Quantitative and population genomics suggest a broad role of stay-green loci in the drought adaptation of sorghum

Jacques M Faye, Eyanawa A Akata, Bassirou Sine, Cyril Diatta, Ndiaga Cisse, Daniel Fonceka, Geoffrey P Morris

Received: 25 June 2021 | Accepted: 8 October 2021

DOI: 10.1002/tpg2.20176

The Plant Genome

ORIGINAL RESEARCH

Quantitative and population genomics suggest a broad role of stay-green loci in the drought adaptation of sorghum

Jacques M. Faye[1,2] | **Eyanawa A. Akata**[2,3] | **Bassirou Sine**[2] | **Cyril Diatta**[2] | **Ndiaga Cisse**[2] | **Daniel Fonceka**[2,4] | **Geoffrey P. Morris**[1,5]

[1] Dep. of Agronomy, Kansas State Univ., Manhattan, KS, USA

[2] Centre d'Étude Régional pour l'Amélioration de l'Adaptation à la Sécheresse, Institut Sénégalais de Recherches Agricoles, Thiès, Senegal

[3] Institut Togolais de Recherche Agronomique, Kara, Togo

[4] AGAP, Univ. Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

[5] Dep. of Soil & Crop Science, Colorado State Univ., Ft. Collins, CO 80523, USA

Correspondence
Geoffrey P. Morris, Dep. of Agronomy, Kansas State Univ., Manhattan, KS, USA.
Email: Geoff.Morris@colostate.edu

Assigned to Associate Editor Emma Mace.

Abstract

Drought is a major constraint on plant productivity globally. Sorghum [*Sorghum bicolor* (L.) Moench] landraces have evolved in drought-prone regions, but the genetics of their adaptation is poorly understood. Here we sought to identify novel drought-tolerance loci and test hypotheses on the role of known loci including those underlying stay-green (*Stg*) postflowering drought tolerance. We phenotyped 590 diverse sorghum accessions from West Africa in 10 environments, under field-based managed drought stress [preflowering water stress (WS1), postflowering water stress (WS2), and well-watered (WW)] and rainfed (RF) conditions over 4 yr. Days to 50% flowering (DFLo), aboveground dry biomass (DBM), plant height (PH), and plant grain yield components (including grain weight [GrW], panicle weight [PW] and grain number [GrN] per plant, and 1000-grain weight [TGrW]) were measured, and genome-wide association studies (GWAS) was conducted. Broad-sense heritability for biomass and plant grain yield was high (33–92%) across environments. There was a significant correlation between stress tolerance index (STI) for GrW per plant across WS1 and WS2. Genome-wide association studies revealed that *SbZfl1* and *SbCN12*, orthologs of maize (*Zea mays* L.) flowering genes, likely underlie flowering time variation under these conditions. Genome-wide association studies further identified associations ($n = 134$; common between two GWAS models) for STI and drought effects on plant yield components including 16 putative pleiotropic associations. Thirty of the associations colocalized with *Stg1*, *Stg2*, *Stg3*, and *Stg4* loci and had large effects. Seven lead associations, including some within *Stg1*, overlapped

Abbreviations: DBM, aboveground dry biomass; DFLo, days to 50% flowering; FTSW, fraction of transpirable soil water; GBS, genotyping-by-sequencing; GLM, general linear model; GLM+Q, general linear model with principal components; GrN, grain number; GrW, grain weight; GWAS, genome-wide association studies; MLM, mixed linear model; PC, principal component; PH, plant height; PVE, phenotypic variance explained; PW, panicle weight; QTL, quantitative trait loci; RDBM, reduction of aboveground dry biomass; RF, rainfed; RGrN, reduction of grain number; RPH, reduction of plant height; RPW, reduction of panicle weight; RTGrW, reduction of 1000-grain weight; SNP, single-nucleotide polymorphism; STI, stress tolerance index; TGrW, 1000-grain weight; TPE, target population of environment; WASAP, West African sorghum association panel; WS1, preflowering water stress; WS2, postflowering water stress; WW, well-watered

with positive selection outliers. Our findings reveal previously undescribed natural genetic variation for drought-tolerance-related traits and suggest a broad role of *Stg* loci in drought adaptation of sorghum.

1 | INTRODUCTION

Unpredictable rainfall and drought are major limitations to plant productivity worldwide. Improving crop adaptation to water limitation is critical for establishing food security in developing countries where smallholder farmers are vulnerable to climate change (Mundia et al., 2019). From an agronomic perspective, drought adaptation is the ability to maintain yield under agronomic water limitation (Blum, 2010). An understanding of the genetic architecture of grain yield and its components across various drought scenarios can facilitate crop breeding to increase production. However, collecting good phenotypic data under well-managed water stress environments and integrating phenotypes with genotypes remain major constraints. The genetic dissection of yield components under various drought scenarios would provide favorable natural variants for drought tolerance.

Sorghum [*Sorghum bicolor* (L.) Moench] is a staple cereal crop in drought-prone regions worldwide including many developing countries of the semi-arid tropics as well as industrialized countries in the temperate latitudes. Sorghum is among the most drought-resilient crops, but the physiological and genetic basis of drought tolerance in sorghum landraces is not yet understood (Mullet et al., 2014). Several quantitative trait loci (QTL) associated with drought-tolerance variation in sorghum have been identified but no genes have been cloned. The best studied of these QTL are stay-green (*Stg*) loci (*Stg1–Stg4)* linked to postflowering drought tolerance in temperate-adapted breeding lines (Borrell et al., 2014b; Harris et al., 2007; Hayes et al., 2016; Tuinstra et al., 1997; Xu et al., 2000). The *Stg* loci influence several aspects of sorghum development including canopy architecture, water use, and grain yield (Borrell et al., 2014b). *Stg* alleles were identified in a temperate-adapted breeding line BTx642 (formerly B35) that is derived from a tropically adapted Ethiopian durra landrace (IS12555). However, the prevalence of the *Stg* alleles in sub-Saharan Africa or their role in drought adaptation of landraces (if any) is not understood. Understanding the genetic basis of drought adaptation in sorghum could elucidate the process of environmental adaptation and facilitate breeding of drought-tolerant cultivars.

Local cultivars have been under natural and farmers' selection for adaptation to environmental conditions and farming systems. Local cultivars of sorghum have adapted to various environmental conditions since their domestication (Harlan & De Wet, 1972; Wendorf et al., 1992). Consequently, positive pleiotropic loci for combined pre- and postflowering drought tolerance might exist in locally adapted cultivars. West African sorghum is extremely diverse and there have been few cycles of selection in breeding programs (Leiser et al., 2014; Mauboussin et al., 1977). The West African sorghum association panel (WASAP), including landraces and breeding lines that consist of working collections of breeding programs, was assembled and genotyped using genotyping-by-sequencing (GBS) technology. However, the genetic architecture underlying grain yield and its components under various drought scenarios remains largely unknown in the germplasm. We hypothesized that positively pleiotropic QTL confer combined pre- and postflowering drought tolerance in the West African sorghum.

Genome-wide association studies (GWAS) contribute to the identification of natural variants, taking advantage of historical recombinations within diverse germplasm (Hu et al., 2019; Tao et al., 2020; Zhao et al., 2019). A grass species such as sorghum is suitable to identify natural variants underlying complex agronomic traits partly because of its small genome size and moderate linkage disequilibrium (LD) (Mace et al., 2013; McCormick et al., 2018; Paterson et al., 2009). Disentangling positive pleiotropic effects of drought-yield QTL through GWAS can contribute to detect and characterize the natural allelic variation existing within locally adapted populations. In this study, we performed GWAS on 590 sorghum accessions of the WASAP under 10 different environments using the previous GBS single-nucleotide polymorphism (SNP) dataset. We (a) characterize the genetic variation of biomass and yield components at the plant level under various water stress environments, (b) identify genetic variants at known and novel drought-tolerance loci with high productivity under pre- and postflowering water stress, (c) investigate the pleiotropic effect of drought-tolerance QTL associated with stress tolerance index (STI) and reduction of biomass and plant grain yield components under various drought scenarios, and (d) determine signatures of selection overlapping with identified drought-tolerance QTL. The present study provides knowledge of the genetic architecture of yield components under various drought scenarios.

2 | MATERIALS AND METHODS

2.1 | Plant materials

The WASAP consists of 756 genotyped accessions from the four West African countries of Senegal (118 accessions), Mali (123), Togo (156), and Niger (359) (Faye et al., 2021) (Figure 1a). The botanical types were defined based on a priori classification for the majority of the accessions (Faye et al., 2021). The collection does not include kafir sorghum type, which is mostly found in southern Africa. The panel includes predominantly landraces along with some local breeding lines and local improved cultivars. Five local breeding lines were used as checks for use in augmented design: T1 (IRAT 204/CE151-262), T2 (CE145-266), T3 (ISRA-621B/Faourou), T4 (CE180-33), and T5 (53-49). Two international drought-response reference lines, Tx7000 (preflowering drought-tolerant, postflowering drought susceptible) and BTx642 (preflowering drought susceptible, postflowering drought tolerant), were used as controls (Borrell et al., 2014a; Burke et al., 2013). A total of 590 accessions were evaluated in field-based managed drought-stress environments based on seed availability.

2.2 | Field trials

Field experiments were performed over 4 yr (2014–2017) in Senegal at the Bambey Research Station, Centre National de Recherche Agronomique (14.42˚N, 16.28˚W) in the Soudano–Sahelian zone (Figure 1a). The average annual precipitation is ~600 mm, which occurs strictly in the rainy season (*hivernage*) of July to October, with maximum monthly precipitation typically occurring in August (Figure 1b). In total, 10 experiments were performed in an incomplete randomized block design (augmented block design) across the 4 yr (Table 1; Supplemental Figure S1a–f). The experimental set-up followed a column–row field layout with 30 blocks for 2014 experiments or 25 blocks for 2015–2017 experiments, with 19 genotypes and the five local check cultivars (present in each block for spatial variation analysis) within each block. Each entry was sown in a 3-m row with 0.6-m space between rows and 0.2-m space between plants (or hills) within a row. Each entry was flanked on each side by one row of a standard cultivar (IRAT 204). Ten days after planting, plants were thinned to keep only one plant per hill for a density of ~84,000 plants ha^{-1}. Two experiments were carried out under rainfed (RF) conditions during the rainy season in 2014 with 1-mo planting date interval: RF1 (planted in August) and RF2 (planted in September). Managed-drought-stress experiments were conducted in the off-season to take advantage of the complete lack of precipitation during the Sahelian dry season (Figure 1b).

Core Ideas

- Sorghum is famously drought tolerant but the underlying genetics remains poorly understood
- We studied drought response of a large diverse West African panel in managed field stress.
- There is evidence of drought adaptation across all botanical types and novel pleiotropic QTL.
- Stay-green drought-tolerance loci may have broad role in drought adaptation across Africa.
- A global role of florigen *CN12* in flowering variation and evidence of a role in drought escape was found.

