



Sorghum genetic, genomic, and breeding resources

Zhanguo Xin¹ · Mingli Wang² · Hugo E. Cuevas³ · Junping Chen¹ · Melanie Harrison² · N. Ace Pugh¹ · Geoffrey Morris⁴

Received: 16 April 2021 / Accepted: 28 September 2021 / Published online: 5 November 2021
This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021

Abstract

Main conclusion Sorghum research has entered an exciting and fruitful era due to the genetic, genomic, and breeding resources that are now available to researchers and plant breeders.

Abstract As the world faces the challenges of a rising population and a changing global climate, new agricultural solutions will need to be developed to address the food and fiber needs of the future. To that end, sorghum will be an invaluable crop species as it is a stress-resistant C₄ plant that is well adapted for semi-arid and arid regions. Sorghum has already remained as a staple food crop in many parts of Africa and Asia and is critically important for animal feed and niche culinary applications in other regions, such as the United States. In addition, sorghum has begun to be developed into a promising feedstock for forage and bioenergy production. Due to this increasing demand for sorghum and its potential to address these needs, the continuous development of powerful community resources is required. These resources include vast collections of sorghum germplasm, high-quality reference genome sequences, sorghum association panels for genome-wide association studies of traits involved in food and bioenergy production, mutant populations for rapid discovery of causative genes for phenotypes relevant to sorghum improvement, gene expression atlas, and online databases that integrate all resources and provide the sorghum community with tools that can be used in breeding and genomic studies. Used in tandem, these valuable resources will ensure that the rate, quality, and collaborative potential of ongoing sorghum improvement efforts is able to rival that of other major crops.

Communicated by Anastasios Melis.

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

✉ Zhanguo Xin
Zhanguo.Xin@ARS.USDA.GOV

- ¹ Plant Stress and Germplasm Development Unit, Crop Systems Research Laboratory, USDA-ARS, 3810, 4th Street, Lubbock, TX 79424, USA
- ² Plant Genetic Resources Conservation Unit, USDA-ARS, Griffin, GA 30223, USA
- ³ Tropical Agriculture Research Station, USDA-ARS, Mayagüez 00680, Puerto Rico
- ⁴ Crop Quantitative Genomics, Soil and Crop Sciences, Colorado State University, Plant Sciences Building, Fort Collins, CO 80523, USA

Introduction

Sorghum (*Sorghum bicolor* L. Moench) is a stress-resilient crop with highly productive NADP-ME type C₄ photosynthesis and highly efficient nitrogen and water utilization (Wang et al. 2009). It is the fifth-most important cereal crop globally, serving as a staple food for over 500 million people in the arid and semi-arid regions of the world. It is also an increasingly important crop for animal feed, forage, and bioenergy feedstock for production of biofuel and bioproducts. As an annual crop, most sorghum varieties in the temperate region are photoperiod-insensitive and can complete life cycle in about 4 months. With help of winter nursery, two crops can be grown in a year. In greenhouse, it is easy to grow three generations in a year. Sorghum lines from the tropical origin are photoperiod-sensitive and need short days to flower. Sorghum is transformable, but only a few lines, such as RTx430 and P898012, have adequate transformation efficiency (Gurel et al. 2009; Howe et al. 2006; Che et al. 2018). A genotype-independent efficient transformation

protocol remains the main obstacle to apply genome editing efficiently in sorghum.

Despite its value, sorghum breeding and genomic studies have lagged other crops like rice and maize. With the completion of a sorghum reference genome sequenced over a decade ago (Paterson et al. 2009), the construction of several sorghum association panels (SAP) (Casa et al. 2008; Morris et al. 2013b; Upadhyaya et al. 2009), the establishment of mutant libraries, the recent completion of sorghum pan-genomes, the availability of sorghum gene expression atlas and the development of sorghumbase online (Addo-Quaye et al. 2018; Jiao et al. 2016; Xin et al. 2008; Tao et al. 2021; Makita et al. 2015; Shakoor et al. 2014), sorghum research has entered an exciting new age. With the vast resources available to producers and researchers, sorghum will become a critical crop for addressing global food and energy security in a changing global climate.

Resource of induced mutant libraries

A well-categorized mutant library generated from an inbred line with uniform genetic background and detailed annotation of phenotypes provides a powerful resource to isolate independent alleles of mutants with relevant traits. In the past, bulk mutagenesis followed by selection of a phenotype of interest has been very effective in the identification of genes underlying specific phenotypes; however, other mutations or phenotypes, which can be important for processes other than the current interest, are often ignored (Oria et al. 2000; Peters et al. 2009; Porter et al. 1978; Singh and Drolsom 1974). To preserve all mutations for functional genomic studies, Xin et al. designed a systematic approach to develop a pedigreed mutant library (Xin et al. 2008). Individual seeds from an elite inbred line, BTx623, were subjected to chemical mutagenesis by soaking them in various concentrations of ethyl methane sulfonate (EMS) from 0.1 to 0.25% (v/w). BTx623 was chosen for this purpose because it was the sorghum line used to generate the first sorghum reference genome. Next, the mutagenized M_1 seeds were propagated to the M_3 generation by single-seed descent. Ten panicles from the M_3 plants were bulk selected to serve as one pool of M_4 seeds. Thus, each pool of seeds was derived from a single independently mutagenized M_1 seed by pedigree. The pedigreed M_4 seed pools were propagated from M_1 seeds without selection; therefore, most mutations, including recessive lethal mutations, are preserved in the mutant library. Because the M_4 seed pools can be replenished by planting and pooling 20 or more panicles from the remaining seeds, the pedigreed mutant library can serve as a permanent resource for screening relevant mutants under a variety of growth conditions. The library now consists of approximately 6400 independent seed pools that can be used

to select mutants with traits that are potentially useful for sorghum improvement. A similar approach has been used to generate approximately 10,000 pedigreed seed pools in Purdue University (Addo-Quaye et al. 2017, 2018). Within the pedigreed mutant library, a wide range of phenotypes relevant to sorghum improvement have been observed and the mutants have been selected (Jiao et al. 2016). Furthermore, both forward and reverse genetic resources to efficiently explore the mutant library for traits of potentially useful for sorghum improvement and genomic studies have been developed (Jiao et al. 2016, 2017; Wang et al. 2021). Here, we discuss these resources and how to use them to facilitate sorghum studies. We also discussed a few classes of mutants in detail so that we may demonstrate the utility of the mutant library.

Fast forward genetics—identify causal mutations from the pedigreed mutant library via bulk segregant analysis and next-generation sequencing

Identification of causal gene mutation underlying a mutant phenotype of interest, termed forward genetic, is instrumental in understanding the mechanisms governing growth and development, signal transduction, and metabolic pathways (<http://thearabidopsisbook.org>). The conventional method to identify the causal mutation from a mutant is map-based cloning, which requires the analysis of thousands of DNA markers in a large segregating individual F_2 population (usually over 1000 individual F_2 plants) to delimit the mutation to a region harboring only a few genes (Jander et al. 2002). The conventional map-based cloning is lengthy and complicated. It often takes a skilled researcher three or more years to identify one gene in model plants, like *Arabidopsis* (Jander et al. 2002).

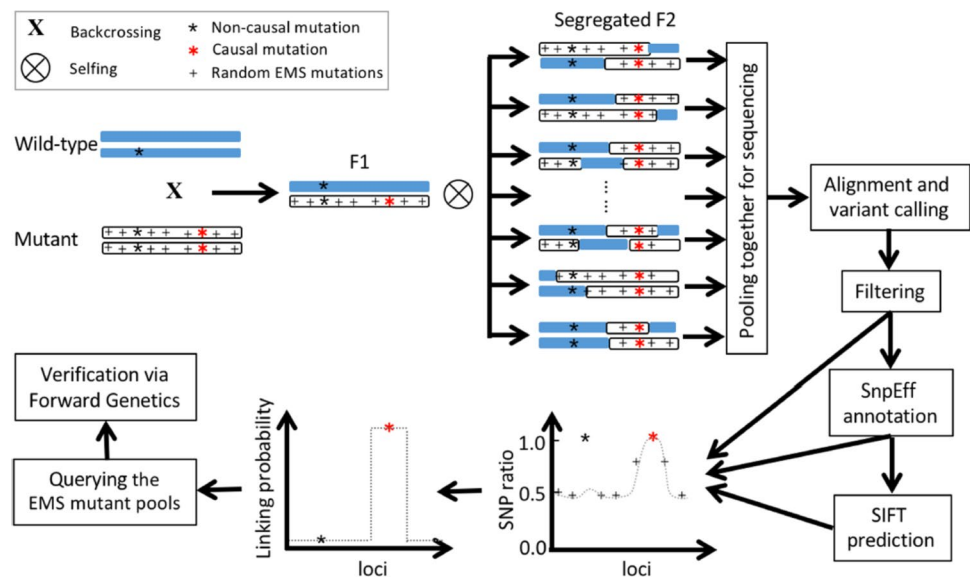
The rapid innovation in next-generation sequencing (NGS) techniques provides large number of DNA markers at affordable cost (Metzker 2010). Two general approaches have been developed to use the single-nucleotide polymorphic (SNP) markers that are annotated from NGS to map and identify the causal mutations underlying mutant phenotypes of interesting. The first strategy is represented by mapping-by-sequencing (ShoreMap) (Schneeberger et al. 2009; Hartwig et al. 2012) or next-generation map (NGM) (Austin et al. 2011). This approach is very similar to conventional map-based cloning by crossing the mutant to a line that has extensive DNA polymorphism from the mutant. Equal amount of DNA from many homozygous mutants selected from the F_2 population is pooled and sequenced. The co-segregation of SNP markers with the mutant phenotype is analyzed following the principle of bulk segregant analysis (BSA) (Michelmore et al. 1991). The second

strategy is isogenic mapping-by-sequencing, also called MutMap, a variation of the ShoreMap, designed to identify the causal mutations from bulked F₂ mutants isolated from a cross of the mutant to its original un-mutated parent (Abe et al. 2012; Hartwig et al. 2012; Zhu et al. 2012). If the F₂ population segregates for a mutant phenotype to the wild-type phenotype at an approximately 1 to 3 ratio, or consistent with recessive mutation, genomic DNA from 50 homozygous of the mutants selected from the F₂ population is pooled for NGS. The causal mutation for the phenotype is expected to be a 100% of the mutated SNP due to the selection. The mutations that are unlinked with the causal mutation will have a SNP ratio (the number of mutated SNPs/total SNPs) of 0.5, approximately 50% the of the SNPs are the mutated type.

Conceptually, MutMap is a straightforward process; however, in practice, it is often very difficult to distinguish causative mutations from background mutations present in the parental lines before mutagenesis or from sequencing errors in the reference genome, which also show SNP ratios of 1. Furthermore, MutMap requires the SNP ratio to be exactly 1; consequently, the phenotype of the mutant must be 100% correct. To overcome these potential pitfalls in MutMap, Wang et al. (2021) developed a web-based workflow, BSAsseq, an interactive and integrated bioinformatic pipeline for identification of causal mutations (Fig. 1). The first step of BSAsseq is to cross a mutant of interest to the WT parent. Next, the F₁ plants are self-pollinated to produce an F₂ generation. After confirming that the mutant segregation ratio is consistent with a single recessive Mendelian trait at a ratio of approximately 1:3, the genomic DNA is extracted from a pooled sample of 20 homozygous mutants and sequenced to 15× coverage of the genome by pair-end sequencing on Illuminia platforms. Short reads are aligned to the reference

assembly using Bowtie2 (Langmead and Salzberg 2012), and SNPs are called and filtered using Bcftools (Li 2011) to keep the canonical EMS-induced mutations (G → A or C → T) with the desired coverage (e.g., 5–100). SnpEff (Cingolani et al. 2012) is used to annotate and select the mutations with a large predicted effect (missense, nonsense, splice site acceptor, or splice site donor), and SIFT 4G (Vaser et al. 2016) is used to predict whether the mutations are deleterious to the gene. The SIFT score ranges from 0 to 1 and a score of less than 0.05 is generally considered deleterious. These steps, from aligning the short sequences with reference genes to the prediction of deleterious mutations, are integrated into the BSAsseq pipeline. All that is required of a user is to input the two sequencing files (pair-ends) into the workflow, and the output of the workflow is the BSA-viewer that consists of two interactive graphs. The first graph is the SNP ratio plot, from which one can identify mutations with a SNP ratio close to 1. Within this plot, there are two types of mutations represented by red or gray circles, wherein the red circles are those with SIFT scores < 0.05 and the gray circles are those with SIFT scores > 0.05, or non-deleterious mutations. The other graph is a linkage probability plot, from which one can determine the probability of the candidate mutation with the predicted SNP ratio occurring by chance. The horizontal line on the plot is the significance threshold for this SNP ratio and can be used to ultimately determine if the SNP ratio is significantly different from the ratio of unlinked SNPs (50%). The possible candidate SNP must be above this threshold to be considered significant. The final and most important criterium is the biological knowledge, i.e., the function of the gene that makes biological sense to the phenotype of interest. To further validate if the mutation is indeed the causative mutation for the phenotype of interest, one can search the gene harboring the putative causative

Fig. 1 Illustration of the BSAsseq Workflow. Creation of the BSAsseq mapping population is similar to MutMap. To use the workflow in sorghum, the user only need input the two pair-end sequencing file into the workflow using “Browse” to select the files in the Cyverse Datastore. All other steps are automated. The output is the SNP ratio and linking probability plots, from which, the candidate genes can be selected based on SNP ratio close to 1 in the region with a probability above the cutoff line (the horizontal line in the linking probability plot)



mutation for a second independent allele from the sequenced sorghum mutants (or mutation-indexed mutant library) available on Gramene (<http://www.gramene.org/>). If the mutant line containing the independent mutation in the same gene also segregates for the phenotype of the mutant of interest, the causative mutations is then confirmed.

Compared with MutMap, the BSAsseq requires pooling of only 20 F_2 mutants, a sequencing depth of 15 \times coverage, and is tolerant to the inclusion a few mis-phenotyped F_2 plants. At current sequencing prices, it takes approximately \$200 USD to sequence and analyze one bulked F_2 population. The BSAsseq workflow is publicly available at (<https://www.sciapps.org/>), and a free account at Cyverse to hold the sequencing data and access the workflow can be obtained. Once the account is activated, the user only needs to upload the two pair-sequencing data to Cyverse Datastore using the free data transfer software Cyberduck (<https://cyberduck.io/>) or IRODS (<https://irods.org/>). In addition, an instructional video is available online to help users learn to use the BSAsseq workflow (<https://www.sciapps.org/>). Because there are resources and tutorials available to users, no prior training in bioinformatics techniques is required for a researcher to use this workflow. Although the workflow is designed with data obtained from the sorghum pedigreed mutant library, it can be easily applied to other crops by replacing the sorghum reference genome with the reference genome of the crop from which the mutants are selected.

Reverse genetics—identification of mutation series from sequenced mutant library

Identification of mutation series in a gene with known sequence to deduce its function by studying the phenotypes of the series mutants is an important genetic research approach called reverse genetics (McCallum et al. 2000a; Till et al. 2003; Wienholds et al. 2003; Gilchrist and Haughn 2005; Winkler et al. 2005; Gilchrist et al. 2006; Xin et al. 2008; Tsai et al. 2011). A technique, dubbed as TILLING (Targeting Induced Local Lesions IN Genomes (TILLING) (McCallum et al. 2000a, b), is usually used to identify the mutant series. The TILLING technique begins with the development of a mutant population by chemical mutagenesis (McCallum et al. 2000b). Genomic DNA is extracted from individual M_2 plants and pooled by one or multiple dimensions with no more 8 samples per pool. PCR primers are designed to cover the regions of interest. After the PCR amplification of the targeted regions, the heteroduplex formation through denaturation and re-annealing, the mutation within the amplified region is detected through cleavage of the heteroduplex by the endonucleases that can recognize single base pair mismatch, such as Cell (Till et al. 2003).

In TILLING, each pair of primers requires meticulous optimization. Thus, it can be cost-prohibitive to analyze a large number of genes.

