

1 **Natural variation further increases resilience of sorghum bred for chronically**  
2 **drought-prone environments**

3 Hongxu Dong<sup>1,\*</sup>, Techale Birhan<sup>2</sup>, Nezif Abajebel<sup>2</sup>, Misganu Wakjira<sup>2</sup>, Tesfaye Mitiku<sup>2</sup>,  
4 Cornelia Lemke<sup>1</sup>, Vincent Vadez<sup>3</sup>, Andrew H. Paterson<sup>1</sup>, Kassahun Bantte<sup>2,□</sup>

5 <sup>1</sup>Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602, USA

6 <sup>2</sup>Department of Horticulture and Plant Science, Jimma University, Ethiopia

7 <sup>3</sup>ICRISAT, Patancheru 502324, Andhra Pradesh, India

8 \*Current address: Department of Plant and Soil Sciences, Mississippi State University, Starkville,  
9 MS 39762, USA

10 □Corresponding author: Kassahun Bantte, Department of Horticulture and Plant Science, Jimma  
11 University, Ethiopia: Email: [kassahunb@gmail.com](mailto:kassahunb@gmail.com)

12 **Running title:** drought resilience of sorghum

13

14 **Abstract**

15 Climate-change-associated shifts in rainfall distribution together with a looming worldwide  
16 water crisis make drought resilience of central importance to food security. Even for relatively  
17 drought resilient crops such as sorghum, moisture stress is nonetheless one of the major  
18 constraints for production. Here, we explore the potential to use natural genetic variation to build  
19 on the inherent drought tolerance of an elite cultivar (Teshale) bred for Ethiopian conditions  
20 including chronic drought, evaluating a backcross nested-association mapping (BC-NAM)  
21 population using 12 diverse founder lines crossed with Teshale under three drought-prone  
22 environments in Ethiopia. All twelve populations averaged higher head exertion and lower leaf

23 senescence than the recurrent parent in the two highest-stress environments, reflecting new  
24 drought resilience mechanisms from the donors. A total of 154 QTLs were detected for eight  
25 drought responsive traits – the validity of these were supported in that 100 (64.9%) overlapped  
26 with QTLs previously detected for the same traits, concentrated in regions previously associated  
27 with ‘stay-green’ traits as well as the flowering regulator *Ma6* and drought resistant gene *P5CS2*.  
28 Allele effects show that some favorable alleles are already present in the Ethiopian cultivar,  
29 however the exotic donors offer rich scope for increasing drought resilience. Using model-  
30 selected SNPs associated with eight traits in this study and three in a companion study,  
31 phenotypic prediction accuracies for grain yield were equivalent to genome-wide SNPs and were  
32 significantly better than random SNPs, indicating that these studied traits are predictive of  
33 sorghum grain yield. Rich scope for improving drought resilience even in cultivars bred for  
34 drought-prone regions, together with phenotypic prediction accuracy for grain yield, provides a  
35 foundation to enhance food security in drought-prone areas like the African Sahel.

36

37 **Keywords:** sorghum, food security, African Sahel, drought resilience, genetic mapping, genomic  
38 prediction

39

## 40 **Introduction**

41 In drier parts of the world, more than a half-century of climate change coupled with ongoing  
42 population growth has already pushed agriculture increasingly into ecologically marginal regions  
43 and is expected to severely impact agricultural production especially in sub-Saharan Africa  
44 (Kurukulasuriya et al., 2006; Pironon et al., 2019). In the African Sahel, years of above-average

45 rainfall followed by drought starting in the 1960's (Haywood, Jones, Bellouin, & Stephenson,  
46 2013), were closely associated with more than doubling of the area devoted to crops (Kandji,  
47 Verchot, & Mackensen, 2006). Such unsustainable land management practices are well known to  
48 deplete topsoil (Desta, Tamene, Abera, Amede, & Whitbread, 2021) and contribute to the Sahel  
49 ranking among the regions with most severe soil degradation worldwide (ISRIC, 1990). Indeed,  
50 depleting soil organic matter not only increases atmospheric carbon but also sacrifices the  
51 potential of soil as a carbon reservoir (Amelung et al., 2020).

52 Its unusual tolerance of low inputs, especially water, make the cereal crop sorghum  
53 essential in rain-fed low-input cropping regions that were inadequately addressed in the 'Green  
54 Revolution' (Cantrell & Hettel, 2004; Conway, 1997; Paterson & Li, 2011), being a failsafe crop  
55 that remains productive in conditions that are too dry for other cereals. Despite its inherent  
56 drought tolerance, moisture stress is nonetheless one of the major constraints for sorghum  
57 production (Vadez et al., 2011).

58 Phenotypic selection for drought tolerance is challenging as it is conditioned by complex  
59 inheritance and involve a plethora of component traits. A recent study catalogued more than 600  
60 QTL underlying drought and heat stress from ~150 publications in sorghum from 1995 to present  
61 (Mace et al., 2019). Transcriptomic analysis comparing stay-green and senescent sorghum  
62 genotypes identified more than 2,000 differentially expressed genes, of which 289 were within  
63 known stay-green QTL (Johnson, Cummins, Lim, Slabas, & Knight, 2015). Further, plant height  
64 and flowering time are extensively studied adaptive traits in sorghum (Bouchet et al., 2017;  
65 Brown et al., 2006; Kong et al., 2018; Lin, Schertz, & Paterson, 1995; Morris et al., 2013; Zhang  
66 et al., 2015), and along with variation in stay-green and other drought-related traits, are also  
67 predictive of grain yield performance in drought prone environments (Jordan et al., 2003).

