Justus Liebig University, Giessen
Institute of Agronomy and Plant Breeding I
Department of Plant Breeding

Genetic dissection of cold tolerance in sorghum (*Sorghum bicolor* L. Moench)

Inaugural Dissertation for a Doctorate Degree in Agricultural Sciences
in the Faculty of Agricultural Sciences, Nutritional Sciences and
Environmental Management

Examiners

Prof. Dr. Rod Snowdon

Prof. Dr. Matthias Frisch

Submitted by

Subhadra Chakrabarty

Giessen 2022

This thesis is dedicated to my mother,

my source of inspiration, support, and guidance.

Table of contents

1 General introduction

1.1 Sorghum: origin, evolution, and genetic diversity

Sorghum (*Sorghum bicolor* L. Moench) is ranked as the 5th most important cereal crop in the world, after maize, rice, wheat, and barley, with an annual global production of 57.8 million tons (FAO STAT 2022) (https://www.fao.org). The ancient history of sorghum origin and domestication is not known and has been long debated (De Wet and Harlan 1971) but it is believed to have originated around 7000–5000 BC in the north-eastern part of Africa (Smith and Frederiksen 2000) where a significant amount of diversity still exists (Dahlberg 2000). Sorghum transitioned from a wild pluvial plant in northeastern Africa to the ancestral domesticated form *S. bicolor* type bicolor in Central Eastern Sudan around 5,000 years ago, while cultivation is inferred to have begun by 6000 years ago (Winchell et al. 2017). From the center of origin, domesticated sorghum spread across Africa and through the Middle East as a result of human migration (Mann et al. 1983). It arrived in India via trade routes, from where it finally reached China. Eventually, sorghum also reached the Americas, Europe and later Australia (Venkateswaran et al. 2014; Klein et al. 2015).

The evolutionary history of sorghum is difficult to reconstruct from modern data sets because of numerous complex genomic interactions and rigorous selection pressure (Smith et al. 2019) that took place over the years. But the diversity of various sorghum types can be associated with movement of people, geographic isolation, diverse selection, gene flow and recombination of different types in different environments (de Wet and Huckabay 1967; Doggett 1988). De Wet and Harlan (1971) identified three species of the genera *Sorghum* (*S. halepense*, *S. propinquum*, and *S. bicolor*) representing all annual wild, weedy and cultivated taxa. *S. bicolor* was further classified into the three subspecies *S. bicolor* subsp. *Drummondii*, *bicolor*, and *verticilliflorum*, respectively (De Wet 1978; Mann et al. 1983). *S. bicolor* subsp. *bicolor* contains all the cultivated sorghums. Finally, the primary gene pool of *S. bicolor* L. Moench was partitioned into five basic races (designated bicolor, guinea, caudatum, kafir, and durra) and ten intermediate races from the combinations of the five basic races (De Wet and Harlan 1971).

Sorghum is widely grown as small-grain cereal, cultivated between 40°N and 40°S of the equator (Srinivasa Rao et al. 2014). It is mainly rainfed and cultivated in lowlands and semi-arid areas of the tropics and sub-tropics. Its adaptation to a wide range of environments and selection for various end uses (grain, fodder, sugar, fiber) led to increased genetic and morphological diversity. For example, in African countries, sorghum is essential as food grain and stalk, and leaves are valued as forage and building material. In China, it is a popular bioenergy feedstock and is used to manufacture traditional alcoholic beverages. Although sorghum has displayed immense genetic variation based on origin, race, seed type, photoperiod sensitivity, agronomic traits and molecular markers (Assar et al. 2005; Deu et al. 2006; Kayodé et al. 2006), it has not been fully harnessed for breeding purposes.

Knowledge regarding the role of socio-economic, cultural and environmental factors in shaping crop diversity is crucial to assessing the resilience of sustainable agriculture in the face of adverse

environmental changes across the globe. The predicted negative impact on yield traits due to changing climate (Ramirez-Villegas et al. 2013) draw attention towards both availability of appropriate genetic resources and ability of breeding programs to develop the required adaptations promptly (Burke et al. 2009). Understanding of local seed systems can contribute to better utilization of resources and improving agriculture in the face of climate change and human insecurity (McGuire and Sperling 2013).

1.2 Genomic resources in sorghum

Sorghum is a member of the grass family Poaceae, tribe Andropogoneae, subtribe Sorghinae and genus Sorghum Moench (Dillon et al. 2007). It is a diploid organism (2n = 20) with a relatively small genome (~ 730 Mbp). According to phylogenetic studies, sorghum is closely related to other crops like maize, rice, and especially sugarcane. During evolution, sorghum, a C4 grass, diverged from rice and maize and finally sugarcane (*Saccharum officinarum*) around 5 million years ago (mya) (Paterson et al. 2009). However, due to genetic relatedness between sorghum and sugarcane, intergeneric hybrids have been created (Bowers et al. 2003). High-density genetic maps, one intraspecific and another from an interspecific cross (between *S. bicolor* and *S. propinquum*) show a high degree of collinearity, indicating the point of divergence between sorghum species (Feltus et al. 2006). These comparative maps serve as bridges for information exchanges across the Andropogoneae tribe for breeding, evolution, ecology, and molecular biology studies. With a completely assembled genome and its African origin, sorghum is considered an important model crop along with maize (America) and rice (Asia). In particular, it can be a useful model to study polyploid crops like sugarcane because of its small, less complex diploid genome and close genetic relationship (Lawrence and Walbot 2007).

Comparative genetic studies in angiosperms revealed the presence of nearly 94% (25,875) of high-confidence sorghum gene orthologues in rice, Arabidopsis and/or poplar. The adaptability of sorghum can be the result of cytochrome P450 domain-containing genes, which occur abundantly in sorghum (326) as compared to rice (228) and/or expansins, a class of enzymes responsible for various growth responses which are present in 82 copies in sorghum versus 58 in rice and 40 each in Arabidopsis and poplar. The sorghum genome has not been re-duplicated like other grasses and, hence, may play a critical role in understanding the evolution of repeatedly duplicated grass genomes (Paterson et al. 2004). Understanding reproductive biology processes like flowering, pollination and seed development are essential for designing effective sorghum breeding strategies. The transformation from the vegetative to reproductive phase is marked by floral initiation and is highly influenced by genotype and environment. Sorghum is mostly self-pollinated, but varied frequencies of outcrossing (10% to 73%) are known (Ellstrand and Foster 1983; Djè et al. 2004; Barnaud et al. 2008). As a result, spontaneous hybridization between cultivated and wild weedy sorghum often results in intermediary forms (Ejeta and Grenier 2005). Sorghum can be divided into primary, secondary, and tertiary gene pools based on cross-compatibility among its relatives (Ananda et al. 2020).

According to Grootboom et al. (2010), most improvements to sorghum have been achieved through conventional breeding. A turning point in sorghum breeding arrived when Stephens and Holland (1954) discovered cytoplasmic male sterility (CMS) in sorghum. The first CMS source, known as milo cytoplasm, was found in the progenies of two cultivars milo (female) and combine kafir (male). Use of milo CMS proved particularly useful for hybrid cultivar production to exploit high yields caused by heterosis. Enhancement of grain production is one of the keys to solving the global hunger issue. The most critical requirement for the application of genomics in the breeding of complex traits is the availability of suitable platforms and genetic markers. Previously, sorghum genetic linkage maps were constructed using early generation markers like RFLP (Restriction fragment length polymorphism), AFLP (Amplified fragment length polymorphism), RAPD (Random amplified polymorphic DNA) and simple sequence repeat (SSR) (Hulbert et al. 1990; Tao et al. 1998; Singh and Lohithaswa 2006; Ejeta and Knoll 2007). However, there were major drawbacks of these technologies. In the past decade, crop breeding has been revolutionized by the invention of molecular genetic tools like next-generation and third-generation sequencing technologies, easing the development of molecular markers like single nucleotide polymorphisms (SNP) (Han et al. 2015). In sorghum, molecular markers were first used by Hamblin et al. (2004) to study genetic variation and linkage disequilibrium. Markers like SSR and SNPs rely on genomic sequence information, because of which cost per data point is remarkably high. Diversity Arrays Technology (DArT™) was introduced as a cost-effective hybridization marker technology that did not require prior sequence information. Mace et al. (2008) first developed DArT markers for *S. bicolor* and created a medium-density linkage map for different molecular breeding and genomics applications.

In addition, targeted genome modifications and reverse genomics approaches have accelerated the understanding of gene functions for underlying traits (Liu and Godwin 2012; Nida et al. 2016). Currently, biofortified sorghum with qualities like reduced phytic acid, higher lysine, and threonine, increased protein digestibility, increased level of iron and zinc as well as higher amount of vitamin A precursor beta carotene is being produced to tackle hunger issues in many under-developed or developing nations like in Africa or India (Rao et al. 2006; Hokanson et al. 2010). Transgenic traits have been added to *S. bicolor* using Agrobacterium-mediated transformation and particle bombardment (Casas et al. 1993; Zhao et al. 2000). The genetic transformation efficiency of *S. bicolor* was extremely low (> 10%) until Liu and Godwin (2012) optimized the process and reported efficiencies of over 20% in an inbred sorghum line. Several sorghum genomic databases and large-scale resource platforms are now available which are invaluable to molecular breeding programs and research (Goodstein et al. 2012; Luo et al. 2016; Tian et al. 2016; Jiao et al. 2016; Mace et al. 2019).

The first assembled and annotated sorghum genome, published by Paterson et al. (2009) and later updated by McCormick et al. (2018), opened new avenues in sorghum genomic research. It is characterized by ~ 34,129 annotated genes, 58.8% retrotransposons, and 8.7% DNA transposons. Based on this reference genome, numerous sequence variations like single nucleotide polymorphism (SNPs), genomic copy number variation (CNV), insertions and deletions (InDels),

and presence-absence (PAV) have been identified (Zheng et al. 2009; Mace et al. 2013). Most of the sorghum domestication studies are based on archeological evidence of fossil remains (Winchell et al. 2017). However, this method has some obvious limitations, like complications involving analysis of fossilized data, shifts in domestication events and genomic characterization (Gaut et al. 2018). Smith et al. (2019) used sequencing data for ancient sorghum accessions to study the effects of domestication. Results indicate an overall increase of deleterious mutation load instead of the appearance of a genetic bottleneck. This may be a result of different end-uses of sorghum and gene flow across different subpopulations (Ohadi et al. 2017).

With the improvement of sequencing technologies and reduction in costs, multiple sorghum reference genomes have been reported over the past few years. Recently, a chromosome-scale *de novo* genome assembly of Tx430, a grain sorghum accession, was generated using a combination of Oxford Nanopore, Bionano Genomics Direct Label long-read sequencing and Stain (DLS) optical maps (Deschamps et al. 2018). In comparison to the previous BTx623 genome assembly (McCormick et al. 2018), this new assembly consists of a shorter median length and contains a higher number of predicted genes. The following year, Cooper et al. (2019) published the sequence of the sweet sorghum 'Rio', to study possible genomic differences between sweet and grain types. Although there is a high similarity between grain and sweet at the genetic structural level, key differences in regulatory genes, including potential deletions and loss-of-function mutations in sugar metabolism genes crucial for stem sugar accumulation were discovered. These assemblies allow an in-depth understanding of underlying genetic diversity at the SNP level. However, it was soon discovered that single-reference assemblies do not represent the complete species-wide genomic space (Springer et al. 2009; Anderson et al. 2014; Yang et al. 2014). Existence of several SVs like PAV, CNV, and chromosomal rearrangements. Access to multiple assembled reference genomes allow opportunities for the identification of SVs in a non-reference-biased manner. The sorghum pan-genome includes 44,079 gene families with 222.6 Mb of newly identified sequence (Tao et al. 2021). Extensive genetic variation within the pan-genome suggests its influence on the phenotypic outcome and its contribution to crop improvement. Results show that adaptation of sorghum lines to different abiotic and biotic stresses might be a result of dispensable genes (variable fraction of genome) as compared to core genes (genomic fraction common to all individuals within a species). The proportion of dispensable genes in the sorghum pan-genome (64%) was found to be higher than rice (54%) (Wang et al. 2018), soybean (49%) (Li et al. 2014), and *B. distachyon* (45%) (Gordon et al. 2017) showing its high genetic diversity and adaptability. Ruperao et al. (2021) developed another sorghum pan-genome using the reference genome as well as 354 genetically diverse sorghum lines from different races. Around two million SNPs developed through the pan-genome were used in different association studies. The availability of the first sequenced sorghum genome (Paterson et al. 2009) and genetic markers like SNP provided a basis for the construction of high-density genetic maps, mapping of quantitative trait loci (QTL) for multiple stress and quality traits. Although the regulation mechanism for most of the underlying genes remains unknown, the knowledge has helped design "super sorghum" for various end uses (Hao et al. 2021). Mace et al. (2019) provided the 'Sorghum QTL Atlas' as an open-access research platform to facilitate gene discovery across cereal species.

Since domestication, sorghum has been cultivated in a wide range of agro-climatic zones. Utilizing population genomics for studying complexities of genetic relatedness, population structure, selection patterns, genetic adaptation, linkage disequilibrium (LD), etc. has always been a preferred choice. Genome-wide association studies (GWAS) and genotype-by-environment (G × E) interaction studies have been conducted for sorghum to understand the effect of environment and genetics on a particular phenotype. Complex traits have been dissected using markers-based models to design enhanced breeding programs for genetic evaluation Several studies on genome-wide dissection associated genes and loci underlying important traits like grain yield and quality, flowering, plant height, and stress tolerance has been conducted (Multani et al. 2003; Murphy et al. 2014; Zhang 2015; Boyles et al. 2017; Spindel et al. 2018). Bernal et al. (2014) studied the effect of G × E interaction on sugar accumulation in improved varieties of sweet sorghum and identified genotypes with high potential for the biofuel agroindustry. A total of 336 sorghum RILs (Recombinant Inbred Lines) was used to study G × E interaction for grain iron and zinc concentrations across multiple environments (Phuke et al. 2017). Results showed a significant positive correlation across the environments for both traits, indicating the prospect of simultaneous effective selection. Genomic selection (GS), a form of marker-assisted selection, has been particularly useful in reducing evaluation cost and generation interval (Meuwissen et al. 2001). Decreasing genotyping costs or set of preselected markers like genotyping-by-sequencing GBS combined with new genome fingerprinting techniques like shallow sequencing (Gorjanc et al. 2017) facilitates an early generation marker-assisted and genomic selection for multiple complex agronomic traits and thus reduces the duration breeding cycle.

GS has been reported in sorghum for different traits like grain yield, drought adaptation, improvement of biomass, and other adaptability traits (Fernandes et al. 2018; Velazco et al. 2019a, 2019b; Bernardino et al. 2021). Extensive genomic knowledge of adaptive traits would ease the performance prediction of sorghum lines under different environments.

Understanding the genetic basis of wide adaptability and speciation of sorghum would facilitate molecular breeding programs and studies of other C4 crops. Multiple well-established diverse association panels are now available for sorghum research. Mace et al. (2013) used resequencing data of 44 diverse genotypes to study genetic diversity and evolution through domestication events. Results suggest the existence of a strong racial structure and a complex domestication history involving at least two distinct domestication events. The rate of LD decay of wild sorghums was found to be greater than those of landrace and improved varieties, suggesting that genetic diversity decreased in the case of the latter. The most widely used diversity panel genotyped GBS and consisting of 265,487 SNPs was created by (Morris et al. 2013). Soon after, Ji et al. (2017) used genome-wide specific-locus amplified fragments (SLAF) markers, which are available and evenly distributed across the genome to create high-density genetic maps which would facilitate gene exploration.

In recent times, transcriptome and proteome profiling technologies have emerged as revolutionary tools in genomic research. While transcriptome analysis can be used for finding novel transcripts and gene expression analysis (Vicente et al. 2019), proteomics has been used to study expression

patterns and functions of proteins (Uberegui et al. 2015). Multiple sorghum transcriptome research focusing on traits like growth, development, sugar accumulation, etc. have been reported (Kebrom and Mullet 2016; McKinley et al. 2016; McCormick et al. 2018). Jedmowski et al. (2014) studied the adaptive response of *S. bicolor* landraces to drought stress and its recovery by analysis of alterations in protein content. Results showed drought tolerance and recovery were induced by joint activities of several protein groups. To understand the background of sorghum grain mold, Nida et al. (2021) studied the transcriptome profile of grain at different developing stages for both susceptible and resistant genotypes. Previously undescribed defensin genes showed higher expression patterns in resistant genotypes as compared to the susceptible ones. Further integration of third generation (long-read) RNA sequencing will improve transcriptome analysis and help identify all relevant splice forms.

1.3 Biotic and abiotic stress tolerance in sorghum

To increase the productivity and quality of sorghum, the reduction of crop damage due to different biotic stresses is crucial. Pest and fungus attacks are usually common in tropical and subtropical climates. Sorghum head bug (*Eurystylus oldi*), shoot fly (*Atherigona soccata*), the midge (*Stenodiplosis sorghicola*), the bollworm (*Helicoverpa armigera*), and maize weevil (*Sitophilus zeamais*) are a few of the common insects attacking sorghum (Chandrashekar and Satyanarayana 2006). Infestation by *Striga hermonthica* and *S. asiatica*, types of parasitic weeds are common in semiarid parts of Africa and Asia. (Sleper and Poehlman 2006). Fungal diseases like ergot (*Claviceps sorghi*), grain mold (*Aspergillus spp*.), smut (*Sporisorium sorghi*), anthracnose (*Colletotrichum sublineola*), downy mildew (*Sclerospora sorghi*), and leaf blight (*Setosphaeria turcica*) are a few of the diseases rampant in the humid regions of east African countries (Hulluka and Esele 1992). As compared to other parts of the world, like northern parts of the USA and Canada, disease pressure on sorghum remains low throughout Europe as the crop area is still low. Some fungal disease incidents have been reported across Europe, but without any major yield losses (Forbes 1992). Sorghum is also known to be unsusceptible towards certain infections occurring in maize, like Western Corn Rootworm (*Diabrotica virgifera*) (Oyediran et al. 2004), making it a promising alternative for Europe.

The key to improving the stress tolerance of sorghum for increasing its productivity is the understanding of how various abiotic stresses impact physiological processes, development stages, and the mechanism associated with tolerance. Sorghum is known for its excellent tolerance towards most abiotic stresses (Tari et al. 2013). Although water deficiency can cause adverse effects on plant development, many sorghum cultivars are adapted to semi-arid conditions (Patanè et al. 2013). It is also sometimes called “camel crop” because of its water use efficiency. However, drought stress is known to occur in sorghum during both pre- and post-flowering development stages (Kebede et al. 2001). While in the case of susceptible genotypes, pre-flowering drought stress results in delayed flowering, early termination of the floret, reduced seed-set and panicle

size, there is the occurrence of lodging, reduced grain weight, and premature senescence under post-flowering drought stress (Borrell et al. 2000a, b). Xu et al. (2000) reported tolerance to drought at this stage by a stay-green phenotype. Antifreeze proteins, heat shock proteins and dhurrin were found to be strongly associated with drought stress (Spindel et al. 2018; Rosati et al. 2019).

In contrast to drought, flooding stress or soil waterlogging stress causes an inadequate supply of oxygen to the submerged tissues, causing decreased growth and yield (Setter and Waters 2003). Sorghum grown in tropical and subtropical climates are often faced with heavy rains causing deleterious effects (Tari et al. 2013). Although sorghum shows moderate resistance towards salinity, the occurrence of genotypic differences is known (Krishnamurthy et al. 2007). Soil salinity causing sorghum yield losses usually occurs in arid and semi-arid regions (Koca et al. 2007), particularly affecting plants during the seedling emergence stage (Macharia et al. 1994). Reports suggest that the air temperature is increasing at the rate of 0.2°C, causing extreme weather changes, including severe droughts and erratic patterns of elevated temperature (Shukla et al. 2019). The optimum temperature for sorghum at the vegetative stage is 26–34 °C (Maiti 1996), while for the reproductive growth stage it is 31 °C (Prasad et al. 2006). Temperature deviating from these values can have a significant negative impact on sorghum growth and yield. Delays in panicle emergence (Prasad et al. 2006) and a decrease in floret fertility are often observed under extreme heat stress (Prasad et al. 2015).

Having tropical origin, sorghum is sensitive to low temperatures. Lower temperatures (<20°C) at early developmental stages result in reduced germination and emergence, poor seedling growth, and reduced vigor (Tiryaki and Andrews 2001a, b; Knoll and Ejeta 2008). Cold stress (<13 °C) at the reproductive stage is also detrimental to yield (Chakrabarty et al. 2021a). Reduction of pollen fertility (Downes and Marshall 1971; Osuna-Ortega et al. 2003) or total loss of seed set (Maulana and Tesso 2013) has been reported. According to Ortiz et al. (2017) thioredoxin, carotenoids, components of PSI, phytohormones, and antioxidants play a crucial role in cold tolerance. Regions associated with seedling survival and final emergence percentage under cold stress were found on chromosomes SB-01, SB-02, SB-03, SB-06, SB-09 and SB-10 (Parra-Londono et al. 2018). Over the years, multiple studies on sorghum chilling tolerance have been conducted (Upadhyaya et al. 2015; Schaffasz et al. 2019a; Marla et al. 2019) to identify diversity for temperate climate breeding programs. Kaoling from China (Franks et al. 2006), highland races from Yemen, Burundi, Uganda, and Ethiopia (Singh 1985) are a few of the promising sources. Osuna-Ortega et al. (2003) has developed varieties with good seed sets which can withstand night temperatures as low as 6°C during flowering time. Several major loci associated with chilling tolerance were found to co-localize with the classical grain tannin (Tan1 and Tan2) and dwarfing genes (Dw1 and Dw3) suggesting that chilling sensitivity was inadvertently selected due to coinheritance with the desired non-tannin and dwarfing alleles (Marla et al. 2019).

1.4 Sorghum breeding for temperate Europe

Natural factors like geography, climate, environment, and human activities shape the adaptability of a crop to a certain region (Chloupek et al. 2004). Sorghum, having a subtropical origin, requires extensive breeding for necessary adaptation to cooler climates. According to (Dahlberg et al. 2011), the earliest record of sorghum production in Europe dates to 1204 in Italy for broomcorn. However, the relative lack of research and knowledge compared to other crops like maize arrested the expansion of sorghum. Sorghum still holds minor importance in Europe, which contributes only around 2.2% of the total global sorghum production (FAO STAT 2022) (https://www.fao.org). So far, European production is concentrated in areas with hot summers, like southern France, Italy, Hungary followed by Romania, Ukraine, and southern Russia (Chakrabarty et al. 2021b). With the changing climate, the demand for alternative crops for silage and grains is steadily growing. Trends show that, over the past decade, sorghum production in Europe has increased 2-fold, as its importance as a resilient, "fail-proof" crop is being more recognized. In Europe sorghum is mostly used as fodder and for bioenergy. It is an excellent alternative to maize in terms of production, genetic diversity, abiotic and biotic stress resilience as well as end-use.

Sorghum has some clear advantages like tolerance of drought (Schittenhelm and Schroetter 2014) and flooding (Promkhambut et al. 2010) stress, but especially complete resistance against the western corn rootworm (Diabrotica virgifera) (Oyediran et al. 2004), which is a major cause of devastating damage to European maize production. Moreover, the production costs can be lowered by using less fertilization and crop protection and it also serves as a good preceding crop for spring crops such as sunflower or maize.

1.5 Germplasm mining and gene identification for expanding cold tolerance in European sorghum

Genetic resources are crucial to plant breeders to create novel gene combinations and selecting enhanced crop varieties better suited for different end goals. An enormous germplasm collection of more than 7 million plant accessions are stored across 1,750 genebanks worldwide FAO (2013). However due to lack of genetic information these collections are barely tapped (less than 1%) by breeders and researchers (Upadhyaya et al. 2006) and can be exploited to break crop productivity bottlenecks to accelerate yield gains.

The major obstacles in exploiting germplasm are the overwhelming scale of collection and our lack of resources and knowledge to use it, namely: a) characterization of novel germplasm at large scale and b) identify gene alleles and transfer them into cultivars successfully. With the advancement of high throughput phenotyping and genotyping technologies along with data management systems now have opened new avenues to use exotic plant materials in crop breeding and research (Wambugu et al. 2018; Mascher et al. 2019).

Africa is the home to the largest diversity of cultivated and wild sorghum (De Wet and Harlan 1971; De Wet 1978). Uganda, an East African country is situated at the center of sorghum genetic diversity (Doggett 1953; Mukuru 1993). The Plant Genetic Resources Centre at the Uganda National Gene Bank has conserved over 3000 sorghum accessions, the majority of which are genetically diverse landraces and wild relatives whose diversity has yet to be tapped for use in breeding programs. A wide range of variation among phenotypic traits exists for the Ugandan landraces. However, its genetic diversity, especially at a molecular level, has not been studied in depth (Mbeyaga et al. 2012).

With global warming, it can be seen that sorghum production areas in Europe are constantly increasing. Despite the advantages that the crop possesses like lower need for water and inputs, good hardiness to name a few, it is susceptible to low temperatures. The Ugandan sorghum germplasm is known to possess various morphological and agronomically important traits, such as panicle and grain architecture, maturity, plant height, and pigmentation (Mukuru 1993). Considerable variation in socio-cultural conditions and geographical variation like soil, temperature, and rainfall have shaped the immense genetic diversity of local sorghum varieties (Doggett 1970). For example, samples from Northern and Eastern Uganda are tolerant towards drought and salt toxicity. On the other hand, the accessions collected from the cold highlands of the Kigezi region in the southwestern part of Uganda are adapted to the low temperatures and can carry potential genes of interest, especially for Germany and other temperate climate zones. Inevitably, the sorghum genome of local accessions has undergone strong selection at gene levels, responsible for important adaptive and agronomic traits. Population diversity not only affects the potential of genetic gain and plays a key role in breeding programs (Allier et al. 2019) but also facilitates breeders to classify germplasms into heterotic groups to enhance hybrid breeding (Menz et al. 2004). Therefore, assessment of genetic diversity, population structure, and selection signatures is crucial for the improvement and conservation of valuable resources and utilization of its potential in different breeding programs. To date, the Ugandan sorghum germplasm has not yet been fully characterized and assessed at a molecular level, therefore it is under-utilized in breeding programs in Uganda and elsewhere (Chakrabarty et al. 2022).

1.6 Scope and aims

Addressing the constantly changing constraints in crop production requires a compound breeding approach that involves assistance of researchers, collaboration of breeding programs and use of modern technology. One of the crucial factors for the success of this approach is the access and utilization of germplasm with important traits like yield and tolerance to biotic and abiotic stresses. After identification of these traits, they can be introgressed into breeding programs to create new cultivars.

Over the years, germplasm resources have been used extensively to enhance crop production and create climate-smart crop varieties. Most of the breeding achievements across various crop

species started with extracting diversity from the landraces and incorporating new alleles into breeding programs.

Compared to most of the cereal crops, sorghum is an extremely versatile crop with a highly beneficial reservoir of untapped genetic resources. Studies indicate the increasing importance of sorghum, especially in areas that are predicted to become hotter with scarcity of resources. However, there needs to be more research in this area to deal with the changing environmental factors and utilizing sorghum to its full capacity in Europe.

The overall aim of this thesis was to characterize novel sorghum germplasm and identify genomic regions potentially involved in cold tolerance. The following major goals were pursued:

(i) Describe existing germplasm resources and developments for identification and implementation of useful diversity for stress adaptation traits in sorghum and comparing different genomics-based methods for breeding of complex, low-heritability traits. (Chapter 2)

(ii) The novel sorghum germplasm collection conserved at the Uganda National Gene Bank was characterized using whole genome SNP markers. Discriminant analysis of principal components (DAPC) was implemented to racially classify the accessions and characterize the sub-groups present within the diversity set. This dataset was also used to perform genome-wide association studies for identifying cold tolerance in sorghum at juvenile stage as an adaptive trait for temperate climate. (Chapter 3)

(ii) Evaluation of genetic architecture of reproductive cold tolerance in sorghum for expanding its production into temperate areas via genome-wide association studies and identification of superior and environmentally stable cold tolerance accessions (Chapter 4)

(iii) Provide insights and ideas about conservation and improvement of genebank resources through modern genomic technologies and the current and future state of sorghum breeding though implementations of genomics and bioinformatics approaches, especially towards its adaptation to temperate climate (Chapter 5)

2 Improving abiotic stress tolerance to adapt sorghum to temperate climatic regions

Subhadra Chakrabarty, André Schaffasz, Harmeet Singh Chawla, Maria Laura Federico, Rod J. Snowdon and Steffen Windpassinger

Molecular breeding in wheat, maize and sorghum: *strategies for improving abiotic stress tolerance and yield* (eds M.A. Hossain et al.). CABI, Wallingford, pp 444–462 (2021)
doi: 10.1079/9781789245431.0026

26 Improving Abiotic Stress Tolerance to Adapt Sorghum to Temperate Climatic Regions

Subhadra Chakrabarty[1], André Schaffasz[1], Harmeet Singh Chawla[1], Maria Laura Federico[2,3], Rod J. Snowdon[1] and Steffen Windpassinger[1]*
[1]*Justus Liebig University, Giessen, Germany;* [2]*CONICET, Buenos Aires, Argentina;* [3]*INTA, Famaillá, Argentina*

26.1 Introduction

Sorghum is a genus consisting of many species in different levels of domestication, comprising wild (e.g. *Sorghum arundinaceum*), weedy (e.g. *Sorghum halapense*) and domesticated ones (e.g. *Sorghum* × *almum*, *Sorghum sudanense* and *Sorghum bicolor*). *S. bicolor* (L.) Moench is the fifth most important cereal crop globally. It shows a remarkable diversity, including five different subspecies and their intermediates, and several crop types like grain, forage, sweet and broomcorn (Hariprasanna and Patil, 2015). Although it originates in the tropics of Africa, the remarkable scope of genetic diversity among the different subspecies has conferred an extraordinarily broad adaptability and a highly versatile range of end uses (Boyles *et al*., 2019). Sorghum has a particular advantage over crops like maize in extreme climates, where it achieves superior yields to maize (Farré and Faci, 2006; Staggenborg *et al*., 2008). Although sorghum is a major subsistence crop worldwide and an important component of industrial agriculture, it is frequently considered an orphan crop (Boyles *et al*., 2019) and breeding progress has been considerably slower than for most major global crops.

Despite its enormous potential and broad adaptive diversity, breeding to adapt sorghum for agricultural use in temperate Europe has progressed only slowly so far. Until recently, conventional breeding methods were the primary approach for genetic improvement of sorghum. These were mainly based on selection by visual phenotyping accompanied by introgression of desirable traits into elite germplasm. As in the early stages of classical breeding in most crops, these classical approaches present strong challenges to overcome linkage drag and maintain useful diversity in chromosome regions carrying essential adaptation genes under strong selection.

In recent times, modern molecular breeding tools like high-throughput molecular markers and genomic selection (GS) have shown great potential. Both can be combined to great effect with conventional breeding schemes to improve ecogeographical adaptation to new growing environments and further increase genetic gain for essential agronomic traits. For example, the major factors limiting adaptation of sorghum as

*Corresponding author: steffen.m.windpassinger@agrar.uni-giessen.de

 Molecular Breeding in Wheat, Maize and Sorghum: Strategies for Improving Abiotic Stress Tolerance and Yield (eds M.A. Hossain *et al*.)
DOI: 10.1079/9781789245431.0026

a cereal or biofuel crop in Northern Europe are its photoperiod sensitivity, along with a strong susceptibility to cool or cold temperatures at sowing and flowering (Windpassinger, 2016). Whereas photoperiod sensitivity is relatively simple to solve by conventional or marker-assisted backcrossing, adaptation to other abiotic stress is much more difficult to deal with in breeding programmes. This chapter reviews germplasm sources for and new developments in the identification and implementation of useful genetic diversity for temperate climate adaptation, along with genomics-based methods for breeding of complex, low-heritability traits like abiotic stress tolerance.

26.2 Origin and Genetic Diversity of Sorghum

The origin of early domestication of sorghum is hypothesized to have taken place in sub-Saharan Africa (present Ethiopia and Sudan) around 5000–8000 years ago (Mann *et al.*, 1983). From here, the species spread into different climatic zones of Africa, India, the Middle East and East Asia, initially via anthropogenic migration and later via human trade routes (Morris *et al.*, 2013). The lengthy process of adaptation to these different environments and ongoing selection for different kinds of human agricultural uses (grain, fodder, sugar, fibre) have created a vast genetic and phenotypic diversity among cultivated sorghum forms. Based on panicle and spikelet morphology the cultivated germplasm can be classified into five major races (*bicolor*, *caudatum*, *durra*, *guinea* and *kafir*) and ten intermediate races (Harlan and de Wet, 1972), representing the adaptation into different agroclimatic zones. Phylogenetic studies showed that genetic relatedness in sorghum accessions is predominantly based on geographic origin and race (Morris *et al.*, 2013). Similar to other crops, the identification and implementation of useful genetic diversity remain a paramount goal in sorghum breeding to ensure maintenance of genetic gain and constant replenishment of resistances to important pathogens. Recent studies have underlined the importance of a continuous integration of diverse, exotic material into modern breeding programmes (Jordan *et al.*, 2011), particularly the value represented by the secondary gene pool in the wild species *S. bicolor* subsp. *verticilliflorum* and *Sorghum propinquum* (Dillon *et al.*, 2007; Mace *et al.*, 2013; Muraya, 2014; Venkateswaran *et al.*, 2014).

26.3 Existing Molecular Tools to Enhance Abiotic Stress Tolerance

Since the origins of plant breeding as a systematic science, after the rediscovery of Mendel's laws just over a century ago, plant breeders have successfully developed new improved varieties by crossing and selection (Smýkal *et al.*, 2016). In recent decades, integration of new molecular breeding and biotechnological techniques into the process of crop improvement has contributed further to enhanced genetic gain and breeding has increased the productivity and sustainability of major crops (Voss-Fels *et al.*, 2019). The availability of various genomic tools and resources has led to a new revolution in plant breeding, as they facilitate the study of the genotype and its relationship with the phenotype, in particular for complex traits under strong environmental influence. The availability of a reference genome for sorghum (Paterson *et al.*, 2009), along with extensive additional genome sequence resources (e.g. Mace *et al.*, 2013; Morris *et al.*, 2013), greatly facilitates genomic-based breeding approaches. The relatively small diploid genome ($2n = 20$, ~730 Mb) and the high degree of synteny to maize and rice (Paterson *et al.*, 2009) greatly facilitate comparative genomics to take advantage of vast genomics resources from related monocot crops. Progress in this field for sorghum, however, has been relatively slow and limited due to its lower economic importance compared with other cereals.

Reduction in the cost of next-generation sequencing (NGS) technologies has led to mass sequencing of crop genomes and transcriptomes, which has facilitated the discovery of new genes and the development of vast collections of molecular markers like single-nucleotide polymorphisms (SNPs), which in turn are the basis for construction of high-density genetic maps. These facilitate mapping of genes and quantitative trait loci (QTLs) in bi-parental populations or via genome-wide association studies (GWAS),

helping to dissect genomic regions associated with complex phenotypes. In addition, reverse genomics approaches including mutagenesis approaches make it possible to screen mutant and germplasm collections for allelic variants of target genes. All these tools and resources facilitate exploration of the genetic diversity and the developed markers can be incorporated into breeding programmes for marker-assisted selection (MAS). Common strategies include marker-assisted backcrossing with trait-linked foreground markers and genome-wide background markers, 'breeding by design' or new strategies like GS. Sorghum was one of the first major crops whose genome sequence was assembled. More recently, large-scale sequencing of transcriptomes (Makita *et al*., 2015) has opened up new avenues for the global sorghum research community and has drastically simplified SNP marker identification and development. Sorghum genomic databases like SorghumFDB (Tian *et al*., 2016) have combined various genomic data and functional annotations to create multidimensional biological relationships which could assist in sorghum functional genomics analyses and help in effective crop improvement.