Due to the increased output of high-quality DNA sequences and the decreased price of next-generation sequencing technologies, it is possible to conduct whole-genome sequencing of the pedigreed mutants as a resource for reverse genetics by searching gene mutations online. Jiao et al. (2016) sequenced 256 mutant lines to an average coverage of 16 \times of the genome, an effort that produced 1.8 million EMS-induced mutations (i.e., G \rightarrow C or A \rightarrow T mutations). The average mutation rate is 11.2/Mb, or about one SNP per 344 bp in the 256 lines. All 10 chromosomes were evenly covered with mutations (Fig. 2). About 236,000 (6.2%) of the SNPs are in 30,294 genes, which covers about 92% of the genes in the sorghum genome (Table 1). A total of 111,850 non-synonymous SNPs are in the exons of 25,605 (77.5%) genes, with an average of 4 mutations per gene. There were 8043 stop-gain or splice junction mutations that potentially produce a knockout mutation of the genes. Addo-Quaye et al (2018) sequenced 586 mutants and discovered 1,275,872 homozygous and 477,531 heterozygous EMS-induced mutations (Addo-Quaye et al. 2017, 2018). There are 56,514 homozygous missense mutations in 23,227 genes, among which 4035 are high-impact homozygous mutations in 3637 genes, and each of these sequencing results can be searched on Gramene (<http://www.gramene.org/>). A collaboration is currently underway with Joint Genome Institute (JGI) of the Department of Energy to sequence additional 1000 pedigreed mutant lines. Based on two published sequencing results, we would obtain a total of > 10 million canonical EMS-induced mutations after completion of the sequencing of additional 1000 lines. More importantly, we would obtain over 150,000 deleterious mutations (SIFT score < 0.05), including > 35,000 knockout mutations (stop-gained and splice junction mutations). With these collections of sequenced mutants, most genes in the sorghum genome would have knockout or deleterious mutations, making the sequenced data a useful resource for reverse genetics; Moreover, these collections could allow for the validation of candidate genes in biparental and genome-wide association mapping of QTLs of important agronomic traits and for the identification of additional independent alleles for genes predicted from the BSAsseq.

New dwarf mutants

The semi-dwarf trait is generally considered the single most impactful trait in the development of modern high yielding crops. The wave of adoption of semi-dwarf genes in plant breeding that was coupled with increased usage of irrigation and chemical fertilizers, a period known as the “Green

Fig. 2 Distribution of EMS-induced mutations. The circle b is the distribution of the EMS-induced mutations (SNPs), which evenly cover the 10 chromosomes except the centromere regions

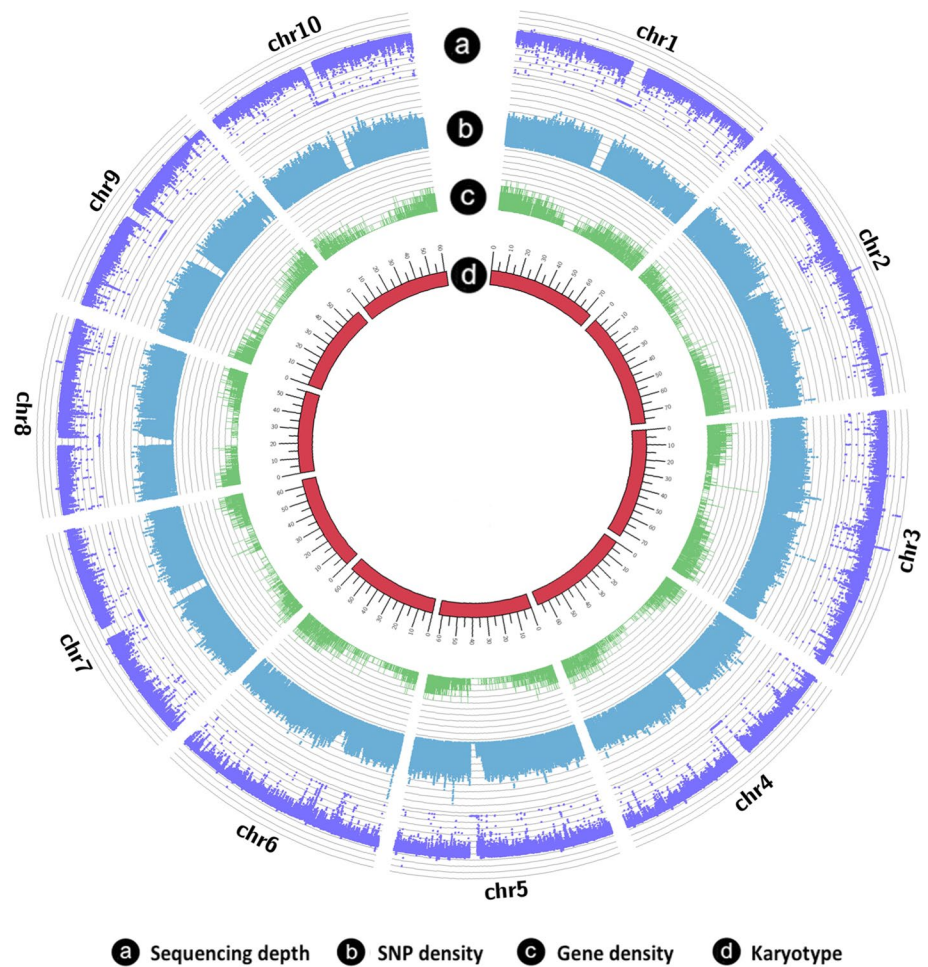


Table 1 Annotation of mutations

Effect type	SNPs	Genes
UTR 3'	44,120	18,772
Start gained	5237	4326
UTR 5'	29,452	14,055
Non-synonymous coding	86,070	25,311
Start lost	136	136
Stop gained	4652	4043
Synonymous coding	138,416	22,035
Splice site acceptor	862	834
Splice site donor	636	621
Splice site region	6894	5260
Intron	195,916	21,443
Upstream 5 kb	480,763	31,358
Downstream 5 kb	464,091	31,141
Intergenic	1,457,491	–

The EMS-induced mutations (SNPs) were classified according to their location and potential effect on gene expression and protein structure

Revolution”, dramatically increased world grain production and successfully averted a catastrophic food shortage that was predicted would occur (Hedden 2003). Unlike wheat (*Triticum aestivum* L.) and rice (*Oryza sativa* L.), which require only a single dwarf locus to breed for semi-dwarf plants (Peng et al. 1999; Sasaki et al. 2002; Wang et al. 2017), sorghum requires multiple dwarf loci stacked together to breed semi-dwarf plants because any single locus is insufficient to create the desired reduction in height (Quinby and Karper 1954). All four known dwarf loci used since the beginning of breeding semi-dwarf sorghum can be traced back to the same sources described in the early 1950s (Quinby 1975; Quinby and Karper 1954), and the causal gene mutations for three of the four dwarf loci have since been identified. The *dw3* locus encodes P-glycoprotein modulating auxin polar transport (Multani et al. 2003). The *dw1* locus encodes a novel protein (Hilley et al. 2016; Yamaguchi et al. 2016). It has recently been suggested that *dw1* may play a role in the brassinosteroid signaling pathway (Hirano et al. 2017). The *dw2* locus encodes a protein kinase, but the signal pathway in which the *DW2* gene functions is not known (Hilley et al. 2017). The gene corresponding to the

dw4 locus has not been identified at the time of this writing. In addition to the four known dwarf loci, other dwarf loci or modifying loci are also inadvertently selected during the process of sorghum improvement (Hashimoto et al. 2021). An in-depth analysis of plant height variations in the SAP with known dwarf loci as covariates shows that at least 10 loci are involved in sorghum height variation (Li et al. 2015). At this time, none of the dwarf loci currently used in sorghum breeding are mapped near the canonical Green Revolution dwarf genes that have been identified in other cereal crops (Wang et al. 2017).

To make semi-dwarf hybrids, both seed parent and pollinator parents must have the same set of dwarf genes, which makes it likely that both parents may carry the same genome segments linked to the dwarf loci. There are strong linkages between the dwarf loci and various maturity loci that are required for flowering in temperate regions, which can make it difficult to utilize dwarf genes in sorghum. The large scale conversion of exotic germplasm from Africa into semi-dwarf and early maturing (photoperiod-insensitive) lines, known as the Sorghum Conversion Program, created thousands of new lines with unique traits and provided the sorghum industry with a large supply of raw breeding material (Stephens et al. 1967). The efficiency of the Sorghum Conversion Program is retrospectively evaluated with over 50,000 SNP markers for 800 pairs of converted lines and their original exotic parents (Thurber et al. 2013). BTx406, a four-dwarf line, is the donor parent for dwarf and maturity loci in most of the Sorghum Conversion Program. One surprising discovery from this retrospective analysis is that large segments of chromosomes 1, 2, 9, 10, and nearly all of the markers on chromosome 6 are consistent with the donor parent BTx406 (Thurber et al. 2013). These homozygous regions may be present in most sorghum hybrids, and this could be a major contributing factor to the prevention of full expression of hybrid vigor in sorghum.

Researchers and plant breeders may wonder why improvement in sorghum requires pyramiding three or more recessive dwarf loci to breed semi-dwarf varieties while wheat and rice only require a single locus to bring plants to their desired height. One hypothesis to explain this key difference could be that researchers have not yet found the optimal dwarf gene for breeding dwarf sorghum. As noted earlier, none of the dwarf genes that have been identified in sorghum appear to be involved in gibberellic acid (GA) signaling or biosynthesis, while the critical genes in other cereal crops are involved in these processes. Since none of the currently known dwarf genes can singularly reduce the height of sorghum to desired levels, breeders must stack multiple dwarf loci until the plant reaches the desired height for combine harvesting. Unfortunately, pyramiding multiple recessive dwarf genes in sorghum is complicated, expensive, and introduces the possibility of reduced hybrid vigor.

We have initiated a systematic approach to identify superior dwarf mutations that can be used alone to breed semi-dwarf sorghum varieties. Dwarf mutant phenotypes are frequently observed in the mutant library; indeed, from approximately 6400 pedigreed lines, over 300 independent dwarf mutants have been isolated (Fig. 3). Some of the dwarf mutants exhibit valuable agronomic traits, such as large panicles and increased seed size, and we are currently in the process of identifying the causative mutations for these dwarf mutants. The genes involved in GA signaling or biosynthesis and the genes identified from mutants that have good agronomic traits will be prioritized for further characterization. We will introduce these mutations into tall sorghum lines that currently have no dwarf genes to determine if any of these mutations can be used to reduce resulting sorghum hybrids to a desirable height without needing to stack multiple dwarf loci.

Several mutations in genes involved GA signaling or biosynthesis were identified from 256 sequenced lines; however, they either have no immediately apparent effect

Fig. 3 Phenotypes of a few dwarf mutants identified from the sorghum pedigreed mutant library. The wild-type BTx623 carries three homozygous conventional dwarf loci *dw1*, *dw3*, *dw4*. The new dwarf mutants carry a novel dwarf locus in addition to the three conventional loci



BTx623

Representative dwarf mutants isolated from the mutant library

on height or are too short to be considered useful for breeding purposes. Fortunately, efforts are underway to sequence an additional 1000 mutant lines, which may reveal leaky mutations in these genes which can alter function enough to achieve a more ideal plant height.

Mutants with altered tolerance to high temperature

High-temperature (HT) stress negatively impacts crop yield worldwide. It is urgent to enhance HT tolerance in crops due to the escalated trend of global warming. It has been well established in model plants that HT has many negative impacts on plants (Burke and Chen 2006; Iba 2002; Kotak et al. 2007; Larkindale et al. 2005; Djanaguiraman et al. 2014). Mechanisms of HT tolerance discovered from model organisms include heat shock proteins (HSPs) (Gurley 2000; Hong and Vierling 2000; Nieto-Sotelo et al. 2002), membrane stability under HT through adjustments of membrane lipid composition and fatty acid unsaturation levels (Alfonso et al. 2001; Chen et al. 2006b; Falcone et al. 2004; Marcum 1998; Sung et al. 2003), and production of antioxidants (Burke and Chen 2006; Chen et al. 2006a; Larkindale and Huang 2004; Wang et al. 2006).

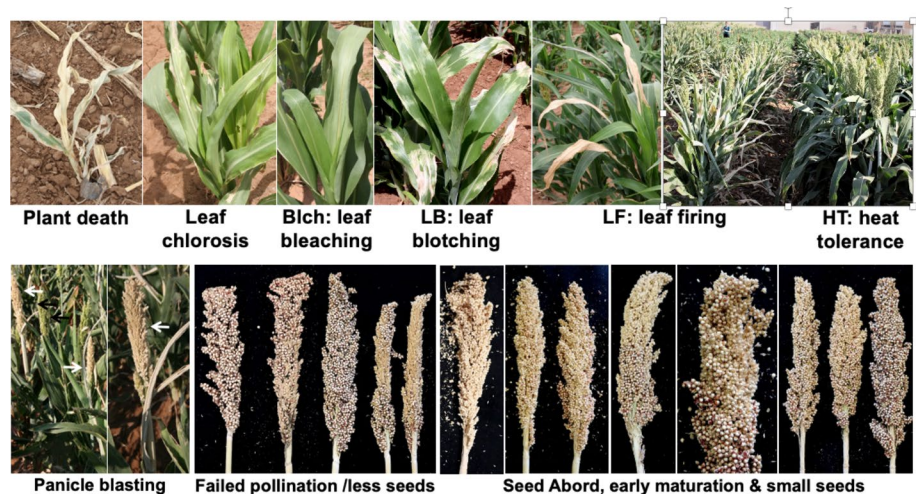
Crops under field conditions often experience a combination of HT and drought stresses. Little is known about the mechanisms of how field crops cope with HT stress, probably more complex than that observed from model organisms under controlled laboratory conditions (Chen et al. 2012). Chen et al. (2010) observed that all known mechanisms of HT tolerance, such as expression of heat shock proteins, are very similar between two maize lines with contrast HT tolerance in the field under both normal or HT stress. Instead, the two maize lines differ significantly in phosphatidic acid (PA), a minor phospholipid acting as a signal molecule in

plants, before or after a high-temperature stress (Chen et al. 2010).

The hot summer seasons that often occur in Lubbock, Texas provide an appropriate environment to evaluate HT injury in maize. Maize plants display a diverse HT injury phenotype from leaf firing (death of leaves), tassel blast, to reduction in pollen production (Chen et al. 2012). However, sorghum, one of the most HT tolerant crops, shows little damage under similar environmental conditions (Chen et al. 2017). Thus, sorghum could serve as an excellent source for isolating HT tolerant genes that can subsequently be used to improve HT tolerance in crops prone to HT injury.

The sorghum mutant library has displayed a variety of HT injury phenotypes that nearly phenocopy the heat-sensitive phenotypes observed in field grown maize (Chen et al. 2012). These phenotypes include plant death, leaf rolling, leaf bleach/blotch, leaf firing, panicle blasting, reduction in pollen production, and a reduction in seed set and seed size (Fig. 4). Because maize and sorghum share a common ancestor as recently as 11 million years ago (Paterson et al. 2004), the high-temperature-sensitive mutants may serve as tools to dissect the mechanism of high-temperature tolerance of maize under field conditions. For example, if a heat-sensitive mutation in sorghum is mapped close to high-temperature tolerance quantitative trait loci (QTL) in maize, the sorghum mutant may be used to clone the sorghum gene through bulked segregant analysis coupled with next-generation sequencing (BSAseq) (Wang et al. 2021). Once a causative mutation/gene is identified, a natural allele of the gene with enhanced HT tolerance may be searched from the vast collection of sorghum germplasm for improving HT tolerance in sorghum. Similarly, the causative gene(s) can be overexpressed in maize or other crops to improve their HT tolerance.

