthocyanidins (Boddu et al., 2005; Ibraheem et al., 2010). In maize, *p1* an R2R3-MYB, which also activates the expression of CHS, CHI, and DFR and regulates the synthesis of phlobaphenes (Grotewold et al., 1994). The *y1* gene in sorghum is the orthologue of *p1* gene in maize, both shown to have 99% sequence similarity with each other (Boddu et al., 2006). *Y1* gene in sorghum was shown to bring resistance to anthracnose leaf blight (Ibraheem et al., 2010), so this *y1* gene was transferred to maize to control the anthracnose leaf blight disease (Ibraheem et al., 2015). Therefore, the gene Sb04g010280 which encode MYB domain might be the potential candidate for 3-DA biosynthesis in sorghum. This gene could be exploited for altering the phenylpropanoid pathway to gain valuable insights to produce crops for better biotic and abiotic stress tolerance via marker assisted breeding and genomic selection.

GWAS analysis revealed the gene Sb09g004660, which encodes peroxidase (luteolin triglucuronide degradation) as a probable candidate for 3-DA regulation in sorghum. Peroxidase is a known free radical scavenger, which detoxify the hydrogen peroxide (H_20_2) into water (H_20). The expression of peroxidase genes are regulated in response to biotic and abiotic stresses and the sorghum peroxidase gene, Sb09g004660 (gene symbol: LOC8071200; locus tag: SORBI_3009G055300) is involved in the oxidation of toxic reductants (Hiraga et al., 2001), biosynthesis and degradation of lignin (Sasaki et al., 2004; Zhao et al., 2013; Koutaniemi et al., 2015), suberization (Espelie et al., 1986; Quiroga et al., 2000), auxin catabolism (Jansen et al., 2001), response to environmental stresses such as wounding (Lagrimini and Rothstein, 1987; Lagrimini et al. 1990), pathogen attack (Kristensen et al., 1999; Choi et al., 2007; Wally and Punja, 2010; Daudi et al., 2012), temperature stress (Kawakami et al., 2002; Song et al., 2005), high light stress (Yoshimura et al., 2000; Danna et al., 2003), salt stress (Shi et al., 2001), drought stress (D'Arcy-Lameta et al., 2006), and heavy metal toxicity (Sharma and Dubey, 2007). It was reported that peroxidase mediate the degradation of luteolin triglucuronide from vacuole to cell wall whenever plant experiences the stress (Anhalt and Weissenbock, 1992). Basically, peroxidases play a role in transporting the flavonoid metabolites from storage compartment (vacuole) to the cell in need. These evidences suggest that peroxidases are probable candidate for 3-DA transport from vacuole to epidermis whenever plant cell perceive stress signals from external stimuli.

Next potential candidate from our GWAS analysis was Sb07g005820, which encodes SHR5-receptor-like kinase (leucine rich repeats (LRR)) in sorghum. Plants have a large

family of receptor-like kinases (RLKs) that have been implicated in mechanisms of perception and transduction of extracellular signals into the cell (Shiu and Bleecker, 2001). In sorghum, a new class of LRR domain was reported by Hipskind et al. (1996) which participates in the signal transduction mechanism during sorghum- anthracnose fungus interactions. Members of the LRR subfamily were known to play roles in diverse processes related to plant growth/development, stress, defense against pathogens, and symbiosis (Shiu et al., 2004; Vinagre et al., 2006). A novel gene in sugarcane namely SHR5, involved in the association with endophytic nitrogen-fixing bacteria (Vinagre et al., 2006). Sequence analyses suggest that the SHR5 gene encodes a protein that belongs to a subclass of the LRR-RLK protein family. Biruma et al. (2012) identified two loci correspond to NB-LRR domains in sorghum (Sb09g027470 and Sb09g027520), which impart resistant against anthracnose disease in sorghum. Currently those two loci are being utilized in developing anthracnose resistant sorghum cultivars. LRR domains also reported to impart resistance to shoot fly in sorghum (Satish et al., 2012), ascochyta blight in chickpea (Sagi et al., 2017), and anthracnose in common bean (Wu et al., 2017). Hence the identified LRR gene (Sb07g005820) acts as a messenger in perceiving and feeding back the stress response signals from pathogens to host to activate the defense response pathway (induction of R genes or 3-DAs) in sorghum.

Finally, we have identified genes involved in autophagy and hypersensitive response mediated programmed cell death (HR-PCD) such as ASC1 (Sb01g040200) and GTPase activator protein-like/Rab-GTPase-TBC domain (Sb07g006550) as candidates for 3-DA regulation in sorghum. ASC1 is a plant resistance gene (yeast homologue of LAG1

(longevity assurance factor 1)), which prevents the host-specific AAL-toxin induced programmed cell death (PCD) in tomato by synthesizing the sphingolipids sphingolipids (Spassieva et al., 2006; Tsuge et al., 2013). The host specific AAL-toxin is produced upon *Alternaria alterata f. sp. lycopersici* infection, which ultimately causes PCD in tomato. Rab-GTPase was shown to regulate autophagy and promotes HR cell death in response to avirulent bacterial pathogens in Arabidopsis (Kwon et al., 2013). Rapid and localized induction of HR-PCD ensures that pathogen invasion is prohibited further. Rab-GTPases are also involved in metabolite trafficking in post-Golgi traffic to the plasma membrane and vacuoles (Woollard and Moore, 2008). These evidences suggest that 3-DA is involved in PCD pathway and getting transported to the site of infection via Golgi associated vesicles.

Conclusion

GWAS identified 11 candidate genes for 3-DA bio-synthesis, secretion and transport in sorghum. The role of candidate genes was discussed. This study identified novel genes associated with signal transduction pathway, PCD pathway, vesicular trafficking, and sugar (major) transporters as candidates for 3-DAs in sorghum. The genes like LRR (Sb07g005820), MYB (Sb04g010280), ASC1 (Sb01g040200), Rab-GTPase (Sb07g006550), peroxidase (Sb09g004660), and GBPs (Sb04g013160) could be exploited in marker assisted breeding as targets for improving crop tolerance to biotic and abiotic stresses especially for developing anthracnose disease resistant cultivars. Further gene validation pertaining to 3-DA induction upon fungal infection on selected candidates needs to be performed.

Literature Cited

Adeyanju, A., Little, C., Yu, J., and Tesso, T. (2015). Genome-Wide Association Study on Resistance to Stalk Rot Diseases in Grain Sorghum. *G3 (Bethesda)* **5**, 1165-75.

Awika, J. M., and Rooney, L. W. (2004). Sorghum phytochemicals and their potential impact on human health. *Phytochemistry* **65**, 1199-221.

Biruma, M., Martin, T., Fridborg, I., Okori, P., and Dixelius, C. (2012). Two loci in sorghum with NB-LRR encoding genes confer resistance to Colletotrichum sublineolum. *Theor Appl Genet* **124**, 1005-15.

Boddu, J., Jiang, C., Sangar, V., Olson, T., Peterson, T., and Chopra, S. (2006). Comparative structural and functional characterization of sorghum and maize duplications containing orthologous myb transcription regulators of 3-deoxyflavonoid biosynthesis. *Plant Mol Biol* **60**, 185-99.

Boddu, J., Svabek, C., Ibraheem, F., Jones, A. D., and Chopra, S. (2005). Characterization of a deletion allele of a sorghum Myb gene yellow seed1 showing loss of 3-deoxyflavonoids. *Plant Science* **169**, 542-552.

Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., Franks, C. D., and Kresovich, S. (2008). Community Resources and Strategies for Association Mapping in Sorghum. *Crop Science* **48**.

Chen, L. Q., Hou, B. H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X. Q., Guo, W. J., Kim, J. G., Underwood, W., Chaudhuri, B., Chermak, D., Antony, G., White, F. F., Somerville, S. C., Mudgett, M. B., and Frommer, W. B. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**, 527-32.

Chopra, S., Gevens, A., Svabek, C., Wood, K. V., Peterson, T., and Nicholson, R. L. (2002). Excision of the Candystripe1 transposon from a hyper-mutable Y1-cs allele shows that the sorghumY1 gene controls the biosynthesis of both 3-deoxyanthocyanidin phytoalexins and phlobaphene pigments. *Physiological and Molecular Plant Pathology* **60**, 321-330.

Ibraheem, F., Gaffoor, I., and Chopra, S. (2010). Flavonoid phytoalexin-dependent resistance to anthracnose leaf blight requires a functional yellow seed1 in Sorghum bicolor. *Genetics* **184**, 915-26.

Ibraheem, F., Gaffoor, I., Tan, Q., Shyu, C. R., and Chopra, S. (2015). A sorghum MYB transcription factor induces 3-deoxyanthocyanidins and enhances resistance against leaf blights in maize. *Molecules* **20**, 2388-404.

Kim, B. H., Chee, J. D., Bradfield, C. J., Park, E. S., Kumar, P., and MacMicking, J. D. (2016). Interferon-induced guanylate-binding proteins in inflammasome activation and host defense. *Nat Immunol* **17**, 481-9.

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., and Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397-9.

Lo, S. C. C., and R. L. Nicholson. (1998). Reduction of light-induced anthocyanin accumulation in inoculated sorghum mesocotyls: implications for a compensatory role in the defense response. *Plant Physiol* **116**, 979–989.

Morkunas, I., and Ratajczak, L. (2014). The role of sugar signaling in plant defense responses against fungal pathogens. *Acta Physiologiae Plantarum* **36**, 1607-1619.

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., and Kresovich, S. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc Natl Acad Sci U S A* **110**, 453-8.

Nicholson, R. L., S. S. Kollipara, J. R. Vincent, P. C. Lyons and G. Cadena-Gomez. (1987). Phytoalexin synthesis by the sorghum mesocotyl in response to infection by pathogenic and nonpathogenic fungi. Proc. Natl. Acad. Sci. 84, 5520–5524.

Nielsen, K. A., Gotfredsen, C. H., Buch-Pedersen, M. J., Ammitzbøll, H., Mattsson, O., Duus, J. Ø., and Nicholson, R. L. (2004). Inclusions of flavonoid 3-deoxyanthocyanidins in Sorghum bicolor self-organize into spherical structures. *Physiological and Molecular Plant Pathology* **65**, 187-196.

Reddy, B. V. S., Reddy S, Sadananda A.R, Elango Dinakaran, Kumar A.A, Deshpande S.P, Srinivasa Rao, Sharma H.C, Rajan Sharma, Krishnamurthy L, Patil J.V. (2012). Postrainy season sorghum: Constraints and breeding approaches. SAT eJournal 10.

Rhodes, D. H., Hoffmann, L., Jr., Rooney, W. L., Ramu, P., Morris, G. P., and Kresovich, S. (2014). Genome-wide association study of grain polyphenol concentrations in global sorghum [Sorghum bicolor (L.) Moench] germplasm. *J Agric Food Chem* **62**, 10916-27.

Sagi, M. S., Deokar, A. A., and Tar'an, B. (2017). Genetic Analysis of NBS-LRR Gene Family in Chickpea and Their Expression Profiles in Response to Ascochyta Blight Infection. *Front Plant Sci* **8**, 838.

Satish, K., Madhusudhana, R., Padmaja, P. G., Seetharama, N., and Patil, J. V. (2012). Development, genetic mapping of candidate gene-based markers and their significant association with the shoot fly resistance quantitative trait loci in sorghum [Sorghum bicolor (L.) Moench]. *Molecular Breeding* **30**, 1573-1591.

Spassieva, S., Seo, J. G., Jiang, J. C., Bielawski, J., Alvarez-Vasquez, F., Jazwinski, S. M., Hannun, Y. A., and Obeid, L. M. (2006). Necessary role for the Lag1p motif in (dihydro)ceramide synthase activity. *J Biol Chem* **281**, 33931-8.

Sukumaran, S., Xiang, W., Bean, S. R., Pedersen, J. F., Kresovich, S., Tuinstra, M. R., Tesso, T. T., Hamblin, M. T., and Yu, J. (2012). Association Mapping for Grain Quality in a Diverse Sorghum Collection. *The Plant Genome Journal* **5**.

Snyder, B. A., and Nicholson, R.L. (1990). Synthesis of Phytoalexins in Sorghum as a Site-Specific Response to Fungal Ingress. *Science* **248**, 1637-1639.

Styles, E. D., and O. Ceska. (1977). Genetic control of flavonoid synthesis in maize. *Can. J. Genet. and Cytol.* **19**, 289–302.

Thakur, R.P. and K. Mathur, 2000. Anthracnose, in Compendium Sorghum Diseases, R.A. Frederiksen and G.N. Odovody, Editors. APS Press: St. Paul, MN. p. 10-12.

Tsuge, T., Harimoto, Y., Akimitsu, K., Ohtani, K., Kodama, M., Akagi, Y., Egusa, M., Yamamoto, M., and Otani, H. (2013). Host-selective toxins produced by the plant pathogenic fungus Alternaria alternata. *FEMS Microbiol Rev* **37**, 44-66.

Upadhyaya, H. D., Pundir, R. P. S., Dwivedi, S. L., Gowda, C. L. L., Reddy, V. G., and Singh, S. (2009). Developing a Mini Core Collection of Sorghum for Diversified Utilization of Germplasm. *Crop Science* **49**.

Venard, C. and L. Vaillancourt, Penetration and colonization of unwounded maize tissues by the maize anthracnose pathogen Colletotrichum graminicola and the related nonpathogen C. sublineolum. Mycologia, 2007. 99(3): p. 368-377.

Vestal, D. J., and Jeyaratnam, J. A. (2011). The guanylate-binding proteins: emerging insights into the biochemical properties and functions of this family of large interferon-induced guanosine triphosphatase. *J Interferon Cytokine Res* **31**, 89-97.

Vinagre, F., Vargas, C., Schwarcz, K., Cavalcante, J., Nogueira, E. M., Baldani, J. I., Ferreira, P. C., and Hemerly, A. S. (2006). SHR5: a novel plant receptor kinase involved in plant-N2-fixing endophytic bacteria association. *J Exp Bot* **57**, 559-69.

Wheeler, H., D.J. Politis, and C.G. Poneleit., 1974. Pathogenicity, host range, and distribution of Colletotrichum graminicola on corn. *Phytopathology* **64**, 293-296.

Woollard, A. A., and Moore, I. (2008). The functions of Rab GTPases in plant membrane traffic. *Curr Opin Plant Biol* **11**, 610-9.

Wu, J., Zhu, J., Wang, L., and Wang, S. (2017). Genome-Wide Association Study Identifies NBS-LRR-Encoding Genes Related with Anthracnose and Common Bacterial Blight in the Common Bean. *Front Plant Sci* **8**, 1398.

Zhang, D., Li, J., Compton, R. O., Robertson, J., Goff, V. H., Epps, E., Kong, W., Kim, C., and Paterson, A. H. (2015). Comparative Genetics of Seed Size Traits in Divergent Cereal Lineages Represented by Sorghum (Panicoidae) and Rice (Oryzoidae). *G3 (Bethesda)* **5**, 1117-28.

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M., and Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat Genet* **42**, 355-60.

Figure Legends

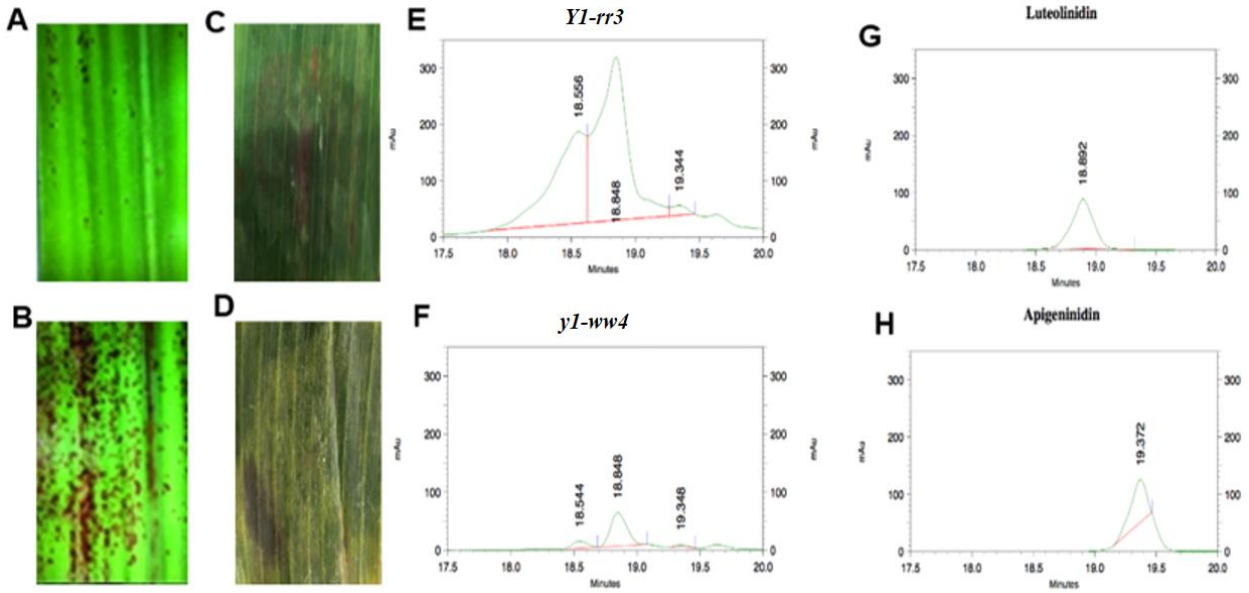

Figure 5. The 3-DA variation in two near-isogenic lines of *Y1-rr3* and *y1-ww4*, differ in functional *y1* (*yellow seed1*) gene. 1A & 1B. Anthracnose disease severity in *Y1-rr3* and *y1-ww4*; 1C & 1D. Anti-fungal compound (3-DAs) synthesizes in response to pathogen infection in *Y1-rr3* and *y1-ww4*; 1E &1F. HPLC chromatograms of *Y1-rr3* and *y1-ww4*; 1G & H. Chromatograms of 3-DA standards luteolinidin and apigeninidin.

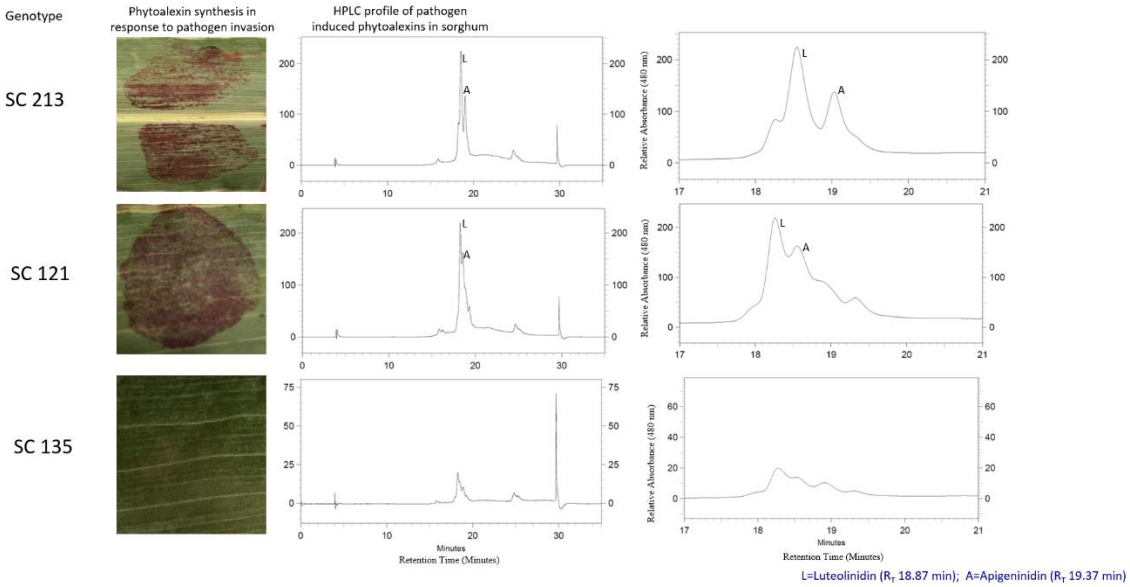

Figure 6. Chromatographic profiles of a high and low 3-DA producing sorghum lines from the sorghum association mapping panel.

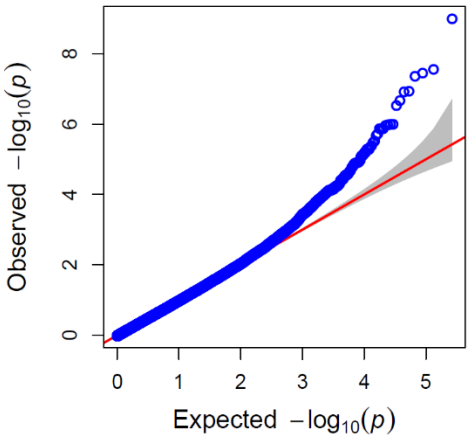

Figure 7. Quantile-quantile (QQ) –plot of P-values for fungal induced 3-DAs in sorghum association mapping and ICRISAT mini-core panel. The Y-axis is the observed negative base 10 logarithm of the P-values, and the X-axis is the expected negative base 10 logarithm of the P-values under the assumption that the P-values follow a uniform (0,1) distribution. The dotted lines show the 95% confidence interval for the QQ-plot under the null hypothesis of no association between the SNP and the trait.