2.3 | Managed drought stress

Well-watered (WW) and preflowering water stress (WS1) experiments were planted during the hot off-season in 2015 (March–August). Three experiments—under WW, WS1, and postflowering water stress (WS2)—were planted during the cool off-season in 2015–2016 (October 2015 to March 2016, hereafter referred to as 2016 experiments) and 2016–2017 (October 2016 to March 2017; hereafter referred to as 2017 experiments). During the rainy season of 2014, the cumulative rainfall recorded was 395 mm. The average daily temperature ranges between 22.4 and 35 ˚C and average relative humidity between 42 and 89%. For WW experiments, irrigation was applied twice a week (30 mm each time) until physiological maturity. For WS1 experiments, water limitation was applied 30 d after planting, to mimic a 1-mo preflowering drought, and irrigation was restarted 60 d after planting until physiological maturity. For the WS2 experiments, water limitation was applied when 75% of plants in a maturity group flowered and was maintained until physiological maturity. Three maturity groups were defined based on accession phenology characterized during 2014 experiments for water deficit application in WS2. The fraction of transpirable soil water (FTSW) in different managed drought-stress experiments was determined using a DIVINER 2000 soil moisture monitor (Sentek Pty Ltd).

2.4 | Phenotypic measurements

In each environment, phenological, physiological, and yield component traits were measured. Days to 50% flowering (DFLo) of plants in a plot (one row), aboveground dry biomass (DBM), plant height (PH), and grain yield components including grain weight (GrW), panicle weight (PW) and grain number (GrN) per plant, and 1000-grain weight (TGrW)

TABLE 1 Details of field experiments

Treatment[a]	Season	Year[b]	Period	Trial code[c]
Rainfed	Rainy season	2014	July–October	RF1
Rainfed	Rainy season	2014	July–October	RF2
Well-watered	Hot off-season	2015	March–August	WW_15
Preflowering	Hot off-season	2015	March–August	WS1_15
Well-watered	Cool off-season	2016	October–February	WW_16
Preflowering	Cool off-season	2016	October–February	WS1_16
Postflowering	Cool off-season	2016	October–February	WS2_16
Well-watered	Cool off-season	2017	October–February	WW_17
Preflowering	Cool off-season	2017	October–February	WS1_17
Postflowering	Cool off-season	2017	October–February	WS2_17

[a]Preflowering and postflowering denote the pre- or postflowering water stress environments.
[b]The year where the majority of the experiment took place.
[c]RF, rainfed condition; WW, well-water; WS1, preflowering water stress; WS2, postflowering water stress.

were measured and used for association mapping studies. For each trait except for DFLo and TGrW, three plants from the middle row of each plot, including tillers, were used for measurements. Note that observations of grain yield components are based on a per-plant basis, including tillering, rather than on a per-plot basis. The drought STI (Li et al., 2018a; Yuan et al., 2019) for GrW per plant was calculated from the GrW under WW and WS1 or WS2 as follows:

$$\text{STI} = \frac{(Y_{\text{ww}})(Y_{\text{ws}})}{Y^2_{\text{m,ww}}}$$

where Y_{ww} and Y_{ws} are the GrW of a given genotype in WW and water stress environments, respectively, and $Y_{\text{m, ww}}$ is the mean value of GrW in the WW environment. For the STI, the higher the value, the more tolerant the genotype to the stress. The drought reduction of each yield component relative to the control environment was calculated as follows:

$$R_i(\%) = \frac{(Y_{\text{ww}})(Y_{\text{ws}})}{Y_{\text{ww}}} \times 100$$

where R_i is the drought response of a genotype for trait i, and Y_{ww} and Y_{ws} are the performance of the genotype in the control environment and water stress environments, respectively.

2.5 | Statistical analysis of phenotypes

Each year–treatment combination is considered an environment. Statistical analysis was performed using the R program (R Core Team, 2016). Spatial variation within each environment was analyzed based on the check cultivars in each block using the *SpATS* package (Rodríguez-Álvarez et al., 2018) to obtain genotype-adjusted means. The variance components were estimated by fitting the mixed linear model with random effects for all genotypes (G), water regimes, years (Y), and G × Y interaction effects using the *lme4* package (Bates et al., 2015). Broad-sense heritability was calculated based on variance components derived from the mixed effect model. Broad-sense heritability was estimated for each trait across environments based on the genotypic variance and the total phenotypic variance. Phenotypic correlations among traits were calculated using the Pearson correlation coefficient of the PerformanceAnalytics package (Peterson et al., 2014). Tukey's honest significant difference test in the Agricolae package (Mendiburu, 2009) was used to test the difference of genotype performance between environments or botanical types. The best linear unbiased predictors (BLUPs) values of the phenotypes were calculated by combining data for a given water regime across years or across all environments. The phenotypic BLUPs and genotype-adjusted means were used for the GWAS across environments.

2.6 | Genome-wide association studies

To identify drought-yield QTL, GWAS was performed using the general linear model (GLM) with principal component (PC) eigenvalues and mixed linear model (MLM) in the GAPIT package (Lipka et al., 2012) using previously published GBS SNPs (Faye et al., 2021). These two GWAS models were used as complementary because the GLM may identify false–positive associations while MLM may lead to false–negative associations when controlling for false–positive significant associations. The SNP dataset was filtered for minor allele frequency >0.02 (130,708 SNPs used for GWAS), which corresponds to >15 observations of the minor allele within the panel of $N = 756$ genotyped accessions. The first five PCs were used in the GLM to account for population structure. The first five PCs and the kinship

matrix were used in the MLM to account for population structure and genetic relatedness effects, respectively. The significance level of GWAS associations were defined based on Bonferroni-corrected p value at alpha 0.05 for the GLM with PC (referred to as GLM+Q hereafter) or at least top five SNPs above $p < 10^{-5}$ cutoff for the MLM. The effect and proportion of variance explained by these SNPs were considered in the downstream analysis to retain associations with reasonable biological significance. The most highly associated SNP (lead SNP) within a 150-kb genomic region defined based on average LD decay in global sorghum germplasm (Morris et al., 2013) was chosen to represent the association. A list of a priori candidate genes of cloned cereal flowering times from a previous study (Faye et al., 2019) was used for colocalization analysis between lead SNP and candidate genes.

2.7 | Locus-specific analyses

Linkage disequilibrium heatmaps were constructed using the R package Ldheatmap 0.99-4 (Shin et al., 2006). The BLUP values of phenotypes across water stress environments were used for the estimation of the proportion of phenotypic variance explained (PVE) by lead SNPs from the GWAS. The PVE was estimated using linear models with fractions of ancestry inferred by ADMIXTURE (Alexander et al., 2009) used as fixed covariates. Statistical enrichment analysis for colocalization between GWAS lead SNPs and all *Stg* QTL from the sorghum QTL Atlas (Mace et al., 2019) was performed based on 1,000 permutation tests. Statistical significance was assessed with a two-sample t test with $\alpha = 0.05$. Geographic distribution of the associated lead SNP alleles with DFLo or putative drought tolerance was determined using an existing set of georeferenced global sorghum landraces (Lasky et al., 2015). Lead associations within *Stg1–3* QTL were selected based on their association with drought-tolerance variables, LD with other lead associations within a locus, contribution to the phenotypic variation, and availability in the GBS data for global sorghum landraces.

2.8 | Genome-wide selection scans

For selection scans, we included 550 worldwide sorghum accessions including wild relative sorghum accessions with available sequencing data (Morris et al., 2013). Genome-wide selection scans were performed based on 100-kb sliding windows using the vcftools program (Danecek et al., 2011). Decreased genome-wide nucleotide diversity (π) in durra–caudatum, durra, and guinea cultivars relative to wild relatives was performed to assess domestication and diversification selections for drought responses to dry (in durra–caudatum and durra genome) vs. humid (in guinea genome) regions. Statistical enrichment analysis for colocalization between π outlier regions and *Stg1–4* loci was performed based on 1,000 permutation tests. Statistical significance of mean differences were based on two-sided, two-sample t tests with $\alpha = 0.05$.

3 | RESULTS

3.1 | Phenotypic variation for drought-tolerance related traits

A total of 590 WASAP accessions were evaluated for phenological, physiological, and yield component traits under 10 environments across 4 yr in Senegal (Figure 1a,b; Supplemental Figure S1). To assess the level of drought stress applied, we estimated the FTSW in the WW, WS1, and WS2 environments (Figure 1c–f). Fraction of transpirable soil water was estimated to be 0.6 in both WW and stressed treatment before water deficit treatment then dropped to ~0.2 and 0.3 in WS1 and WS2 environments, respectively. To assess the effect of each water condition, we characterized the grain yield components and days to flowering of genotypes. Cross-over G × E interaction ($p < .04$, one-tailed ANOVA) was observed between the two drought-tolerance reference lines, BTx642 and Tx7000, in WS1 and WS2 (Figure 1g). As expected, the average GrW and number of genotypes was significantly reduced in WS1 and WS2 relative to WW treatment (Supplemental Figure S1a,b). Overall, DFLo was significantly delayed in 2015 hot off-season environments, whereas it was reduced in cool off-seasons of 2016 and 2017 relative to RF conditions (Supplemental Figure S1c). The DFLo was delayed in WS1, whereas it was not different in WS2 relative to the WW controls. The DBM was significantly reduced in all stressed environments except in WS1 of 2015 relative to RF (Supplemental Figure S1d). Average GrW was not significantly different between RF and WS2 (Supplemental Figure S1e). The average GrN was significantly lower in WS1 than in WS2 (Supplemental Figure S1f).