Fig. 4 Various high-temperature-sensitive phenotypes observed from the mutant library. The top panel shows the high-temperature-sensitive phenotypes during vegetative stage. The bottom panel shows the phenotypes during reproductive stage



Multi-seeded mutants

Grain yield per acre is determined by the number of plants per acre, the number of effective tillers per plant, the number of grains per panicle, and the grain weight, which is usually expressed as the weight of a thousand grains. Grain number per panicle is a major determinant of grain yield in sorghum and other cereal crops (Saeed et al. 1986; Ashikari et al. 2005; Duggan et al. 2000; Reynolds et al. 2009; Richards 2000). The sorghum panicle bears many primary branches, upon which several secondary branches can develop. Sometimes, tertiary branches can even develop from the secondary branches (Brown et al. 2006; Burow et al. 2014). Sorghum panicles produce two types of spikelets: the sessile spikelets that are directly attached to flower branches and the pedicellate spikelets that are attached to flower branches through a short petiole called a pedicel (Walters and Keil 1988). In the wild-type BTx623 and other natural accessions of sorghum, the sessile spikelets bear hermaphrodite flowers with both female and male floral organs that will eventually develop into seeds (Fig. 5). The pedicellate spikelets are either sterile or staminate and will eventually abort before a seed is produced (Karper and Stephens 1936).

We have isolated and characterized a novel class of sorghum mutants, referred to as multi-seeded (*msd*) mutants, in which the developmental arrest of the pedicellate spikelets does not occur. In the *msd* mutants, all sessile or pedicellate spikelets develop hermaphrodite flowers and could produce seeds. In addition, the *msd* mutants have shown an increased number and length of the primary inflorescence branches. Thus, the *msd* mutants have the potential to produce up to

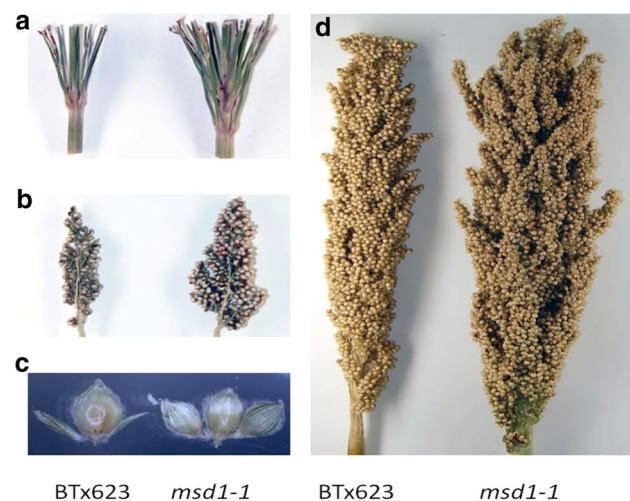


Fig. 5 Panicle architecture of a *msd* mutant. The *msd* mutants has increased size and numbers of primary branches. In addition, both sessile and pedicellate spikelets are fertile and capable of producing seeds

double the grain number per panicle as compared to the wild-type BTx623 (Burow et al. 2014; Jiao et al. 2018). Over 100 independent *msd* mutants have been isolated from the mutant library, and incomplete genetic analyses indicated that these *msd* mutants represent at least 8 complement groups of recessive mutations. In addition, we have isolated several *msd* mutations that are dominant, although further investigation will be required to confirm this and explore its use in sorghum improvement.

Three *MSD* genes have been identified by BSAseq (Wang et al. 2021). The *MSD1* gene (SORBI_3007G135700) encodes the TCP (Teosinte Branched/Cycloidea/proliferating cell nuclear antigen factor) transcription factor, which can activate several enzymes in the jasmonic acid (JA) biosynthetic pathway (Jiao et al. 2018). The *MSD2* gene (SORBI_3006G095600) encodes 13-lipoxygenase (LOX), which catalyzes the conversion of free linolenic acid (18:3, a fatty acid with 18-carbon chain and three double bonds) to hydroperoxy octadecadienoic acids (HPODE), which is the first step of JA biosynthesis (Gladman et al. 2019). The *MSD3* gene (Sobic.3001G407600) encodes a major linoleic (18:2) desaturase FAD7, which is required for biosynthesis of linolenic acid, the substrate for JA biosynthesis (Damp-anaboina et al. 2019). Based on the three *MSD* genes identified, it appears that JA plays an important role in the development of the pedicellate spikelets. Both male and female floral organs initiate normally in the pedicellate spikelets of the wild-type BTx623 and the *msd* mutants (Jiao et al. 2018). These floral organs are subsequently aborted in the wild-type pedicellate spikelets but continue to develop in the *msd* mutants. The current hypothesis is that a developmental signal in the pedicellate spikelets activates *MSD1* shortly after floral organ initiation. Then, the *MSD1* triggers the burst production of JA that ultimately leads to the abortion of the floral organs through JA-mediated programmed cell death. In the pedicellate spikelets of the *msd* mutants, the lack of this burst production of JA cannot activate the programmed cell death pathways, which results in floral organs that continue development into fertile flowers and grains. It is currently unclear whether other processes are involved in the development of the pedicellate spikelets; however, characterization of the other *msd* mutants will likely uncover these processes.

Other mutants

The mutant library contains many other classes of mutants that cannot be fully described in this review. The mutants with altered agronomic traits are listed in Table 2. One class of mutants with erected leaf (*erl*) architecture particularly worth mentioning because the erect leaf architecture in maize has been considered as an important trait

Table 2 A selection of mutants observed from the sorghum pedigreed mutant library

Mutants	Number of lines	Potential application of the trait
Erect leaf (<i>erl</i>)	23	Biomass and grain yield
Brown midrib (<i>bmr</i>)	30	Biomass conversion efficiency
Multiple tiller (<i>mtl</i>)	120	Biomass yield
Early flowering (<i>efl</i>)	7	Biomass and grain yield
Late flowering	5	Biomass yield
Stiff stem (<i>stf</i>)	13	Biomass and grain yield
Early senescence (<i>esn</i>)	7	Early dehydration of biomass
Multiple seed (<i>msd</i>)	> 100	Grain yield
Dominant msd (<i>msd/d</i>)	6	Grain yield
Large seed (<i>lsd</i>)	18	Grain quality
bloomless (<i>blm</i>)	107	Water use efficiency
Heat sensitive (<i>hs</i>)	> 200	Tolerance to high temperature
Nuclear male-sterile	> 200	Develop new hybrid breeding system
Dwarf (<i>dwf</i>)	> 300	New super dwarf to improve hybrid vigor

Updated from Jiao et al. (2016)

leading to the eightfold yield gains after the Green Revolution. Duvick and Cassman (1999) compared 10 morphological and agronomical traits in 36 maize hybrids released from 1936 to 1991 and found that leaf angle score of the new hybrids displayed an improvement of 122% over the old ones (Duvick and Cassman 1999). The modern maize hybrids with more acute (erect) leaf angle can be planted at higher density to capture more solar radiation per unit land area (Duvick and Cassman 1999). Erect leaf mutants in rice is also considered as an important trait to increase biomass and grain yield (Sakamoto et al. 2006).

Compared with modern maize hybrids, sorghum exhibits an open canopy with wide leaf angles that are nearly parallel to the ground. The only reported sorghum mutant that has erect leaf angle is the ligule-less mutant (Singh and Drolsom 1973). Due to the presence of other undesirable agronomic characteristics, this mutant has not been used to improve leaf angle in sorghum breeding programs.

To identify new genetic resources for improving leaf angle, we designed a systematic approach to search for sorghum mutants with erect leaves. Among the 6,400 M₃ plots in the field, over 50 plots segregated for leaf angles that differ from the wild-type BTx623 (Xin et al. 2009). Eleven of these mutants were confirmed in the next generation (M₄) and several mutants have similar or slightly larger panicles as compared to the wild type. Most of the erect leaf mutants are caused by a recessive mutation on a single nuclear gene, and efforts are currently underway to identify the causative mutations for the erect leaf mutants. Once the causative mutations are identified, it should not be difficult to breed the erect leaf phenotype into elite sorghum lines using either marker-assisted backcrossing or genome editing methodologies.

We have also isolated novel brown midrib (*bmr*) mutants that can be used to improve digestibility of sorghum stalks, bloomless (*blm*) mutants that can be used to dissect the mechanisms of drought tolerance and high water use efficiency of sorghum, mutants with altered root morphology, nuclear male-sterile mutants for developing a two-line breeding system, and other mutants that could be useful in future studies (Chen et al. 2019a; Jiao et al. 2017; Saballos et al. 2012; Sattler et al. 2014; Scully et al. 2016; Tishchenko et al. 2020; Xin et al. 2017). The mutant library may possess many other valuable mutations that could eventually be identified and exploited to introduce novel traits for genomic studies and breeding.

Genome editing and mutagenesis

Since the publication of a seminal study using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated (Cas) systems to induce mutations in genomes at precise locations, genome editing technology has advanced rapidly (Jinek et al. 2012). With these improved genome editing technologies, many mutations of any type can be introduced into genomes at pre-determined positions simultaneously (Anzalone et al. 2020, 2019; Zuo et al. 2019). Genome-editing technologies promise to transform breeding in both animals and plants (Chen et al. 2019b). The first genome editing in sorghum was performed on immature embryos using CRISPR/Cas9/sgRNA-mediated targeted gene modification but did not produce stably-edited plants (Jiang et al. 2013). Afterward, several studies of genome editing have been reported on sorghum (Li et al. 2018; Char et al. 2020; Liu et al. 2019; Che et al. 2018).

Although transformation techniques have been developed to apply genome-editing in a few sorghum lines, only RTx430 and P898012 have reasonable transformation efficiency (Gurel et al. 2012; Liu and Godwin 2012). Development of genotype-independent transformation methods remains the primary obstacle to applying genome editing technologies in sorghum.

Although genome-editing tools can almost introduce any type of mutations in genomes at precise locations, it is critical to know the target gene to be edited, or ideally, the causal mutations in the gene to achieve successful outcome quickly. With the exciting development of improved editing technologies and new ways to apply these technologies that circumvent tissue culture, it is still a challenge to edit all genes in a genome in multiple alterations to determine the impact of these alterations on plant health, productivity, and quality (Wang et al. 2021). On the other hand, mutant libraries derived from induced mutagenesis provide efficient resources to identify the causal mutations through reverse and forward genetics. Given each mutant line carrying thousands of induced mutations (Jiao et al. 2016), the desired traits can be obtained by screening a limited number of mutant lines and in a reasonable time frame (Addo-Quaye et al. 2018; Jiao et al. 2016). The causal mutation underlying the traits of interesting can be identified with BSAsseq at a cost less than \$200 in sorghum. Although the BSAsseq is developed in sorghum, it can be used in other crops by simply changing the reference genome (Wang et al. 2021). We anticipate that BSAsseq will become even more affordable for identifying causative mutations, providing promising targets for genome editing. However, induced mutant populations often have a high density of background mutations (Addo-Quaye et al. 2018; Jiao et al. 2016), which must be removed before the mutants are used in breeding. Although recurrent backcrosses can remove unlinked background mutations, it will take several generations to remove 90% of the unlinked mutations. Furthermore, no effective genetic method currently exists for removing the linked mutations. On the other hand, genome editing can introduce precise mutations with few or no offsite mutations, allowing rapid introduction of superior traits into elite germplasm for breeding. The combination of affordable and fast target gene discovery using BSAsseq with precise genome editing will truly revolutionize plant breeding.

Genome sequences

Genome sequences are an important resource for understanding gene contents and functions and are critical to many genomic studies. The first sorghum genome sequence was completed in an elite grain sorghum inbred line, BTx623, over 12 years ago (Paterson et al. 2009). This genome has

served as the reference or “blueprint” for genome-wide analyses to identify genetic variations from other sorghum accessions. For comparative genomics, two sweet sorghums (Keller and E-tian) and one grain sorghum (Ji2731) were sequenced using the whole-genome shotgun strategy and Illumina Genome Analyzer sequencing technology and compared with BTx623 by a group of scientists in China (Zheng et al. 2011). A large number of genetic variations were identified among these four genomes including 1,057,018 single-nucleotide polymorphisms (SNPs), 99,948 insertion/deletions (Indels), 16,487 present/absent variations (PAVs), and 17,111 copy number variations (CNVs). Variations in 1442 genes were identified between the sweet and grain sorghums. Sweet sorghums are enriched with genes related to starch and sucrose metabolism as well as lignin- and coumarine-biosynthesis pathways. Subsequently, 44 sorghum accessions, including a wild relative (*Sorghum propinquum*) were sequenced (Mace et al. 2013b). This effort identified untapped genetic potential in sorghum accessions indigenous to the African continent.

Ten years after the first sorghum reference genome was published, a high-quality reference genome sequence (version 3) was assembled with an integrated approach of deep sequencing, genetic linkage analysis, and transcriptome data. Compared with the original reference genome, the new reference genome size added 29.6 Mb, the number of annotated genes increased 24% to 34,211, and the sequencing error rate was reduced tenfold to 1 per 100 kbp (McCormick et al. 2018). Thanks to the development of new technologies, a chromosome-level de novo assembly was generated from a sorghum inbred line, Tx430 (Deschamps et al. 2018). In this genome assembly, 90% of the genome was covered with merely 29 scaffolds. A chromosome-level reference genome has also been generated in a sweet sorghum Rio (Cooper et al. 2019). Despite the apparent difference in agronomic traits between sweet and grain sorghums, the genomes are highly similar in structure (Cooper et al. 2019). Comparative transcriptomics has indicated that the variation in expression of genes related to high stalk sugar content, changes in the activity and localization of transporters, and the timing of sugar metabolism all may play a critical role in sweet sorghum.

Compared with high-quality genome sequencing from individual lines, pan-genome sequences, which consist of genome sequences assembled de novo from many representative genetic diverse lines, can capture the full genomic diversity of a species. Tao et al. (2021) performed a de novo assembly of 13 diverse lines representing the cultivated sorghum and its wild relatives. Along with the publicly available genome sequences, 16 genome sequences were used to construct the sorghum pan-genome sequence. This pan-genome sequence has 955 Mb and contains 44,079 gene families. When the other 15 genome assemblies were

anchored to the reference BTx623 genome, 0.3 to 0.5 million insertions/deletions and 15,293,465 SNPs were identified. Gene copy number variations were also identified in 429 and 1118 genes. Surprisingly, 64% of gene families show presence/absence variation (PAV) among genomes. Within this sorghum pan-genome sequence, the genes are classified as either core genes that are present in at least 15 genomes, shell genes that are present in 2–14 genomes, or cloud genes that are present in only one genome. In sorghum, core genes only represent 36% (15,867), shell genes represent 63.6% (28,026), and cloud genes represent 0.4% (186) of the total gene families. The proportion of dispensable genes (shell plus cloud) in the sorghum pan-genome is much higher than that reported in *Oryza sativa* (54%), *Glycine max* (49%) and *Brachypodium distachyon* (45%), indicating that sorghum may have greater genetic diversity.

RNA-seq data and functional enrichment analyses indicate that core genes are enriched in basic and critical functions (such as developing process, RNA processing, reproductive system development, leaf and seed development, cell differentiation, and chloroplast organization), while dispensable genes are enriched in adaptive biological processes (such as secondary metabolic processes, cellular metabolic processes, and amino acid transport). The pan-genome sequence provides the sorghum community a solid genetic foundation for understanding trait biology and a platform for mining genetic diversity for future sorghum improvement and gene discovery.

Sorghum germplasm

World collections and gaps

Sorghum has a rich worldwide collection of conserved germplasm, with at least 20 sorghum gene banks in the world, among which are four major centers. The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India holds 37,949 accessions collected from 92 countries (ICRISAT 2021). The United States Department of Agriculture (USDA) National Plant Germplasm System (NPGS) maintains over 45,000 accessions at the Plant Genetic Resources Conservation Unit in Griffin, GA (USDA 2021). The Institute of Crop Science, Chinese Academy of Agricultural Sciences (ICS-CAAS), China, holds 18,263 accessions. The National Bureau of Plant Genetic Resources (NBPGR) of India holds 20,221 accessions (NBPGR 2021). In total, over 240,000 accessions are safeguarded in ex-situ gene banks. Cultivated accessions and wild weedy relatives account for 98.3% and 1.7% of the collection, respectively (Upadhyaya et al. 2016). These collections may have redundancies because of the frequency of exchange among the gene banks; however, the level of this redundancy has yet

to be fully understood. To increase the efficiency of conservation and utilization, the redundant accessions must be identified in future studies.