Early genetic linkage maps for sorghum were constructed using labour-intensive or dominant markers such as RFLP (restriction fragment length polymorphism), AFLP (amplified fragment length polymorphism) and RAPD (random amplified polymorphic DNA) (Hulbert *et al*., 1990; Berhan *et al*., 1993; Boivin *et al*., 1999; Peng *et al*., 1999; Singh and Lohithaswa, 2006; Ejeta and Knoll, 2007). Although these maps played an important role in early sorghum genetic research, they were later superseded by more informative markers like simple sequence repeat (SSR) microsatellite markers, which were commonly used for sorghum gene mapping, genome evolutionary studies, molecular genetics and marker-assisted breeding (Tao *et al*., 1998; Xu *et al*., 2001; Yonemaru *et al*., 2009; Kong *et al*., 2013).

However, these technologies were limited by their restricted genome coverage and the relatively low number of polymorphic markers. Hamblin *et al*. (2004) first integrated SNP markers in sorghum to study genetic variation. The Diversity Arrays Technology (DArT; Canberra, Australia) was later applied to integrate multiple-component sorghum genetic maps to create a consensus map (Mace *et al*., 2009). Presently, SNP markers are widely used in sorghum for studies tackling the genetic control of abiotic stress tolerance (see below).

26.4 History of Temperate Adaptation in Sorghum

Sorghum is originally a photosensitive short-day plant, conferring the best adaptation to its centre of origin in the semi-arid tropics of the Sahel zone. In these environments, the rainy season ends quite reliably at a latitude-specific time in autumn, with a day length below 12 h, whereas the onset of the new rainy season can vary strongly from year to year. Hence, local landraces are best adapted when flowering starts around 20 days before onset of the dry season, regardless of their sowing time, to allow for sufficient water supply during anthesis and grain-filling period and dry conditions during ripening (Guitton *et al*., 2015). The relatively simple genetic architecture of photoperiodism in sorghum, which is controlled by four major maturity loci designated *Ma1*, *Ma2*, *Ma3* and *Ma4* (Quinby and Karper, 1945; Quinby, 1966), facilitated the generation of useful mutants. Dominant alleles at these loci induce photosensitivity, with *Ma1* having the largest impact (Klein *et al*., 2008). A recessive mutation at this locus alone is sufficient to allow flowering under longer days in extra-tropical environments, and corresponding mutations occurred independently in different parts of the world following geographic dispersal of sorghum. By tracing allelic variants of the underlying gene *PRR37*, new insights were obtained into the historical expansion of sorghum into temperate areas of South Africa, China, Europe and the USA (Klein *et al*., 2015). Sorghum was probably introduced into China as early as AD 400 via trade routes from India. Subsequently, under strong selection pressure for photoperiod insensitivity and early vigor, grain type *kaoliang* and broomcorn diverged there. From China, sorghum was brought to Europe (Klein *et al*., 2015), where its first description dates back to 1204 in the Piemont region of Italy (Becker-Dillingen, 1927). However, in contrast to new-world crops such as maize and potato, which arrived several centuries later

but were rapidly adopted, the importance of sorghum for European agriculture remained limited for a considerable time. It was relatively widespread from the 16th to 18th centuries in Southern and South-Eastern Europe; however, during that time its utilization was confined to broomcorn (Dahlberg *et al*., 2012), implying low selection pressure for grain yield and further adaptive traits. The introduction of sorghum into North America occurred as broomcorn from Europe during the 1750s (Berenji *et al*., 2011) and, more importantly, as grain and sweet sorghum arriving from Africa on slave ships during the first half of the 19th century (Sleper and Poehlman, 2006).The number of founder cultivars was low and sorghum cultivation was initially limited to subtropical areas of Texas. However, farmers soon selected early-maturing mutant plants, corresponding to the previously described mutations at the *Ma* loci, with a short stature. Similarly to photoperiodism, plant height in sorghum is determined by four major *dwarf* loci (designated $Dw_{1–4}$) with dominant alleles conferring tallness (Quinby and Karper, 1945; Multani *et al*., 2003; Hilley *et al*., 2016, 2017).

Systematic, modern sorghum breeding started in the USA during the first half of the 20th century and the combination of different desirable mutants facilitated the release of early-maturing cultivars that enabled cultivation as far north as Nebraska (Klein *et al*., 2008). Impressive yield gains were achieved in the 1950s by changing from line to hybrid breeding (see below). Nevertheless, the narrow genetic base of photoperiod-insensitive breeding lines was soon recognized as a bottleneck for further yield gains and improvements in abiotic and biotic stress tolerance. To broaden the genetic diversity for temperate sorghum breeding, the 'Sorghum Conversion Program' was initiated in 1963 by the US Department of Agriculture (USDA). Taking into consideration the inheritance of photosensitivity and plant height in sorghum, a backcrossing programme was conducted to convert genetically diverse tropical accessions to early-maturing lines suitable for combine harvesting. Through this programme, about 850 converted and partially converted lines have been developed (Stephens *et al*., 1967). Owing to the huge impact of this programme, most temperate sorghum hybrids today have conversion lines in their pedigree (Gabriel, 2005).

In theory, fully converted sorghum genotypes were expected to consist of 97% recurrent tropical parent genome. However, the ability to visualize and characterize genome introgressions using whole-genome sequencing or SNP genotyping revealed that the recovery of the exotic genome in backcrossed progenies containing desirable *dwarf* and *maturity* alleles was not as complete as assumed. Extensive stretches of the donor genome remained in linkage drag, for example on sorghum chromosome Sb06, which harbours crucial adaptive loci (*ma1* and *dw2* genes) (Klein *et al*., 2008). As a result, little functional diversity in temperate sorghum genotypes has remained on this chromosome, which contains roughly 10% of all sorghum genes. This severely limits the adaptive potential especially for complex traits (Thurber *et al*., 2013), presenting a key target for genomics-assisted breeding.

26.5 Potential of Sorghum in Temperate Climates

Underlining its adaptive potential, a substantial proportion of the world's sorghum harvest is today produced far away from its origin in tropical Africa, with countries like the USA, Mexico, Argentina, China and Australia among the main producers. However, in these countries, sorghum is mainly still grown in hot and dry, predominantly subtropical environments, with little expansion of production into temperate high-latitude areas. In Europe, sorghum has not yet achieved more than minor importance and so far the production is also concentrated in areas with hot summers, like southern France, Italy, Hungary, Romania, Ukraine and southern Russia. Altogether, the sorghum acreage in Europe of around 465,000 ha comprises only 1% of the global acreage. In contrast, maize is planted on 18,000,000 ha in Europe, spanning as far north as southern Scandinavia. This comparison with maize is relevant for two main reasons. First, maize is also a C_4 crop derived from the tropics, so its adaptation and expansion into cool temperate areas may serve as a blueprint for sorghum. Second, maize has a highly similar production technique and provides the same range of major uses as sorghum (grain, silage,

bioenergy). Presently, maize is the crop of choice for biofuel and feed in temperate zones and the production system is very well established. However, there are several factors favouring an increase in sorghum acreage at the expense of maize in temperate Europe. Most importantly, the continuing expansion of maize monoculture on to vast production areas in Europe presents serious phyto-pathological threats like the western corn rootworm *Diabrotica virgifera* (Wesseler and Fall, 2010), especially in light of European policies on non-use of genetically modified organisms which preclude cultivation of transgenic *Bt* maize as an effective solution to this pest. Sorghum as a non-host (Oyediran *et al*., 2004) may become one of the most compatible, viable alternatives to maize in quarantine areas. Furthermore, sorghum has a greater tolerance against most types of abiotic stress than maize. In the face of growing environmental and agro-ecological concerns with regard to climate change and agricultural sustainability, the high nutrient efficiency of sorghum can be a key asset, allowing for significant reduction of N fertilization. Its high water-use efficiency and drought tolerance are becoming more important in the context of climate change, enabling satisfactory yields in recent hot and dry European summers where other crops have frequently failed. Surprisingly, sorghum also tolerates temporary waterlogging considerably better than maize (Promkhambut *et al*., 2011), ensuring good yield stability even in extreme or fluctuating environments. Despite its potential, as a tropical C_4 plant sorghum needs extensive breeding effort for adaptation to temperate climates before it can be broadly established in areas like Central Europe (Windpassinger *et al*., 2015). In particular, its sensitivity to chilling represents a major constraint. Again, however, the history of North American and European maize breeding demonstrates that successful adaptation of a highly diverse tropical C_4 plant into temperate environments is feasible, suggesting similar possibilities for sorghum. However, the prerequisites for maize were arguably more advantageous, with a far longer history of temperate adaptation than sorghum. Already in pre-Columbian times, maize had spread into temperate North America up to what is today southern Canada (Matsuoka *et al*., 2002), and recent results indicate a divergence between tropical and temperate maize as early as 3400 years ago (Liu *et al*., 2015). Successful maize introductions to Central Europe from the 16th century onwards consisted of these 'pre-adapted' temperate maize types (*flint* variety group). Subsequently, adapted *flint* landraces developed under a strong selection pressure for early maturity and tolerance to cool spring temperatures. The fast expansion of maize acreage in Central Europe during the second half of the 20th century benefited considerably from (i) the existence of adapted *flint* germplasm and (ii) their heterotic pattern with North American *dents*, which allowed an optimal exploitation of heterosis in well-adapted and high-yielding hybrids. Both of these prerequisites exist only vaguely in sorghum, although Chinese *kaoliang* forms might eventually play a similar role as European *flint* in terms of early ecogeographic adaptation to temperate climatic zones. However, thanks to the modern molecular breeding techniques presented in the following sections, enhancement of quantitative traits like abiotic stress tolerance can potentially be achieved at a significantly faster pace today than during the early days of European maize breeding in the last century, provided sufficient investment is possible. Furthermore, access to extremely diverse tropical sorghum materials, including germplasm with pre-existing cool-temperature adaptation from growth at altitude in the highlands of Central and East Africa, is facilitated by a simple inheritance of photosensitivity. This allows for rapid backcrossing programmes, as successfully proven by the Sorghum Conversion Program in the USA.

26.6 Breeding Goals for Sorghum Temperate Adaptation and Their Present State-of-the-Art

Enhancements in abiotic stress tolerances, especially chilling tolerance, are the principal breeding goals required to establish stable sorghum productivity in temperate areas. Obviously, yield and adequate maturity, as the final outcome of genotype × environment interactions, have greater economic importance; however, in temperate sorghum production systems these traits are highly intertwined with chilling tolerance.

26.6.1 Juvenile chilling tolerance

Due to its tropical origin, sorghum generally does not tolerate frost and requires temperatures of more than 20°C for optimal growth. Lower temperatures induce different grades of chilling stress (see Fig. 26.1) and are especially problematic during emergence and seedling establishment (Pinthus and Rosenblum, 1961; Peacock, 1982). An improved juvenile chilling tolerance is thus mandatory for a successful adaptation to higher latitudes, since it would allow for earlier sowing, enhancing yield potential and maturity due to a longer growth period. Presently, in temperate areas such as Central Europe, sorghum is still sown several weeks later than maize, implying a loss of growth days which explains most of its present yield penalty in comparison to maize. However, improved early chilling tolerance can also be beneficial for some subtropical regions where sorghum is already well established, since earlier sowing in spring can potentially allow a better utilization of winter moisture (Patane *et al.*, 2006).

Several studies have been undertaken to mine sources of chilling tolerance in sorghum (Singh, 1985; Salas Fernandez *et al.*, 2014). Basically, these studies coincide in the identification of chilling-tolerant germplasm among Chinese *kaoliangs* and tropical highland accessions (e.g. from Ethiopia, Uganda and Yemen). Singh (1985) studied juvenile and pre-flowering cold tolerance in 380 accessions, with a focus on materials from China, Ethiopia, Uganda and the USA. The highest cold tolerance was observed among the Ethiopian and Ugandan accessions Alemaya70, Jewegere 935, Muyra, Mabere, Magune and Nyundo. The Chinese accession PI 610727 was highlighted as highly cold-tolerant by Franks *et al.* (2006) and further used by Burow *et al.* (2011) as the cold-tolerant parent in a bi-parental QTL mapping population. Salas Fernandez *et al.* (2014) tested 38 *kaoliangs* and 18 non-*kaoliangs* and identified new tolerant material from China, Korea and Russia. *Kaoliangs* are often highly cold-tolerant but have poor agronomic characteristics which limit their direct use in breeding programmes (Franks *et al.*, 2006).

In a recent study conducted under Central European field conditions, Schaffasz *et al.* (2019a) revealed that valuable sources for chilling tolerance and early vigour can also be found among existing US sorghum conversion lines. The accessions SC614 and SC1201 performed best for emergence and SC702 best for early shoot biomass. In contrast to *kaoliang*, these conversion lines are more amenable for breeding of grain sorghum. In recent years, increasing research efforts have begun to dissect the genetic architecture of juvenile chilling tolerance in

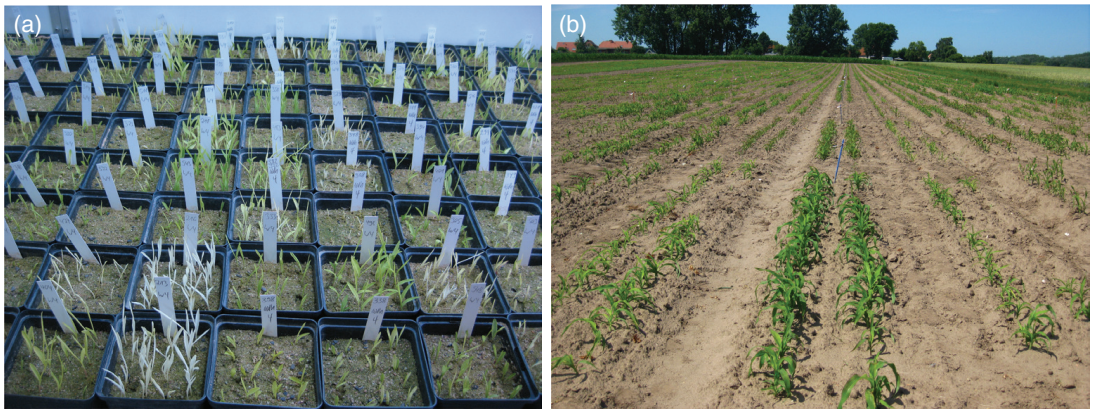

Fig. 26.1. Variation for chilling tolerance and early vigour among different sorghum accessions. (a) Different reaction of sorghum accessions to controlled chilling-stress conditions (13°C day/10°C night during emergence and subsequent growth). Susceptible genotypes show poor emergence and development, and/or complete chlorophyll degradation (white leaves), while tolerant accessions show satisfying emergence and maintenance of photosynthetic apparatus (green leaves). (b) Sorghum genotypes (sown in two-rowed plots) show remarkable variation for establishment and early vigour in a field experiment in Germany.

sorghum by QTL studies in segregating bi-parental populations.

Knoll *et al*. (2008) dissected the early-season cold tolerance in sorghum using a recombinant inbred line (RIL) population derived from a cross between Shan Qui Red (SQR, cold-tolerant) and SRN39 (cold-sensitive). They identified two QTLs for germination under cold stress on linkage group SBI-03a (on chromosome Sb03) and on group SBI-07b (chromosome Sb07), both of which showed significant trait associations under cold temperatures. A region on chromosome Sb01 derived from SQR showed strong associations with seedling emergence and seedling vigour scores under early and late field plantings. One QTL for both early and late emergence and another QTL for early vigour were identified on Sb02 and Sb04, respectively. Shortly after, 14 QTLs associated with different cold tolerance traits were detected on chromosomes Sb01, Sb02, Sb04, Sb07 and Sb09 (Burow *et al*., 2011). In particular, Sb09 was shown to harbour four QTLs for field emergence that co-localized with QTLs for cold germinability. Bekele *et al*. (2014) identified highly interactive epistatic QTL hotspots, including a previously unknown QTL on Sb06 with a significant effect on prolonged chilling survival, which were found to regulate different physiological mechanisms contributing to maintenance of growth and development even under chilling temperatures.

During the past decade, genetic association studies using various diversity panels have accelerated the identification of genome regions and promising candidate genes highly influencing cold tolerance during emergence. Fiedler *et al*. (2012, 2014, 2016) reported multiple cold tolerance QTLs and identified gene *Cold-Shock Domain Protein 1* (*CSDP1*) as a potential positional and functional candidate. One QTL region on chromosome Sb06 was identified as a putative hotspot for temperature-mediated seedling emergence and survival. This region was later independently verified (Parra-Londono *et al*., 2018). Recently, Schaffasz *et al*. (2019a) presented the first study to jointly analyse both agronomical and cold tolerance traits on a broad diversity set under Central European conditions. The findings from these studies show the potential of GWAS to help dissect the genetic complexity of cold temperature susceptibility, an important prerequisite for development of temperature-resilient sorghum cultivars and further characterization of genomic regions responsible for adaptation to thermal stresses (Chopra *et al*., 2017).

Recently, Marla *et al*. (2019) applied a nested association mapping (NAM) approach to investigate chilling tolerance in a multi-parental population, identifying ten loci explaining 20–41% of the phenotypic variation within these US sorghum accessions. Surprisingly, the results showed the co-inheritance of chilling tolerance loci with wild-type alleles of classical tannin (*Tan1* and *Tan2*) and dwarfing genes (*Dw1* and *Dw3*), four of the five most important genes under selection by US sorghum breeders in the 20th century. The fifth of these, *Maturity1*, did not co-localize with chilling tolerance QTLs. Because there is no clear evidence to suggest that *Tan1*, *Tan2*, *Dw1* or *Dw3* is directly involved in cold tolerance responses, this association seems more likely to be caused by the substantial linkage drag surrounding these gene variants rather than a negative pleiotropic association between the traits. In other words, this result indicates that strong selection for essential recessive adaptation alleles in modern grain sorghum resulted in loss of early-season chilling tolerance by linkage in repulsion. In consequence, the authors suggest revision of the original model of sorghum chilling sensitivity to be only caused by its tropical origin, since African sorghums introduced into the USA harboured basal chilling tolerance which was subsequently lost by breeding. Altogether, the results of the previously discussed mapping studies are concordant and coincide in a strongly quantitative character for all juvenile chilling tolerance-related traits. For practical breeding, these results imply strong limitations regarding the possibilities of MAS. Nevertheless, these studies provide valuable insights into the underlying physiological mechanisms of abiotic stress tolerance. Of special interest are the results of Marla *et al*. (2019), which provide explanations on how the co-inheritance of chilling sensitivity with desired traits has hampered breeding efforts during the last decades. The recently developed Sorghum QTL Atlas platform (Mace *et al*., 2019), which facilitates identification of candidate genes in sorghum and their comparison across related species like maize and rice, provides important new data about linkage relationships among genome-wide QTLs for important agronomic traits and can provide a

starting point (foreground and background markers) for endeavours to identify useful recombinants that disrupt linkage in repulsion between key adaptation and chilling tolerance loci.

26.6.2 Reproductive chilling tolerance

Pre-flowering reproductive stage is the second sensitive developmental phase in sorghum affected by temperatures below 15°C (Singh, 1985). Depending on duration and intensity of the stress, male sterility can be induced in sorghum, leading to a reduction or even, in extreme cases, a complete loss of seed yield (see Fig. 26.2) (Downes and Marshall, 1971; Osuna-Ortega *et al.*, 2003). For adaptation of sorghum to temperate climates, reproductive chilling tolerance is at least equally as important as juvenile chilling tolerance. While farmers can choose later sowing dates to reduce juvenile chilling stress (albeit at the expense of yield potential), there is no escape strategy for cold nights during the critical reproductive stage in summer. Tropical high-altitude environments could also benefit from sorghum varieties with improved reproductive chilling tolerance. In contrast to temperate high-latitude environments, where chilling stress at pre-flowering stage occurs rather infrequently followed by intervals of warmer weather, tropical highlands tend to have constantly cool nights, making them suitable selection environments also for temperate breeding programmes. Reproductive chilling tolerance is also considered to be potentially beneficial for sorghum cultivation in the Indian post-rainy season (*rabi*) (Krishnamurthy *et al.*, 2014).

The first scientific description of male sterility in sorghum after a cold treatment (13°C) was provided by Downes and Marshall (1971), who were originally seeking a new crossing method. Brooking (1976) described problems with meiosis in mother-spore cells as a possible reason for this phenomenon. Singh (1985) scored reproductive chilling tolerance (along with juvenile chilling tolerance, see above) in a set of 380 accessions, identifying several tolerance sources. Notable breeding efforts were undertaken to develop sorghum varieties with enhanced

Fig. 26.2. Variation for sorghum reproductive cold tolerance in a field experiment in Germany: tolerant line with high pollen viability also at cool temperatures, and hence full seed set (left); susceptible line with cold-induced male sterility, resulting in almost no seed set (right).

reproductive chilling tolerance for the Mexican High Valleys (>2000 m above sea level) (Mendoza, 1988; Osuna-Ortega *et al*., 2000, 2003; Leon-Velasco *et al*., 2009; Cisneros-López *et al*., 2010), using cold-tolerant accessions from Africa and India donated by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)–International Maize and Wheat Improvement Center (CIMMYT) as base material (Leon-Velasco *et al*., 2009). As a result, varieties with a satisfying seed set and yield even at night temperatures of 6°C during the critical stages could be developed (Osuna-Ortega *et al*., 2003), underlining the feasibility of genetic improvement of this trait. Maulana and Tesso (2013) recommended the accession Shan Qui Red (a Chinese *kaoliang* known for good early chilling tolerance) as a tolerance source also for reproductive chilling tolerance. In contrast to juvenile chilling tolerance, surprisingly little is known about the genetic architecture of reproductive chilling tolerance in sorghum to date. Both Singh (1985) and Schaffasz *et al*. (2019b) described a more or less dominant inheritance. Interestingly, the observed heritability for seed set traits under stress (Schaffasz *et al*., 2019b) was notably higher than for juvenile chilling tolerance traits (Windpassinger, 2016; Schaffasz *et al*., 2019a). However, GWAS or QTL mapping for this trait still needs to be performed to unravel its quantitative genetic control.

26.7 Breeding Methods

The type of sorghum and the purpose for its production vary widely depending on the region where it is grown. Sorghum is a predominantly self-pollinating crop. The level of cross-pollination depends on panicle architecture and weather conditions and increases under stress (Osuna-Ortega *et al*., 2003). Thus, breeding procedures applicable to both self- and cross-pollinated crops can be deployed to sorghum improvement (Rakshit and Bellundagi, 2019). However, for both pedigree breeding and hybrid breeding the primary goals of sorghum breeders throughout the world are always grain/biomass yield, adaptation, stress tolerance and product quality.

Heterosis (hybrid vigour) in sorghum was already described by Conner and Karper (1927) but unlike monoecious maize, the perfect flowers of sorghum prevented hybrid seed production on a commercial scale until 1952, when both cytoplasmic-male sterility (CMS) and fertility restorers possessing dominant *Rf* alleles were discovered in the USA (Stephens and Holland, 1954). Commercial CMS-based hybrid seed production began in 1956 and only four years later, the proportion of sorghum production from hybrid cultivar in the USA reached 95%, resulting in a doubling of grain yield compared with 1952 (Quinby, 1974; Smith and Frederiksen, 2000). Presently, sorghum production in regions with an industrialized, commercial agriculture (USA, Latin America, Australia and Europe) relies almost completely on hybrids, while open-pollinated landraces are still predominantly used for subsistence agriculture in Africa.

26.7.1 Enhancements of abiotic stress tolerance via heterosis and hybrid breeding

Heterosis in sorghum is not only expressed for grain and biomass yield, but also for maturity (Kirby and Atkins, 1968) and abiotic stress tolerance (see Fig. 26.3), including juvenile (Pinthus and Rosenblum, 1961; Yu and Tuinstra, 2001; Windpassinger, 2016) and reproductive chilling stress (Leon-Velasco *et al*., 2009; Schaffasz *et al*., 2019b). Efficient and successful hybrid breeding requires the development of complementary heterotic pools with a sufficiently high genetic distance between them. Well-designed heterotic pools ensure a consistent exploitation of heterosis by increasing the relative contribution of general combining ability (GCA) effects in comparison to effects from specific combining ability (SCA) (Reif *et al*., 2005; Schnable and Springer, 2013). Sorghum heterotic pools are not yet defined as clearly as they are in maize (Monk *et al*., 2014). However, the availability of cost-effective molecular markers provides new opportunities to evaluate the phylogenetic and genomic structure of accessions to establish genetically distinct pools for temperate sorghum breeding programmes. Pre-existing heterotic patterns can be easily compromised by arbitrary crosses, a frequent occurrence prior to the implementation of molecular marker

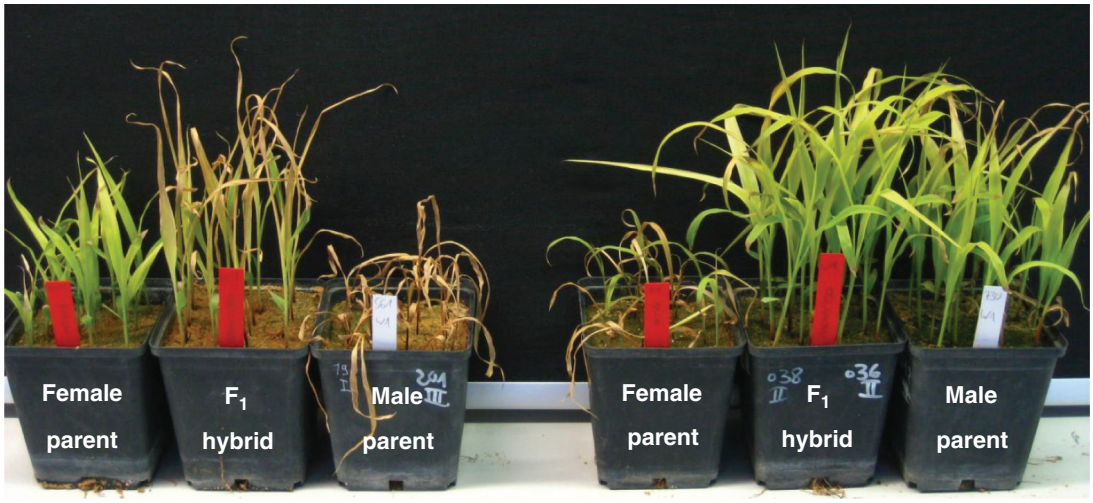

Fig. 26.3. Reaction of two different sorghum F_1 hybrids and their respective parental lines to prolonged chilling stress (13°C day/10°C night in a climate chamber experiment) induced after emergence at warm temperatures. While heterosis for early biomass production is clearly visible, the ability to survive prolonged chilling (expressed in the maintenance of green leaves) seems to be a rather additive trait.

techniques in sorghum hybrid breeding (Menz *et al.*, 2004). On the other hand, genome-wide markers can also help to rapidly characterize genetic diversity and design genome-assisted cross schemata for separation of heterotic pools.

The development of heterotic pools can potentially make a significant contribution to enhancement of abiotic stress tolerance in sorghum hybrids. An important question for breeders is to what extent hybrid performance can be predicted based on per se line performance. For juvenile chilling tolerance, Windpassinger (2016) showed that line performance per se is in fact a poor predictor of hybrid performance. Hence, an overly strict selection of hybrid parents based on the per se tolerance is in fact counterproductive, whereas GCA tests seem to be a more efficient and useful approach. For emergence and early heterotrophic growth of hybrids, the impact of the female parent is known to be higher (Yu and Tuinstra, 2001; Windpassinger 2016), suggesting priority should be given to improvement of the female pool in a hybrid breeding programme. For reproductive chilling tolerance, correlations observed between per se and hybrid performance were somewhat higher than for juvenile chilling tolerance (Schaffasz *et al.*, 2019b). Nevertheless, GCA tests also appear to be the preferable selection method for this trait. Due to a high GCA:SCA ratio and low GCA × environment interaction, robust enhancements of reproductive chilling tolerance via hybrid breeding seem to be feasible (Schaffasz *et al.*, 2019b) and a more systematic exploitation of heterosis using genomic tools may simultaneously help improve genetic gain for chilling tolerance. In consequence, future association studies and GS approaches for chilling tolerance should focus rather on GCA than per se performance (as done in the past) as a basis for successful hybrid breeding towards more robust and stable plant establishment, pollen fertility and seed set under cool-temperature conditions.

26.8 Advancement and Use of Genomics and Bioinformatics Approaches

26.8.1 High-throughput genotyping tools

The major prerequisite for application of genomics in genetic analyses or breeding of complex quantitative traits is the availability of suitable platforms and SNP marker panels for rapid, cost-effective, genome-wide marker screening. In sorghum, large population genomics studies have been achieved with sequencing-based marker techniques (see below), but molecular

breeding efforts for most other major crops have generally been based on dedicated SNP array genotyping platforms. SNP arrays have a number of advantages for breeding in comparison to sequencing approaches, not least the ability of service providers who can deliver low-cost genotype data sets without the need for breeders to have access to their own sophisticated molecular genetics laboratories or bioinformatics facilities. With an SNP array a fixed marker panel is genotyped for every individual and the customer/breeder is provided with a simple spreadsheet containing genotype calls. This means that considerably less bioinformatics analysis is required than for derivation of SNP variants from genotyping-by-sequencing (GBS) data, for example. In contrast to other major crops like wheat, maize, canola, barley or soybean, however, until recently there has not been a major push to develop a community-driven public SNP genotyping array platform. Bekele *et al*. (2013) developed a small-scale Illumina Infinium 3K SNP genotyping array with 2620 SNP markers and demonstrated its implementation for genetic mapping, diversity analyses and GWAS (Bekele *et al*., 2014). Later, large-scale genomic resequencing data were used to generate a 90K SNP Affymetrix Axiom genotyping array for GWAS (Parra-Londono *et al*., 2018); however, to date, no commercial sorghum SNP chip has been made available for public use. In 2019, efforts were initiated to establish a private–public consortium for development of a low-cost Illumina SNP array for sorghum breeding, but to date most large-scale genotyping efforts have been carried out using sequencing-based genotyping technologies.

26.8.2 Use of next-generation-sequencing genotyping techniques in sorghum

Incredible progress has been made in modern DNA sequencing technologies and accompanying bioinformatics methods in recent years. NGS technologies can be utilized for identifying the genetic basis of agriculturally important traits and for predicting the breeding value of individuals in a plant breeding population (Varshney *et al*., 2014). Detailed phenotyping of multi-parental NAM populations (Jordan *et al*., 2012; Marla *et al*., 2019) has been applied in association with high-resolution sequencing-based sorghum genotype data to dissect different abiotic stress traits.

Sequence-based genotyping platforms that have been applied for sorghum include restriction site-associated DNA sequencing (RAD-seq) (Nelson *et al*., 2011) and GBS (Morris *et al*., 2013). These two similar, reduced-representation, genome-wide resequencing methods are capable of identifying, sequencing and genotyping thousands of markers across the genome at low cost in large populations, making them highly suitable for genome-wide analyses of complex traits. In a more targeted approach, Ji *et al*. (2017) implemented genome-wide specific-locus amplified fragments (SLAF) markers, which are highly abundant and evenly distributed across the genome and thus facilitate the scanning of the sorghum genome for gene mining. The main advantage of NGS-based genotyping platforms compared with arrays is their lack of ascertainment bias in the markers assayed, improving their potential for discovery of novel variants of interest for trait improvement. The large numbers of low-cost SNP markers generated by NGS-based genotyping systems make them an economical option for GS in modern breeding programmes. Several studies have shown the potential of GS to enhance abiotic stress tolerance traits like heat and drought tolerance in major cereal crops like rice and maize (Yuan *et al*., 2018; Bhandari *et al*., 2019; Trachsel *et al*., 2019), and recently Velazco *et al*. (2019) demonstrated the efficacy of GS to enhance the stay-green trait in sorghum. To date, however, GS is still at its infancy in sorghum for improvement of abiotic stress traits.

26.8.3 Transcriptome analysis

Coupled with precise phenotyping and proper gene annotations, functional genomics can provide crucial information regarding complex biological processes like abiotic stress responses. The impact of transcriptome analysis in sorghum increased rapidly after the completion of the first reference genome and with the advent of next-generation molecular tools. Remarkable progress has been made in regard to transcriptome

analysis of traits like drought tolerance, cold stress, heat and salinity in sorghum (Fracasso *et al.*, 2016; Bashir *et al.*, 2019). It has been shown that cold stress induces osmotic stress and the expression of transcription factors for protein kinase genes is altered (Bashir *et al.*, 2019). Kadier *et al.* (2017) identified upregulation of the sorghum NAC-transcription factor family genes *SbNAC17* and *SbNAC73* in leaf tissues under cold stress conditions. The role of NAC transcription factors in general abiotic stress-response regulation has been described in many plant species (see Shao *et al.*, 2015 for a review).

Woldesemayat *et al.* (2018) recently introduced an integrated approach to mine for candidate stress genes across species by combining ontology-based semantic data integration with expression profiling, comparative genomics, phylogenomics, functional gene enrichment and gene enrichment network analysis. As a result, 221 cold stress genes were identified in sorghum and were validated using ontology mapping. In addition, a phylogenetic tree was constructed to infer the evolutionary relationship of the sorghum orthologues.

26.8.4 Genetic transformation

To intensify the plant development, genetic transformation has proved to be a powerful tool for gene induction, modulation and expression (Gurel *et al.*, 2009). However, sorghum has been classified as one of the most challenging plant species to perform tissue culture and genetic transformation (Zhu *et al.*, 1998). *Agrobacterium*-mediated and particle bombardment transformation are the two main approaches that have been exploited for the development of transgenic sorghum (Ahmed *et al.*, 2018). Over the years, there have been a few studies where sorghum was genetically transformed to dissect gene complexity of traits like soil salinity tolerance, protein and tannin content (Yellisetty *et al.*, 2015; Kuriyama *et al.*, 2019; Liu *et al.*, 2019) but the efficiency is still behind other major crops like rice, maize and barley (Che *et al.*, 2018). Since traits like juvenile cold tolerance are controlled by a large number of loci and exhibit low heritability (Bekele *et al.*, 2014), it remains challenging to design genetic transformation strategies for improving these traits.

26.8.5 TILLING

Although great progress has been made in sorghum genomics, the availability of mutant lines for functional studies via reverse genetics is limited. Chemical mutagenesis of sorghum germplasm, followed by screening for mutants altered in important agronomic traits by targeting induced local lesion in genomes (TILLING), represents a rapid and effective means for studying agronomically important genes (Xin *et al.*, 2008). Jiao *et al.* (2016) identified potential genes involved in drought tolerance using TILLING mutants. Similarly, other abiotic stress traits like cold tolerance and photoperiodism can also be studied. Using EcoTILLING, Bharathi *et al.* (2016) screened naturally occurring mutations in potential candidate genes to study several agronomically important traits in sorghum. Undoubtedly, a large-scale resource of well-characterized mutants and naturally occurring genetic variation would provide an efficient platform for functional validation of genes in sorghum, thereby accelerating sorghum breeding.

26.9 Future Prospects

Sorghum is an important failsafe crop which provides food, feed, fuel and fodder in many countries around the globe. It can be used as a model for other C_4 crops because of its extensive collection of diverse germplasm, genetic and genomic resources, and breeding information. Recent advancements in NGS, high-throughput phenotyping and bioinformatics tools are helping to accelerate genetic gain in sorghum across different climatic zones. Despite efforts to improve genetic and genomic resources, many such resources are still decentralized and independent. Increased efforts to coordinate cooperation among complementary public research programmes could further help to integrate research platforms available for meta-analysis of complex adaptive traits like chilling tolerance. Using information from related monocot crops can critically increase the power of comparative genomics and help dissect adaptive traits to enhance crop improvement. New molecular breeding methods and tools are already promoting considerable progress in plant breeding, including fast-track

adaptation and the genetic dissection and breeding for complex abiotic stress traits (Pérez-de-Castro *et al*., 2012). The integration of large-scale molecular marker data sets, high-density genetic maps, genome and transcriptome sequences with quantitative genetics can help translate functional genomics knowledge into genome-based improvements in modern breeding populations. This will help to further accumulate base knowledge to develop a framework for implementing GS for sorghum improvement. As a consequence, the historical division between breeding and genomics is becoming increasingly blurred (Deshpande *et al*., 2017). Crops like sorghum, which to date have been bred by traditional means but for which exceptional genome resources are available, stand to benefit greatly from genomics-based breeding applications.