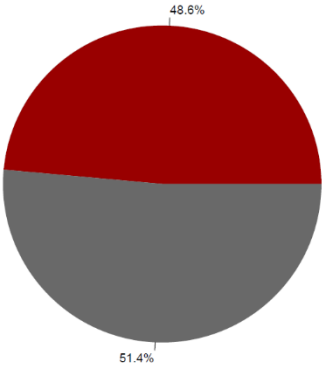

Figure 8. Proportion of genetic and residual variance observed in SAP and ICRISAT mini-core panels for 3-DA's

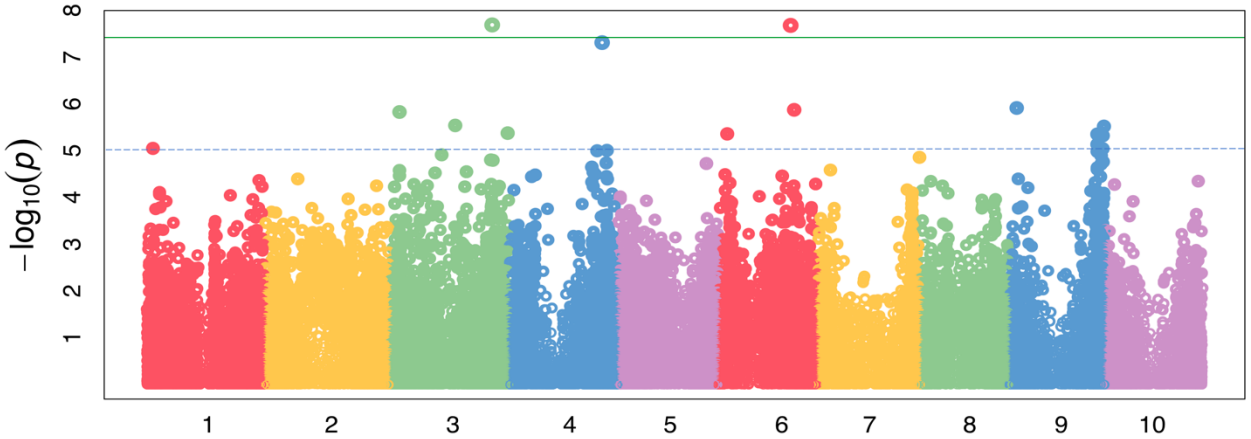

Figure 9. Manhattan plot of association mapping result from a CMLM analysis using 502 accessions for 3-DA content in sorghum. Each point represents a SNP, with the -log10 p-values plotted against the position on each chromosome. The horizontal dashed line represents the genome-wide significance threshold at 5% FDR.

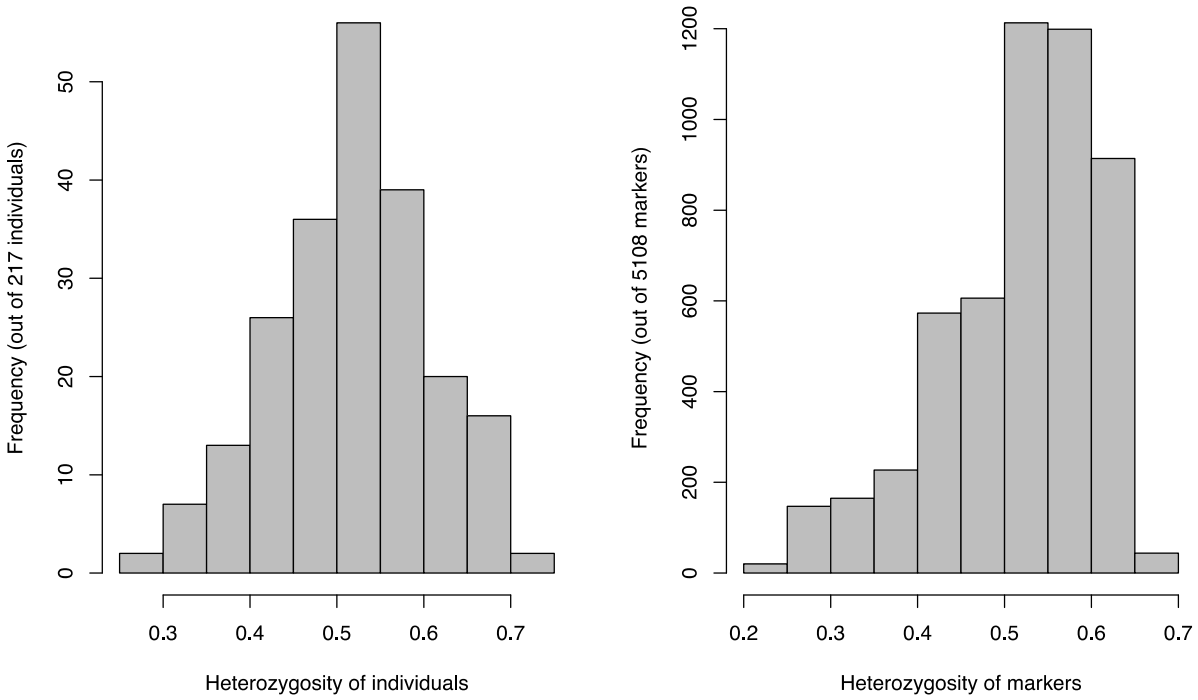

Figure 10. Phenotypic and genotypic (SNP markers) diversity for the NAM RIL population of SC 265 x RTx 430 for fungal induced 3-deooxyanthocyanidins.

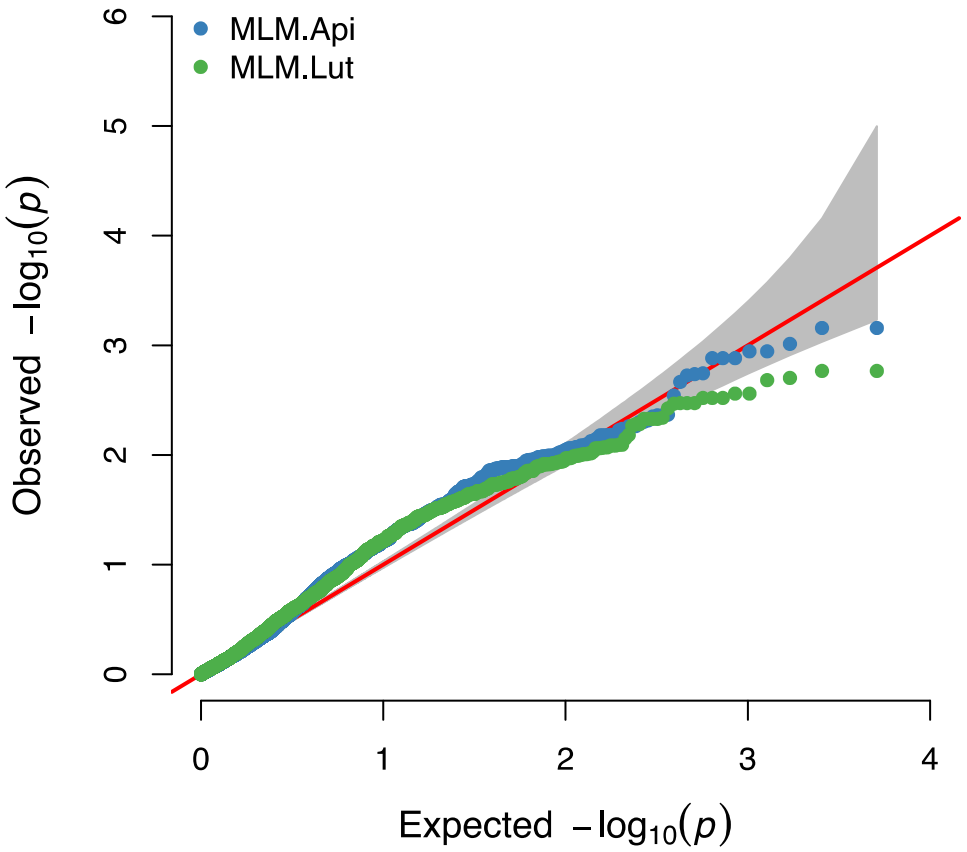

Figure 11. Q-Q plot of total 3-deoxyanthocyanidins in the NAM RIL population of SC 265 x RTx 430.

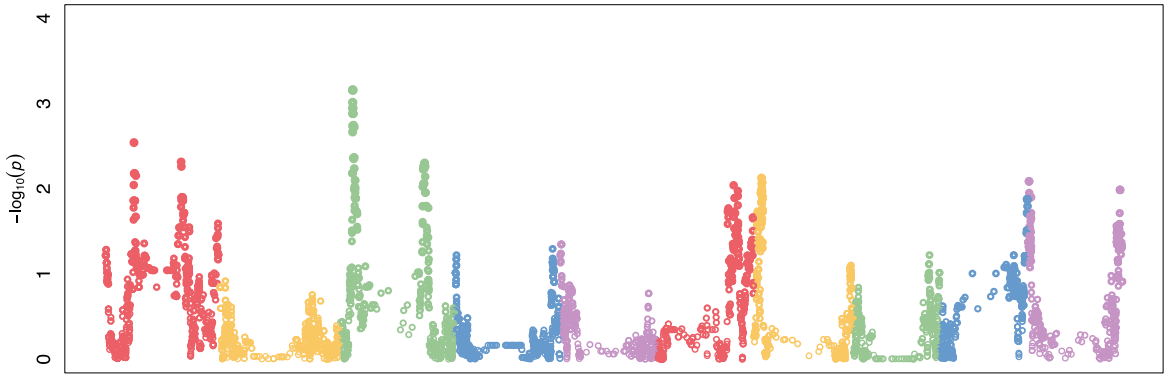

Figure 12. Manhattan plot of association mapping result from a MLM analysis using 220 NAM RILs of SC 265 x RTx 430 for total 3-DA content in sorghum. Each point represents a SNP, with the -log10 p-values plotted against the position on each chromosome. The horizontal dashed line represents the genome-wide significance threshold at 5% FDR.

Tables

Table 2. 3-DA variation in standard sorghum lines

Genotype	Disease rating (1-5 scale)	Category	3-DA type
Y1-rr3	2.2	Resistant	Luteolinidin + Apigeninidin
IS 19153	2.7	Resistant	Luteolinidin + Apigeninidin
H 112	3.0	Moderately resistant	Apigeninidin
y1-ww2	3.5	Moderately susceptible	Did not produce 3-DAs
y1-ww4	3.8	Susceptible	Did not produce 3-DAs

Table 3. Genetic diversity of fungal induced 3-deoxyanthocyanidins in sorghum by race, seed color and geographic origin

Race, Seed color and Origin	n	Total 3-DAs (λ_{max} 480nm/mg tissue)	
		Range (low to high)	Mean
Race			
Bicolor	29	0.08 - 0.91	0.44
Caudatum	108	0.07 - 2.92	0.56
Caudatum-bicolor	35	0.07 - 1.23	0.45
Durra	51	0.08 - 1.39	0.47
Durra-bicolor	27	0.09 - 2.99	0.72
Durra-caudatum	22	0.12 - 1.86	0.48
Guinea	48	0.07 - 1.46	0.47
Guinea-bicolor	5	0.02 - 0.95	0.40
Guinea-caudatum	37	0.14 - 2.00	0.47
Guinea-durra	5	0.21 - 0.48	0.36
Guinea-kafir	5	0.14 - 0.53	0.32
Kafir	25	0.13 - 1.68	0.45
Kafir-bicolor	2	0.37 - 0.63	0.50
Kafir-caudatum	34	0.09 - 1.62	0.49
Kafir-durra	4	0.19 - 1.03	0.62
Seed color			
Colored	340	0.07 - 2.99	0.52
Non-colored	194	0.02 - 2.00	0.52
Origin			
Africa	286	0.07 - 2.99	0.54
America	86	0.10 - 1.49	0.54
Arabia	11	0.20 - 0.72	0.44
Asia	89	0.02 - 1.40	0.47
Australia	1	0.64 - 0.64	0.64

Table 4. High and low 3-DA producing sorghum accessions across mini-core and sorghum association panels

Genotype	Panel	3-DAs (ƛmax 480nm/mg tissue)	Origin	Race	Seed color
High					
IS 11619	Mini-core	2.99	Ethiopia	Durra-bicolor	Reddish brown
IS 9108	Mini-core	2.92	Kenya	Caudatum	Reddish brown
IS 11919	Mini-core	2.20	Ethiopia	Durra-bicolor	Reddish brown
PI 597967	SAP	2.00	Ethiopia	Caudatum	White
IS 15478	Mini-core	2.00	Cameroon	Guinea-caudatum	White
Low					
IS 1219	Mini-core	0.02	China	Guinea-bicolor	White
PI 533869	SAP	0.07	Tanzania	Guinea	White
PI 656093	SAP	0.07	Nigeria	Guinea	White
PI 533760	SAP	0.07	Sudan	Caudatum	White
PI 533965	SAP	0.07	Uganda	Caudatum-bicolor	Red

Table 5. List of candidate genes for fungal induced 3-DAs in sorghum

Gene Name (Version 3)	Gene ID (Version 2)	Description	SNP	*P*-value	PV (%)	FDR
Sobic.004G134200	Sb04g013160	Guanylate-binding protein	S4_22015607	1.02E-09	9.586	0.000
Sobic.001G469500	Sb01g044000	Sugar transporter	S1_67119515	3.49E-08	8.136	0.003
Sobic.002G319766	Sb02g034830	BCS1 AAA-type ATPase	S2_69298227	4.22E-08	8.058	0.003
Sobic.001G049200	Sb01g004510	Acyl-activating enzyme	S1_3611958	1.00E-06	6.785	0.023
Sobic.004G123200	Sb04g010280	Myb-like DNA-binding domain	S4_13892549	1.00E-06	6.784	0.023
Sobic.009G055300	Sb09g004660	Peroxidase	S9_5629470	1.07E-06	6.759	0.023
Sobic.009G114000	Sb09g018180	Glycosyl hydrolase family	S9_45610963	3.95E-06	6.241	0.052
Sobic.010G120100	Sb10g010580	Putative expansin-B14 precursor	S10_13574485	5.13E-06	6.138	0.057
Sobic.007G073400	Sb07g005820	SHR5-receptor-like kinase	S7_8220310	6.25E-06	6.060	0.064
Sobic.001G427900	Sb01g040200	ASC1-like protein 3	S1_63581937	7.57E-06	5.985	0.071
Sobic.007G081500	Sb07g006550	GTPase activator protein-like/Rab-GTPase-TBC domain	S7_9640975	7.73E-06	5.976	0.071

Table 6. 3-deoxyanthocyanidin variability in the sorghum nested association mapping recombinant inbred line cross of SC 265 × RTx 430. The top and bottom five lines are listed in the below table.

NAM RIL, SC 265 × RTx 430	3-deoxyanthocyanidin (Apigeninidin /g tissue)
PR1314_2772	65.66
PR1314_2936	60.73
PR1314_2860	59.21
PR1314_2940	51.02
PR1314_2868	48.88
PR1314_2804	1.89
PR1314_2780	2.21
PR1314_2787	2.25
PR1314_2777	2.32
PR1314_2813	2.38
Mean ± SD	13.31 ± 11.12

Table 7. List of candidate genes identified in NAM RILs for fungal induced 3-DAs in sorghum.

Gene Name (Version 3)	Gene ID (Version 2)	Description	SNP	*P*-value	PV (%)	MAF
Sobic.003G099000	Sb03g008290	Helix-loop-helix DNA-binding domain	S3_8782734	0.001	6.73	0.18
Sobic.003G099300	Sb03g008320	RAP55 related	S3_8806240	0.001	6.60	0.18
Sobic.003G106200	Sb03g008920	Integrator complex subunit 7	S3_9697964	0.002	6.27	0.19
Sobic.001G200800	Sb01g017700	Nonspecific lipid-transfer protein	S1_18269654	0.003	5.89	0.21
Sobic.003G106400	Sb03g008940	Sterol-8,7-isomerase	S3_9719016	0.004	5.53	0.18
Sobic.001G283200	Sb03g008680	26S proteasome regulatory subunit N9	S3_9360233	0.004	5.50	0.19
Sobic.003G212300	Sb03g027200	BCS1 AAA-type ATPase	S3_54725238	0.006	5.27	0.13
Sobic.003G097000	Sb03g008160	Serine/threonine protein kinase	S3_8643920	0.007	5.15	0.18
Sobic.001G203900	Sb01g018010	Glutaredoxin-related protein	S1_18675800	0.007	5.15	0.21

Chapter 3

Genome wide mapping of frost induced flavonoid metabolites and antioxidant activity in global sorghum germplasm

Abstract

Sorghum is a climate resilient multipurpose crop. Moreover, it exhibits poor tolerance to cold and frost. A number of flavonoid secondary metabolites are induced during the plant–stress cross talk, and they play a major role in imparting stress tolerance to plants. Our objective of this study is to quantify the flavonoid induction before and after frost stress, and determine its anti-oxidant activity in the global sorghum panel and also identify the key candidate genes involved in the process through genome wide association mapping. GWAS identified 51 genes for DPPH (anti-oxidant activity) and 20 genes for total phenolic content (TP) before frost. Whereas, 19 genes were identified 3-deoxyanthocyanidins (3-DAs) and 6 genes for TPC under after frost stress. Most of the identified genes are involved in plant defense pathways for biotic and abiotic resistance. The probable candidates for after stress were peroxidase (sb01g041770), vesicle fusing ATPase (sb01g041930), leucine rich repeats (sb08g001430), flavonol reductase (sb01g025770), cytosine deaminase (sb02g041770), and anthocyanidin reductase (sb02g038520) for 3-DAs; and leucine rich repeat (Sb08g023030) and UDP-glucosyl transferase (Sb06g021900) for TP. Anti-fungal gene (Sb06g014340) was the most probable candidate for DPPH before frost stress. The identified candidate genes can be used in breeding climate resilient sorghum.

Keywords: 3-DAs, DPPH, Frost, SNPs, Total phenolic content.

Abbreviations

3-DAs	- 3-deoxyanthocyanidins
CMLM	- Compressed Mixed Linear Model
DPPH	- 1,1-diphenyl-2-picrylhydrazyl
GAPIT	- Genomic Association and Prediction Integrated Tool
GWAS	- Genome Wide Association Study
ICRISAT	- International Crops for the Semi-Arid Tropics
MNC	- ICRISAT Mini-core Collection
PCA	- Principal Component Analysis
Q-Q plot	- Quantile-Quantile plot
SAP	- Sorghum Association Panel
SNPs	- Single Nucleotide Polymorphisms
SUPER	- Settlement of MLM under Progressively Exclusive Relationship
TPC	- Total Phenolic Content

Introduction

Sorghum is a major cereal crop grown worldwide for use as human food and livestock feed [1]. Millions of people in sub-Saharan Africa and Asia survive on food made from sorghum grain [1, 2]. In the United States (US), sorghum is primarily used as animal feed [3], but it is exploited as an alternate source of bio-energy. US remains world leader in sorghum export, farmers planted 5.6 million acres of sorghum and harvested 9.2 million metric tons of sorghum grain in 2017 across the US [4]. During the 2015-16 fiscal year alone US exported $1.68 billion worth of sorghum grains to China [5]. Sorghum is well suited to drought prone environments and requires less input and also can be grown on marginal soils [6-11]. Due to these reasons, sorghum is continually expanding to newer areas [12]. Therefore, we need to develop sorghum cultivars that fit in all environments.

Sorghum is a warm-season crop mainly grown in semiarid, sub-humid, and humid tropical and subtropical parts of the world where average minimum temperatures generally stay above 18ºC during the growing season [13]. Although, it is highly susceptible to chilling, frost, and freezing stress. Chilling stress greatly affect the seed germination, frost reduces the grain size [14-16], nutritional content, increases the cyanogenic (HCN) content [17], and induces male sterility [18]. In northern, northwestern, and northeastern USA, southern Canada, and highlands of Latin America the peak flowering of sorghum coinciding with the cooler days in winter [13]. In this environment, reproductive success and ultimately grain yield is limited by poor seed setting [19]. Moreover, early or late planting is not a viable option that affect either the shoot biomass productivity or partitioning; because below-normal temperatures that occur late in the growing-season, can delay grain sorghum

maturity, and delaying maturity increases the probability of a freeze occurring before the crop reaches physiological maturity [14]. Thus, low temperature and frost tolerance is very much important for sorghum adaptation to high altitude regions.

Low temperature decreases the crop productivity, increases the reactive oxygen species (ROS) production that in turn activate the synthesis of anti-oxidants, also increases the unsaturated fatty acid content [20]. A killing frost before maturity stops the growth of the seed, creating small, lightweight grain, which may be shrivelled and difficult to thresh. Signs of frost damage include watery patches on plant leaves, followed by necrotic white lesions on the affected area. A light frost may kill the leaves, but the grain may continue to fill until the stalk dies. Death does not occur until the stalks have been frozen, breaking the flow of nutrients to the grain. Many of the biological roles of flavonoids are attributed to their potential cytotoxicity and antioxidant abilities [21-24]. A number of flavonoid class secondary metabolites are induced during frost stress. They play a major role in crop growth and protection.

Flavonoid, a class of low-weight phenolic compounds, are widely distributed in plants [25]. Attributed to their structural diversity, flavonoid perform varied biological functions, such as acting as pigments to attract pollinators, as auxin transport regulators and as molecular signals for the interaction of plants with microorganisms. Several lines of evidence substantiate the premise that flavonoids have antioxidant functions in higher plants that are challenged with a range of environmental stresses [23, 26-28]. First, polyphenol-oxidases and peroxidases occur ubiquitously in higher plants [29]. Second, plants undergoing severe

stress conditions preferentially accumulate dihydroxy B-ring-substituted flavonoids, which are effective scavengers of ROS [23]. There are reports indicating that bio-synthesis of antioxidant flavonoids increases more in stress-sensitive plants than in stress-tolerant plants [23]. Under severe stress conditions the stress sensitive plants might inactivate antioxidant enzymes, while up regulating the biosynthesis of flavonols [30]. As a result, flavonoid compounds in turn act as a potential anti-oxidative agent in the absence of primary ROS scavengers [30].

Due to the advent of low-cost sequencing technologies genome wide association studies (GWAS) are gaining attention in recent days to understand and dissect the most complex traits in plants and humans. GWAS offers high allelic diversity and high recombination rate compared to bi-parental mapping. Successful GWA studies in sorghum identified genomic regions associated with trait of interest for grain quality [31], anthracnose resistance [32], agroclimatic traits [33], aluminium tolerance, grain polyphenol concentrations [34], adaptive traits [35], seed size [36], Fe and Zn density, stalk rot resistance [37], and cold and heat stress [38]. Until now, no studies have been carried out on the effect of frost stress on the flavonoid metabolism and total anti-oxidant activity in sorghum. Therefore, a study was conducted to understand whether there is common allelic variation that could explain differences under before and after frost stress for the induction of different secondary metabolites in sorghum. Thus, our objectives of this study were 1) to assess the variability of stress induced flavonoid metabolites and 2) to identify genes/QTLs associated with antioxidant activity, total phenolic content, and 3-

deoxyanthocyanidins using phenotype data generated through wet chemistry with a larger sorghum population.