3.2 | Genetic variation in drought response

Broad-sense heritability estimates varied from moderate to high with values ranging from 33% for GrN to 92% for PH in the whole WASAP (Supplemental Table S1). The average GrW was not significantly different between caudatum accessions and durra and guinea accessions within each water regime in terms of production under drought stress (Figure 2a). The durra–caudatum intermediates had significantly higher average GrW than caudatum (13%, $P < .05$) and guinea (16%, $P < .05$) accessions but not with durra (7%,

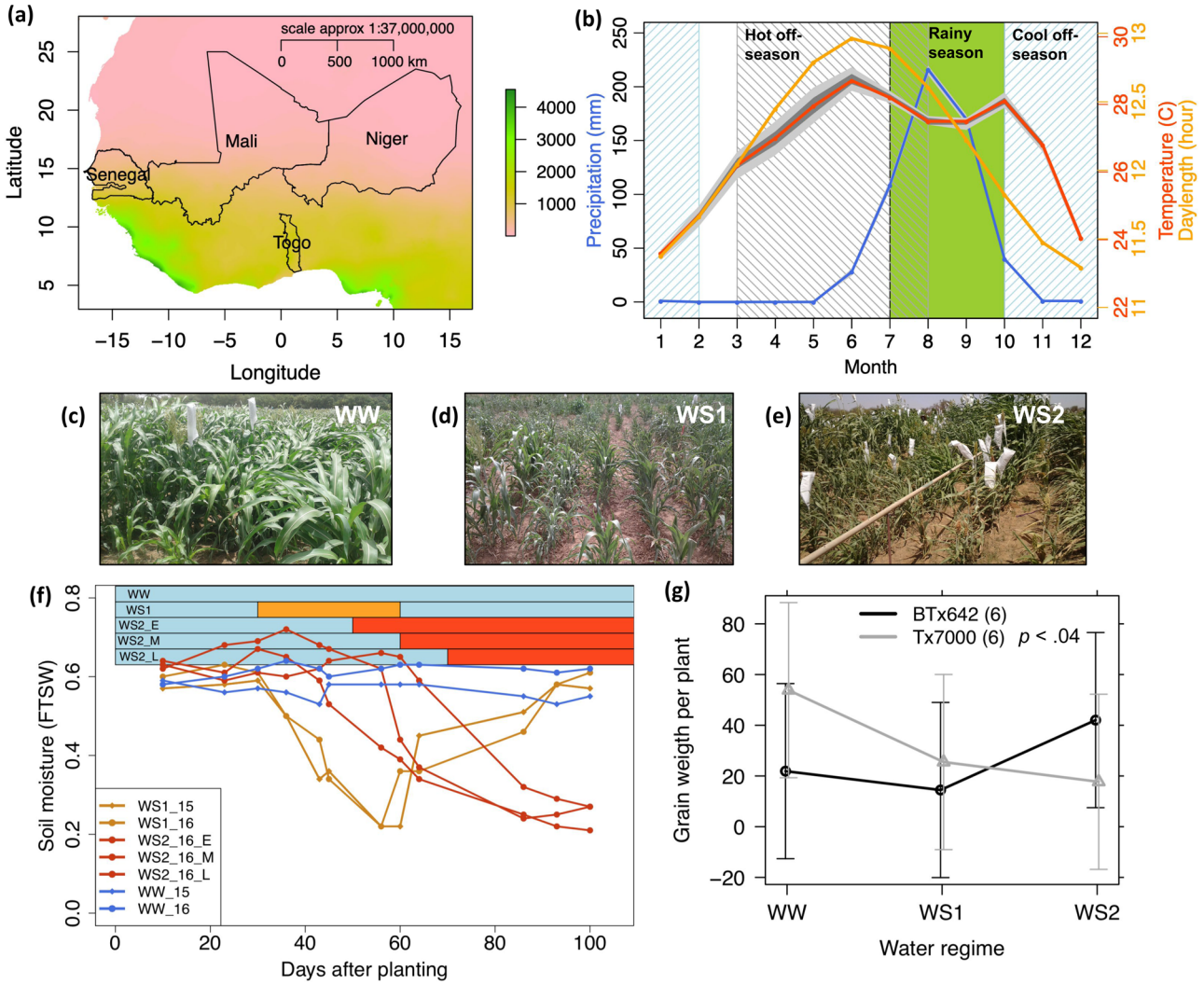

FIGURE 1 Experimental system to study drought-stress response of diverse sorghum germplasm. (a) The four countries of origin for sorghum accessions in the West African sorghum association panel (WASAP) with the West African precipitation gradient noted by the color scale. (b) Average monthly precipitation, temperature and day length at the experimental station in Bambey, Senegal. The green block represents the rainy season (*hivernage*) when farmers grow crops and when we conducted rainfed experiments. The gray-striped block indicates the hot off-season and the blue-striped block indicates the cool off-season when we conducted managed drought stress. (c-e) Plants under (c) well-watered (WW), (d) preflowering water stress (WS1), and (e) postflowering water stress (WS2) environments. (f) Fraction of transpirable soil water in WW, WS1, and WS2 during 2015 and 2016 off-seasons. Horizontal bars indicate the water stress application periods. The three red bars and lines for WS2 represent three maturity groups (E, early maturity; M, medium maturity; L, late maturity) that the panel was divided into so that postflowering water stress could be applied consistently relative to flowering. (g) Cross genotype × environment interaction ($p < .04$; interaction term in one-tailed ANOVA) of the preflowering (Tx7000) and postflowering (BTx642) drought-tolerance checks across WW, WS1, and WS2 environments. The error bars represent 95% confidence intervals

$P < .1$) accessions. The average GrN was not significantly different between botanical types (Figure 2b). Significant correlations were observed among yield components, including GrW, DBM, and STI for GrW, across WS1 and WS2 regimes (Supplemental Figure S2a). High positive correlation was observed between BLUPs of GrW, PW, DBM, and GrN, while TGrW was negatively correlated with GrN (Supplemental Figure S2b). Significant correlations were observed between DBM in WS1 and WS2 and other yield components—GrW, GrN, DFLo, and PH—across RF conditions (Supplemental Figure S2c,d). Overall, genetic differences contributed to the phenotypic variation in managed water stress conditions.

3.3 | GWAS of flowering time

To identify loci potentially underlying quantitative trait variation in West African sorghum, we carried out GWAS using 130,709 SNP markers. First, we considered DFLo under

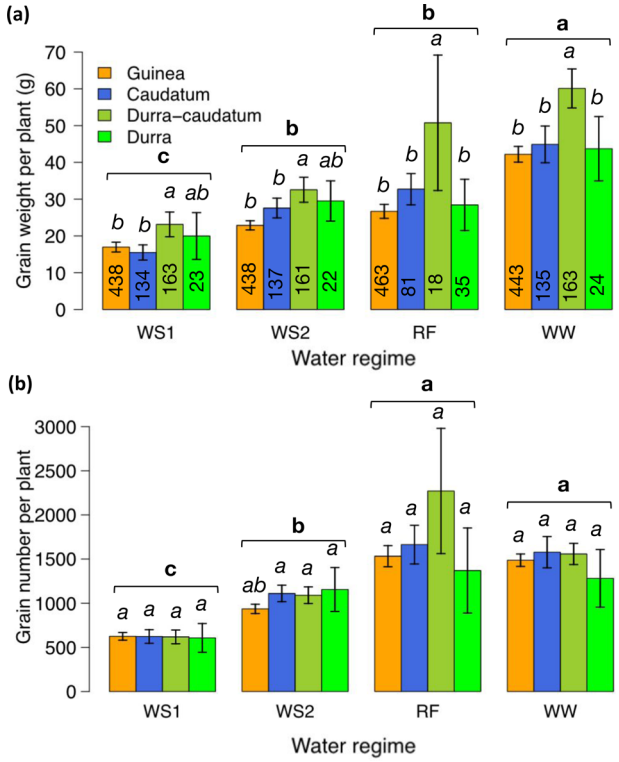

FIGURE 2 Effects of managed drought stress on grain yield components. (a) Differences in grain weight among botanical types within each water regime, rainfed condition (RF), well-watered (WW), preflowering water stress (WS1), and postflowering water stress (WS2). Numbers within bar plots indicate the number of genotypes per botanical type in each water regime (two environments in each). (b) Differences in grain number among botanical types within each water regime. The letters indicate Tukey honest statistical difference at $\alpha = 0.05$, with bold letters indicating the across water regime comparison and italics letters representing the across botanical type comparison within the water regime

WW off-season environments of 2015, 2016, and 2017 and BLUPs across all off-season environments to map known flowering time candidate genes using GLM+Q. No significant peak above the Bonferroni-corrected P value of .05 was identified for DFLo of the 2015 data, but significant associations were identified for DFLo of the 2016 and 2017 data (Supplemental Figure S3). Two SNPs, S6_55280640 and S3_62811196, were significantly associated with DFLo in both years and colocalized with a priori candidate flowering time genes *SbZfl1* (Sobic.006G201600; 9 kb away) and *SbCN12* (Sobic.003G295300; 61 kb away), respectively. In both 2016 and 2017, S6_55280640 was the lead SNP ($P < 10^{-10}$ in 2016; $P < 10^{-10}$ in 2017) of the associated region on chromosome 6. A third SNP, S2_67812515, was significantly associated with DFLo in 2017 data and colocalized with the a priori candidate gene *Maturity2* (Sobic.002G302700; 70 kb away). Significant associations were not identified above the Bonferroni threshold ($P > 10^{-5}$) when the MLM with PC analysis and kinship matrix were used to account for both population structure and genetic relatedness effects (Supplemental Figure S3).

The same associated SNPs near *SbZfl1* and *SbCN12*, noted above, were observed for flowering time BLUPs across all off-season environments (Figure 3a; Supplemental File S1). Lead SNP S6_55280640 was located one gene away from *SbZfl1* (Figure 3c). The *T* allele of S6_55280640, associated with shorter flowering times under RF conditions (Figure 3d), had a wide geographic distribution and was found at high frequency in accessions of the Sahel, Ethiopia, and western India (Figure 3e). Lead SNP S3_62811196 was the top association near *SbCN12* (Figure 3f). The *T* allele of S3_62811196, associated with short flowering times under RF conditions (Figure 3g), is globally rare and found mostly in accessions from Niger and northern Nigeria (Figure 3h).

3.4 | GWAS for drought tolerance

To generate hypotheses on the loci that underlie drought-tolerance variation in sorghum, we performed GWAS for GrW STI and the reduction of PW (RPW), DBM (RDBM), GrN (RGrN), PH (RPH), TGrW (RTGrW) in water stress environments. We considered water stress scenarios separately (WW vs. WS1 and WW vs. WS2) and together (WW vs. WS1 and WS2). In total, 222 and 214 associations were identified by the GLM+Q and MLM, respectively, for drought-response variables and STI in all drought-stress environments (Supplemental File S2; Supplemental Figure S4). Among the associations, 134 were commonly identified by both GWAS methods.

To determine QTL that have positive pleiotropic effect on pre- and postflowering drought tolerance among the associations above, we looked for common associations across different water stress environments. We defined a pleiotropic QTL as one lead SNP or locus being mapped in both pre- and postflowering drought scenarios or associated with several drought-response variables. Among the associations, 16 putative pleiotropic associations for drought responses were observed across water stress environments (Supplemental Table S2). For example, the SNP S4_67777846 was associated with STI under WS1 of 2016 and 2017 and WS2 of 2017 using both GLM+Q and MLM. The SNPs S3_13763609 and S1_74186408 were associated with RPW in WS1 and WS2 of 2017 using both GLM+Q and MLM. The identified pleiotropic lead SNPs showed significant allelic effect and significantly ($P < 10^{-8}$) explained 11–25% of STI for GrW across water-deficit treatments (Supplemental Table S2). Of the 16 putative pleiotropic associations, six associations (S4_67777846, S2_18195896, S9_57781496, S10_1402513, and S6_55048997) overlapped with associations identified for the STI BLUPs across water stress environments (Figure 4; Supplemental File S3).