References

Ahmed, R.I., Ding, A., Xie, M. and Kong, Y. (2018) Progress in optimization of *Agrobacterium*-mediated transformation in sorghum (*Sorghum bicolor*). *International Journal of Molecular Sciences* 19(10), 2983. Available at: https://doi.org/10.3390/ijms19102983

Bashir, K., Matsui, A., Rasheed, S. and Seki, M. (2019) Recent advances in the characterization of plant transcriptomes in response to drought, salinity, heat, and cold stress. *F1000Research* 8, 658. Available at: https://doi.org/10.12688/f1000research.18424.1

Becker-Dillingen, J. (1927) *Handbuch des Getreidebaues einschließlich Mais, Hirse und Buchweizen*. Paul Parey, Berlin.

Bekele, W.A., Wieckhorst, S., Friedt, W. and Snowdon, R.J. (2013) High-throughput genomics in sorghum: from whole-genome resequencing to a SNP screening array. *Plant Biotechnology Journal* 11(9), 1112–1125. Available at: https://doi.org/10.1111/pbi.12106

Bekele, W.A., Fiedler, K., Shiringani, A., Schnaubelt, D., Windpassinger, S. *et al*. (2014) Unravelling the genetic complexity of sorghum seedling development under low-temperature conditions. *Plant, Cell & Environment* 37(3), 707–723. Available at: https://doi.org/10.1111/pce.12189

Berenji, J., Dahlberg, J., Sikora, V. and Latkovi, D. (2011) Origin, history, morphology, production, improvement, and utilization of broomcorn [*Sorghum bicolor* (L.) Moench] in Serbia. *Economic Botany* 65(2), 190–208. Available at: https://doi.org/10.1007/s12231-0119155-2

Berhan, A.M., Hulbert, S.H., Butler, L.G. and Bennetzen, J.L. (1993) Structure and evolution of the genomes of *Sorghum bicolor* and *Zea mays*. *Theoretical and Applied Genetics* 86(5), 598–604. Available at: https://doi.org/10.1007/BF00838715

Bhandari, A., Bartholomé, J., Cao-Hamadoun, T.-V., Kumari, N., Frouin, J., Kumar, A. and Ahmadi, N. (2019) Selection of trait-specific markers and multi-environment models improve genomic predictive ability in rice. *PLoS ONE* 14(5), e0208871. Available at: https://doi.org/10.1371/journal.pone.0208871

Bharathi, R.R., Agasimani, S., Anusheela, V., Thiruvengadam, V., Chibbar, R.N. and Ganesh R.S. (2016) TILLING and EcoTILLING for discovery of induced and natural variations in sorghum genome. In: Rakshit, S. and Wang, Y.-H. (eds) *The Sorghum Genome*. Compendium of Plant Genomes. Springer, Cham, Switzerland, pp. 257–267.

Boivin, K., Deu, M., Rami, J.-F., Trouche, G. and Hamon, P. (1999) Towards a saturated sorghum map using RFLP and AFLP markers. *Theoretical and Applied Genetics* 98(2), 320–328. Available at: https://doi.org/10.1007/s001220051076

Boyles, R.E., Brenton, Z.W. and Kresovich, S. (2019) Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments. *The Plant Journal* 97(1), 19–39. Available at: https://doi.org/10.1111/tpj.14113

Brooking, I.R. (1976) Male sterility in *Sorghum bicolor* (L.) Moench induced by low night temperature, I. Timing of the stage of sensitivity. *Australian Journal of Plant Physiology* 3(5), 589–596. Available at: https://doi.org/10.1071/pp9760589

Burow, G., Burke, J.J., Xin, Z. and Franks, C.D. (2011) Genetic dissection of early-season cold tolerance in sorghum (*Sorghum bicolor* (L.) Moench). *Molecular Breeding* 28(3), 391–402. Available at: https://doi.org/10.1007/s11032-010-9491-4

Che, P., Anand, A., Wu, E., Sander, J.D., Simon, M.K. *et al*. (2018) Developing a flexible, high-efficiency *Agrobacterium*-mediated sorghum transformation system with broad application. *Plant Biotechnology Journal* 16(7), 1388–1395. Available at: https://doi.org/10.1111/pbi.12879

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 Biology* 17(1), 12. Available at: https://doi.org/10.1186/s12870-016-0966-2

Cisneros-López, M.E., Mendoza-Onofre, L.E., Zavaleta-Mancera, H.A., González-Hernández, V.A., Mora-Aguilera, G., Córdova-Téllez, L. and Hernández-Martínez, M. (2010) Pollen–pistil interaction, pistil histology and seed production in A × B grain sorghum crosses under chilling field temperatures. *The Journal of Agricultural Science* 148(1), 73–82.

Conner, A.B. and Karper, R.E. (1927) *Hybrid Vigor in Sorghum*. Bulletin No. 359. Texas Agricultural Experiment Station, Texas A&M University, College Station, Texas. Available at: http://hdl.handle.net/1969.1/4015 (accessed 5 March 2021).

Dahlberg, J., Berenji, J., Sikora, V. and Latković, D. (2012) Assessing sorghum [*Sorghum bicolor* (L) Moench] germplasm for new traits: food, fuels & unique uses. *Maydica* 56(2). Available at: https://journals-crea.4science.it/index.php/maydica/article/view/688

Deshpande, S., Rakshit, S., Manasa, K.G., Pandey, S. and Gupta, R. (2017) Genomic approaches for abiotic stress tolerance in sorghum. In: Rakshit, S. and Wang, Y.-H. (eds) *The Sorghum Genome*. Compendium of Plant Genomes. Springer, Cham, Switzerland, pp. 169–187.

Dillon, S.L., Shapter, F.M., Henry, R.J., Cordeiro, G., Izquierdo, L. and Lee, L.S. (2007) Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). *Annals of Botany* 100(5), 975–989. Available at: https://doi.org/10.1093/aob/mcm192

Downes, R.W. and Marshall, D.R. (1971) Low temperature induced male sterility in *Sorghum bicolor*. *Australian Journal of Experimental Agriculture* 11(50), 352–356.

Ejeta, G. and Knoll, J.E. (2007) Marker-assisted selection in sorghum. In: Varshney R.K. and Tuberosa R. (eds) *Genomics-assisted Crop Improvement*. Springer, Dordrecht, the Netherlands, pp. 187–205.

Farré, I. and Faci, J.M. (2006) Comparative response of maize (*Zea mays* L.) and sorghum (*Sorghum bicolor* L. Moench) to deficit irrigation in a Mediterranean environment. *Agricultural Water Management* 83(1–2), 135–143. Available at: https://doi.org/10.1016/j.agwat.2005.11.001

Fiedler, K., Bekele, W.A., Friedt, W., Snowdon, R., Stützel, H., Zacharias, A. and Uptmoor, R. (2012) Genetic dissection of the temperature dependent emergence processes in sorghum using a cumulative emergence model and stability parameters. *Theoretical and Applied Genetics* 125(8), 1647–1661. Available at: https://doi.org/10.1007/s00122-012-1941-4

Fiedler, K., Bekele, W.A., Duensing, R., Gründig, S., Snowdon, R. *et al*. (2014) Genetic dissection of temperature-dependent sorghum growth during juvenile development. *Theoretical and Applied Genetics* 127(9), 1935–1948. Available at: https://doi.org/10.1007/s00122-014-2350-7

Fiedler, K., Bekele, W.A., Matschegewski, C., Snowdon, R., Wieckhorst, S., Zacharias, A. and Uptmoor, R. (2016) Cold tolerance during juvenile development in sorghum: a comparative analysis by genomewide association and linkage mapping. *Plant Breeding* 135(5), 598–606. Available at: https://doi.org/10.1111/pbr.12394

Fracasso, A., Trindade, L.M. and Amaducci, S. (2016) Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE. *BMC Plant Biology* 16(1), 115. Available at: https://doi.org/10.1186/s12870-016-0800-x

Franks, C.D., Burow, G.B. and Burke, J.J. (2006) A comparison of US and Chinese sorghum germplasm for early season cold tolerance. *Crop Science* 46(3), 1371–1376.

Gabriel, K. (2005) A study of heterotic relationships in sorghum. PhD dissertation, Texas A&M University, College Station, Texas. Available at https://oaktrust.library.tamu.edu/bitstream/1969.1/3226/1/etd-tamu-2005C-PLBR-Gabriel.pdf (accessed 3 March 2021).

Guitton, B., Vaksman, M., Rami, J.-F., Weltzein, E., Rattunde, F. *et al*. (2015) Enhancing sorghum grain yield and quality for the Sudano-Sahelian zone of West Africa using the backcross nested association mapping (BCNAM) approach. *Presented at XXIIIrd EUCARPIA Maize and Sorghum Conference of the Vth Session – New Insights in the Genetics of Traits (2), Montpellier, France, 10–11 June 2015*.

Gurel, S., Gurel, E., Kaur, R., Wong, J., Meng, L., Tan, H.-Q. and Lemaux, P.G. (2009) Efficient, reproducible *Agrobacterium*-mediated transformation of sorghum using heat treatment of immature embryos. *Plant Cell Reports* 28(3), 429–444. Available at: https://doi.org/10.1007/s00299-008-0655-1

Hamblin, M.T., Mitchell, S.E., White, G.M., Gallego, J., Kukatla, R. *et al*. (2004) Comparative population genetics of the panicoid grasses: sequence polymorphism, linkage disequilibrium and selection in a diverse sample of sorghum bicolor. *Genetics* 167(1), 471–483.

Hariprasanna, K. and Patil, J.V. (2015) Sorghum: origin, classification, biology and improvement. In: Madhusudhana, R., Rajendrakumar, P. and Patil, J.V. (eds) *Sorghum Molecular Breeding*. Springer India, New Delhi, pp. 3–20.

Harlan, J.R. and de Wet, J.M.J. (1972) A simplified classification of cultivated sorghum. *Crop Science* 12(2), 172–176. Available at: https://doi.org/10.2135/cropsci1972.0011183X001200020005x

Hilley, J., Truong, S., Olson, S., Morishige, D. and Mullet, J. (2016) Identification of *Dw1*, a regulator of sorghum steminternodelength.*PLoSONE*11(3),e0151271.Availableat:https://doi.org/10.1371/journal.pone.0151271

Hilley, J.L., Weers, B.D., Truong, S.K., McCormick, R.F., Mattison, A.J. *et al*. (2017) Sorghum *Dw2* encodes a protein kinase regulator of stem internode length. *Scientific Reports* 7(1), 4616. Available at: https://doi.org/10.1038/s41598-017-04609-5

Hulbert, S.H., Richter, T.E., Axtell, J.D. and Bennetzen, J.L. (1990) Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. *Proceedings of the National Academy of Sciences USA* 87(11), 4251–4255. Available at: https://doi.org/10.1073/pnas.87.11.4251

Ji, G., Zhang, Q., Du, R., Lv, P., Ma, X. *et al*. (2017) Construction of a high-density genetic map using specific-locus amplified fragments in sorghum. *BMC Genomics* 18(1), 51. Available at: https://doi.org/10.1186/s12864-016-3430-7

Jiao, Y., Burke, J., Chopra, R., Burow, G., Chen, J. *et al*. (2016) A sorghum mutant resource as an efficient platform for gene discovery in grasses. *The Plant Cell* 28(7), 1551–1562. Available at: https://doi.org/10.1105/tpc.16.00373

Jordan, D.R., Mace, E.S., Cruickshank, A.W., Hunt, C.H. and Henzell, R.G. (2011) Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. Crop Science 51(4), 1444–1457. Available at: https://doi.org/10.2135/cropsci2010.06.0326

Jordan, D.R., Hunt, C.H., Cruickshank, A.W., Borrell, A.K. and Henzell, R.G. (2012) The relationship between the stay-green trait and grain yield in elite sorghum hybrids grown in a range of environments. *Crop Science* 52(3), 1153–1161. Available at: https://doi.org/10.2135/cropsci2011.06.0326

Kadier, Y., Zu, Y.-Y., Dai, Q.-M., Song, G., Lin, S.-W. *et al*. (2017) Genome-wide identification, classification and expression analysis of NAC family of genes in sorghum [*Sorghum bicolor* (L.) Moench]. *Plant Growth Regulation* 83(2), 301–312. Available at: https://doi.org/10.1007/s10725-017-0295-y

Kirby, J.S. and Atkins, R.E. (1968) Heterotic response for vegetative and mature plant characters in grain sorghum, *Sorghum bicolor* (L.) Moench. *Crop Science* 8(3), 335–339. Available at: https://doi.org/10.2135/cropsci1968.0011183x000800030022x

Klein, R.R., Mullet, J.E., Jordan, D.R., Miller, F.R., Rooney, W.L. *et al*. (2008) The effect of tropical sorghum conversion and inbred development on genome diversity as revealed by high-resolution genotyping. *Crop Science* 48 (S1), S12–S26.

Klein, R.R., Miller, F.R., Dugas, D.V., Brown, P.J., Burrell, A.M. and Klein, P.E. (2015) Allelic variants in the *PRR37* gene and the human-mediated dispersal and diversification of sorghum. *Theoretical and Applied Genetics* 128(9), 1669–1683.

Knoll, J., Gunaratna, N. and Ejeta, G. (2008) QTL analysis of early-season cold tolerance in sorghum. *Theoretical and Applied Genetics* 116(4), 577–587. Available at: https://doi.org/10.1007/s00122-007-0692-0

Kong, W., Jin, H., Franks, C.D., Kim, C., Bandopadhyay, R. *et al*. (2013) Genetic analysis of recombinant inbred lines for *Sorghum bicolor* × *Sorghum propinquum*. *G3: Genes, Genomes, Genetics* 3(1), 101–108. Available at: https://doi.org/10.1534/g3.112.004499

Krishnamurthy, L., Dinakaran, E., Kumar, A.A. and Reddy, B.V.S. (2014) Field technique and traits to assess reproductive stage cold tolerance in sorghum (*Sorghum bicolor* (L.) Moench. *Plant Production Science* 17(3), 218–227. Available at: https://doi.org/10.1626/pps.17.218

Kuriyama, T., Shimada, S. and Matsui, M. (2019) Improvement of agrobacterium-mediated transformation for tannin-producing sorghum. *Plant Biotechnology* 36(1), 43–48. Available at: https://doi.org/10.5511/plantbiotechnology.19.0131a

Leon-Velasco, H., Mendoza-Onofre, L.E., Castillo-Gonzalez, F., Cervantes-Santana, T. and Martínez-Garza, A. (2009) Evaluation of two generations of cold tolerant sorghum hybrids and parental lines. II: Combining ability, heterosis and heterobeltiosis. *Agrociencia* 43(6), 609–623.

Liu, H., Wang, X., Warburton, M.L., Wen, W., Jin, M. *et al*. (2015) Genomic, transcriptomic, and phenomic variation reveals the complex adaptation of modern maize breeding. *Molecular Plant* 8(6), 871–884. Available at: https://doi.org/10.1016/j.molp.2015.01.016

Liu, G., Gilding, E.K., Kerr, E.D., Schulz, B.L., Tabet, B., Hamaker, B.R. and Godwin, I.D. (2019) Increasing protein content and digestibility in sorghum grain with a synthetic biology approach. *Journal of Cereal Science* 85, 27–34. Available at: https://doi.org/10.1016/j.jcs.2018.11.001

Mace, E.S., Rami, J.-F., Bouchet, S., Klein, P.E., Klein, R.R. *et al*. (2009) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. *BMC Plant Biology* 9(1), 13. Available at: https://doi.org/10.1186/1471-2229-9-13

Mace, E.S., Tai, S., Gilding, E.K., Li, Y., Prentis, P.J. *et al*. (2013) Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature Communications* 4(1), 2320. Available at: https://doi.org/10.1038/ncomms3320

Mace, E., Innes, D., Hunt, C., Wang, X., Tao, Y. *et al*. (2019) The Sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. *Theoretical and Applied Genetics* 132(3), 751–766. Available at: https://doi.org/10.1007/s00122-018-3212-5

Makita, Y., Shimada, S., Kawashima, M., Kondou-Kuriyama, T., Toyoda, T. and Matsui, M. (2015) MOROKOSHI: transcriptome database in *Sorghum bicolor*. *Plant & Cell Physiology* 56(1), e6. Available at: https://doi.org/10.1093/pcp/pcu187

Mann, J.A., Kimber, C.T. and Miller, F.R. (1983) *The Origin and Early Cultivation of Sorghums in Africa*. *Bulletin No. 1454. Texas Agricultural Experiment Station, Texas A&M University*, College Station, Texas. Available at: http://hdl.handle.net/1969.1/128074 (accessed 3 March 2021).

Marla, S.R., Burow, G., Chopra, R., Hayes, C., Olatoye, M.O. *et al*. (2019) Genetic architecture of chilling tolerance in sorghum dissected with a nested association mapping population. *bioRxiv* 130, 622894. Available at: https://doi.org/10.1101/622894

Matsuoka, Y., Vigouroux, Y., Goodman, M.M., Sanchez, G.J., Buckler, E. and Doebley, J. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences USA* 99(9), 6080–6084. Available at: https://doi.org/10.1073/pnas.052125199

Maulana, F. and Tesso, T.T. (2013) Cold temperature episode at seedling and flowering stages reduces growth and yield components in sorghum. *Crop Science* 53(2), 564–574.

Mendoza, O.L.E. (1988) Formación de híbridos de sorgo para grano: II. Comportamiento per se de las líneas y su aptitud combinatoria general. *Revista Fitotecnia Mexicana* 11, 39–47.

Menz, M.A., Klein, R.R., Unruh, N.C., Rooney, W.L., Klein, P.E. and Mullet, J.E. (2004) Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. *Crop Science* 44(4), 1236–1244. Available at: https://doi.org/10.2135/cropsci2004.1236

Monk, R., Franks, C., Dahlberg, J., Smith, S., Diers, B., Specht, J. and Carver, B. (2014) Sorghum. In: Smith, S., Diers, B., Specht, J. and Carver, B. (eds) *Yield Gains in Major U.S. Field Crops*. CSSA Special Publication, Vol. 33. American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, Madison, Wisconsin, pp. 293–310. Available at: https://doi.org/10.2135/cssaspecpub33.c11

Morris, G.P., Ramu, P., Deshpande, S.P., Hash, C.T., Shah, T. *et al*. (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences USA* 110(2), 453–458. Available at: https://doi.org/10.1073/pnas.1215985110

Multani, D.S., Briggs, S.P., Chamberlin, M.A., Blakeslee, J.J., Murphy, A.S. and Johal, G.S. (2003) Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. *Science* 302(5642), 81–84. Available at: https://doi.org/10.1126/science.1086072

Muraya, M.M. (2014) Sorghum genetic diversity. In: Wang, Y.-H., Upadhyaya, H.D. and Kole, C. (eds) *Genetics, Genomics and Breeding of Sorghum*. CRC Press, Boca Raton, Florida, pp. 136–162.

Nelson, J.C., Wang, S., Wu, Y., Li, X., Antony, G., White, F.F. and Yu, J. (2011) Single-nucleotide polymorphism discovery by high-throughput sequencing in sorghum. *BMC Genomics* 12, 352. Available at: https://doi.org/10.1186/1471-2164-12-352

Osuna-Ortega, J., Mendoza-Onofre, L.E., González-Hernández, V.A., Castillo-González, F., del Mendoza-Castillo, M.C. and Williams-Alanís, H. (2000) Potential of cold tolerant germplasm in the adaptation and adaptability of sorghum in México: I. *High valleys. Agrociencia* 34(5), 561–572.

Osuna-Ortega, J., Mendoza-Castillo, M.D.C. and Mendoza-Onofre, L. (2003) Sorghum cold tolerance, pollen production. *Maydica* 48, 125–132.

Oyediran, I.O., Hibbard, B.E. and Clark, T.L. (2004) Prairie grasses as hosts of the western corn rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology* 33(3), 740–747. Available at: https://doi.org/10.1603/0046-225X-33.3.740

Parra-Londono, S., Fiedler, K., Kavka, M., Samans, B., Wieckhorst, S., Zacharias, A. and Uptmoor, R. (2018) Genetic dissection of early-season cold tolerance in sorghum: genome-wide association studies for seedling emergence and survival under field and controlled environment conditions. *Theoretical and Applied Genetics* 131(3), 581–595. Available at: https://doi.org/10.1007/s00122-017-3021-2

Patane, C., Cavallaro, V., Avola, G. and D'Agosta, G. (2006) Seed respiration of sorghum [*Sorghum bicolor* (L.) Moench] during germination as affected by temperature and osmoconditioning. *Seed Science Research* 16(4), 251–260.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J. *et al*. (2009) The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457(7229), 551–556. Available at: https://doi.org/10.1038/nature07723

Peacock, J.M. (1982) Response and tolerance of sorghum to temperature stress. In: Mertin, J.V. (ed.) *Sorghum in the Eighties: Proceedings of the International Symposium on Sorghum, ICRISAT Center, Patancheru, India, 2–7 November 1981*. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India, pp. 143–159.

Peng, Y., Schertz, K.F., Cartinhour, S. and Hart, G.E. (1999) Comparative genome mapping of *Sorghum bicolor* (L.) Moench using an RFLP map constructed in a population of recombinant inbred lines. *Plant Breeding* 118(3), 225–235. Available at: https://doi.org/10.1046/j.1439-0523.1999.118003225.x

Pérez-de-Castro, A.M., Vilanova, S., Cañizares, J., Pascual, L., Blanca, J.M. *et al*. (2012) Application of genomic tools in plant breeding. *Current Genomics* 13(3), 179–195. Available at: https://doi.org/10.2174/138920212800543084

Pinthus, M.J. and Rosenblum, J (1961) Germination and seedling emergence of sorghum at low temperatures. *Crop Science* 1, 293–296.

Promkhambut, A., Polthanee, A., Akkasaeng, C. and Younger, A. (2011) Growth, yield and aerenchyma formation of sweet and multipurpose sorghum (*Sorghum bicolor* L. Moench) as affected by flooding at different growth stages. *Australian Journal of Crop Science* 5(8), 954–965.

Quinby, J.R. (1966) Fourth maturity gene locus in sorghum. *Crop Science* 6(6), 516–518. Available at: https://doi.org/10.2135/cropsci1966.0011183X000600060005x

Quinby, J.R. (1974) *Sorghum Improvement and the Genetics of Growth*. Texas A&M University Press, College Station, Texas.

Quinby, J.R. and Karper, R.E. (1945) The inheritance of three genes that influence time of floral initiation and maturity date in Milo. *Agronomy Journal* 37(11), 916–936. Available at: https://doi.org/10.2134/agronj1945.00021962003700110006x

Rakshit, S. and Bellundagi, A. (2019) Conventional breeding techniques in sorghum. In: Aruna, C., Visarada, K.B.R.S., Venkatesh Bhat, B. and Tonapi, V.A. (eds) *Breeding Sorghum for Diverse End Uses*. Woodhead Publishing, Cambridge, pp. 77–91.

Reif, J.C., Hallauer, A.R. and Melchinger, A.E. (2005) Heterosis and heterotic patterns in maize. *Maydica* 50(3/4), 215.

Salas Fernandez, M.G., Schoenbaum, G.R. and Goggi, A.S. (2014) Novel germplasm and screening methods for early cold tolerance in sorghum. *Crop Science* 54(6), 2631–2638.

Schaffasz, A., Windpassinger, S., Friedt, W., Snowdon, R. and Wittkop, B. (2019a) Sorghum as a novel crop for central Europe: using a broad diversity set to dissect temperate-adaptation. *Agronomy* 9(9), 535. Available at: https://doi.org/10.3390/agronomy9090535

Schaffasz, A., Windpassinger, S., Snowdon, R. and Wittkop, B. (2019b) Reproductive cold stress tolerance in sorghum F_1 hybrids is a heterotic trait. *Agronomy* 9(9), 508. Available at: https://doi.org/10.3390/agronomy9090508

Schnable, P.S. and Springer, N.M. (2013) Progress toward understanding heterosis in crop plants. *Annual Review of Plant Biology* 64, 71–88.

Shao, H., Wang, H. and Tang, X. (2015) NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Frontiers in Plant Science* 6, 902. Available at: https://doi.org/10.3389/fpls.2015.00902

Singh, H.P. and Lohithaswa, H.C. (2006) Sorghum. In: Kole, C. (ed.) *Cereals and Millets*. Springer, Berlin/Heidelberg, Germany, pp. 257–302.

Singh, S.P. (1985) Sources of cold tolerance in grain sorghum. *Canadian Journal of Plant Science* 65(2), 251–257. Available at: https://doi.org/10.4141/cjps85-037

Sleper, D.A. and Poehlman, J.M. (2006) *Breeding Field Crops*. Blackwell Publishing, Oxford.

Smith, C.W. and Frederiksen, R.A. (2000) *Sorghum: Origin, History, Technology, and Production*. Wiley, New York.

Smýkal, P., Varshney, R.K., Singh, V.K., Coyne, C.J., Domoney, C., Kejnovský, E. and Warkentin, T., (2016) From Mendel's discovery on pea to today's plant genetics and breeding: commemorating the 150th anniversary of the reading of Mendel's discovery. *Theoretical and Applied Genetics* 129(12), 2267–2280. Available at: https://doi.org/10.1007/s00122-016-2803-2

Staggenborg, S.A., Dhuyvetter, K.C. and Gordon, W.B. (2008) Grain sorghum and corn comparisons: yield, economic, and environmental responses. *Agronomy Journal* 100(6), 1600–1604. Available at: https://doi.org/10.2134/agronj2008.0129

Stephens, J.C. and Holland, R.F. (1954) Cytoplasmic male sterility for hybrid sorghum seed production. *Agronomy Journal* 46(1), 20–23.

Stephens, J.C., Miller, F.R. and Rosenow, D.T. (1967) Conversion of alien sorghums to early combine genotypes. *Crop Science* 7(4), 396. Available at: https://doi.org/10.2135/cropsci1967.0011183x000700040036x

Tao, Y.Z., Jordan, D.R., Henzell, R.G. and McIntyre, C.L. (1998) Construction of a genetic map in a sorghum recombinant inbred line using probes from different sources and its comparison with other sorghum maps. *Australian Journal of Agricultural Research* 49(4), 729–736. Available at: https://doi.org/10.1071/a97112

Thurber, C.S., Ma, J.M., Higgins, R.H. and Brown, P.J. (2013) Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biology* 14(6), R68. Available at: https://doi.org/10.1186/gb-2013-14-6-r68

Tian, T., You, Q., Zhang, L., Yi, X., Yan, H., Xu, W. and Su, Z. (2016) SorghumFDB: sorghum functional genomics database with multidimensional network analysis. *Database* 2016, baw099. Available at: https://doi.org/10.1093/database/baw099

Trachsel, S., Dhliwayo, T., Gonzalez Perez, L., Mendoza Lugo, J.A. and Trachsel, M. (2019) Estimation of physiological genomic estimated breeding values (PGEBV) combining full hyperspectral and marker data across environments for grain yield under combined heat and drought stress in tropical maize (*Zea mays* L.). *PLoS ONE* 14(3), e0212200. Available at: https://doi.org/10.1371/journal.pone.0212200

Varshney, R.K., Terauchi, R. and McCouch, S.R. (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. *PLoS Biology* 12(6), e1001883. Available at: https://doi.org/10.1371/journal.pbio.1001883

Velazco, J.G., Jordan, D.R., Mace, E.S., Hunt, C.H., Malosetti, M. and van Eeuwijk, F.A. (2019) Genomic prediction of grain yield and drought-adaptation capacity in sorghum is enhanced by multi-trait analysis. *Frontiers in Plant Science* 10, 997. Available at: https://doi.org/10.3389/fpls.2019.00997

Venkateswaran, K., Muraya, M., Dwivedi, S.L. and Upadhyaya, H.D. (2014) Wild sorghums – their potential use in crop improvement. In: Wang, Y.-H., Upadhyaya, H.D. and Kole, C. (eds) *Genetics, Genomics and Breeding of Sorghum*. CRC Press, Boca Raton, Florida, pp. 78–111.

Voss-Fels, K.P., Stahl, A., Wittkop, B., Lichthardt, C., Nagler, S. *et al*. (2019) Breeding improves wheat productivity under contrasting agrochemical input levels. *Nature Plants* 5(7), 706–714. Available at: https://doi.org/10.1038/s41477-019-0445-5

Wesseler, J. and Fall, E.H. (2010) Potential damage costs of *Diabrotica virgifera virgifera* infestation in Europe – the 'no control' scenario. *Journal of Applied Entomology* 134(5), 385–394. Available at: https://doi.org/10.1111/j.1439-0418.2010.01510.x

Windpassinger, S. (2016) Breeding strategies for the adaptation of sorghum (*Sorghum bicolor* L. Moench) as a novel crop for temperate Europe. PhD thesis, Justus Liebig University, Giessen, Germany.

Windpassinger, S., Friedt, W., Frauen, M., Snowdon, R. and Wittkop, B. (2015) Designing adapted sorghum silage types with an enhanced energy density for biogas generation in temperate Europe. *Biomass and Bioenergy* 81(3), 496–504. Available at: https://doi.org/10.1016/j.biombioe.2015.08.005

Woldesemayat, A.A., Modise, D.M., Gemeildien, J., Ndimba, B.K. and Christoffels, A. (2018) Cross-species multiple environmental stress responses: an integrated approach to identify candidate genes for multiple stress tolerance in sorghum (*Sorghum bicolor* (L.) Moench) and related model species. *PLoS ONE* 13(3), e0192678. Available at: https://doi.org/10.1371/journal.pone.0192678

Xin, Z., Wang, M.L., Barkley, N.A., Burow, G., Franks, C., Pederson, G. and Burke, J. (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biology* 8(1), 103. Available at: https://doi.org/10.1186/1471-2229-8-103

Xu, J.C., Weerasuriya, Y.M. and Bennetzen, J.L. (2001) Construction of genetic map in sorghum and fine mapping of the germination stimulant production gene response to Striga asiatica. *Yi Chuan Xue Bao* 28(9), 870–876.

Yellisetty, V., Reddy, L.A. and Mandapaka, M. (2015) *In planta* transformation of sorghum (*Sorghum bicolor* (L.) Moench) using *TPS1* gene for enhancing tolerance to abiotic stresses. *Journal of Genetics* 94(3), 425–434. Available at: https://doi.org/10.1007/s12041-015-0540-y

Yonemaru, J.-I., Ando, T., Mizubayashi, T., Kasuga, S., Matsumoto, T. and Yano, M. (2009) Development of genome-wide simple sequence repeat markers using whole-genome shotgun sequences of sorghum (*Sorghum bicolor* (L.) Moench). *DNA Research* 16(3), 187–193. Available at: https://doi.org/10.1093/dnares/dsp005

Yu, J. and Tuinstra, M.R. (2001) Genetic analysis of seedling growth under cold temperature stress in grain sorghum. *Crop Science* 41(5), 1438–1443. Available at: https://doi.org/10.2135/cropsci2001.4151438x

Yuan, Y., Cairns, J.E., Babu, R., Gowda, M., Makumbi, D. *et al*. (2018) Genome-wide association mapping and genomic prediction analyses reveal the genetic architecture of grain yield and flowering time under drought and heat stress conditions in maize. *Frontiers in Plant Science* 9, 1919. Available at: https://doi.org/10.3389/fpls.2018.01919

Zhu, H., Jeoung, J.M., Liang, G.H., Muthukrishnan, S., Krishnaveni, S. and Wilde, G. (1998) Biolistic transformation of sorghum using a rice chitinase gene [*Sorghum bicolor* (L.) Moench–*Oryza sativa* L.]. *Journal of Genetics & Breeding* 52, 243–252.

3 Genetic and genomic diversity in the sorghum genebank collection of Uganda

Subhadra Chakrabarty, Raphael Mufumbo, Steffen Windpassinger, David Jordan, Emma Mace, Rod J. Snowdon, Adrian Hathorn
BMC Plant Biology, Vol 22, pp 378 (2022).
doi: https://doi.org/10.1186/s12870-022-03770-y

BMC Plant Biology

RESEARCH

Open Access

Genetic and genomic diversity in the sorghum gene bank collection of Uganda

Subhadra Chakrabarty[1], Raphael Mufumbo[1,2], Steffen Windpassinger[1], David Jordan[3], Emma Mace[3], Rod J. Snowdon[1*] and Adrian Hathorn[3]

Abstract

Background: The Plant Genetic Resources Centre at the Uganda National Gene Bank houses has over 3000 genetically diverse landraces and wild relatives of *Sorghum bicolor* accessions. This genetic diversity resource is untapped, under-utilized, and has not been systematically incorporated into sorghum breeding programs. In this study, we characterized the germplasm collection using whole-genome SNP markers (DArTseq). Discriminant analysis of principal components (DAPC) was implemented to study the racial ancestry of the accessions in comparison to a global sorghum diversity set and characterize the sub-groups present in the Ugandan (UG) germplasm.

Results: Population structure and phylogenetic analysis revealed the presence of five subgroups among the Ugandan accessions. The samples from the highlands of the southwestern region were genetically distinct as compared to the rest of the population. This subset was predominated by the caudatum race and unique in comparison to the other sub-populations. In this study, we detected QTL for juvenile cold tolerance by genome-wide association studies (GWAS) resulting in the identification of 4 markers associated ($-\log10p > 3$) to survival under cold stress under both field and climate chamber conditions, located on 3 chromosomes (02, 06, 09). To our best knowledge, the QTL on Sb09 with the strongest association was discovered for the first time.

Conclusion: This study demonstrates how genebank genomics can potentially facilitate effective and efficient usage of valuable, untapped germplasm collections for agronomic trait evaluation and subsequent allele mining. In face of adverse climate change, identification of genomic regions potentially involved in the adaptation of Ugandan sorghum accessions to cooler climatic conditions would be of interest for the expansion of sorghum production into temperate latitudes.

Keywords: *Sorghum bicolor*, Genetic diversity, Population structure, Cold tolerance, Temperate climate adaptation, Genome-wide association study, Genebank

Background

Sorghum bicolor [L.] Moench (sorghum) is the fifth most important cereal crop globally and shows remarkable diversity, including five different races, their intermediates, and several crop forms classified as grain, forage, sweet and broomcorn types [1]. Sorghum has extraordinary untapped variation in grain type, plant type, adaptability, productive capacity, and underutilized genetic potential [2]. Because of its wide adaptability to drought and heat, sorghum's importance is expected to increase with the changing global climate and an ongoing increase in the use of marginal lands for agriculture [3].

In Eastern Africa sorghum is traditionally grown as a food-fodder crop by smallholder farmers in low-input agricultural systems spanning highlands, lowlands, and semi-arid cropping regions. Uganda is located in eastern Africa, south of South Sudan. Archaeological evidence

*Correspondence: Rod.Snowdon@agrar.uni-giessen.de

[1] Department of Plant Breeding, Justus Liebig University, Giessen, Germany
Full list of author information is available at the end of the article

suggests that sorghum's center of domestication is likely the Ethiopia-Sudan region in the north-east [4]. For ages, sorghum breeders have classified cultivated sorghum into various races (mainly bicolor, guinea, caudatum, kafir, and durra) based on morphological characteristics [5, 6]. Uganda is one of only three countries where all of the five basic races and ten intermediate races of *S. bicolor* are native [7], making it extremely rich in genetic diversity.

The country's broad sorghum diversity reflects the variety of environments where the crop is grown, mainly on marginal agricultural lands ranging from extremely arid and semi-arid zones in eastern and northern Uganda to cool highlands in south-western regions. The Uganda National Genebank houses a large collection of *S. bicolor* accessions, including a vast range of landraces whose diversity has yet to be capitalized for use in breeding. This germplasm has not yet been fully characterised and evaluated, limiting its utilisation to date in sorghum improvement programs in Uganda and elsewhere. The considerable geographical and topological diversity of Uganda makes this genetic resource a potentially interesting reservoir for genetic analysis and diversity for adaptive traits of interest for sorghum breeding. For example, cool highland areas in the southwestern Uganda are potential sources of diversity for cold-tolerance traits that could help improve sorghum adaptation in temperate cropping regions.

Effective and efficient management of germplasm from a genebank collection is an essential prerequisite for farmers and breeders to identify, extract and exploit the extensive diversity. Genome-wide characterization of untapped genetic resources using genome sequencing technologies provides new opportunities for sustainable breeding and efficient usage of material.