Materials and Methods

Plant materials and genotype data

A total of 551 global sorghum accessions from sorghum association panel (SAP#315) and ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) mini-core (MNC#236) collection were used in the study. The germplasm collections were received from GRIN (Germplasm Resources Information Network: www.ars-grin.gov). The germplasm accessions used in this study were diverse geographically and racially, which represents the major races of kafir, durra, caudatum, guinea, bicolor and working groups of sorghum. A total of 265,487 SNPs (Single Nucleotide Polymorphisms) (http://www.morrislab.org/data) developed through genotype-by-sequencing (GBS) method was used in this study, which corresponds to variation in 27,412 annotated genes in sorghum [33].

Growth conditions

Sorghum diversity panels were grown during summer 2014 and 2015 at the Penn state agronomy farm at Rock Springs, Pennsylvania (40° 43` N, 77° 56` W, 372 m elevation) under rain-fed conditions. Triplicate leaf samples were collected at booting stage (45 days after planting) for before stress, and 75 days after planting (soft-dough stage) for a day after natural frost. All the leaf samples were subjected to quantify the plant stress response to frost stress towards their bio-chemical changes before and after frost stress. We have

quantified total phenolic content (TPC), anti-oxidant activity (DPPH) and 3-deoxyanthocyanidins (3-DAs) in all 551 global sorghum germplasm lines under before and after frost stress.

Reagents

Methanol was purchased from EMD Millipore, MA. Hydrochloric acid was obtained from Ricca Chemical Company, Texas. 1,1-diphenyl-2-picrylhydrazyl (DPPH) was purchased from Alfa Aesar, MA and Trolox and gallic acid were obtained from Acros Organics, NJ. Folin-Ciocalteu reagent were purchased from Sigma. Sodium phosphate (monobasic and dibasic, anhydrous) and sodium carbonate were purchased from Fisher Scientific.

Extraction

Sorghum leaf samples were extracted in methanol acidified with 0.1% (v/v) hydrochloric acid. The leaf discs were left in methanol for 48 hours at -20°C, after which they were vortexed on ice for 30 min. The samples were then centrifuged for 15 minutes at 3500 rpm at 4°C. The resulting supernatant was collected and transferred to new micro plates for quantitative analyses of total phenolic content (TPC), anti-oxidant activity (DPPH) and 3-deoxyanthocyanidins (3-DAs).

Total phenolic content

The total phenolic content of sorghum was assayed using the colorimetric Folin-Ciocalteu method. Briefly, 35 µl of a known dilution of the sample was added to 150 µl of a 1:10 (v/v) dilution of Folin-Ciocalteu reagent and the plate was mixed on a platform vortex for

30 seconds at 400 rpm. The reaction was allowed for 5 minutes. Then, 115 µl of a 7.5% Sodium Carbonate solution was added to all wells and the plate mixed again as mentioned above and incubated at 45°C for 30 minutes and then cooled to room temperature for an additional 60 minutes. The absorbance was read at 765 nm using Biotek Synergy 2 plate reader. The observed values were compared to a standard curve of known dilutions of gallic acid and expressed as milligrams of gallic acid equivalents per gram of sorghum leaf sample.

DPPH antioxidant assay

The antioxidant potential of the samples were determined by using DPPH assay. DPPH is a stable free radical that gives a UV-vis absorption maximum at 515 nm. The observed values are compared to a standard curve obtained by plotting the net AUC of different known Trolox concentrations. Analyses are run in triplicates and are reported as Trolox equivalents (TE).

3-deoxyanthocyanidins (3-DAs)

Acidified methanol extracted was used to quantify the 3-DAs level. All the samples were transferred to 96 well plate with a volume of 350 µl. Cytation 5 (BioTek, USA) plate reader was used to quantify the 3-DAs at the 480nm (λ_{max}) along with acidified methanol as a blank. Subsequently, the blank values were subtracted from the observed values to determine the total 3-DAs.

Association analyses

Genome-wide association analysis was performed using CMLM (compressed mixed linear model) in combination with SUPER (Settlement of MLM Under Progressively Exclusive Relationship) model to determine marker-trait associations. The compressed mixed linear model (CMLM) involves genetic marker–based kinship matrix modeling of random effects used jointly with population structure estimated by principal components analysis (PCA) to model fixed effects [40]. SUPER model (Wang et al. 2014) extracts a small subset of SNPs from CMLM and uses them in FaST-LMM method. This method increases statistical power and is computationally advantageous. The compression level and optimal number of principal components that adequately explain population structure were previously determined by the Genomic Association and Prediction Integrated Tool (GAPIT) [41]. Log quantile-quantile (QQ) p-value plots for 265,487 single SNP tests of association implied that there were few systematic sources of spurious association using CMLM with SUPER model. As several studies have suggested varying number of linkage disequilibrium (LD) blocks in sorghum, LD range of 2–5 kb was used for the analysis [33].

Genes in QTL regions

Ensembl sorghum expression data (http://plants.ensembl.org/Sorghum_bicolor) was used to determine the genes nearby the SNPs from the obtained GWAS result using bio-conductor packages in R package. The genes and their functions with genomic coordinates were obtained using gramene biomart (www.gramene.org) and phytozome v.10 (www.phytozome.net). Closest genes in the upstream or downstream of 2 to 5 kb were selected and the respective annotations were obtained for each of the genes.

Statistical Analysis

The total measurements included two years (2014 and 2015) field × two treatments (before frost and after frost) × 551 lines × three different assays (TPC, DPPH, and 3-DAs) × three biological replicates on each assay for a total of 19,836 for all the assays. Covariation was determined with Pearson's correlation co-efficient. A simple linear regression was used to model the effect of total phenolic content, anti-oxidant activity and 3-DAs. P-values of < 0.05 were considered statistically significant. Statistical analysis was conducted with R, a statistical computing language and environment (R Core Team 2015). Data obtained were expressed as mean ± standard deviation.

Results and Discussion

Our objective was to investigate the bio-chemical changes that occur in sorghum plants once the frost arrives and identify the significant marker-trait associations to improve frost tolerance in sorghum. Thus, we have conducted a two-year field experiment using the wide collection of sorghum materials at two-time points (before and after frost) during the growing season in central PA. Here the results are presented treatment wise and discussed.

Phenotypic variation and correlations among the traits

Before frost stress: We have quantified the level of antioxidant activity (DPPH assay), total phenolic content (TP), and 3-deoxyanthocyanidins (3-DAs) and their covariation with each other (Table 8). The DPPH values for the sorghum population before frost ranged from 0.5 μg of TE/mg (sorghum leaf tissue) in PI535882 to 88.6 μg of TE/mg in PI23590, with a mean of 22.7 μg of TE/mg (Fig. 13, and Table 8). TP in the 551 sorghum accessions

before frost ranged from 0.02 µg of GAE/mg (sorghum leaf tissue) in IS31186 and IS32349 to 46.3 µg of GAE/mg in PI533866, with a mean of 16.6 µg of GAE/mg (Fig. 13, Table 8). The 3-DAs before frost ranged from 0.01 abs/mg in IS2379 to 9.7 abs/mg in IS12706, with a mean of 1.9 abs/mg (Fig. 13, Table 8). Moderate heritability values were observed for 3-DAs (0.33) and DPPH (0.51) and high heritability was observed for TP (0.75) under before frost condition (Table 8).

After frost stress: The DPPH values of the sorghum population after frost ranged from 2.1 µg of TE/mg (sorghum leaf tissue) in IS25249 to 71.1 µg of TE/mg in IS13893, with a mean of 16.0 µg of TE/mg (Fig. 13, and Table 8). TP in the 551 sorghum accessions after frost ranged from 1.9 µg of GAE/mg (sorghum leaf tissue) in IS13782 to 90.6 µg of GAE/mg in IS13893, with a mean of 24.8 µg of GAE/mg (Fig. 13, Table 8). The 3-DAs after frost ranged from 0.1 abs/mg in PI656048 and IS13782 to 56.4 abs/mg in IS12706, with a mean of 2.9 abs/mg (Fig. 13, Table 8). Low heritability values were observed for DPPH (0.16) and TP (0.23) and moderate heritability for 3-DAs (0.48) under after frost stress (Table 8). The level of 3-DAs and TP was increased significantly after the frost stress (Table 8). However, DPPH level was decreased after the frost stress. The level of variation observed for various traits assessed is reflected in the histograms under before and after frost stress (Fig. 13).

Particularly, under after frost stress condition all the metabolites had positive relationship with each another (Fig. 13). Under after frost stress condition, the correlation between the DPPH and TP was higher (0.60) while the correlation of 3-DAs with the other traits such

as DPPH (0.18) and TPC (0.34) were moderate (Fig. 13). Whereas, we have observed significant negative correlations between the traits (-0.22) except for 3-DAs and TP (0.58), which has shown positive linear relationship under before frost (Fig. 13). The correlation analysis revealed the networking action of different metabolites after the frost stress. These results suggest that the observed genetic variations underlying before and after frost stress responses are biologically meaningful. Thus, molecular markers associated with the traits could be applicable in sorghum breeding for improving frost tolerance.

Genome wide association analyses

Before frost stress: A total of 71 SNPs were significantly associated with the three traits measured with FDR values ≤0.0001 (Table 9). 51 SNPs were significantly associated with DPPH and 20 SNPs were significantly associated with TP (Table 9). Notably, the significant SNPs that indicate marker-trait associations (MTA) for the DPPH and TP content were mainly on Chr06 and Chr09 (Fig. 14). We have observed significant association for DPPH across all chromosomes except the chromosome 7 (Fig. 14). Majority of the SNPs associated with the traits subjected to before frost stress were either within the gene or a few thousand nucleotides apart (Table 11). A total of 65 putative genes near the 71 associated loci were identified and are presented with gene annotations including possible functions, in Table 12.

After frost stress: A total of 25 SNPs were significantly associated with the three traits measured with FDR values ≤0.0001 (Table 9). 19 SNPs were significantly associated with 3-DAs on Chr01, Chr02, Chr04, Chr05, Chr06, Chr08 and Chr09 (Fig. 14). 6 SNPs for TP

on Chr03 and Chr06. Majority of the SNPs associated with the traits subjected to after frost stress were either within the gene or a few thousand nucleotides apart. Twenty putative genes near the 25 associated loci were identified and are presented with gene annotations including possible functions, in Table 13.

Candidate genes

GWAS identified following candidate genes for DPPH, TP and 3-DAs under before and after frost stress. The roles of the identified candidate genes in frost tolerance are delineated here.

3-DAs: GWAS identified 19 candidate genes for 3-DAs under after frost stress (Table 11 and 13). Most of the identified SNPs were located in Chr01 (Fig. 14). The top candidate gene from our GWAS analysis was peroxidase (sb01g041770) and followed by vesicle fusing ATPase (sb01g041930), leucine rich repeats (sb08g001430), flavonol reductase (sb01g025770), cytosine deaminase (sb02g041770), sugar transporter (sb02g021040), anthocyanidin reductase (sb02g038520) and WRKY DNA-binding domain (sb04g005520) (Table 6). Peroxidase was detected all across the traits measured under after frost stress. They are potential scavengers of free radical [39]. Peroxidase quench the singlet oxygen molecule from the hydrogen peroxidase (H_20_2) and convert them into water (H_20) in mitochondria, chloroplast and peroxisomes [39]. Peroxidase is one of the most important candidate gene that we detected for frost stress. We have also identified genes most likely involved in the production of flavonoid secondary metabolites such as flavonol reductase and anthocyanidin reductase. They play a major role in flavonoid metabolite synthesis in

the upstream and downstream pathways of flavonoid bio-synthesis upon stress induction. Finally, the genes like cytosine deaminase, leucine rich repeats, sugar transporter and WRKY DNA-binding domain they all play a role in induction of defence genes and transport of nutrients under biotic and abiotic stress conditions. These identified candidate genes will be utilized in marker assisted selection and genomic selection for crop improvement.

DPPH (Anti-oxidant activity): GWAS identified 51 candidate genes for DPPH under before frost stress (Table 10 and 12). Whereas, GWAS did not detect any significant candidates for DPPH under after frost stress. The possible reason for this discrepancy may be due to all the flavonoid compounds induced under stress might be acted as potential anti-oxidants to quench the free radicals and ROS [23, 30]. The possible candidate genes for DPPH under before frost stress were salt stress response/anti-fungal (Sb06g014340), transcriptional regulator (Sb09g025910), predicted Na^+-dependent cotransporter (Sb09g028950), aluminium activated malate transporter (Sb06g025640) and heat shock protein (Sb06g000660). However, we have detected peroxidase (Sb06g033860) and cell wall invertase (Sb04g021810) as top candidates for DPPH under after frost stress.

TP: GWAS identified 20 candidate genes for TPC under before frost stress (Table 10 and 12) and 6 candidate genes under after frost stress (Table 11 and 13). We have got enrichment in Chr09 for TPC under before frost (Fig. 14). Whereas under after frost stress we have observed most of the significant SNPs in Chr03 (Fig. 14). The top candidate gene in our GWAS analysis under after frost was apoptotic ATPase/ leucine rich repeat/NB-

ARC domain (Sb08g023030) followed by UDP-glucoronosyl and UDP-glucosyl transferase (Sb06g021900) and mitochondrial chaperone BCS1 (Sb02g005660).

Conclusions

A comprehensive investigation involving large sorghum accessions for frost induced metabolite profiling of genome wide association mapping revealed a number of key associations and their potential role in regulating the secondary metabolite bio-synthesis upon stress induction. Many of the associated SNPs for frost induced metabolites are in very close proximity to stress responsive genes. Particularly, flavonoid class genes are detected as candidate genes. The genes like peroxidase and apoptotic ATPase (leucine rich repeats) are enriched in all the traits measured under after frost stress conditions. The identified SNPs will be used in marker assisted breeding and genomic selection for crop improvement to develop frost tolerant sorghum cultivars.

Acknowledgements

We thank GRIN germplasm network and Dr. Wilfred Vermerris (University of Florida) for providing the sorghum mapping panels to conduct this study. We also want to thank Mr. Scott Harkcom, agronomy farm manager at penn state rock springs for facilitating us in field preparation and necessary managements. Thanks to Jin Cui, Debamalya Chatterje, Kameroon Wittemeyer and Bin Liu for helping us in sample collection and phenotyping.

Literature Cited

1. Reddy BVS, Reddy PS, Sadananda AR, Dinakaran E, Kumar AA, Deshpande SP, Rao PS, Sharma HC, Sharma R, Krishnamurthy L *et al*: **Postrainy season sorghum: Constraints and breeding approaches**. *Journal of SAT Agricultural Research* 2012, **10**:1-12.
2. FAOSTAT: **Food and Agriculture Organization of the United Nations. FAOSTAT Statistics Database.** 2016.
3. Rhodes D, Gadgil P, Perumal R, Tesso T, Herald TJ: **Natural Variation and Genome-Wide Association Study of Antioxidants in a Diverse Sorghum Collection**. *Cereal Chemistry Journal* 2017, **94**(2):190-198.
4. **All about sorghum** [http://www.sorghumcheckoff.com/all-about-sorghum]
5. **Market opportunities international** [http://www.sorghumcheckoff.com/market-opportunities/international/]
6. Martin JH: **The comparative drought resistance of sorghums**. *Agron J* 1930, **22**(993–1003).
7. Martin JH, Leonard WH, Stamp DL: **Principles of field crop production**. New York: Macmillan Publishing Co Inc; 1976.
8. Lamm FR, Stone LR, O'Brien DM: **Crop production and economics in Northwest Kansas as related to irrigation capacity**. *Appl Eng Agric* 2007, **23**:737– 745.
9. Merrill SD, Tanaka DL, Krupinsky JM, Liebig MA, Hanson JD: **Soil water depletion and recharge under ten crop species and application to the principles of dynamic cropping systems**. *Agron J* 2007, **99**:931–938.
10. Howell TA, Steiner JL, Schneider AD, Evett SR, Tolk JA: **Seasonal and maximum daily evapotranspiration of irrigated wheat, sorghum, and corn—Southern High Plains**. *Trans Am Soc Agric Eng* 1997, **40**:623–634.
11. Howell TA, Yazar A, Schneider AD, Dusek DA, Copeland KS: **Yield and water use efficiency of corn in response to LEPA irrigation**. *Trans Am Soc Agric Eng* 1995, **38**:1737–1747.
12. Liu C, Ketterings Q, Czymmek K, Kilcer T, Cherney J, Hunter M, Neil KO, Lyons S: **Guidance for growing BMR brachytic dwarf forage sorghum**. *Cornell University Cooperative Extension Agronomy Fact Sheet Series* 2016, **92**:1-2.
13. Singh SP: **Sources of cold tolerance in grain sorghum**. *Can J Plant Sci* **1985**, **65**:251-257.
14. Staggenborg SA, Vanderlip RL: **Sorghum grain yield reductions caused by duration and timing of freezing temperatures**. *Agron J* 1995, **88** (3):473-477.
15. Staggenborg SA, Vanderlip RL, Roggenkamp GJ, Kofoid KD: **Methods of simulating freezing damage during sorghum grain fill**. *Agronomy Journal* 1999, **91**(1):46-53.
16. Roozeboom K, McKinney L: **Harvesting grain from freeze-damaged sorghum**. In.: Kansas State University; 2009.

17. Wattenbarger DW, Gray E, Rice JS, Reynolds JH: **Effects of frost and freezing on hydrocyanic acid potential of sorghum plants**. *Crop Science* 1968, **8**(5):526-528.
18. Wood AW, Tan DKY, Mamun EA, Sutton BG: **Sorghum can compensate for chilling-induced grain loss**. *J Agronomy and Crop Science* 2006, **192**:445—451.
19. Jain M, Prasad PV, Boote KJ, Hartwell AL, Jr., Chourey PS: **Effects of season-long high temperature growth conditions on sugar-to-starch metabolism in developing microspores of grain sorghum (Sorghum bicolor L. Moench)**. *Planta* 2007, **227**(1):67-79.
20. Guo Y, Liu S, Yang Z, Tian S, Sui N: **Responses of Unsaturated Fatty Acid in Membrane Lipid and Antioxidant Enzymes to Chilling Stress in Sweet Sorghum (Sorghum bicolor (L.) Moench) Seedling**. *Journal of Agricultural Science* 2016, **8**(9).
21. Kawahigashi H, Kasuga S, Sawada Y, Yonemaru J, Ando T, Kanamori H, Wu J, Mizuno H, Momma M, Fujimoto Z *et al*: **The Sorghum Gene for Leaf Color Changes upon Wounding (P) Encodes a Flavanone 4-Reductase in the 3-Deoxyanthocyanidin Biosynthesis Pathway**. *G3 (Bethesda)* 2016, **6**(5):1439-1447.
22. Guidi L, Brunetti C, Fini A, Agati G, Ferrini F, Gori A, Tattini M: **UV radiation promotes flavonoid biosynthesis, while negatively affecting the biosynthesis and the de-epoxidation of xanthophylls: Consequence for photoprotection?** *Environmental and Experimental Botany* 2016, **127**:14-25.
23. Tattini M, Guidi L, Morassi-Bonzi L, Pinelli P, Remorini D, Degl'Innocenti E, Giordano C, Massai R, Agati G: **On the role of flavonoids in the integrated mechanisms of response of Ligustrum vulgare and Phillyrea latifolia to high solar radiation**. *New Phytol* 2005, **167**(2):457-470.
24. Ibraheem F, Gaffoor I, Chopra S: **Flavonoid phytoalexin-dependent resistance to anthracnose leaf blight requires a functional yellow seed1 in Sorghum bicolor**. *Genetics* 2010, **184**(4):915-926.
25. Dao TT, Linthorst HJ, Verpoorte R: **Chalcone synthase and its functions in plant resistance**. *Phytochem Rev* 2011, **10**(3):397-412.
26. Pollastri S, Tattini M: **Flavonols: old compounds for old roles**. *Ann Bot* 2011, **108**(7):1225-1233.
27. Agati G, Azzarello E, Pollastri S, Tattini M: **Flavonoids as antioxidants in plants: location and functional significance**. *Plant Sci* 2012, **196**:67-76.
28. Agati G, Brunetti C, Di Ferdinando M, Ferrini F, Pollastri S, Tattini M: **Functional roles of flavonoids in photoprotection: new evidence, lessons from the past**. *Plant Physiol Biochem* 2013, **72**:35-45.
29. Pourcel L, Routaboul JM, Cheynier V, Lepiniec L, Debeaujon I: **Flavonoid oxidation in plants: from biochemical properties to physiological functions**. *Trends Plant Sci* 2007, **12**(1):29-36.
30. Fini A, Guidi L, Ferrini F, Brunetti C, Di Ferdinando M, Biricolti S, Pollastri S, Calamai L, Tattini M: **Drought stress has contrasting effects on antioxidant enzymes activity and phenylpropanoid biosynthesis in Fraxinus ornus leaves: an excess light stress affair?** *J Plant Physiol* 2012, **169**(10):929-939.

31. Sukumaran S, Xiang W, Bean SR, Pedersen JF, Kresovich S, Tuinstra MR, Tesso TT, Hamblin MT, Yu J: **Association Mapping for Grain Quality in a Diverse Sorghum Collection**. *The Plant Genome Journal* 2012, **5**(3).
32. Upadhyaya HD, Wang YH, Sharma R, Sharma S: **Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis**. *Theor Appl Genet* 2013, **126**(6):1649-1657.
33. Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE *et al*: **Population genomic and genome-wide association studies of agroclimatic traits in sorghum**. *Proc Natl Acad Sci U S A* 2013, **110**(2):453-458.
34. Rhodes DH, Hoffmann L, Jr., Rooney WL, Ramu P, Morris GP, Kresovich S: **Genome-wide association study of grain polyphenol concentrations in global sorghum [Sorghum bicolor (L.) Moench] germplasm**. *J Agric Food Chem* 2014, **62**(45):10916-10927.
35. Jessy L, Upadhaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J, Juenger TE, Hyma K, Acharya C, Mitchell SE *et al*: **Genome-environment associations in sorghum landraces predict adaptive traits**. *Science Advances* 2015, **1**(6):1-13.
36. Zhang D, Li J, Compton RO, Robertson J, Goff VH, Epps E, Kong W, Kim C, Paterson AH: **Comparative Genetics of Seed Size Traits in Divergent Cereal Lineages Represented by Sorghum (Panicoidae) and Rice (Oryzoidae)**. *G3 (Bethesda)* 2015, **5**(6):1117-1128.
37. Adeyanju A, Little C, Yu J, Tesso T: **Genome-Wide Association Study on Resistance to Stalk Rot Diseases in Grain Sorghum**. *G3 (Bethesda)* 2015, **5**(6):1165-1175.
38. Chopra R, Burow G, Burke JJ, Gladman N, Xin Z: **Genome-wide association analysis of seedling traits in diverse Sorghum germplasm under thermal stress**. *BMC Plant Biol* 2017, **17**(1):12.
39. Caverzan A, Passaia G, Rosa SB, Ribeiro CW, Lazzarotto F, Margis-Pinheiro M: **Plant responses to stresses: Role of ascorbate peroxidase in the antioxidant protection**. *Genetics and Molecular Biology* 2012, **35**(4):1011-1019.
40. Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM *et al*: **Mixed linear model approach adapted for genome-wide association studies**. *Nat Genet* 2010, **42**(4):355-360.
41. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z: **GAPIT: genome association and prediction integrated tool**. *Bioinformatics* 2012, **28**(18):2397-2399.