Although many sorghum accessions have been collected and deposited in the world gene banks, geographical and taxonomic gaps still exist. Among 6415 accessions from West and Central Africa (WCA) analyzed with passport and characterization data, clear geographical gaps were identified in Burkina Faso and Nigeria. In addition to these geographical gaps, taxonomic gaps also existed. For example, 22 species within the sorghum genus have been reported, but the WCA collection has only three species (Upadhyaya et al. 2017a). Geographical and taxonomic gaps were also identified at the ICRISAT sorghum gene bank from the South Asia collection (Upadhyaya et al. 2017b). Within the sorghum genus, there are five subgenera. *Sorghum bicolor* belongs to the subgenera of *Eu Sorghum*, within which there are three species: *Sorghum bicolor*, *Sorghum halepense* (Johnson grass), and *Sorghum propinquum*. Within the *S. bicolor* species, there are three subspecies: *subsp. bicolor*, *subsp. verticilliflorum*, and *subsp. drummondii* (Sudan grass). The *subsp. bicolor* has five races (bicolor, caudatum, durra, guinea, and kafir) and ten intermediate races (Ananda et al. 2020; Harlan and Wet 1972; Lazarides et al. 1991). There are about 24 species within the *Sorghum* genus, 17 of which are found in Australia. As with other major cereal crops, use of the sorghum wild relatives (CWR) in breeding is a promising approach for enhancing the genetic diversity and sorghum yield (Mammadov et al. 2018). Due to human activities and their impacts, such as urbanization, cultivation of new varieties, and climate change, the habitats of some sorghum wild species are disappearing (CIAT 2020). To prevent the permanent loss of sorghum wild species, collection activities are needed in regions of Africa and Australia.

Germplasm preservation

Proper storage conditions help to ensure the optimal preservation of sorghum germplasm. The USDA NPGS (National Plant Germplasm System) sorghum collection is divided into multiple inventories. Distribution inventories are stored at 4 °C with 25% humidity while long-term inventories are stored at –18 °C in sealed foil bags. A back-up inventory is stored at the USDA National Laboratory for Germplasm Preservation in Fort Collins, Colorado to safeguard germplasm against unforeseeable events, such as weather-related disasters, that may impact the active collection stored in Griffin, GA. A select number of accessions (~20%) are further safeguarded in the Svalbard Global Seed Vault in Norway.

Germplasm regenerations are required to maintain viable seed in these collections. Viability testing is conducted periodically to monitor seed health and accessions low in seed

quantity or with reduced germination rates or viability are prioritized for regeneration. For the USDA collection, regenerations are conducted in Mayaguez, Puerto Rico annually as this location allows for flowering of short-day sorghum accessions. To prevent cross-pollination, the panicles are bagged prior to anthesis. During regeneration, basic descriptor data are recorded when possible and made available on the Germplasm Resources Information Network (GRIN)-Global for the USDA, NPGS collection.

Core and mini-core collections

To facilitate the evaluation of plant germplasm collections, the core collection concept was adopted from studies in humans (Frankel 1984). A core collection can be developed for any species and typically represents 10% of the entire collection (Brown 1989a, b). Sorghum core collections have been established for the ICRISAT and USDA NPGS collections. The first ICRISAT sorghum core collection (3475 accessions) was established using seven morphological traits from 33,100 accessions (Rao et al. 1996). The second core collection (2247 accessions) was established using photoperiod-sensitivity grouping and logarithmic random sampling from 22,473 landraces at ICRISAT (Grenier et al. 2001a, b). The USDA NPGS sorghum core collection (3011 accessions) represents 77 countries and was established by Dahlberg et al. (Dahlberg et al. 2004). The number of accessions in a core collection is greatly reduced and can be useful for large institutions, but the sheer amount of material can still prove challenging for small institutions to handle. Thus, the mini-core concept was postulated, representing 10% of the core collection or 1% of the entire germplasm collection, of a sorghum gene bank (Upadhyaya and Ortiz 2001). The first sorghum mini-core (containing 242 accessions) collection was developed using 21 morpho-agronomic traits and passport information from 2246 accessions of the core collection (Upadhyaya et al. 2009). The sorghum core and mini-core collections have been successfully used for identification of accessions with different traits of interest (Upadhyaya et al. 2016).

Subset collections orientated by geographic locations

For purposes of identifying specific adaptive traits, evaluation of germplasm accessions collected from particular geographic locations may be the most efficient approach, because these accessions have evolved and developed adaptive traits due to selection pressure under local environments. For example, many accessions with superior resistance to Anthracnose disease were identified from a small subset of accessions from the geographic region of Mali where Anthracnose is highly prevalent (Erpelding 2012).

Two subsets from the Ethiopia sorghum germplasm collections at NPGS USDA were constructed and evaluated by two independent research groups. A subset of 374 accessions from Ethiopia was phenotyped and genotyped by sequencing (GBS), and these studies revealed that the NPGS Ethiopia germplasm is comprised of 11 populations with high levels of admixture (Cuevas et al. 2017). A subset of 1,425 Ethiopia landrace accessions was also phenotyped and genotyped by sequencing, and the genetic architecture and natural variation of the Ethiopian germplasm was identified, and the associated SNPs can be used as markers for molecular breeding (Girma et al. 2019). To identify sources of resistance to anthracnose and grain mold, 158 Senegal accessions from NPGS collection of USDA were screened in two locations (Cuevas et al. 2018b). From these screenings, eight accessions were identified as resistant to both diseases and 14 accessions were identified as resistant to grain mold alone. In another study, 55 accessions resistant to anthracnose were identified from 318 accessions of Sudan germplasm curated in the NPGS core collection (Cuevas and Prom 2020).

Recently, a subset of 387 lines of Ethiopian germplasm was assembled using a comprehensive phenotypic and genomic characterization (Girma et al. 2020). First, a subset of 2010 accessions were selected from 9000 total accessions in the Ethiopia sorghum germplasm bank and phenotyped at three locations for six different quantitative traits. Then, 1628 accessions were genotyped by sequencing (GBS), and a high level of genetic diversity and rare natural variation were identified in the collection. Finally, the core subset representing the Ethiopia landrace germplasm was selected and assembled following posteriori grouping of genotypes, genetic clusters, and stratified random sampling using quantitative traits. The above studies demonstrate that core collections from geographical regions are very efficient resources for evaluation and utilization of sorghum germplasm.

Sorghum linkage and association mapping resources

Linkage and association mapping are widely used approaches to understand the genetic basis of natural variation of plant traits (Huang and Han 2014). These approaches are used to characterize the genetic architecture of traits (number of loci, distribution of loci, gene action, linkage, and allele frequency), generate hypotheses on the genes that underlie trait variation, and test hypotheses on the conservation of trait genetics across plant species. The use of natural variation, or standing polymorphisms, for genetic studies has several strengths (Brachi et al. 2011) and can complement mutant analysis. First, natural variation has been filtered by natural and/or artificial selection, so the variants

identified are more likely to have direct ecophysiological relevance and agronomic utility. Second, genetic studies of natural variation can leverage eco-geographic and historical information. Finally, genetic analysis of natural variation can illuminate evolutionary processes that led to extant patterns of variation. Given the great diversity in sorghum germplasm, it is well-suited to natural variation-based approaches (Boyles et al. 2019). Conversely, some challenges of natural variation-based (Morris et al. 2013c) approaches, such as the complexity of the germplasm or the relative subtlety of phenotypic variation, are addressed by mutant analysis.

All genetic mapping approaches depend on recombination to shuffle genomic variation, so that statistical associations of genotypic variation (i.e., markers) with phenotypic variation (i.e., traits) will occur preferentially near the causal variant. Given high-density genome-wide markers, the resolution of mapping depends primarily on the density of recombination (Korte and Farlow 2013). In linkage mapping, crosses among parents are used to generate recombination events, while historical recombination events are used in association mapping. For both approaches, immortalized inbred lines are necessary for mapping of most complex traits, to allow genotypes to be replicated, grown under contrasting environments, and grown in agronomically relevant plant stands. Sorghum is well suited for development of immortalized mapping populations because it can be both readily crossed and self-fertilized (Barnaud et al. 2008). Here, we focus on linkage and association mapping panels that are freely available to the community via the NPGS and describe the marker genotype resources associated with each panel.

Linkage mapping resources

Many linkage mapping studies have been conducted in sorghum in the past 25 years that have investigated the genetic basis of traits, such as height (Hilley et al. 2017; Lin et al. 1995; Yamaguchi et al. 2016), flowering time (Casto et al. 2019; Guitton et al. 2018; Mace et al. 2013a), pigmentation (Wu et al. 2012; Morris et al. 2013b), drought tolerance (Hayes et al. 2016; Subudhi et al. 2000; Tuinstra et al. 1996), cold tolerance (Burow et al. 2011; Knoll et al. 2008; Marla et al. 2019), and disease resistance (Upadhyaya et al. 2013). Unfortunately, much of the germplasm and associated marker data for earlier work is not readily available in public repositories (Subudhi et al. 2000; Tuinstra et al. 1996; Wu et al. 2012). Thankfully, an increasing number of sorghum RIL families are being deposited and made easily accessible to the community via NPGS-GRIN.

In the NPGS-GRIN system, sorghum RIL families are listed under the crop designation of “SORGHUM-GEN-STOCKS”. Four of these RIL families were developed by

the USDA Plant Stress and Germplasm Development unit in Lubbock, Texas. The first of these, BTx623*IS3620C, is derived from a cross between the genome reference line BTx623 and the highly-divergent (and possibly independently domesticated) Nigerian guinea margaritifera line IS3620C (Camjin) (Burow et al. 2011). Given (i) the wide divergence between the parents, (ii) the large family size, and (iii) the availability of high-density GBS markers, this population can be a powerful resource for the dissection of many traits (Kong et al. 2018). In addition, three RIL families developed to study cold tolerance, derived from crosses of cold-susceptible US parent lines with cold-tolerant Chinese parent lines, are currently available from GRIN along with GBS markers (RTx430*Gai Gaoliang, BTx623*Hong Ke Zi, BTx623*Niu Sheng Zi) (Burow et al. 2011; Marla et al. 2019). While these families were primarily developed for studies of cold tolerance, they segregate for many other traits of interest, such as pigmentation, flowering time, plant architecture, and inflorescence architecture (Marla et al. 2019). Other RIL families available from GRIN are a part of the NAM resource described below.

Association mapping resources

With the advent of methods to generate high-density genome-wide markers, much of the focus of genetic mapping in crops has shifted from linkage mapping and candidate gene association to genome-wide association studies (GWAS) (Huang and Han 2014). For the purposes of this review, we will refer to association mapping “panels” rather than populations, to distinguish the germplasm sets from the population they were derived from. The most widely used sorghum GWAS resource is the sorghum association panel (SAP). The SAP global diversity panel with GBS marker data (Hu et al. 2019; Morris et al. 2013b) that is available from GRIN. The SAP was designed to capture sorghum’s global diversity of form, function, and end uses (Casa et al. 2008), and includes predominantly grain sorghum types, along with some sweet sorghums, forages, and broomcorns. Importantly, all SAP lines flower in temperate latitudes, either because they are temperate-adapted genotypes or because they are sorghum conversion (SC) lines, which are tropical accessions converted to photoperiod insensitivity and short stature (Klein et al. 2008; Stephens et al. 1967). Many traits have been subjects of GWAS using the SAP, as recently reviewed (Boyles et al. 2019; Mural et al. 2020).

Another global diversity panel available from GRIN is the bioenergy association panel (BAP) (Brenton et al. 2016), consisting predominantly of photoperiod-sensitive tropical accessions, along with sweet and forage sorghums. The BAP has associated SNP data from GBS and whole-genome resequencing (Bellis et al. 2020; Brenton et al. 2016; Lozano

et al. 2021). A collection of over 2000 georeferenced sorghum landraces, many of which are available from GRIN, was genotyped using GBS and is available for phenotypic and environmental association studies (Bellis et al. 2020; Lasky et al. 2015; Wang et al. 2020). In addition, GBS SNP data are available for a large number of landrace accessions from Africa, many of which are georeferenced, including a core collection for Ethiopia (Cuevas et al. 2017), Sudan (Cuevas and Prom 2020), sweet sorghum (Cuevas et al. 2019b) and the entire GRIN collections from Niger (Maina et al. 2018), Nigeria (Olatoye et al. 2018), and Senegal (Faye et al. 2019).

It is important to note that these resources were developed to capture diversity, not to conform to the assumptions of association mapping (i.e., a lack of population structure), so key caveats apply to any GWAS using these association resources. The availability of genomic resources, such as panels and markers, has made GWAS a popular approach for sorghum in recent years (Boyles et al. 2019; Morris et al. 2013b); however, GWAS may or may not be effective depending on the target traits and germplasm. Because association mapping depends on historical recombination generated by uncontrolled population processes over evolutionary timescales, assortative mating and selection cause the genetic structure of association mapping panels to deviate from what is ideal for association models (Korte and Farlow 2013). Not only can these deviations lead to false positives and false negatives, but the false findings can also be misleading (i.e., more data lead to more statistically significant false results) (Platt et al. 2010).

Under some conditions, general linear models with population terms (i.e., Q) or mixed linear models with kinship terms (i.e., K or Q + K) can account for polygenic background effects and reduce false positive associations while retaining true positives. However, when oligogenic variants (i.e., QTL) are colinear with polygenic and/or neutral background variation (which is presumably common for environmentally adaptive traits) the polygenic term can account for oligogenic variation and leads to false negatives, which has been observed in sorghum (Lasky et al. 2015; Morris et al. 2013c). Given the typical focus on avoiding false positives, the possibility of false negative is arguably an underappreciated challenge of association studies in sorghum.

Interpretation of association mapping requires thorough consideration of the population and evolutionary genetics of the source germplasm, to a greater degree than linkage mapping (Brachi et al. 2011). Those interested only in generating candidate genes for fine mapping may find that a biparental linkage approach avoids the complexity and ambiguity of GWAS interpretation (Casto et al. 2019; Hilley et al. 2017). By contrast, those interested the broader patterns of diversity and the forces that generate them may find that challenges of GWAS interpretation are worth surmounting, as in some

studies of sorghum global adaptation (Lasky et al. 2015; Wang et al. 2020; Wu et al. 2019).

Multi-parent mapping resources

Given the tradeoffs between linkage and association mapping, there has been increasing interest in developing alternative genetic mapping approaches that combine the strengths linkage mapping (power, sensitivity) with those of association mapping (diversity, resolution) (Brachi et al. 2011; Korte and Farlow 2013). Several such multi-parent mapping approaches including nested association mapping (NAM), backcross NAM, and multi-parent advanced generation intercross (MAGIC), have been applied in sorghum (Boyles et al. 2019).

The NAM approach was pioneered in maize and now used in many crops, and it uses multiple RIL families that have diverse founders but share a common parent (Gage et al. 2020; Yu et al. 2008). A global grain sorghum NAM resource is available in sorghum, developed using RTx430 as the common founder line and 10 diverse global founders, for a total of > 2200 RILs in 10 families (Bouchet et al. 2017; Yu et al. 2013). RTx430 was selected as the common parent because it has been the most important public pollinator line in recent decades (Miller 1984; Smith and Frederiksen 2000). Importantly, all the NAM parents are a part of the SAP (Yu et al. 2013), and this overlap of germplasm between these mapping resources facilitates the cross-comparison and validation of findings between NAM and SAP studies (Olatoye et al. 2020b). SNP markers for the sorghum NAM were originally developed using the v.2 BTx623 reference genome (Bouchet et al. 2017), but an updated version of the marker data are now available based on the v.3 BTx623 reference genome (Hu et al. 2019). This NAM resource has been used to map flowering time, height, inflorescence morphology, and vegetative morphology (Bouchet et al. 2017; Hu et al. 2019; Olatoye et al. 2020a, b). Notably, it has been demonstrated that the NAM resource has greater power to map some adaptive traits that would be confounded by population structure in GWAS (Bouchet et al. 2017; Olatoye et al. 2020b).