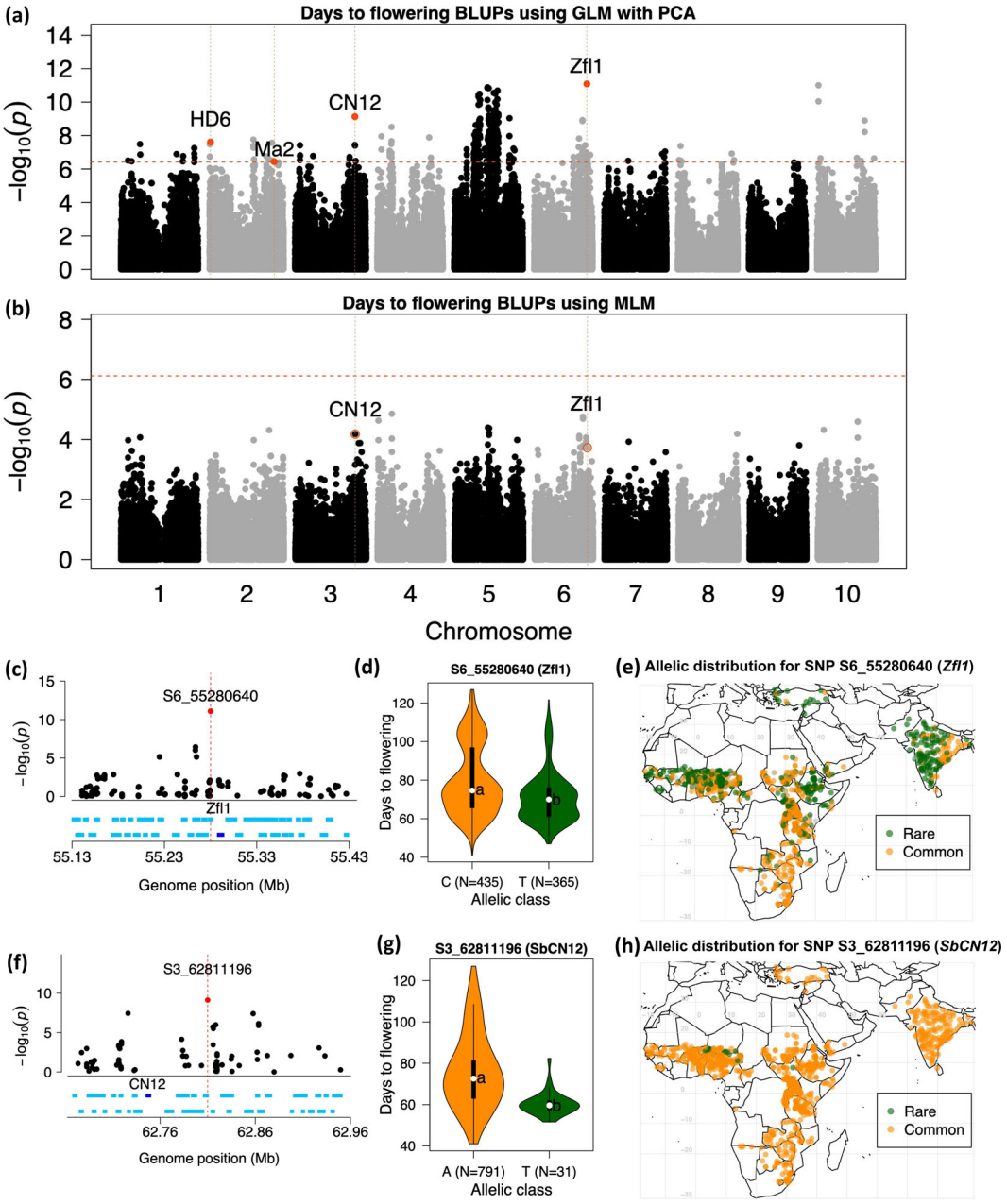

FIGURE 3 Genome-wide associations for flowering time. (a, b) Manhattan plots for days to flowering (DFLo) for best linear unbiased predictors (BLUPs) across all off-season environments over 3 yr using (a) general linear model with principal components (GLM+Q) and (b) mixed-linear model (MLM). Horizontal dashed line indicates the Bonferroni correction at 0.05. Red dots indicate peak single-nucleotide polymorphisms (SNPs) colocalizing (based on 150-kb cutoff) with a priori candidate genes for flowering time. (c) Zoomed-in Manhattan plot for the GLM+Q of a 150 kb region on chromosome 6 around the lead associated SNP, S6_55280640, that colocalizes with a priori candidate gene *Zfl1* (dark blue segment). (d) Days to flowering across rainfed environments by allelic classes of S6_55280640 in the West African sorghum association panel (WASAP). Letters within violin pots indicate Tukey's honest significant difference at $\alpha = 0.05$. (e) Geographic distribution of early (T) and late (C) flowering-associated alleles of S6_55280640 in global sorghum landraces. (f) Zoomed-in Manhattan plot for the GLM+Q of a 150 kb region on chromosome 3 around the lead associated SNP, S3_62811196, that colocalizes with a priori candidate gene *SbCN12* (dark blue segment). (g) Days to flowering across rainfed environments by allelic classes of S3_62811196 in the WASAP. (h) Geographic distribution of the early (T) and late (A) flowering-associated alleles of S3_62811196 in global sorghum landraces

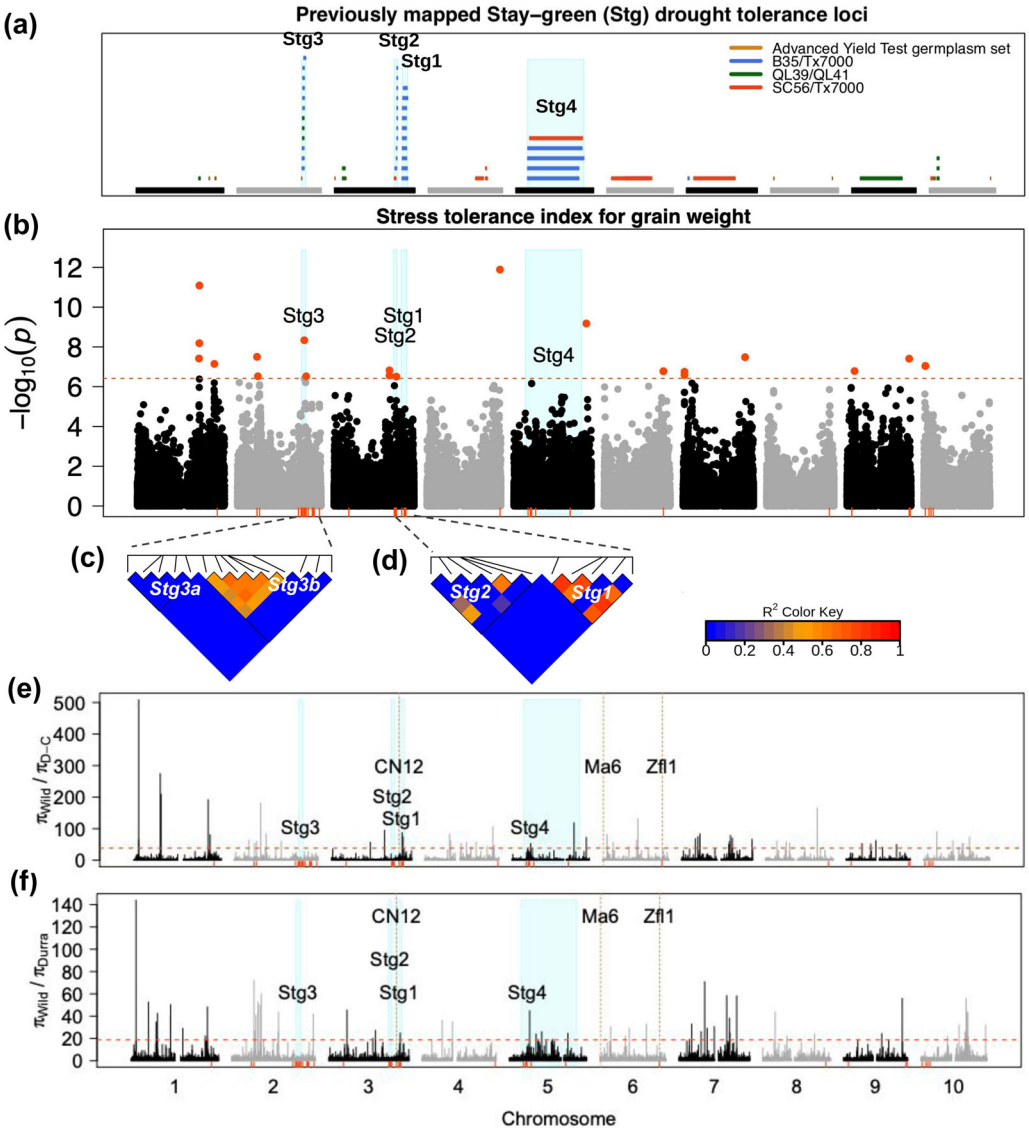

FIGURE 4 Genome-wide associations for drought tolerance and genome scans for adaptation. (a) Genomic location of the different stay-green quantitative trait loci (QTL), including *Stg1–4*, obtained from the Sorghum QTL Atlas. Light blue bars indicate the genomic position of *Stg1–4* intervals. (b) Manhattan plots of best linear unbiased predictors of stress tolerance index (STI) for grain weight across preflowering (WS1) and postflowering (WS2) water-stressed treatments of 2015–2017 based on general linear model with principal components (GLM+Q). The horizontal red dashed line represents the Bonferroni significance threshold at 0.05 and red dots indicate lead single-nucleotide polymorphisms (SNPs) above the threshold. Some lead SNPs colocalize within *Stg1, Stg2, Stg3*, and *Stg4* loci that are represented by light blue bars. Rug plots indicate the genomic position of the putative pleiotropic lead SNPs and lead SNPs at *Stg1–4* loci, significantly associated with grain weight STI and drought response variables: reduction of panicle weight (RPW), aboveground dry biomass (RDBM), grain number (RGrN), plant height (RPH), and 1,000-grain weight (RTGrW). (c) Linkage disequilibrium heatmap for lead SNPs at *Stg3a* (left triangle) and *Stg3b* (right triangle). (d) Linkage disequilibrium heatmap for lead SNPs at *Stg2* (left triangle) and *Stg1* (right triangle). (e, f) Reduction of nucleotide diversity, based on 100-kb sliding windows in (e) durra–caudatum (D-C) and (f) durra landraces relative to wild sorghums. Red dashed horizontal lines indicate the 99th percentile threshold for signatures of selection outliers. Rug plots in red indicate lead SNPs for putative pleiotropic lead SNPs and lead SNPs within *Stg1–4* loci associated with drought response variables. Light blue bars indicate the genomic position of *Stg1–4* intervals

3.5 | Drought-response associations colocalizing with *Stg* loci

To test the hypothesis that *Stg* loci identified from Ethiopia, we characterized the colocalization of GWAS peak SNPs with previously defined *Stg* QTL intervals as summarized in the sorghum QTL Atlas. The interval of *Stg3* (*Stg3a* and *Stg3b*) was defined based on the introgressed region by the ICRISAT breeding program (Vadez et al., 2013). Of the total lead SNPs associated with STI for GrW and drought-response variables, 78 overlapped with 54 QTL of the published *Stg* QTL (Supplemental File S4, Figure 4a,b), which represents a significant

enrichment ($P < 10^{-16}$). Thirty lead SNPs colocalized with known *Stg1–4* loci (Supplemental Table S3). The lead SNPs colocalizing with each locus could explain up to 16% ($P < 10^{-10}$, *Stg1*), 20% ($P < 10^{-13}$, *Stg2*), 19% ($P < 10^{-13}$, *Stg3a*), 27% ($P < 10^{-16}$, *Stg3b*), and 21% ($P < 10^{-15}$, *Stg4*) of the phenotypic variation across WS1 and WS2 over years based on STI BLUPs. At *Stg2*, SNP S3_56094063 was the top association ($P_{GLM} < 10^{-19}$, $P_{MLM} < 10^{-13}$) for STI in WS2 and WS1. At *Stg3b*, S2_62095163 was the top association ($P_{GLM} < 10^{-18}$, $P_{MLM} < 10^{-13}$) with high effect for STI in WS2. The remaining lead SNP associations did not colocalize with *Stg* loci.