In this present study, the diverse Uganda National Genebank *S. bicolor* collection, representing different agro-ecological zones of the country, was investigated using genome-wide SNP markers and population genetic analysis. The primary objective was to genetically characterize Ugandan (UG) sorghum germplasm in the context of global sorghum diversity, in the absence of morphological data from flowering and maturity for classification into racial groups. As a case study for the value of this resource in adaptive breeding, we furthermore used the available genome data, in association with phenotypic data for juvenile cold tolerance traits, to identify genomic regions enriched with genetic variants associated with low temperature adaptation. The results demonstrate how genebank genomics can help facilitate discovery of economically or biologically important plant diversity and genes as a prerequisite for crop genetic improvement and climatic adaptation.

Results

Genetic diversity and population structure analysis

In order to efficiently conserve and utilize the novel UG germplasm, we studied the underlying genetic variation and diversity. The population structure for the UG set was studied using all the 3333 samples and a subset of 12,742 markers from the complete dataset.

Five sub-populations within the germplasm were inferred by the DAPC analysis based on Bayesian Inference Criterion (BIC) (Fig. 1a). Four hundred PCs (27.4% of variance conserved by first three PCs) and four discriminant eigenvalues were retained in the DAPC analysis in order to capture maximum underlying variance. According to the population structure analysis (Fig. 1a, b, Table S1) the accessions from the central, eastern, northern and north western geographical regions did not cluster with any of the described racial groups, whereas the southwestern region showed a distinct genetic pattern with a predominance of caudatum race. The former genetic pattern can be a result of admixed races or it could be a possibility that the samples belonged to the bicolor race, which was not included in the reference global panel because of its dispersed nature. From the results presented in this study we can note that the geographical grouping does not explain the pattern of diversity observed in the germplasm and genetic clustering with the global panel would most likely explain the racial classification better. The distinct pattern of the samples from southwestern region as identified from the DAPC analysis was also confirmed using the phylogenetic tree analysis (Fig. 1c).

With the aim of genetically characterizing the novel Ugandan germplasm and racially comparing it to a well-established global diversity set of geo-referenced sorghum accessions a DAPC co-analysis was performed between the two. As Ugandan accessions outnumbered the global (UQ) set by almost three to one ratio, a random set of UG samples (10 samples from 10 random clusters) were selected for the comparison. All the underlying

(See figure on next page.)

Fig. 1 Discriminant analysis of principal components (DAPC) and phylogenetic analysis of 3333 Ugandan accessions. **a** Scatterplot showing the clustering pattern within the population. The axes represent the first two Linear Discriminants (LD) that were retained. Each circle represents an identified cluster whereas each dot represents an accession. **b** Barplot showing assignment of individuals to the five clusters recovered by the DAPC and their racial composition. Each vertical bar represents one individual accession (**c**) Neighbour-joining clustering of the Ugandan germplasm. The numbers are based on the geographical origins and the colours are based on the racial background

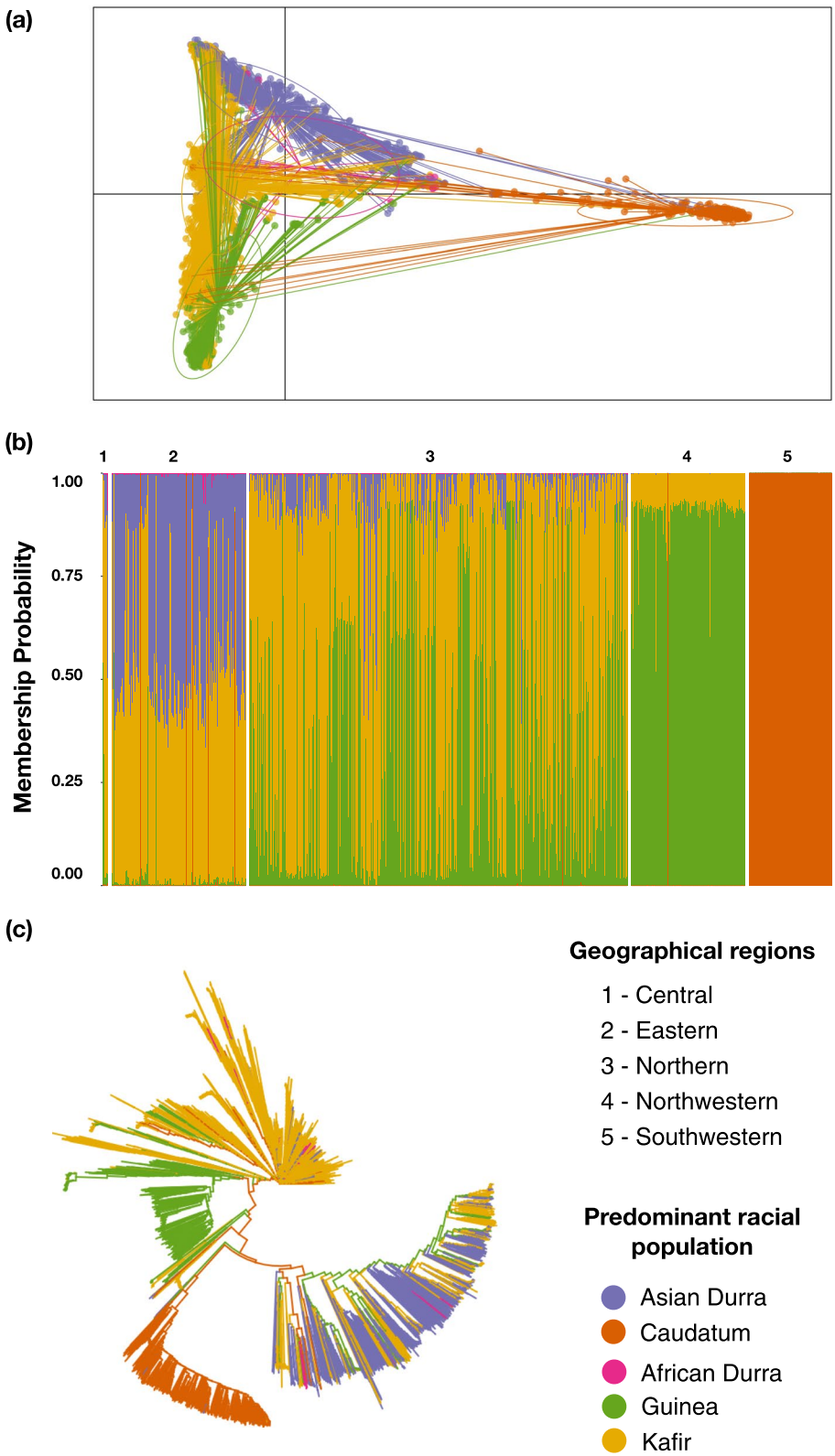

Fig. 1 (See legend on previous page.)

racial groups from the global collection were identified in the Ugandan germplasm (Fig. 2). Relative to the global set the UG samples were well distributed indicating that majority of the racial diversity was covered by the latter germplasm (Fig. 2a). Based on a threshold of 70%, majority of the samples from this UG subset selected for co-analysis could be assigned to a particular racial group (Fig. 2b) that were used in this study. However, around 5% could not be classified into any single racial category from the global set.

To confirm the racial distribution and diversity in the UG germplasm the DAPC co-analysis with the global set was replicated three times with three random unrelated UG samples. The racial clustering present among the UG set was found to be well dispersed (Fig. 2c, Fig. S1).

Due to significant adaptability differences between the lowland and highland races from the Ugandan diversity set, genetic diversity among the sub groups within the population were tested. Fst between each UG subpopulation from the various geographical regions compared to the entire population revealed genetic variance ranging from −0.608 to 0.525 (Table S2), indicating presence of significant genetic structuring or genotype variability within each subpopulation within the UG sorghum germplasm.

Phenotypic variation for juvenile cold stress survival, association mapping and haplotype analysis

A genome-wide association study (GWAS) was conducted for identification of regions of interest associated to juvenile survival under cold stress in two temperate-climate field environments (GG19, GG20) and one controlled-environment climate chamber experiment (CC). Highly significant differences for cold tolerance among genotypes ($p = 0.000^{***}$, Table S3) were found in all experiments. Comparing lowland- and highland genotypes as groups by a one-way ANOVA, highland genotypes showed a superior cold tolerance in all experiments as expected ($p = 0.015^{*}$ for GG19, $p = 0.000^{***}$ for GG20, $p = 0.000^{***}$ for CC).

Significant marker-trait associations to juvenile survival under cold stress, consistent across all test environments, were identified on chromosomes Sb02, Sb06, and Sb09 (Fig. 3 a, b, Table S4). We compared these selected genomic regions with previously curated QTL in sorghum QTL-Atlas [8], based on physical position (v3.0) and filtered using category resistance abiotic and subcategory cold tolerance. A sum of nineteen overlapping QTL involved in juvenile cold tolerance were identified (Table S5) in these regions. This result affirms the important role of these genomic regions towards cold adaptation. Furthermore, a novel QTL was discovered on Sb09 which showed the strongest association to juvenile cold tolerance in this UG germplasm.

Comparing protein sequence of orthologous gene segments of sorghum, maize and rice, six genes involved in various cold acclimatization and tolerance were identified (Table S6). For example, the candidate gene associated with the significant peak on chromosome Sb09 is *Sobic.009G260500*, annotated as a tetratricopeptide repeat (TPR)-like superfamily protein coding gene and as a cleavage stimulation factor subunit 3 (Cstf3). This gene family has been shown to have anti-cold response functions in tomato [9] and to be critically involved in heat stress responses in *Arabidopsis thaliana* [10] .

Discussion

Plant genetic resources like the sorghum collection of the National Genebank of Uganda are extremely important public germplasm resources for local breeders in crop centres of origin and the agricultural and crop research community worldwide. The substantial variation we identified in the Ugandan *S. bicolor* germplasm reflects the highly diverse environments where sorghum grows in Uganda. Accessions belonging to the arid regions of northern and eastern Uganda are well adapted to extreme drought and heat stress, whereas the ones from cold highlands of the southwestern region (Fig. S2) tolerant towards cold temperatures. According to [11] many sorghum landraces exist in this region which has previously not been characterized. To date, there is no available literature demonstrating genetic characterization of sorghum in the southwestern highlands and its role in cold tolerance.

As a result of adaptation to higher altitudes, the accessions from southwestern area are of major interest for breeding programs aiming to expand sorghum production into temperate climates in North America, Asia and Europe, whereas heat and drought stress tolerance are becoming increasingly important for global sorghum production in the face of climate change. The rich genetic resource of the Ugandan genebank sorghum collection has not yet been fully characterised and evaluated,

(See figure on next page.)

Fig. 2 Racial composition and classification of UG samples in comparison to global diversity set. **a** DA loadings (LD) displaying UG clustering (black dots) in comparison to global racial clusters. Each dot represents an individual and the colour code is displayed in the index. **b** composition of the races amongst the UG subset selected for co-analysis with global set. The x axis represents the different races and the y axis indicates the number of individuals that were assigned to the particular race. **c** DA loading of Ugandan representative accessions based on races

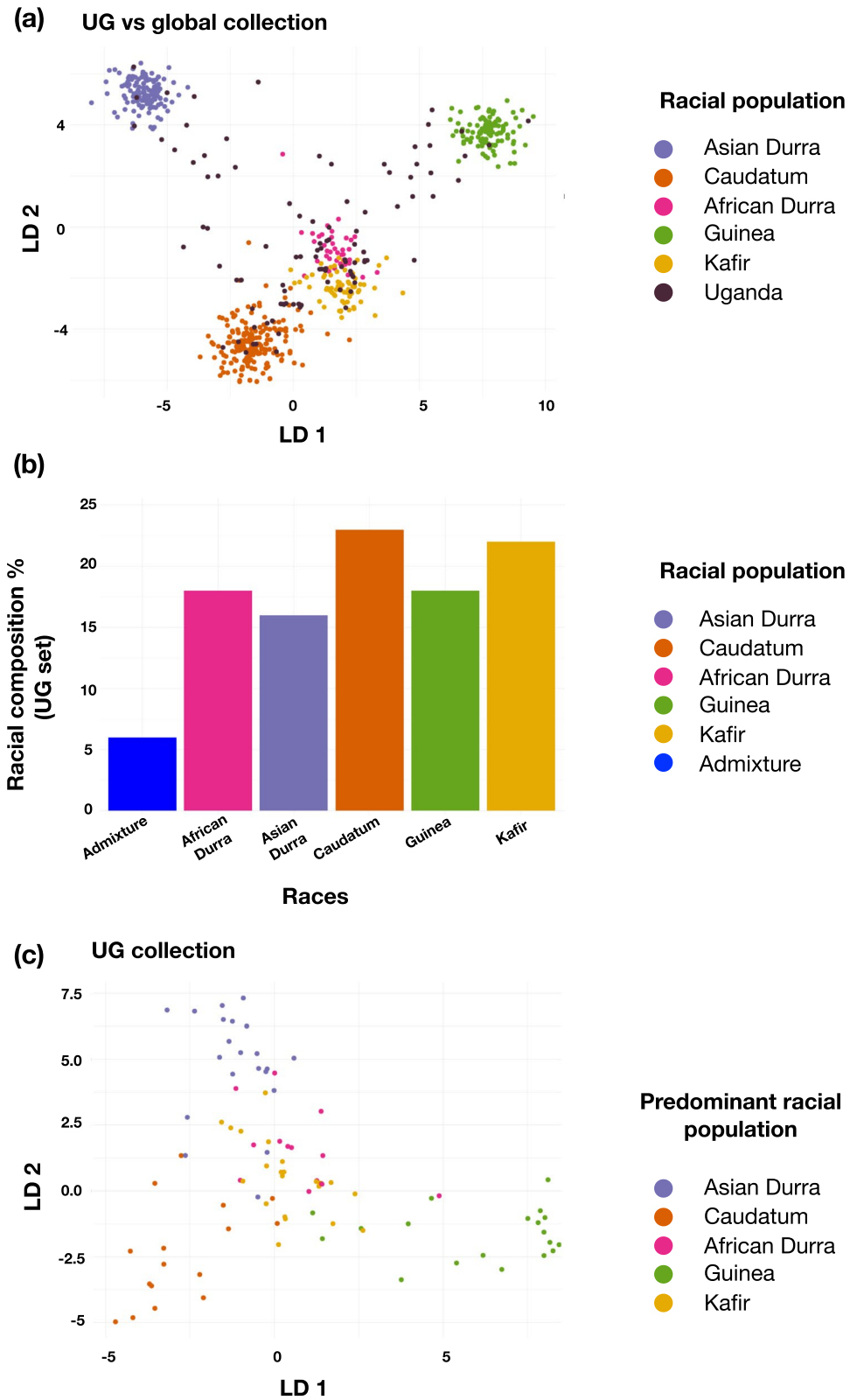

Fig. 2 (See legend on previous page.)

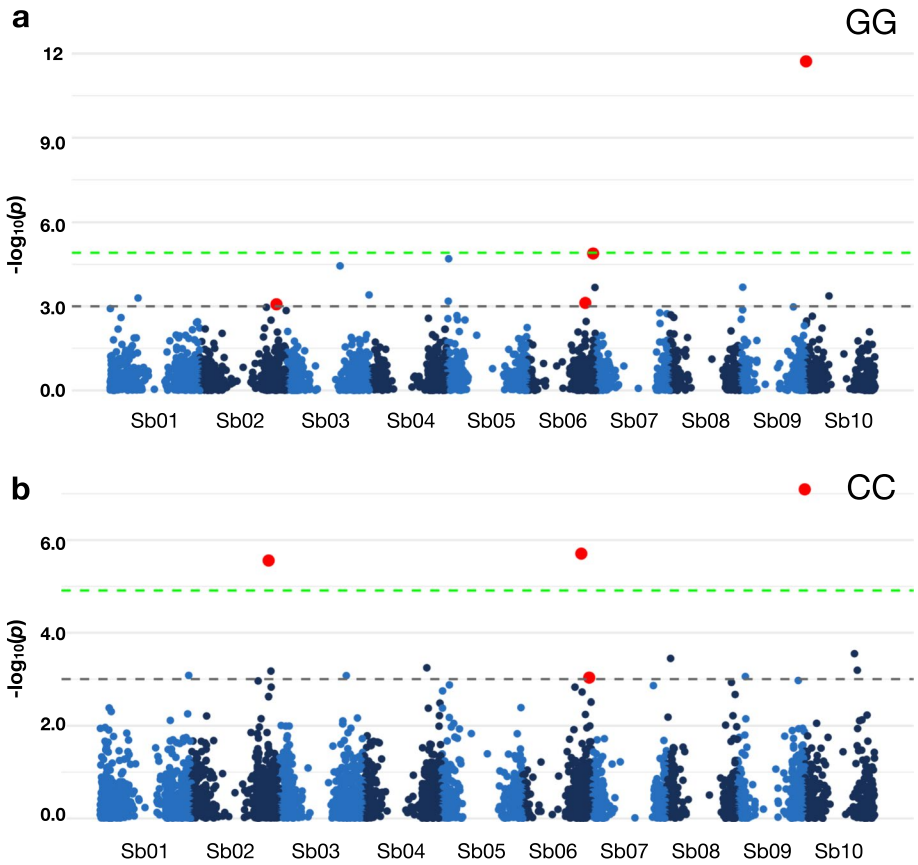

Fig. 3 Association mapping of survival under cold stress. The experiments were conducted under (**a**) field condition (Gross Gerau; GG) and (**b**) climate chamber (CC). The grey dotted horizontal line indicates a threshold of genome-wide cut-off at -log10(*p*) > 3.0 while the green line indicates the Bonferroni threshold at -log10(*p*) > 4.9. The selected associated markers which overlapped under both conditions are marked in red

limiting its utilisation to date in sorghum improvement programs in Uganda and elsewhere.

The genetic structure of sorghum has been previously documented in multiple studies using a variety of germplasm collections of different sizes [12]. To our knowledge this is the first report on sorghum genetic diversity using a substantial novel population of 3333 samples from Uganda, a key sorghum centre of origin in comparison to a well characterized global diversity panel [13].

This diversity panel consisted of accessions from the races caudatum, guinea, kafir and two sub populations of race durra, Asian and African. Accessions belonging to race bicolor were not included in this study as they did not form any coherent cluster. The lack of clear differentiation and clustering of *S. bicolor ssp. bicolor* races is not novel and has been previously reported in multiple studies [14–16]. According to literature, durra sorghums reached India before 3000 B.P. From there, they were subsequently introduced around 615 A.D. into Arabic Muslim states [17, 18] and over a period of time developed as distinct groups. Accessions belonging to two subpopulations of durra were included in this study because of their divergent characteristics.

The DAPC method is a good alternative to other population structure analysis software such as STRUCTURE because of its ability to deal with large datasets. Clustering of genotypes presented in this study provide interesting leads for increasing diversity in breeding programs and germplasm utilization. Usually, sorghum is racially classified based on phenotypic and/or morphological data. However, in the absence of appropriate phenotypic information for the collection under the long-day conditions in our field trials, we classified the germplasm solely based on genetic data. Overall, results from DAPC, Kmean and phylogenetic tree analyses were in agreement and provided evidence that the global racial diversity was well-covered in the UG germplasm. However, considerable admixture between racial groups were identified, particularly between genotypes from lowland areas in central, eastern and northern Uganda. This could be caused by movements by local farmers, resulting in infrequent gene flow between these regions,

or due to ancestral polymorphisms. According to [19] occurrence of numerous complex genomic interactions involving introgression between the different sorghum races shaped the current genomic diversity and structure within the species. Uganda, being located at the centre of origin and genetic diversity, can be presumed to accommodate accessions possessing patterns of genetic diversity that have been ancestrally inherited. The findings of admixture patterns reported in this study were therefore most likely caused by ancient recombination events and not recent crossovers between the closely related races. However, this phenomenon of non-clustering could also occur if these samples belong to bicolor race, which was not included in the current study. Resolution of this question requires further verification using more phenotypic information from plants grown to maturity under tropical short-day conditions.

Southwestern Uganda comprises predominantly the highland region, which in contrast to the other lowland agricultural areas has a lower average mean temperature [20]. Unlike sorghum accessions from the other geographic regions in Uganda, the accessions from the southwestern region appeared genetically distinct and were comprised predominantly of caudatum race. Because this region includes cool-temperature highland areas, it may contain interesting adaptive diversity for juvenile or reproductive cold tolerance. Population structure and phylogenetic analysis revealed the distinctness of this group compared to the other eco-geographical regions, presumably due to the dissimilarity of the highland cultivation environment and a relatively low exchange of germplasm between high and lowland farmers. Similarly, [21] reported the existence of caudatum and its intermediate races in the highlands of Ethiopia. This confers with the theory that the spread and diversification of crops to different locations can lead to new variants, a process influenced by genotype-by-environment interactions and geographical isolation [22]. According to [23], sorghum accessions from southwestern Uganda tend to have semi-compact elliptic panicles, a well-known characteristic of caudatum and its intermediate races.

Sorghum has a high potential for adaptation to a wide range of environmental conditions. Besides yield and other agronomic traits, the improvement of cold tolerance at juvenile and reproductive stages [24–26] is a major breeding objective for sorghum temperate cropping regions. Early seedling vigour is critical for crop establishment in any environment [27] and vigorous germination and growth under low temperatures is essential for early establishment and weed competition in temperate climates. Improving cold tolerance in the early juvenile stage allows higher yield potential and better maturity. The yield of sorghum is highly temperature dependent, especially between sowing and flowering time [28, 29]. Hence, breeding for juvenile cold tolerance is of utmost importance, especially for temperate European climates. In this study, juvenile survival under low temperature was studied for multi environments. In contrast to emergence and juvenile biomass under cold conditions which have been extensively studied in several publications (e. g. 23,28,29), the trait juvenile survival has received much less attention so far. Though, it is of utmost importance, because a satisfying emergence is worthless if the seedlings later succumb to cold stress. Promising candidates for cold tolerance during juvenile development were identified. Since population stratification was accounted for before performing GWAS, we have reduced the likelihood that genetic background effects are generating spurious associations. We also learned that multiple QTL identified in previous studies [8] and genes known to be involved in cold stress endurance were physically co-located with our QTL, suggesting the importance of our selected genomic regions in this regard. The QTL on Sb09 identified in our study did not overlap with previously reported loci, indicating the importance of this understudied germplam in association to cold tolerance. This draws attention towards the importance of studying novel genetic resources and genome regions containing alleles which can be mined for improvement of cold temperature adaptation.

Given the complex genetics underlying juvenile cold tolerance [29–31], promising approaches like genomic prediction or a genotype-to-phenotype modelling approach can be implemented to assess performance of promising accessions [32]. In addition, after narrowing down genomic region involved in cold tolerance, precise gene editing tools such as CRISPR-Cas9 system can potentially be implemented to validate genes with possible positive effects on abiotic stress tolerance like low temperatures [33]. It would also be interesting to study differential gene expression networks of identified candidate genes in order to to understand their role and elucidate molecular mechanism towards adaptation to abiotic stress responses.

Conclusions

To our knowledge, this is the first extensive study of the unique and large sorghum germplasm collection conserved at the National Genebank of Uganda. This study focuses on two important aspects, (i) genetic characterization of an underutilized novel germplasm and (ii) dissection of cold tolerance trait within this dataset. The population structure results indicate immense genetic and racial diversity within the germplasm predominated by admixed accessions. Contrast to other geographical regions in Uganda the accessions from the south-western

highlands displayed a unique pattern, composed mainly of the caudatum race and genetically isolated from the other subpopulations. This manuscript can be used a base for precisely characterizing novel germplasm based on genetic data. The genomic and phenotypic data collected in this study provide an objective criterion for the selection of accessions for genetic diversity preservation and management, utilization in breeding programs and genetic relationship analysis with other germplasm collections.

A comprehensive investigation involving survival traits, identified multiple key associations and genes underlying the response of juvenile sorghum seedlings to cold stress. The identification of multiple QTL associated to juvenile cold stress reported here can be used by breeders to enhance early-stage chilling tolerance in sorghum. The results provide important new insight for adaptive crops breeding to support the expansion and stability of sorghum production in the face of increasing abiotic stress constraints and climatic change.

Methods

Plant materials

A total of 3333 diverse Ugandan *S. bicolor* germplasm accessions (UG set) collected by the Plant Genetic Resources Centre at the Uganda National Genebank were used in this study (Table S1). This germplasm collection represents the entire sorghum diversity from all eco-geographical regions of Uganda ranging from arid and semi-arid areas in eastern parts to the cold highlands of the Kigezi region in southwestern Uganda (Fig. S2). The landraces are well adapted to different local agroecological conditions in terms of elevation, climate, soil and usage. However, these samples are sensitive to photoperiodism and fail to transition from vegetative to reproductive stage if photoperiods are longer than 12 hours, so that morphological classification into racial groups based on floral morphology was not possible in our field trials in Germany. Instead, for racial composition analysis, the collection was compared to a global sorghum germplasm collection of 1033 genotypes (global set) which was previously described by [13].

DNA extraction and genotyping

Diversity Arrays Technology Pty Ltd. (www.diversityarrays.com) for DNA extraction. The DNA samples were then genotyped using DArTseq, an efficient genotyping-by-sequencing (GBS) platform which enables discovery of genome-wide markers through genome complexity reduction using restriction enzymes. Genotyping was performed fundamentally as described in references [34–36] using the PstI+BanII complexity reduction method. The resulting microarrays were scanned to analyse and score markers by dedicated software DArTsoft (DArT P/L, Canberra, Australia). The sorghum reference genome version v3.1.1 [37] was used for sequence alignment and single nucleotide polymorphism (SNP) calling.

SNP data filtration

A total of 40,290 SNP markers were reported for the global and UG sets. Firstly, all nonspecific markers and those belonging to supercontigs were removed. The remaining 34,469 were used for further analysis. For racial ancestry analysis, an extremely high stringency was then applied to remove markers and genotypes which exhibited greater than 1% missing data. The global diversity set comprised conversion lines containing introgressed chromosome regions from the Sorghum Conversion Program conducted by Texas Agricultural Experiment Station [38]. Hence, to reduce disparity with UG samples in the co-analysis with the global set, all markers from the specific genomic regions impacted by the conversion program were excluded as follows: Sb06: all markers, Sb07: all markers beyond 40 Mb, Sb09: all markers beyond 46 Mb. Markers with minor allele frequency (MAF) less than 0.01 were also excluded. This set was then imputed using Beagle 5.1 [39] to infer the remaining missing data values. A total of 2331 common markers between the UG and global set were used for the racial ancestry analysis.

Population structure and genetic diversity study

In order to understand the racial classification and the population structure of the UG sorghum collection, principal component analysis (PCA) and Discriminant Analysis of Principal Components (DAPC) were implemented using the R package Adegenet (2.1.3) [40]. To avoid bias caused by the large size of the UG set, we used a representative subset of the UG set to assign racial groupings in comparison to the global collection. Usually, sorghum racial classification is mainly based on morphological and phenotyping data. However, because the collection is not long-day adapted, the vast majority of the accessions do not reach maturity under the growth conditions in Germany, hence we were unable to classify into racial groups based on morphological data in the field trials performed in this study. However, ongoing analysis of the materials in tropical environments of Uganda will likely provide this missing information in future and enable validation of the genetic classifications.

To select the UG subset representing the meta-population, we initially clustered the entire set into 10 groups followed by randomly sampling 10 accessions from each group. These randomly-selected samples were combined

with the global dataset for DAPC co-analysis. To validate the racial assignment of the UG set groups, the DAPC co-analysis was repeated three times, each time using a different set of 10 random genotypes from each of the 10 clusters. The SNP data was converted to the genlight object bit-level genotype coding scheme using the function 'vcfR2genlight' of the vcfR tool (https://github.com/knausb/vcfR). After an initial transformation using the PCA analysis, racial composition was subsequently identified for the clusters using discriminant analysis (DA).

To describe the population structure of the UG germplasm and evaluate the racial ancestry of each group in relation to geographical origin of the accessions, the filtered marker set (34,466) was pruned to exclude SNPs which were in strong LD using PLINK software [41]. Pruning was performed using a window of 50 SNPs, step size of 5 makers and r^2 threshold of 0.5. Finally, a total of 12,742 markers for all UG lines were used to analyse population structure using Discriminant analysis (DA). To elaborate the genetic relationship among the accessions a pairwise distance matrix was established using the tool VCF2Dis (https://github.com/BGI-shenzhen/VCF2Dis), which was then converted to a neighbour-joining phylogenetic tree using the R package ape (5.5) [42] and visualized with R package ggtree (3.0.4) [43]. To study patterns of genetic differentiation, Fst values were calculated for each of the five geographical UG subpopulations against the whole set. This was implemented using the 'popgen' function of R package snpReady (0.9.6) [44].

Phenotyping and association mapping

Juvenile survival under cold stress was evaluated through two field trials at Gross Gerau (GG), Germany (spring 2019 and 2020) and one climate chamber experiment. A UG subset of 444 (field trials) and 255 (climate chamber) accessions representing all agro-ecological zones were used for the association study. For the field experiments (Table S7), all genotypes were sown in microplots consisting of single rows (2.5×0.7 m) using an alpha lattice block design with two replications. While the recommended sowing time for sorghum in southern Germany is mid-May, the goal of our field experiments was to induce cold stress. Hence, they were sown notably earlier in spring, on April 8 in 2019 and on April 22 in 2020. Even though the mean soil temperatures were relatively high during the course of the experiments (13.8 and 16.2 °C, respectively) and permitted satisfying emergence, occurrence of several cold nights (up to −1.5 °C) during both the years implied strong stress on the seedlings. Around 4 weeks after emergence, after a period of 7 days from the last occurrence of cold event, the number of surviving plants was scored per plot. For subsequent GWAS analysis, the alpha-lattice adjusted mean value of both years was used. For the climate chamber (CC) experiment, 16 seedlings per genotype were established in $12 \times 12 \times 12$ cm pots. Experiments were designed as randomized complete block design with four replications (Table S8) and the number of surviving plants per pots was scored after 60 days of cold period as a measure for juvenile cold stress tolerance.

After eliminating markers and genotypes with more than 25% missing data points, the dataset was imputed using Beagle 5.1 [39]. It was further corrected for MAF lower than 5%. A total of 4099 markers were used for genome wise association study (GWAS) implemented in R package GenABEL (1.8–0) [45] for juvenile cold tolerance traits. Prior to performing GWAS, population stratification was accounted for by including principal components of the genotypes and genomic kinship matrix [46]. In an effort to reduce the type II error rate and classify a marker–trait association as significant, a threshold of $-\log10\ (p\ \text{value}) \geq 3.0$ was defined [47]. Linkage disequilibrium across the entire genome was calculated using the squared allele frequency correlations (r^2) between each pair of SNPs. Haplotype blocks were calculated using an LD threshold of $r^2 > 0.7$ implemented in the tool LDBlockShow [48], as described by [49].

Candidate genes were selected based on the *Sorghum bicolor* reference genome *v3.1.1* hosted by Phytozome 12 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sbicolor). Orthologous genes within selected haploblocks were identified by homology comparisons of the genomic sequences in maize and rice. Protein sequence alignments were conducted by using blastp option of DIAMOND [50].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-022-03770-y.

Additional file 1: Supplementary Figure 1. Scatter plot representing DA loading of Ugandan accessions based on races. (a) and (b) are replicate 2 and 3 respectively. Each dot represents an individual and the colour code is displayed in the index.

Additional file 2: Supplementary Figure 2. Topographic map of Uganda showing the elevation.

Additional file 3: Table S1. DAPC analysis of Ugandan lines based on races and geographical origin. **Table S2.** Population genetics indices for the sorghum accessions from different districts of geographical regions in the Uganda National GeneBank. **Table S3.** Statistical data of the field experiments and the climate chamber experiment. **Table S4.** Genome-wide association study reveals the genetic basis of juvenile cold tolerance trait in sorghum. **Table S4.** Summary of different cold tolerance traits QTL overlapping the haploblock regions. **Table S6.** Summary table of genes identified in the haploblock regions involved in cold stress tolerance. **Table S7.** Conditions of the field experiments conducted in Gross-Gerau, Germany (49°55′ N, 8°29′ E) in 2019 and 2020. **Table S8.** Conditions of the climate chamber experiment.

Acknowledgements
The authors thank Mario Tolksdorf for assistance with field trials along with Annette Plank and Birgit Keiner for the assistance in greenhouse. The computational analysis was supported by the BMBF-funded de. NBI Cloud within the German Network for Bioinformatics Infrastructure (de. NBI).

Authors' contributions
SC generated the data, conducted the data analysis and wrote the manuscript. RM curated the material and assisted in data collection. SW planned and oversaw the field trials and data collection and assisted in data analysis. DJ and EM provided comparative data and assisted in data analysis. RJS conceived the study and edited the manuscript. AH designed joint data analysis concepts and assisted in the data analysis. The author(s) read and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. The study was funded by grant number SN 14/21–1 from the German Research Foundation (DFG) to RJS.

Availability of data and materials
The raw data generated and/or analysed during current study has been deposited to the NCBI short-read archive under the Bio-project number PRJNA779225. All variants reported for the Ugandan material as reported by DArTseq are available at https://doi.org/10.5281/zenodo.6535431. Seeds from the collection are deposited in the Uganda National GeneBank in Entebbe and available upon request according to ITPGRFA procedures by contacting the genebank via https://www.pgrc.go.ug/index.php/contactuspgrc.

Declarations

Ethics approval and consent to participate
All experimental research, including the collection of plant material, complies with relevant institutional, national, and international guidelines and legislation.

Consent for publication
Not applicable.

Competing interests
The authors declare no conflicts of interest.

Author details
[1]Department of Plant Breeding, Justus Liebig University, Giessen, Germany. [2]Uganda National Gene Bank, National Agricultural Research Laboratories, Kampala, Uganda. [3]Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Warwick, QLD 4370, Australia.