68           Here, we explore the potential to build on the inherent drought tolerance of an elite cultivar bred  
69 for Ethiopian conditions, by crossing 12 diverse founder lines with different drought adaptations  
70 (Vadez et al., 2011) to an Ethiopian elite cultivar, and phenotyping the resulting 12 populations  
71 under natural drought environments in Ethiopia, taking advantage of the ‘nested association  
72 mapping’ (NAM) strategy (Yu *et al.* 2008) to combine high mapping resolution and high  
73 statistical power in studying variation from diverse parental inbreds. Head exertion provides a  
74 valuable indicator of pre-flowering drought stress (Rosenow & Clark, 1981; Seetharama, Reddy,  
75 Peacock, & Bidinger, 1982) -- driven by elongation of the uppermost internode of the stem  
76 (peduncle) shortly before anthesis, increased peduncle exertion was also associated with  
77 drought resistance in rice (*Oryza sativa* L.) (O’Toole & Namuco, 1983); and was positively  
78 correlated with yield, length of head, and number of productive tillers in pearl millet (*Pennisetum*  
79 *americanum* L.) (Ibrahim, Marcarian, & Dobrenz, 1986). The ‘stay-green’ trait, reflecting post-  
80 flowering drought tolerance by the ability to maintain green leaf area, resist premature plant  
81 senescence, resist lodging, and fill grain normally (Rosenow, Quisenberry, Wendt, & Clark,  
82 1983; Thomas & Ougham, 2014), has been a focus in sorghum ( Borrell et al., 2000, 2001, 2014;  
83 Crasta et al., 1999; Harris et al., 2007; Kassahun et al., 2010; Sanchez et al., 2002; Subudhi et al.,  
84 2000; Xu et al., 2000; Vadez et al., 2011), wheat (*Triticum aestivum*) (Christopher, Manschadi,  
85 Hammer, & Borrell, 2008; Lopes & Reynolds, 2012), maize (*Zea mays*) (Campos, Cooper,  
86 Habben, Edmeades, & Schussler, 2004; Zheng et al., 2009), and rice (*Oryza sativa*) (Cha et al.,  
87 2002; Jiang et al., 2007).

88

## 89 **Materials and Methods**

### 90 ***Plant materials and methods***

91 Twelve donor sorghum lines from eight countries including Nigeria (IS10876), South Africa  
92 (IS14298), Sudan (IS14446, IS3583, IS9911), Ethiopia (IS14556), Cameroon (IS15428,  
93 IS16044, IS16173), India (IS2205), Botswana (IS22325), and Yemen (IS23988) (Table 1) were  
94 obtained from ICRISAT, crossed to the Ethiopian elite cultivar Teshale and the F<sub>1</sub> plants  
95 backcrossed to manually emasculated Teshale to produce BC<sub>1</sub>F<sub>1</sub> families which were selfed by  
96 single seed descent to produce BC<sub>1</sub>F<sub>4</sub> populations of 32-153 lines. For simplicity, we use the  
97 name of the alternate parent to refer to an individual BC<sub>1</sub>F<sub>4</sub> population (*e.g.*, the IS10876  
98 population).

### 99 ***Experimental design and field managements***

100 Parental lines and the BC-NAM population were evaluated at three locations that represent major  
101 sorghum growing regions in Ethiopia (Table 1, Figure S1): Kobo (12°09'N, 39°38'E) and Meiso  
102 (09°14'N, 40°45'E) in the 2017 cropping season, and Sheraro (14°23'N, 37°46'E) in the 2018  
103 cropping season. These three locations are typical of the range of moisture stress conditions in  
104 the target region and had no irrigation, subjecting the study populations to natural moisture stress  
105 conditions. An alpha lattice design with two replications was used at each location. Each field  
106 trial was conducted from July to December. Populations IS2205, IS14556, IS16044, IS32234  
107 were not included in Sheraro due to space constraints. Seeds of each sorghum line were sown  
108 into a 4 m row and thinned three weeks after sowing to spacings of 75 cm and 20 cm between  
109 and within rows, respectively. Therefore, an individual plot without plant loss would consist of  
110 20 plants. During sowing and three weeks after sowing, inorganic fertilizers DAP and Urea were  
111 added at the rate of 100 and 50 kg ha<sup>-1</sup>, respectively. Field weeding and other cultural practices  
112 were carried out as needed.

### 113 ***Phenotypic data analysis***

114 A total of eight traits were evaluated. Leaf senescence (LS) was scored on a plot basis at  
115 physiological maturity on a 1-5 scale based on degree of leaf death, with 1 = least senescent  
116 (10%), 2 = slightly senescent (25%), 3 = intermediate senescent (50%), 4 = mostly senescent  
117 (75%), and 5 = completely senescent (100%). Number of leaves (NL) per plant was averaged  
118 from five plants per plot. As the lower leaves of sorghum tend to die and fall off as plants  
119 mature, all leaves on the main stem were tagged when plants were at five or six leaf stage, so  
120 they could be reflected in counts even if they had been shed. Head exertion (HX) beyond the  
121 ligule of the flag leaf was recorded in centimeters. Number of basal tillers (NT) per plant was  
122 counted. Thousand seed weight (TSW) was determined from 1,000 randomly selected seeds in  
123 grams. Single plant-based grain yield (SGY) was recorded as the total grain weight in grams of  
124 threshed panicles of one plant, averaged from five plants per plot. However, due to plant loss in  
125 many plots, particularly at Kobo and Meiso fields, individual plant grain yield estimates were  
126 confounded by plant density among plots (i.e., plants from low density plots would have  
127 received more resources such as light and moisture due to less competition than plants from high  
128 density plots). Therefore, counts of plants per plot (CPP) were recorded and used as covariate to  
129 adjust for SGY as follows:

$$SGY_{ij} = \mu + \beta(CPP_{ij} - \overline{CPP}) + G_i + R_j + \varepsilon_{ij}$$

130 where  $SGY_{ij}$  represents the original single plant-based grain yield,  $\mu$  is the grand mean,  $\beta$  is the  
131 regression coefficient of the covariate on the genotype,  $CPP_{ij}$  is the value of the covariate for a  
132 particular plot, and  $\overline{CPP}$  is the overall mean of the covariate,  $G_i$  represents genotype,  $R_j$   
133 represents replication, and  $\varepsilon_{ij}$  is random error. Plot-based grain yield (PGY) was calculated by  
134 multiplying adjusted single plant-based grain yield (adjusted SGY) by the number of plants per  
135 plot (CPP). Thus, plot-based grain yield would reflect the actual grain yield potential of sorghum

136 genotypes under the targeted natural environments by considering planting density. In Sheraro,  
137 LS, HX, and NT did not express symptoms or exhibited limited phenotypic variations and thus  
138 were not phenotyped (K. Bantte, personal communications).