3.6 | Putative pleiotropic drought-response associations colocalizing with *Stg* loci

At each of the *Stg1–4* loci, there were several associations observed across two or more drought scenarios or drought-response variables (Supplemental Table S3). The *Stg3a* and *Stg3b* (which are next to each other) region covered associations for STI in WS1 of 2015 and 2016, STI in WS2 of 2016 and 2017, RPW in WS1 of 2017, and RDBM in WS1 of 2016. There was a strong LD among the lead SNPs within *Stg3b* but no LD among lead SNPs within *Stg3a* (Figure 4c). The SNP S2_62095163 was in strong LD with other lead SNPs in *Stg3b* but not in LD with lead SNPs in *Stg3a*. The *Stg2* locus colocalized with putative pleiotropic associations for STI in WS1 of 2015 and 2017, WS2 of 2017, RGrN in WS1 of 2017, and RDBM in WS1 of 2016. The *Stg1* locus covered associations for RPW in WS1 and WS2 of 2017 and associations for STI in WS1 of 2017. At both *Stg1* and *Stg2*, there was strong LD among several lead SNPs (Figure 4d). At the *Stg4* locus, there were associations for RPW in WS1 of 2017 and for STI in WS1 of 2015 and in WS2 of 2017 and moderate LD among lead SNPs (Supplemental Figure S4f).

3.7 | Evolutionary signals around drought-tolerance loci

To investigate the possibility of positive selection for drought tolerance at *Stg* loci, we conducted a genome scan of pairwise nucleotide diversity (π) ratios for West African sorghums relative to wild relatives (i.e. outliers with high $\pi_{sorghum}/\pi_{wild}$ ratio), considering 95th and 99th percentile outliers (Figure 4e,f; Supplemental Figure S5). Twelve of the lead SNPs associated with RPW and GrW STI overlapped with π ratio outliers in durra–caudatums and durras but not in guineas (Table 2). Colocalizations of π ratio outliers with *Stg1–4* loci were significantly enriched ($P < 10^{-16}$). In durra–caudatums and durras, but not in guineas, some 99th percentile π ratio outliers were localized within *Stg1* (Figure 4e,f; Supplemental Figure S5). We characterized the geographic distribution of two selected lead associations within each *Stg* locus to determine whether the *Stg* alleles are rare and only involved in local adaptation or are broadly involved in adaptation across sorghum landraces beyond known sources in Ethiopia sorghums (Figure 5a,b). The rare alleles associated with increased STI at a few selected lead SNPs within *Stg1–3* were broadly geographically distributed in sorghum landraces (Figure 5c–h) (*Stg4* was excluded because of its large interval). However, the increased STI-associated allele at lead SNPs that overlapped with high π ratio outliers were found mostly in West African sorghums (Figure 5h) except for S3_66366589 (Figure 5g).

4 | DISCUSSION

4.1 | How well do managed stress trials reveal the genetics of drought tolerance in sorghum?

In this study, we sought to better understand genetics of drought adaptation in sorghum, a crop that is well known for drought tolerance but whose mechanisms of drought adaptation are not yet understood (Choudhary et al., 2020). We characterized a diverse panel of West African sorghum germplasm in common-garden managed drought-stress field trials with the aim of balancing experimental repeatability (via the use of irrigation in off-season) with biological and breeding relevance (via the use of a field environment) (Cooper et al., 2014). The usefulness of managed-stress experiments to understand crop evolution and improve crop resilience depends on several criteria we consider in turn:

1. Was the intended stress applied? Two lines of evidence support the contention that the intended drought stress was successfully imposed via irrigation management in the off-season. First, the measured soil moisture was consistently high in WW control treatment (FTSW $\approx$ 0.6) but dropped to ~0.2 at the intended times in pre- or postflowering drought-stress treatments (Figure 1f). The FTSW values achieved in WS1 and WS2 were similar to the critical values (~0.2–0.5) for decreases in transpiration and leaf expansion in diverse sorghum lines (Choudhary et al., 2020), suggesting that a physiologically relevant stress was experienced by the plants. Second, we observed a substantial (but not complete) reduction of yield components (~50%; Figure 2) under managed drought stress (WS1 and WS2 relative to WW), suggesting the stress was also agronomically relevant (Blum, 2010).

2. Was the stress comparable to previous stress experiments? To be able to address this question, we included two international drought-tolerance check lines, which are the canonical postflowering (BTx642) and preflowering (Tx700) drought-tolerant genotypes based on many studies in the

TABLE 2 Pair-wise wide nucleotide diversity (π) ratio outliers overlapping with some lead single-nucleotide polymorphism (SNP) associations and their allele frequency in each botanical type

Chr.	Locus	Lead SNP	Nucleotide diversity ratio[a]			Common allele frequency		
			πW/πDC	πW/πD	πW/πG	DC	Durra	Guinea
1	–	S1_74186408	**18.2**	3	–	0.5	0.5	0.46
2	–	S2_18195896	2.4	**9**	–	0.49	0.5	0.47
2	–	S2_20558788	**61.3**	**72.2**	**54.6**	0.49	0.5	0.45
2	*Stg3b*	S2_71386056	2.1	**7.8**	–	0.5	0.5	0.44
2	–	S2_76213690	**13.4**	3.3	–	0.45	0.26	–
3	*Stg1*	S3_62836558	2.8	2.2	**7.5**	0.5	0.5	0.45
3	*Stg1*	S3_65137990	**15.1**	1.7	–	0.5	0.5	0.46
3	*Stg1*	S3_65430305	**88.2**	**15.5**	**8.4**	0.5	0.5	0.46
3	*Stg1*	S3_66366589	**74.6**	**25**	–	0.5	0.46	0.46
5	*Stg4*	S5_15215761	**32**	6.8	**10.1**	0.5	0.48	0.48
5	*Stg4*	S5_16480120	**27**	**13**	–	0.45	0.41	0.48
5	*Stg4*	S5_20251208	1	1.5	**8.7**	0.45	0.46	0.5

Note. Bold numbers are among the 95th percentile of π ratios.

[a]DC, durra–caudatum landraces; D, durra landraces; G, guinea landraces; W, wild sorghums.

United States and Australia (Borrell et al., 2014b; Burke et al., 2013; Tuinstra et al., 1996). Consistent with the idea that our managed drought stress was comparable to natural and managed drought stress in the United States and Australia, a crossover G × E interaction ($p < .04$; one-tailed ANOVA) for grain yield of Tx7000 vs. BTx642 under pre- vs. postflowering drought in the expected direction (Figure 1g). These two genotypes did not perform well on a per-plant basis in the different environments relative to local cultivars in the present study. This lack of yield performance may be explained by their adaptation to temperate regions relative to tropical regions.

3. Was the timing and severity of stress comparable to that in the target population of environment (TPE)? Among the three criteria, this is the most difficult to assess. A formal envirotyping study, which quantifies the frequency of particular water deficits relative to crop phenology, would be necessary to address this question (Chenu et al., 2011; Cooper et al., 2014). One particular concern for off-season managed stress would be that differences in photoperiod regime relative to the TPE (i.e. the rainy season) could change in growth or developmental dynamics in a way that alters the drought response (Blum, 2010; Gano et al., 2021). However, the overall similarity of plant grain yield components in the rainy season (RF) and off-season experiments (Supplemental Figures S1a,e and S2c,d) suggest that the managed drought stress is broadly comparable to drought in the TPE. Ultimately, to rigorously test hypotheses on the similarity of off-season managed drought to the drought in the TPE, a comparison of phenotypes under managed stress to multienvironment field trials under natural drought stress will be necessary (Cooper et al., 1995).

4.2 | Evidence for a role of *SbZfl1* and *SbCN12* in flowering time variation and for *SbCN12* in drought adaptation

Flowering time is a critical component of geographic adaptation (Castelletti et al., 2020; Lasky et al., 2015) and a potential contributor to drought adaptation via early flowering drought escape (Blum, 2010). Among the six canonical sorghum photoperiodic flowering genes (*Maturity1–Maturity6*) characterized in U.S. germplasm (Casto et al., 2019; Murphy et al., 2011, 2014; Yang et al., 2014), we identified colocalization of associations only at *Ma2* (Figure 3a). Instead, the top QTL mapped two a priori flowering time candidate loci, *SbZfl1* (chr6: 55.289–55.293 Mb) and *SbCN12* (chr3: 62.747–62.749 Mb), which are not known to underlie genetic variation in U.S. germplasm (Figure 3; Supplemental Figure S3).