Received: 22 February 2022 Accepted: 21 July 2022
Published online: 29 July 2022

References

1. Hariprasanna K, Patil JV. In: Madhusudhana R, Rajendrakumar P, Patil JV, editors. Sorghum: origin, classification, biology and improvement BT - Sorghum molecular breeding. New Delhi: Springer India; 2015. p. 3–20. https://doi.org/10.1007/978-81-322-2422-8_1.
2. National Research Council . Lost crops of Africa: volume I. Washington, DC: grains: National Academies Pres; 1996.
3. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, et al. The Sorghum bicolor genome and the diversification of grasses. Nature. 2009;457(7229):551–6. https://doi.org/10.1038/nature07723.
4. Sorghum DH. Longman scientific & technical. New York: Wiley; 1988.
5. Snowden JD. The cultivated races of sorghum. The cultivated races of sorghum. 1936.
6. De Wet JMJ, Harlan JR. The origin and domestication of Sorghum bicolor. Econ Bot. 1971;25(2):128–35. https://doi.org/10.1007/BF02860074.
7. Reddy VG, Rao NK, Reddy BVS, Rao KEP. Sorghum research reports geographic distribution of basic and inter- sorghum germplasm; 2002. p. 15–7.
8. Mace E, Innes D, Hunt C, Wang X, Tao Y, Baxter J, et al. The Sorghum QTL atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. Theor Appl Genet. 2019;132(3):751–66.
9. Zhou X, Zheng Y, Cai Z, Wang X, Liu Y, Yu A, et al. Identification and functional analysis of tomato TPR gene family. Int J Mol Sci. 2021;22(2):758 Available from: https://pubmed.ncbi.nlm.nih.gov/33451131.
10. Kim M, Swenson J, McLoughlin F, Vierling E. A temperature sensitive mutation in the CstF77 subunit of the polyadenylation complex reveals the critical function of mRNA 3′ end formation for a robust heat stress response in plants. bioRxiv. 2021, 2021:10.31.466691 Available from: http://biorxiv.org/content/early/2021/11/01/2021.10.31.466691.abstract.
11. Mbabwine Y, Sabiiti EN, Kiambi D. Assessment of the status of plant genetic resources in Kabale Highlands, Uganda; a case of cultivated crop species. Plant Genet Resour Inst (IPGRI) Bioversity Int Via dei Tre Denari. 2004;472:71.
12. Boyles RE, Brenton ZW, Kresovich S. Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments. Plant J. 2019;97(1):19–39 Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/tpj.14113. Cited 2022 May 23.
13. Tao Y, Zhao X, Wang X, Hathorn A, Hunt C, Cruickshank AW, et al. Large-scale GWAS in sorghum reveals common genetic control of grain size among cereals. Plant Biotechnol J. 2020;18(4):1093–105. https://doi.org/10.1111/pbi.13284.
14. Ritter KB, McIntyre CL, Godwin ID, Jordan DR, Chapman SC. An assessment of the genetic relationship between sweet and grain sorghums, within Sorghum bicolor ssp. bicolor (L.) Moench, using AFLP markers. Euphytica. 2007;157(1–2):161–76.
15. Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, et al. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc Natl Acad Sci. 2013;110(2):453–8 Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1215985110. Cited 2022 May 6.
16. Brenton ZW, Cooper EA, Myers MT, Boyles RE, Shakoor N, Zielinski KJ, et al. A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. Genetics. 2016;204(1):21–33 Available from: /pmc/articles/PMC5012387/. Cited 2022 May 11.
17. Stemler ABL, Harlan JR, de Wet JMJ. Evolutionary history of cultivated sorghums (Sorghum bicolor [Linn.] Moench) of Ethiopia. Bull Torrey Bot Club. 1975;102(6):325.
18. Mann JA, Kimber CT, Miller FR. The origin and early cultivation of sorghums in Africa. Bull B Texas Agric Exp Stn. 1983;1454:21.
19. Smith O, Nicholson WV, Kistler L, Mace E, Clapham A, Rose P, et al. A domestication history of dynamic adaptation and genomic deterioration in Sorghum. Nat Plants. 2019;5(4):369–79. https://doi.org/10.1038/s41477-019-0397-9.
20. Rutayisire A, Lubadde G, Mukayiranga A, Edema R. Response of Sorghum to cold stress at early developmental stage. Int J Agronomy. 2021;2021:10.
21. Mekbib F. Farmers' breeding of sorghum in the center of diversity, Ethiopia: I. socioecotype differentiation, varietal mixture and selection efficiency. J New Seeds. 2008;9(1):43–67. https://doi.org/10.1080/15228860701879299.
22. Pressoir G, Berthaud J. Population structure and strong divergent selection shape phenotypic diversification in maize landraces. Heredity (Edinb). 2004;92(2):95–101. https://doi.org/10.1038/sj.hdy.6800388.
23. Akatwijuka R, Rubaihayo PR, Odong TL. Genetic diversity among sorghum landraces of southwestern highlands of Uganda. Afr Crop Sci J. 2016;24(2):179.
24. Schaffasz A, Windpassinger S, Friedt W, Snowdon R, Wittkop B. Sorghum as a novel crop for Central Europe: using a broad diversity set to dissect temperate-adaptation. Agronomy. 2019;9(9):535.
25. Schaffasz A, Windpassinger S, Snowdon R, Wittkop B. Reproductive cold stress tolerance in sorghum F1 hybrids is a heterotic trait. Agronomy. 2019;9(9):508.
26. Chakrabarty S, Kravcov N, Schaffasz A, Snowdon RJ, Wittkop B, Windpassinger S. Genetic architecture of novel sources for reproductive cold tolerance in Sorghum. Front Plant Sci. 2021:2574.

27. Xie L, Tan Z, Zhou Y, Xu R, Feng L, Xing Y, et al. Identification and fine mapping of quantitative trait loci for seed vigor in germination and seedling establishment in rice. J Integr Plant Biol. 2014;56(8):749–59.
28. Craufurd PQ, Mahalakshmi V, Bidinger FR, Mukuru SZ, Chantereau J, Omanga PA, et al. Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. Theor Appl Genet. 1999;99(5):900–11. https://doi.org/10.1007/s001220051311.
29. Fiedler K, Bekele WA, Duensing R, Gründig S, Snowdon R, Stützel H, et al. Genetic dissection of temperature-dependent sorghum growth during juvenile development. Theor Appl Genet. 2014;127(9):1935–48. https://doi.org/10.1007/s00122-014-2350-7.
30. Burow G, Burke JJ, Xin Z, Franks CD. Genetic dissection of early-season cold tolerance in sorghum (Sorghum bicolor (L.) Moench). Mol Breed. 2011;28(3):391–402. https://doi.org/10.1007/s11032-010-9491-4.
31. Bekele WA, Fiedler K, Shiringani A, Schnaubelt D, Windpassinger S, Uptmoor R, et al. Unravelling the genetic complexity of sorghum seedling development under low-temperature conditions. Plant Cell Environ. 2014;37(3):707–23.
32. Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, et al. Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. Crop Pasture Sci. 2014;65(4):311.
33. Zeng Y, Wen J, Zhao W, Wang Q, Huang W. Rational improvement of Rice yield and cold tolerance by editing the three genes OsPIN5b, GS3, and OsMYB30 with the CRISPR–Cas9 system, vol. 10: Front Plant Sci; 2020. Available from: https://www.frontiersin.org/article/10.3389/fpls.2019.01663
34. Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, et al. A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics. 2006;7(1):1–22 Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-7-206. Cited 2022 May 23.
35. Akbari M, Wenzl P, Caig V, Carling J, Xia L, Yang S, et al. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theor Appl Genet. 2006;113(8):1409–20 Available from: https://link.springer.com/article/10.1007/s00122-006-0365-4. Cited 2022 May 23.
36. Mace ES, Xia L, Jordan DR, Halloran K, Parh DK, Huttner E, et al. DArT markers: diversity analyses and mapping in Sorghum bicolor. BMC Genomics. 2008;9(1):26. https://doi.org/10.1186/1471-2164-9-26.
37. McCormick RF, Truong SK, Sreedasyam A, Jenkins J, Shu S, Sims D, et al. The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant J. 2018;93(2):338–54. https://doi.org/10.1111/tpj.13781.
38. Thurber CS, Ma JM, Higgins RH, Brown PJ. Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. Genome Biol. 2013;14(6):1–13.
39. Browning BL, Zhou Y, Browning SR. A one-penny imputed genome from next-generation reference panels. Am J Hum Genet. 2018;103(3):338–48. https://doi.org/10.1016/j.ajhg.2018.07.015.
40. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010;11(1):94. https://doi.org/10.1186/1471-2156-11-94.
41. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
42. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289–90.
43. Yu G. Using ggtree to visualize data on tree-like structures. Curr Protoc Bioinforma. 2020;69(1):e96. https://doi.org/10.1002/cpbi.96.
44. Granato ISC, Galli G, de Oliveira Couto EG, e Souza MB, Mendonça LF, Fritsche-Neto R. snpReady: a tool to assist breeders in genomic analysis. Mol Breed. 2018;38(8):102.
45. Aulchenko YS, de Koning DJ, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics. 2007;177(1):577–85. https://doi.org/10.1534/genetics.107.075614.
46. Stich B, Möhring J, Piepho HP, Heckenberger M, Buckler ES, Melchinger AE. Comparison of mixed-model approaches for association mapping. Genetics. 2008;178(3):1745–54.
47. Gabur I, Chawla HS, Snowdon RJ, Parkin IAP. Connecting genome structural variation with complex traits in crop plants. Theor Appl Genet. 2019;132(3):733–50. https://doi.org/10.1007/s00122-018-3233-0.
48. Dong SS, He WM, Ji JJ, Zhang C, Guo Y, Yang TL. LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. Brief Bioinform. 2021;22(4). https://doi.org/10.1093/bib/bbaa227.
49. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science (80-). 2002;296(5576):2225–9.
50. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12(1):59–60.

Publisher's Note

4 Genetic architecture of novel sources for reproductive cold tolerance in sorghum

Subhadra Chakrabarty, Natalja Kravcov, André Schaffasz, Rod J. Snowdon, Benjamin Wittkop and Steffen Windpassinger
Frontiers in Plant Science, Vol 12, (2021)
doi: 10.3389/fpls.2021.772177

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 November 2021
doi: 10.3389/fpls.2021.772177

Genetic Architecture of Novel Sources for Reproductive Cold Tolerance in Sorghum

*Subhadra Chakrabarty, Natalja Kravcov, André Schaffasz, Rod J. Snowdon, Benjamin Wittkop and Steffen Windpassinger**

Department of Plant Breeding, IFZ Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University Giessen, Giessen, Germany

Enhancements in reproductive cold tolerance of sorghum are essential to expand growing areas into both high-latitude temperate areas and tropical high-altitude environments. Here we present first insights into the genetic architecture of this trait *via* genome-wide association studies in a broad genetic diversity set (n = 330) phenotyped in multi-location field trials including high-altitude tropical (Mexico) and high-latitude temperate (Germany) environments. We observed a high degree of phenotypic variation and identified several novel, temperate-adapted accessions with superior and environmentally stable cold tolerance. Good heritability indicates strong potential for implementation of reproductive cold tolerance in breeding. Although the trait was found to be strongly quantitative, promising genomic regions with multiple-trait associations were found, including hotspots on chromosomes 3 and 10 which contain candidate genes implicated in different developmental and survival processes under abiotic stress conditions.

Keywords: sorghum, GWAS, reproductive cold tolerance, temperate climate adaptation, genetic diversity

OPEN ACCESS

Edited by:
Sukhjiwan Kaur, Department of Jobs, Precincts and Regions, Agriculture Victoria, Australia

Reviewed by:
Daniela Marone, Cereal Research Centre, CRA-CER, Italy
Cheng-Ruei Lee, National Taiwan University, Taiwan

***Correspondence:**
Steffen Windpassinger
steffen.m.windpassinger@agrar.uni-giessen.de

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *07 September 2021*
Accepted: *19 October 2021*
Published: *24 November 2021*

Citation:
Chakrabarty S, Kravcov N, Schaffasz A, Snowdon RJ, Wittkop B and Windpassinger S (2021) Genetic Architecture of Novel Sources for Reproductive Cold Tolerance in Sorghum. Front. Plant Sci. 12:772177. doi: 10.3389/fpls.2021.772177

INTRODUCTION

Sorghum (*Sorghum bicolor* L. Moench, 2n = 20) is of vital importance for global food and feed supply. Due to its tolerance to drought and low-input conditions, it represents an essential staple crop and commodity especially in semi-arid regions of Africa, India, Australia, and both Americas. However, as a tropical C_4 plant, its sensitivity to temperatures below 15°C is a substantial obstruction to successful implementation in both high-latitude temperate climates and tropical high-altitude areas (Singh, 1985). Early juvenile development (e.g., Maulana et al., 2017) and pre-flowering reproductive stage (Brooking, 1976) are considered the most cold-sensitive growth stages. While numerous studies have targeted enhancements of juvenile cold tolerance and analyzed their genetic architecture (e.g., Chopra et al., 2017; Parra-Londono et al., 2018; Schaffasz et al., 2019), comparably little research has focused on reproductive cold tolerance of sorghum to date. Cool temperatures before anthesis are known to induce male sterility in sorghum, leading to complete failure of seed set and grain yield in sensitive genotypes. This phenomenon was first described scientifically by Downes and Marshall (1971). For successful adaption of sorghum to temperate climates as Central Europe, enhancements in reproductive cold tolerance are at least equally important to juvenile cold tolerance (Windpassinger et al., 2015). While farmers can avoid juvenile

cold stress by later sowing (albeit at the expense of maturity and yield potential), there is no escape strategy for unpredictable cold spells during reproductive stages.

Brooking (1976, 1979) reported the pre-leptotene and the leptotene as the most cold-sensitive developmental stages, suggesting meiotic problems in microspore mother cells as a possible reason for this phenomenon. Reproductive cold tolerance in a set of 380 sorghum accessions, and identified tolerance sources originating mainly from tropical highlands (e.g., Ethiopia, Uganda), United States and China (Singh, 1985). Further, this study also provided preliminary information on the inheritance of this trait using factorial F_1 hybrids. Due to the need for cold-tolerant sorghum in the Mexican High Valleys, Mendoza-Onofre (1988) developed cold-tolerant grain sorghum lines with a good local adaption. Over time a steady development of sorghum lines and hybrids with enhanced reproductive cold tolerance has taken place in Mexico (Osuna-Ortega et al., 2000, 2001, 2003; Cisneros-López et al., 2009, 2010; León-Velasco et al., 2009). Furthermore, evaluation of reproductive cold tolerance in the Indian post-rainy season by Krishnamurthy et al. (2014) suggested that *Panicle Harvest Index* (PHI) can be efficiently scored as a proxy for spikelet fertility. In a recent study, Schaffasz et al. (2019) analyzed several traits related to reproductive cold tolerance in a line × tester design, showing a dominant inheritance and heterotic effect in F_1 hybrids.

However, in spite of the above mentioned progress in breeding cold tolerant varieties and first information about the inheritance of reproductive cold tolerance, there is still no knowledge available about genomic regions involved in its inheritance or the detailed genetic architecture of the trait. This has prevented the implementation of molecular breeding approaches like marker assisted selection, genomic prediction and genome editing so far. In the present study, a broad diversity set (n = 330) genotyped using 20K Dartseq markers was utilized for extensive phenotyping of reproductive cold tolerance in multi-environment field trials, aiming at the identification of novel tolerance sources and underlying genomic regions in order to accelerate breeding progress. To our best knowledge, this is the first genome wide association study (GWAS) to date with regard to reproductive cold tolerance in sorghum.

MATERIALS AND METHODS

Germplasm

A *S. bicolor* diversity set consisting of n = 330 inbred lines of different origin, type of use (grain, dual-purpose, and forage) and subspecies, comprising all five sorghum races (*bicolor, caudatum, durra, guinea*, and *kafir*) and most of their respective intermediaries, was utilized for the present study. All these lines are photo-insensitive and show a similar, early maturity, allowing for an adequate scoring of seed set traits under cool environments with a short growing season. A representative selection of early-maturing *sorghum conversion lines*, obtained in the 1960s by repeated backcrossing of genetically diverse tropical accessions to a short, photoperiod-insensitive cultivar (Stephens et al., 1967) accounts for 35% (n = 117) of the set. A further 22% (n = 73) consist of other publicly available accessions from temperate countries, mainly China, Russia, the United States, Turkey, and Hungary. The original germplasm of these accessions and the conversion lines was received from the *United States Department of Agriculture Agricultural Research Service* (USDA-ARS). The remainder of the diversity set consists of 140 diverse breeding lines (grain and dual-purpose sorghum) from a joint breeding program of Justus Liebig University Giessen, Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (Hohenlieth, Germany) and Deutsche Saatveredelung AG (Lippstadt, Germany). The composition of the diversity set is shown in **Supplementary Table 1**.

Field Trials

Field trials were conducted at five locations (three in Germany and two in Mexico, **Table 1**). At Texcoco (TEX17, TEX18), one of the Mexican locations, the diversity set was scored in two subsequent years (2017 and 2018), while at all other locations, the experiments were carried out in 1 year (2017 or 2018). Hence, data from six environments (location × year combination) are available for analyses. The locations of this study represent different mega-environments contrasting in climate, radiation and day length. Among the German locations, Asendorf (AS; located in NW-Germany) and Poel (PL; a small island in the Baltic Sea) have a cool, maritime climate, usually providing harsh conditions for sorghum. In contrast, Gross Gerau (GG) is located in the Upper Rhine Valley and characterized by a warm and sunny climate, being a suitable control environment without cold stress. The locations in Mexico differ strongly from Germany, having shorter days during the growing season but much stronger radiation. While San Juan del Río (SJR; 1,920 m, federal state Querétaro) is considered to be at the altitude limit for commercial sorghum cultivation in Mexico, Texcoco (2,250 m, federal state México) is a tropical high-altitude stress environment for sorghum, providing the lowest minimum temperatures of all locations.

To avoid shading effects by neighbors of different plant height, the inbred lines were split into three subgroups, based on previous scorings of plant height. These subgroups were planted in adjacent but separate blocks at all sites. Within these subgroups, an un-replicated randomized complete block design was used. Entries were grown in microplots, consisting of single rows (2.5 × 0.7 m) at Gross Gerau and double rows (2.5 × 1.4 m) at all other locations, with 0.7 m row spacing and a plant density of approx. 20 plants/m^2. Plant protection and fertilizer application were executed according to site specific best practice. Per entry, the primary panicles of five healthy plants were covered before anthesis with a transparent Cryovac® bag (330 mm × 750 mm, 15 μm) to avoid cross pollination. These five self-pollinated panicles were considered as biological replications for further analyses. However, due to lodging and insufficient maturity, some plants and entries had to be excluded site-specifically. At maturity, the panicles were harvested with secateurs and dried. The peduncles of each panicle were cut just below the first branches before determining the panicle dry weight. Seed yield (SY) per panicle was measured after threshing.

TABLE 1 | Overview and weather data of the different environments during the duration of the experiments (from sowing until harvest of the panicles).

Environment	Coordinates	Altitude	Soil type	Year	Mean temp. (°C)	Mean max. temp. (°C)	Mean min. temp. (°C)	Precipitation (mm)
PL	53°99'N, 11°47'E	19 m	Loamy sand	2017	16.7	20.2	13.1	301
AS	52°46'N, 9°01'E	49 m	Loamy sand	2017	16.1	21.0	11.6	611
GG	49°55'N, 8°29'E	90 m	Sand	2018	21.3	28.9	13.7	94 (+150 irrigation)
SJR	20°25'N, 99°56'W	1,920 m	Loam	2017	23.0	31.4	14.7	326
TEX17	19°31'N, 98°51'W	2,250 m	Loam	2017	16.4	24	8.7	360
TEX18				2018	16.3	23.2	9.4	537

Environments: Poel (PL), Asendorf (AS), Gross-Gerau (GG), San Juan del Río (SJR), Texcoco 2017 (TEX17), and Texcoco 2018 (TEX18).

Subsequently, PHI was calculated according to Krishnamurthy et al. (2014) as follows:

$$\text{PHI} = \text{grain dry weight (i.e., seed yield per panicle)} / \text{panicle dry weight (before threshing)}$$

Consequently, a PHI value of 0 implies absolutely no seed set, while values close to 1 indicate a high seed set. However, even assuming complete spikelet fertility, PHI will be <1 due to the panicle raw weight. Moreover, seed number (SN) was measured using a Contador seed-counter (Pfeuffer, Germany). In addition to these seed set traits, plant height and start of flowering (in all environments except SJR) were scored on a plot basis.

Statistical Analyses of Phenotype Data

For statistical analyses of SY, seed number (SN) and PHI, a general linear model was used (IBM SPSS Statistics version 27, IBM Software, Armonk, NY, United States), in which entries (genotypes) and environments (combination of location and year) were considered as fixed effects and replicates (individual plants) as random effects:

$$Y_{ijk} \sim \mu + G_i + E_j + GE_{ij} + R_{kj} + e$$

where μ represents the population mean, G_i is the genotypic effect, E_j is the environmental effect, GE_{ij} is the genotype-by-environment interaction, R_{kj} is the replicate effect, and e is the residual effect.

To compare the levels of genotypic variance obtained in the different environments, ANOVA was also computed separately for each environment, using the following general linear model, where genotypes were considered as fixed and replicates as random effects:

$$Y_{ij} \sim \mu + G_i + R_j + e$$

To determine which environments were significantly different from one another, Student-Newman-Keuls *post hoc*-test was applied.

The heritability was calculated as proposed by Piepho and Möhring (2007) using the following formula:

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G + \frac{1}{2}\overline{vd}}$$

where h^2 represents broad-sense heritability, $\sigma^2{}_G$ is the genotypic variance calculated by a random effect model considering genotype and environment as random factors, and $\overline{vd}$ is the average variance of the difference between two means.

The phenotypic stability of the best performing inbred lines with a complete data set in all stress environments was analyzed using the coefficient of variation (CV) across the mean trait values of a line in each environment.

Association Mapping and Candidate Gene Identification

The sorghum diversity set was genotyped using DArTseq reduced-representation sequencing[1] to identify genome-wide single-nucleotide polymorphism (SNP) markers at high resolution. SNPs with more than 20% missing data or a minor allele frequency lower than 5% were removed from the final dataset. After filtering, a total of 21,520 high-quality SNPs remained and were used for downstream analyses.

Analysis of phylogenetic relatedness was conducted with TASSEL version 5.0 (Bradbury et al., 2007) using the neighbor-joining method (Saitou and Nei, 1987). Dendroscope 3.7.3 (Huson et al., 2007) was utilized to visualize the genetic relatedness among accessions in a phylogram. The R package GenABEL (Aulchenko et al., 2007) was used to perform a GWAS for the target traits. To adjust population stratification, a mixed linear model approach was implemented by using the kinship matrix as covariates (Stich et al., 2008). We used a threshold of $-\log10(p) \geq 3$ to minimize type II error and identify SNP-trait association (Gabur et al., 2020).

Candidate genes were identified based on the *S. bicolor* reference genome *v3.1.1* hosted by Phytozome 12[2], the same reference genome used by Diversity Array Technology to call SNPs. Linkage disequilibrium and haplotype blocks across the entire genome were calculated using the squared allele frequency correlations (r2) between each pair of SNPs, using an LD threshold of $r^2 > 0.7$ block as described by Gabriel et al. (2002) and implemented using the tool LDBlockShow (Dong et al., 2021).

Genes within haploblocks surrounding trait-associated SNPs with the maximum $-\log10(p)$ values from GWAS analysis were selected and annotation was checked using the sorghum

[1]https://www.diversityarrays.com/

[2]https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sbicolor

reference assembly. The gene sequences were further blasted against maize and rice for validation and identification of potential candidates. Additionally, we used the high resolution, open access sorghum QTL Atlas (Mace et al., 2019) to identify and characterize overlapping abiotic stress QTL.

We compared the sequence of *S. bicolor* reference genome *v3.1*.1 and the sequence based on the alternate allele discovered at the SNP position Sb10-5394955 with the recently published sorghum pangenome data (Tao et al., 2021). The gene sequence was selected from the reference genome (*v3.1*.1) and blasted (blastn: 2.2.31; Camacho and Madden, 2013) against the individual thirteen assembled genomes from the sorghum pangenome dataset. The fasta sequence of the results were individually trimmed from the respective genome sequences using bedtools (v2.30.0; Quinlan and Hall, 2010) and aligned using MUSCLE (v3.8.31; Edgar, 2004). The alignment was visualized using SeaView5 (Galtier et al., 1996).

RESULTS

Phenotypic Variation for Seed Set Traits

The temperature conditions and consequently the level of cold stress differed strongly among the phenotyping environments (**Table 1**), which is also reflected in a high degree of environmental variance on the traits (**Table 2**). Environments AS, PL, TEX17, and TEX18 showed a significant reduction of PHI, SY, and SN compared to the environments without thermal stress (GG and SJR; **Figure 1** and **Table 2**). In consequence, GG and SJR were regarded as control environments for further analyses, whereas all other environments were considered and pooled as cold stress environments. Highly significant differences among the entries for the scored seed set traits were observed in both environmental groups. However, as expected, the CV was higher for the group "stress environments," especially for the trait PHI. While genotype × environment interaction between the two control environments GG and SJR was high and of a similar magnitude as genotypic variance, it was lower for the stress environments (depending on the trait, 29–40% of genotypic variance). Heritability estimates for the groups "stress environments" and "all environments" were high for both PHI (h^2 = 0.72 and 0.69, respectively) and SN (h^2 = 0.68 and 0.63, respectively), and somewhat lower for SY (h^2 = 0.60 and 0.51, respectively).

The race (i.e., panicle architecture) of the genotypes did not influence SY, SN, or PHI ($p \geq 0.063$), regardless of whether one-way ANOVA was applied for the pooled groups of stress environments, control environments or all environments. Comparing the public material (conversion lines and other public accessions) and the breeding lines developed under Central European conditions as groups, the latter were superior ($p < 0.001$) for all considered traits under the pooled stress environments, while under the control environments a difference between the groups was only detected for SY ($p = 0.004$). The phenotype data of all entries are provided in **Supplementary Table 1**.

Trait Correlations

As expected, SY, SN, and PHI were highly correlated with each other (**Table 3**). However, the correlation of PHI to SY and SN was stronger for the group of stress environments (r = 0.880*** and 0.780***, respectively) than for the control environments (r = 0.647*** and 0.578***, respectively). In contrast, TKW was only weakly related to these traits, implying SN as the predominant factor for SY and PHI in all environmental groups. Plant height showed a significant and positive correlation to SY, PHI and – at a lower level especially in control environments- also SN. Whereas TKW correlated with plant height in the control environments, no relation between these traits was found in the stress environments.

Trait correlations were inconsistent among the different single environments (**Supplementary Table 2**). While moderate correlations were observed between AS, PL, TEX18 and SJR, TEX17 showed only weak correlations to the other environments. Surprisingly, GG showed no correlation to any of the other environments ($p \geq 0.103$). Pearson's correlation coefficients of r = 0.343*** for SY, r = 0.442*** for SN and r = 0.385*** for PHI were observed between the pooled groups of stress and control environments.

TABLE 2 | Variances (mean squares) for entries (genotypes), environments (Env), entries × environment (Env) interaction and heritability estimates for the traits seed yield per panicle (SY), seed number (SN), and panicle harvest index (PHI).

Items	All environments				Stress environments				Control environments			
	d.f.	SY (g)	SN	PHI	d.f.	SY (g)	SN	PHI	d.f.	SY (g)	SN	PHI
Entries	328	736.18***	1.756×10^6 ***	0.393***	326	434.26***	1.257×10^6***	0.456***	327	755.21***	1.228×10^6 ***	0.076***
CV entries		211.81	229.68	127.89		272.05	286.79	177.77		116.34	115.23	38.83
Env	5	104,344***	143.878×10^6 ***	53.03***	3	7432.54***	19.985×10^6***	9.22***	1	71.88	17.463×10^6 ***	0.359***
Entries × Env	1,520	367.52***	0.654×10^6 ***	0.12***	877	172.99***	0.402×10^6 ***	0.130***	318	809.98***	1.223×10^6 ***	0.064***
Error	6,847	43.2	0.074×10^6	0.014	4,421	25.05	0.054×10^6	0.017	2,426	76.28	0.11×10^6	0.010
Heritability (H^2)		0.51	0.63	0.69		0.60	0.68	0.72		0	0.01	0.17

*Stress environments: Poel (PL), Asendorf (AS), Texcoco 2017 (TEX17), Texcoco 2018 (TEX18); control environments: Gross-Gerau (GG), and San Juan del Río (SJR). ***P ≤ 0.001.*

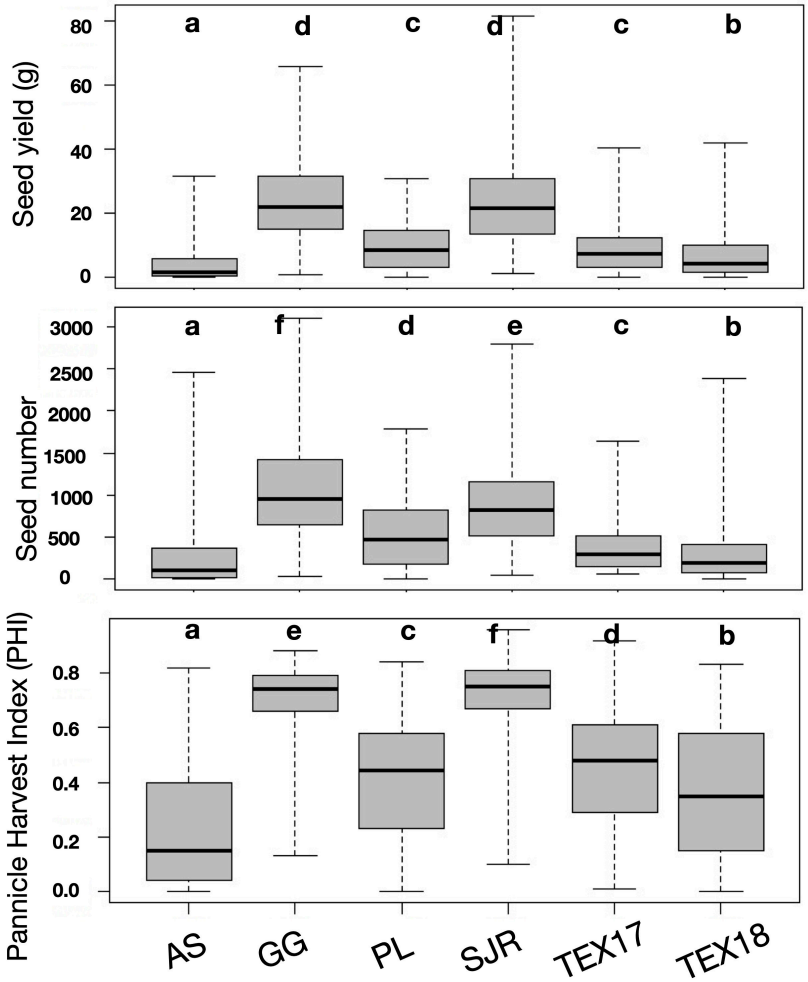

FIGURE 1 | Boxplots showing the variation for the traits seed yield, seed number and panicle harvest index (PHI) across the different environments. Denotation with different letters indicates significant differences as revealed by a Student-Newman-Keuls (SNK) *Post hoc*-Test ($\alpha = 0.05$).

TABLE 3 | Pearson's correlation coefficient (*r*) among the traits seed yield (SY), seed number (SN), panicle harvest index (PHI), thousand kernel weight (TKW), and plant height (PH) across the different environmental groups [stress environments: Poel (PL), Asendorf (AS), Texcoco 2017 (TEX17), Texcoco 2018 (TEX18); control environments: Gross-Gerau (GG), San Juan del Río (SJR)], based on accession-level mean trait values.

	All environments				Stress environments				Control environments			
Traits	**SY**	**SN**	**PHI**	**TKW**	**SY**	**SN**	**PHI**	**TKW**	**SY**	**SN**	**PHI**	**TKW**
SN	0.902***				0.914***				0.880***			
PHI	0.804***	0.723***			0.880***	0.780***			0.647***	0.578***		
TKW	0.169**	-0.157*	0.291***		0.253***	-0.051	0.321***		0.259***	-0.166**	0.259***	
PH	0.380***	0.289***	0.401***	0.122	0.325***	0.306***	0.382***	0.013	0.296***	0.158*	0.269***	0.283***

*$P \leq 0.05$, **$P \leq 0.01$, and ***$P \leq 0.001$.

The effect of flowering time on the yield-related traits was generally weak and inconsistent over the environments (**Supplementary Table 3**), as expected due to the vast ecogeographical variation between the testing locations. For the control environment GG, there was no effect of flowering time at all ($p \geq 0.053$). Whereas a significant and moderately negative correlation between days to flowering and TKW was observed for the temperate, short-season stress environments of AS and PL, this effect was much weaker for the tropical stress environments (insignificant for TEX17 and weak for TEX18). Also for PHI, later flowering genotypes showed a slight disadvantage at three out of four stress environments. In contrast, later flowering resulted in no penalty for SN nor SY (with the exception of location PL).

A substantial proportion (n = 110) of the diversity set had previously been extensively scored for juvenile cold tolerance and early vigor, in independent field trials and climate chamber experiments (Schaffasz et al., 2019). However, only few, weak and inconsistent correlations between juvenile and reproductive cold tolerance could be found. Seedling emergence of an early-sown field experiment in Giessen, Germany in 2014 showed a positive, but weak correlation to mean SN (r = 0.197*) and PHI (r = 0.200*) of the pooled stress environments in the present study. Similarly, seedling emergence under controlled cold stress (13/10°C) was also positively related (r = 0.204*) to mean PHI of stress environments. In contrast, mean juvenile biomass of four field trials (Schaffasz et al., 2019) correlated negatively with mean SN (r = −0.221*) and mean PHI (r = −0.198*) of the pooled stress environments in the present study. Furthermore, leaf greenness, a score for photosynthetic activity under controlled and prolonged cold stress, was negatively related to PHI as well (r = −0.238*).

Accessions With Superior Reproductive Cold Tolerance

To identify accessions with a superior and stable reproductive cold tolerance, the 5% (n = 12) best performing genotypes for each trait, based on their mean over all four stress environments [for which only entries with a complete dataset (n = 240) were considered], were further dissected for their stability, expressed by the CV across the mean trait values of a line in each environment. **Figure 2** shows that breeding lines SB14109 and SB14122 attained the highest PHI scores. Among the public material, accessions PI53578101 (from United States), PI550586 and PI550599 (both from Russia) showed the best results. Regarding SN, the conversion line SC37 (originating from Ethiopia) reached the highest mean value, but a rather low environmental stability due to a poor performance in Texcoco2017. In contrast, SC20 (also originating from Ethiopia) showed a very high environmental stability at the expense of just a slight reduction in mean performance, hence appearing to be the preferable tolerance source. Further, public lines SC701 (from Sudan) and PI176767 (from Turkey) ranked among the top 5% with regard to reproductive cold tolerance. As expected due to the high trait correlations, SC37, SC701, and SC20 (again at a superior environmental stability) were among the top-performers for SY. Breeding line SB14022 reached the highest mean SY, but showed a rather high environmental fluctuation. Moreover, public lines SC1069 and SC566 (both from Nigeria) were among the top performers for SY under cold stress. Chinese *kaoliang* types (n = 15), which are known for their superior juvenile cold tolerance, showed only moderate levels of reproductive cold tolerance (**Supplementary Table 1**). This finding is concordant with the low and inconsistent correlation between juvenile and reproductive cold tolerance (see section "Trait Correlations").

As presented in **Table 4**, ten lines ranked among the top 5% of performers for more than one trait. Of these, line SB14022 excelled in all three traits simultaneously and is a strong candidate for further focus in breeding.

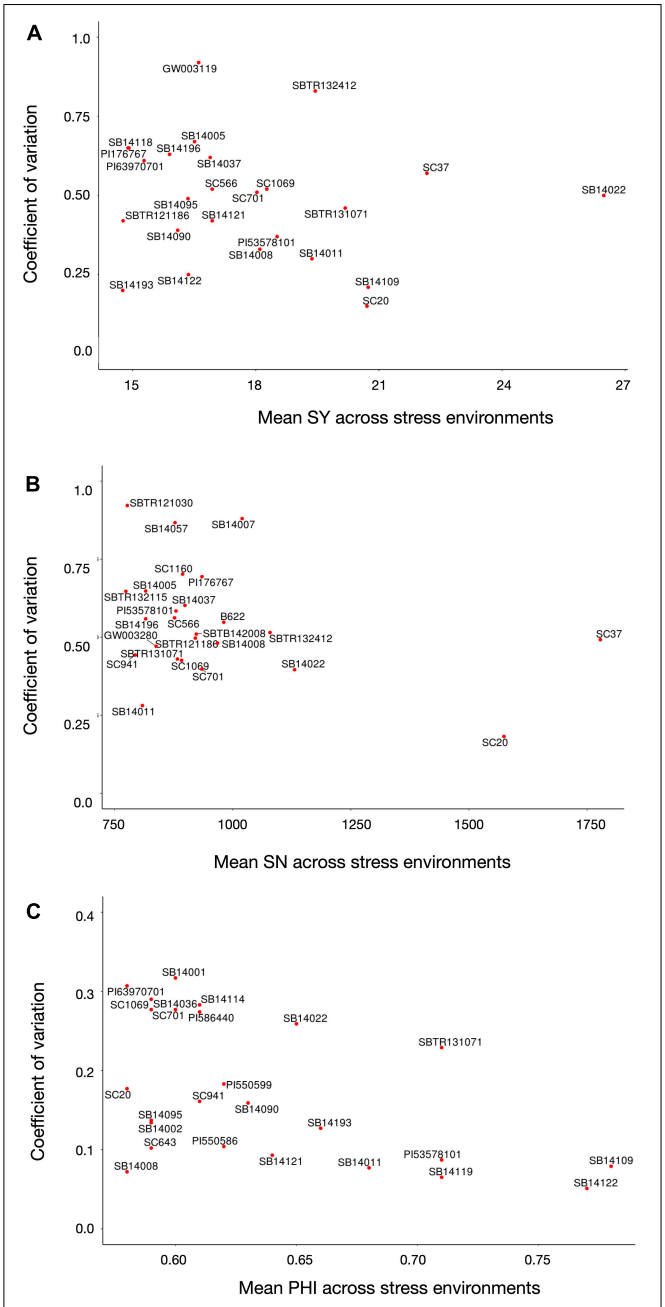

FIGURE 2 | Mean values and stability, depicted as coefficient of variation across the mean trait values of a line in each environment (low values indicate a higher stability), of the top 5% (n = 12) best performing genotypes with complete dataset over all four stress environments for **(A)** seed yield per panicle, **(B)** seed number, and **(C)** and PHI.