139 Raw data from each location were checked and outlier phenotypes were discarded. HX at  
140 Meiso exhibited left-skewed data distribution (Figure S2) and values used in genetic analyses  
141 were transformed to approximate normal distribution using Box-Cox method (Box & Cox,  
142 1964), calculated as  $(y^\lambda - 1)/\lambda$ , where  $y$  is the original phenotypic value, and  $\lambda$  is the optimal  
143 value which results in the best approximation of normalization.  $\lambda$  was determined as 0.4 in this  
144 case. Many genotypes had no exertion (HX=0), and a constant of one was added because Box-  
145 Cox transformation only works if all data are positive. To interpret results such as allelic effects,  
146 transformed data were converted back to the original values.

147 In order to evaluate the effects of environment, family, genotype nested in family, family  
148 by environment interaction, and genotype nested in family by environment interaction, a multi-  
149 environment analysis of variance (ANOVA) was conducted for each trait across all three  
150 environments. Due to the strong environment and genotype-by-environment effects (Table S1),  
151 trait best linear unbiased predictions (BLUPs) were estimated for each line within each  
152 environment using a mixed linear model implemented in the lme4 package (Bates, Mächler,  
153 Bolker, & Walker, 2015). All model terms were treated as random effects in the following  
154 model:

$$Y_{ijk} = \mu + F_i + G(F)_{ij} + R_k + \varepsilon_{ijk}$$

155 where  $Y$  represents raw phenotypic data,  $F$  is the individual family in BC-NAM population,  
156  $G(F)$  is genotype nested within family,  $R$  is replication, and  $\varepsilon$  is random error. Pearson

157 correlation coefficients between traits were calculated using line BLUPs. Broad-sense  
158 heritability in each environment was determined as the proportion of total phenotypic variance  
159 explained by the combined family and line terms using equation 1 from Wallace *et al.* (2016):

$$H^2 = \frac{\sigma_F^2 + \sigma_{G(F)}^2}{\sigma_F^2 + \sigma_{G(F)}^2 + \sigma_\varepsilon^2/2}$$

160 where  $\sigma_F^2$  the variance explained by family,  $\sigma_{G(F)}^2$  is the variance explained by individual lines,  
161 and  $\sigma_\varepsilon^2$  is the random error variance.

### 162 ***Marker-trait association***

163 A total of 4395 SNPs generated using a *PstI-MspI* protocol were used for marker-trait  
164 association analyses. Details of marker development and filtering were described as follows:  
165 GBS libraries were sequenced on a NextSeq500 (Illumina, San Diego, CA) with 150-bp single-  
166 end reads at the Genomics and Bioinformatics Core of the University of Georgia; SNP calling  
167 was conducted using TASSEL GBS v2 pipeline (Bradbury et al., 2007) and *Sorghum bicolor*  
168 genome v2.1 (Paterson et al., 2009); then raw SNPs were filtered based on tag coverage > 5%,  
169 minor allele frequency > 0.03, single marker missing data < 5%, yielding a set of 4395 SNPs for  
170 analysis. QTL mapping considered both joint-linkage (JL) (Buckler et al. 2009) and GWAS  
171 models. JL analysis only included eight large populations (N > 80, Table 1) by removing  
172 IS14556, IS16044, IS2205, and IS23988, which contain many monomorphic SNPs and pose a  
173 problem in accurately estimate marker effects due to their small size. JL mapping used the  
174 forward-backward stepwise regression model implemented in the Stepwise Plugin of TASSEL 5,  
175 with family as a cofactor and marker nested within family, suggested to be the most powerful  
176 procedure for JL mapping (Würschum et al., 2012). The family main effect was fit first, then



177 marker effects were selected to enter or leave the model based on the E-03  $P$ -value threshold for  
178 the marginal  $F$ -test of that term. GWAS was performed using the `lm` function in R (R  
179 Development Core Team 2016) by fitting SNPs as non-nested effects with family as a cofactor.  
180 All 12 populations were included in GWAS because SNPs were non-nested and treated as  
181 biallelic, thus small size of individual populations was less a problem. Loci of the GWAS model  
182 that had  $P$ -value below the E-03 threshold were considered QTL.

183 All JL QTLs identified in this study were compared to the Sorghum QTL Atlas database  
184 [Mace *et al.* (2019)], which collated the projected locations of *c.* 6000 QTL or GWAS loci from  
185 150 publications in sorghum from 1995 to present. Peak SNP associations from JL and GWAS  
186 were also compared to *c.* 2000 genes differentially expressed between stay-green (B35) and  
187 senescent (R16) sorghum genotypes (Johnson *et al.*, 2015). Sorghum genes containing or directly  
188 adjacent to SNP associations were also searched using BEDOPS (Neph *et al.*, 2012) and standard  
189 UNIX tools.

### 190 ***Genome-wide prediction***

191 Drought-related traits are predictive of grain yield in drought prone sorghum production  
192 environments (Jordan *et al.*, 2003). We performed genomic predictions for plot-based grain yield  
193 (PGY) and 1000-seed weight (TSW) using marker-trait associations for the eight drought  
194 responsive traits in this study and three adaptive traits mapped in a companion study of this  
195 sorghum BC-NAM population including days to flowering, days to maturity, and plant height  
196 (Dong *et al.*, unpublished data). Specifically, we considered the following three scenarios for  
197 genomic prediction: (i) All 4395 segregating SNPs across the genome; (ii) the peak SNPs  
198 identified from JL and GWAS of all eleven traits from three environments; (iii) an equal number  
199 of random SNPs that have the similar allele frequencies as those detected SNPs from JL and

200 GWAS. All models were run with 50 five-fold cross validations, in which 20% of phenotypes  
201 were randomly masked. The resulting predictions were compared to the masked (true) values to  
202 generate a prediction accuracy estimate, measured using a Pearson correlation coefficient. *P*-  
203 values for the differences in distribution of correlations were determined via a two-sided *t*-test.