SbZfl1 is the ortholog of maize *ZFL1/2* and rice (*Oryza sativa* L.) *RLF*, which induce early flowering by activating vegetative-to-reproductive transition (Bomblies & Doebley, 2006; Rao et al., 2008). While *SbZfl1* variation has not been previously identified via linkage mapping (Mace et al., 2019), *SbZfl1* was identified as a top candidate in a recent GWAS of photoperiodic flowering rating in a Senegal regional panel (Faye et al., 2019). The minor allele frequency of the *SbZfl1* QTL was high (>0.4) in both WASAP and global georeferenced landraces (Figure 3e), suggesting a common, moderate-effect variant exists at *SbZfl1*. Sorghum is a short-day species, so under short days (e.g., the cool off-season in West Africa; Figure 1b) it is expected to flower early regardless of photoperiodism. Given *SbZfl1* was the top flowering time association under short days, *SbZfl1* may be a regulator of basic

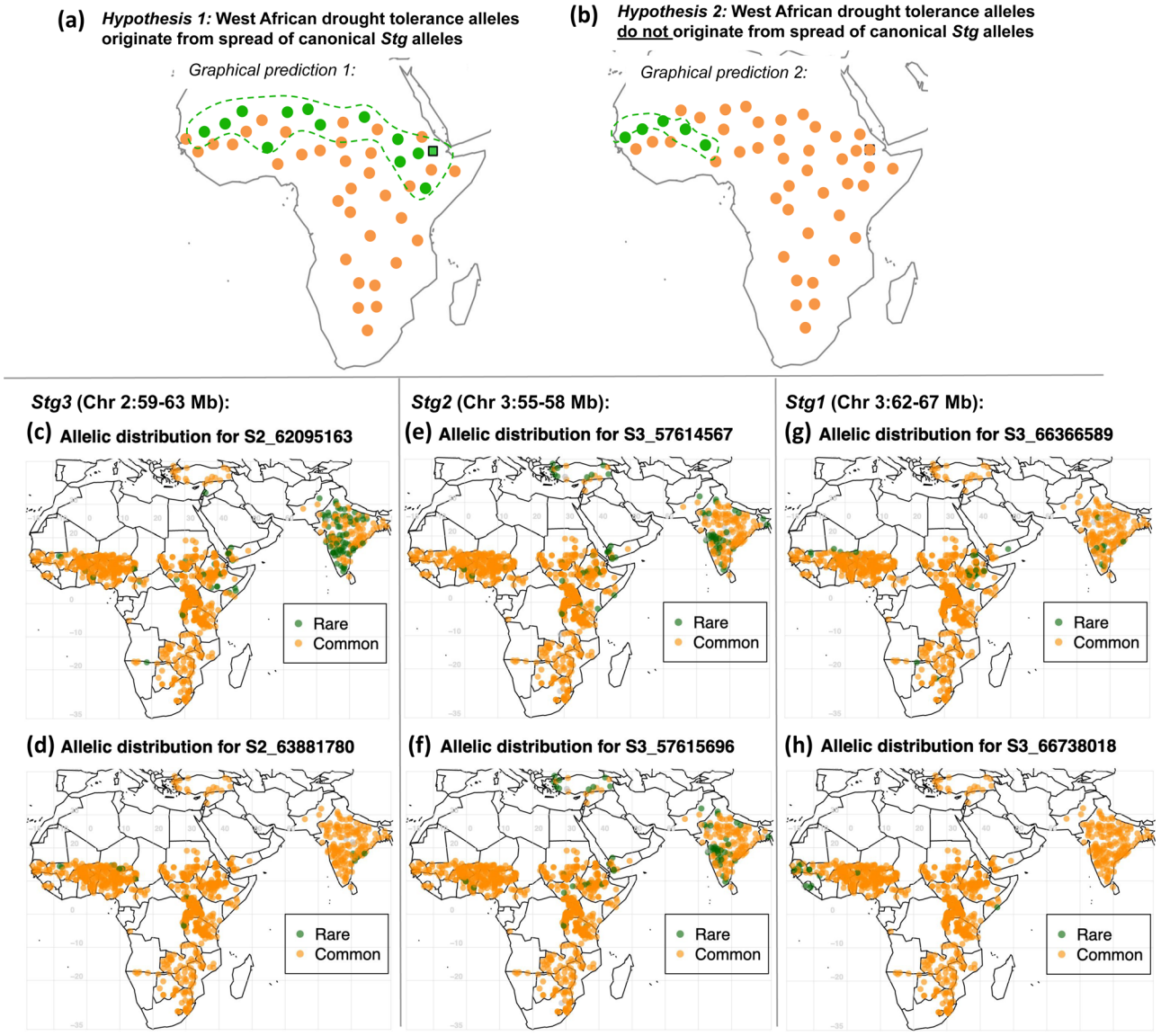

FIGURE 5 Evidence for a broad role of canonical stay-green alleles in drought adaptation. (a, b) Competing hypotheses on the origin of West African drought-tolerance alleles and relationship with the canonical *Stg* alleles (titles) and graphical predictions under each hypothesis (diagrams). Under Hypothesis 1 (panel a) some West African drought-tolerance alleles represented *Stg* alleles that have diffused from eastern Africa, while under Hypothesis 2 (panel b) these alleles are unrelated to *Stg* alleles. The location of the known *Stg* allele source, accession IS12555 from Ethiopia, is noted by the black square. (c–h) Observed global allelic distributions at some West African drought-tolerance MTA that colocalize with *Stg1–3*. (c, d) Geographic distribution of the common allele (orange) and rare allele (green) associated with increased drought tolerance of genome-wide association studies (GWAS) lead SNPs, (c) S2_62095163 and (d) S2_63881780, in the *Stg3* locus. (e, f) Geographic distribution of the common allele (orange) and rare allele (green) associated with increased drought tolerance of GWAS lead SNPs, (e) S3_57614567 and (f) S3_57615696, in the *Stg2* locus. (g, h) Geographic distribution of the common allele (orange) and rare allele (green) associated with increased drought tolerance of GWAS lead SNPs, (g) S3_66366589 and (h) S3_66738018 in the *Stg1* locus, respectively. The different GWAS lead SNPs above were selected based on their association with drought tolerance, proportion of variance explained, colocalization within *Stg1–3* loci, linkage disequilibrium with other lead SNPs within each *Stg1-3* locus, and availability in the genotype-by-sequencing data for global sorghum landraces. Note, lead SNPs in *Stg4* locus were not included because of the large interval for this locus

vegetative phase, the thermal time component of flowering regulation that acts independently of photoperiodic flowering regulation (Guitton et al., 2018). This hypothesis could explain the lack of a flowering time QTL at *SbZfl1* in a previous GWAS under long days (rainy season) in the WASAP (Faye et al., 2021); subtle basic vegetative phase variation due to *SbZfl1* could have been masked by large-effect photoperiodic variants at *Ma6*, *SbCN8*, or other loci. However, this hypothesis would not explain the photoperiod flowering association at *SbZfl1* previously observed in Senegalese

germplasm (Faye et al., 2019). Given inherent limitations of association studies (Korte & Farlow, 2013) and the complexity of photothermal flowering (Li et al., 2018b), linkage mapping and ecophysiological modeling will be needed to test these hypotheses on the role of *SbZfl1* in flowering time adaptation (Guitton et al., 2018; Li et al., 2018b).

SbCN12 (also known as *SbFT8*) is a co-ortholog of the canonical florigen *FT* in *Arabidopsis thaliana* (L.) Heynh and ortholog of maize *ZCN12* (Castelletti et al., 2020; Yang et al., 2014), which was identified as a likely sorghum florigen based on conserved sequence and expression dynamics (Wolabu et al., 2016; Yang et al., 2014). The current GWAS findings, along with the previous finding that *SbCN12* explained up to 12% of the variation in the global nested association mapping population (Bouchet et al., 2017; Hu et al., 2019), provide strong support for the hypothesis that functional allelic variation exists at *SbCN12*. Given that the early flowering associated allele near *SbCN12* is globally rare (Figure 3h), it may be a useful new allele for sorghum breeding programs targeting earlier flowering for stress escape. Molecular cloning of causative variants at *SbCN12* and *SbZfl1* could shed light on their role in flowering time evolution (Bomblies & Doebley, 2006; Castelletti et al., 2020) and facilitate the development of molecular markers to recover locally adaptive flowering time.

The evidence for a role of these flowering time genes in drought adaptation (e.g. via drought escape) is mixed. On one hand, *SbCN12* colocalized with a drought response association (RPW for WS1 vs. WW; S3_62836558; 64 kb away; Supplemental Table S3), so could plausibly underlie some variation for preflowering drought response of this yield component. Also, the same SNP near *SbCN12* was in a window of reduced π in guinea genotypes (Table 2), suggesting selection at this locus (note, this is not the same SNP as the rare flowering time associated variant S3_62811196, but a common variant 25 kb away). On the other hand, *SbZfl1* did not colocalize with the drought response QTL (STI, RPW, etc.; the nearest association with STI, S6_55048997, was at ~240 kb away) and there was no evidence of positive selection around *SbZfl1* based on the π ratios (Figure 4; Supplemental Figure S5). Given that causative variants at *SbCN12* and *SbZfl1* are not yet known, hypotheses on the role of these genes in drought adaptation remain speculative but could be tested using near-isogenic lines.

4.3 | Insights on the genetics of drought adaptation in sorghum

The botanical types of sorghum vary strikingly in their morphology and geographic distribution and based on a phytogeographic adaptation model (Vavilov, 2009). It has long been hypothesized that they vary in their drought adaptedness (Harlan & De Wet, 1972; Lasky et al., 2015; Tao et al., 2020; Wang et al., 2020). For instance, durra sorghums, which predominate in arid regions, are thought to be the most drought tolerant (Harlan & De Wet, 1972), while guinea sorghums, which predominate in humid regions, are thought to be adapted to high humidity (De Wet et al., 1972). However, previous studies of large sorghum diversity panels under managed drought stress have not directly tested this hypothesis, for instance, by comparison of drought response for yield among botanical types (Lasky et al., 2015; Upadhyaya et al., 2017; Vadez et al., 2011). Surprisingly, in this study, we find no evidence of overall differences in drought tolerance among botanical types based on the drought response of yield components (Figure 2). These findings could be explained by one of two competing hypotheses. First, it is possible that the drought scenarios we applied do not correspond to the drought scenarios in the TPE, such that true differences in drought tolerance among botanical types were not reflected in the phenotypes. Alternatively, it may be that the major botanical types in West Africa all harbor substantial drought tolerance, presumably because droughts are common even in the higher precipitation portions of the sorghum range (Traore et al., 2014). It could be also that drought tolerance, as measured on per-plant basis, observations may miss yield differences on a per-plot basis between botanical types. In either case, our findings suggest that long-held views on differential drought adaptation among botanical types in sorghum require more formal testing.

Theoretical considerations on water use tradeoffs (Tardieu, 2012) and the apparent lack of sorghum genotypes harboring both pre- and postflowering drought tolerance (Burke et al., 2013) suggest that a tradeoff might exist between early- vs. late-stage tolerance mechanisms. However, the moderate positive correlation of the grain yield estimates under pre- and postflowering drought (e.g. for GrW or STI; Supplemental Figure S2a) suggests no major physiological tradeoff for tolerance to these contrasting drought scenarios, at least at this broad scale of diversity. Colocalization of MTA for drought-tolerance-related traits in WS1 and WS2 would provide further evidence for genetic factors that contribute pleiotropically to both pre- and postflowering drought tolerance. Consistent with this hypothesis, 16 distinct MTAs (Supplemental Table S2) were detected for drought-related traits (mostly STI) under both the pre- and postflowering drought treatments. Among the positive pleiotropic associations, the STI MTA at S4_67777846 may be the most interesting candidate for further study, given that it had the highest PVE estimate (25%) in both pre- and postflowering water stress over 2 yr. This MTA did not colocalize with *Stg* QTL or other a priori candidate genes, and there were no obvious post hoc candidate genes near the SNP. If confirmed, positive pleiotropic drought-tolerance QTL, which could contribute to yield stability across drought scenarios, would be of great interest for breeding of broadly adapted climate-resilient cultivars and

help elucidate mechanisms that circumvent potential tradeoffs (Tardieu, 2012).