Figure Legends

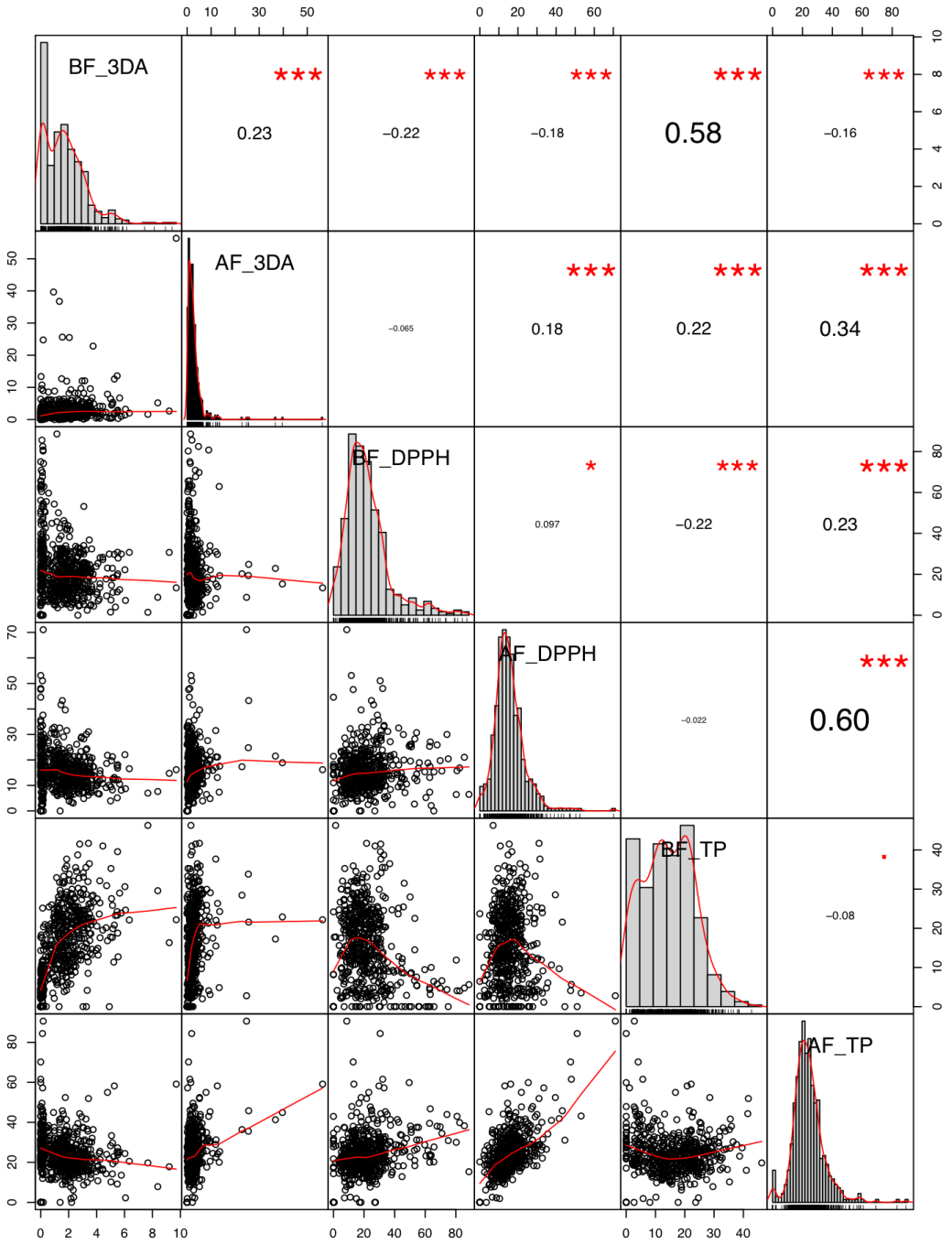

Figure 13. Variation and pearson pairwise correlations of 3-deoxyanthocyanidin (3-DA), anti-oxidant activity (DPPH), and total phenolic content (TP) under before and after frost stress. Histograms for 3-DA, DPPH, and TP levels are displayed along the diagonal. *BF- Before frost; AF- After frost.

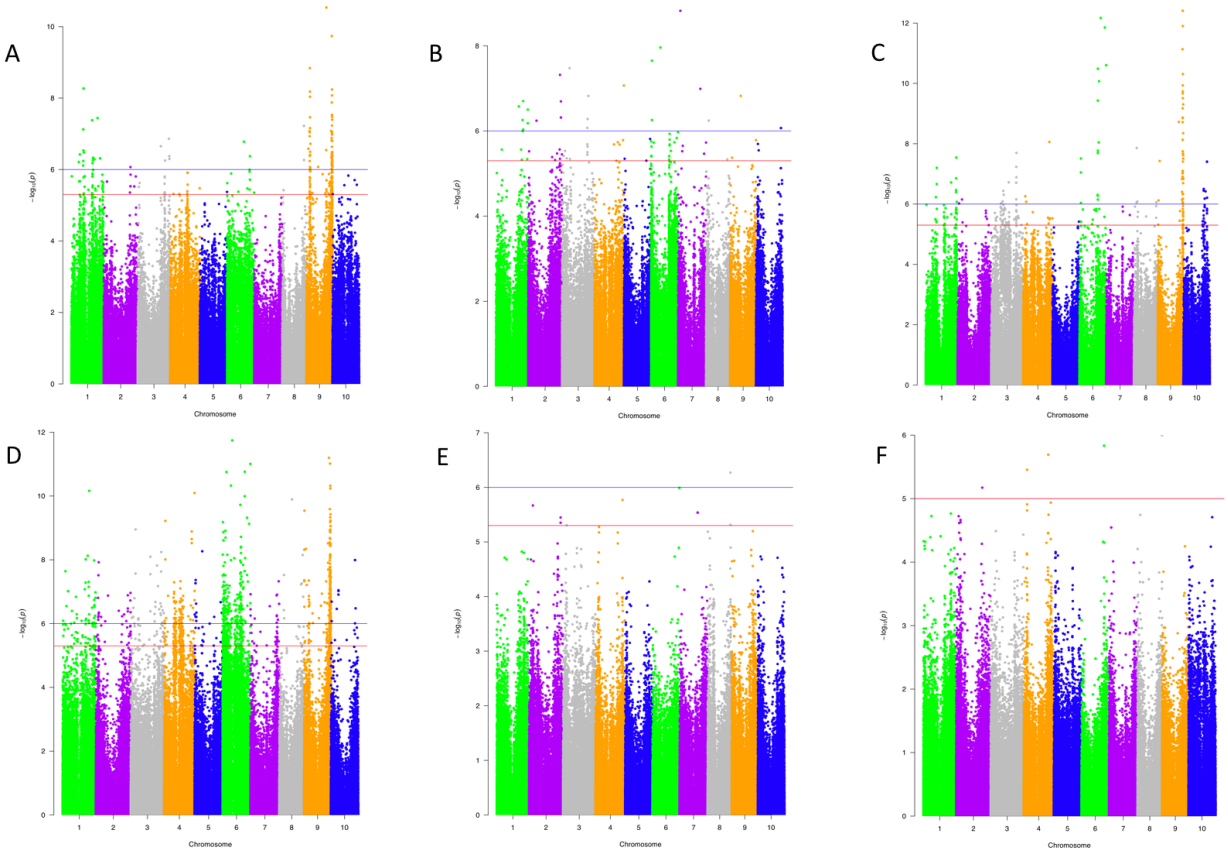

Figure 14. Genome-wide association study (GWAS) under before and after frost stress for 3-deoxyanthocyanidin (3-DA), antioxidant activity (DPPH), and total phenolic content (TP) of sorghum leaf extracts. Scatter plot of GWAS results for 3-DA (A), DPPH (B), and TP (C) values under before frost stress, and 3-DA (D), DPPH (E), and TP (F) values under after frost stress. Negative log10-transformed *P* values (y axis) are plotted against the physical single nucleotide polymorphism (SNP) position on each chromosome (x axis). Each circle represents a SNP. The solid horizontal blue line represents the genome-wide significance threshold as determined by Bonferroni correction, and solid red line represents the suggested significance threshold line. Regions with negative log10 *P* values above the threshold contain quantitative trait loci candidates.

Tables

Table 8. Summary statistics expressed as mean and heritability for traits evaluated for GWAS under frost stress

Stress	Traits measured	Low	High	Mean (±SD)	Heritability (%)
Before frost (BF)	*3-DAs (480nm mg^{-1})	0.01	9.70	1.90(±1.47)	33.4
	†DPPH (µg TE mg^{-1})	0.51	88.59	22.69(±14.30)	51.3
	$TPC (µg GAE mg^{-1})	0.02	46.27	16.55(±8.85)	75.4
After frost (AF)	*3-DAs (480nm mg^{-1})	0.06	56.38	2.94(±4.26)	47.7
	†DPPH (µg TE mg^{-1})	2.11	71.06	15.97(±7.75)	15.6
	$TPC (µg GAE mg^{-1})	1.88	90.60	24.82(±9.90)	23.0

*3-deoxyanthocyanidins
†1,1-diphenyl-2-picrylhydrazyl
$Total phenolic content

Table 9. List of the numbers of significant association for the traits measured under no frost and frost stress including the lowest probability value and false discovery rate (FDR) analysis

Trait (and designation)	Numbers of significantly associated SNPs	Lowest P-value	Lowest FDR Adjusted values
BF- 3-DAs	0	2.91×10^{-11}	7.71×10^{-6}
BF - DPPH	51	1.80×10^{-12}	4.77×10^{-7}
BF - TPC	20	3.90×10^{-13}	5.18×10^{-8}
Total	71		
AF - 3-DAs	19	1.50×10^{-9}	0.0004
AF - DPPH	0	5.38×10^{-7}	0.0964
AF - TPC	6	1.01×10^{-6}	0.1078
Total	25		

Table 10. List of significant SNP associations, the genes tagged by significant SNPs and information from the genome wide analysis of sorghum germplasm under before frost (BF) stress

Trait (BF)	SNP	FDR Adjusted values*	Nucleotide variation	Distance to nearest gene (bp)**	Nearest Gene ID
3-DAs	S9_57836978	2.43205E-05	T/C	2135	Sb09g029090
	S1_27516968	0.00025064	T/C	8730	Sb01g022310
	S9_8049583	0.000270391	C/G	2219	Sb09g005800
	S9_58279046	0.000612454	T/C	7351	Sb09g029640
	S9_58469834	0.000612454	T/C	1109	Sb09g029830
	S1_59686442	0.000612454	A/C	4475	Sb01g036050
DPPH	S9_55291420	6.54E-07	T/C	4291	Sb09g025910
	S9_57712210	6.54E-07	A/C	6615	Sb09g028950
	S6_61977643	6.54E-07	C/G	3257	Sb06g034070
	S6_49533289	7.80E-07	T/A	2598	Sb06g019990
	S1_59232186	1.82E-06	T/G	4936	Sb01g035610
	S4_67863046	1.94E-06	G/A	689	Sb04g038490
	S6_49516846	2.24E-06	T/G	6154	Sb06g019970
	S6_39652889	3.59E-06	A/C	4996	Sb06g014340
	S9_57236791	4.55E-06	G/A	340	Sb09g028340
	S9_592147	4.79E-06	A/G	2230	Sb09g000800
TPC	S6_49516846	5.96E-08	T/G	6154	Sb06g019970
	S9_57963489	7.42E-08	C/A	1226	Sb09g029220
	S6_59090312	7.42E-08	T/G	6363	Sb06g030650
	S9_57561138	3.26E-07	G/A	1342	Sb09g028720
	S6_61977643	9.44E-07	C/G	3257	Sb06g034070
	S9_57836978	4.38E-06	T/C	2135	Sb09g029090
	S9_57941670	5.86E-06	G/C	2801	Sb09g029180

*False Discovery Rate (FDR) values for the significantly associated SNP

**Distance to the nearest annotated gene coordinates in the reference genome

Table 11. List of significant SNP associations, the genes tagged by significant SNPs and information from the genome wide analysis of sorghum germplasm under after frost (AF) stress

Trait (AF)	SNP	FDR Adjusted values*	Nucleotide variation	Distance to nearest gene (bp)**	Nearest Gene ID
3-DAs	S7_5220307	0.000399496	G/T	4595 1340 689 3782	Sb07g004120
	S3_17341041	0.002221416	T/C		Sb03g013650
	S4_67863046	0.003778294	G/A		Sb04g038490
	S7_51110695	0.003873861	A/G		Sb07g019860
DPPH	S8_52274973	0.096366176	C/T	7500	Sb08g020827
	S6_61790504	0.096366176	T/G	1583	Sb06g033860
	S2_8822447	0.096366176	G/A	3531	Sb02g006886
	S8_52442295	0.104948424	G/C	686	Sb08g020990
	S9_48400310	0.104948424	C/G	3130	Sb09g019550
	S4_8066080	0.104948424	C/T	7246	Sb04g007700
	S4_50951550	0.104948424	A/T	4716	Sb04g021810
	S8_5164741	0.115555588	C/T	1834	Sb08g004360
TPC	S8_55167098	0.107847708	T/C	5126	Sb08g023030
	S6_51122953	0.107847708	A/G	1517	Sb06g021900
	S4_55434798	0.107847708	C/G	1586	Sb04g025660
	S2_58868022	0.223021599	C/T	1991	Sb02g024480

*False Discovery Rate (FDR) values for the significantly associated SNP

**Distance to the nearest annotated gene coordinates in the reference genome

Table 12. List of candidate genes identified in the study based on proximity to the significant markers identified through GWAS analysis and their description or function obtained from phytozome (www.phytozome.net) database under no frost stress

Gene ID	Annotation	Trait (no frost)
Sb09g029090	PPR repeat	3-DA
Sb01g022310	Expressed protein	
Sb09g005800	Two-component phosphorelay intermediate involved in MAP kinase cascade regulation	
Sb09g029640	Sister chromatid cohesion complex Cohesin, subunit RAD21/SCC1	
Sb09g029830	Transcriptional repressor, ovate	
Sb01g036050	Osmotic stress potassium transporter	
Sb09g025910	Transcriptional regulator	DPPH
Sb09g028950	Predicted Na+-dependent cotransporter	
Sb06g034070	Plant phosphoribosyltransferase C-terminal	
Sb04g038490	similar to RING zinc finger protein-like	
Sb06g019970	TELOMERIC REPEAT BINDING PROTEIN	
Sb06g014340	Salt stress response/antifungal	
Sb09g000800	EamA-like transporter family	
Sb06g025640	Aluminium activated malate transporter	
Sb09g030160	[Fructose-bisphosphate aldolase]-lysine N-methyltransferase.	
Sb04g002870	similar to Putative uncharacterized protein	
Sb06g000660	Heat shock protein 90	
Sb06g019970	Telomeric repeat binding protein 1	TP
Sb06g030650	Ring finger domain	
Sb09g028720	Chlorophyll A-B binding protein	
Sb09g029180	Beta catenin-related armadillo repeat-containing	

Table 13. List of candidate genes identified in the study based on proximity to the significant markers identified through GWAS analysis and their description or function obtained from phytozome (www.phytozome.net) database under frost stress

Gene ID	Annotation	Trait (under frost)
Sb07g004120	similar to Putative TPA: Cgi67 serine protease	3-DA
Sb03g013650	SRF-type transcription factor	
Sb04g038490	RING zinc finger protein-like	
Sb07g019860	Histone H3 (Lys9) methyltransferase SUV39H1/Clr4, required for transcriptional silencing	
Sb06g033860	Peroxidase (luteolin triglucuronide degradation)	DPPH
Sb04g007700	Dihydrolipoamide S-acetyltransferase (superpathway of cytosolic glycolysis (plants), pyruvate dehydrogenase and TCA cycle)	
Sb04g021810	similar to Cell wall invertase (sucrose degradation V (sucrose α-glucosidase))	
Sb08g004360	transcription factor gata (gata binding factor)	
Sb08g023030	Leucine-rich repeat-containing protein (NB-ARC domain)	TP
Sb06g021900	UDP-glucoronosyl and UDP-glucosyl transferase	
Sb04g025660	B-box zinc finger	
Sb02g024480	Putative uncharacterized protein P0431A03.17	

Chapter 4

Genome wide association study reveals novel genes associated with epi-cuticular wax biosynthesis and transport in sorghum

Abstract

Sorghum is one of the most important cereal crops mainly used for food, feed and fodder. It accumulates epi-cuticular wax (EW) or bloom in plant surfaces such as leaves, sheaths, and culm. EW reduces the non-transpiration water loss and protects the plant from severe drought stress and also imparts resistance against various insect pests. We present here results from the analysis of epi-cuticular wax (EW) content of 387 diverse sorghum accessions and its genome-wide association study (GWAS). EW content in sorghum leaves ranged from 0.1 mg cm^{-2} to 29.7 mg cm^{-2} with a mean value of 5.1 mg cm^{-2}. GWAS using 265,487 single nucleotide polymorphisms (SNPs) identified thirty-seven putative genes that were associated ($p < 9.89E-06$) with EW biosynthesis and transport in sorghum; out of these, Sobic.002G310400 (3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal), Sobic.004G154200 (Ankyrin repeats), and Sobic.003G004500 (bHLH-MYC and R2R3-MYB transcription factors N-terminal) have been shown to be involved in EW biosynthesis; Sobic.001G447300 (ABC transporter), Sobic.004G089400 (Lipid exporter ABCA1 and related proteins, ABC superfamily), Sobic.002G311200 (Multidrug resistance protein), Sobic.001G447200 (Inositol 1, 3, 4-trisphosphate 5/6-kinase), and Sobic.005G217500 (Cytochrome p450) are involved in EW regulation or transport in sorghum. GWAS results from this study demonstrate the potential for genetic manipulation of EW content in sorghum and other grain cereals for better adaptation to biotic and abiotic stresses.

Keywords: ABC transporters, ACP synthase, Epi-cuticular wax, GWAS, SNPs, and Very long chain fatty acids.

Abbreviations

ABC transporters	- ATP Binding Cassette transporters
ACP	- Acyl Carrier Protein
BLMC	- BLooM Cuticle
CMLM	- Compressed Mixed Linear Model
EW	- Epi-cuticular Wax
FAS	- Fatty Acid Synthase
GAPIT	- Genomic Association and Prediction Integrated Tool
GWAS	- Genome Wide Association Study
ICRISAT	- International Crops for the Semi-Arid Tropics
LD	- Linkage Disequilibrium
LTPs	- Lipid Transfer Proteins
MATE	- Multi-antimicrobial extrusion protein
MNC	- ICRISAT Mini-core Collection
PCA	- Principal Component Analysis
Q-Q plot	- Quantile-Quantile plot
SAP	- Sorghum Association Panel
SNPs	- Single Nucleotide Polymorphisms
SUPER	- Settlement of MLM under Progressively Exclusive Relationship
VLCFAs	- Very Long Chain Fatty Acids

Introduction

Sorghum is an important cereal crop which is mainly used for food, feed, and forage (Reddy et al., 2012). Recently it has also been exploited as a viable source for bio-energy purposes (Mathur et al., 2017). Sorghum is a drought resilient crop and offers a model system to study role of waxes in this stress. The cork cells of sorghum secrete long filaments of epi-cuticular wax (EW) structures (Jenks et al., 1994). EW provides the outermost barrier between the plant and environment (Kunst and Samuels, 2003; Samuels et al., 2008; Kunst and Samuels, 2009). Apart from water loss, EW has been shown to impart tolerance to a number of other environmental stresses, including freezing (Thomas and Barber, 1974), insects (Edwards, 1982; Eigenbrode, 2004), pathogens (El-Otmani et al., 1989), mechanical damage (Eglinton and Hamilton, 1967), air pollutants and acid rain (Percy and Baker, 1990), excessive ultraviolet radiation (Eglinton and Hamilton, 1967; Reicosky and Hanover, 1978), heat (Jefferson et al., 1989), and drought (Blum, 1975; Bengston et al., 1978). However, reduced wax (sparse bloom) or no wax (bloomless) sorghum plants were also shown to exhibit non-preference effect to greenbugs (Peiretti et al., 1980) and resistance to sheath blight disease (Jenks et al., 1994).

The proportion of the major wax composition tends to vary among tissues in sorghum (Avato et al., 1984). EW deposition is highly visible on the abaxial side of the leaf blade, culm, and peduncle of sorghum (Burow et al., 2009). In sorghum, the EW production begins to appear at the seedling stage and reaches at peak in-between pre-flowering to the soft-dough stage. The wax pathway starts from precursors of fatty acids that proposed to be derived from *de novo* synthesis in plastids (Kunst and Samuels, 2003; Shepherd and

Griffiths, 2006). The *de novo* fatty acid biosynthesis is catalyzed by a series of enzymatic steps, collectively referred to as fatty acid synthase (FAS). The long chain fatty acids (C_{16}/C_{18}) in plastids get elongated into very long chain fatty acids (VLCFAs) by fatty acid elongase (FAE) reactions. FAE consists of four enzymatic steps, which includes two rounds of reduction, one round of condensation and one round of dehydration (Bernard and Joubes, 2013). VLCFAs (mostly C_{30}s) are then transported to endoplasmic reticulum (ER). From there, all the major wax components such as wax esters and ketones are produced by VLCFAs by two different pathways; alcohol-forming pathway (accounts 20% of waxes) and alkane-forming pathway (accounts 80% of waxes). The synthesized wax components from ER to plasma membrane (PM) are transported via osmiophilic globules present in the golgi bodies. Then the different wax constituents are transported from PM to cell wall through ABC transporters (Pighin et al., 2004; Yazaki, 2006). From cell wall to cuticle, the wax components are transported by lipid transfer proteins (LTPs) (Jacq et al., 2017).