204

## 205 **Results**

### 206 *Phenotypic variation within and between families*

207 Different natural drought conditions provided opportunities to study this BC-NAM population  
208 across three environments. Multi-environment ANOVAs revealed strong and significant effects  
209 of environment, family, genotype nested in family, family by environment interaction, and  
210 genotype by environment interaction for all traits, except that genotype by environment effect  
211 was not significant for head exertion (Table S1). During the 2017 cropping season, Kobo and  
212 Meiso experienced severe moisture stress during sowing season (July). Daily data from nearby  
213 weather stations showed cumulative precipitation before sowing (Jan-June) of 194.4 mm and  
214 313.6 mm in Kobo and Meiso, respectively, and 401.2 mm in Sheraro (Figure S1). At Kobo, a  
215 total of 347 lines (34%) across all 12 families had plant loss, ranging from 22% in IS9911 to  
216 56% in IS2205 (Table 1). At Meiso, 189 lines (16%) lost one of the two replicates, and plant loss  
217 percentage ranged from 5% in IS16044 to 24% in IS23988. Between plant loss at Kobo and  
218 Meiso there were 58 lines in common, however neither the Pearson's correlation coefficient ( $r =$   
219  $0.021$ ,  $p = 0.9484$ ) nor the rank-based Kendall's correlation statistic ( $\tau = 0.030$ ,  $p = 0.9466$ ) were  
220 significant, indicating that sorghum response to drought is dynamic and labile under different  
221 stressed environments. However, during the 2018 cropping season in Sheraro, seeds germinated  
222 well and virtually all plants survived (Table 1).

223 Phenotypic distribution and variance differed greatly among the three environments  
224 (Figure 1). Five traits including leaf number, single plant-based grain yield, count of plants per  
225 plot, plot-based grain yield, and 1000-seed weight were evaluated over three locations; three  
226 traits including head exertion, leaf senescence, and tiller number were evaluated at Kobo and  
227 Meiso only (Figure 1). Plants grew fewer leaves in Kobo (avg. leaf number: 8.76) and Meiso  
228 (avg. leaf number: 9.95) than Sheraro (avg. leaf number: 13.60) (Table S2, Figure 1). Head  
229 exertion and leaf senescence are two interesting traits to evaluate drought tolerance in sorghum,  
230 the former trait showed larger phenotypic variation in Meiso, whereas the latter trait exhibited  
231 more phenotypic variations in Kobo (Figure 1), presumably reflecting somewhat different  
232 drought regimes. Phenotypic variation of tiller number was quite limited in this BC-NAM  
233 population, especially in Meiso (Figure 1).

234 Plot-based grain yield was highest in Sheraro, followed by Meiso, and lowest in Kobo  
235 (Figure 1), reflecting relative levels of plant loss and plant density at maturity. In contrast, single  
236 plant-based grain yield and 1000-seed weight ranked highest in Kobo, followed by Meiso, and  
237 lowest in Sheraro (Figure 1, Table S2). The nonuniform plant density across three environments  
238 might also lead to the differences of length of grain filling period (anthesis to maturity) at three  
239 locations -- subtracting days to flowering from days to maturity, plants in Kobo (48.89 d.) and  
240 Meiso (37.61 d) had significantly longer grain filling periods than in Sheraro (25.23 d; Table S2).

241 While many phenotypes of individual families were similar to that of the recurrent parent,  
242 Teshale (Figure S3, File S1), consistent with the backcross breeding scheme, all families showed  
243 higher head exertion and lower leaf senescence indicating improved pre- and post-flowering  
244 drought tolerance in the most stressed tests. In Kobo, population means for head exertion ranged  
245 from 0.57 cm in IS2205 to 2.3 cm in IS22325 (Figure S3, File S1), whereas that of recurrent

246 parent Teshale was only 0.03 cm; in Meiso, population means for head exertion ranged from  
247 4.95 cm for IS9911 to 7.60 cm in IS15428, whereas Teshale was 3.92 cm (Figure S3, File S1).

248 All 12 families also showed lower leaf senescence mean values than the recurrent parent  
249 at Kobo and Meiso (Figure S3), based on an ordinal scale 1-5 to evaluate leaf senescence with 1  
250 being least senescent and 5 being complete senescent. In Kobo, population means for leaf  
251 senescence ranged from 2.30 in IS14298 to 2.64 in IS14556, whereas that of Teshale was 2.67;  
252 in Meiso, population means for leaf senescence ranged from 2.79 cm for IS2205 and IS9911 to  
253 2.84 in IS14556 and IS16044, whereas that of Teshale was 2.86 (File S1). Therefore, phenotype  
254 data of head exertion and leaf senescence indicated that drought tolerance alleles from diverse  
255 founder lines could have been introgressed into the recurrent parent.

256 Broad-sense heritabilities of the eight traits across the BC-NAM population within each  
257 environment varied from 0.08 for leaf senescence in Meiso to 0.63 for 1000-seed weight in  
258 Sheraro (Table 2). Head exertion showed moderate heritabilities at Kobo (0.58) and Meiso  
259 (0.59). Heritability of leaf senescence was moderate in Kobo (0.42) but very low in Meiso (0.28),  
260 which was primarily due to the limited variation in Meiso (Figure 1). Number of tillers was  
261 similarly low heritable at both Kobo (0.27) and Meiso (0.26). Number of leaves showed  
262 moderate heritabilities across three environments (0.44-0.60). Plot-based grain yield was less  
263 heritable than 1000-seed weight, and heritabilities of these two traits ranged from 0.26-0.39 and  
264 0.31-0.63, respectively. Count of plants per plot, which was used as the covariate to adjust for  
265 planting density in the estimation of plot-based grain yield, exhibited low heritabilities across  
266 three environments (0.16-0.38).