Phylogenetic Diversity of Cold Tolerant Lines

The phylogenetic relatedness of the diversity panel, and the localization of genotypes with superior reproductive cold tolerance is depicted as a phylogram in **Figure 3**, and as a principal component analysis in **Supplementary Figure 1**.

TABLE 4 | Overview of the lines ranking among the top 5% performers for more than one trait related to reproductive cold tolerance.

Genotype	Description	Origin	Among top 5% for traits
PI53578101	Public accession	United States	PHI, SY
SC20	Conversion line	Ethiopia	SN, SY
SC37	Conversion line	Ethiopia	SN, SY
SC701	Conversion line	Sudan	SN, SY
SB14008	Breeding line	Germany	SN, SY
SB14011	Breeding line	Germany	PHI, SY
SB14022	Breeding line	Germany	PHI, SN, SY
SB14109	Breeding line	Germany	PHI, SY
SBTR131071	Breeding line	Netherlands	PHI, SY
SBTR132412	Breeding line	Netherlands	SN, SY

Genotypes with superior reproductive cold tolerance tend to cluster within related branches. However, these branches are rather evenly distributed across the phylogram, indicating that sources of reproductive cold tolerance are not confined to a particular origin.

Association Mapping and Candidate Gene Identification

Using pooled phenotype data from the stress environments AS, PL, TEX17 and TEX18, GWAS identified a total of 55 marker-trait associations for PHI (5 associations), SN (21), SY (9), TKW (9), DTF (1), and PH (10), respectively, (**Figure 4** and **Supplementary Figures 2, 3**). Full details of chromosome positions of identified MTAs (marker trait associations) for each trait, including explained variance (R^2) values and *p*-values, are provided in **Supplementary Table 4**. Due to the high phenotypic correlations among the SY traits, some overlapping associations were detected between the traits under different environments (**Supplementary Table 5**). The average phenotypic variance explained (R^2) was 4.34 for the stress environments.

Considering the most relevant traits in regard to reproductive cold tolerance, namely PHI and SY, marker-trait associations were identified on chromosomes Sb01, Sb05, Sb06, Sb09, and Sb10 (**Supplementary Figure 2** and **Supplementary Table 4**). SNP marker Sb10-5394955 on chromosome Sb10 was associated to both PHI and SY (and also associated with TKW) in the cold-stress environments (**Figure 4**).

The SNP marker Sb10-5394955 on chromosome Sb10, which showed multiple cold-related trait associations to PHI, SY, and TKW, showed relatively weak LD to the adjacent SNPs on either side ($R^2 < 0.8$; **Figure 4**), indicating recombination within the haploblock region. Based on the *S. bicolor* reference genome assembly (*S. bicolor* reference genome *v3.1.1*; https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sbicolor), the haplotype block containing Sb10-5394955 overlapped with the gene Sobic.010G067300 and no other

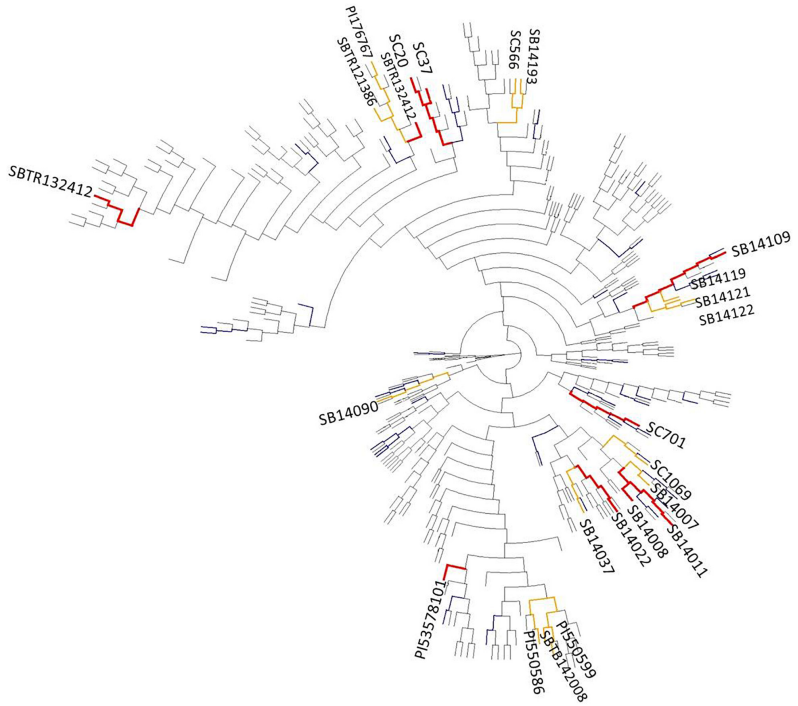

FIGURE 3 | Phylogram constructed using the neighbor-joining method, showing the genetic relatedness in the diversity set (*n* = 330) and among lines with superior reproductive cold tolerance (colored bars). Genotypes ranking among the top 5% of performers for more than one cold tolerance trait (**Table 4**) are highlighted in red, while genotypes that are among the top 5% for a single trait (**Figure 2**) are highlighted in orange. Genotypes with a mean panicle harvest index (PHI) in the upper 5–25% quantile in the pooled stress environments are colored in blue, unless they were previously assigned to one of the two previous groups.

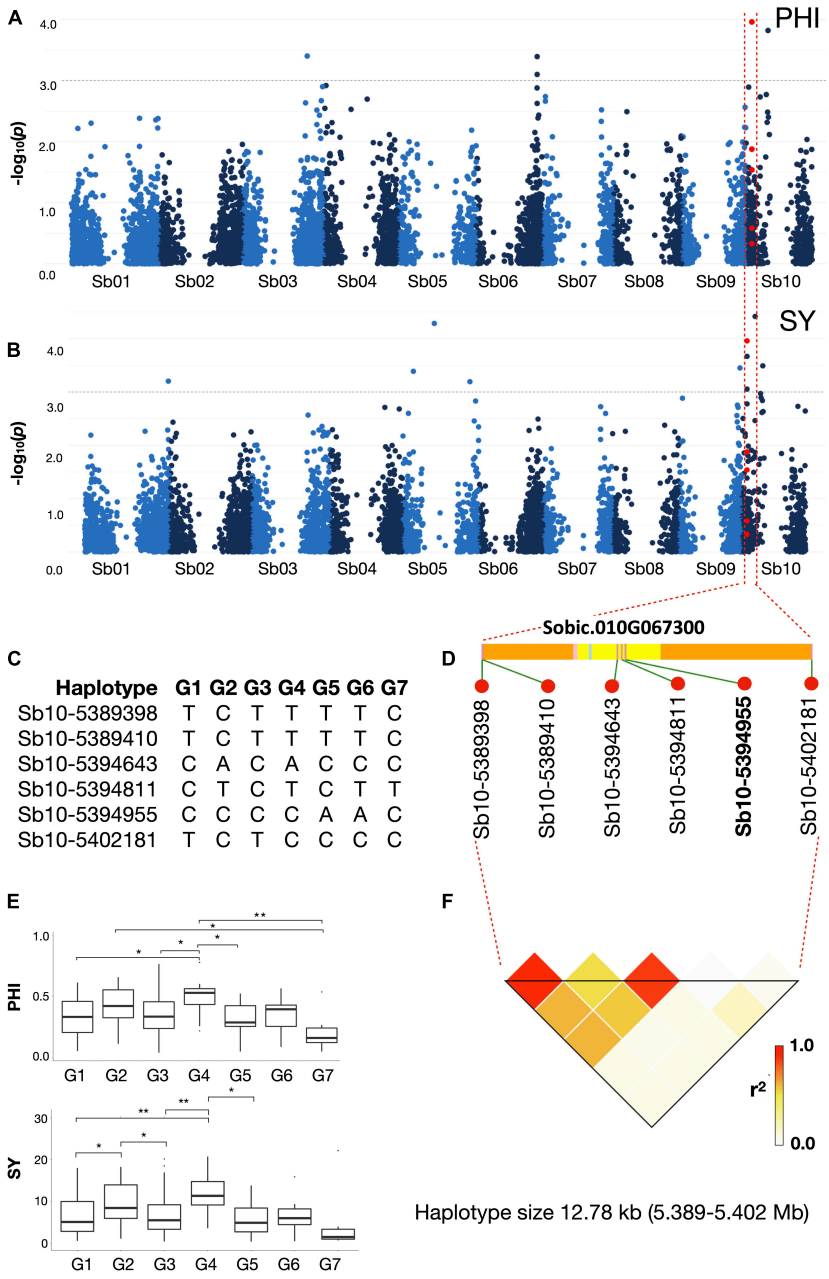

FIGURE 4 | Association mapping of **(A)** panicle harvest index (PHI) and **(B)** seed yield (SY) on chromosome Sb10. Red dots represent a haplotype region spanning just 12.78 kb (5,389,398–5,402,181 bp) including SNP marker Sb10-5394955 which is significantly associated with both PHI and SY under all combined cold stress environments. The gray dotted horizontal line indicates a threshold of genome- wide cut-off at -log10(p) > 3.0. The SNP markers with the highest p values for each trait were used to define a haplotype region with linkage disequilibrium (LD) to the traits. **(C)** shows the seven haplotype groups (G1 – G7) present in the region. **(D)** The SNP Sb10-5394955 (marked in bold) lies in the coding region of gene Sobic.010G067300 (yellow: CDS; light blue: Intron; pink: UTR; orange: Intergenic region) and **(F)** shows the haploblock around this region. **(E)** *Post Hoc* result for the seven haplotype groups for traits PHI and SY. Boxplots indicate presence of significant difference between haplogroups G1-G4, G4-G3 and G5-G4 for both PHI and SY under all cold stress environments combined.

genes were present in the region. Strong synteny to the very well-annotated public reference genomes of rice[3] and maize (B73 RefGen_v5; https://www.maizegdb.org/genome/assembly/Zm-B73-REFERENCE-NAM-5.0) was exploited to obtain putative functional information for this gene. Protein sequence BLAST of the genes present in syntenic rice and maize chromosome regions revealed that *Sobic.010G067300* is a leucine-rich repeat receptor-like protein kinase gene. It has been found to be critically involved in cold response in rice and soybean (Gao and Xue, 2012; Yang et al., 2014).

[3]https://plants.ensembl.org/Oryza_sativa/Info/Index

Haploblocks surrounding each significant SNP were further investigated using the sorghum QTL Atlas (Mace et al., 2019) for overlaps to known QTL for relevant traits in other sorghum populations. QTL overlapping with the associated markers were identified for all traits under both stress and stress+control conditions (**Supplementary Table 6**). On each of chromosomes Sb03 and Sb10 we identified four overlapping, previously identified QTL related to cold stress overlapping the haploblock regions of traits PHI and SY (**Supplementary Figure 4**).

DISCUSSION

Phenotypic Variation for Reproductive Cold Tolerance in Sorghum

The high amount of phenotypic variation for reproductive cold tolerance observed in the present study, along with satisfying heritability estimates and identified tolerance sources with high environmental stability, suggests that a robust breeding progress for this crucial adaptation trait is feasible, even though marker-assisted selection will be challenging due to its quantitative inheritance. The cold stress environments chosen for this study represent contrasting mega-environments (temperate vs. tropical). While the lowest minimum temperatures occurred at the Mexican highland station Texcoco, the most severe stress reaction of the plant material was observed in the maritime, high-latitude environment of Asendorf (Germany). As already outlined by Schaffasz et al. (2019), this shows that other factors besides minimum temperatures also play a role. For locations Asendorf and Poel, lack of radiation, high moisture and suboptimal daily temperatures of frequently <20°C implied additional stressors, while at Texcoco radiation was not limiting and the daily temperatures were also more favorable. Nevertheless, despite these climatic differences, satisfactory heritability estimates under stress were observed for all traits (h^2 = 0.72 for PHI, 0.68 for SN and 0.60 for SY). Indeed, Schaffasz et al. (2019) reported even higher heritability estimates ($h^2 > 0.80$ for PHI, SN, SY) for F_1-hybrids across more or less the same set of environments as those used in the present work.

Panicle harvest index showed the highest heritability and proved to be the most suitable and reliable trait for scoring reproductive cold tolerance in sorghum, representing an efficient proxy for spikelet fertility by reducing the effects of different panicle sizes (i.e., spikelet numbers) considering panicle raw weight. In contrast to our expectations, the results showed that PHI was not influenced by race, which in sorghum is expressed in terms of contrasting panicle architecture. However, the use of PHI as a proxy for spikelet fertility might at least theoretically be distorted to a degree by compensating effects of increased TKW in panicles with poor seed set and, hence, less sink competition. The lack of negative correlation between TKW and SN for the stress environments seems to support this hypothesis, in contrast to the control environments, where a weak negative correlation between these traits was observed as expected (**Supplementary Table 3**). However, comparing the correlation between PHI and TKW (r = 0.321***) vs. PHI and SN (r = 0.780***) reveals that the influence of SN on PHI is much higher. In a study on sorghum heat stress, which can reduce seed set in the same way as cold stress, Singh et al. (2015) also found that the effect of stress on seed-set percentage was much higher than on TKW, excluding relevant compensation effects of an increased seed mass. Furthermore, there are reports suggesting a determination of maximum seed size of sorghum prior to anthesis based on meristem size (Yang et al., 2009).

In consequence, SN itself represents also an efficient score for reproductive cold tolerance. While it is not affected by possible fluctuations of TKW, it is obviously strongly determined by spikelet number, so that for actual spikelet fertility, it seems a less accurate score than PHI. Though, from a breeder's point of view, a high spikelet number is desirable, reflecting satisfying yield potential, so that SN might be preferred over PHI for selection. Altogether, both PHI and SN showed a very high correlation to SY (r = 0.880*** and 0.914***, respectively) in stress environments. Since SY itself had a somewhat lower heritability, PHI and SN may represent preferable surrogate traits for selection.

Differences in flowering time can potentially complicate comparisons of reproductive cold tolerance levels across divergent genotypes, since they can lead to possible escape of earlier or later flowering genotypes from brief cold-weather conditions at the most critical growth stages. In this regard, tropical highland locations like Texcoco are particularly suitable selection environments due to their consistently low and generally stable minimum temperatures during reproductive phases, whereas temperatures tend to be more fluctuating in temperate areas. However, also for Asendorf and Poel, only weak effects of flowering time were observed on stress symptoms. We observed that later flowering entries tended to have a lower TKW in these environments, due to their shorter grain filling period. This also implies a minor disadvantage in terms of PHI, whereas SN was not influenced by flowering time.

The observed positive correlations between plant height and SY traits are in line with earlier reports (e.g., Jordan et al., 2003; George-Jaeggli et al., 2011). Except for TKW, which was probably distorted by the aforementioned compensation effects for reduced spikelet fertility under stress, the correlations with plant height were somewhat higher for the group of stress environments. The increased availability of stem reserves in taller sorghum, which is especially beneficial under stress (Blum et al., 1997), is believed to be the main reason for this phenomenon.

Tolerance Sources

We identified several inbred lines with superior and environmentally stable reproductive cold tolerance. The diversity panel consisted of both publicly available material (sorghum conversion lines and other public accessions) and breeding lines developed under temperate conditions of Central Europe. As expected due to their breeding history with selection in temperate environments, the latter group performed significantly better under stress conditions. However, some public accessions also ranked among the top 5% for the stress-related traits we evaluated. These superior public accessions were of diverse origin, spanning accessions from temperate (Russia, Turkey, United States) to tropical (Ethiopia, Nigeria, Sudan) countries. While the Ethiopian highland is a known diversity

center for cold tolerance (Singh, 1985), the identification of high reproductive cold tolerance in Nigerian sorghum materials may seem surprising.

However, the fact that tolerance sources from both groups (public accessions and private breeding lines) were found among genetically diverse branches over the entire phylogram (**Figure 3**) suggests a polyphyletic origin of reproductive cold tolerance.

Detection of Candidate Genes From Association Mapping and Overlapping QTL

The results of our study indicate that reproductive cold tolerance in sorghum is a quantitative and complex trait. This finding is consistent with the description of reproductive cold tolerance in sorghum as a heterotic trait (Schaffasz et al., 2019) and also in line with studies on the reproductive cold tolerance of rice (Jiang et al., 2010; Mori et al., 2011). Since rice diverged from maize and sorghum 50–70 Mya, it has been used as a model for sorghum using a comparative genomic approach (Roulin et al., 2009). The candidate genes discussed hereinafter have been well studied in rice, maize and other crops, indicating that a similar function in sorghum can be expected.

To detect regions associated with reproductive cold tolerance traits, we applied a combination of GWAS and comparative combined-QTL analysis, utilizing a published Sorghum QTL Atlas (Mace et al., 2019) as a reference. Combined-QTL analysis including results from different environments and populations yields a better overview of the genetic control of a trait than any single study (Rong et al., 2007). Since our study comprised environments with and without cold stress, separate analyses were performed to compare QTL detected only in cold stress environments with those detected in all (stress and control) environments. Considering cold stress environments alone, the QTL Atlas (Mace et al., 2019) revealed QTL related to abiotic stress tolerance (**Supplementary Table 7** and **Supplementary Figure 4**) for traits such as emergence rate, seedling vigor and survival that overlapped with our identified associations. However, when focussing on only the associations consistent between the traits PHI and SY, identified in the present study as the two trait complexes most impacted by reproductive cold tolerance, we found overlaps with known QTL for dry matter growth rate, chlorophyll content and chlorophyll fluorescence. Obviously the relatively low resolution of the QTL Atlas means these findings are preliminary, however, they provide a starting point for future functional studies to dissect the genetic and physiological basis of the identified QTL.

As expected due to the high phenotypic trait correlations, GWAS revealed multiple overlapping loci among the considered SY traits. This strengthens our confidence of associations not being false positives (Shen and Carlborg, 2013; Boyles et al., 2016). While some common associations between SY traits and DTF or plant height were also observed, in line with the significant phenotypic correlations among these traits, the majority of associations for reproductive cold tolerance were not linked to a taller stature or earlier flowering. Hence, the breeding of short-statured grain sorghum with enhanced cold tolerance and maturity adapted to the temperate target environment appears to be feasible using the stress tolerance sources we identified. Presence of SNP marker highly associated to cold tolerance within the gene *Sobic.010G067300*, a leucine-rich repeat receptor-like protein kinase further enhances our hypothesis of this genomic region being involved in reproductive cold tolerance. A novel allele of the gene was found to be enriched in breeding lines selected for adaption to temperate climates of central Europe. Based on the results of this study, it can be speculated that the haploblock region around Sb10-5394955 may be involved in temperature adaptation of *S. bicolor* and could be of interest for future breeding efforts.

CONCLUSION

Reproductive cold tolerance is a crucial trait to expand sorghum production into both high-latitude temperate areas and tropical high-altitude environments. The present study is the first to analyze this trait in a broad diversity set using data from multi-location field trials including tropical high-altitude and temperate environments. Satisfying heritability estimates and novel, temperate-adapted tolerance sources of polyphyletic origin suggest that robust breeding progress to improve reproductive cold tolerance is feasible.

More detailed genetic dissection of reproductive cold tolerance related traits can help to understand the physiological control of the trait and contribute to more targeted selection. In this study, several significant marker-trait associations were identified. Combining local LD analysis with QTL data from previous studies and synteny to potential candidate genes in rice and maize helped narrow down two interesting marker-trait associations to specific genomic regions involved in cold stress response. One of the promising genomic regions include hotspot on Sb10 which contain a functionally-relevant candidate gene involved in developmental and survival processes under abiotic stress conditions. Future studies to dissect the impact of allelic variants of this gene under contrasting environments and conditions can provide more precise information, functional markers and potential gene editing targets for applied breeding.

DATA AVAILABILITY STATEMENT

The raw data generated during and/or analyzed during current study have been deposited in NCBI database (https://www.ncbi.nlm.nih.gov/) under the project PRJNA775860.

AUTHOR CONTRIBUTIONS

SC conducted data analysis, interpreted the results, and wrote the manuscript. NK conducted data curation and contributed to data analysis. AS planned and oversaw field trials and data collection. RS received the funding and edited the manuscript. BW received the funding and contributed to devise the study. SW devised the study, interpreted the results, and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Fachagentur Nachwachsende Rohstoffe e. V., Germany (FNR) grants 22008716 and 22023515.

ACKNOWLEDGMENTS

For excellent field trial management, we thank Bärbel Frenz/Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ), Dörte Schweneker/Deutsche Saatveredelung (DSV), Mario Tolksdorf/JLU Experimental Station Gross-Gerau, and Manuel Velazquez Almaraz, Mexico. We also acknowledge USDA-ARS for providing the public germplasm.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.772177/full#supplementary-material

REFERENCES

Aulchenko, Y. S., Ripke, S., Isaacs, A., and Van Duijn, C. M. (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23, 1294–1296. doi: 10.1093/bioinformatics/btm108

Blum, A., Golan, G., Mayer, J., and Sinmena, B. (1997). The effect of dwarfing genes on sorghum grain filling from remobilized stem reserves, under stress. *Field Crops Res.* 52, 43–54. doi: 10.1016/S0378-4290(96)03462-4

Boyles, R. E., Cooper, E. A., Myers, M. T., Brenton, Z., Rauh, B. L., Morris, G. P., et al. (2016). Genome-wide association studies of grain yield components in diverse Sorghum germplasm. *Plant Genome* 9. doi: 10.3835/plantgenome2015.09.0091

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Brooking, I. R. (1976). Male sterility in *Sorghum bicolor* (L.) moench induced by low night temperature. i. timing of the stage of sensitivity. *Funct. Plant Biol.* 3, 589–596. doi: 10.1071/pp9760589

Brooking, I. R. (1979). Male sterility in *Sorghum bicolor* (L.) moench induced by low night temperature. ii. genotypic differences in sensitivity. *Funct. Plant Biol.* 6, 143–147. doi: 10.1071/pp9790143

Camacho, C., and Madden, T. (2013). *BLAST+ Release Notes. BLAST® Help [Internet]*. Available online at: https://www.ncbi.nlm.nih.gov/books/NBK131777/#Blast_ReleaseNotes.BLAST_2_2_31_May_18_2 (accessed April 1, 2021).

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.* 17:12. doi: 10.1186/s12870-016-0966-2

Cisneros-López, M. E., Mendoza-Onofre, L. E., Zavaleta-Mancera, H. A., González-Hernández, V. A., Mora-Aguilera, G., órdova-Téllez, L. C., et al. (2010). Pollen–Pistil interaction, pistil histology and seed production in A× B grain sorghum crosses under chilling field temperatures. *J. Agric. Sci.* 148, 73–82. doi: 10.1017/s0021859609990396

Cisneros-López, M. E., Mendoza-Onofre, L. E., Zavaleta-Mancera, H. A., González-Hernández, V. A., Córdova-Téllez, L., Hernández-Martínez, M., et al. (2009). Floral traits and seed production of sorghum A-, B-and R-lines under chilling field temperatures. *J. Agron. Crop Sci.* 195, 464–471.

Dong, S.-S., He, W.-M., Ji, J.-J., Zhang, C., Guo, Y., and Yang, T.-L. (2021). LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Brief. Bioinformatics* 22:bbaa227. doi: 10.1093/bib/bbaa227

Downes, R. W., and Marshall, D. R. (1971). Low temperature induced male sterility in *Sorghum bicolor*. *Aust. J. Exp. Agric.* 11, 352–356. doi: 10.1071/ea9710352

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., et al. (2002). The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229. doi: 10.1126/science.1069424

Gabur, I., Chawla, H. S., Lopisso, D. T., von Tiedemann, A., Snowdon, R. J., and Obermeier, C. (2020). Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. *Sci. Rep.* 10:4131. doi: 10.1038/s41598-020-61228-3

Galtier, N., Gouy, M., and Gautier, C. (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Bioinformatics* 12, 543–548. doi: 10.1093/bioinformatics/12.6.543

Gao, L.-L., and Xue, H.-W. (2012). Global analysis of expression profiles of rice receptor-like kinase genes. *Mol. Plant* 5, 143–153. doi: 10.1093/mp/ssr062

George-Jaeggli, B., Jordan, D. R., van Oosterom, E. J., and Hammer, G. L. (2011). Decrease in sorghum grain yield due to the Dw3 dwarfing gene is caused by reduction in shoot biomass. *Field Crops Res.* 124, 231–239. doi: 10.1016/j.fcr.2011.07.005

Huson, D. H., Richter, D. C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8:460. doi: 10.1186/1471-2105-8-460

Jiang, W., Lee, J., Chu, S.-H., Ham, T.-H., Woo, M.-O., Cho, Y.-I., et al. (2010). Genotype×environment Interactions for chilling tolerance of rice recombinant inbred lines under different low temperature environments. *Field Crops Res.* 117, 226–236. doi: 10.1016/j.fcr.2010.03.007

Jordan, D., Tao, Y., Godwin, I., Henzell, R., Cooper, M., and McIntyre, C. (2003). Prediction of hybrid performance in grain sorghum using RFLP markers. *Theor. Appl. Genet.* 106, 559–567. doi: 10.1007/s00122-002-1144-5

Krishnamurthy, L., Dinakaran, E., Kumar, A. A., and Reddy, B. V. S. (2014). Field technique and traits to assess reproductive stage cold tolerance in Sorghum (*Sorghum bicolor* (L.) Moench). *Plant Prod. Sci.* 17, 218–227. doi: 10.1626/pps.17.218

León-Velasco, H., Mendoza-Onofre, L. E., Castillo-González, F., Cervantes-Santana, T., and Martínez-Garza, Á (2009). Evaluation of two generations of cold tolerant sorghum hybrids and parental lines. II: combining ability, heterosis and heterobeltiosis. *Agrociencia* 43, 609–623.

Mace, E. S., Innes, D., Hunt, C. H., Wang, X., Tao, Y., Baxter, J., et al. (2019). The sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. *Theor. Appl. Genet.* 132, 751–766. doi: 10.1007/s00122-018-3212-5

Maulana, F., Weerasooriya, D., and Tesso, T. (2017). Sorghum landrace collections from cooler regions of the world exhibit magnificent genetic differentiation and early season cold tolerance. *Front. Plant Sci.* 8:756. doi: 10.3389/fpls.2017.00756

Mendoza-Onofre, L. E. (1988). Formation of sorghum hybrids for grain. II. Behavior per se of lines and their general combining ability. *Revis. Fitotec. Mex* 11, 39–47.

Mori, M., Onishi, K., Tokizono, Y., Shinada, H., Yoshimura, T., Numao, Y., et al. (2011). Detection of a novel quantitative trait locus for cold tolerance at the booting stage derived from a tropical japonica rice variety Silewah. *Breed. Sci.* 61, 61–68. doi: 10.1270/jsbbs.61.61

Osuna-Ortega, J., Mendoza-Castillo, M. D. C., and Mendoza-Onofre, L. (2003). Sorghum cold tolerance, Pollen Production. *Maydica* 48, 125–132. doi: 10.1038/s41598-021-87450-1

Osuna-Ortega, J., Mendoza-Onofre, L. E., González-Hernández, V. A., Castillo-González, F., del Mendoza-Castillo, M. C., and Williams-Alanís, H. (2000). Potential of cold tolerant germplasm in the adaptation and adaptability of sorghum in México: I. High valleys. *Agrociencia* 34, 561–572.

Osuna-Ortega, J., Onofre, L. E., González, F. C., Hernández, V. A., Castillo, M. D., Alanís, H. W., et al. (2001). Potencial del germoplasma tolerante Al Frío en la adaptación y adaptabilidad del sorgo granífero en México: II. Río Bravo, Tamaulipas; y Celaya, Guanajuato. *Agrociencia* 35, 625–636.

Parra-Londono, S., Fiedler, K., Kavka, M., Samans, B., Wieckhorst, S., Zacharias, A., et al. (2018). Genetic dissection of early-season cold tolerance in sorghum: genome-wide association studies for seedling emergence and survival under field and controlled environment conditions. *Theor. Appl. Genet*. 131, 581–595. doi: 10.1007/s00122-017-3021-2

Piepho, H.-P., and Möhring, J. (2007). Computing heritability and selection response from unbalanced plant breeding trials. *Genetics* 177, 1881–1888. doi: 10.1534/genetics.107.074229

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Rong, J., Feltus, F. A., Waghmare, V. N., Pierce, G. J., Chee, P. W., Draye, X., et al. (2007). Meta-Analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. *Genetics* 176, 2577–2588. doi: 10.1534/genetics.107.074518

Roulin, A., Piegu, B., Fortune, P. M., Sabot, F., D'Hont, A., Manicacci, D., et al. (2009). Whole genome surveys of rice, maize and sorghum reveal multiple horizontal transfers of the LTR-Retrotransposon Route66 in poaceae. *BMC Evol. Biol.* 9:58. doi: 10.1186/1471-2148-9-58

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Schaffasz, A., Windpassinger, S., Snowdon, R., and Wittkop, B. (2019). Reproductive cold stress tolerance in sorghum F1 hybrids is a heterotic trait. *Agronomy* 9:508. doi: 10.3390/agronomy9090508

Shen, X., and Carlborg, Ö. (2013). Beware of risk for increased false positive rates in genome-wide association studies for phenotypic variability. *Front. Genet.* 4:93. doi: 10.3389/fgene.2013.00093

Singh, S. P. (1985). Sources of cold tolerance in grain sorghum. *Can. J. Plant Sci.* 65, 251–257. doi: 10.4141/cjps85-037

Singh, V., Nguyen, C. T., van Oosterom, E. J., Chapman, S. C., Jordan, D. R., and Hammer, G. L. (2015). Sorghum genotypes differ in high temperature responses for seed set. *Field Crops Res.* 171, 32–40. doi: 10.1016/j.fcr.2014.11.003

Stephens, J. C., Miller, F. R., and Rosenow, D. T. (1967). Conversion of alien sorghums to early combine genotypes 1. *Crop Sci.* 7:396. doi: 10.2135/cropsci1967.0011183x000700040036x

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

Tao, Y., Luo, H., Xu, J., Cruickshank, A., Zhao, X., Teng, F., et al. (2021). Extensive variation within the pan-genome of cultivated and wild sorghum. *Nat. Plants* 7, 766–773. doi: 10.1038/s41477-021-00925-x

Windpassinger, S., Friedt, W., Frauen, M., Snowdon, R., and Wittkop, B. (2015). Designing adapted sorghum silage types with an enhanced energy density for biogas generation in temperate Europe. *Biomass Bioenergy* 81, 496–504. doi: 10.1016/j.biombioe.2015.08.005

Yang, L., Wu, K., Gao, P., Liu, X., Li, G., and Wu, Z. (2014). GsLRPK, a novel cold-activated leucine-rich repeat receptor-like protein kinase from glycine Soja, is a positive regulator to cold stress tolerance. *Plant Sci.* 215–216, 19–28. doi: 10.1016/j.plantsci.2013.10.009

Yang, Z., van Oosterom, E. J., Jordan, D. R., and Hammer, G. L. (2009). Pre-anthesis ovary development determines genotypic differences in potential kernel weight in sorghum. *J. Exp. Bot.* 60, 1399–1408. doi: 10.1093/jxb/erp019

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

5 Discussion

5.1 Genomics assisted preservation and improvement of genebank material

The establishment and preservation of ex-situ plant genetic resources (PGRs) were influenced by erosion of genetic diversity in crop plants as systematic breeding programs involving elite cultivars surpassed traditional landraces (Peres 2016). Genebanks were designed to not just conserve the material but also facilitate the incorporation of novel alleles from landraces into elite breeding programs. However, this has only been successfully achieved in limited cases, for example the genes of the Green Revolution (Hedden 2003) or the *mlo* alleles in barley (Jørgensen 1992). In 1967, Krull and Borlaug stated that "the problem at present is less a lack of genetic variation but rather of efficiency in identifying and incorporating it" (Pistorius 1997). As pointed out by Li et al. (2018), two major limitations are preventing appropriate utilization of genebank genetic resources for breeding programs: 1) time and available resources for thorough characterization of accessions at a large scale; and 2) identifying and introducing the allelic variance into elite breeding materials.

The traditional method of assigning identities to genebank material is based on passport data and botanic information associated with phenotypes. However, with the ongoing development of inexpensive, high-throughput genome profiling technologies, large-scale genomic characterization of gene bank collections is within reach (Yu et al. 2016) and has already been shown in maize (Romay et al. 2013) and barley (Milner et al. 2019). In this thesis, a reduced-representation sequencing to identify previously unknown racial classification and genetic diversity of the Ugandan sorghum germplasm was conducted. This genotypic information can now be used as molecular passport data where it complements, corrects, and corroborates existing traditional records. As pointed out by (McFerson et al. 1996; Debouck 2003) the more information (both morphological and genotypic) available for the materials in a germplasm collection, the more beneficial it is for farmers and/or researchers. This vast amount of genetic data would also enable the genebanks to identify genetically diverse accessions within their germplasm collection.

Genotypic information related to accessions is not only crucial in context of genebank management, but also has multiple purposes. The genomic data can be utilized to address the issue of duplicated accessions (van Hintum and Visser 1995) which not only increases the workload of genebank managers but also misguides end users. Genotypic characterization can further be extended to create a complete and accurate genotype-to-phenotype map (Alberch 1991) for individuals stored in the genebank. This map not only promotes the understanding of the predicted quantitative values but also improves our knowledge about mechanisms governing plant performance at a molecular level (Davidson 2010).

Apart from germplasm characterization and connecting them to phenotypes, management and utilization of resources are also key factors for breeding success. However, large germplasms are difficult to maintain, assess and utilize (Holden 1984). This issue could be resolved by establishing

a core set for identification and utilization of novel variation in genetic resources (Hodgkin et al. 1995). Core sets are representative collections, constructed by using phenotypic (Rincent et al. 2018), genomic (Wambugu et al. 2018) and geographical information of accessions for certain crop traits. Combining genetic knowledge as presented in this thesis with detailed phenotype data will assist the generation of representative core collections that span the entire diversity of a specific collection (Figure 1). These selected accessions can be refined through consecutive phenotyping or calculating genome-wide prediction of phenotypes (Mascher et al. 2019) and be further trained as prediction models to infer phenotypes of the remaining or independent population (Yu et al. 2016) in a cost-effective and accelerated manner. Genotypes with the highest breeding values can then enter pre-breeding programs. Whole-genome sequencing of these representative lines can further aid association studies and in best cases identification of candidate genes (Yano et al. 2016). Various alternative types of collections to boost diversity capture or utilization include mini core sets (Upadhyaya and Ortiz 2001), nested core collections (McKhann et al. 2004) and composite collections (Furman 2006).