There has been a long-standing interest to characterize the genetics of EW pathway in sorghum. Sorghum variants for EW production were first reported by Ayyangar and Ponnaiya (1941) and designated either bloomless (lacking visible epi-cuticular wax) or sparse bloom (reduced visible epi-cuticular wax). Allelism tests indicated that each of these phenotypes was determined by individual homozygous recessive genes designated *bmbm* (bloomless) and *hh* (sparse bloom). Peterson et al. (1982) performed allelism tests on four sparse bloom and five bloomless mutants in sorghum. Two loci *bm1*, and *bm2* were shown to control the expression of bloom, and at least three loci *h1*, *h2*, and *h3* governed the expression of sparse bloom phenotype. A major QTL *Dw4* for epi-cuticular wax synthesis

was identified from a bi-parental cross between 296B and IS18551(Madhusudhana and Patil, 2012). Another locus *BLMC* (BLooMCuticle) responsible for the production of profuse wax in sorghum was identified from a population derived from BTx623 (profuse wax donor) and KFS2021 (sparse wax) (Burow et al. (2009). Further allelism test revealed *BLMC* is allelic to *Bm2* (Punnuri et al., 2017). A genomic region associated with bloom inhibition (bloomless production) have been identified from the gamma irradiated mutant populations (Mizuno et al., 2013). It was reported that the inhibition of wax secretion was due to the inversion of Sb06g023280 gene. Further, gene annotation revealed that, Sb06g023280 belongs to a family of ABC transporters. Recently, GDSL-like lipase/acylhydrolase (Sobic.001G269200) was identified as a probable candidate gene for EW production in sorghum (Jiao et al., 2017; Uttam et al., 2017). Additionally, sorghum EMS mutant lines which harbors defect in VLCFAs pathway have been identified by Jiao et al. (2016). These bloomless mutant (*blm*) lines are valuable resources for cross-validation and to deduce the function of wax pathway genes in sorghum and other crop species.

Although some key genes involved in the biosynthesis of epi-cuticular wax (EW) have been identified in sorghum and model plant species, the complete understanding of the EW pathway remains elusive in sorghum. Therefore, our objective of this study was to assess the EW variability present in the large genetic pool and to map gene/genic regions associated with the EW production in sorghum through genome wide association mapping.

Materials and methods

Plant materials and genotype data

Global sorghum accessions from sorghum association panel (SAP) and ICRISAT mini-core (MNC) collection were used in the study. The SAP germplasm collections were received from GRIN germplasm network (www.ars-grin.gov) and MNC were received from ICRISAT, India via an MTA through Dr. Wilfred Vermerris, University of Florida, Gainesville. The germplasm accessions used in this study are geographically and racially diverse, representing the major races of kafir, durra, caudatum, guinea, bicolor and working groups of sorghum. A total of 265,487 SNPs (http://www.morrislab.org/data) developed through genotype-by-sequencing (GBS) which corresponds to variation in 27,412 annotated genes in sorghum was used in this study (Morris et al., 2013). GWAS studies conducted using these panels and genotypic data have previously successfully mapped the QTL regions for grain quality (Sukumaran et al., 2012), anthracnose resistance (Upadhyaya et al., 2013), stalk rot (Adeyanju et al., 2015), grain polyphenol concentrations (Rhodes et al., 2014), agro-climatic traits (Morris et al., 2013), and cold and heat stress (Chopra et al., 2017).

Growth conditions and leaf sampling

Sorghum panels were grown in summers of 2014 and 2015 at the Penn State agronomy farm, Rock Springs, Pennsylvania. Photosynthetically active leaves (third leaf from top) were collected as triplicates for each genotype at soft dough stage (75 DAP). Gravimetric determination of wax accumulation in leaves was conducted using 8 leaf discs (0.8 cm diam.) from each diversity lines. The samples were transported to the laboratory on ice,

and epicuticular wax was extracted according to previously described method (Ebercon et al., 1977). Epicuticular wax was extracted using 2 mL of gas-chromatography grade chloroform by swirling the discs in the solvent for 30 seconds. Measurement of gravimetric amount of wax was performed using a balance with sensitivity of 0.1 mg (Mettler Toledo model AB104-S, Mettler Toledo Inc., Columbus, OH).

Association analyses

Genome-wide association analysis was performed using CMLM in combination with SUPER model to determine marker-trait associations. The compressed mixed linear model (CMLM) involves genetic marker–based kinship matrix modeling of random effects used jointly with population structure estimated by principal components analysis (PCA) to model fixed effects (Zhang et al., 2010). SUPER model extracts a small subset of SNPs from CMLM and uses them in FaST-LMM method (Wang et al., 2014). This method increases statistical power and is computationally advantageous. The compression level and optimal number of principal components that adequately explain population structure were previously determined by the genomic association and prediction integrated tool (GAPIT) (Lipka et al., 2012). Log quantile-quantile (QQ) p-value plots for 265,487 single SNP tests of association implied that there were few systematic sources of spurious association using CMLM with SUPER model. As several studies have suggested varying number of LD blocks in sorghum, LD range of 3-5 kb was used for the analysis (Morris et al., 2013).

Genes in QTL regions

Associated SNPs were scanned for the nearby genes in the genomic coordinates using Gramene biomart (www.gramene.org) and Phytozome v.10 (www.phytozome.net). Closest genes in the 5 kp upstream or downstream region were selected and the respective annotations were obtained for each of these genes.

Statistical Analysis

The total measurements included two years field × 387 lines × 3 replicates for a total of 2328 assays. Covariation was determined with Pearson's correlation co-efficient. P values of < 0.05 were considered statistically significant. Statistical data analysis was conducted with R package (R-Core-Team, 2013). Data obtained were expressed as mean ± standard deviation.

Results

Genetic variation for epi-cuticular wax

The significant amount of variability for epi-cuticular wax was observed among the genotypes in SAP and MNC; Figure 15 indicates the frequency of wax variability. The genotype IS 30450 accumulated highest amounts of wax (29.7 mg cm^{-2}), and the genotype PI 655992 accumulated very low wax level (0.1 mg cm^{-2}), with an overall mean of 5.1 mg cm^{-2} across panels. Genotypes with high EW are particularly useful in breeding for abiotic (drought) and biotic (insect pest) tolerant crop varieties for better adaptation to varied agro-climatic conditions. Among the five races and ten intermediate races, durra exhibited high level of EW content with a mean of 13.6 mg cm^{-2}, followed by guinea-bicolor (12.3 mg cm^{-2}), kafir-caudatum (12.1 mg cm^{-2}), bicolor (10.8 mg cm^{-2}), guinea (10.7 mg cm^{-2}) and

caudatum (10.6 mg cm^{-2}). Among the geographical distribution, the genotypes that came from tropically adapted environments had more EW content compared to the genotypes evolved or photo-period converted for temperate adaptation. The genotypes from semi-arid regions of the world exhibited higher EW content with a mean value of 12.9 mg cm^{-2} (Sudan), 12.6 mg cm^{-2} (Nigeria), 11.6 mg cm^{-2} (India), 9.9 mg cm^{-2} (Ethiopia) and 9.9 mg cm^{-2} (Uganda) compared to the temperate region germplasm accessions (United States of America, with a mean of 4.5 mg cm^{-2}). Both SAP and MNC exhibited greater amount of diversity for EW, therefore these panels could be exploited to map the genes/genic regions responsible for the EW biosynthesis and transport in sorghum.

GWAS analysis and identification of candidate genes

Significant positive association was observed between the SNP markers and EW content in sorghum (Fig. 16), with a moderate level of heritability (32.4%) (Fig. 17). GWAS identified 37 significant SNPs with the p-values less than 9.89E-06 (Fig. 18; Table 14), and the candidate genes were recognized within 5 kb distance (up and downstream) from the SNPs. The Chr05 and Chr03 regions were shown to have significant enrichment of SNPs for EW biosynthesis (Fig. 18). The list of putative candidate genes and their position are outlined in Table 14 and 15. Among the genes identified, following are the most probable candidates for biosynthesis, secretion and transport of EW. These genes are Sobic.001G447300 (ABC transporter), Sobic.002G310400 (3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal), Sobic.004G154200 (Ankyrin repeats), Sobic.003G004500 (bHLH-MYC and R2R3-MYB transcription factors N-terminal), Sobic.002G311200 (Multidrug resistance protein), Sobic.001G447200 (Inositol 1,3,4-

trisphosphate 5/6-kinase), Sobic.004G089400 (Lipid exporter ABCA1 and related proteins, ABC superfamily), and Sobic.005G217500 (Cytochrome p450).

Discussion

GWAS was carried out using sorghum mapping panels (SAP and MNC) to identify key genes involved in the wax biosynthesis and transport. The wax composition and accumulation in plants varies according to their response to biotic and abiotic stresses (Eigenbrode and Espelie, 1995). Several wax genes are upregulated during plant-stress cross talk (Xue et al., 2017). Particularly, wax genes are concurrently upregulated with defense genes such as cytochrome p450, WRKY and MYB transcription factors, and flavonoid pathway genes such as chalcone synthase during plant-pathogen or insect interactions. In our study, we have identified several 'R' genes and transcription factors as candidates for EW biosynthesis in sorghum. Here we discuss the role of identified candidate genes in wax pathway.

Among the candidate genes identified, the Sobic.001G447300 (ABC transporter) and Sobic.004G089400 (Lipid exporter ABCA1 and related proteins, ABC superfamily) are the most probable candidate genes for EW secretion and transport in sorghum (Table 1). Wax biosynthesis initiates in the endoplasmic reticulum (ER) and then transported plasma membrane (PM) through golgi complex and white brown complex (WBC) transporters (Bird et al., 2007; Bird, 2008; Samuels et al., 2008; McFarlane et al., 2014). For transport from PM to epidermis, it is exported via ABC transporters and lipid transfer proteins (Samuels et al., 2008). In our GWAS analysis, we particularly found enrichment of several

ABC family lipid exporter proteins (Table 14). Adrenoleukodystrophy (ALD) is x-linked neurodegenerative disease in humans, which is due to the defect in ABC transporter pathway. ABC transporter is responsible for the transport of saturated VLFC (very long fatty acid chain) into peroxisome for β-oxidation (Powell et al., 1975). Arabidopsis *cer5* mutant which lacks ABC transporter found to localize the wax constituents in to the vacuole similar to ALD disorder. Corresponding wild type CER5 functions normally due to the presence of functional copies of ABC transporters, therefore it transports lipid from ER to epidermis or vacuole to epidermis (Pighin et al., 2004). This study suggests that, indeed a family of ABC transporters are necessary for wax export from vacuole to the outer membrane in plants. Furthermore, several sub-family domains of ABC transporters have been shown to play a major role in phospholipid translocation, cholesterol efflux and transport of long chain acyl-COAs (Gottesman and Pastan, 1993; Mosser et al., 1994; Liu et al., 1999; Berge et al., 2000). Also, Xu et al. (2003) reported that an ABC transporter family gene AtTGD1 is involved in the inter-organelle lipid transfer in Arabidopsis. Mizuno et al. (2013) reported an ABC transporter gene, Sb06g023280 is responsible for epi-cuticular wax biosynthesis in sorghum. All these evidences suggest that lipid export is happening via ABC transporters and it is essential for transporting lipids from inner cellular membranes to outer most layer of the plant.

GWAS mapped the gene Sobic.002G310400, which encodes for 3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal (Abbadi et al., 2000) as a top candidate for EW biosynthesis in sorghum (Table 1 and 2). Fatty acid synthesis (FAS) is occurring via two distinct pathways (FAS I and II). ACP synthase is the prime catalyst for moderating the

FAS II pathway in mitochondria, bacteria, plants and parasites. The 3-Oxoacyl-[acyl carrier protein (ACP)] synthase III catalyzes the first condensation step within the FAS II pathway, using acetyl-CoA as the primer and malonyl-ACP as the acceptor. The resultant oxoacyl-ACP formed by this reaction subsequently enters the elongation cycle, where the acyl chain is progressively lengthened by the combined activities of several enzymes. Furthermore, the GO term GO:0008610 is associated with the ACP synthase (Sobic.002G310400), encodes for fatty acid biosynthetic process in plants and bacteria (Table 2). Thus, Sobic.002G310400 involved in the formation of very long chain fatty acids (VLCFs). Further VLCFs produce the downstream products in wax biosynthesis. Hence, Sobic.002G310400 is another strong candidate we have identified in this GWAS study for EW biosynthesis in sorghum.

Ankyrin repeats (Sobic.004G154200) were shown to have specific affinity towards lipid binding domains, especially the lipid domains (monogalactosyldiacylglycerol and phosphatidylglycerol) of plastids (Kim et al .2014). The initial synthesis of FAS occurs in either plastids or ER. It has been established that ankryin repeats are involved in one or another way in lipid biosynthesis. Sobic.003G004500 (bHLH-MYC and R2R3-MYB transcription factors N-terminal) also involved in EW regulation in Arabidopsis (At5G62470). Another important candidate (Sobic.002G311200) is a multidrug resistance protein (MATE) involved in the EW regulation. It transports lipids from cytoplasm to plasma membrane. Sobic.001G447200 (Inositol 1, 3, 4-trisphosphate 5/6-kinase) is involved in the formation of lipid layer in outer most plant tissues (Tang et al., 2013) and Sobic.005G217500 (Cytochrome p450) is also reported to have a role in EW biosynthesis

and regulation for pollen fertility (Li et al., 2010). GDSL-like lipase/acylhydrolase has been reported to be involved in production of VLCFAs in the wax biosynthesis pathway. Uttam et al. (2017) and Jiao et al. (2017) mapped GDSL-like lipase locus as a probable candidate for EW in sorghum and mutations in the GDSL-like lipase gene reported to reduce the synthesis of epi-cuticular wax in sorghum (Jiao et al., 2016; Jiao et al., 2017). We have also mapped the GDSL-like lipase/ acylhydrolase (Sobic.008G133200) gene, but it was not significant. However, we imply that Sobic.008G133200 is another strong candidate for EW biosynthesis in sorghum.

We have also observed upregulation of several stress (biotic and abiotic) responsive genes. Drought responsive genes such as Sobic.003G273800 (DUF4005), Sobic.005G225500 (DUF1618), Sobic.010G065300 (DUF1086), and Sobic.005G217900 (DUF4409); defense genes such as Sobic.003G296300 (WRKY transcription factor), Sobic.004G153100 (NB-ARC domain), and Sobic.005G217500 (Cytochrome P450); flavonoid genes such as Sobic.002G195700 (Cinnamyl-alcohol dehydrogenase), and Sobic.010G181100 (Kynurenine 3-monooxygenase and related flavoprotein monooxygenases) were detected as top candidates (Table 1 and 2). This analyses thus revealed that the coordinated regulation of EW along with other stress responsive genes in sorghum.

Conclusions

Epi-cuticular wax (EW) content in sorghum leaves of 387 diverse cultivars varied widely. Genome-wide association analysis identified thirty-seven genetic associations for EW content, which have the potential as molecular markers to manipulate EW content in

sorghum. This study identified several genes that play a major role in wax biosynthesis and transport in sorghum. Especially, the major regulators like ABC transporters, acyl carrier protein synthase and multidrug resistance protein as candidates. We have also observed the upregulation of several defense genes such as WRKY transcription factor, NB-ARC domain, Cytochrome P450 and flavonoid genes such as cinnamyl-alcohol dehydrogenase, and flavoprotein monooxygenases.

Acknowledgments

We thank GRIN germplasm network and Dr. Wilfred Vermerris (University of Florida) for providing the sorghum mapping panels to conduct this study. We also want to thank Mr. Scott Harkcom, agronomy farm manager at penn state rock springs for facilitating us in field preparation and necessary managements. Thanks to Jin Cui, Debamalya Chatterje, Iffa Gaffoor and Bin Liu for helping us in sample collection and phenotyping.

Literature Cited

Abbadi, A., Brummel, M., Schutt, B.S., Slabaugh, M.B., Schuch, R., and Spener, F. (2000). Reaction mechanism of recombinant 3-oxoacyl-(acyl-carrier-protein) synthase III from Cuphea wrightii embryo, a fatty acid synthase type II condensing enzyme. *Biochem Journal* 345**,** 153–160.

Adeyanju, A., Little, C., Yu, J., and Tesso, T. (2015). Genome-Wide Association Study on Resistance to Stalk Rot Diseases in Grain Sorghum. *G3 (Bethesda)* 5(6)**,** 1165-1175. doi: 10.1534/g3.114.016394.

Avato, P.N., Bianchi, G., and Mariani, G. (1984). Epicuticular waxes of sorghum and some compositional changes with plant age. *Phytochemistry* 23(12)**,** 2843-2846.

Ayyangar, G.N., and Ponnaiya, B.W.X. (1941). The occurrence and inheritance of a bloomless sorghum. *Current Science* 10**,** 408-409.

Bengston, C., Larsson, S., and Liljenberg, C. (1978). Effects of water stress on cuticular transpiration rate and amount and composition of epicuticular wax in seedlings of six oat varieties. *Physiologia Plantarum* 44(4)**,** 319-324.

Berge, K.E., Tian, H., Graf, G.A., Yu, L., Grishin, N.V., Schultz, J., et al. (2000). Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* 290(5497)**,** 1771-1775. doi: 10.1126/science.290.5497.1771.

Bernard, A., and Joubes, J. (2013). Arabidopsis cuticular waxes: advances in synthesis, export and regulation. *Prog Lipid Res* 52(1)**,** 110-129. doi: 10.1016/j.plipres.2012.10.002.

Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R., et al. (2007). Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *Plant J* 52(3)**,** 485-498. doi: 10.1111/j.1365-313X.2007.03252.x.

Bird, D.A. (2008). The role of ABC transporters in cuticular lipid secretion. *Plant Science* 174(6)**,** 563-569. doi: 10.1016/j.plantsci.2008.03.016.

Blum, A. (1975). Effect of the Bm gene on epicuticular wax deposition and the spectral characteristics of sorghum leaves. *SABRAO Journal* 7**,** 45-52.

Burow, G.B., Franks, C.D., Acosta-Martinez, V., and Xin, Z. (2009). Molecular mapping and characterization of BLMC, a locus for profuse wax (bloom) and enhanced cuticular features of Sorghum (Sorghum bicolor (L.) Moench.). *Theor Appl Genet* 118(3)**,** 423-431. doi: 10.1007/s00122-008-0908-y.

Chopra, R., Burow, G., Burke, J.J., Gladman, N., and Xin, Z. (2017). Genome-wide association analysis of seedling traits in diverse Sorghum germplasm under thermal stress. *BMC Plant Biol* 17(1)**,** 12. doi: 10.1186/s12870-016-0966-2.

Ebercon, A., Blum, A., and Jordan, W.R. (1977). A rapid colorimetric method for epicuticular wax content of sorghum leaves. *Crop Science* 17**,** 179-180.

Edwards, P.B. (1982). Do waxes on juvenile Eucalyptus leaves provide protection from grazing insects? *Australian Journal of Ecology* 7**,** 347-352.

Eglinton, G., and Hamilton, R.J. (1967). Leaf epicuticular waxes. *Science* 156**,** 1322-1335.

Eigenbrode, S.D. (2004). The effects of plant epicuticular waxy blooms on attachment and effectiveness of predatory insects. *Arthropod Struct Dev* 33(1)**,** 91-102. doi: 10.1016/j.asd.2003.11.004.

Eigenbrode, S.D., and Espelie, K.E. (1995). Effects of plant epicuticular lipids on insect herbivores. *Annual Review of Entomology* 40**,** 171–194.

El-Otmani, M., Arpaia, M.L., Jr, C.W.C., Jr, J.E.P., and O'Connell, N.V. (1989). Developmental changes in 'Valencia' orange fruit epicuticular wax in relation to fruit position on the tree. *Scientia Horticulturae* 41**,** 69-81.

Gottesman, M.M., and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annual Review of Biochemistry* 62**,** 385-427.

Jacq, A., Pernot, C., Martinez, Y., Domergue, F., Payre, B., Jamet, E., et al. (2017). The Arabidopsis Lipid Transfer Protein 2 (AtLTP2) Is Involved in Cuticle-Cell Wall Interface Integrity and in Etiolated Hypocotyl Permeability. *Front Plant Sci* 8**,** 263. doi: 10.3389/fpls.2017.00263.

Jefferson, P.G., Johnson, D.A., and Asay, K.H. (1989). Epi cuticular wax production, water status and leaf temperature in Triticeae range grasses of contrasting visible glaucousness. *Canadian Journal of Plant Science* 69**,** 513-519.

Jenks, M.A., Rich, P.J., and Ashworth, E.N. (1994). Involvement of cork cells in the secretion of epicuticular wax filaments on Sorghum bicolor (L.) Moench. *International Journal of Plant Sciences* 155(5)**,** 506-518.

Jiao, Y., Burke, J., Chopra, R., Burow, G., Chen, J., Wang, B., et al. (2016). A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses. *Plant Cell* 28(7)**,** 1551-1562. doi: 10.1105/tpc.16.00373.

Jiao, Y., Burow, G., Gladman, N., Acosta-Martinez, V., Chen, J., Burke, J., et al. (2017). Efficient Identification of Causal Mutations through Sequencing of Bulked F 2 from Two Allelic Bloomless Mutants of Sorghum bicolor. *Front Plant Sci* 8**,** 2267. doi: 10.3389/fpls.2017.02267.

Kunst, L., and Samuels, A.L. (2003). Biosynthesis and secretion of plant cuticular wax. *Progress in Lipid Research* 42**,** 51-80.

Kunst, L., and Samuels, L. (2009). Plant cuticles shine: advances in wax biosynthesis and export. *Curr Opin Plant Biol* 12(6)**,** 721-727. doi: 10.1016/j.pbi.2009.09.009.