267 *Genetic architecture of adaptive traits*

268 Genetic architecture of complex traits is defined by the number, effect size, frequency, and gene  
269 action of the QTLs that affect it. We conducted association analyses for eight adaptation traits. A  
270 total of 154 JL QTLs were detected using joint-linkage model across the three environments  
271 (Table 2, Figure 3A, File S2), of which 100 (64.9%) overlapped with previous QTLs detected for  
272 the same trait from 72 different studies (File S2). Large number of common QTLs of same traits  
273 between this study and many previous studies demonstrated the power of this sorghum BC-NAM  
274 population in dissecting drought tolerance traits by sampling more allelic diversity than possibly  
275 within any single biparental population. In order to leverage the investment, we incorporated the  
276 four small populations (IS2205, IS14556, IS16044, and IS32234) into GWAS analyses. A total  
277 of 233 GWAS associations were detected, with 103, 90, and 40 in Kobo, Meiso, and Sheraro,  
278 respectively (Figure 3A, File S3). Despite the different statistical frameworks, high  
279 correspondence between JL QTLs and GWAS hits was observed (Figure 3A). Exact overlap of  
280 linkage mapping and GWAS is not expected, as linkage mapping tests markers within an  
281 individual population whereas GWAS tests marker effects across populations, leading to  
282 different strengths and weakness in each approach (Tian et al., 2011). Joint linkage analysis  
283 produces many more small effects than GWAS analysis as an artifact of the model fitting  
284 process, which assigns a separate effect to all populations at each QTL. Moreover, the addition  
285 of those four small populations also played a role in the discrepancy of associations between  
286 these two methods.

287 Genes within stay-green QTLs were enriched for proximity to drought-related trait loci  
288 identified in this study (Figure 3). The 154 JL QTLs and 233 GWAS hits were distributed across  
289 all ten chromosomes in a non-random fashion (Figure 3A). We compared these associations with  
290 the transcriptomic analysis comparing stay-green and senescent sorghum lines (Johnson *et al.*

291 2015), which detected 2,036 differentially expressed genes between the stay-green (B35) and  
292 senescent (R16) sorghum genotypes. In particular, the associations detected in this study  
293 clustered around 289 genes within known stay-green QTLs catalogued in the Comparative  
294 Saccharinae Genome Resource (File S4) (Zhang et al., 2013), which further corroborated the  
295 reliability of the associations detected in this study. Due to the moisture stress during sowing  
296 season (Figure S1), 58 sorghum lines lost one of their two replicates at both Kobo and Meiso  
297 (Table 1), indicating that these lines may be more sensitive to moisture stress than the remaining  
298 1113 lines. Intriguingly, fixation index ( $F_{ST}$ ) analysis revealed elevated peaks at stay-green QTL  
299 regions between these 58 drought-sensitive lines and the other 1113 lines (Figure 3B), on  
300 chromosome 1 ( $F_{ST} = 0.093$ , S01\_11158453), chromosome 4 ( $F_{ST} = 0.077$ , S04\_9538999),  
301 chromosome 5 ( $F_{ST} = 0.094$ , S05\_363813), and chromosome 7 ( $F_{ST} = 0.077$ , S07\_56408236).  
302 Many associations from joint-linkage and GWAS analyses were also detected near these  $F_{ST}$   
303 peaks (Figure 3A).

304 Allelic effect of QTLs is an important facet of genetic architecture of traits. For the 154  
305 JL QTLs, marker effects within each individual population were estimated (i.e. 154 QTLs  $\times$  8  
306 populations = 1,232 QTL effects). Since most QTLs were not present in all families, many of  
307 these effects were near zero and generally normally distributed (Figure S4, File S2). Fisher  
308 (1930) provided a simple theoretical justification for these observations. For a well-adapted  
309 organism close to its fitness optimum, only small effects can increase fitness. This explanation  
310 conforms to the breeding design of this BC-NAM population, that all founder lines and the  
311 common recurrent parent are generally adapted to semiarid environments. In addition, Orr (1998)  
312 showed that regardless of the distance from the fitness optimum, the expected distribution of  
313 effect sizes progressively fixed during adaptation is exponential, with a small number of large-

314 effect loci fixed first, followed by progressively larger numbers of loci with small effects  
315 becoming fixed. Moreover, the genetic architecture of intraspecific variation like BC-NAM  
316 population in this study consists of many loci with small effects because loci with larger effects  
317 tend to be only briefly polymorphic.

318         Among the seven head exertion JL QTLs at Kobo, allelic effects ranged from -2.9 cm at  
319 S01\_14044648 in IS22325 to 2.8 cm at S01\_12672806 in IS22325 (Figure S4, File S2); 34  
320 showed positive effects (i.e., increase head exertion ability) versus 22 negative allelic effects.

321         Among the 12 JL QTLs for head exertion in Meiso, allelic effects ranged from -1.3 cm at  
322 S05\_55829561 in IS14446 to 1.1 cm at S01\_71318983 in IS22325 (Figure S4, File S2). Leaf  
323 senescence allelic effects were of small magnitude, ranging from -0.6 (i.e., negative number  
324 means less senescent) at S01\_49654172 in IS3583 to 0.5 (i.e., positive number means more

325 senescent) at S08\_5990097 in IS10876 at Kobo; allelic effects of leaf senescence ranged from -  
326 0.11 at S09\_54402475 in IS9911 to 0.08 at S10\_439084 in IS16173 in Meiso (File S2). Smaller  
327 allelic effects of leaf senescence in Meiso compared to those in Kobo corresponds to the

328 narrower trait phenotypic variation in Meiso (Figure 1). For number of leaves, allelic effects  
329 ranged from -1.3 (i.e., negative number means less leaves than recurrent parent) at

330 S08\_51534279 in IS22325 to 0.8 at S10\_10472812 in IS22325 at Kobo; and allelic effects  
331 ranged from -0.7 at S07\_2397109 in IS9911 to 0.8 at S06\_769807 in IS14298. For number of

332 tillers, allelic effects ranged from -0.3 (i.e., negative number means reduced tillering compared to  
333 recurrent parent alleles) at S04\_62057354 in IS13428 to 0.3 at S05\_11960542 in IS10876 at

334 Kobo; whereas in Meiso, allelic effects for number of tillers were very small, ranging from -0.1  
335 to 0.2, which again corresponds to the very limited trait phenotypic variation in Meiso (Figure 1).

336 Such approximately normally distributed allelic effects were observed for all traits studied

337 (Figure S4). Indeed, there were QTLs exhibiting large allelic effects within specific families  
338 (Figure S4). For example, at one plot-based grain yield QTL detected in Kobo on chromosome 4  
339 (S04\_50497689), the IS9911 allele contributed a positive effect of 97.44 g relative to the  
340 recurrent parent (File S2), which is considerably very large given that plot-based grain yield  
341 across the whole BC-NAM population ranged from 205.81 g to 512.03 g in Kobo (File S1).  
342 Large effect alleles in this study were of great breeding value in improving Ethiopian adapted  
343 cultivars.