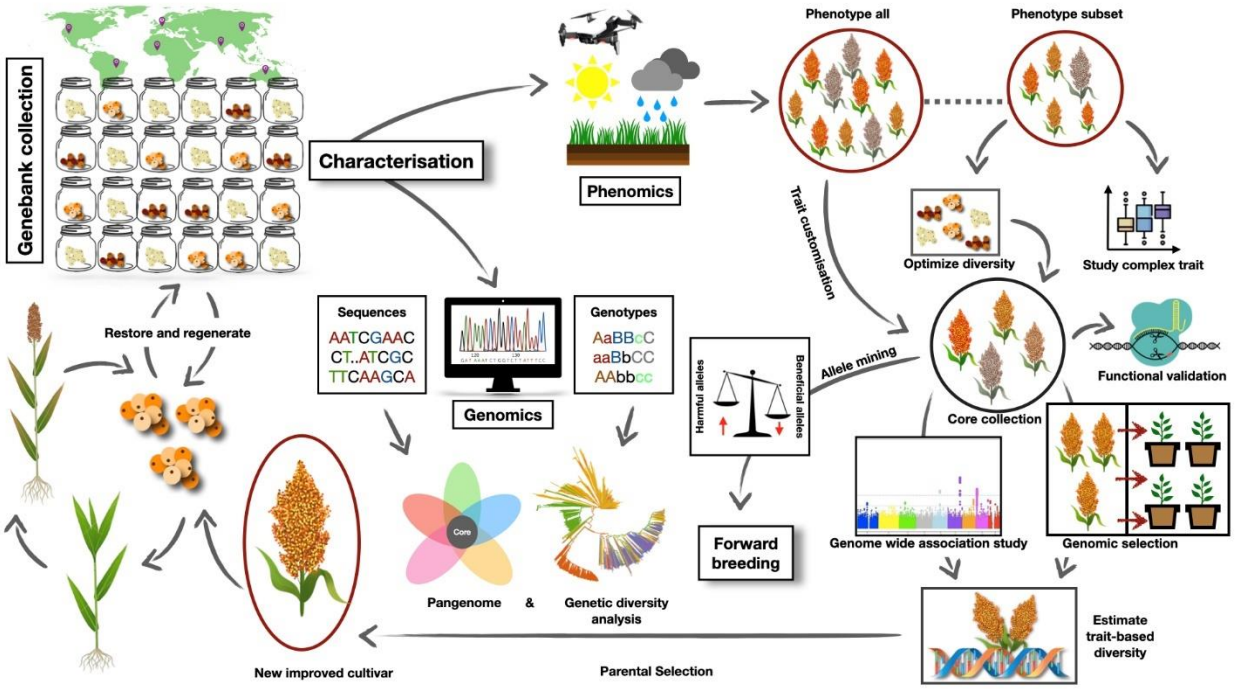

Figure 1: Plant genetic resources archived in genebanks serve as indispensable source of genetic variation. In-depth genomic characterization combined with new age phenomic approaches will not only facilitate genebank material management but also better utilization and conservation of genetic diversity. New genes/haplotypes discovered from analyzing genomic data and precise phenotyping data will open avenues for enhanced crop improvement. Developing a representative core collection, which can be phenotyped and used for genome-wide prediction of the entire

population. Genotypes with the highest breeding values can be selected to enter pre-breeding programs aimed towards development of improved cultivars.

Another efficient method to explore plant agro-biodiversity, especially in the context of climate change, is the "focused identification of germplasm strategy" (FIGS). Unlike a core collection, which is designed to work with fewer accessions that represents, "with a minimum of repetitiveness, the genetic diversity of a crop species and its relatives" (Frankel 1984), FIGS is designed to improve the efficiency with which specific adaptive traits are identified from genetic resource collections. It focuses on the concept that crops are likely to evolve under environmental selection pressure and adapt as a response to changing climate. The FIGS approach has been successfully used to identify several core sets for traits such as drought tolerance in faba bean (Khazaei et al. 2013) or stem rust and powdery mildew in wheat (Bhullar et al. 2009; Endresen et al. 2012). Haupt and Schmid (2020) reported the use of combined approaches to enhance identification of targeted accessions in soybean by using a combination of FIGS approach and SNP genotypic markers.

According to McCouch et al. (2013), to overcome the food shortages in the predicted future, the biodiversity in seed banks needs to be mined using a three-step process. (i) 'Fingerprinting' all samples in a genebank by sequencing the genomes. (ii) Analyze the phenotypes of genebank accessions for trait and performance evaluation, and finally, (iii) create an internationally accessible informatics platform cataloging the world's seed collection diversity. Although the use of genomics to study genebank accessions has significantly advanced in recent years (Yu et al. 2016), there is a considerable lag in the case of genebank phenomics. It is not only intellectually challenging but also overly complex, costly, and time-consuming (McCouch 2013; Philipp et al. 2018).

Current High-throughput phenotyping (HTP) tools using sensors and imagers is a promising, efficient, and cost-effective approach to collect phenotypic data from large-scale trials for multiple traits. These phenotyping tools have already been applied successfully to study different traits such as yield prediction in wheat (Pantazi et al. 2016), plant height and biomass of rice (Bendig et al. 2015a) and barley (Bendig et al. 2015b), crop biomass of field peas (Nguyen et al. 2018), seed characteristics of lentils (Lemasurier et al. 2014), and prediction of seed longevity in oilseed rape from chemical compositions (Nagel et al. 2018). The application of low-cost HTP methods to assess huge genetic resources will significantly improve the gains of pre-breeding or breeding programs with marginal extra expenses. A strategic phenomics approach has already been deployed at a few of the genebanks around the world, like the Australian grains genebank (AGG), Horsham, Victoria. HTP platforms such as automated phenotyping of Plant Phenomics Victoria, Horsham (Nguyen et al. 2019) is currently being applied at AGG to capture various important traits from seed regeneration cycles.

5.2 Molecular breeding for enhanced sorghum adaptability to a temperate climate

Currently, molecular breeding is an indispensable tool for the improvement of many crop species. However, in sorghum, recently heralded as an indispensable crop in face of climate change, progress in this field has been slow and limited as compared to major cereal crops like wheat, rice, etc. Despite its potential, the major factors limiting the adaptation of sorghum as a cereal or biofuel crop in temperate zones are its photoperiod sensitivity, as well as strong susceptibility to low temperatures at sowing and flowering (Schaffasz et al. 2019a; Chakrabarty et al. 2021a, 2022).

Sorghum is a short-day plant with a strong photoperiod response. The simple genetic architecture of photoperiodism in sorghum is controlled by four major maturity loci i.e., Ma1, Ma2, Ma3, and Ma4 (Quinby and Karper 1945; Quinby 1966). In 1963 the 'Sorghum Conversion Program' was initiated by the US Department of Agriculture (USDA) to expand the genetic diversity for temperate sorghum breeding programs. Stephens et al. (1967) implemented backcrossing to convert genetically diverse tropical accessions to early-maturing lines which turned out to be an enormous success and was later incorporated into most of the current sorghum hybrids for temperate zones. According to Cuevas et al. (2016), a cross between tropical and temperate sorghums [*S. propinquum* (Kunth) Hitchc. ×*S. bicolor* (L.) Moench], revealed the FlrAvgD1 quantitative trait locus (QTL) which accounts for 85.7% of the variation in flowering time under long days. Over the years there have been several studies regarding the floral transition in sorghum (Wolabu et al. 2016; Wolabu and Tadege 2016; Abdul-Awal et al. 2020). Hence, the use of sorghum for grain or forage and feedstock production depends on the selection of photoperiod insensitive mutants or photoperiod sensitive wild genotypes respectively.

Early juvenile development (Maulana et al. 2017) and the pre-flowering reproductive stage (Brooking 1976) are considered the most critical growth stages. Sensitivity to cold stress at early juvenile stages poses serious threats for successful adaptation. To avoid cold nights, the sowing is delayed by around 3-4 weeks as compared to maize limiting its potential growth period. However, improving chilling tolerance at the juvenile stage would also be beneficial for certain subtropical regions where sorghum is already well established. Early sowing in the spring season can potentially increase utilization of winter moisture giving better yields (Patane et al. 2006).

With increasing research efforts, the dissection of the genetic architecture of juvenile chilling tolerance in sorghum by QTL studies in segregating bi-parental populations has been conducted. For example, Knoll et al. (2008) used a recombinant inbred line (RIL) population, cross between Shan Qui Red (SQR, cold-tolerant) and SRN39 (cold-sensitive) to dissect early-season cold tolerance in sorghum. A nested association mapping (NAM) approach was used by Marla et al. (2019) to investigate chilling tolerance in a multi-parental population. The results indicate strong selection for important recessive adaptation alleles in modern grain sorghum ensued with the loss of early-season chilling tolerance by linkage in repulsion. Various genetic association studies using diversity panels have also enhanced the identification of genome regions and promising candidate genes having a strong influence on early-stage low-temperature tolerance Cold-Shock Domain Protein 1 (CSDP1) has been described as a potential functional candidate gene underlining these

QTLs (Fiedler et al. 2012, 2014, 2016). Recent studies analyzing both agronomical and cold tolerance traits were conducted under Central European field conditions (Schaffasz et al. 2019a; Chakrabarty et al. 2022) revealing valuable sources for chilling tolerance and early vigor.

The pre-flowering reproductive stage is the second sensitive and crucial developmental phase in sorghum affected by low temperatures (Singh 1985). Even though effects of juvenile chilling stress can be reduced by later sowing dates (albeit at the expense of yield potential), there is no escape strategy for cold nights during the critical reproductive stage in summer. Remarkable breeding efforts have led to the development of sorghum varieties with improved reproductive chilling tolerance for the Mexican high valleys (Mendoza-Onofre 1988; Osuna-Ortega et al. 2000, 2003; León-Velasco et al. 2009; Cisneros-López et al. 2010). Unlike chilling tolerance at juvenile stage, not much is known about the genetic architecture of tolerance to low temperature at reproductive stage in sorghum. Recently, genetic dissection and enhancement of this trait have been reported (Schaffasz et al. 2019b; Chakrabarty et al. 2021a). Both Singh (1985) and Schaffasz et al. (2019b) described a dominant inheritance for this trait.

The availability of cost-effective molecular markers in recent times makes it easier to evaluate the phylogenetic and genomic structure of accessions which can be used to establish genetically distinct heterotic pools for temperate sorghum breeding programs. Schaffasz et al. (2019b) observed that correlations between *per se* and hybrid performance were higher in case of reproductive cold tolerance as compared to juvenile chilling tolerance. Upcoming association studies and GS approaches for chilling tolerance should focus rather on general combining ability (GCA) than per se performance (as done in the past) for successful hybrid breeding towards improved plant establishment, pollen fertility and seed set under cool-temperature conditions (Chakrabarty et al. 2021b).

5.3 Genomics and bioinformatic approaches in sorghum research

The most crucial aspect of genomics application for genetic analysis or breeding of quantitative traits is the availability and utilization of genetic markers and bioinformatic platforms for efficient, cost-effective genome-wide screening. Unlike other major crops where molecular screening is generally based on dedicated SNP arrays, in sorghum, population genomics studies have been mostly achieved with sequencing-based marker techniques. Significantly less bioinformatics analysis is required in the case of SNP arrays as compared to the derivation of information from GBS data (Chakrabarty et al. 2021b). In case of sorghum, until recently there has not been a significant effort in developing a community-driven public SNP genotyping array platform. A small-scale Illumina Infinium 3K SNP array was developed by Bekele et al. (2013) which was also implemented for building genetic maps, diversity analyses, and GWAS (Bekele et al. 2014).

In recent years, there has been a boom in NGS technologies and their utilization in identifying the underlying genetic basis of agronomically important traits. Two types of sequence-based genotyping platforms involving restriction site-associated DNA sequencing are: GBS (Morris et

al. 2013) and RAD-seq (Nelson et al. 2011). Over the years, the GBS panels have been used in numerous studies to dissect traits like cold and heat stress (Chopra et al. 2017; Ortiz et al. 2017), grain quality traits (Boyles et al. 2017) and grain mold resistance (Cuevas et al. 2019). Studies have reported a prominent level of genetic differentiation in local germplasm populations. A diverse accession set of 1425 Ethiopian landraces indicated the presence of 11 subgroups with high genetic diversity which was shaped due to local agro-climatic conditions (Girma et al. 2019). A diversity panel of 607 Nigerian accessions was compared to 1785 geo-referenced global accessions to study genetic diversity. The results showed the clear distinction of the former in comparison to the global set and it clustered with West African accessions (Olatoye et al. 2018). Recently, Kajiya-Kanegae et al. (2020) created high-density linkage maps using RAD-seq to study different physiological traits in Takakibi, a Japanese landrace.

Although sorghum is a common bioenergy and cereal crop, the depth of research in comparison to other grain crops such as maize, rice and wheat is far more limited. Recent studies have facilitated the inclusion of sorghum information in databases like Ensembl Plants (https://plants.ensembl.org), Gramene (Youens-Clark et al. 2011), PlantsDB (Nussbaumer et al. 2012), Phytozome (Goodstein et al. 2012), National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/), PLAZA (Proost et al. 2015) and SorghumBase (Gladman et al. 2022). Makita et al. (2015) created a sorghum transcriptome database, MOROKOSHI, which integrated functional annotations and specific RNA-seq data to construct a co-expression network to study in-depth expression profile variations. Like popular agricultural databases TIGR, MaizeGDB and SIFGD, SorghumFDB was developed as a comprehensive platform for the genome functional annotation of sorghum (Tian et al. 2016).

Advances in different fields of crop genomics would not have been possible without the help of bioinformatics providing new tools, comprehensive databases, and the other resources required to analyze the ever-increasing amounts of data. Different analysis and workflow platforms like Galaxy (https://usegalaxy.org/) and de.NBI (https://www.denbi.de/cloud) are now available for users with limited or low bioinformatic knowledge. With increasing technology and advancements in the field of bioinformatics, deep learning is of unparalleled value to understand connections between genetic information and molecular processes. Its application in sorghum breeding programs is fairly new with very few reported studies. Recently, traditional machine learning combined with deep learning algorithms has been implemented to detect and count sorghum panicles and seeds (Guo et al. 2018; Ghosal et al. 2019). The systematic management of scientific and research data is equally important as producing it, allowing its accessibility to both national and international researchers and users. Kumari et al. (2021) used the KBase platform, a free, open-source software, and data platform to study transcriptomics of sorghum in response to drought stress. This not only illustrates the power of online data analysis platforms for biologists but also promotes FAIR (Findable, Accessible, Interoperable, and Reusable) principles for scientific data management and usage.

Sorghum remains one of the crucial staple crops for many underdeveloped countries and regions in Africa. In these parts, limitations in resources and infrastructure make it difficult to the adoption

of recent technologies. Publicly funded breeding programs like the Genomic Open-source Breeding informatics initiative (GOBii) (www.gobiiproject.org) and Excellence in Breeding Platform (EiB) (www.excellenceinbreeding.org) are working towards improving breeding strategies for sorghum by enabling routine use of innovative technologies.

5.4 Conclusions

The conservation and utilization of diverse plant genetic resources are critical at the current climatic, eco-geographical scenario to ensure global food security. The future success of global agriculture depends on its ability to adapt to new forms of agriculture and maintain increased productivity to feed a growing world population. Although recently the genome sequencing prices have gone down dramatically, most genebanks, especially in developing or underdeveloped countries do not have the resources to screen and perform in-depth characterization of their collections. The methods presented in this study can be implemented in future research and pre-breeding programs for genomic and population genetic characterization of large, highly-diverse germplasm collections. This knowledge can be implemented in creating a representative collection or a core set which can be further phenotyped and be used as a training set in genomic selection to predict a larger population.

In the face of climate change, sorghum has proved to be a cereal for the future and although it is increasingly gaining ground in temperate climate zones like Europe, it still has a long way to go to become one of the major crops. The plant material, genomic tools and evaluated genetic data presented in this thesis can be used as a base material for breeding programs aiming towards sorghum adaptation in a temperate climate and other cooler regions. The novel, previously uncharacterized material can serve as an exotic cold resistance source that can be incorporated into existing breeding programs. The genetic data would also facilitate marker-assisted breeding programs and when combined with phenotypic data, it can be used as a basis for genomic selection. The complex quantitative nature of cold tolerance traits, especially for big populations, often makes it difficult to study under controlled conditions. This makes the genomic breeding strategy for early generation selection a viable alternative. Screening for other traits, like root architecture can help in decoding genetic control and other stress responses crucial for adaptation (Hatzig et al. 2015; Voss-Fels et al. 2017) screens as a basis for potential genomic prediction of stress responses. A key for linking genotypes to phenotypes is gene expression and understanding its effect on the process of local adaptation. It would be interesting to identify genes whose expression is shaped by selection pressure. Blanc et al. (2021) were able to show how to effectively detect differentially expressed genes influenced by local adaptation. A crucial step in the adaptation of tropical-origin sorghum to temperate climate is cold acclimation. As early as 1970, Weiser (1970) followed by other studies revealed changes at protein levels after cold treatment (Meza-Basso et al. 1986; Guy and Haskell 1987). A transcriptome profile and validation of genes can be created for this material

to construct a putative transcriptional network aiding dissection of abiotic stress impact at the molecular level (Chopra et al. 2015). However, the sheer number of genes showing differential expression patterns under stress makes this process harder. Thus, we can combine the expression patterns and prior molecular data from this study, like association mapping or racial classification to elucidate molecular mechanisms involved in thermal and other abiotic stress responses in cereal crops.

Apart from the more traditional methods to interpret and evaluate underlying molecular mechanisms, the importance of genomic structural differences has come to light in recent times. They are source of genetic variation and are known to play key roles in plant evolution, diversification, and creating phenotyping alteration to adapt to different environmental conditions. Chawla et al. (2021) has elucidated the importance of studying SVs for unraveling novel genetic and genomic diversity associated with important traits in crop plants. A feasible option when an enormous number of genotypes are available can be whole-genome resequencing of selected lines from this thesis based on their performance under cold stress to study existing genomic structural variation (SVs) underlying key phenotypic differences. This may lead to the detection of previously unknown levels of functional diversity and enhance the breeding process and crop adaptation in the future.

In recent years, plant 3D genomics has become a hot topic, and its combined application with multi-omics data has given a sneak peek into the complex interaction of regulatory elements during the plant life cycle. This opens up new doorways in understanding the regulatory mechanism of chromatin conformation on gene expression. According to Pei et al. (2021) understanding the dynamics of the 3D genome on gene transcription regulations stimulating important agronomic traits would lead the way towards further crop improvement. This concept has already been applied to some of the major crop plants like rice (Zhao et al. 2019), maize (Peng et al. 2019), cotton (Wang et al. 2017) and brassicas (Xie et al. 2019).

Recently, Dong et al. (2017) studied 3D chromatin architecture in sorghum revealing interactions between the adjacent loci. This would be particularly useful to study (i) Is different chromatin structure of different cell types influenced under changing environmental conditions like adaptation and/or different developmental stages (ii) To what extent, if at all transcriptional memory alters gene expression in response to changing environment.

Looking forward, it can be expected that the most novel and exciting discoveries in near future will be possible by integrating multiple “omics” tools and “big data” analytics in an integrative framework, in the process of achieving genotype-to-phenotype transformations (Ghosh et al. 2011). Even with the expansion into modern science, it has been proven that crop evolution can potentially improve agriculture (Meyer and Purugganan 2013). It can therefore be expected that studying the molecular footprint of selection and adaptation process in wild crop relatives will enhance bridging the gap between phenotyping and genotyping in genebanks (Dempewolf et al. 2017). Open access and data sharing are crucial to better utilize and characterize the isolated diversity hotspots maintained at the different genebanks (McCouch et al. 2020).

The goal is to establish a climate-smart and resilient cropping system in temperate regions which can sustain the environmental pressure and resolve hunger issues in the near future. To achieve this aim, different methods for designing and assessing climate-smart crops mentioned in the thesis can used as initial steps before adapting and mitigating solutions in the real world at the field and farm level.

6 Summary

Sorghum (*Sorghum bicolor* L. Moench), the 5th most important cereal crop belongs to the Poaceae grass family. According to records, it originated around 7000-5000 BC in the north-eastern part of Africa from where it traveled across the different parts of Africa, the middle east, India, China and eventually Americas, Australia, and finally Europe. As a result of multiple complex genomic interactions and selection, pressure can be broadly subdivided into the botanical classification of *S. halepense*, *S. propinquum*, and *S. bicolor*, and classification of cultivated forms of *S. bicolor* subsp. bicolor into five races, namely bicolor, guinea, caudatum, kafir, and durra. It is mainly grown in the lowlands and semi-arid regions of the tropics and subtropics and has been adapted to various contrasting environments for different end-products leading to an increase in morphological and genetic diversity. For example, in African countries, it is primarily grown as food grain and stalk, and leaves are valued as forage and building material. In the USA, its mainly used for livestock feed and ethanol production, whereas, in China, it is popular for manufacturing traditional alcoholic beverages.

Sorghum's remarkable ability to survive and produce yields under extreme climatic conditions compared to most other grain crops makes it an important 'failsafe' source of food, feed, fiber, and fuel in the global agroecosystem. Future projections regarding changing climate and its negative effect on yield traits, highlight the urgency to harness new genetic resources and the ability of breeding programs to develop the required adaptations promptly. Sorghum has numerous agro-ecological advantages over other crops like maize when it comes to temperate climates. But few of the biggest hurdles for sorghum adaptation to cooler climates are juvenile and reproductive cold stress.

The present study reported a genetic characterization of the diverse, previously uncharacterized *S. bicolor* collection of the Uganda National Genebank, representing different agro-ecological zones of the country. High-resolution genotyping using reduced representation sequencing was used to characterize the material and study population genetics. More than 3000 *S. bicolor* accessions were genotyped using a panel of around 20,000 genome-wide DArTseq SNP markers and co-analyzed with a global sorghum collection genotyped previously with the same panel of markers for genetic diversity analysis and studying different interesting traits. The results revealed the presence of extensive genetic and racial diversity in predominantly admixed accessions and a unique, genetically isolated group of accessions from the southwestern Ugandan highlands, a region which low temperatures which potentially harbors genes of interest for breeding of sorghum in Germany and other temperate climate zones.

A representative core set of the novel Ugandan sorghum germplasm was analyzed to study juvenile cold tolerance. Data was collected from multi-year field trials and controlled climate chamber experiments. Genome-wide association studies were used to identify genomic regions involved in adaptation to cooler climatic conditions that could be of interest for the expansion of sorghum production into temperate latitudes. This thesis can be considered as a case study to illustrate the potential of genebank genomics to screen valuable, underutilized germplasm collections for evaluation of various biological and agro-economical traits and alleles.

While farmers can avoid early-stage cold stress by later sowing (albeit at the expense of maturity and yield potential), there is no escape strategy for reproductive stage cold stress. This trait was analysed in another broad diversity panel consisting of 330 inbred lines of different origin, types of use, and subspecies from multi-location field trials including tropical high-altitude and temperate environments. In this study, several significant marker-trait associations were identified. This was further combined with local LD analysis, previously curated QTL data, and synteny to potential candidate genes in rice and maize to narrow down to interesting marker-trait associations to specific genomic regions involved in cold stress response.

This thesis can be considered as a basis for the selection of accessions for genetic diversity preservation and management, utilization in breeding programs, and establishing genetic relationships with other existing germplasm collections. The results provide important new insights for adaptive crop breeding in the face of climate change and the expansion of sorghum production to different regions. This will facilitate sorghum from being a "plant of the future" to transforming into a real-life major agricultural alternative.

7 Zusammenfassung

Sorghum (*Sorghum bicolor* L. Moench), die fünftwichtigste Getreideart, gehört zur Familie der Poaceae-Gräser. Aufzeichnungen zufolge entstand es etwa 7000-5000 v. Chr. im nordöstlichen Teil Afrikas, von wo aus es über verschiedene Teile Afrikas, den Nahen Osten, Indien, China und schließlich Amerika, Australien Europa verbreitet wurde. Als Ergebnis zahlreicher komplexer genomischer Interaktionen und von Selektion lässt sich der Druck grob in die botanische Klassifizierung von S. halepense, S. propinquum und S. bicolor sowie die Klassifizierung der Kulturformen von S. bicolor subsp. bicolor in fünf Rassen, nämlich bicolor, guinea, caudatum, kafir und durra, unterteilen. Es wird hauptsächlich im Tiefland und in halbtrockenen Regionen der Tropen und Subtropen angebaut und hat sich an verschiedene kontrastreiche Umgebungen für unterschiedliche Endprodukte angepasst, was zu einer Zunahme der morphologischen und genetischen Vielfalt geführt hat. In den afrikanischen Ländern wird es beispielsweise hauptsächlich als Nahrungsmittel angebaut, während die Blätter als Futtermittel und Baumaterial geschätzt werden. In den USA wird sie hauptsächlich als Viehfutter und zur Herstellung von Ethanol verwendet, während es in China für die Herstellung traditioneller alkoholischer Getränke beliebt ist.

Die bemerkenswerte Fähigkeit von Sorghum, unter extremen klimatischen Bedingungen zu überleben und Erträge zu erzielen, macht es im Vergleich zu den meisten anderen Getreidearten zu einer wichtigen "ausfallsicheren" Quelle für Lebensmittel, Futtermittel, Fasern und Brennstoffe im globalen Agrarökosystem. Künftige Prognosen über den Klimawandel und seine negativen Auswirkungen auf die Ertragsmerkmale machen deutlich, dass es dringend notwendig ist, neue genetische Ressourcen zu nutzen und die Fähigkeit von Züchtungsprogrammen zur raschen Entwicklung der erforderlichen Anpassungen zu verbessern. Sorghum hat zahlreiche agrarökologische Vorteile gegenüber anderen Kulturen wie Mais, wenn es um gemäßigte Klimazonen geht. Einige der größten Hürden für die Anpassung von Sorghum an kühlere Klimazonen sind jedoch juveniler und reproduktiver Kältestress.

In der vorliegenden Studie wurde eine genetische Charakterisierung der vielfältigen, bisher uncharakterisierten S. bicolor-Sammlung der Nationalen Genbank Ugandas vorgenommen, die verschiedene agro-ökologische Zonen des Landes repräsentiert. Zur Charakterisierung des Materials und zur Untersuchung der Populationsgenetik wurde eine hochauflösende Genotypisierung mittels Sequenzierung in reduzierter Darstellung durchgeführt. Mehr als 3000 S. bicolor-Akzessionen wurden mit einem Panel von rund 20.000 genomweiten DArTseq-SNP-Markern genotypisiert und gemeinsam mit einer globalen Sorghum-Sammlung analysiert, die zuvor mit demselben Panel von Markern zur Analyse der genetischen Vielfalt und zur Untersuchung verschiedener interessanter Merkmale genotypisiert worden war. Die Ergebnisse zeigten das Vorhandensein einer umfangreichen genetischen und rassischen Vielfalt in überwiegend gemischten Akzessionen und einer einzigartigen, genetisch isolierten Gruppe von Akzessionen aus dem südwestlichen ugandischen Hochland, einer Region mit niedrigen Temperaturen, die potenziell Gene beherbergt, die für die Züchtung von Sorghum in Deutschland und anderen gemäßigten Klimazonen interessant sind.

Ein repräsentativer Kernset des neuen ugandischen Sorghum-Kollektionen wurde analysiert, um die juvenile Kältetoleranz zu untersuchen. Die Daten wurden aus mehrjährigen

Feldversuchen und kontrollierten Klimakammerexperimenten gewonnen. Mit Hilfe von genomweiten Assoziationsstudien wurden genomische Regionen identifiziert, die an der Anpassung an kühlere Klimabedingungen beteiligt sind und für die Ausweitung des Sorghumanbaus in gemäßigten Breiten von Interesse sein könnten. Diese Arbeit kann als Fallstudie betrachtet werden, die das Potenzial der Genbankgenomik für das Screening wertvoller, wenig genutzter Sorghum-Kollektionenssammlungen zur Bewertung verschiedener biologischer und agroökonomischer Merkmale und Allele veranschaulicht.

Während Landwirte Kältestress im Frühstadium durch eine spätere Aussaat vermeiden können (wenn auch auf Kosten der Reife und des Ertragspotenzials), gibt es für Kältestress im in der reproduktiven Phase keine Ausweichstrategie. Dieses Merkmal wurde in einem weiteren breit angelegten Diversitätspanel analysiert, das aus 330 Inzuchtlinien unterschiedlicher Herkunft, Nutzungsarten und Unterarten aus Feldversuchen an mehreren Standorten, darunter in tropischen Höhenlagen und gemäßigten Klimazonen, besteht. In dieser Studie wurden mehrere signifikante Marker-Merkmals-Assoziationen identifiziert. Diese wurden mit lokalen LD-Analysen, zuvor kuratierten QTL-Daten und Syntenie zu potenziellen Kandidatengenen in Reis und Mais kombiniert, um interessante Marker-Merkmals-Assoziationen zu spezifischen genomischen Regionen einzugrenzen, die an der Reaktion auf Kältestress beteiligt sind.

Diese Arbeit kann als Grundlage für die Auswahl von Akzessionen für die Erhaltung und das Management der genetischen Vielfalt, die Nutzung in Züchtungsprogrammen und die Herstellung genetischer Beziehungen zu anderen bestehenden Sorghum-Kollektionensmasammlungen dienen. Die Ergebnisse liefern wichtige neue Erkenntnisse für eine anpassungsfähige Pflanzenzüchtung angesichts des Klimawandels und der Ausweitung der Sorghum-Produktion auf verschiedene Regionen. Dies wird es Sorghum erleichtern, sich von einer "Pflanze der Zukunft" zu einer echten landwirtschaftlichen Alternative zu entwickeln.

8 References

Abdul-Awal SM, Chen J, Xin Z, Harmon FG (2020) A sorghum gigantea mutant attenuates florigen gene expression and delays flowering time. Plant Direct 4:e00281. https://doi.org/https://doi.org/10.1002/pld3.281

Alberch P (1991) From genes to phenotype: dynamical systems and evolvability. Genetica 84:5–11

Allier A, Lehermeier C, Charcosset A, et al (2019) Improving Short- and Long-Term Genetic Gain by Accounting for Within-Family Variance in Optimal Cross-Selection. Frontiers in genetics 10:1006. https://doi.org/10.3389/fgene.2019.01006

Ananda GKS, Myrans H, Norton SL, et al (2020) Wild Sorghum as a Promising Resource for Crop Improvement . Frontiers in Plant Science 11:1108

Anderson JE, Kantar MB, Kono TY, et al (2014) A roadmap for functional structural variants in the soybean genome. G3: Genes, Genomes, Genetics 4:1307–1318

Assar AHA, Uptmoor R, Abdelmula AA, et al (2005) Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop science 45:1636–1644

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

Bekele WA, Fiedler K, Shiringani A, et al (2014) Unravelling the genetic complexity of sorghum seedling development under low-temperature conditions. Plant, cell & environment 37:707–723

Bekele WA, Wieckhorst S, Friedt W, Snowdon RJ (2013) High-throughput genomics in sorghum: from whole-genome resequencing to a SNP screening array. Plant biotechnology journal 11:1112–1125. https://doi.org/10.1111/PBI.12106

Bendig J, Willkomm M, Tilly N, et al (2015a) Very high resolution Crop Surface Models (CSM) from UAV-based stereo images for rice growth monitoring in Northeast China. gisScience - Die Zeitschrift fur Geoinformatik XL-1/W2:1–9. https://doi.org/10.5194/isprsarchives-xl-1-w2-45-2013

Bendig J, Yu K, Aasen H, et al (2015b) Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation 39:79–87. https://doi.org/10.1016/j.jag.2015.02.012

Bernal JH, Ligarreto M GA, Hernández RS (2014) Effects of the genotype and environment interaction on sugar accumulation in sweet sorghum varieties (*Sorghum bicolor*-{L.}-Moench) grown in the lowland tropics of Colombia. Agronomía Colombiana 32:307–314

Bernardino KC, de Menezes CB, de Sousa SM, et al (2021) Association mapping and genomic selection for sorghum adaptation to tropical soils of Brazil in a sorghum multiparental random mating population. Theoretical and Applied Genetics 134:295–312. https://doi.org/10.1007/s00122-020-03697-8

Bhullar NK, Street K, Mackay M, et al (2009) Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the Pm3 resistance locus. Proceedings of the National Academy of Sciences of the United States of America 106:9519–9524. https://doi.org/10.1073/pnas.0904152106

Blanc J, Kremling KAG, Buckler E, Josephs EB (2021) Local adaptation contributes to gene expression divergence in maize. G3 Genes|Genomes|Genetics 11:. https://doi.org/10.1093/g3journal/jkab004

Borrell AK, Hammer GL, Douglas ACL (2000a) Does maintaining green leaf area in sorghum improve yield under drought? I. Leaf growth and senescence. Crop science 40:1026–1037

Borrell AK, Hammer GL, Henzell RG (2000b) Does maintaining green leaf area in sorghum improve yield under drought? II. Dry matter production and yield. Crop science 40:1037–1048

Bowers JE, Abbey C, Anderson S, et al (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

Boyles RE, Pfeiffer BK, Cooper EA, et al (2017) Genetic dissection of sorghum grain quality traits using diverse and segregating populations. Theoretical and applied genetics 130:697–716

Brooking IR (1976) Male sterility in *Sorghum bicolor* (L.) Moench induced by low night temperature. I. Timing of the stage of sensitivity. Functional Plant Biology 3:589–596

Burke MB, Lobell DB, Guarino L (2009) Shifts in African crop climates by 2050, and the implications for crop improvement and genetic resources conservation. Global Environmental Change 19:317–325. https://doi.org/10.1016/j.gloenvcha.2009.04.003

Casas AM, Kononowicz AK, Zehr UB, et al (1993) Transgenic sorghum plants via microprojectile bombardment. Proceedings of the National Academy of Sciences 90:11212–11216

Chakrabarty S, Kravcov N, Schaffasz A, et al (2021a) Genetic Architecture of Novel Sources for Reproductive Cold Tolerance in Sorghum. Frontiers in Plant Science 12:. https://doi.org/10.3389/fpls.2021.772177

Chakrabarty S, Mufumbo R, Windpassinger S, et al (2022) Genetic and genomic diversity in the sorghum gene bank collection of Uganda. BMC Plant Biol 22:378. https://doi.org/10.1186/s12870-022-03770-y

Chakrabarty S, Schaffasz A, Chawla HS, et al (2021b) Improving abiotic stress tolerance to adapt sorghum to temperate climatic regions. In: Molecular breeding in wheat, maize and sorghum: strategies for improving abiotic stress tolerance and yield. CABI, Wallingford, pp 444–462

Chawla HS, Lee H, Gabur I, et al (2021) Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant. Plant Biotechnology Journal 19:240–250. https://doi.org/10.1111/pbi.13456

Chloupek O, Hrstkova P, Schweigert P (2004) Yield and its stability, crop diversity, adaptability and response to climate change, weather and fertilisation over 75 years in the Czech Republic in comparison to some European countries. Field Crops Research 85:167–190

Chopra R, Burow G, Burke JJ, et al (2017) Genome-wide association analysis of seedling traits in diverse Sorghum germplasm under thermal stress. BMC plant biology 17:1–15

Chopra R, Burow G, Hayes C, et al (2015) Transcriptome profiling and validation of gene based single nucleotide polymorphisms (SNPs) in sorghum genotypes with contrasting responses to cold stress. BMC Genomics 16:1040. https://doi.org/10.1186/s12864-015-2268-8

Cisneros-López ME, Mendoza-Onofre LE, Zavaleta-Mancera HA, et al (2010) Pollen–pistil interaction, pistil histology and seed production in A× B grain sorghum crosses under chilling field temperatures. The Journal of Agricultural Science 148:73–82

Cooper EA, Brenton ZW, Flinn BS, et al (2019) A new reference genome for *Sorghum bicolor* reveals high levels of sequence similarity between sweet and grain genotypes: implications for the genetics of sugar metabolism. BMC Genomics 20:420. https://doi.org/10.1186/s12864-019-5734-x

Cuevas HE, Fermin-Pérez RA, Prom LK, et al (2019) Genome-Wide Association Mapping of Grain Mold Resistance in the US Sorghum Association Panel. The plant genome 12:180070

Cuevas HE, Zhou C, Tang H, et al (2016) The Evolution of Photoperiod-Insensitive Flowering in Sorghum, A Genomic Model for Panicoid Grasses. Molecular Biology and Evolution 33:2417–2428. https://doi.org/10.1093/molbev/msw120

Dahlberg J, Wolfrum E, Bean B, Rooney WL (2011) Compositional and agronomic evaluation of sorghum biomass as a potential feedstock for renewable fuels. Journal of Biobased Materials and Bioenergy 5:507–513

Dahlberg JA (2000) Classification and characterization of sorghum. Sorghum: origin, history, technology, and production 99–130

Davidson EH (2010) The regulatory genome: gene regulatory networks in development and evolution. Elsevier

de Wet JMJ (1978) Special paper: systematics and evolution of sorghum sect. Sorghum (Gramineae). American journal of botany 65:477–484

de Wet JMJ, Harlan JR (1971) The origin and domestication of*Sorghum bicolor*. Economic Botany 25:128–135. https://doi.org/10.1007/BF02860074

de Wet JMJ, Huckabay JP (1967) The Origin of *Sorghum bicolor*. II. Distribution and Domestication. Evolution 21:787. https://doi.org/10.2307/2406774

Debouck DG (2003) Managing Plant Genetic Diversity. Crop Science 43:749–750. https://doi.org/10.2135/cropsci2003.749a