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., et al. (2010). Cytochrome P450 family member CYP704B2 catalyzes the {omega}-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *Plant Cell* 22(1)**,** 173-190. doi: 10.1105/tpc.109.070326.

Lipka, A.E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P.J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28(18)**,** 2397-2399. doi: 10.1093/bioinformatics/bts444.

Liu, L.X., Janvier, K., Lecellier, V.B., Cartier, N., Benarous, R., and Aubourg, P. (1999). Homo and heterodimerization of peroxisomal ATP-binding cassette half-transporters. *The Journal of Biological Chemistry* 274(46)**,** 32738–32743.

Madhusudhana, R., and Patil, J.V. (2012). A major QTL for plant height is linked with bloom locus in sorghum [Sorghum bicolor (L.) Moench]. *Euphytica* 191**,** 259-268.

Mathur, S., Umakanth, A.V., Tonapi, V.A., Sharma, R., and Sharma, M.K. (2017). Sweet sorghum as biofuel feedstock: recent advances and available resources. *Biotechnol Biofuels* 10**,** 146. doi: 10.1186/s13068-017-0834-9.

McFarlane, H.E., Watanabe, Y., Yang, W., Huang, Y., Ohlrogge, J., and Samuels, A.L. (2014). Golgi- and trans-Golgi network-mediated vesicle trafficking is required for wax secretion from epidermal cells. *Plant Physiol* 164(3)**,** 1250-1260. doi: 10.1104/pp.113.234583.

Mizuno, H., Kawahigashi, H., Ogata, J., Minami, H., Kanamori, H., Nakagawa, H., et al. (2013). Genomic inversion caused by gamma irradiation contributes to downregulation of a WBC11 homolog in bloomless sorghum. *Theoretical and Applied Genetics* 126**,** 1513-1520.

Morris, G.P., Ramu, P., Deshpande, S.P., Hash, C.T., Shah, T., Upadhyaya, H.D., et al. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc Natl Acad Sci U S A* 110(2)**,** 453-458. doi: 10.1073/pnas.1215985110.

Mosser, J., Lutz, Y., Stoeckel, M.E., Sarde, C.O., Kretz, C., Douar, A.M., et al. (1994). The gene responsible for adrenoleukodystrophy encodes a peroxisomal membrane protein. *Hum Mol Genet* 3(2)**,** 265-271.

Peiretti, R.A., Amini, I., Weibel, D.E., Starks, K.J., and McNew, R.W. (1980). Relationship of "Bloomless" (bm bm) sorghum to greenbug resistance. *Crop Science* 20**,** 173-176.

Percy, K.E., and Baker, E.A. (1990). Effects of simulated acid rain on epicuticular wax production, morphology, chemical composition and on cuticular membrane thick ness in two clones of Sitka spruce [Picea sitchensis (Bong.) Carr.]. *New Phytologist* 116**,** 79-87.

Peterson, G.C., Krittika, S., and Weibel, D.E. (1982). Inheritance of some bloomless and sparse-bloom mutants in sorghum. *Crop Science* 22**,** 63-67.

Pighin, J.A., Zheng, H., Balakshin, L.J., Goodman, I.P., Western, T.L., Jetter, R., et al. (2004). Plant cuticular lipid export requires an ABC transporter. *Science* 306(5696)**,** 702-704. doi: 10.1126/science.1102331.

Powell, M.B.H., Tindall, R., Schultz, P., Paa, D., O'Brien, J., and Lampert, P. (1975). Adrenoleukodystrophy electron microscopic findings. *Arch Neurol* 32(4)**,** 256-260. doi: 10.1001/archneur.1975.00490460066009.

Punnuri, S., Harris-Shultz, K., Knoll, J., Ni, X., and Wang, H. (2017). The Genes and that Affect Epicuticular Wax Deposition in Sorghum are Allelic. *Crop Science* 57(3). doi: 10.2135/cropsci2016.11.0937.

R-Core-Team (2013). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

Reddy, B.V.S., Reddy, P.S., Sadananda, A.R., Dinakaran, E., Kumar, A.A., Deshpande, S.P., et al. (2012). Postrainy season sorghum: Constraints and breeding approaches. *Journal of SAT Agricultural Research* 10**,** 1-12.

Reicosky, D.A., and Hanover, J.W. (1978). Physiological effects of surface waxes. I. Light reflectance for glaucous and non-glaucous Picea pungens. *Plant Physiology* 62**,** 101- 104.

Rhodes, D.H., Hoffmann, L., Jr., Rooney, W.L., Ramu, P., Morris, G.P., and Kresovich, S. (2014). Genome-wide association study of grain polyphenol concentrations in global sorghum [Sorghum bicolor (L.) Moench] germplasm. *J Agric Food Chem* 62(45)**,** 10916-10927. doi: 10.1021/jf503651t.

Samuels, L., Kunst, L., and Jetter, R. (2008). Sealing plant surfaces: cuticular wax formation by epidermal cells. *Annu Rev Plant Biol* 59**,** 683-707. doi: 10.1146/annurev.arplant.59.103006.093219.

Shepherd, T., and Griffiths, D.W. (2006). The effects of stress on plant cuticular waxes. *New Phytol* 171(3)**,** 469-499. doi: 10.1111/j.1469-8137.2006.01826.x.

Sukumaran, S., Xiang, W., Bean, S.R., Pedersen, J.F., Kresovich, S., Tuinstra, M.R., et al. (2012). Association Mapping for Grain Quality in a Diverse Sorghum Collection. *The Plant Genome Journal* 5(3). doi: 10.3835/plantgenome2012.07.0016.

Tang, Y., Tan, S., and Xue, H. (2013). Arabidopsis inositol 1,3,4-trisphosphate 5/6 kinase 2 is required for seed coat development. *Acta Biochim Biophys Sin (Shanghai)* 45(7)**,** 549-560. doi: 10.1093/abbs/gmt039.

Thomas, D.A., and Barber, H.N. (1974). Studies on leaf characteristics of a cline of Eucalyptus urnigera from mount Wellington, Tasmania. II. Reflection, transmission and absorption of radiation. *Australian Journal of Botany* 22(4)**,** 701-707.

Upadhyaya, H.D., Wang, Y.H., Sharma, R., and Sharma, S. (2013). Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. *Theor Appl Genet* 126(6)**,** 1649-1657. doi: 10.1007/s00122-013-2081-1.

Uttam, G.A., Praveen, M., Rao, Y.V., Tonapi, V.A., and Madhusudhana, R. (2017). Molecular mapping and candidate gene analysis of a new epicuticular wax locus in sorghum (Sorghum bicolor L. Moench). *Theor Appl Genet* 130(10)**,** 2109-2125. doi: 10.1007/s00122-017-2945-x.

Wang, Q., Tian, F., Pan, Y., Buckler, E.S., and Zhang, Z. (2014). A SUPER Powerful Method for Genome Wide Association Study. *Plos One* 9(9)**,** e107684. doi: https://doi.org/10.1371/journal.pone.0107684.

Xu, C., Fan, J., Riekhof, W., J.E.Froehlich, and Benning, C. (2003). A permease-like protein involved in ER to thylakoid lipid transfer in Arabidopsis. *The EMBO Journal* 22(10)**,** 2370-2379.

Xue, D., Zhang, X., Lu, X., Chen, G., and Chen, Z.H. (2017). Molecular and Evolutionary Mechanisms of Cuticular Wax for Plant Drought Tolerance. *Front Plant Sci* 8**,** 621. doi: 10.3389/fpls.2017.00621.

Yazaki, K. (2006). ABC transporters involved in the transport of plant secondary metabolites. *FEBS Lett* 580(4)**,** 1183-1191. doi: 10.1016/j.febslet.2005.12.009.

Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat Genet* 42(4)**,** 355-360. doi: 10.1038/ng.546.

Figure Legends

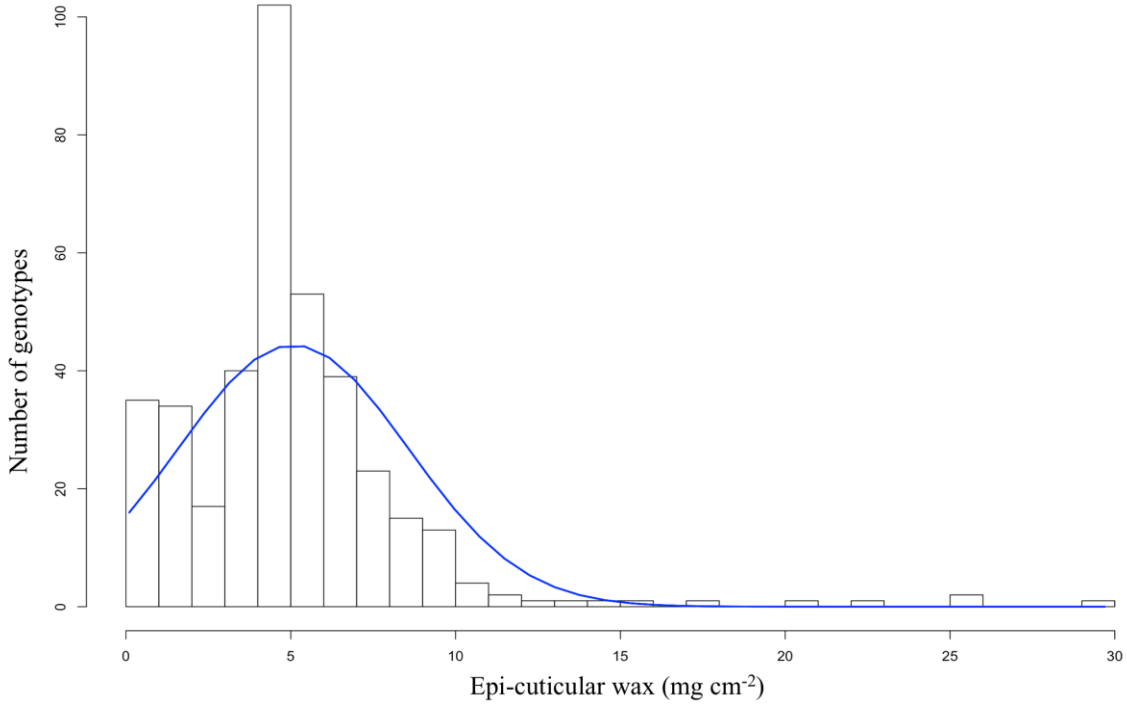

Figure 15. Frequency of epi-cuticular wax (mg cm^{-2}) distribution among genotypes (387) in sorghum diversity panels (sorghum association and ICRISAT mini-core panels).

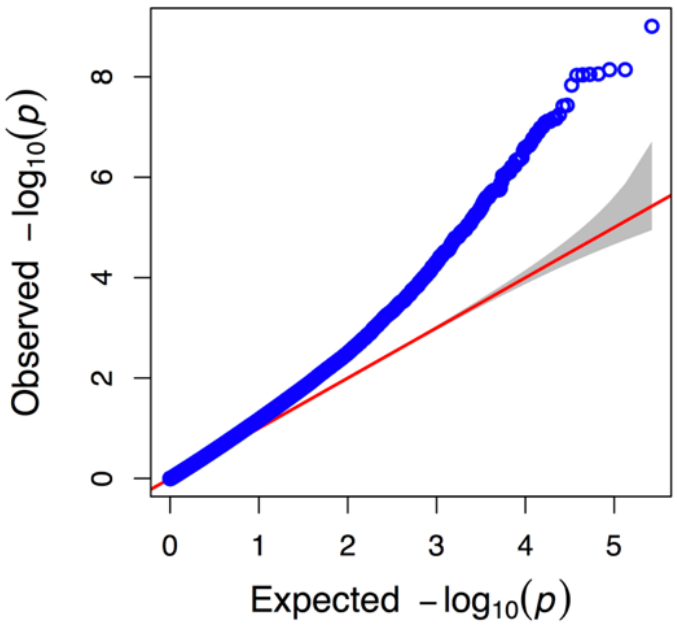

Figure 16. Q-Q plot reveals significant marker trait associations for epi-cuticular wax biosynthesis in sorghum

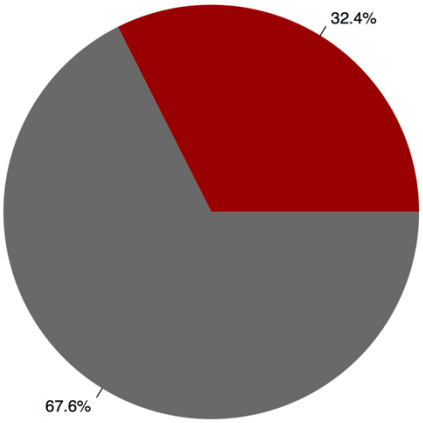

Figure 17. Optimum compression reveals moderate level of heritability for epi-cuticular wax biosynthesis in sorghum

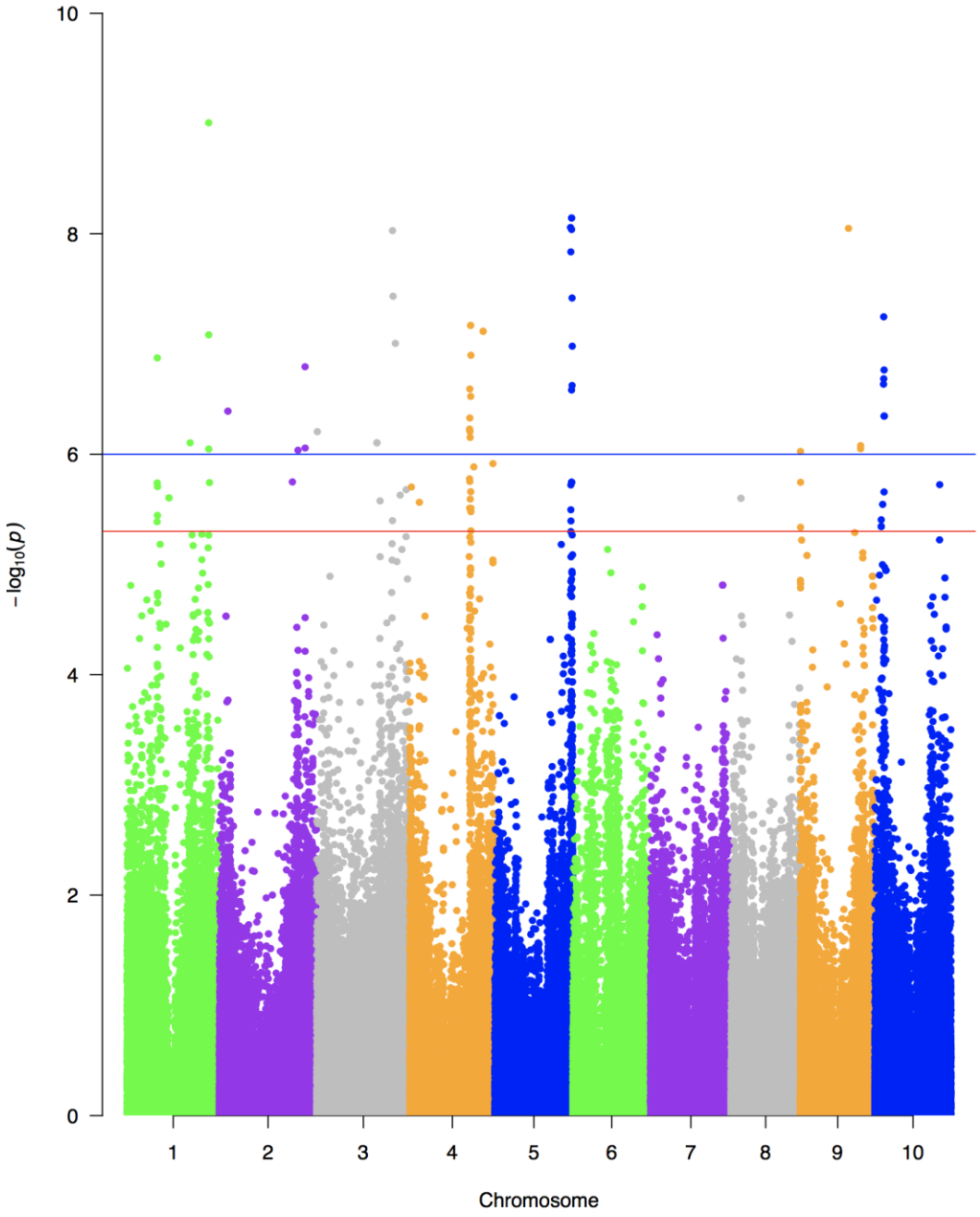

Figure 18. Manhattan plot reveals genome wide significant SNPs associated with epicuticular wax biosynthesis in sorghum.

Tables

Table 14. Priori list of candidate genes for epi-cuticular wax biosynthesis and transport in sorghum

SNP		Allele	P-value	Minor Allele Frequency	R^2 value	FDR Adjusted P-values	Distance from the SNP (bp)	Sorghum Gene ID (Version 3)
Chromosome	Scaffold Position							
1	65299690	G/A	9.48×10^{-10}	0.05	20.70	0.0003	3072	Sobic.001G447100
5	61354951	G/C	7.20×10^{-09}	0.05	19.72	0.0004	7607	Sobic.005G222000
5	60481090	G/C	8.76×10^{-09}	0.06	19.62	0.0004	1723	Sobic.005G213600
3	61008198	T/C	3.69×10^{-08}	0.05	18.92	0.0010	2206	Sobic.003G273800
5	61722619	T/C	3.82×10^{-08}	0.07	18.90	0.0010	7107	Sobic.005G225500
10	7071935	T/C	5.66×10^{-08}	0.05	18.71	0.0014	9259	Sobic.010G083900
4	58590545	T/G	7.66×10^{-08}	0.01	18.57	0.0014	3736	Sobic.004G244600
1	65315947	T/C	8.26×10^{-08}	0.08	18.53	0.0014	5973	Sobic.001G447300
3	62843242	G/A	9.87×10^{-08}	0.02	18.45	0.0015	1917	Sobic.003G296300
5	61870739	T/C	1.05×10^{-07}	0.04	18.42	0.0015	878	Sobic.005G226600
4	48837079	C/T	1.27×10^{-07}	0.05	18.33	0.0018	5538	Sobic.004G156500
1	24186828	C/T	1.34×10^{-07}	0.05	18.30	0.0018	1059	Sobic.001G237400
2	68533966	C/T	1.61×10^{-07}	0.09	18.21	0.0020	1202	Sobic.002G310400
10	7383995	T/C	1.72×10^{-07}	0.06	18.18	0.0021	8628	Sobic.010G087100
10	6935192	T/G	2.32×10^{-07}	0.05	18.04	0.0025	1732	Sobic.010G082100
4	47969201	T/C	2.56×10^{-07}	0.06	17.99	0.0026	1200	Sobic.004G153700
4	48681211	C/A	2.99×10^{-07}	0.06	17.91	0.0028	3821	Sobic.004G156000
10	7472583	G/C	4.51×10^{-07}	0.06	17.72	0.0036	956	Sobic.010G088400
4	47807327	G/A	6.03×10^{-07}	0.05	17.58	0.0044	2786	Sobic.004G153400
4	48079754	G/A	6.23×10^{-07}	0.05	17.56	0.0044	2880	Sobic.004G154200
3	419429	T/C	6.24×10^{-07}	0.03	17.56	0.0044	2320	Sobic.003G004500
4	48219493	G/T	7.04×10^{-07}	0.04	17.51	0.0048	5771	Sobic.004G154300
9	48231938	T/A	8.36×10^{-07}	0.13	17.42	0.0052	851	Sobic.009G126800
2	68595229	C/T	8.79×10^{-07}	0.09	17.40	0.0052	7814	Sobic.002G311200
1	65311952	G/T	8.97×10^{-07}	0.06	17.39	0.0052	5902	Sobic.001G447200
2	58407855	G/A	1.78×10^{-06}	0.37	17.07	0.0081	2209	Sobic.002G195700
9	172074	C/T	1.80×10^{-06}	0.11	17.06	0.0081	3007	Sobic.009G001500

5	60737536	C/T	1.90×10^{-06}	0.05	17.04	0.0081	1495	Sobic.005G216833
4	7640058	T/A	2.73×10^{-06}	0.14	16.87	0.0097	6860	Sobic.004G089400
4	47785237	C/T	3.09×10^{-06}	0.07	16.81	0.0105	3887	Sobic.004G153100
10	5081054	G/A	3.93×10^{-06}	0.06	16.69	0.0127	14138	Sobic.010G065300
5	60827716	C/G	4.03×10^{-06}	0.05	16.68	0.0127	1848	Sobic.005G217500
1	65292389	C/A	5.42×10^{-06}	0.09	16.54	0.0150	15962	Sobic.001G447000
10	51767382	C/T	5.98×10^{-06}	0.10	16.50	0.0160	1563	Sobic.010G181100
9	49961009	C/T	7.83×10^{-06}	0.09	16.37	0.0191	2261	Sobic.009G140500
5	60888751	G/A	8.53×10^{-06}	0.10	16.33	0.0199	4386	Sobic.005G217900
3	60272435	G/A	9.13×10^{-06}	0.04	16.30	0.0204	2397	Sobic.003G265800

Table 15. Candidate gene description and their biological role in epi-cuticular wax biosynthesis and transport in sorghum

Gene name	GO ID	Description	Biological function
Sobic.001G447100	GO:0003950	Poly polymerase catalytic domain containing protein	Wnt signaling pathway
Sobic.005G222000		Importin-7, 8, 11	
Sobic.005G213600		Sulfotransferase	
Sobic.003G273800	GO:0005515	Protein of unknown function (DUF4005)	Protein binding
Sobic.005G225500		Protein of unknown function (DUF1618)	
Sobic.010G083900		Alpha/beta hydrolase family	
Sobic.004G244600	GO:0005975	Glycosyl hydrolase family 9	Carbohydrate metabolic process
Sobic.001G447300	GO:0043190	ABC transporter	Transmembrane transport
Sobic.003G296300	GO:0003700	WRKY transcription factor	
Sobic.005G226600	GO:0008270	Ring finger domain	Zinc ion binding
Sobic.004G156500		Telomeric repeat binding protein	
Sobic.001G237400	GO:0046872	Heavy-metal-associated domain	Metal ion binding
Sobic.002G310400	GO:0016747	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal	Fatty acid biosynthetic process
Sobic.010G087100	GO:0003677	DNA replication licensing factor MCM1	DNA binding
Sobic.010G082100	GO:0042546	Xyloglucan fucosyltransferase	Cell wall biogenesis
Sobic.004G153700		Cotton fibre expressed protein	
Sobic.004G156000		11-beta-hydroxysteroid dehydrogenase	
Sobic.010G088400		Phosphate-induced protein 1 conserved region	
Sobic.004G153400	GO:0005515	Similar to Kelch repeat-containing F-box protein-like	Protein binding
Sobic.004G154200		Ankyrin repeats (3 copies)	
Sobic.003G004500	GO:0046983	bHLH-MYC and R2R3-MYB transcription factors N-terminal	Signal transduction
Sobic.004G154300	GO:0008641	Ubiquitin-activating enzyme E1	Cellular protein modification process
Sobic.009G126800		3'-5' Exonuclease ERI1-related	
Sobic.002G311200	GO:0055085	Multidrug and toxic compound extrusion family protein	Transmembrane transport
Sobic.001G447200	GO:0052726	Inositol 1, 3, 4-trisphosphate 5/6-kinase	Inositol trisphosphate metabolic process
Sobic.002G195700	GO:0055114	Cinnamyl-alcohol dehydrogenase.	Phenylpropanoid biosynthesis
Sobic.009G001500	GO:0006457	Peptidylprolyl isomerase.	