344 Total phenotypic variance explained by the final JL model was generally low for these  
345 eight adaptive traits, ranging from 0.10 for CPP to 0.35 for PGY in Kobo, from 0.14 for CPP to  
346 0.37 for HE in Meiso, and from 0.06 for CPP to 0.31 for TSW in Sheraro (Table 2). Note that  
347 heritability imposes an approximate upper limit to the  $R^2$  of a QTL model (Yu et al., 2008).  
348 Therefore, this is not unexpected given the low to moderate heritabilities of these traits under  
349 drought conditions. By taking the broad-sense heritability into account, proportion of genetic  
350 variance explained by the final joint-linkage model of each trait ranged from 26% for CPP to  
351 90% for PGY in Kobo, 37% for NL to 88% for LS and PGY in Meiso, and 32% for NL to 64%  
352 for PGY in Sheraro (Table 2).

353 Some of the strongest and (or) most consistent associations of various traits locate near  
354 known genes. Green leaf is the hub of photosynthesis and leaf number has been of particular  
355 interest in drought tolerance studies in sorghum. JL model detected 13, 6, and 8 QTLs for leaf  
356 number in Kobo, Meiso, and Sheraro, respectively (Table 2, Figures 3, 4). Significant  
357 associations for leaf number were located near the *Ma6* region at Kobo and Meiso, with QTL  
358 peaks (S06\_782670 at Kobo and S06\_769807 at Meiso) 109 kb and 96 kb away from *Ma6*,  
359 respectively (Figure 4A and 4B). However, this association was not detected in Sheraro (File



360 S2). In addition to flowering regulators, associations were also detected near known stress  
361 regulation genes. One *DREB1A* gene (Sobic.007G181500), which encodes a dehydration-  
362 responsive element-binding transcription factor, was located at the  $F_{ST}$  peak region on  
363 chromosome 7 ( $F_{ST} = 0.077$ , S07\_56408236), and GWAS association for leaf number at Meiso  
364 (S07\_59324541; Figure 4B, File S3), and JL-QTL for 1000-seed weight at Kobo  
365 (S07\_60171225; File S2) were detected near this gene. Several traits including number of leaves  
366 (JL-QTL S03\_67359281 at Kobo, Figure 4A, File S2), head exertion at Meiso (S03\_59654236,  
367 Figure 4C, File S2), and number of tillers at Meiso (S03\_67407460 at Meiso, File S2) also  
368 showed strong associations near the *P5CS2* gene (delta1-pyrroline-5-carboxylate synthase 2;  
369 Sobic.003G356000, File S2). Among these, the association for number of tillers in Meiso  
370 (S03\_67407460) was only 33 kb from *P5CS2*, encoding an enzyme responsible for proline  
371 biosynthesis (Kishor, Hong Zonglie, Miao Guo-Hua, Hu Chein-An, & Verma, 1995) that is  
372 highly expressed in the stay-green sorghum line compared with the senescent line (Johnson et al.,  
373 2015). Several traits including head exertion, number of leaves, number of tillers, and 1000-  
374 seed weight had associations on chromosome 5 (Figure 4, File S2, File S3) near the *LEA* gene  
375 (Sobic.005G021800) encoding hydrophilic proteins with major roles in drought and other abiotic  
376 stresses tolerance in plants (Magwanga et al., 2018). The *LEA* gene on chromosome 5 was also  
377 located near the  $F_{ST}$  peak ( $F_{ST} = 0.094$ , S05\_363813, Figure 3) distinguishing the 58 sorghum  
378 lines lost one of their two replicates at both Kobo and Meiso from the remaining 1113 lines.  
379 Moreover, significant association for head exertion in Meiso (S01\_69085453) was 217 kb from  
380 the *KCS20* gene (3-ketoacyl-CoA synthase 20; Sobic.001G495500, Figure 4C), found to be up-  
381 regulated by drought and osmotic stress (Joubès et al., 2008). In *Arabidopsis*, expression of  
382 *KCS20* was up-regulated twofold by drought (Lee et al., 2009).

383

## 384 **Discussion**

385 Drought-related traits are very complex and polygenic as evidenced by our association analyses.  
386 In genetic dissection of yield and drought-related traits under drought conditions, a major and  
387 universally recognized obstacle is properly phenotyping in a high-throughput fashion (Tuberosa,  
388 2012). Phenomics-enabled breeding using high-throughput phenotyping (HTP) may enable  
389 researchers to better understand the complexities of trait development and to better optimize  
390 genotypes through selection in breeding programs (Jiang et al. 2018; Pugh et al. 2018; Xu et al.  
391 2018).

392 Most of the associations detected for these eight traits located within known stay-green  
393 QTL regions (Figure 3). In sorghum, four major stay-green quantitative trait loci (QTL) have  
394 been consistently identified (Crasta et al., 1999; Subudhi et al., 2000; Xu et al., 2000; Sanchez et  
395 al., 2002; Harris et al., 2007), with a number of physiological studies testing their effects across  
396 different genetic backgrounds (Borrell et al., 2000, 2001, 2014; Kassahun et al., 2010; Vadez et  
397 al., 2011). Nearby candidates for these associations could be of great interest for future  
398 functional studies. Limited knowledge of drought resilience restricts the power of candidate  
399 approaches, but it is nonetheless notable that several previously known drought tolerance genes  
400 were associated with drought resilience in this study.