Another multi-parent mapping approach that balances some tradeoffs of linkage and association mapping is MAGIC (Huang et al. 2015). Like NAM, MAGIC seeks to increase allelic diversity compared to biparental linkage families while increasing power and specificity of QTL detection relative to GWAS. Unlike NAM, however, MAGIC aims to increase recombination and mapping resolution using advanced intercrosses and to balance allele frequencies using equal contributions from all founder lines. To date, one sorghum MAGIC resource has been described, which can be requested from the authors that developed it

(Ongom and Ejeta 2018). The backcross NAM (BCNAM) approach, which uses backcrossing and selection to recover lines that are phenotypically similar to the elite common parent, can have the advantage of greater agronomic and breeding relevance compared to NAM (Jordan et al. 2011).

Future mapping resources needed

The rapid advances in sorghum mapping resources in recent years has certainly facilitated various genetic studies. However, there are two notable gaps that remain. First, causative variants underlying sorghum natural variation have only been identified for a few traits (Casto et al. 2019; Hilley et al. 2017; Wu et al. 2019; Zhang et al. 2018; Zou et al. 2020). Generally, sorghum mapping studies have identified QTL and suggested candidate genes but have not identified the causative variants. Future resource development should concentrate on facilitating identification of causative variants (as opposed to capturing additional diversity). For instance, high-quality whole-genome resequencing or de novo genome sequencing of the NAM parents and the SAP could facilitate the discovery of causative variants from sorghum NAM and SAP-GWAS studies. The second major gap that exists is that there are cases where genetic mapping has led directly to advances in sorghum improvement. This issue is not unique to sorghum (Bernardo 2016), but future efforts may need to focus on approaches designed for direct relevance of sorghum improvement. For example, future initiatives to develop mapping resources may have a greater impact on sorghum improvement if they focus on BCNAM or other approaches that put diversity in the context of elite genetic backgrounds.

Resources in plant diseases resistance

The resilience and adaptability of sorghum to multiple agriculture systems located in the tropical, sub-tropical, and temperate regions of the world result in the crop facing challenges from several different biotic constraints. The different end uses of the crop for food, animal feed, forage, and bioenergy determine the relevance of each disease. Sorghum productivity and profitability worldwide are limited by several diseases, such as anthracnose (*Colletotrichum sublineolum*), stalk rot (*Fusarium tapersinum* and *Macrophomina phaseolina*), head smut (*Sporisorium reilianum*), downy mildew (*Peronosclerospora sorghi*), rust (*Puccinia pupurea*), leaf blight (*Exserohilum turcicum*), and grain mold (multiple fungal species), among others, which reduce biomass and seed yield and quality.

Currently, the preferred strategy to control sorghum diseases is the identification and incorporation of resistance

genes. The narrow genetic diversity among improved sorghum varieties and temperate-adapted germplasm predisposes the crop to a devastating epidemic. Over the last 30 years, multiple disease resistant accessions have been identified that are adapted to tropical and temperate climates; however, the genetic control underlying these resistance responses is unknown for most of these accessions. The recent genomics and germplasm resources available for the crop are providing insights to the genetic control of most important diseases and the molecular tools for the effective use of resistance germplasm through marker-assisted selection.

Disease resistance germplasm in diversity panels

The two most widely used sorghum association panels used to study disease resistance responses are the SAP (Casa et al. 2008) and the ICRISAT mini-core collection (Upadhyaya et al. 2009). Both association panels enclose a large degree of genetic diversity and are genetically characterized based on GBS marker data (Lasky et al. 2015; Morris et al. 2013a) providing the research tools to elucidate disease resistance responses.

Stalk rots, a disease that causes significant damage to root and stalk tissue, was the first one to be studied with the SAP (Adeyanju et al. 2016). The GWAS based on the stalk rots resistance response of 300 accessions and 79,132 SNPs found 14 loci in chromosome 2, 3, 4, 7, 8, and 9. Remarkably, a 2-Mb genomic region at chromosome 9 is associated with stalk rots resistant across multiple environments and explains up to 16% of the observed variation. Anthracnose is a disease that affects both grain and biomass and causes significant yield losses in humid production regions (Thakur and Mathur 2000). The analysis of the disease resistance response of the SAP in Puerto Rico, Georgia, U.S. and Texas, U.S. (Prom et al. 2019; Cuevas et al. 2018a) identified 40 resistant accessions across locations. These resistance accessions clustered within different sorghum races, suggesting the presence of multiple resistance sources. Genome-wide association analysis based on binary resistance response (i.e., resistant and susceptible) identified three loci at the distal region of chromosome 5 (Cuevas et al. 2018b). These three loci explained only a limited portion of the phenotypic variation, indicating the presence of other resistance sources that have not been detected due to their low frequency or due to an overcorrection for population structure. Grain mold is a disease caused by multiple pathogenic and opportunistic fungi that affect the plant from anthesis to harvest and reduce the grain yield and quality (Bandyopadhyay et al. 2000). The SAP was evaluated for grain mold resistance in Puerto Rico (Cuevas et al. 2019a) and Texas, U.S. (Prom et al. 2020) over three consecutive years and 18 and 3 highly resistance accessions were identified, respectively.

Three loci in chromosomes 1, 8, and 10 were associated to resistance response in Puerto Rico, while in Texas, U.S. multiple candidate genes were identified based on the top 10 significance SNPs. Pathogens and environments differing among both locations had large effects upon the resistance response and the GWAS results. These GWAS studies demonstrated the efficacy of SAP for the genomic dissection of disease resistance response but also showed the necessity to use other family-based approaches.

The ICRISAT mini-core collection was first evaluated for grain mold and downy mildew resistance response in India (Sharma et al. 2010). This two-year screening identified 50 and six accessions resistance to grain mold and downy mildew, respectively. Likewise, the evaluation for resistance to anthracnose, leaf blight and rust found that 13, 27, and 6 accessions exhibited resistance response to these diseases, respectively (Sharma et al. 2012). A subsequent GWAS analysis using 14,739 SNPs and the anthracnose resistance response identified 8 resistance loci and several candidate genes related to the plant immune system (Upadhyaya et al. 2013). The mini-core was also evaluated for anthracnose, downy mildew, and head smut resistance response against multiple U.S. isolates from each disease in greenhouse (Ahn et al. 2019). Nevertheless, the GWAS analysis could not found significant associations with the variation in disease resistance. Most of the accessions in the ICRISAT mini-core are originally from tropical regions; therefore, its evaluation in temperate regions might be limited to greenhouse studies due to its photoperiod sensitivity.

Disease resistance germplasm in ex-situ collection

The narrow genetic diversity among temperate-adapted and breeding germplasm encourages the search for new disease resistance sources in tropical germplasm collections. Over the last six years, GBS techniques have been used to genetically characterize collections from Ethiopia, Sudan, Nigeria, and Senegal (Cuevas et al. 2017; Cuevas and Prom 2020; Girma et al. 2019; Olatoye et al. 2018; Faye et al. 2019). Mapping Anthracnose resistance in two core sets from U.S. NPGS Ethiopian and Sudan germplasm collection identified two resistance loci in chromosome 9 and 5, respectively (Cuevas et al. 2019b; Cuevas and Prom 2020). These two loci explain a limited portion of the observed phenotypic variation, and both resistance alleles are present in the SAP. A large GWAS analysis for grain mold resistance using 1425 Ethiopian landraces from the Ethiopian Biodiversity Institute identified a major locus at chromosome 1 (Nida et al. 2019) that was associated with the biosynthetic pathway of flavonoid compounds in the seeds. Gene expression analysis of candidate genes was based on resistant and susceptible temperate-adapted germplasm; thus, this resistant allele is also present in the SAP. While multiple disease

resistance sources have been identified in tropical sorghum germplasm, further research efforts are needed to identify which resistance alleles are present in temperate-adapted and breeding germplasm. Realizing the untapped potential of new resistance alleles in tropical germplasm will require the development of segregating mapping populations using multiple genetically diverse resistance sources to overcome the limitations of the GWAS analysis.

Challenges and further breeding approaches

Multiple disease resistance sources are present in tropical and temperate-adapted germplasm. The genomic resources in both groups of germplasm encourage screening to identify and delimit genomic regions associated with disease resistance responses. Elucidation of the disease resistance sources present in temperate-adapted germplasm is necessary to confirm the identification of new sources in tropical germplasm. In parallel, the screening of tropical germplasm could be focused on new sources of resistance or alleles that could also enhance the genetic diversity of temperate-adapted germplasm. The creation and use of core sets of tropical germplasm might help to reduce the number of accessions that need to be evaluated for disease resistance in the greenhouse and field. Similarly, the improvement and development of new disease screening assays based on *omics* technology (e.g., phenomics and metabolomics) will be relevant for the identification of resistance germplasm.

Sorghum hybrid seeds production

The production of sorghum hybrid seed relies on the cytoplasmic male sterility system (CMS). Three different CMS systems have been identified in sorghum, but the original system, designated as A1, is the most commonly used system for producing commercial sorghum hybrids (Kuhlman et al. 2006). The CMS breeding system consist of three lines and two parental groups (A/B and R). The A/B parental group involves the female parent, or A line (cytoplasm male-sterile), which is pollinated by its isocyttoplasmic maintainer, or B line (normal cytoplasm), to regenerate seeds of the A line (i.e., two lines to maintain cytoplasm male sterility). The R parental group includes the restorer, or R lines, that are used to produce fertile progeny by its cross to the female parent A line (Rooney 2004). Hence, heterosis in sorghum is determined by the complementary genes among the A/B-lines (female) and R lines (restorer/male) parental pools in breeding programs (Mindaye et al. 2015). Genetic studies based upon molecular markers show that the genetic diversity in the R parental gene pool is significantly larger than in the A/B parental gene pool (Menz et al. 2004; Silva et al. 2021). Evidently, this imbalance in diversity between these

two parental groups constrains the development of higher-yielding sorghum hybrids.

Broadening the genetic base of the A/B parental gene pool requires the identification of tropical germplasm that maintains the cytoplasm male sterility. This assignment is determined by testcrossing tropical germplasm to the A line to evaluate the fertility of the F₁ progeny. When the F₁ progeny is male-sterile, the tropical accession is classified as a maintainer (i.e., B line) and assigned to the A/B parental gene pool. In addition to the low frequency of maintainer lines among temperate-adapted and tropical germplasm (Madugula et al. 2018), new maintainer lines also need to be able of produce F₁ progeny male-sterile under broad environments where sorghum breeding is possible. Today, less than 2% of the NPGS tropical germplasm has been classified into the A/B parental gene pools. The integration of multiple *omics* techniques may provide insight into the phenomena of CMS and fertility restoration in crops (Bohra et al. 2016). Nevertheless, it is imperative for sorghum breeding programs to identify additional germplasm belonging to the A/B parental pool to construct favorable heterotic groups that permit the creation of superior hybrids.

Perspectives

New sequencing technologies are continuously developed that achieve higher throughputs with lower associated costs. It is now already possible to map and identify causal genes represented by interesting mutants using BSAsq, and it will soon become routine to rapidly identify causative mutations using the BSAsq workflow that is freely available online. Several SAPs have been established, and it is now feasible to sequence every line in the SAP, producing high density and high quality of DNA markers and greatly enhancing the power of mapping the causative genes under QTL of agronomic traits. In the years ahead, the sequenced pedigree mutant libraries will be crucial to validate the causative genes under QTL. The rapid increase in genomic data will be integrated online in Sorghumbase (Gladman et al. 2021 in this collection issue), providing a valuable resource to accelerate breeding throughput. Mutant libraries, QTL mapping through SAP, and biparental mapping will lead to the discovery of many causative genes and provide valuable potential targets for genome editing. These rapid advances in many frontiers of sorghum research make it impractical, if not impossible, for one laboratory to perfect and specialize in all areas of research. It is therefore certain that a true revolution in sorghum breeding will rely on the close collaboration of the entire global sorghum community.

Author contribution statement ZX conceived the idea. All authors participate the writing and agree with the final draft.

Acknowledgements The authors thank the support of ARS CRIS projects 3096-21000-021-00D, 3096-21000-022-00D and Dr. Doreen Ware for discussion and helpful suggestion to improve the manuscript.

Data availability statements Dataset(s) derived from public resources and made available with the article.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Tamiru M, Innan H, Cano L, Kamoun S, Terauchi R (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30(2):174–178. <https://doi.org/10.1038/nbt.2095>
- Addo-Quaye C, Buescher E, Best N, Chaikam V, Baxter I, Dilkes BP (2017) Forward genetics by sequencing EMS variation-induced inbred lines. *G3* 7(2):413–425. <https://doi.org/10.1534/g3.116.029660>
- Addo-Quaye C, Tuinstra M, Carraro N, Weil C, Dilkes BP (2018) Whole-genome sequence accuracy is improved by replication in a population of mutagenized sorghum. *G3* 8(3):1079–1094. <https://doi.org/10.1534/g3.117.300301>
- Adeyanju A, Yu JM, Little C, Rooney W, Klein P, Burke J, Tesso T (2016) Sorghum RILs segregating for stay-green QTL and leaf dhurrin content show differential reaction to stalk rot diseases. *Crop Sci* 56(6):2895–2903. <https://doi.org/10.2135/cropsci2015.10.0628>
- Ahn E, Hu Z, Perumal R, Prom LK, Odvody G, Upadhyaya HD, Magill C (2019) Genome wide association analysis of sorghum mini core lines regarding anthracnose, downy mildew, and head smut. *PLoS ONE* 14(5):e0216671. <https://doi.org/10.1371/journal.pone.0216671>
- Alfonso M, Yruela I, Almarcegui S, Torrado E, Perez MA, Picorel R (2001) Unusual tolerance to high temperatures in a new herbicide-resistant D1 mutant from *Glycine max* (L.) Merr. cell cultures deficient in fatty acid desaturation. *Planta* 212(4):573–582. <https://doi.org/10.1007/s004250000421>
- Ananda GKS, Myrans H, Norton SL, Gleadow R, Furtado A, Henry RJ (2020) Wild sorghum as a promising resource for crop improvement. *Front Plant Sci* 11:1108. <https://doi.org/10.3389/fpls.2020.01108>
- Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR (2019) Search-and-replace genome editing without double-strand breaks