Another question that motivated our study was whether canonical *Stg* alleles, which were originally discovered in Ethiopia-derived materials (BTx642) (Borrell et al., 2014a), are also present in West African landraces (Figure 5a). The hypothesis that canonical *Stg* alleles have a broad role in drought adaptation across Africa is plausible since it is well established that durra sorghums diffused from Ethiopia across the Sahel to West Africa (Harlan & Stemler, 1976; Morris et al., 2013). Indeed, we observe a statistically significant enrichment of drought-tolerance-related MTA colocalized with canonical *Stg* QTL intervals, which provides preliminary support for the shared *Stg* hypothesis File (Supplemental S4; Table S3). Most notable among these are the highly significant association for GrW STI under postflowering drought at *Stg3* (S2_62095163) and *Stg2* (S3_57614567). Further, the West Africa drought-tolerance-associated alleles in the *Stg* intervals are found in Ethiopia, as would be expected if they were shared across Africa. While these findings are suggestive, they are not sufficient to exclude the alternate hypothesis (Figure 5b) that West African drought-tolerance loci are unrelated to Ethiopian-origin *Stg* alleles. Testing this hypothesis conclusively would require positional cloning of the West Africa drought-tolerance QTL and *Stg* alleles.

The final hypothesis we considered was that drought-tolerance alleles underlie drought adaptation per se and were subject to positive selection. This finding was supported by significant enrichment for colocalization of selection outliers with *Stg* QTL and common allele frequencies of lead SNPs overlapping with selection outliers in durra–caudatums and durras relative to guineas (Figure 4e,f; Supplemental Figure S5; Table 2). As with the other findings, the development of near-isogenic lines and the validation of major-effect QTL in breeding populations (Borrell et al., 2014a,b) will be crucial to rigorously test the proposed role of QTL in these genomic regions for drought adaptation. Overall, our findings support the long-standing hypothesis that genetic variation for drought tolerance exists in West African sorghum and provide preliminary evidence for a broad role of canonical *Stg* drought-tolerance alleles across Africa.

DATA AVAILABILITY STATEMENT

All phenotype and genotype data, candidate genes, genome scan results, and analysis scripts are available at Dryad (https://doi.org/10.5061/dryad.1jwstqjw8).

ACKNOWLEDGMENTS

This study is made possible by the support of the American People provided to the Feed the Future Innovation Lab for Collaborative Research on Sorghum and Millet through the United States Agency for International Development (USAID) under Cooperative Agreement No. AID-OAA-A-13-00047. The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. Analyses were made possible by Beocat High-Performance Computing Cluster at Kansas State University.

AUTHOR CONTRIBUTIONS

Jacques M. Faye: Formal analysis, Visualization, Writing-original draft, Writing-review & editing. Eyanawa A. Akata: Data curation, Formal analysis, Investigation. Bassirou Sine: Conceptualization, Data curation, Investigation, Methodology, Supervision. Cyril Diatta: Investigation. Ndiaga Cisse: Conceptualization, Funding acquisition, Project administration, Supervision. Daniel Fonceka: Conceptualization, Funding acquisition, Project administration, Supervision. Geoffrey P. Morris: Conceptualization, Funding acquisition, Project administration, Supervision, Writing-original draft, Writing-review & editing.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Geoffrey P. Morris https://orcid.org/0000-0002-3067-3359

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Blum, A. (2010). *Plant breeding for water-limited environments*. Springer Publishing.

Bomblies, K., & Doebley, J. F. (2006). Pleiotropic effects of the duplicate maize *FLORICAULA/LEAFY* genes *zfl1* and *zfl2* on traits under selection during maize domestication. *Genetics*, *172*, 519–531. https://doi.org/10.1534/genetics.105.048595

Borrell, A. K., Mullet, J. E., George-Jaeggli, B., van Oosterom, E. J., Hammer, G. L., Klein, P. E., & Jordan, D. R. (2014a). Drought adaptation of stay-green sorghum is associated with canopy development, leaf anatomy, root growth, and water uptake. *Journal of Experimental Botany*, *65*, 6251–6263. https://doi.org/10.1093/jxb/eru232

Borrell, A. K., van Oosterom, E. J., Mullet, J. E., George-Jaeggli, B., Jordan, D. R., Klein, P. E., & Hammer, G. L. (2014b). Stay-green alleles individually enhance grain yield in sorghum under drought by modifying canopy development and water uptake patterns. *New Phytologist*, *203*, 817–830. https://doi.org/10.1111/nph.12869

Bouchet, S., Olatoye, M. O., Marla, S. R., Perumal, R., Tesso, T., Yu, J., Tuinstra, M., & Morris, G. P. (2017). Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. *Genetics*, *206*, 573–585. https://doi.org/10.1534/genetics.116.198499

Burke, J. J., Chen, J., Burow, G., Mechref, Y., Rosenow, D., Payton, P., Xin, Z., & Hayes, C. M. (2013). Leaf dhurrin content is a quantitative measure of the level of pre- and post-flowering drought tolerance in sorghum. *Crop Science*, *53*, 1056–1065. https://doi.org/10.2135/cropsci2012.09.0520

Castelletti, S., Coupel-Ledru, A., Granato, I., Palaffre, C., Cabrera-Bosquet, L., Tonelli, C., Nicolas, S. D., Tardieu, F., Welcker, C., & Conti, L. (2020). Maize adaptation across temperate climates was obtained via expression of two florigen genes. *PLOS Genetics*, *16*, e1008882. https://doi.org/10.1371/journal.pgen.1008882

Casto, A. L., Mattison, A. J., Olson, S. N., Thakran, M., Rooney, W. L., & Mullet, J. E. (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*, e0212154. https://doi.org/10.1371/journal.pone.0212154

Chenu, K., Cooper, M., Hammer, G. L., Mathews, K. L., Dreccer, M. F., & Chapman, S. C. (2011). Environment characterization as an aid to wheat improvement: Interpreting genotype–environment interactions by modelling water-deficit patterns in North-Eastern Australia. *Journal of Experimental Botany*, *62*, 1743–1755. https://doi.org/10.1093/jxb/erq459

Choudhary, S., Guha, A., Kholova, J., Pandravada, A., Messina, C. D., Cooper, M., & Vadez, V. (2020). Maize, sorghum, and pearl millet have highly contrasting species strategies to adapt to water stress and climate change-like conditions. *Plant Science*, *295*, 110297. https://doi.org/10.1016/j.plantsci.2019.110297

Cooper, M., Messina, C. D., Podlich, D., Totir, L. R., Baumgarten, A., Hausmann, N. J., Wright, D., & Graham, G. (2014). Predicting the future of plant breeding: Complementing empirical evaluation with genetic prediction. *Crop and Pasture Science*, *65*, 311–336https://doi.org/10.1071/CP14007

Cooper, M., Woodruff, D. R., Eisemann, R. L., Brennan, P. S., & DeLacy, I. H. (1995). A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat: Managed-environments for selection among genotypes. *Theoretical and Applied Genetics*, *90*, 492–502. https://doi.org/10.1007/BF00221995

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R., & Group, 1000 Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Wet, J. M. J., Harlan, J. R., & Kurmarohita, B. (1972). Origin and evolution of guinea sorghums. *East African Agricultural and Forestry Journal*, *38*, 114–119

Faye, J. M., Maina, F., Akata, E. A., Sine, B., Diatta, C., Mamadou, A., Marla, S., Bouchet, S., Teme, N., Rami, J.-F., Fonceka, D., Cisse, N., & Morris, G. P. (2021). A genomics resource for genetics, physiology, and breeding of West African sorghum. *The Plant Genome*, *14*, e20075. https://doi.org/10.1002/tpg2.20075

Faye, J. M., Maina, F., Hu, Z., Fonceka, D., Cisse, N., & Morris, G. P. (2019). Genomic signatures of adaptation to Sahelian and Soudanian climates in sorghum landraces of Senegal. *Ecology and Evolution*, *9*, 1–14. https://doi.org/10.1002/ece3.5187

Gano, B., Dembele, J. S. B., Tovignan, T. K., Sine, B., Vadez, V., Diouf, D., & Audebert, A. (2021). Adaptation responses to early drought stress of West Africa sorghum varieties. *Agronomy*, *11*, 443. https://doi.org/10.3390/agronomy11030443

Guitton, B., Théra, K., Tékété, M. L., Pot, D., Kouressy, M., Témé, N., Rami, J. -F., & Vaksmann, M. (2018). Integrating genetic analysis and crop modeling: A major QTL can finely adjust photoperiod-sensitive sorghum flowering. *Field Crops Research*, *221*, 7–18. https://doi.org/10.1016/j.fcr.2018.02.007

Harlan, J. R., & De Wet, J. J. M. (1972). A simplified classification of cultivated sorghum. *Crop Science*, *12*, 172–176. https://doi.org/10.2135/cropsci1972.0011183x001200020005x

Harlan, J. R., & Stemler, A. (1976). The races of sorghum in Africa. In J. R. Harlan, J. M. J. D. Wet, & A. B. L. Stemler (Eds.), *Origins of African plant domestication* (pp. 465–478). DE GRUYTER MOUTON.