Sobic.005G216833	GO:0051260	Potassium Channel tetramerization domain-containing	
Sobic.004G089400	GO:0016887	Lipid exporter ABCA1 and related proteins, ABC superfamily	
Sobic.004G153100	GO:0043531	NB-ARC domain/Leucine-rich repeat-containing protein/Apoptotic ATPase	
Sobic.010G065300	GO:0003677	DNA helicase	
Sobic.005G217500	GO:0055114	Cytochrome P450	Momilactone A biosynthesis
Sobic.001G447000	GO:0005737	O-methyltransferase	
Sobic.010G181100	GO:0071949	Kynurenine 3-monooxygenase and related flavoprotein monooxygenases	
Sobic.009G140500	GO:0006355	No apical meristem (NAM) protein	
Sobic.005G217900		Domain of unknown function (DUF4409)	
Sobic.003G265800		PPR repeat family	

Chapter 5

Effect of seeding and N fertilization rate on yield and quality of brachytic BMR forage sorghum hybrids

Abstract

Brown mid-rib (BMR) forage sorghum (*Sorghum bicolor* (L.) Moench) silage is a reasonable alternative to corn silage for areas with limited soil moisture. Traditional forage sorghum varieties are tall and prone to lodging with lower forage quality. Brachtyic dwarf BMR forage lines are shorter, lodging resistant and have higher forage quality. Newer, earlier hybrids have expanded the potential adaptation of forage sorghums to more northern areas. A two-year study was conducted during the 2014 and 2015 growing seasons using newly available brachytic dwarf BMR forage hybrids to determine the effects of different seeding rates and N (nitrogen) fertilization rates on forage dry matter (DM) yield and quality for two hybrids. The experimental design was split-split-plot with four replications. In each replication, main plots were two hybrids (AF7202 and AF7401), subplots were two seeding rates (198,000 and 296,400 seeds ha^{-1}), and sub-subplots were two N rates (123 and 168 kg ha^{-1}). DM yield and forage quality parameters were measured for each treatment. We observed significant varietal differences for all the parameters except neutral detergent fiber digestibility (NDFD) in 2015. The early maturity line, AF7202, had higher yields, higher starch and net energy for lactation (NE_L) levels than AF7401. The dwarf line, AF7401, had higher crude protein (CP) and NDFD than AF7202. AF7202 was more responsive to the higher N rate than AF7401. CP was increased with increased N for both varieties. Other forage quality traits were unaffected by N rates. Neither variety responded to an increase in seeding rate. This study showed that the earlier brachytic dwarf forage sorghums such as AF7202, managed with recommended seeding rates and possibly higher N rates have good potential for high forage yield and quality in central PA.

Keywords: Brachytic dwarf, brown mid rib, forage sorghum, forage yield and quality.

Abbreviations

NDFD	- Neutral detergent fiber digestibility
NE_L	- Net energy for lactation
CP	- Crude protein
BMR	- Brown mid-rib
DAP	- Days after planting
ESC	- Ethanol soluble carbohydrates
TDN	- Total digestible nitrogen
NDF	- Neutral detergent fiber
NIRS	- Near infrared reflectance spectroscopy
DMY	- Dry matter yield
FDMY	- Forage dry matter yield
FDMQ	- Forage dry matter quality

Introduction

Forage sorghum has become an alternative forage for dairy producers in several regions in the US. The lower water requirement of forage sorghum (Martin 1930; Merrill et al. 2007; Miron et al. 2007; Lamm, Stone, and O'Brien 2007; Howell et al. 2008) has made it an attractive alternative not only in arid irrigated regions such as Texas, New Mexico and California, but also in the mid-Atlantic region where corn production can be limited on shallow or coarse textured soils. Forage sorghums can also be double cropped following small grains harvested for forage in the Mid Atlantic US, which can add up to 7 t ha^{-1} to the seasonal forage yield. In addition, they provide farms a late spring opportunity for manure spreading following corn planting. Forage sorghums, however, have traditionally suffered from lodging and lower feed quality than corn silage and this has limited their adoption (Bean et al. 2003; Oliver et al. 2005; Howell et al. 2008; Marsalis et al. 2009; Contreras-Govea et al. 2011).

BMR mutants of sorghum were identified by Porter et al. (1978) and this led eventually to commercial hybrids with improved neutral detergent fiber digestibility (NDFD) compared to conventional sorghum (Grant et al. 1995; Oliver et al. 2005; Bean and McCollum 2006). The brachytic dwarf trait was also incorporated into forage sorghum hybrids recently (Liu et al. 2016). The brachytic dwarf plant types have shorter internode lengths and maintain the equal number of leaves as taller sorghum varieties, making them less vulnerable to lodging (Liu et al. 2016). With the combination of brachytic dwarf and BMR traits, new more lodging resistant hybrids with improved fibre digestibility have become commercially available (Alta Seeds 2013).

The initial introductions of brachytic dwarf BMR forage sorghums were late maturity hybrids such as AF7401. This limited the adoption of forage sorghum in some shorter season environments with drought prone soils and they were at risk for frost damage in the fall and had limited double crop potential. Eventually, newer shorter season brachytic BMR lines such as AF7202 became available as an option for some of these regions such as Pennsylvania and New York (Alta Seeds 2013), but the yield and forage quality potential of the earlier hybrids was not well established.

In some trials, yields from the BMR lines have not been as high as the taller forage sorghum lines (Bean and McCollum 2006). A potential way to overcome these lower yields of BMR genotypes could be through more intensive management such as the use of higher N rates or seeding rates, especially with reduced potential for lodging. Current N fertilization recommendations for forage sorghum in PA are 141 kg ha^{-1} for a crop yielding 11.8 t ha^{-1} (Beegle 2016). Seed industry recommendations have been approximately 110 kg ha^{-1} to 132 kg ha^{-1} for forage sorghums in this region (Kings Agriseeds 2017).

Previous research on N rates on forage sorghum is limited. Bean et al. (2003) reported that high amounts of N application do not significantly increase the yields but may increase the incidence of lodging. Marsalis, Angadi, and Contreras-Govea (2010) found no benefit from increasing N rates on conventional or BMR forage sorghums from 106 to 140 kg ha^{-1}. Buxton, Anderson, and Hallam (1999) showed that no benefit on forage sorghum dry matter (DM) yields by increasing the N levels from 140 to 280 kg ha^{-1}. In some trials, the

maximum DM yields of forage sorghum were obtained at 125 kg N ha^{-1} (Beyaert and Roy 2005). Moreover, Ketterings et al. (2007) reported that optimum N levels was between 125 and 145 kg N ha^{-1} for BMR-sorghum sudangrass in the northeastern U.S.

Jahanzad et al. (2013) studied the response of forage sorghum yield and quality in varied seeding and water regimes and showed that increased seeding rates can positively influence forage DM yield. The recommended seeding rate for forage sorghum hybrids from industry is 198,000 seeds ha^{-1} (Kings Agriseeds 2017). According to Sanderson et al. (1994), forage sorghum growers generally believe that high plant densities improve the forage quality via thinner stems and higher leaf percentage. However, Caravetta, Cherney, and Johnson (1990a) showed decreased tillering, stem diameter, height and leaf/stem ratio with increasing plant density. Moreover, Caravetta, Cherney, and Johnson (1990b) also demonstrated that higher plant densities decrease the forage quality, but increases the forage DM yield. Previous work with grain sorghum has shown no advantage to high plant densities on above ground biomass and grain production (Staggenborg et al. 1999; Berenguer and Faci 2001).

The objectives of this study were 1) to evaluate the yield and forage quality of the newer early season brachytic dwarf sorghum relative to the full season hybrid and 2) to evaluate the yield and forage sorghum quality response of these lines to increasing both the seeding rate and N rate above currently recommended levels.

Materials and methods

Experiment location, weather and methodology

Field experiments were conducted in 2014 and 2015 at the Russell Larson Research Farm of the Pennsylvania State University (40° 43` N, 77° 56` W, 372 m elevation). Sorghum was planted using no-till techniques in the first week of June during both years of the study. Seasonal precipitation totaled 623 mm in 2014 and 553 mm in 2015 (Table 16). No supplemental irrigation was applied. The experimental design was a split-split-plot with four replications. In each replication, main plots were two varieties, subplots were two plant populations, and sub-subplots were two N rates. Individual subplots consisted of four 6 m long rows with a row spacing of 76 cm. The two forage sorghum hybrids were AF7202 and AF7401. AF7202 is a semi-dwarf early maturity hybrid and AF7401 is a dwarf medium maturity hybrid. Both hybrids carry the brown mid-rib 6 (*bmr-6*) gene. The two plant populations evaluated were 198,000 (recommended) and 296,400 (high) plants ha^{-1}. The two N rates were 123 (recommended) and 168 (high) kg ha^{-1}. Basal N rates consisted of a pre-emerge application of 112 kg ha^{-1} as UAN broadcast along with 11 kg ha^{-1} N as a starter. A second side-dress N application of 45 kg ha^{-1} N as UAN dribbled between the rows at 45 DAP (days after planting) was used for the higher rate N treatment.

Observations recorded

The days to fifty percent flowering and plant height were recorded before harvest. All the observations were collected from the middle two rows in each plot. The fifty percent flowering was recorded based on the flowering in fifty percent of the plants in each plot. Plant height was recorded to the top of the inflorescence on three randomly selected plants

in each plot. Plant stands were counted when plants were approximately 20 cm tall. Plant stand counts were recorded in 2015 but not 2014.

Harvesting and sample collection

The middle two rows were harvested at 90-95 days after planting at the soft dough stage using a research forage harvester (Wintersteiger, Salt Lake City, UT) that finely chopped the plant material and recorded the fresh weight from each plot with the sensors attached to the harvester. Two sets of 900 g of sub samples were collected in each treatment to determine dry matter and to asses forage quality analysis. The lab test sub-samples were immediately transferred to a cold storage unit to freeze and preserve the sample quality. The dry weight for each treatment were determined by weighing the fresh sub-samples at harvest and then again after a week of drying at 35°C.

Forage quality analysis

The forage quality analysis was performed by Cumberland Valley Analytical Services (Waynesboro, PA) using near infrared reflectance (NIR) spectroscopy to predict the levels of neutral detergent fiber (NDF), crude protein (CP), lignin, starch, ethanol soluble carbohydrates (ESC), total digestible nutrients (TDN), net energy for lactation (NE_L) and other parameters relevant to forage quality. NDFD was determined with wet chemistry analysis using a 30-hour incubation.

Statistical analysis

All the agronomic, yield and quality parameters were analyzed using the split-split-plot analysis of variance program in the GENSTAT software 14th edition of Rothamsted experimental station (Payne et al. 2011). PROC MIXED (SAS 2008) was used to generate type 3 fixed effects for selected variables and mean separation using Tukeys test to differentiate the mean by ranking for main effects and interaction terms (Steel and Torrie 1981).

Results and Discussion

Plant growth was generally good in both years of the study. Precipitation was near normal except in July and September of 2014 and August of 2015. Mean temperatures were near normal except in September 2014. In both seasons, all treatments reached the soft dough harvest maturity before a killing frost in the fall.

Flowering dates were influenced by variety and seeding rates in both years. The earlier hybrid AF7202 flowered 7 days earlier in 2014 and 9 days earlier in 2015 than AF7401. The effect of seeding rate on flowering date was significant in both years. On average higher seeding rates caused flowering to occur 2 days earlier (data not shown) but this was not consistent over varieties and seeding rates as we found significant interactions with varieties, nitrogen rates and seeding rates in both years. In 2014, lower N rate caused flowering dates to be earlier, but this was not consistent over varieties and seeding rates. In 2015, a significant variety by nitrogen interaction was observed, lower N rate delayed the flowering for 2 days in AF7202, whereas the higher N rate delayed the flowering in AF7401 for a day (data not shown).

Plant height was impacted primarily by the variety with the AF7202 averaging 105 and 110 cm in height in 2014 and 2015, respectively compared to 81 and 78 cm for AF7401 in 2014 and 2015. This difference in height was expected because AF7202 is an early season semi-dwarf hybrid, whereas AF7401 is a full season dwarf hybrid (Alta Seeds 2013). In 2014 a variety by seeding rate interaction for plant height was significant. This occurred because of the good emergence of the lines.

Forage dry matter yield

Forage DM yields averaged across all treatments were 13.2 and 13.6 t ha^{-1} in 2014 and 2015, respectively, demonstrating that these crops have a good yield potential in this region (Table 3). Significant differences were observed among hybrids for both years for forage DM yield (Table 2 and 3). The DM yield for AF7202 was higher than AF7401 in both years (Table 3) and averaged over both years the AF7202 yield was 1.9 t ha^{-1} higher than AF7401. The higher yield of AF7202 could be due to its taller height and ability to mature under more favourable and warmer conditions. The later hybrid AF7401 reached the soft dough stage each year just prior to frost and matured during the cool days of late September.

There were no significant differences due to the main effects of seeding and N rates on forage DM yield in either year. In 2014, a significant variety × N rate interaction as well as a seeding rate × N rate interaction occurred (Table 17). The yield of the AF7202 hybrid was increased in 2014 with the higher N rate, whereas AF7401 was not (Table 18). The

seeding rate x N interaction occurred because yields tended to increase with more N but not at the low seeding rate (data not shown). This was likely because both varieties tillered well at both seeding rates. In 2015, no differences due to N or seeding rate were observed. Our results are similar to those by Marsalis et al. (2009) and Marsalis, Angadi, and Contreras-Govea (2010). Bean et al. (2003) and Carmi et al. (2006) also reported no effect of increasing either seeding or N rate above recommended levels on forage sorghum DM yield.

Crude protein

We observed significant differences for CP due to varieties in both years, and significant differences between years with 78 g kg^{-1} in 2014 and 85 g kg^{-1} in 2015 (Table 17 and 19). The hybrid AF7401 had higher CP in both years, with 81 g kg^{-1} in 2014 and 90 g kg^{-1} in 2015. Whereas, AF7202 had only 75 g kg^{-1} in 2014 and 81 g kg^{-1} in 2015 (Table 19). On average, the CP of AF7401 was 8 g kg^{-1} higher than AF7202. N rate increased CP content in 2014 but not 2015 (Table 19). This may be due to differences in the plant population stand. In the year 2015, there was a less than an ideal plant stand due to poor emergence (data not shown). Individual plants likely had less competition for N which improved CP content and reduced response to additional N. Generally, increasing N rates have resulted in higher CP (Oliver et al. 2005; Miron et al. 2007). Seeding rate did influence the CP values across years (Table 4). A significant seeding rate × N rate interaction effect was also observed for CP for both years (data not shown). When increasing from the low to high seeding rate with increasing N rate the CP content was also improved in 2014. However, the high seeding rate with high N rate had a negative influence on CP content in 2015. The

same trend also observed by Cox and Cherney (2001) in corn where they showed reduced CP in high plant populations (116,000 plants ha^{-1} vs. 80,000 plants ha^{-1}) at several different N levels with consistent results across different environments.

Starch

Starch content was primarily affected by hybrid and this was significant in both years. The hybrid AF7202 accumulated more starch than AF7401 in both years (Table 20). Averaged over both years, the starch concentration of AF7202 was 295 or 49 g kg^{-1} higher than AF7401 which averaged 246 g kg^{-1}. This was due to the earliness of the hybrid AF7202, approximately 7 to 9 days earlier than the hybrid AF7401 (data not shown). This led to good grain fill in AF7202 compared to AF7401 during harvest. N rate did not influence starch content in either year of the study (Table 2 and 5). A significant seeding rate, seeding rate × N rate, and variety × N rate interaction occurred in 2014 but not 2015 (Table 17). This resulted because starch content increased in AF7401 in each year with additional N but this did not occur with AF7202.

Neutral detergent fiber digestibility

NDFD was only affected by variety in 2014 (Table 17) when AF7401 had a 19 g kg^{-1} advantage over AF7202 (Table 21). Although not significant, there was a trend for slightly higher NDFD levels in AF7401 in 2015 as well. Averaged over both years the differences between NDFD for AF7401 and AF 7202 was 12 g kg^{-1}, which may not be biologically significant in animal nutrition. Neither seeding rate nor N rate had significant effects on NDFD for in either year (Table 17 and 21). Larger differences in NDFD were evident

between years as in 2015, when NDFD averaged 521 g kg^{-1}, whereas in 2014 NDFD averaged 464 g kg^{-1} (Table 21). Carmi et al. (2006) reported that increasing the seeding rate from 200,000 to 260,000 plants ha^{-1} increased the NDFD with an additional irrigation. The higher NDFD in the BMR-forage sorghums may contribute to higher milk yield compared to conventional sorghum (Bean and McCollum 2006; Marsalis et al. 2009). Based on the current study, the recommended seeding rate (198000 plants ha^{-1}) and N rate (123 kg ha^{-1}) should be sufficient to achieve the optimum NDFD levels in BMR-forage sorghums for silage.

Net energy for lactation

The only variable that affected NE_L in this study was variety and this was significant in both years (Table 17). In both years, AF7202 had a slightly higher NE_L, which was likely due to the higher starch content in this variety. NE_L is calculated from the acid detergent fiber content, which is inversely related to starch content. Seeding rate and N rate had no impact on NE_L (Table 17 and 22). The higher NE_L was likely also due to the earliness of the hybrid AF7202, which facilitated the accumulation of more starch in these environments.

Conclusions

This study has demonstrated that the development of an earlier semi dwarf BMR forage sorghum variety such as AF7202 can have a positive impact on the potential of the crop in the Northeast US where the growing season can be marginal for some of the longer season forage sorghum varieties such as AF7401. The earlier line had higher yields, higher starch

content and $NE_{L.}$ There was some indication that in some environments AF7202 might be more responsive to N so this should be evaluated further. Seeding rate recommendations appear to be adequate for both varieties as both appear to tiller well when emergence is less than desired. Of the factors evaluated in this study, the improved genetics of AF7202 had the largest impact on forage sorghum yield and quality. Future development of earlier BMR forage sorghum varieties could focus on continuing to improve the NDFD of the early lines to match or be higher than later lines.

Funding

This work was funded by in part by a grant from the United States Department of Agriculture – National Institute of Food and Agriculture – Agriculture and Food Research Initiative – Project Number 2011-67009-30017.

Literature Cited

Alta Seeds. 2013. "Sorghums forage and grain." In, edited by Alta Seeds, 201 East John Carpenter Freeway, Irving, Texas. http://altaseeds.advantaus.com.

Bean, B., T. McCollum, M. Rowland, and K. McCuistion. 2003. "Seeding and nitrogen rate effect on forage yield, lodging, and nutrient composition of two brown midrib forage sorghums." http://agrilife.org/amarillo/files/2010/11/Seedingnitrogenrate2003.pdf.

Bean, B., and T. McCollum. 2006. "Summary of six years of forage sorghum variety trials." http://amarillo.tamu.edu/files/2010/11/SixYearForageSorghumSummary.pdf.

Beegle, D.B. 2016. *Soil fertility management*. Vol. 1, *The Agronomy Guide*. Cooperative Extension, The Pennsylvania State University, University Park, PA.

Berenguer, M.J., and J.M. Faci. 2001. "Sorghum (Sorghum bicolor (L.) Moench) yield compensation processes under different plant densities and variable water supply." *European Journal of Agronomy* 15:43-55.

Beyaert, Ronald P., and Robert C. Roy. 2005. "Influence of nitrogen fertilization on multi-cut forage sorghum–sudangrass yield and nitrogen use." *Agronomy Journal* 97 (6). doi: 10.2134/agronj2005.0079.

Buxton, D.R., I.C. Anderson, and A. Hallam. 1999. "Performance of sweet and forage sorghum grown continuously, double-cropped with winter rye, or in rotation with soybean and maize." *Agronomy Journal* 91:93-101.

Caravetta, G.J. , J.H. Cherney, and K.D. Johnson. 1990a. "Within-row spacing influences on diverse sorghum genotypes: I. Morphology." *Agronomy Journal* 82:206-10.

Caravetta, G.J., J.H. Cherney, and K.D. Johnson. 1990b. "Within-row spacing influences on diverse sorghum genotypes: II. Dry matter yield and forage quality." *Agronomy Journal* 82:210-5.

Carmi, Avner, Yohav Aharoni, Menahem Edelstein, Nakdimon Umiel, Amir Hagiladi, Edith Yosef, Moses Nikbachat, Abraham Zenou, and Joshua Miron. 2006. "Effects of irrigation and plant density on yield, composition and in vitro digestibility of a new forage sorghum variety, Tal, at two maturity stages." *Animal Feed Science and Technology* 131 (1-2):121-33. doi: 10.1016/j.anifeedsci.2006.02.005.

Contreras-Govea, Francisco, Mark Marsalis, Sangamesh Angadi, Gerald Smith, L. M. Lauriault, and Dawn VanLeeuwen. 2011. "Fermentability and nutritive value of corn and forage sorghum silage when in mixture with lablab bean." *Crop Science* 51 (3). doi: 10.2135/cropsci2010.05.0282.