- or donor DNA. *Nature* 576(7785):149–157. <https://doi.org/10.1038/s41586-019-1711-4>
- Anzalone AV, Koblan LW, Liu DR (2020) Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* 38(7):824–844. <https://doi.org/10.1038/s41587-020-0561-9>
- Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. *Science* 309(5735):741–745. <https://doi.org/10.1126/science.1113373>
- Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, McCourt P, Guttman DS (2011) Next-generation mapping of Arabidopsis genes. *Plant J* 67(4):715–725. <https://doi.org/10.1111/j.1365-313X.2011.04619.x>
- Bandyopadhyay R, Butler DR, Chandrashekar A, Reddy RK, Navi SS (2000) Biology, epidemiology and management of sorghum grain mold. In: Chandrashekar ABR, Hall AH (eds) Technical and institutional options for sorghum grain mold management: proceedings of an international consultation, ICRISAT, Patancheru, Andhra Pradesh, India, 18–19 May 2000, pp 34–71
- Barnaud A, Trigueros G, McKey D, Joly HI (2008) High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity* (edinb) 101(5):445–452. <https://doi.org/10.1038/hdy.2008.77>
- Bellis ES, Kelly EA, Lorts CM, Gao H, DeLeo VL, Rouhan G, Buden A, Bhaskara GB, Hu Z, Muscarella R, Timko MP, Nebie B, Runo SM, Chilcoat ND, Juenger TE, Morris GP, dePamphilis CW, Lasky JR (2020) Genomics of sorghum local adaptation to a parasitic plant. *Proc Natl Acad Sci USA* 117(8):4243–4251. <https://doi.org/10.1073/pnas.1908707117>
- Bernardo R (2016) Bandwagons I, too, have known. *Theor Appl Genet* 129(12):2323–2332. <https://doi.org/10.1007/s00122-016-2772-5>
- Bohra A, Jha UC, Adhimoolam P, Bisht D, Singh NP (2016) Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. *Plant Cell Rep* 35(5):967–993. <https://doi.org/10.1007/s00299-016-1949-3>
- Bouchet S, Olatoye MO, Marla SR, Perumal R, Tesso T, Yu J, Tuinstra M, Morris GP (2017) Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. *Genetics* 206(2):573–585. <https://doi.org/10.1534/genetics.116.198499>
- Boyles RE, Brenton ZW, Kresovich S (2019) Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments. *Plant J* 97(1):19–39. <https://doi.org/10.1111/tpj.14113>
- Brachi B, Morris GP, Borevitz JO (2011) Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol* 12(10):232. <https://doi.org/10.1186/gb-2011-12-10-232>
- Brenton ZW, Cooper EA, Myers MT, Boyles RE, Shakoor N, Zielinski KJ, Rauh BL, Bridges WC, Morris GP, Kresovich S (2016) A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. *Genetics* 204(1):21–33. <https://doi.org/10.1534/genetics.115.183947>
- Brown A (1989a) The case for core collections. In: Brown AHD, Frankel OH, Marshall DR, Williams JT (eds) The use of plant genetic resources. Cambridge University Press, Cambridge, pp 136–156
- Brown AHD (1989b) Core collections—a practical approach to genetic-resources management. *Genome* 31(2):818–824. <https://doi.org/10.1139/g89-144>
- Brown PJ, Klein PE, Bortiri E, Acharya CB, Rooney WL, Kresovich S (2006) Inheritance of inflorescence architecture in sorghum. *Theor Appl Genet* 113(5):931–942. <https://doi.org/10.1007/s00122-006-0352-9>
- Burke JJ, Chen J (2006) Changes in cellular and molecular processes in plant adaptation to heat stress. In: Huang B (ed) Plant-environment interactions. CRC Press, Boca Raton, pp 27–46
- Burow G, Burke JJ, Xin ZG, Franks CD (2011) Genetic dissection of early-season cold tolerance in sorghum (*Sorghum bicolor* (L.) Moench). *Mol Breed* 28(3):391–402. <https://doi.org/10.1007/s11032-010-9491-4>
- Burow G, Xin ZG, Hayes C, Burke J (2014) Characterization of a multiseeded (*msd1*) mutant of sorghum for increasing grain yield. *Crop Sci* 54(5):2030–2037. <https://doi.org/10.2135/cropsci2013.08.0566>
- Casa AM, Pressoir G, Brown PJ, Mitchell SE, Rooney WL, Tuinstra MR, Franks CD, Kresovich S (2008) Community resources and strategies for association mapping in sorghum. *Crop Sci* 48(1):30–40. <https://doi.org/10.2135/cropsci2007.02.0080>
- Casto AL, Mattison AJ, Olson SN, Thakran M, Rooney WL, Mullet JE (2019) Maturity2, a novel regulator of flowering time in *Sorghum bicolor*, increases expression of SbPRR37 and SbCO in long days delaying flowering. *PLoS ONE* 14(4):e0212154. <https://doi.org/10.1371/journal.pone.0212154>
- Char SN, Wei J, Mu Q, Li X, Zhang ZJ, Yu J, Yang B (2020) An Agrobacterium-delivered CRISPR/Cas9 system for targeted mutagenesis in sorghum. *Plant Biotechnol J* 18(2):319–321. <https://doi.org/10.1111/pbi.13229>
- Che P, Anand A, Wu E, Sander JD, Simon MK, Zhu W, Sigmund AL, Zastrow-Hayes G, Miller M, Liu D, Lawit SJ, Zhao ZY, Albertsen MC, Jones TJ (2018) Developing a flexible, high-efficiency Agrobacterium-mediated sorghum transformation system with broad application. *Plant Biotechnol J* 16(7):1388–1395. <https://doi.org/10.1111/pbi.12879>
- Chen J, Burke JJ, Velten J, Xin Z (2006a) FtsH11 protease plays a critical role in Arabidopsis thermotolerance. *Plant J* 48(1):73–84. <https://doi.org/10.1111/j.1365-313X.2006.02855.x>
- Chen J, Burke JJ, Xin Z, Xu C, Velten J (2006b) Characterization of the Arabidopsis thermosensitive mutant atts02 reveals an important role for galactolipids in thermotolerance. *Plant Cell Environ* 29(7):1437–1448. <https://doi.org/10.1111/j.1365-3040.2006.01527.x>
- Chen JP, Xu WW, Burke JJ, Xin ZG (2010) Role of phosphatidic acid in high temperature tolerance in maize. *Crop Sci* 50(6):2506–2515. <https://doi.org/10.2135/cropsci2009.12.0716>
- Chen J, Xu W, Velten J, Xin Z, Stout J (2012) Characterization of maize inbred lines for drought and heat tolerance. *J Soil Water Conserv* 67(5):354–364. <https://doi.org/10.2489/jswc.67.5.354>
- Chen J, Chopra R, Hayes C, Morris G, Marla S, Burke J, Xin Z, Burow G (2017) Genome-wide association study of developing leaves' heat tolerance during vegetative growth stages in a sorghum association panel. *Plant Genome*. <https://doi.org/10.3835/plantgenome2016.09.0091>
- Chen J, Jiao Y, Laza H, Payton P, Ware D, Xin Z (2019a) Identification of the first nuclear male sterility gene (male-sterile 9) in sorghum. *Plant Genome* 12(3):1–12. <https://doi.org/10.3835/plantgenome2019.03.0020>
- Chen K, Wang Y, Zhang R, Zhang H, Gao C (2019b) CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu Rev Plant Biol* 70(1):667–697. <https://doi.org/10.1146/annurev-arplant-050718-100049>
- CIAT (2020) Surprising trove of sorghum diversity discovered in Australia—but it is disappearing fast. *Research News*, Nov 30, 2020. International Center for Tropical Agriculture. https://www.eurekalert.org/pub_releases/2020-11/icft-sto113020.php
- Cingolani P, Platts A, le Wang L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* (austin) 6(2):80–92. <https://doi.org/10.4161/fly.19695>

- Cooper EA, Brenton ZW, Flinn BS, Jenkins J, Shu S, Flowers D, Luo F, Wang Y, Xia P, Barry K, Daum C, Lipzen A, Yoshinaga Y, Schmutz J, Saski C, Vermerris W, Kresovich S (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 Genom* 20(1):420. <https://doi.org/10.1186/s12864-019-5734-x>
- Cuevas HE, Prom LK (2020) Evaluation of genetic diversity, agronomic traits, and anthracnose resistance in the NPGS Sudan Sorghum Core collection. *BMC Genom* 21(1):88. <https://doi.org/10.1186/s12864-020-6489-0>
- Cuevas HE, Rosa-Valentin G, Hayes CM, Rooney WL, Hoffmann L (2017) Genomic characterization of a core set of the USDA-NPGS Ethiopian sorghum germplasm collection: implications for germplasm conservation, evaluation, and utilization in crop improvement. *BMC Genom* 18(1):108. <https://doi.org/10.1186/s12864-016-3475-7>
- Cuevas HE, Prom LK, Cooper EA, Knoll JE, Ni X (2018a) Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in the US Sorghum Association Panel. *Plant Genome*. <https://doi.org/10.3835/plantgenome2017.11.0099>
- Cuevas HE, Prom LK, Rosa-Valentin G (2018b) Population structure of the NPGS Senegalese sorghum collection and its evaluation to identify new disease resistant genes. *PLoS ONE* 13(2):e0191877. <https://doi.org/10.1371/journal.pone.0191877>
- Cuevas HE, Fermin-Perez RA, Prom LK, Cooper EA, Bean S, Rooney WL (2019a) Genome-wide association mapping of grain mold resistance in the US Sorghum Association Panel. *Plant Genome*. <https://doi.org/10.3835/plantgenome2018.09.0070>
- Cuevas HE, Prom LK, Cruet-Burgos CM (2019b) Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in NPGS Ethiopian sorghum germplasm. *G3* 9(9):2879–2885. <https://doi.org/10.1534/g3.119.400350>
- Dahlberg JA, Burke JJ, Rosenow DT (2004) Development of a sorghum core collection: refinement and evaluation of a subset from Sudan. *Econ Bot* 58(4):556–567. [https://doi.org/10.1663/0013-0001\(2004\)058\[0556:Doasc\]2.0.Co;2](https://doi.org/10.1663/0013-0001(2004)058[0556:Doasc]2.0.Co;2)
- Dampaboina L, Jiao Y, Chen J, Gladman N, Chopra R, Burow G, Hayes C, Christensen SA, Burke J, Ware D, Xin Z (2019) Sorghum MSD3 encodes an omega-3 fatty acid desaturase that increases grain number by reducing jasmonic acid levels. *Int J Mol Sci*. <https://doi.org/10.3390/ijms20215359>
- Deschamps S, Zhang Y, Llaca V, Ye L, Sanyal A, King M, May G, Lin H (2018) A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping. *Nat Commun* 9(1):4844. <https://doi.org/10.1038/s41467-018-07271-1>
- Djanaguiraman M, Prasad PVV, Murugan M, Perumal R, Reddy UK (2014) Physiological differences among sorghum (*Sorghum bicolor* L. Moench) genotypes under high temperature stress. *Environ Exp Bot* 100:43–54. <https://doi.org/10.1016/j.envexpbot.2013.11.013>
- Duggan BL, Domitruk DR, Fowler DB (2000) Yield component variation in winter wheat grown under drought stress. *Can J Plant Sci* 80(4):739–745. <https://doi.org/10.4141/P00-006>
- Duvick DN, Cassman KG (1999) Post-green revolution trends in yield potential of temperate maize in the north-central United States. *Crop Sci* 39(6):1622–1630. <https://doi.org/10.2135/cropsci1999.3961622x>
- Erpelding J (2012) Anthracnose resistance in sorghum germplasm from the Segou region of Mali. *J Crop Improv* 26(3):397–414
- Falcone DL, Ogas JP, Somerville CR (2004) Regulation of membrane fatty acid composition by temperature in mutants of *Arabidopsis* with alterations in membrane lipid composition. *BMC Plant Biol* 4:17. <https://doi.org/10.1186/1471-2229-4-17>
- Faye JM, Maina F, Hu Z, Fonceka D, Cisse N, Morris GP (2019) Genomic signatures of adaptation to Sahelian and Soudanian climates in sorghum landraces of Senegal. *Ecol Evol* 9(10):6038–6051. <https://doi.org/10.1002/ece3.5187>
- Frankel O (1984) Genetic perspectives of germplasm conservation. In: Arber W, Llimensee K, Peacock WJ, Starlinger P (eds) Genetic manipulations: impact on man and society. Cambridge University Press, Cambridge, pp 161–170
- Gage JL, Monier B, Giri A, Buckler ES (2020) Ten years of the maize nested association mapping population: impact, limitations, and future directions. *Plant Cell* 32(7):2083–2093. <https://doi.org/10.1105/tpc.19.00951>
- Gilchrist EJ, Haughn GW (2005) TILLING without a plough: a new method with applications for reverse genetics. *Curr Opin Plant Biol* 8(2):211–215. <https://doi.org/10.1016/j.pbi.2005.01.004>
- Gilchrist EJ, O'Neil NJ, Rose AM, Zetka MC, Haughn GW (2006) TILLING is an effective reverse genetics technique for *Caenorhabditis elegans*. *BMC Genom* 7:262. <https://doi.org/10.1186/1471-2164-7-262>
- Girma G, Nida H, Seyoum A, Mekonen M, Nega A, Lule D, Dessalegn K, Bekele A, Gebreyohannes A, Adeyanju A, Tirfessa A, Ayana G, Taddese T, Mekbib F, Belete K, Tesso T, Ejeta G, Mengiste T (2019) A large-scale genome-wide association analyses of Ethiopian sorghum landrace collection reveal loci associated with important traits. *Front Plant Sci* 10:691. <https://doi.org/10.3389/fpls.2019.00691>
- Girma G, Nida H, Tirfessa A, Lule D, Bejiga T, Seyoum A, Mekonen M, Nega A, Dessalegn K, Birhanu C, Bekele A, Gebreyohannes A, Ayana G, Tesso T, Ejeta G, Mengiste T (2020) A comprehensive phenotypic and genomic characterization of Ethiopian sorghum germplasm defines core collection and reveals rich genetic potential in adaptive traits. *Plant Genome* 13(3):e20055. <https://doi.org/10.1002/tpg2.20055>
- Gladman N, Jiao Y, Lee YK, Zhang L, Chopra R, Regulski M, Burow G, Hayes C, Christensen SA, Dampaboina L, Chen J, Burke J, Ware D, Xin Z (2019) Fertility of pedicellate spikelets in sorghum is controlled by a jasmonic acid regulatory module. *Int J Mol Sci*. <https://doi.org/10.3390/ijms20194951>
- Grenier C, Bramel-Cox P, Hamon P (2001a) Core collection of sorghum: I. Stratification based on eco-geographical data. *Crop Sci* 41(1):234–240
- Grenier C, Hamon P, Bramel-Cox PJ (2001b) Core collection of sorghum: II. Comparison of three random sampling strategies. *Crop Sci* 41(1):241–246. <https://doi.org/10.2135/cropsci2001.411241x>
- Guitton B, Thera K, Tekete ML, Pot D, Kouressy M, Teme N, Rami JF, Vaksman M (2018) Integrating genetic analysis and crop modeling: a major QTL can finely adjust photoperiod-sensitive sorghum flowering. *Field Crop Res* 221:7–18. <https://doi.org/10.1016/j.fcr.2018.02.007>
- Gurel S, Gurel E, Kaur R, Wong J, Meng L, Tan HQ, Lemaux PG (2009) Efficient, reproducible *Agrobacterium*-mediated transformation of sorghum using heat treatment of immature embryos. *Plant Cell Rep* 28(3):429–444. <https://doi.org/10.1007/s00299-008-0655-1>
- Gurel S, Gurel E, Miller TI, Lemaux PG (2012) *Agrobacterium*-mediated transformation of *Sorghum bicolor* using immature embryos. *Methods Mol Biol* 847:109–122. https://doi.org/10.1007/978-1-61779-558-9_10
- Gurley WB (2000) HSP101: a key component for the acquisition of thermotolerance in plants. *Plant Cell* 12(4):457–460. <https://doi.org/10.1105/tpc.12.4.457>
- Harlan JR, Wet JM (1972) A simplified classification of cultivated sorghum. *Crop Sci* 12(2):172–176. <https://doi.org/10.2135/cropsci1972.0011183X001200020005x>
- Hartwig B, James GV, Konrad K, Schneeberger K, Turck F (2012) Fast isogenic mapping-by-sequencing of ethyl methanesulfonate-induced mutant bulks. *Plant Physiol* 160(2):591–600. <https://doi.org/10.1104/pp.112.200311>