401 While it is clear that the drought resilience of an elite Ethiopian cultivar can be improved  
402 by incorporation of different adaptations from diverse germplasm, none of the 12 individual  
403 populations produced higher population mean yield than the recurrent parent Teshale (Figure S3,  
404 File S1). The estimated plot-based grain yield of Teshale were 401.05 g and 539.42 g at Kobo  
405 and Meiso, respectively; while population mean yield ranged from 309.99 g in IS22325 to

406 377.63 g in IS16044 at Kobo, and from 506.41 g in IS22325 to 536.99 g in IS16173 at Meiso  
407 (File S1). Since the parents were not tested at Sheraro, similar plot yield comparison could not be  
408 conducted. Transgressive segregants found in each population (Figure S3, File S1) suggest that  
409 the adaptive phenotype of Teshale can be recapitulated or perhaps even further improved,  
410 strengthening drought resilience of cultivars that stand between vulnerable geographies and  
411 famine. For all traits, rich variation was reflected by mixtures of ‘positive’ and ‘negative’ QTL  
412 alleles (Figure S4, File S2), suggesting the 12 diverse founder lines with different drought  
413 adaptations (Vadez et al., 2011) complementing and supplementing the inherent drought  
414 resilience of Teshale. Correlations of plot-based grain yield across all three environments with  
415 days to flowering (-0.20 to -0.42); and plant height (0.14-0.39) exemplify scope for the sorts of  
416 adjustments that may be needed to re-establish locally adaptive phenotypes. Indeed, with the  
417 enormous altitudinal variation of a country such as Ethiopia, somewhat different lines may be  
418 needed for different locales.

419 Genomic prediction for plot-based grain yield and 1000-seed weight based on our  
420 findings support the assertion that drought-related traits are predictive of grain yield performance  
421 in drought prone sorghum production environments (Jordan et al., 2003). We conducted genomic  
422 prediction across the entire BC-NAM population within each environment under three scenarios:  
423 using all 4395 genome-wide markers, using 434 model-selected SNPs from JL and GWAS for  
424 11 traits (eight in this study and three in a companion study including days to flowering, days to  
425 maturity and plant height; File S2, File S3), and using an equivalent number (434) of randomly  
426 selected SNPs that are at least 260 kb based on LD decay away from JL QTL or GWAS hits.  
427 Prediction accuracy was low to modest, with average correlations of 0.16-0.34 for plot-based  
428 grain yield and 0.25-0.41 for 1000-seed weight, but the selected 434 SNPs had similar accuracy

429 as all 4395 genome-wide SNPs and were significantly more accurate than 434 random SNPs  
430 (Figure 5, File S5). Low to moderate genomic prediction accuracies were not unexpected -- the  
431 theoretical upper limit of genomic prediction accuracy cannot exceed narrow-sense heritability  
432 because rrBLUP uses the additive kinship matrix among genotypes to perform genomic  
433 prediction. In this study, most adaptive traits showed low to moderate heritabilities especially  
434 under drought conditions (Table 2).

435         With a worldwide water crisis looming (Rosegrant, Cai, & Cline, 2002; Serageldin, 2004;  
436 UNESCO, 2002), increasing the inherent drought resilience of sorghum is vital in rain-fed  
437 environments, exemplified by but not limited to sub-Saharan Africa. The sorghum BC-NAM  
438 population reported here demonstrates proof of principle, as well as a foundation for breeding  
439 improved cultivars for East Africa and a valuable community resource for understanding  
440 drought-related traits. Clustering of associations within stay-green QTL regions highlighted the  
441 merit of more in-depth exploration of these genomic regions. Given the complex and labile  
442 nature of drought-related traits, breeding crops with improved yield and resilience to  
443 environmental stresses may greatly benefit from the integration of phenomics technologies such  
444 as HTP with genomics.

445

#### 446 **Acknowledgments**

447 This work was funded in whole or part by the United States Agency for International Development  
448 (USAID) Bureau for Resilience and Food Security under Agreement # AID-OAA-A-13-00044 as part of  
449 Feed the Future Innovation Lab for Climate Resilient Sorghum. Any opinions, findings, conclusions, or  
450 recommendations expressed here are those of the authors alone. The authors thank Jimma University,  
451 Ethiopian Institute of Agricultural Research, Amhara Region Agricultural Research Institute and

452 Tigray Region Agricultural Research Institute for providing land and technical support for the  
453 field trials, and the Paterson Lab for valuable help and discussions. We also thank the Georgia  
454 Genomics and Bioinformatics Core for sequencing service.

#### 455 ***Data availability statement***

456 Sequencing data are available in the NCBI Sequence Read Archive under project accession  
457 BioProject ID PRJNA687679. Data analysis scripts have been deposited to GitHub  
458 (<https://github.com/hxdong-genetics/Ethiopian-Sorghum-BC-NAM>). Please contact  
459 corresponding author for other data.

#### 460 **References**

- 461 Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., ...  
462 Chabbi, A. (2020). Towards a global-scale soil climate mitigation strategy. *Nature*  
463 *Communications*. <https://doi.org/10.1038/s41467-020-18887-7>  
464 Bates, D., Mächler, M., Bolker, B., & Walker, S. C. (2015). Fitting linear mixed-effects models  
465 using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
466 Borrell, A., Hammer, G., & Van Oosterom, E. (2001). Stay-green: A consequence of the balance  
467 between supply and demand for nitrogen during grain filling? *Annals of Applied Biology*,  
468 138(1), 91–95.  
469 Borrell, A. K., Hammer, G. L., & Henzell, R. G. (2000). Does maintaining green leaf area in  
470 sorghum improve yield under drought? II. Dry matter production and yield. *Crop Science*,  
471 40, 1037–1048.  
472 Borrell, A. K., van Oosterom, E. J., Mullet, J. E., George-Jaeggli, B., Jordan, D. R., Klein, P. E.,  
473 & Hammer, G. L. (2014). Stay-green alleles individually enhance grain yield in sorghum  
474 under drought by modifying canopy development and water uptake patterns. *New*  
475 *Phytologist*, 203(3), 817–830.  
476 Bouchet, S., Olatoye, M. O., Marla, S. R., Perumal, R., Tesso, T., Yu, J., ... Morris, G. P. (2017).  
477 Increased power to dissect adaptive traits in global sorghum diversity using a nested  
478 association mapping population. *Genetics*, 206(June), 573–585.  
479 Box, G. E. P., & Cox, D. R. (1964). An analysis of transformation. *Journal of the Royal*  
480 *Statistical Society. Series B (Methodological)*, 26(2), 211–252.  
481 Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S.  
482 (2007). TASSEL: Software for association mapping of complex traits in diverse samples.  
483 *Bioinformatics*, 23(19), 2633–2635.  
484 Brown, P. J., Klein, P. E., Bortiri, E., Acharya, C. B., Rooney, W. L., & Kresovich, S. (2006).  
485 Inheritance of inflorescence architecture in sorghum. *Theoretical and Applied Genetics*,  
486 113(5), 931–942.  
487 Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., ...