Cox, W.J., and D.J.R. Cherney. 2001. "Row spacing, plant density, and nitrogen effects on corn Silage." *Agronomy Journal* 93 (3):597-602.

Grant, R. J., S.G. Haddad, K.J. Moore, and J.F. Pedersen. 1995. "Brown midrib sorghum silage for midlactation dairy cows." *Journal of Dairy Science* 78:1970-80.

Howell, T.A. , S.R. Evett, J.A. Tolk, K.S. Copeland, P.D. Colaizzi, and P. Gowda. 2008. "Evapotranspiration of corn and forage sorghum for silage." *World environmental and water resources congress* 10 (1):1-10.

Jahanzad, E., M. Jorat, H. Moghadam, A. Sadeghpour, M. R. Chaichi, and M. Dashtaki. 2013. "Response of a new and a commonly grown forage sorghum cultivar to limited irrigation and planting density." *Agricultural Water Management* 117:62-9. doi: 10.1016/j.agwat.2012.11.001.

Ketterings, Quirine M., Jerry H. Cherney, Greg Godwin, Tom F. Kilcer, Peter Barney, and Sam Beer. 2007. "Nitrogen management of brown midrib sorghum × sudangrass in the northeastern USA." *Agronomy Journal* 99 (5). doi: 10.2134/agronj2006.0350.

Kings Agriseeds. 2017. "Product information guide " In *Product information guide*, edited by Kings Agriseeds, 1828 Freedom Rd, Lancaster, PA. http://www.kingsagriseeds.com.

Lamm, F.R., L.R. Stone, and D.M. O'Brien. 2007. "Crop production and economics in northwest kansas as related to irrigation capacity." *Applied Engineering in Agriculture* 23 (6):737-45.

Liu, Chutao., Quirine. Ketterings, Karl. Czymmek, Tom. Kilcer, Jerry. Cherney, Mike. Hunter, Kitty. O' Neil, and Sarah. Lyons. 2016. "Guidance for growing BMR brachytic dwarf forage sorghum." *Cornell University Cooperative Extension Agronomy Fact Sheet Series* 92:1-2.

Marsalis, M. A., S. V. Angadi, and F. E. Contreras-Govea. 2010. "Dry matter yield and nutritive value of corn, forage sorghum, and BMR forage sorghum at different plant populations and nitrogen rates." *Field Crops Research* 116 (1-2):52-7. doi: 10.1016/j.fcr.2009.11.009.

Marsalis, M.A., S. V. Angadi, F.E. Contreras-Govea, and R.E. Kirksey. 2009. "Harvest timing and by product addition effects on corn and forage sorghum silage grown under water stress." *Agricultural Experiment Station Bulletin* 799.

Martin, John H. 1930. "The comparative drought resistance of sorghums." *Agronomy Journal* 22:993-1003.

Merrill, S. D., D. L. Tanaka, J. M. Krupinsky, M. A. Liebig, and J. D. Hanson. 2007. "Soil water depletion and recharge under ten crop species and applications to the principles of dynamic cropping systems." *Agronomy Journal* 99 (4):931-8. doi: 10.2134/agronj2006.0134.

Miron, Joshua, Ephraim Zuckerman, Gabriel Adin, Ran Solomon, Ezra Shoshani, Moshe Nikbachat, Edith Yosef, et al. 2007. "Comparison of two forage sorghum varieties with corn and the effect of feeding their silages on eating behavior and lactation performance of dairy cows." *Animal Feed Science and Technology* 139 (1-2):23-39. doi: 10.1016/j.anifeedsci.2007.01.011.

Oliver, A. L., J. F. Pedersen, R. J. Grant, and T. J. Klopfenstein. 2005. "Comparative effects of the sorghum bmr-6 and bmr-12 genes: I. Forage sorghum yield and quality." *Crop Science* 45 (6). doi: 10.2135/cropsci2004.0644.

Payne, R.W., D.A. Murray, S.A. Harding, D.B. Baird, and D.M. Soutar. 2011. "An introduction to GenStat for windows (14th Edition)." *VSN International, Hemel Hempstead, UK*.

Porter, K. S., J. D. Axtell, V. L. Lechtenberg, and V. F. Colenbrander. 1978. "Phenotype, fiber composition, and in vitro dry matter disappearance of chemically induced brown midrib (bmr) mutants of sorghum." *Crop Science* 18:205-8.

Sanderson, M.A. , G. Ali, M.A. Hussey, and F. Miller. 1994. "Plant density effects on forage quality and morphology of four sorghum cultivars." *Proceedings of American Forage Grassland Council* 3:237-41.

SAS. 2008. "The SAS system for windows." *SAS Institute, Cary, NC.* Release 9.2.

Staggenborg, S.A., D.L. Fjell, D.L. Devlin, W.B. Gordon, and B.H. Marsh. 1999. "Grain sorghum response to row spacings and seeding rates in Kansas." *Journal of Production Agriculture* 12 (3):390-5.

Steel, R.G.D. , and J.H. Torrie. 1981. *Principles and procedures of statistics: A biometrical approach*. Edited by R.G.D. Steel and J.H. Torrie. Vol. 2: McGraw-Hill, New York.

Tables

Table 16. Monthly rainfall and mean temperature during the growing seasons of 2014 and 2015 at Rock Springs, PA.

	Rainfall (mm)		Minimum temperature (°C)		Maximum temperature (°C)		Mean temperature (°C)	
	2014	2015	2014	2015	2014	2015	2014	2015
June	123	166	14	14	25	24	19	19
July	80	158	15	15	26	27	20	21
August	120	51	13	13	24	26	19	19
September	34	75	11	12	22	25	16	18

Table 17. Analysis of variance for flowering, plant height, dry matter (DM) yield, crude protein (CP), starch, neutral detergent fiber digestibility (NDFD), and net energy for lactation (NE_L) for the summer 2014 and 2015 with two varieties, two seeding rates (SR), and two N rates (NR).

Source	d.f.	Flowering (days)	Plant height (cm)	DM yield (t ha^{-1})	CP (g kg^{-1})	Starch (g kg^{-1})	NDFD (g kg^{-1})	NE_L (Mcal kg^{-1})
2014								
Variety (Var)	12	***	***	***	*	**	*	***
Seeding rate (SR)	3	***	***	NS	NS	**	NS	NS
Nitrogen rate (NR)	6	***	NS	NS	***	NS	NS	NS
SR × NR	6	**	NS	*	*	*	NS	NS
Var × SR	12	NS	**	NS	NS	NS	NS	NS
Var × NR	12	NS	NS	**	NS	*	NS	NS
Var × SR × NR	12	*	NS	NS	NS	NS	NS	NS
2015								
Variety (Var)	12	**	***	***	***	***	NS	***
Seeding rate (SR)	3	*	NS	NS	NS	NS	NS	NS
Nitrogen rate (NR)	6	NS	NS	NS	NS	NS	NS	NS
SR × NR	6	NS	NS	NS	**	NS	NS	NS
Var × SR	12	*	NS	NS	NS	NS	NS	NS
Var × NR	12	*	NS	NS	NS	NS	NS	NS
Var × SR × NR	12	NS	NS	NS	NS	NS	NS	NS

*significant at p = 0.05.
**significant at p = 0.01.
***significant at p = 0.001.
NS, non-significant at p = 0.05.

Table 18. Total seasonal dry matter (DM) yield (t ha^{-1}) of sorghum forage varieties at two seeding rates (SR) and two N rates (NR) for two cropping seasons at Rock Springs, PA.

Forage dry matter yield (t ha^{-1})	2014			2015		
	Cultivar			Cultivar		
	AF7401	AF7202	Means	AF7401	AF7202	Means
Variety	$12.3^{B\dagger}$	$14.1^{A\dagger}$	13.2***	$12.6^{B\dagger}$	$14.6^{A\dagger}$	13.6***
LSD	0.6			0.6		
N rate (kg ha^{-1})						
123	12.6	13.7	$13.1^{a\ddagger}$	12.9	14.7	$13.8^{a\ddagger}$
168	12.0	14.6	$13.3^{a\ddagger}$	12.3	14.5	$13.4^{a\ddagger}$
LSD	0.5			0.5		
Seeding rate (ha^{-1})						
198000	12.3	14.1	$13.2^{a\ddagger}$	12.1	14.3	$13.2^{a\ddagger}$
296400	12.3	14.1	$13.2^{a\ddagger}$	13.1	14.9	$14.0^{a\ddagger}$
LSD	0.4			0.8		

†Means in a row with the same uppercase letter are not significantly different from each other at p = 0.05.

‡ Means in a column with the same lowercase letter are not significantly different from each other at p = 0.05.

***significant at p = 0.001.

Table 19. Crude protein content (g kg^{-1}) of sorghum forage varieties at two seeding rates (SR) and two N rates (NR) for two cropping seasons at Rock Springs, PA.

Crude protein (g kg^{-1})	2014			2015		
	Cultivar			Cultivar		
	AF7401	AF7202	Means	AF7401	AF7202	Means
Variety	$81^{A\dagger}$	$75^{B\dagger}$	78*	$90^{A\dagger}$	$81^{B\dagger}$	85***
LSD	5			2		
N rate (kg ha^{-1})						
123	75	78	$73^{b\ddagger***}$	90	81	$86^{a\ddagger NS}$
168	87	80	$84^{a\ddagger***}$	90	81	$85^{a\ddagger NS}$
LSD	4			2		
Seeding rate (ha^{-1})						
198000	80	72	$76^{a\ddagger}$	92	81	$86^{a\ddagger}$
296400	82	78	$80^{a\ddagger}$	89	81	$85^{a\ddagger}$
LSD	6			3		

†Means in a row with the same uppercase letter are not significantly different from each other at p = 0.05.

‡ Means in a column with the same lowercase letter are not significantly different from each other at p = 0.05.

*significant at p = 0.05.

***significant at p = 0.001.

NS, non-significant at p = 0.05.

Table 20. Starch content (g kg^{-1}) of sorghum forage varieties at two seeding rates (SR) and two N rates (NR) for two cropping seasons at Rock Springs, PA.

Starch (g kg^{-1})	2014			2015		
	Cultivar			Cultivar		
	AF7401	AF7202	Means	AF7401	AF7202	Means
Variety	$218^{B\dagger}$	$273^{A\dagger}$	245**	$273^{B\dagger}$	$317^{A\dagger}$	295***
LSD	22			7		
N rate (kg ha^{-1})						
123	204	276	$251^{a\ddagger}$	268	322	$295^{a\ddagger}$
168	232	270	$240^{a\ddagger}$	278	312	$295^{a\ddagger}$
LSD	17			13		
Seeding rate (ha^{-1})						
198000	228	284	$256^{a\ddagger}$	264	315	$290^{a\ddagger}$
296400	208	262	$235^{a\ddagger}$	282	319	$300^{a\ddagger}$
LSD	11			20		

†Means in a row with the same uppercase letter are not significantly different from each other at p = 0.05.

‡ Means in a column with the same lowercase letter are not significantly different from each other at p = 0.05.

**significant at p = 0.01.

***significant at p = 0.001.

Table 21. NDFD (g kg^{-1}) of sorghum forage varieties at two seeding rates (SR) and two N rates (NR) for two cropping seasons at Rock Springs, PA.

NDFD (g kg^{-1})	2014			2015		
	Cultivar			Cultivar		
	AF7401	AF7202	Means	AF7401	AF7202	Means
Variety	473^{A†}	454^{B†}	464*	523^{A†}	519^{A†}	521^{NS}
LSD	21			21		
N rate (kg ha^{-1})						
123	467	455	461^{a‡}	526	517	521^{a‡}
168	480	453	466^{a‡}	520	522	521^{a‡}
LSD	16			14		
Seeding rate (ha^{-1})						
198000	466	458	462^{a‡}	527	518	518^{a‡}
296400	480	450	465^{a‡}	519	521	524^{a‡}
LSD	26			19		

†Means in a row with the same uppercase letter are not significantly different from each other at p = 0.05.

‡ Means in a column with the same lowercase letter are not significantly different from each other at p = 0.05.

*significant at p = 0.05.

NS, non-significant at p = 0.05.

Table 22. NE_L (Mcal kg^{-1}) of sorghum forage varieties at two seeding rates (SR) and two N rates (NR) for two cropping seasons at Rock Springs, PA.

NE_L (Mcal kg^{-1})	2014			2015		
	Cultivar			Cultivar		
	AF7401	AF7202	Means	AF7401	AF7202	Means
Variety	1.55[B†]	1.56[A†]	1.56***	1.54[B†]	1.55[A†]	1.55***
LSD	0.02			0.01		
N rate (kg ha^{-1})						
123	1.55	1.56	1.55[a‡]	1.54	1.56	1.55[a‡]
168	1.55	1.57	1.56[a‡]	1.54	1.55	1.55[a‡]
LSD	0.01			0.01		
Seeding rate (ha^{-1})						
198000	1.55	1.56	1.55[a‡]	1.54	1.55	1.55[a‡]
296400	1.54	1.56	1.55[a‡]	1.54	1.55	1.55[a‡]
LSD	0.01			0.01		

†Means in a row with the same uppercase letter are not significantly different from each other at p = 0.05.

‡ Means in a column with the same lowercase letter are not significantly different from each other at p = 0.05.

***significant at p = 0.001.

Chapter 6

Summary and Perspective

Sorghum is one of the most important cereal crops mainly used for food, feed and fodder. It is important to identify the resistance sources for various biotic and abiotic stresses in sorghum. We have used the association mapping approach to identify the genetic factors that contribute towards crop resistance for biotic (anthracnose) and abiotic (drought and frost) stresses in sorghum. *Colletotrichum sublineola* is an aggressive hemi-biotrophic fungal pathogen that causes anthracnose leaf blight and stem rot in sorghum. In severe conditions, it cause substantial yield loss up to 45%. Often major gene resistance fails due to varying degree of virulent pathotypes across sorghum growing locations and availability of other host species for pathogen survival. One mechanism to control anthracnose leaf blight in sorghum is to identify genetic factors associated with anti-fungal production. It is known that *Colletotrichum sublineola* induce the production of anti-fungal compounds known as 3-deoxyanthocyanidins (3-DAs) up on the fungal ingress in sorghum. These compounds are initially colorless and move to the site of infection. After 36 hours, they turn into brick red color and kill the invading fungus and the cell. This mechanism prevents the fungal pathogen spread from one area to other, thus effectively preventing the disease spread. Our goal of this investigation was to identify the genetic factors associated with 3-DA bio-synthesis and regulation in sorghum. GWAS identified 11 candidate genes (*p*-value $\leq$7.73E-06) for 3-DA bio-synthesis and regulation in sorghum. We have identified novel genes associated with signal transduction pathway, programmed cell death pathway,

vesicular trafficking, and sugar (major) transporter pathway genes as candidates for 3-DAs in sorghum. The genes like LRR (Sb07g005820), MYB (Sb04g010280), ASC1 (Sb01g040200), Rab-GTPase (Sb07g006550), peroxidase (Sb09g004660), and GBPs (Sb04g013160) could be exploited in marker assisted breeding as targets for improving crop tolerance to biotic and abiotic stresses especially for developing anthracnose disease resistant cultivars. Further gene validation pertaining to 3-DA induction upon fungal infection on selected candidates needs to be performed. Our results provide basis for anthracnose disease improvement through marker assisted selection and genomic selection for plant secondary metabolites that act as fungicides in sorghum.

Sorghum is a climate resilient multipurpose crop. Moreover, it exhibits poor tolerance to cold and frost. A number of flavonoid secondary metabolites are induced during the plant–stress cross talk, and they play a major role in imparting stress tolerance to plants. Our aim is to quantify the flavonoid induction before and after frost stress, and determine its anti-oxidant activity in the global sorghum panel and also identify the key candidate genes involved in the process through genome wide association mapping. GWAS identified 51 genes for DPPH (anti-oxidant activity) and 20 genes for total phenolic content (TP) before frost. Whereas, 19 genes were identified 3-deoxyanthocyanidins (3-DAs) and 6 genes for TPC under after frost stress. Most of the identified genes are involved in plant defense pathways for biotic and abiotic resistance. The probable candidates for after stress were peroxidase (sb01g041770), vesicle fusing ATPase (sb01g041930), leucine rich repeats (sb08g001430), flavonol reductase (sb01g025770), cytosine deaminase (sb02g041770), and anthocyanidin reductase (sb02g038520) for 3-DAs; and leucine rich repeat

(Sb08g023030) and UDP-glucosyl transferase (Sb06g021900) for TP. Anti-fungal gene (Sb06g014340) was the most probable candidate for DPPH before frost stress. The identified candidate genes can be used in breeding climate resilient sorghum

Sorghum accumulates epi-cuticular wax (EW) or bloom in plant surfaces such as leaves, sheaths, and culm. EW reduces the non-transpiration water loss and protects the plant from severe drought stress and also imparts resistance against various insect pests. We present here results from the analysis of epi-cuticular wax (EW) content of 387 diverse sorghum accessions and its genome-wide association study (GWAS). EW content in sorghum leaves ranged from 0.1 mg cm^{-2} to 29.7 mg cm^{-2} with a mean value of 5.1 mg cm^{-2}. GWAS using 265,487 single nucleotide polymorphisms (SNPs) identified thirty-seven putative genes that were associated ($p < 9.89E-06$) with EW biosynthesis and transport in sorghum; out of these, Sobic.002G310400 (3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal), Sobic.004G154200 (Ankyrin repeats), and Sobic.003G004500 (bHLH-MYC and R2R3-MYB transcription factors N-terminal) have been shown to be involved in EW biosynthesis; Sobic.001G447300 (ABC transporter), Sobic.004G089400 (Lipid exporter ABCA1 and related proteins, ABC superfamily), Sobic.002G311200 (Multidrug resistance protein), Sobic.001G447200 (Inositol 1, 3, 4-trisphosphate 5/6-kinase), and Sobic.005G217500 (Cytochrome p450) are involved in EW regulation or transport in sorghum. GWAS results from this study demonstrate the potential for genetic manipulation of EW content in sorghum and other grain cereals for better adaptation to biotic and abiotic stresses.

Brown mid-rib (BMR) forage sorghum (*Sorghum bicolor* (L.) Moench) silage is a reasonable alternative to corn silage for areas with limited soil moisture. Traditional forage sorghum varieties are tall and prone to lodging with lower forage quality. Brachtyic dwarf BMR forage lines are shorter, lodging resistant and have higher forage quality. Newer, earlier hybrids have expanded the potential adaptation of forage sorghums to more northern areas. A two-year study was conducted during the 2014 and 2015 growing seasons using newly available brachytic dwarf BMR forage hybrids to determine the effects of different seeding rates and N (nitrogen) fertilization rates on forage dry matter (DM) yield and quality for two hybrids. The experimental design was split-split-plot with four replications. In each replication, main plots were two hybrids (AF7202 and AF7401), subplots were two seeding rates (198,000 and 296,400 seeds ha^{-1}), and sub-subplots were two N rates (123 and 168 kg ha^{-1}). DM yield and forage quality parameters were measured for each treatment. We observed significant varietal differences for all the parameters except neutral detergent fiber digestibility (NDFD) in 2015. The early maturity line, AF7202, had higher yields, higher starch and net energy for lactation (NE_L) levels than AF7401. The dwarf line, AF7401, had higher crude protein (CP) and NDFD than AF7202. AF7202 was more responsive to the higher N rate than AF7401. CP was increased with increased N for both varieties. Other forage quality traits were unaffected by N rates. Neither variety responded to an increase in seeding rate. This study showed that the earlier brachytic dwarf forage sorghums such as AF7202, managed with recommended seeding rates and possibly higher N rates have good potential for high forage yield and quality in central PA. This study has demonstrated that the development of an earlier semi dwarf BMR forage sorghum variety such as AF7202 can have a positive impact on the potential of the crop in the Northeast US

where the growing season can be marginal for some of the longer season forage sorghum varieties such as AF7401. The earlier line had higher yields, higher starch content and $NE_{L.}$ There was some indication that in some environments AF7202 might be more responsive to N so this should be evaluated further. Seeding rate recommendations appear to be adequate for both varieties as both appear to tiller well when emergence is less than desired. Of the factors evaluated in this study, the improved genetics of AF7202 had the largest impact on forage sorghum yield and quality. Future development of earlier BMR forage sorghum varieties could focus on continuing to improve the NDFD of the early lines to match or be higher than later lines.

VITA

Dinakaran Elango

Educational Qualification

2013-2018 Ph.D. in Agronomy, The Pennsylvania State University, USA
2006-2008 MS in Plant Breeding & Genetics, Tamil Nadu Agricultural University
2001-2005 BS (Agriculture), Tamil Nadu Agricultural University, India

Awards, Honors and Fellowships

- College of Agricultural Sciences – 2016 graduate student competitive grant
- Penn State Entrepreneurship & Innovation grant "2014 & 2016"
- College of Agricultural Sciences "Tag Along Awardee" for the year "2014"

Professional body membership

- Association for Promotion of DNA Fingerprinting & Other DNA Technologies.

Research Experience

2013-2018 - Graduate Research Assistant at Maize Genetics Lab, Penn State University
2011-2013 - Scientific Officer in Sorghum Breeding, ICRISAT
2010-2011 - Research Associate, Pioneer-HiBred International
2008-2010 - Senior Research Fellow, Tamil Nadu Agricultural University
2005-2006 -Junior Research Fellow, Tamil Nadu Agricultural University

Publications

1. Dinakaran Elango et al., 2015. Determination of oligosaccharide fraction in a worldwide germplasm collection of chickpea using high performance liquid chromatography. Australian Journal of Crop Science 9 (7): 605-613.

2. Krishnamurthy L, Dinakaran Elango. 2014. Field technique and traits to assess reproductive stage cold tolerance in Sorghum. Plant Production Science, 17 (3): 218-227.

3. Dinakaran Elango, S Thirumeni. 2012. Yield and fibre quality components analysis in upland cotton under salinity. Annals of Biological Research, 3 (8): 3910-3915.

4. Belum VS Reddy, Dinakaran Elango, JV Patil. 2012. Postrainy season sorghum: Constraints and breeding approaches. Journal of SAT Agricultural Research, 10.

5. Belum VS Reddy, Dinakaran Elango. 2012. Sorghum improvement: Status and Way forward. Journal of SAT Agricultural Research, 10.