- Hashimoto S, Wake T, Nakamura H, Minamiyama M, Araki-Nakamura S, Ohmae-Shinohara K, Koketsu E, Okamura S, Miura K, Kawaguchi H, Kasuga S, Sazuka T (2021) The dominance model for heterosis explains culm length genetics in a hybrid sorghum variety. *Sci Rep* 11(1):4532. <https://doi.org/10.1038/s41598-021-84020-3>
- Hayes CM, Weers BD, Thakran M, Burow G, Xin ZG, Emendack Y, Burke JJ, Rooney WL, Mullet JE (2016) Discovery of a dhurrin QTL in sorghum: co-localization of dhurrin biosynthesis and a novel stay-green QTL. *Crop Sci* 56(1):104–112. <https://doi.org/10.2135/cropsci2015.06.0379>
- Hedden P (2003) The genes of the green revolution. *Trends Genet* 19(1):5–9. [https://doi.org/10.1016/s0168-9525\(02\)00009-4](https://doi.org/10.1016/s0168-9525(02)00009-4)
- Hilley J, Truong S, Olson S, Morishige D, Mullet J (2016) Identification of Dw1, a regulator of sorghum stem internode length. *PLoS ONE* 11(3):e0151271. <https://doi.org/10.1371/journal.pone.0151271>
- Hilley JL, Weers BD, Truong SK, McCormick RF, Mattison AJ, McKinley BA, Morishige DT, Mullet JE (2017) Sorghum Dw2 encodes a protein kinase regulator of stem internode length. *Sci Rep* 7(1):4616. <https://doi.org/10.1038/s41598-017-04609-5>
- Hirano K, Kawamura M, Araki-Nakamura S, Fujimoto H, Ohmae-Shinohara K, Yamaguchi M, Fujii A, Sasaki H, Kasuga S, Sazuka T (2017) Sorghum DW1 positively regulates brassinosteroid signaling by inhibiting the nuclear localization of BRASSINOSTEROID INSENSITIVE 2. *Sci Rep* 7(1):126. <https://doi.org/10.1038/s41598-017-00096-w>
- Hong SW, Vierling E (2000) Mutants of *Arabidopsis thaliana* defective in the acquisition of tolerance to high temperature stress. *Proc Natl Acad Sci USA* 97(8):4392–4397
- Howe A, Sato S, Dweikat I, Fromm M, Clemente T (2006) Rapid and reproducible Agrobacterium-mediated transformation of sorghum. *Plant Cell Rep* 25(8):784–791
- Hu Z, Olatoye MO, Marla S, Morris GP (2019) An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *Plant Genome*. <https://doi.org/10.3835/plantgenom.e2018.06.0044>
- Huang X, Han B (2014) Natural variations and genome-wide association studies in crop plants. *Annu Rev Plant Biol* 65:531–551. <https://doi.org/10.1146/annurev-arplant-050213-035715>
- Huang BE, Verbyla KL, Verbyla AP, Raghavan C, Singh VK, Gaur P, Leung H, Varshney RK, Cavanagh CR (2015) MAGIC populations in crops: current status and future prospects. *Theor Appl Genet* 128(6):999–1017. <https://doi.org/10.1007/s00122-015-2506-0>
- Iba K (2002) Acclimative response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance. *Annu Rev Plant Biol* 53:225–245. <https://doi.org/10.1146/annurev.arplant.53.100201.160729>
- ICRISAT (2021) Genetic resources. International Crops Research Institute for the Semi-Arid Tropics. <http://exploreit.icrisat.org/profile/genetic%20resources/67>. Accessed 30 Mar 2021
- Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, Last RL (2002) *Arabidopsis* map-based cloning in the post-genome era. *Plant Physiol* 129(2):440–450
- Jiang WY, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31(3):233–239. <https://doi.org/10.1038/nbt.2508>
- Jiao Y, Burke J, Chopra R, Burow G, Chen J, Wang B, Hayes C, Emendack Y, Ware D, Xin Z (2016) A sorghum mutant resource as an efficient platform for gene discovery in grasses. *Plant Cell* 28(7):1551–1562. <https://doi.org/10.1105/tpc.16.00373>
- Jiao Y, Burow G, Gladman N, Acosta-Martinez V, Chen J, Burke J, Ware D, Xin Z (2017) Efficient identification of causal mutations through sequencing of bulked F2 from two allelic bloomless mutants of sorghum bicolor. *Front Plant Sci* 8:2267. <https://doi.org/10.3389/fpls.2017.02267>
- Jiao Y, Lee YK, Gladman N, Chopra R, Christensen SA, Regulski M, Burow G, Hayes C, Burke J, Ware D, Xin Z (2018) MSD1 regulates pedicellate spikelet fertility in sorghum through the jasmonic acid pathway. *Nat Commun* 9(1):822. <https://doi.org/10.1038/s41467-018-03238-4>
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816–821. <https://doi.org/10.1126/science.1225829>
- Jordan DR, Mace ES, Cruickshank AW, Hunt CH, Henzell RG (2011) Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. *Crop Sci* 51(4):1444–1457. <https://doi.org/10.2135/cropsci2010.06.0326>
- Karper RE, Stephens JC (1936) Floral abnormalities in sorghum. *J Hered* 27(5):183–194
- Klein RR, Mullet JE, Jordan DR, Miller FR, Rooney WL, Menz MA, Franks CD, Klein PE (2008) The Effect of tropical sorghum conversion and inbred development on genome diversity as revealed by high-resolution genotyping. *Crop Sci* 48(Suppl 1):S12–S26. <https://doi.org/10.2135/cropsci2007.06.0319tpg>
- Knoll J, Gunaratna N, Ejeta G (2008) QTL analysis of early-season cold tolerance in sorghum. *Theor Appl Genet* 116(4):577–587
- Kong W, Kim C, Zhang D, Guo H, Tan X, Jin H, Zhou C, Shuang LS, Goff V, Sezen U, Pierce G, Compton R, Lemke C, Robertson J, Rainville L, Auckland S, Paterson AH (2018) Genotyping by sequencing of 393 sorghum bicolor BTx623 x IS3620C recombinant inbred lines improves sensitivity and resolution of QTL detection. *G3* 8(8):2563–2572. <https://doi.org/10.1534/g3.118.200173>
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29. <https://doi.org/10.1186/1746-4811-9-29>
- Kotak S, Larkindale J, Lee U, von Koskull-Doring P, Vierling E, Scharf KD (2007) Complexity of the heat stress response in plants. *Curr Opin Plant Biol* 10(3):310–316. <https://doi.org/10.1016/j.pbi.2007.04.011>
- Kuhlman LC, Pring DR, Rooney WL, Tang HV (2006) Allelic frequency at the and loci and the genetics of A3 cytoplasmic fertility restoration in converted sorghum lines. *Crop Sci* 46(4):1576–1580. <https://doi.org/10.2135/cropsci2005.10-0380>
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359. <https://doi.org/10.1038/nmeth.1923>
- Larkindale J, Huang B (2004) Thermotolerance and antioxidant systems in *Agrostis stolonifera*: involvement of salicylic acid, abscisic acid, calcium, hydrogen peroxide, and ethylene. *J Plant Physiol* 161(4):405–413. <https://doi.org/10.1078/0176-1617-01239>
- Larkindale J, Hall JD, Knight MR, Vierling E (2005) Heat stress phenotypes of *Arabidopsis* mutants implicate multiple signaling pathways in the acquisition of thermotolerance. *Plant Physiol* 138(2):882–897. <https://doi.org/10.1104/pp.105.062257>
- Lasky JR, Upadhyaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J, Juenger TE, Hyma K, Acharya C, Mitchell SE, Buckler ES, Brenton Z, Kresovich S, Morris GP (2015) Genome-environment associations in sorghum landraces predict adaptive traits. *Sci Adv* 1(6):e1400218. <https://doi.org/10.1126/sciadv.1400218>
- Lazarides M, Hacker J, Andrew M (1991) Taxonomy, cytology and ecology of indigenous Australian sorghums (*Sorghum* Moench: Andropogoneae: Poaceae). *Aust Syst Bot* 4(4):591–635
- Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* 27(21):2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>

- Li X, Li X, Fridman E, Tesso TT, Yu J (2015) Dissecting repulsion linkage in the dwarfing gene *Dw3* region for sorghum plant height provides insights into heterosis. *Proc Natl Acad Sci USA* 112(38):11823–11828. <https://doi.org/10.1073/pnas.1509229112>
- Li A, Jia S, Yobi A, Ge Z, Sato SJ, Zhang C, Angelovici R, Clemente TE, Holding DR (2018) Editing of an alpha-kafirin gene family increases, digestibility and protein quality in sorghum. *Plant Physiol* 177(4):1425–1438. <https://doi.org/10.1104/pp.18.00200>
- Lin Y, Schertz K, Paterson A (1995) Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. *Genetics* 141:391–411
- Liu G, Godwin ID (2012) Highly efficient sorghum transformation. *Plant Cell Rep* 31(6):999–1007. <https://doi.org/10.1007/s00299-011-1218-4>
- Liu G, Li J, Godwin ID (2019) Genome editing by CRISPR/Cas9 in sorghum through biolistic bombardment. *Methods Mol Biol* 1931:169–183. https://doi.org/10.1007/978-1-4939-9039-9_12
- Lozano R, Gazave E, Dos Santos JPR, Stetter MG, Valluru R, Bandillo N, Fernandes SB, Brown PJ, Shakoor N, Mockler TC, Cooper EA, Taylor Perkins M, Buckler ES, Ross-Ibarra J, Gore MA (2021) Comparative evolutionary genetics of deleterious load in sorghum and maize. *Nat Plants* 7(1):17–24. <https://doi.org/10.1038/s41477-020-00834-5>
- Mace ES, Hunt CH, Jordan DR (2013a) Supermodels: sorghum and maize provide mutual insight into the genetics of flowering time. *Theor Appl Genet* 126(5):1377–1395. <https://doi.org/10.1007/s00122-013-2059-z>
- Mace ES, Tai S, Gilding EK, Li Y, Prentis PJ, Bian L, Campbell BC, Hu W, Innes DJ, Han X, Cruickshank A, Dai C, Frere C, Zhang H, Hunt CH, Wang X, Shatte T, Wang M, Su Z, Li J, Lin X, Godwin ID, Jordan DR, Wang J (2013b) Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat Commun* 4:2320. <https://doi.org/10.1038/ncomms3320>
- Madugula P, Uttam AG, Tonapi VA, Ragimasalawada M, Snowdon R (2018) Fine mapping of *Rf2*, a major locus controlling pollen fertility restoration in sorghum A1 cytoplasm, encodes a PPR gene and its validation through expression analysis. *Plant Breed* 137(2):148–161. <https://doi.org/10.1111/pbr.12569>
- Maina F, Bouchet S, Marla SR, Hu Z, Wang J, Mamadou A, Abdou M, Saidou AA, Morris GP (2018) Population genomics of sorghum (*Sorghum bicolor*) across diverse agroclimatic zones of Niger. *Genome* 61(4):223–232. <https://doi.org/10.1139/gen-2017-0131>
- Makita Y, Shimada S, Kawashima M, Kondou-Kuriyama T, Toyoda T, Matsui M (2015) MOROKOSHI: transcriptome database in *Sorghum bicolor*. *Plant Cell Physiol* 56(1):e6. <https://doi.org/10.1093/pcp/pcu187>
- Mammadov J, Buyyarapu R, Guttikonda SK, Parliament K, Abdurakhmonov IY, Kumpatla SP (2018) Wild relatives of maize, rice, cotton, and soybean: treasure troves for tolerance to biotic and abiotic stresses. *Front Plant Sci* 9:886. <https://doi.org/10.3389/fpls.2018.00886>
- Marcum KB (1998) Cell membrane thermostability and whole-plant heat tolerance of Kentucky bluegrass. *Crop Sci* 38(5):1214–1218. <https://doi.org/10.2135/cropsci1998.0011183X003800050017x>
- Marla SR, Burrow G, Chopra R, Hayes C, Olatoye MO, Felderhoff T, Hu Z, Raymundo R, Perumal R, Morris GP (2019) Genetic architecture of chilling tolerance in sorghum dissected with a nested association mapping population. *G3* 9(12):4045–4057. <https://doi.org/10.1534/g3.119.400353>
- McCallum CM, Comai L, Greene EA, Henikoff S (2000a) Targeted screening for induced mutations. *Nat Biotechnol* 18(4):455–457
- McCallum CM, Comai L, Greene EA, Henikoff S (2000b) Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol* 123(2):439–442
- McCormick RF, Truong SK, Sreedasyam A, Jenkins J, Shu S, Sims D, Kennedy M, Amirebrahimi M, Weers BD, McKinley B, Mattison A, Morishige DT, Grimwood J, Schmutz J, Mullet JE (2018) The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J* 93(2):338–354. <https://doi.org/10.1111/tpj.13781>
- Menz MA, Klein RR, Unruh NC, Rooney WL, Klein PE, Mullet JE (2004) Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. *Crop Sci* 44(4):1236–1244
- Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev* 11(1):31–46. <https://doi.org/10.1038/nrg2626>
- Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci USA* 88(21):9828–9832
- Miller FR (1984) Registration of RTx430 sorghum parental line. *Crop Sci* 24(6):1224–1224. <https://doi.org/10.2135/cropsci1984.0011183X002400060074x>
- Mindaye TT, Mace ES, Godwin ID, Jordan DR (2015) Genetic differentiation analysis for the identification of complementary parental pools for sorghum hybrid breeding in Ethiopia. *Theor Appl Genet* 128(9):1765–1775. <https://doi.org/10.1007/s00122-015-2545-6>
- Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S (2013a) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc Natl Acad Sci USA* 110(2):453–458. <https://doi.org/10.1073/pnas.1215985110>
- Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S (2013b) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc Natl Acad Sci USA* 110(2):453–458
- Morris GP, Rhodes DH, Brenton Z, Ramu P, Thayil VM, Deshpande S, Hash CT, Acharya C, Mitchell SE, Buckler ES, Yu J, Kresovich S (2013c) Dissecting genome-wide association signals for loss-of-function phenotypes in sorghum flavonoid pigmentation traits. *G3* 3(11):2085–2094. <https://doi.org/10.1534/g3.113.008417>
- Multani DS, Briggs SP, Chamberlin MA, Blakeslee JJ, Murphy AS, Johal GS (2003) Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. *Science* 302(5642):81–84. <https://doi.org/10.1126/science.1086072>
- Mural RV, Grzybowski M, Miao C, Damke A, Sapkota S, Boyles RE, Fernandez MGS, Schnable PS, Sigmon B, Kresovich S, Schnable JC (2020) Meta-analysis identifies pleiotropic loci controlling phenotypic trade-offs in sorghum. *BioRxiv*. <https://doi.org/10.1101/2020.10.27.355495>
- NBPGR (2021) National Bureau of plant genetic resources. <http://exploreit.icrisat.org/profile/genetic%20resources/67>. Accessed 30 Mar 2021
- Nida H, Girma G, Mekonen M, Lee S, Seyoum A, Dessalegn K, Tadesse T, Ayana G, Senbetay T, Tesso T, Ejeta G, Mengiste T (2019) Identification of sorghum grain mold resistance loci through genome wide association mapping. *J Cereal Sci* 85:295–304. <https://doi.org/10.1016/j.jcs.2018.12.016>
- Nieto-Sotelo J, Martinez LM, Ponce G, Cassab GI, Alagon A, Meeley RB, Ribaut JM, Yang R (2002) Maize HSP101 plays important roles in both induced and basal thermotolerance and primary root growth. *Plant Cell* 14(7):1621–1633
- Olatoye MO, Hu Z, Maina F, Morris GP (2018) Genomic signatures of adaptation to a precipitation gradient in Nigerian sorghum. *G3* 8(10):3269–3281. <https://doi.org/10.1534/g3.118.200551>