Harris, K., Subudhi, P. K., Borrell, A., Jordan, D., Rosenow, D., Nguyen, H., Klein, P., Klein, R., & Mullet, J. (2007). Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. *Journal of Experimental Botany*, *58*, 327–338. https://doi.org/10.1093/jxb/erl225

Hayes, C. M., Weers, B. D., Thakran, M., Burow, G., Xin, Z., Emendack, Y., Burke, J. J., Rooney, W. L., & Mullet, J. E. (2016). Discovery of a dhurrin QTL in sorghum: Co-localization of dhurrin biosynthesis and a novel stay-green QTL. *Crop Science*, *56*, 104–112. https://doi.org/10.2135/cropsci2015.06.0379

Hu, Z., Olatoye, M. O., Marla, S., & Morris, G. P. (2019). An integrated genotyping-by-sequencing polymorphism map for over 10,000 sorghum genotypes. *The Plant Genome*, *12*, 180044. https://doi.org/10.3835/plantgenome2018.06.0044

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

Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., Juenger, T. E., Hyma, K., Acharya, C., Mitchell, S. E., Buckler, E. S., Brenton, Z., Kresovich, S., & Morris, G. P. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances*, *1*, e1400218. https://doi.org/10.1126/sciadv.1400218

Leiser, W. L., Rattunde, H. F., Weltzien, E., Cisse, N., Abdou, M., Diallo, A., Tourè, A. O., Magalhaes, J. V., & Haussmann, B. I. (2014). Two in one sweep: Aluminum tolerance and grain yield in P-limited soils are associated to the same genomic region in West African Sorghum. *BMC Plant Biology*, *14*, 206. https://doi.org/10.1186/s12870-014-0206-6

Li, D., Dossa, K., Zhang, Y., Wei, X., Wang, L., Zhang, Y., Liu, A., Zhou, R., & Zhang, X. (2018a). GWAS uncovers differential genetic bases for drought and salt tolerances in sesame at the germination stage. *Genes*, *9*, 87. https://doi.org/10.3390/genes9020087

Li, X., Guo, T., Mu, Q., Li, X., & Yu, J. (2018b). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 6679–6684. https://doi.org/10.1073/pnas.1718326115

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, *28*, 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Mace, E., Innes, D., Hunt, C., Wang, X., Tao, Y., Baxter, J., Hassall, M., Hathorn, A., & Jordan, D. (2019). The sorghum QTL atlas: A powerful tool for trait dissection, comparative genomics and crop improvement. *Theoretical and Applied Genetics*, *132*, 751–766. https://doi.org/10.1007/s00122-018-3212-5

Mace, E. S., Li, Y., Prentis, P. J., Bian, L., Campbell, B. C., Hu, W., Innes, D. J., Han, X., Cruickshank, A., Dai, C., Frère, C., Zhang, H., Hunt, C.

H., Wang, X., Shatte, T., Wang, M., Su, Z., Li, J., Lin, X., … Wang, J. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms3320

Mauboussin, J.-C., Gueye, J., & N'Diaye, M. (1977). L'amélioration du sorgho au Sénégal [Sorghum improvement in Senegal]. *Agronomie Tropicale*, *32*, 303–310

McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., Kennedy, M., Amirebrahimi, M., Weers, B. D., McKinley, B., Mattison, A., Morishige, D. T., Grimwood, J., Schmutz, J., & Mullet, J. E. (2018). The *Sorghum bicolor* reference genome: Improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *The Plant Journal*, *93*, 338–354. https://doi.org/10.1111/tpj.13781

Mendiburu, F. (2009). Agricolae: Statistical procedures for agricultural research. https://rdrr.io/cran/agricolae/man/agricolae-package.html

Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., Brown, P. J., Acharya, C. B., Mitchell, S. E., Harriman, J., Glaubitz, J. C., Buckler, E. S., & Kresovich, S. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 453–458. https://doi.org/10.1073/pnas.1215985110

Mullet, J., Morishige, D., McCormick, R., Truong, S., Hilley, J., McKinley, B., Anderson, R., Olson, S. N., & Rooney, W. (2014). Energy Sorghum–A genetic model for the design of C4 grass bioenergy crops. *Journal of Experimental Botany*, *65*, 3479–3489. https://doi.org/10.1093/jxb/eru229

Mundia, C. W., Secchi, S., Akamani, K., & Wang, G. (2019). A regional comparison of factors affecting global sorghum production: The case of North America, Asia and Africa's Sahel. *Sustainability*, *11*, 2135. https://doi.org/10.3390/su11072135

Murphy, R. L., Klein, R. R., Morishige, D. T., Brady, J. A., Rooney, W. L., Miller, F. R., Dugas, D. V., Klein, P. E., & Mullet, J. E. (2011). Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 16469–16474. https://doi.org/10.1073/pnas.1106212108

Murphy, R. L., Morishige, D. T., Brady, J. A., Rooney, W. L., Yang, S., Klein, P. E., & Mullet, J. E. (2014). *Ghd7* (*Ma6*) represses sorghum flowering in long days: *Ghd7* alleles enhance biomass accumulation and grain production. *The Plant Genome*, *7*, plantgenome2013.11.0040. https://doi.org/10.3835/plantgenome2013.11.0040

Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A. K., Chapman, J., Feltus, F. A., Gowik, U., … Rokhsar, D. S. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, *457*, 551–556. https://doi.org/10.1038/nature07723

Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., Zivot E., Lestel, M., Balkissoon, K., & Wuertz, D. (2014). Performanceanalytics: Econometric tools for performance and risk analysis. Version 1.4.4000 from r-Forge. https://rdrr.io/rforge/PerformanceAnalytics/

Rao, N. N., Prasad, K., Kumar, P. R., & Vijayraghavan, U. (2008). Distinct regulatory role for *RFL*, the rice *LFY* homolog, in determining flowering time and plant architecture. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 3646–3651. https://doi.org/10.1073/pnas.0709059105

R Core Team, R.C. (2016). *R, A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rodríguez-Álvarez, M. X., Boer, M. P., van Eeuwijk, F. A., & Eilers, P. H. C. (2018). Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spatial Statistics*, *23*, 52–71. https://doi.org/10.1016/j.spasta.2017.10.003

Shin, J.-H., Blay, S., McNeney, B., & Graham, J. (2006). Ldheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *Journal of Statistical Software*, *16*, 1–9. https://doi.org/10.18637/jss.v016.c03

Tao, Y., Zhao, X., Wang, X., Hathorn, A., Hunt, C., Cruickshank, A. W., Oosterom, E. J. van, Godwin, I. D., Mace, E. S., & Jordan, D. R. (2020). Large-scale GWAS in sorghum reveals common genetic control of grain size among cereals. *Plant Biotechnology Journal*, *18*, 1093–1105. https://doi.org/10.1111/pbi.13284

Tardieu, F. (2012). Any trait or trait-related allele can confer drought tolerance: Just design the right drought scenario. *Journal of Experimental Botany*, *63*, 25–31. https://doi.org/10.1093/jxb/err269

Traore, S. B., Ali, A., Tinni, S. H., Samake, M., Garba, I., Maigari, I., Alhassane, A., Samba, A., Diao, M. B., Atta, S., Dieye, P. O., Nacro, H. B., & Bouafou, K. G. M. (2014). AGRHYMET: A drought monitoring and capacity building center in the West Africa region. *Weather and Climate Extremes*, *3*, 22–30. https://doi.org/10.1016/j.wace.2014.03.008

Tuinstra, M. R., Grote, E. M., Goldsbrough, P. B., & Ejeta, G. (1996). Identification of quantitative trait loci associated with pre-flowering drought tolerance in sorghum. *Crop Science*, *36*, 1337–1344. https://doi.org/10.2135/cropsci1996.0011183X003600050043x

Tuinstra, M. R., Grote, E. M., Goldsbrough, P. B., & Ejeta, G. (1997). Genetic analysis of post-flowering drought tolerance and components of grain development in *Sorghum bicolor* (L.) Moench. *Molecular Breeding*, *3*, 439–448. https://doi.org/10.1023/A:1009673126345

Upadhyaya, H. D., Dwivedi, S. L., Vetriventhan, M., Krishnamurthy, L., & Singh, S. K. (2017). Post-flowering drought tolerance using managed stress trials, adjustment to flowering, and mini core collection in sorghum. *Crop Science*, *57*, 310–321. https://doi.org/10.2135/cropsci2016.04.0280

Vadez, V., Deshpande, S., Kholova, J., Ramu, P., & Hash, C. T. (2013). Molecular breeding for stay-green: Progress and challenges in sorghum. In R. K. Varshney & R. Tuberosa (Eds.), *Translational genomics for crop breeding* (pp. 125–141). John Wiley & Sons Ltd.

Vadez, V., Krishnamurthy, L., Hash, C. T., Upadhyaya, H. D., & Borrell, A. K. (2011). Yield, transpiration efficiency, and water-use variations and their interrelationships in the sorghum reference collection. *Crop and Pasture Science*, *62*, 645–655. https://doi.org/10.1071/CP11007

Vavilov, N. I. (2009). *Origin and geography of cultivated plants* (1st ed.). Cambridge University Press.

Wang, J., Hu, Z., Upadhyaya, H. D., & Morris, G. P. (2020). Genomic signatures of seed mass adaptation to global precipitation gradients in sorghum. *Heredity*, *124*, 108–121. https://doi.org/10.1038/s41437-019-0249-4

Wendorf, F., Close, A. E., Schild, R., Wasylikowa, K., Housley, R. A., Harlan, J. R., & Królik, H. (1992). Saharan exploitation of plants 8,000 years BP. *Nature*, *359*, 721–724. https://doi.org/10.1038/359721a0

Wolabu, T. W., Zhang, F., Niu, L., Kalve, S., Bhatnagar-Mathur, P., Muszynski, M. G., & Tadege, M. (2016). Three *FLOWERING*

LOCUS T-like genes function as potential florigens and mediate photoperiod response in sorghum. *New Phytologist*, *210*, 946–959. https://doi.org/10.1111/nph.13834

Xu, W., Subudhi, P. K., Crasta, O. R., Rosenow, D. T., Mullet, J. E., & Nguyen, H. T. (2000). Molecular mapping of QTLs conferring stay-green in grain sorghum (*Sorghum bicolor* L. Moench). *Genome*, *43*, 461–469. https://doi.org/10.1139/g00-003

Yang, S., Murphy, R. L., Morishige, D. T., Klein, P. E., Rooney, W. L., & Mullet, J. E. (2014). Sorghum phytochrome B inhibits flowering in long days by activating expression of *SbPRR37* and *SbGHD7*, repressors of *SbEHD1*, *SbCN8* and *SbCN12*. *PLoS ONE*, *9*, e105352. https://doi.org/10.1371/journal.pone.0105352

Yuan, Y., Xing, H., Zeng, W., Xu, J., Mao, L., Wang, L., Feng, W., Tao, J., Wang, H., Zhang, H., Wang, Q., Zhang, G., Song, X., & Sun, X.-Z. (2019). Genome-wide association and differential expression analysis of salt tolerance in *Gossypium hirsutum* L at the germination stage. *BMC Plant Biology*, *19*, 394. https://doi.org/10.1186/s12870-019-1989-2

Zhao, Y., Qiang, C., Wang, X., Chen, Y., Deng, J., Jiang, C., Sun, X., Chen, H., Li, J., Piao, W., Zhu, X., Zhang, Z., Zhang, H., Li, Z., & Li, J. (2019). New alleles for chlorophyll content and stay-green traits revealed by a genome-wide association study in rice (*Oryza sativa*). *Scientific Reports*, *9*, 2541. https://doi.org/10.1038/s41598-019-39280-5

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Faye, J. M., Akata, E. A., Sine, B., Diatta, C., Cisse, N., Fonceka, D., & Morris, G. P. Quantitative and population genomics suggest a broad role of stay-green loci in the drought adaptation of sorghum. *Plant Genome*, 2022;*15:*e20176. https://doi.org/10.1002/tpg2.20176