Dempewolf H, Baute G, Anderson J, et al (2017) Past and Future Use of Wild Relatives in Crop Breeding. Crop Science 57:1070–1082. https://doi.org/10.2135/cropsci2016.10.0885

Deschamps S, Zhang Y, Llaca V, et al (2018) A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping. Nature Communications 9:4844. https://doi.org/10.1038/s41467-018-07271-1

Deu M, Rattunde F, Chantereau J (2006) A global view of genetic diversity in cultivated sorghums using a core collection. Genome 49:168–180

Dillon SL, Shapter FM, Henry RJ, et al (2007) Domestication to crop improvement: genetic resources for Sorghum and Saccharum (Andropogoneae). Annals of botany 100:975–989. https://doi.org/10.1093/aob/mcm192

Djè Y, Heuertz M, Ater M, et al (2004) In situ estimation of outcrossing rate in sorghum landraces using microsatellite markers. Euphytica 138:205–212

Doggett H (1953) The Sorghums and Sorghum Improvement in Tanganyika. The East African Agricultural Journal 18:155–159. https://doi.org/10.1080/03670074.1953.11664856

Doggett H (1970) Sorghum. Sorghum 403

Doggett Hugh (1988) Sorghum. Longman Scientific & Technical ; Wiley, Harlow, Essex, England; New York

Dong P, Tu X, Chu P-Y, et al (2017) 3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments. Molecular Plant 10:1497–1509. https://doi.org/10.1016/j.molp.2017.11.005

Downes RW, Marshall DR (1971) Low temperature induced male sterility in *Sorghum bicolor*. Australian Journal of Experimental Agriculture 11:352–356

Ejeta G, Grenier C (2005) Sorghum and its weedy hybrids. Crop ferality and volunteerisim CRC Press, Boca Raton, FL 123–135

Ejeta G, Knoll JE (2007) Marker-assisted selection in sorghum. In: Genomics-assisted crop improvement. Springer, pp 187–205

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

Endresen DTF, Street K, Mackay M, et al (2012) Sources of resistance to stem rust (Ug99) in bread wheat and durum wheat identified using focused identification of germplasm strategy. Crop Science 52:764–773. https://doi.org/10.2135/cropsci2011.08.0427

Feltus FA, Hart GE, Schertz KF, et al (2006) Alignment of genetic maps and QTLs between inter-and intra-specific sorghum populations. Theoretical and applied genetics 112:1295–1305

Fernandes SB, Dias KOG, Ferreira DF, Brown PJ (2018) Efficiency of multi-trait, indirect, and trait-assisted genomic selection for improvement of biomass sorghum. Theoretical and applied genetics 131:747–755

Fiedler K, Bekele WA, Duensing R, et al (2014) Genetic dissection of temperature-dependent sorghum growth during juvenile development. Theoretical and Applied Genetics 127:1935–1948. https://doi.org/10.1007/s00122-014-2350-7

Fiedler K, Bekele WA, Friedt W, et al (2012) Genetic dissection of the temperature dependent emergence processes in sorghum using a cumulative emergence model and stability parameters. Theoretical and Applied Genetics 125:1647–1661. https://doi.org/10.1007/s00122-012-1941-4

Fiedler K, Bekele WA, Matschegewski C, et al (2016) Cold tolerance during juvenile development in sorghum: a comparative analysis by genomewide association and linkage mapping. Plant Breeding 135:598–606. https://doi.org/10.1111/pbr.12394

Forbes GA (1992) A Short Communication on Sorghum Diseases in Europe. Sorghum and millets diseases: a second world review 502:89

Frankel OH (1984) Genetic perspectives of germplasm conservation. Genetic manipulation: impact on man and society

Franks CD, Burow GB, Burke JJ (2006) A comparison of US and Chinese sorghum germplasm for early season cold tolerance. Crop science 46:1371–1376

Furman BJ (2006) Methodology to establish a composite collection: case study in lentil. Plant Genetic Resources 4:2–12

Gaut BS, Seymour DK, Liu Q, Zhou Y (2018) Demography and its effects on genomic variation in crop domestication. Nature plants 4:512–520

Ghosal S, Zheng B, Chapman SC, et al (2019) A Weakly Supervised Deep Learning Framework for Sorghum Head Detection and Counting. Plant Phenomics 2019:1–14. https://doi.org/10.34133/2019/1525874

Ghosh S, Matsuoka Y, Asai Y, et al (2011) Software for systems biology: from tools to integrated platforms. Nature Reviews Genetics 12:821–832. https://doi.org/10.1038/nrg3096

Girma G, Nida H, Seyoum A, et al (2019) A large-scale genome-wide association analyses of Ethiopian sorghum landrace collection reveal loci associated with important traits. Frontiers in plant science 10:691

Gladman N, Olson A, Wei S, et al (2022) SorghumBase: a web-based portal for sorghum genetic information and community advancement. Planta 255:1–8. https://doi.org/10.1007/s00425-022-03821-6

Goodstein DM, Shu S, Howson R, et al (2012) Phytozome: a comparative platform for green plant genomics. Nucleic acids research 40:D1178–D1186

Gordon SP, Contreras-Moreira B, Woods DP, et al (2017) Extensive gene content variation in the Brachypodium distachyon pan-genome correlates with population structure. Nature communications 8:1–13

Gorjanc G, Dumasy JF, Gonen S, et al (2017) Potential of low-coverage genotyping-by-sequencing and imputation for cost-effective genomic selection in biparental segregating populations. Crop Science 57:1404–1420. https://doi.org/10.2135/cropsci2016.08.0675

Grootboom AW, Mkhonza NL, O'Kennedy MM, et al (2010) Biolistic mediated sorghum (*Sorghum bicolor* L. Moench) transformation via mannose and bialaphos based selection systems

Guo W, Zheng B, Potgieter AB, et al (2018) Aerial imagery analysis – Quantifying appearance and number of sorghum heads for applications in breeding and agronomy. Frontiers in Plant Science 871:1544. https://doi.org/10.3389/FPLS.2018.01544/BIBTEX

Guy CL, Haskell D (1987) Induction of freezing tolerance in spinach is associated with the synthesis of cold acclimation induced proteins. Plant physiology 84:872–8. https://doi.org/10.1104/pp.84.3.872

Hamblin MT, Mitchell SE, White GM, et al (2004) Comparative Population Genetics of the Panicoid Grasses: Sequence Polymorphism, Linkage Disequilibrium and Selection in a Diverse Sample of *Sorghum bicolor*. Genetics 167:471–483. https://doi.org/10.1534/genetics.167.1.471

Han Y, Lv P, Hou S, et al (2015) Combining Next Generation Sequencing with Bulked Segregant Analysis to Fine Map a Stem Moisture Locus in Sorghum (*Sorghum bicolor* L. Moench). PLOS ONE 10:e0127065

Hao H, Li Z, Leng C, et al (2021) Sorghum breeding in the genomic era: opportunities and challenges. Theoretical and Applied Genetics. https://doi.org/10.1007/s00122-021-03789-z

Hatzig S v., Schiessl S, Stahl A, Snowdon RJ (2015) Characterizing root response phenotypes by neural network analysis. Journal of Experimental Botany 66:5617–5624. https://doi.org/10.1093/jxb/erv235

Haupt M, Schmid K (2020) Combining focused identification of germplasm and core collection strategies to identify genebank accessions for central European soybean breeding. Plant Cell and Environment 43:1421–1436. https://doi.org/10.1111/pce.13761

Hedden P (2003) The genes of the Green Revolution. Trends in Genetics 19:5–9

Hokanson KE, Ellstrand NC, Ouedraogo JT, et al (2010) Biofortified sorghum in Africa: using problem formulation to inform risk assessment. Nature Biotechnology 28:900–903. https://doi.org/10.1038/nbt0910-900

Hulbert SH, Richter TE, Axtell JD, Bennetzen JL (1990) Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proceedings of the National Academy of Sciences 87:4251–4255

Hulluka M, Esele JPE (1992) Sorghum diseases in eastern Africa. Sorghum and Millet Diseases: A Second World Review 36–39

Jedmowski C, Ashoub A, Beckhaus T, et al (2014) Comparative Analysis of *Sorghum bicolor* Proteome in Response to Drought Stress and following Recovery. International Journal of Proteomics 2014:395905. https://doi.org/10.1155/2014/395905

Ji G, Zhang Q, Du R, et al (2017) Construction of a high-density genetic map using specific-locus amplified fragments in sorghum. BMC Genomics 18:51. https://doi.org/10.1186/s12864-016-3430-7

Jiao Y, Burke J, Chopra R, et al (2016) A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses. The Plant cell 28:1551–1562. https://doi.org/10.1105/tpc.16.00373

Jørgensen IH (1992) Discovery, characterization and exploitation of Mlo powdery mildew resistance in barley. Euphytica 63:141–152. https://doi.org/10.1007/BF00023919

Kajiya-Kanegae H, Takanashi H, Fujimoto M, et al (2020) RAD-seq-Based High-Density Linkage Map Construction and QTL Mapping of Biomass-Related Traits in Sorghum using the Japanese Landrace Takakibi NOG. Plant and Cell Physiology 61:1262–1272. https://doi.org/10.1093/pcp/pcaa056

Kayodé APP, Linnemann AR, Nout MJR, et al (2006) Diversity and food quality properties of farmers' varieties of sorghum from Benin. Journal of the Science of Food and Agriculture 86:1032–1039

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

Kebrom TH, Mullet JE (2016) Transcriptome profiling of tiller buds provides new insights into PhyB regulation of tillering and indeterminate growth in sorghum. Plant physiology 170:2232–2250

Khazaei H, Street K, Bari A, et al (2013) The FIGS (Focused Identification of Germplasm Strategy) Approach Identifies Traits Related to Drought Adaptation in Vicia faba Genetic Resources. PLoS ONE 8:e63107. https://doi.org/10.1371/journal.pone.0063107

Klein RR, Miller FR, Dugas D v, et al (2015) Allelic variants in the PRR37 gene and the human-mediated dispersal and diversification of sorghum. Theoretical and Applied Genetics 128:1669–1683. https://doi.org/10.1007/s00122-015-2523-z

Knoll J, Ejeta G (2008) Marker-assisted selection for early-season cold tolerance in sorghum: QTL validation across populations and environments. Theoretical and Applied Genetics 116:541–553

Knoll J, Gunaratna N, Ejeta G (2008) QTL analysis of early-season cold tolerance in sorghum. Theoretical and Applied Genetics 116:577–587. https://doi.org/10.1007/s00122-007-0692-0

Koca H, Bor M, Özdemir F, Türkan İ (2007) The effect of salt stress on lipid peroxidation, antioxidative enzymes and proline content of sesame cultivars. Environmental and Experimental Botany 60:344–351. https://doi.org/https://doi.org/10.1016/j.envexpbot.2006.12.005

Krishnamurthy L, Serraj R, Hash CT, et al (2007) Screening sorghum genotypes for salinity tolerant biomass production. Euphytica 156:15–24

Kumari S, Kumar V, Beilsmith K, et al (2021) A KBase case study on genome-wide transcriptomics and plant primary metabolism in response to drought stress in Sorghum. Current Plant Biology 28:100229. https://doi.org/10.1016/j.cpb.2021.100229

Lawrence CJ, Walbot V (2007) Translational genomics for bioenergy production from fuelstock grasses: maize as the model species. The Plant Cell 19:2091–2094

Lemasurier LS, Panozzo JF, Walker CK (2014) A digital image analysis method for assessment of lentil size traits. Journal of Food Engineering 128:72–78. https://doi.org/10.1016/j.jfoodeng.2013.12.018

León-Velasco H, Mendoza-Onofre LE;, Castillo-González F;, et al (2009) Evaluation of two generations of cold tolerant sorghum hybrids and parental lines. II: Combining ability, heterosis and heterobeltiosis. Agrociencia 43:609–623

Li H, Rasheed A, Hickey LT, He Z (2018) Fast-Forwarding Genetic Gain. Trends in Plant Science 23:184–186

Li Y, Zhou G, Ma J, et al (2014) De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature biotechnology 32:1045–1052

Liu G, Godwin ID (2012) Highly efficient sorghum transformation. Plant cell reports 31:999–1007

Luo H, Zhao W, Wang Y, et al (2016) SorGSD: a sorghum genome SNP database. Biotechnology for biofuels 9:1–9

Mace E, Innes D, Hunt C, et al (2019) The Sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. Theoretical and applied genetics 132:751–766

Mace ES, Tai S, Gilding EK, et al (2013) Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. Nature communications 4:1–9

Mace ES, Xia L, Jordan DR, et al (2008) DArT markers: diversity analyses and mapping in *Sorghum bicolor*. BMC Genomics 9:26. https://doi.org/10.1186/1471-2164-9-26

Macharia JM, Kamau J, Gituanja JN, Matu EW (1994) Effects of sodium salinity on seed germination and seedling root and shoot extension of four sorghum (*Sorghum bicolor* (L.) Moench) cultivars. Sorghum Improvement Conference of North America, USA; University of Georgia …

Maiti R (1996) Sorghum science. Science Publishers, Inc.

Makita Y, Shimada S, Kawashima M, et al (2015) MOROKOSHI: transcriptome database in *Sorghum bicolor*. Plant & cell physiology 56:e6. https://doi.org/10.1093/pcp/pcu187

Mann JA, Kimber CT, Miller FR (1983) The Origin and Early Cultivation of Sorghums in Africa. Texas FARMER Collection

Marla SR, Burow G, Chopra R, et al (2019) Genetic architecture of chilling tolerance in sorghum dissected with a nested association mapping population. G3: Genes, Genomes, Genetics 9:4045–4057. https://doi.org/10.1534/g3.119.400353

Mascher M, Schreiber M, Scholz U, et al (2019) Genebank genomics bridges the gap between the conservation of crop diversity and plant breeding. Nature Genetics 51:1076–1081

Maulana F, Tesso TT (2013) Cold temperature episode at seedling and flowering stages reduces growth and yield components in sorghum. Crop Science 53:564–574

Maulana F, Weerasooriya D, Tesso T (2017) Sorghum landrace collections from cooler regions of the world exhibit magnificent genetic differentiation and early season cold tolerance. Frontiers in plant science 8:756

Mbeyagala EK, Kiambi DD, Okori P, Edema R (2012) Molecular diversity among sorghum (*Sorghum bicolor* (L.) Moench) Landraces in Uganda. International Journal of Botany 8:85–95

McCormick RF, Truong SK, Sreedasyam A, et al (2018) The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. The Plant Journal 93:338–354. https://doi.org/https://doi.org/10.1111/tpj.13781

McCouch S (2013) Agriculture: Feeding the future. Nature 499:23–24

McCouch S, Navabi ZK, Abberton M, et al (2020) Mobilizing Crop Biodiversity. Molecular Plant 13:1341–1344. https://doi.org/10.1016/j.molp.2020.08.011

McFerson JR, Lamboy WF, Kresovich S (1996) Assessing user perceptions of genetic resource collections in crucifer crops. Crop Science 36:831–838

McGuire S, Sperling L (2013) Making seed systems more resilient to stress. Global Environmental Change 23:644–653. https://doi.org/10.1016/j.gloenvcha.2013.02.001

McKhann HI, Camilleri C, Bérard A, et al (2004) Nested core collections maximizing genetic diversity in Arabidopsis thaliana. The Plant Journal 38:193–202

McKinley B, Rooney W, Wilkerson C, Mullet J (2016) Dynamics of biomass partitioning, stem gene expression, cell wall biosynthesis, and sucrose accumulation during development of *Sorghum bicolor*. The Plant Journal 88:662–680

Mendoza-Onofre LE (1988) Formation of sorghum hybrids for grain. II. Behavior per se of lines and their general combining ability. Revis Fitotec Mex 11:39–47

Menz MA, Klein RR, Unruh NC, et al (2004) Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Science 44:1236–1244

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. Nature Reviews Genetics 14:840–852. https://doi.org/10.1038/nrg3605

Meza-Basso L, Alberdi M, Raynal M, et al (1986) Changes in Protein Synthesis in Rapeseed (Brassica napus) Seedlings during a Low Temperature Treatment. Plant physiology 82:733–8. https://doi.org/10.1104/pp.82.3.733

Milner SG, Jost M, Taketa S, et al (2019) Genebank genomics highlights the diversity of a global barley collection. Nature Genetics 51:319–326. https://doi.org/10.1038/s41588-018-0266-x

Morris GP, Ramu P, Deshpande SP, et al (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proceedings of the National Academy of Sciences 110:453–458

Mukuru SZ (1993) Sorghum and millet in Eastern Africa. Sorghum and Millets Commodities and Research Environment(Byth, DE, ed) Patancheru 502:55–62

Multani DS, Briggs SP, Chamberlin MA, et al (2003) Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. Science 302:81–84

Murphy RL, Morishige DT, Brady JA, et al (2014) Ghd7 (Ma6) represses sorghum flowering in long days: Ghd7 alleles enhance biomass accumulation and grain production. Plant Genome 7: 1–10

Nagel M, Holstein K, Willner E, Börner A (2018) Machine learning links seed composition, glucosinolates and viability of oilseed rape after 31 years of long-term storage. Seed Science Research 28:340–348. https://doi.org/10.1017/S0960258518000259

Nelson JC, Wang S, Wu Y, et al (2011) Single-nucleotide polymorphism discovery by high-throughput sequencing in sorghum. BMC Genomics 12:352. https://doi.org/10.1186/1471-2164-12-352

Nguyen GN, Maharjan P, Maphosa L, et al (2019) A Robust Automated Image-Based Phenotyping Method for Rapid Vegetative Screening of Wheat Germplasm for Nitrogen Use Efficiency. Frontiers in Plant Science 10:. https://doi.org/10.3389/fpls.2019.01372

Nguyen GN, Norton SL, Rosewarne GM, et al (2018) Automated phenotyping for early vigour of field pea seedlings in controlled environment by colour imaging technology. PLoS ONE 13:. https://doi.org/10.1371/journal.pone.0207788

Nida H, Blum S, Zielinski D, et al (2016) Highly efficient de novo mutant identification in a *Sorghum bicolor* TILLING population using the ComSeq approach. The Plant Journal 86:349–359

Nida H, Lee S, Li Y, Mengiste T (2021) Transcriptome analysis of early stages of sorghum grain mold disease reveals defense regulators and metabolic pathways associated with resistance. BMC Genomics 22:295. https://doi.org/10.1186/s12864-021-07609-y

Nussbaumer T, Martis MM, Roessner SK, et al (2012) MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Research 41:D1144–D1151. https://doi.org/10.1093/nar/gks1153

Ohadi S, Hodnett G, Rooney W, Bagavathiannan M (2017) Gene Flow and its Consequences in Sorghum spp. Critical Reviews in Plant Sciences 36:367–385

Olatoye MO, Hu Z, Maina F, Morris GP (2018) Genomic Signatures of Adaptation to a Precipitation Gradient in Nigerian Sorghum. G3 Genes|Genomes|Genetics 8:3269–3281. https://doi.org/10.1534/g3.118.200551

Ortiz D, Hu J, Salas Fernandez MG (2017) Genetic architecture of photosynthesis in *Sorghum bicolor* under non-stress and cold stress conditions. Journal of experimental botany 68:4545–4557

Osuna-Ortega J, Mendoza-Castillo MDC, Mendoza-Onofre L (2003) Sorghum cold tolerance, pollen production. Maydica 48:125–132

Osuna-Ortega J, Mendoza-Onofre LE, González-Hernández VA, et al (2000) Potential of cold tolerant germplasm in the adaptation and adaptability of sorghum in México: I. High valleys. Agrociencia 34:561–572

Oyediran IO, Hibbard BE, Clark TL (2004) Prairie grasses as hosts of the western corn rootworm (Coleoptera: Chrysomelidae). Environmental Entomology 33:740–747

Pantazi XE, Moshou D, Alexandridis T, et al (2016) Wheat yield prediction using machine learning and advanced sensing techniques. Computers and Electronics in Agriculture 121:57–65. https://doi.org/10.1016/j.compag.2015.11.018

Parra-Londono S, Fiedler K, Kavka M, et al (2018) Genetic dissection of early-season cold tolerance in sorghum: genome-wide association studies for seedling emergence and survival under field and controlled environment conditions. Theoretical and applied genetics 131:581–595

Patane C, Cavallaro V, Avola G, D'Agosta G (2006) Seed respiration of sorghum [*Sorghum bicolor* (L.) Moench] during germination as affected by temperature and osmoconditioning. Seed Science Research 16:251–260

Patanè C, Saita A, Sortino O (2013) Comparative Effects of Salt and Water Stress on Seed Germination and Early Embryo Growth in Two Cultivars of Sweet Sorghum. Journal of Agronomy and Crop Science 199:30–37. https://doi.org/https://doi.org/10.1111/j.1439-037X.2012.00531.x

Paterson AH, Bowers JE, Bruggmann R, et al (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457:551–556. https://doi.org/10.1038/nature07723

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proceedings of the National Academy of Sciences 101:9903–9908

Pei L, Li G, Lindsey K, et al (2021) Plant 3D genomics: the exploration and application of chromatin organization. New Phytologist 230:1772–1786. https://doi.org/10.1111/nph.17262

Peng Y, Xiong D, Zhao L, et al (2019) Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. Nature Communications 10:2632. https://doi.org/10.1038/s41467-019-10602-5

Peres S (2016) Saving the gene pool for the future: Seed banks as archives. Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences 55:96–104. https://doi.org/10.1016/j.shpsc.2015.09.002

Philipp N, Weise S, Oppermann M, et al (2018) Leveraging the use of historical data gathered during seed regeneration of an ex situ genebank collection of wheat. Frontiers in Plant Science 9:. https://doi.org/10.3389/fpls.2018.00609

Phuke RM, Anuradha K, Radhika K, et al (2017) Genetic Variability, Genotype × Environment Interaction, Correlation, and GGE Biplot Analysis for Grain Iron and Zinc Concentration and Other Agronomic Traits in RIL Population of Sorghum (*Sorghum bicolor* L. Moench) . Frontiers in Plant Science 8:712

Pistorius R (1997) Scientists, plants and politics: a history of the plant genetic resources movement. Bioversity International

Prasad PVV, Boote KJ, Allen Jr LH (2006) Adverse high temperature effects on pollen viability, seed-set, seed yield and harvest index of grain-sorghum [*Sorghum bicolor* (L.) Moench] are more severe at elevated carbon dioxide due to higher tissue temperatures. Agricultural and forest meteorology 139:237–251

Prasad PV v, Djanaguiraman M, Perumal R, Ciampitti IA (2015) Impact of high temperature stress on floret fertility and individual grain weight of grain sorghum: sensitive stages and thresholds for temperature and duration. Frontiers in Plant Science 6:820

Promkhambut A, Younger A, Polthanee A, Akkasaeng C (2010) Morphological and physiological responses of sorghum (*Sorghum bicolor* L. Moench) to waterlogging. Asian Journal of Plant Sciences 9:183

Proost S, Van Bel M, Vaneechoutte D, et al (2015) PLAZA 3.0: an access point for plant comparative genomics. Nucleic Acids Research 43:D974–D981. https://doi.org/10.1093/nar/gku986

Quinby JR (1966) Fourth Maturity Gene Locus in Sorghum 1. Crop Science 6:516–518

Quinby JR, Karper RE (1945) The Inheritance of Three Genes That Influence Time of Floral Initiation and Maturity Date in Milo 1. Agronomy Journal 37:916–936. https://doi.org/10.2134/agronj1945.00021962003700110006x

Ramirez-Villegas J, Jarvis A, Läderach P (2013) Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. Agricultural and Forest Meteorology 170:67–78. https://doi.org/10.1016/j.agrformet.2011.09.005

Rao PP, Birthal PS, Reddy BVS, et al (2006) Diagnostics of sorghum and pearl millet grains-based nutrition in India. International sorghum and millets newsletter 47:93–96

Rincent R, Charpentier JP, Faivre-Rampant P, et al (2018) Phenomic selection is a low-cost and high-throughput method based on indirect predictions: Proof of concept on wheat and poplar. G3: Genes, Genomes, Genetics 8:3961–3972. https://doi.org/10.1534/g3.118.200760

Romay MC, Millard MJ, Glaubitz JC, et al (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biology 14:R55. https://doi.org/10.1186/gb-2013-14-6-r55

Rosati VC, Blomstedt CK, Møller BL, et al (2019) The Interplay Between Water Limitation, Dhurrin, and Nitrate in the Low-Cyanogenic Sorghum Mutant adult cyanide deficient class 1 . Frontiers in Plant Science 10:1458

Ruperao P, Thirunavukkarasu N, Gandham P, et al (2021) Sorghum Pan-Genome Explores the Functional Utility for Genomic-Assisted Breeding to Accelerate the Genetic Gain . Frontiers in Plant Science 12:963

Schaffasz A, Windpassinger S, Friedt W, et al (2019a) Sorghum as a novel crop for Central Europe: Using a broad diversity set to dissect temperate-adaptation. Agronomy 9:535

Schaffasz A, Windpassinger S, Snowdon R, Wittkop B (2019b) Reproductive cold stress tolerance in sorghum F1 hybrids is a heterotic trait. Agronomy 9:508

Schittenhelm S, Schroetter S (2014) Comparison of drought tolerance of maize, sweet sorghum and sorghum-sudangrass hybrids. Journal of Agronomy and Crop Science 200:46–53

Setter TL, Waters I (2003) Review of prospects for germplasm improvement for waterlogging tolerance in wheat, barley and oats. Plant and soil 253:1–34

Shukla PR, Skea J, Calvo Buendia E, et al (2019) IPCC, 2019: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems

Singh HP, Lohithaswa HC (2006) Sorghum. In: Cereals and Millets. Springer, pp 257–302

Singh SP (1985) Sources of cold tolerance in grain sorghum. Canadian journal of plant science 65:251–257

Sleper DA, Poehlman JM (2006) Breeding field crops. Blackwell publishing

Smith CW, Frederiksen RA (2000) Sorghum: Origin, history, technology, and production. John Wiley & Sons

Smith O, Nicholson W v, Kistler L, et al (2019) A domestication history of dynamic adaptation and genomic deterioration in Sorghum. Nature Plants 5:369–379. https://doi.org/10.1038/s41477-019-0397-9

Spindel JE, Dahlberg J, Colgan M, et al (2018) Association mapping by aerial drone reveals 213 genetic associations for *Sorghum bicolor* biomass traits under drought. BMC genomics 19:1–18

Springer NM, Ying K, Fu Y, et al (2009) Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet 5:e1000734

Srinivasa Rao P, Reddy BVS, Nagaraj N, Upadhyaya HD (2014) Sorghum production for diversified uses. Genetics, Genomics and Breeding of Sorghum CRC Press 2014:1–27

Stephens JC, Holland RF (1954) Cytoplasmic male sterility for hybrid sorghum seed production. Agron J 46:20–23

Stephens JC, Miller FR, Rosenow DT (1967) Conversion of alien Sorghums to early combine genotypes 1. Crop Science 7:396

Tao Y, Luo H, Xu J, et al (2021) Extensive variation within the pan-genome of cultivated and wild sorghum. Nature Plants 7:766–773. https://doi.org/10.1038/s41477-021-00925-x

Tao YZ, Jordan DR, Henzell RG, McIntyre CL (1998) Identification of genomic regions for rust resistance in sorghum. Euphytica 103:287–292

Tari I, Laskay G, Takács Z, Poór P (2013) Response of sorghum to abiotic stresses: a review. Journal of Agronomy and Crop Science 199:264–274

Tian T, You Q, Zhang L, et al (2016) SorghumFDB: sorghum functional genomics database with multidimensional network analysis. Database 2016:. https://doi.org/10.1093/database/baw099

Tiryaki I, Andrews DJ (2001a) Germination and seedling cold tolerance in sorghum: II. Parental lines and hybrids. Agronomy journal 93:1391–1397

Tiryaki I, Andrews DJ (2001b) Germination and seedling cold tolerance in sorghum: I. Evaluation of rapid screening methods. Agronomy Journal 93:1386–1391

Uberegui E, Hall M, Lorenzo Ó, et al (2015) An Arabidopsis soluble chloroplast proteomic analysis reveals the participation of the Executer pathway in response to increased light conditions. Journal of Experimental Botany 66:2067–2077

Upadhyaya HD, Ortiz R (2001) A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. Theoretical and Applied Genetics 102:1292–1298. https://doi.org/10.1007/s00122-001-0556-y

Upadhyaya HD, Wang Y-H, Sastry DVSSR, et al (2015) Association mapping of germinability and seedling vigor in sorghum under controlled low-temperature conditions. Genome 59:137–145. https://doi.org/10.1139/gen-2015-0122

van Hintum TJL, Visser DL (1995) Duplication within and between germplasm collections - II. Duplication in four European barley collections. Genetic Resources and Crop Evolution 42:135–145. https://doi.org/10.1007/BF02539517

Velazco JG, Jordan DR, Mace ES, et al (2019a) Genomic Prediction of Grain Yield and Drought-Adaptation Capacity in Sorghum Is Enhanced by Multi-Trait Analysis . Frontiers in Plant Science 10:997

Velazco JG, Malosetti M, Hunt CH, et al (2019b) Combining pedigree and genomic information to improve prediction quality: an example in sorghum. Theoretical and Applied Genetics 132:2055–2067

Venkateswaran K, Muraya M, Dwivedi SL, Upadhyaya HD (2014) Wild sorghums—Their potential use in crop improvement. Genetics, genomics and breeding of sorghum 78–111

Vicente R, Bolger AM, Martínez-Carrasco R, et al (2019) De Novo Transcriptome Analysis of Durum Wheat Flag Leaves Provides New Insights Into the Regulatory Response to Elevated CO2 and High Temperature . Frontiers in Plant Science 10:1605

Voss-Fels KP, Qian L, Parra-Londono S, et al (2017) Linkage drag constrains the roots of modern wheat. Plant, Cell & Environment 40:717–725. https://doi.org/10.1111/pce.12888

Wambugu PW, Ndjiondjop MN, Henry RJ (2018) Role of genomics in promoting the utilization of plant genetic resources in genebanks. Briefings in functional genomics 17:198–206

Wang M, Tu L, Lin M, et al (2017) Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. Nature Genetics 49:579–587. https://doi.org/10.1038/ng.3807

Wang W, Mauleon R, Hu Z, et al (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 557:43–49

Weiser CJ (1970) Cold Resistance and Injury in Woody Plants. Science 169:1269–1278. https://doi.org/10.1126/science.169.3952.1269

Winchell F, Stevens CJ, Murphy C, et al (2017) Evidence for sorghum domestication in fourth millennium BC eastern Sudan: Spikelet morphology from ceramic impressions of the Butana group. Current Anthropology 58:673–683

Wolabu TW, Tadege M (2016) Photoperiod response and floral transition in sorghum. Plant signaling & behavior 11:e1261232–e1261232. https://doi.org/10.1080/15592324.2016.1261232

Wolabu TW, Zhang F, Niu L, et al (2016) Three FLOWERING LOCUS T-like genes function as potential florigens and mediate photoperiod response in sorghum. New Phytologist 210:946–959. https://doi.org/https://doi.org/10.1111/nph.13834

Xie T, Zhang F-G, Zhang H-Y, et al (2019) Biased gene retention during diploidization in Brassica linked to three-dimensional genome organization. Nature Plants 5:822–832. https://doi.org/10.1038/s41477-019-0479-8

Xu W, Subudhi PK, Crasta OR, et al (2000) Molecular mapping of QTLs conferring stay-green in grain sorghum (*Sorghum bicolor* L. Moench). Genome 43:461–469

Yang L, Wu K, Gao P, et al (2014) GsLRPK, a novel cold-activated leucine-rich repeat receptor-like protein kinase from Glycine soja, is a positive regulator to cold stress tolerance. Plant Science 215–216:19–28. https://doi.org/https://doi.org/10.1016/j.plantsci.2013.10.009

Yano K, Yamamoto E, Aya K, et al (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nature genetics 48:927–34. https://doi.org/10.1038/ng.3596

Youens-Clark K, Buckler E, Casstevens T, et al (2011) Gramene database in 2010: updates and extensions. Nucleic Acids Research 39:D1085. https://doi.org/10.1093/NAR/GKQ1148

Yu X, Li X, Guo T, et al (2016a) Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nature Plants 2:16150. https://doi.org/10.1038/nplants.2016.150

Yu X, Li X, Guo T, et al (2016b) Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nature Plants 2:. https://doi.org/10.1038/nplants.2016.150

Zhang D (2015) Population genomic analysis of inflorescence, plant height and leaf architecture components in sorghum (Panicoidae); and comparative genetics with rice (Oryzoidae)

Zhao L, Wang S, Cao Z, et al (2019) Chromatin loops associated with active genes and heterochromatin shape rice genome architecture for transcriptional regulation. Nature Communications 10:3640. https://doi.org/10.1038/s41467-019-11535-9

Zhao Z, Cai T, Tagliani L, et al (2000) Agrobacterium-mediated sorghum transformation. Plant molecular biology 44:789–798

Zheng X, Chen B, Lu G, Han B (2009) Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. Biochemical and biophysical research communications 379:985–989

Declaration

I declare that the dissertation here submitted is entirely my own work, written without any illegitimate help by any third party and solely with materials as indicated in the dissertation. I have indicated in the text where I have used texts from already published sources, either word for word or in substance, and where I have made statements based on oral information given to me. At all times during the investigations carried out by me and described in the dissertation, I have followed the principles of good scientific practice as defined in the "Statutes of the Justus Liebig University Giessen for the Safeguarding of Good Scientific Practice".

Giessen, Friday 12th August, 2022 Subhadra Chakrabarty

Acknowledgements

The completion of this thesis would not have been possible without the support of many people and I am deeply indebted to them.

Firstly, I would like to thank Prof. Dr. Rod Snowdon, for your immense patience and effort for the completion of my thesis. I still remember your module "Biotechnology Genomics", which felt like science fiction back then, was something that really got me interested in plant breeding. I am forever grateful to you for never saying "no" to any learning opportunities, travel or crazy work ideas. Whenever I felt lost during my PhD, you made sure that I see the light at the end of the tunnel. Thank you Rod.

I am also grateful to Prof. Dr. Matthias Frisch for agreeing to be the second supervisor for my doctoral thesis.

Another person who played a crucial role in shaping my scientific career would be Dr. Steffen Windpassinger. You taught me everything about sorghum and how practical breeding works. Thank you for always being there to answer my questions, teach me and being so supportive.

I want to say a big thank you to all the technicians from our department. A big thank you Stavros, for teaching me everything I know of lab work and organization. You always answered all my questions (even the stupid ones!) and someday I shall give you your "5 euro"! The memories of "Walltorstrasse" will always hold a special place in my heart. Furthermore, I want to thank Birgit, Annette, Regina and Sabine and the entire team at Gross Gerau for the immense help, be it field, greenhouse or lab work. It would not have been possible without you all. In addition, I want to specially mention Lennart, Luisa, Natalja and Price for helping out.

From the bottom of my heart, I want to thank the best secretaries one could imagine, Miss. Sabine Schomber, Miss. Ulla Riedmeier and Miss. Usha Beher. I never had to worry any administrative thing because of you. Thank you, especially Miss Schomber for making me feel so special and loved.

This thesis would not have been possible without the support of my friends and colleagues here in Giessen. Thank you for your help and managing my stress levels. I am grateful for all the discussions, fun coffee breaks and crazy parties, conference trips and the amazing times. A very special "thank you" to Harmeet, my friend philosopher and guide. You have taught me, guided me and supported me at every step of the way. I probably wouldn't have survived my Giessen time without your help.

I am deeply indebted to Raphael Müller for all the help and support. I specially want to mention my uncle, Mr. Aftabul Islam and his family. If it wasn't for you, I might have gone back to India after my masters. Thank you for making sure I have family here.

Last but not the least, I would like to thank my parents and family. Thank you for all your love, patience and support. It wouldn't be possible without you all. I want to specially mention my aunt,

Dr. Snigdha Chakrabarty, for your support and guidance. You always understood the struggles and helped me out. A special “thank you” to my uncle, Mr Kalyan Chakrabarty for believing in me and letting me pursue my dreams. This thesis is as much mine and it is all of yours.

I would also like to acknowledge deNBI (German Network for Bioinformatics Infrastructure) for providing the computational resources needed for this thesis.