- Olatoye MO, Hu Z, Morris GP (2020a) Genome-wide mapping and prediction of plant architecture in a sorghum nested association mapping population. *Plant Genome* 13(3):e20038. <https://doi.org/10.1002/tpg2.20038>
- Olatoye MO, Marla SR, Hu Z, Bouchet S, Perumal R, Morris GP (2020b) Dissecting adaptive traits with nested association mapping: genetic architecture of inflorescence morphology in sorghum. *G3* 10(5):1785–1796. <https://doi.org/10.1534/g3.119.400658>
- Ongom PO, Ejeta G (2018) Mating design and genetic structure of a multi-parent advanced generation intercross (MAGIC) population of sorghum (*Sorghum bicolor* (L.) Moench). *G3* 8(1):331–341. <https://doi.org/10.1534/g3.117.300248>
- Oria MP, Hamaker BR, Axtell JD, Huang CP (2000) A highly digestible sorghum mutant cultivar exhibits a unique folded structure of endosperm protein bodies. *Proc Natl Acad Sci USA* 97(10):5065–5070. <https://doi.org/10.1073/pnas.080076297>
- Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc Natl Acad Sci USA* 101(26):9903–9908. <https://doi.org/10.1073/pnas.0307901101>
- Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberler G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher AA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob ur R, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS (2009) The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457(7229):551–556. <https://doi.org/10.1038/nature07723>
- Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400(6741):256–261. <https://doi.org/10.1038/22307>
- Peters PJ, Jenks MA, Rich PJ, Axtell JD, Ejeta G (2009) Mutagenesis, selection, and allelic analysis of epicuticular wax mutants in sorghum. *Crop Sci* 49(4):1250–1258. <https://doi.org/10.2135/cropsci2008.08.0461>
- Platt A, Vilhjalmsón BJ, Nordborg M (2010) Conditions under which genome-wide association studies will be positively misleading. *Genetics* 186(3):1045–1052. <https://doi.org/10.1534/genetics.110.121665>
- Porter KS, Axtell JD, Lechtenberg VL, Colenbrander VF (1978) Phenotype, fiber composition, and in vitro dry matter disappearance of chemically induced brown midrib (bmr) mutants of sorghum. *Crop Sci* 18:205–208. <https://doi.org/10.2135/cropsci1978.0011183X001800020002x>
- Prom LK, Ahn E, Isakeit T, Magill C (2019) GWAS analysis of sorghum association panel lines identifies SNPs associated with disease response to Texas isolates of *Colletotrichum sublineola*. *Theor Appl Genet* 132(5):1389–1396. <https://doi.org/10.1007/s00122-019-03285-5>
- Prom LK, Cuevas HE, Ahn E, Isakeit T, Rooney WL, Magill C (2020) Genome-wide association study of grain mold resistance in sorghum association panel as affected by inoculation with *Alternaria alternata* alone and *Alternaria alternata*, *Fusarium thapsinum*, and *Curvularia lunata* combined. *Eur J Plant Pathol* 157(4):783–798. <https://doi.org/10.1007/s10658-020-02036-3>
- Quinby JR (1975) The genetics of sorghum improvement. *J Hered* 66(2):56–62
- Quinby JR, Karper RE (1954) Inheritance of height in sorghum. *Agron J* 46:211–216
- Rao SA, Rao KEP, Mengesha MH, Reddy VG (1996) Morphological diversity in sorghum germplasm from India. *Genet Resour Crop Evol* 43(6):559–567
- Reynolds M, Foulkes MJ, Slafer GA, Berry P, Parry MAJ, Snape JW, Angus WJ (2009) Raising yield potential in wheat. *J Exp Bot* 60(7):1899–1918. <https://doi.org/10.1093/jxb/erp016>
- Richards RA (2000) Selectable traits to increase crop photosynthesis and yield of grain crops. *J Exp Bot* 51(suppl_1):447–458. https://doi.org/10.1093/jexbot/51.suppl_1.447
- Rooney WL (2004) Sorghum improvement-integrating traditional and new technology to produce improved genotypes. *Adv Agron* 83:37–109. [https://doi.org/10.1016/s0065-2113\(04\)83002-5](https://doi.org/10.1016/s0065-2113(04)83002-5)
- Saballos A, Sattler SE, Sanchez E, Foster TP, Xin Z, Kang C, Pedersen JF, Vermerris W (2012) Brown midrib2 (Bmr2) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor* (L.) Moench). *Plant J* 70(5):818–830. <https://doi.org/10.1111/j.1365-313X.2012.04933.x>
- Saeed M, Francis CA, Clegg MD (1986) Yield component analysis in grain sorghum. *Crop Sci* 26(2):346–351. <https://doi.org/10.2135/cropsci1986.0011183X002600020028x>
- Sakamoto T, Morinaka Y, Ohnishi T, Sunohara H, Fujioka S, Ueguchi-Tanaka M, Mizutani M, Sakata K, Takatsuto S, Yoshida S, Tanaka H, Kitano H, Matsuoka M (2006) Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. *Nat Biotechnol* 24(1):105–109. <https://doi.org/10.1038/nbt1173>
- Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, Kitano H, Matsuoka M (2002) Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature* 416(6882):701–702. <https://doi.org/10.1038/416701a>
- Sattler SE, Saballos A, Xin Z, Funnell-Harris DL, Vermerris W, Pedersen JF (2014) Characterization of novel Sorghum brown midrib mutants from an EMS-mutagenized population. *G3* 4(11):2115–2124. <https://doi.org/10.1534/g3.114.014001>
- Schneeberger K, Ossowski S, Lanz C, Juul T, Petersen AH, Nielsen KL, Jorgensen JE, Weigel D, Andersen SU (2009) SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat Methods* 6(8):550–551. <https://doi.org/10.1038/nmeth0809-550>
- Scully ED, Gries T, Funnell-Harris DL, Xin Z, Kovacs FA, Vermerris W, Sattler SE (2016) Characterization of novel Brown midrib 6 mutations affecting lignin biosynthesis in sorghum. *J Integr Plant Biol* 58(2):136–149. <https://doi.org/10.1111/jipb.12375>
- Shakoor N, Nair R, Crasta O, Morris G, Feltus A, Kresovich S (2014) A Sorghum bicolor expression atlas reveals dynamic genotype-specific expression profiles for vegetative tissues of grain, sweet and bioenergy sorghums. *BMC Plant Biol* 14(1):35. <https://doi.org/10.1186/1471-2229-14-35>
- Sharma R, Rao VP, Upadhyaya HD, Reddy VG, Thakur RP (2010) Resistance to grain mold and downy mildew in a mini-core collection of sorghum germplasm. *Plant Dis* 94(4):439–444. <https://doi.org/10.1094/PDIS-94-4-0439>
- Sharma R, Upadhyaya HD, Manjunatha SV, Rao VP, Thakur RP (2012) Resistance to foliar diseases in a mini-core collection of sorghum germplasm. *Plant Dis* 96(11):1629–1633. <https://doi.org/10.1094/PDIS-10-11-0875-RE>
- Silva KJ, Pastina MM, Guimarães CT, Magalhães JV, Pimentel LD, Schaffert RE, Pinto MO, Souza VF, Bernardino KC, Silva MJ, Borém A, Menezes CB (2021) Genetic diversity and heterotic grouping of sorghum lines using SNP markers. *Sci Agric*. <https://doi.org/10.1590/1678-992x-2020-0039>
- Singh SP, Drolsom PN (1973) induced recessive mutations affecting leaf angle in *Sorghum bicolor*. *J Hered* 64(2):65–68

- Singh SP, Drolsom PN (1974) Induced early-maturing mutation in sorghum. *Crop Sci* 14(3):377–380
- Smith CW, Frederiksen RA (2000) History of cultivar development in the United States: from “Memoirs of A.B. Maunder—Sorghum Breeder.” In: Smith CWFRA (ed) *Sorghum: origin, history, technology, and production*. Wiley, New York, p 191
- Stephens JC, Miller FR, Rosenow DT (1967) Conversion of alien sorghums to early combine genotypes. *Crop Sci* 7(4):396–400
- Subudhi PK, Rosenow DT, Nguyen HT (2000) Quantitative trait loci for the stay green trait in sorghum (*Sorghum bicolor* L. Moench): consistency across genetic backgrounds and environments. *Theor Appl Genet* 101(5):733–741. <https://doi.org/10.1007/s001220051538>
- Sung D-Y, Kaplan F, Lee K-J, Guy CL (2003) Acquired tolerance to temperature extremes. *Trends Plant Sci* 8(4):179–187. [https://doi.org/10.1016/S1360-1385\(03\)00047-5](https://doi.org/10.1016/S1360-1385(03)00047-5)
- Tao Y, Luo H, Xu J, Cruickshank A, Zhao X, Teng F, Hathorn A, Wu X, Liu Y, Shatte T, Jordan D, Jing H, Mace E (2021) Extensive variation within the pan-genome of cultivated and wild sorghum. *Nat Plants* 7(6):766–773. <https://doi.org/10.1038/s41477-021-00925-x>
- Thakur RP, Mathur K (2000) Anthracnose. In: Frederiksen RA, Odvody GN (eds) *Compendium of sorghum diseases*. The American Phytopathology Society, St. Paul, pp 10–12
- Thurber CS, Ma JM, Higgins RH, Brown PJ (2013) Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biol* 14(6):R68. <https://doi.org/10.1186/gb-2013-14-6-r68>
- Till BJ, Colbert T, Tompa R, Enns LC, Codomo CA, Johnson JE, Reynolds SH, Henikoff JG, Greene EA, Steine MN, Comai L, Henikoff S (2003) High-throughput TILLING for functional genomics. *Methods Mol Biol* 236:205–220
- Tishchenko V, Wang M, Xin Z, Harrison M (2020) Development of root phenotyping platforms for identification of root architecture mutations in EMS-induced and low-path-sequenced sorghum mutant population. *Am J Plant Sci* 11(06):838–850. <https://doi.org/10.4236/ajps.2020.116060>
- Tsai H, Howell T, Nitcher R, Missirian V, Watson B, Ngo KJ, Lieberman M, Fass J, Uauy C, Tran RK, Khan AA, Filkov V, Tai TH, Dubcovsky J, Comai L (2011) Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiol* 156(3):1257–1268. <https://doi.org/10.1104/pp.110.169748>
- Tuinstra MR, Grote EM, Goldsbrough PB, Ejeta G (1996) Identification of quantitative trait loci associated with pre-flowering drought tolerance in sorghum. *Crop Sci*. <https://doi.org/10.2135/cropsci1996.0011183X003600050043x>
- Upadhyaya HD, Ortiz R (2001) A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor Appl Genet* 102(8):1292–1298. <https://doi.org/10.1007/s00122-001-0556-y>
- Upadhyaya HD, Pundir RPS, Dwivedi SL, Gowda CLL, Reddy VG, Singh S (2009) Developing a mini core collection of sorghum for diversified utilization of germplasm. *Crop Sci* 49(5):1769–1780. <https://doi.org/10.2135/cropsci2009.01.0014>
- Upadhyaya HD, Wang YH, Sharma R, Sharma S (2013) Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. *Theor Appl Genet* 126(6):1649–1657. <https://doi.org/10.1007/s00122-013-2081-1>
- Upadhyaya HD, Vetriventhan M, Deshpande S (2016) Sorghum germplasm resources characterization and trait mapping. In: *The sorghum genome*. Springer, pp 77–94
- Upadhyaya HD, Reddy KN, Vetriventhan M, Ahmed MI, Krishna GM, Reddy MT, Singh SK (2017a) Sorghum germplasm from West and Central Africa maintained in the ICRISAT genebank: status, gaps, and diversity. *Crop J* 5(6):518–532. <https://doi.org/10.1016/j.cj.2017.07.002>
- Upadhyaya HD, Reddy KN, Vetriventhan M, Gumma MK, Ahmed MI, Reddy MT, Singh SK (2017b) Status, genetic diversity and gaps in sorghum germplasm from South Asia conserved at ICRISAT genebank. *Plant Genet Resour* 15(6):527–538. <https://doi.org/10.1017/S147926211600023x>
- USDA (2021) Agricultural Research Service, National Plant Germplasm System Germplasm Resources Information Network (GRIN-Taxonomy). National Germplasm Resources Laboratory. <https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomysimple>. Accessed 30 Mar 2021
- Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC (2016) SIFT missense predictions for genomes. *Nat Protoc* 11(1):1–9. <https://doi.org/10.1038/nprot.2015.123>
- Walters DR, Keil DJ (1988) *Vascular plant taxonomy*, 4th edn. Kendall/Hunt Pub. Co., Dubuque
- Wang P, Duan W, Takabayashi A, Endo T, Shikanai T, Ye JY, Mi H (2006) Chloroplastic NAD(P)H dehydrogenase in tobacco leaves functions in alleviation of oxidative damage caused by temperature stress. *Plant Physiol* 141(2):465–474. <https://doi.org/10.1104/pp.105.070490>
- Wang X, Gowik U, Tang H, Bowers JE, Westhoff P, Paterson AH (2009) Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. *Genome Biol* 10(6):R68. <https://doi.org/10.1186/gb-2009-10-6-r68>
- Wang Y, Zhao J, Lu W, Deng D (2017) Gibberellin in plant height control: old player, new story. *Plant Cell Rep* 36(3):391–398. <https://doi.org/10.1007/s00299-017-2104-5>
- Wang J, Hu Z, Upadhyaya HD, Morris GP (2020) Genomic signatures of seed mass adaptation to global precipitation gradients in sorghum. *Heredity (edinb)* 124(1):108–121. <https://doi.org/10.1038/s41437-019-0249-4>
- Wang L, Lu Z, Regulski M, Jiao Y, Chen J, Ware D, Xin Z (2021) BSAseq: an interactive and integrated web-based workflow for identification of causal mutations in bulked F2 populations. *Bioinformatics* 37(3):382–387. <https://doi.org/10.1093/bioinformatics/btaa709>
- Wienholds E, van Eeden F, Kusters M, Mudde J, Plasterk RH, Cuppen E (2003) Efficient target-selected mutagenesis in zebrafish. *Genome Res* 13(12):2700–2707
- Winkler S, Schwabedissen A, Backasch D, Bokel C, Seidel C, Bonisch S, Furthauer M, Kuhrs A, Cobreros L, Brand M, Gonzalez-Gaitan M (2005) Target-selected mutant screen by TILLING in *Drosophila*. *Genome Res* 15(5):718–723
- Wu Y, Li X, Xiang W, Zhu C, Lin Z, Wu Y, Li J, Pandravada S, Ridder DD, Bai G, Wang ML, Trick HN, Bean SR, Tuinstra MR, Tesso TT, Yu J (2012) Presence of tannins in sorghum grains is conditioned by different natural alleles of Tannin1. *Proc Natl Acad Sci USA* 109(26):10281–10286. <https://doi.org/10.1073/pnas.1201700109>
- Wu Y, Guo T, Mu Q, Wang J, Li X, Wu Y, Tian B, Wang ML, Bai G, Perumal R, Trick HN, Bean SR, Dweikat IM, Tuinstra MR, Morris G, Tesso TT, Yu J, Li X (2019) Allelochemicals targeted to balance competing selections in African agroecosystems. *Nat Plants* 5(12):1229–1236. <https://doi.org/10.1038/s41477-019-0563-0>
- Xin Z, Wang ML, Barkley NA, Burow G, Franks C, Pederson G, Burke J (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biol* 8:103. <https://doi.org/10.1186/1471-2229-8-103>
- Xin Z, Wang M, Burow G, Burke J (2009) An induced sorghum mutant population suitable for bioenergy research. *BioEnergy Res* 2(1):10–16
- Xin Z, Huang J, Smith AR, Chen J, Burke J, Sattler SE, Zhao D (2017) Morphological characterization of a new and easily recognizable nuclear male sterile mutant of sorghum (*Sorghum*

- bicolor*). PLoS ONE 12(1):e0165195. <https://doi.org/10.1371/journal.pone.0165195>
- Yamaguchi M, Fujimoto H, Hirano K, Araki-Nakamura S, Ohmae-Shinohara K, Fujii A, Tsunashima M, Song XJ, Ito Y, Nagae R, Wu J, Mizuno H, Yonemaru J, Matsumoto T, Kitano H, Matsuoka M, Kasuga S, Sazuka T (2016) Sorghum Dw1, an agronomically important gene for lodging resistance, encodes a novel protein involved in cell proliferation. Sci Rep 6:28366. <https://doi.org/10.1038/srep28366>
- Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178(1):539–551. <https://doi.org/10.1534/genetics.107.074245>
- Yu J, Hamblin MT, Tuinstra MR (2013) Association genetics strategies and resources. In: Paterson AH (ed) Genomics of the saccharinae. Springer, New York, pp 187–203
- Zhang LM, Leng CY, Luo H, Wu XY, Liu ZQ, Zhang YM, Zhang H, Xia Y, Shang L, Liu CM, Hao DY, Zhou YH, Chu CC, Cai HW, Jing HC (2018) Sweet sorghum originated through selection of dry, a plant-specific NAC transcription factor gene. Plant Cell 30(10):2286–2307. <https://doi.org/10.1105/tpc.18.00313>
- Zheng LY, Guo XS, He B, Sun LJ, Peng Y, Dong SS, Liu TF, Jiang S, Ramachandran S, Liu CM, Jing HC (2011) Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). Genome Biol 12(11):R114. <https://doi.org/10.1186/gb-2011-12-11-r114>
- Zhu Y, Mang HG, Sun Q, Qian J, Hipps A, Hua J (2012) Gene discovery using mutagen-induced polymorphisms and deep sequencing: application to plant disease resistance. Genetics 192(1):139–146. <https://doi.org/10.1534/genetics.112.141986>
- Zuo E, Sun Y, Wei W, Yuan T, Ying W, Sun H, Yuan L, Steinmetz LM, Li Y, Yang H (2019) Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science 364(6437):289–292. <https://doi.org/10.1126/science.aav9973>
- Zou G, Zhai G, Yan S, Li S, Zhou L, Ding Y, Liu H, Zhang Z, Zou J, Zhang L, Chen J, Xin Z, Tao Y (2020) Sorghum qTGW1a encodes a G-protein subunit and acts as a negative regulator of grain size. J Exp Bot 71(18):5389–5401. <https://doi.org/10.1093/jxb/eraa277>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.