- 488 McMullen, M. D. (2009). The genetic architecture of maize flowering time. *Science*,  
489 325(5941), 714–718.
- 490 Campos, H., Cooper, M., Habben, J. E., Edmeades, G. O., & Schussler, J. R. (2004). Improving  
491 drought tolerance in maize: A view from industry. *Field Crops Research*, 90(1), 19–34.
- 492 Cantrell, R. P., & Hettel, G. P. (2004). Rice, biofortification, and enhanced nutrition. In *World*  
493 *Food Prize Symposium*. Des Moines, Iowa.
- 494 Cha, K. W., Lee, Y. J., Koh, H. J., Lee, B. M., Nam, Y. W., & Paek, N. C. (2002). Isolation,  
495 characterization, and mapping of the stay green mutant in rice. *Theoretical and Applied*  
496 *Genetics*, 104, 526–532.
- 497 Christopher, J. T., Manschadi, A. M., Hammer, G. L., & Borrell, A. K. (2008). Developmental  
498 and physiological traits associated with high yield and stay-green phenotype in wheat.  
499 *Australian Journal of Agricultural Research*, 59(4), 354–364.
- 500 Conway, G. (1997). *The doubly green revolution: food for all in the twenty-first century*. Ithaca,  
501 NY: Cornell University Press.
- 502 Crasta, O. R., Xu, W. W., Rosenow, D. T., Mullet, J., & Nguyen, H. T. (1999). Mapping of post  
503 flowering drought resistance traits in grain sorghum: Association between QTLs influencing  
504 premature senescence and maturity. *Molecular and General Genetics*, 262(3), 579–588.
- 505 Desta, G., Tamene, L., Abera, W., Amede, T., & Whitbread, A. (2021). Effects of land  
506 management practices and land cover types on soil loss and crop productivity in Ethiopia: A  
507 review. *International Soil and Water Conservation Research*.  
508 <https://doi.org/10.1016/j.iswcr.2021.04.008>
- 509 Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford University Press.
- 510 Harris, K., Subudhi, P. K., Borrell, A., Jordan, D., Rosenow, D., Nguyen, H., ... Mullet, J.  
511 (2007). Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf  
512 senescence. *Journal of Experimental Botany*, 58(2), 327–338.
- 513 Haywood, J. M., Jones, A., Bellouin, N., & Stephenson, D. (2013). Asymmetric forcing from  
514 stratospheric aerosols impacts Sahelian rainfall. *Nature Climate Change*, 3(7), 660–665.  
515 <https://doi.org/10.1038/nclimate1857>
- 516 Ibrahim, Y. M., Marcarian, V., & Dobrenz, A. K. (1986). Drought tolerance aspects in pearl  
517 millet. *Journal of Agronomy and Crop Science*, 156(2), 110–116.  
518 <https://doi.org/10.1111/j.1439-037X.1986.tb00014.x>
- 519 ISRIC. (1990). The global assessment of human induced soil degradation (GLASOD).  
520 Wageningen, The Netherlands.
- 521 Jiang, H., Li, M., Liang, N., Yan, H., Wei, Y., Xu, X., ... Wu, G. (2007). Molecular cloning and  
522 function analysis of the *stay green* gene in rice. *Plant Journal*, 52, 197–209.
- 523 Jiang, Y., Li, C., Paterson, A. H., Sun, S., Xu, R., & Robertson, J. (2018). Quantitative Analysis  
524 of Cotton Canopy Size in Field Conditions Using a Consumer-Grade RGB-D Camera.  
525 *Frontiers in Plant Science*, 8, 1–20. <https://doi.org/10.3389/fpls.2017.02233>
- 526 Johnson, S. M., Cummins, I., Lim, F. L., Slabas, A. R., & Knight, M. R. (2015). Transcriptomic  
527 analysis comparing stay-green and senescent Sorghum bicolor lines identifies a role for  
528 proline biosynthesis in the stay-green trait. *Journal of Experimental Botany*, 66(22), 7061–  
529 7073.
- 530 Jordan, D. R., Tao, Y., Godwin, I. D., Henzell, R. G., Cooper, M., & McIntyre, C. L. (2003).  
531 Prediction of hybrid performance in grain sorghum using RFLP markers. *Theoretical and*  
532 *Applied Genetics*, 106(3), 559–567.
- 533 Joubès, J., Raffaele, S., Bourdenx, B., Garcia, C., Laroche-Traineau, J., Moreau, P., ... Lessire,



- 534 R. (2008). The VLCFA elongase gene family in *Arabidopsis thaliana*: Phylogenetic  
535 analysis, 3D modelling and expression profiling. *Plant Molecular Biology*, 67(5), 547–566.
- 536 Kandji, S. T., Verchot, L., & Mackensen, J. (2006). Climate Change and Variability in the Sahel  
537 Region: Impacts and Adaptation Strategies in the Agricultural Sector.
- 538 Kassahun, B., Bidinger, F. R., Hash, C. T., & Kuruvinashetti, M. S. (2010). Stay-green  
539 expression in early generation sorghum [*Sorghum bicolor* (L.) Moench] QTL introgression  
540 lines. *Euphytica*, 172(3), 351–362.
- 541 Kishor, P. B., Hong Zonglie, Miao Guo-Hua, Hu Chein-An, A., & Verma, D. P. (1995).  
542 Overexpression of delta1-pyrroline-5-carboxylate synthetase increases proline production  
543 and confers osmotolerance in transgenic plants. *Plant Physiology*, 108(4), 1387–1394.
- 544 Kong, W., Kim, C., Zhang, D., Guo, H., Tan, X., Jin, H., ... Paterson, A. H. (2018). Genotyping  
545 by sequencing of 393 sorghum bicolor BTx623 × IS3620C recombinant inbred lines  
546 improves sensitivity and resolution of QTL detection. *G3 Genes/Genomes/Genetics*, 8,  
547 2563–2572.
- 548 Kurukulasuriya, P., Mendelsohn, R., Hassan, R., Benhin, J., Deressa, T., Diop, M., ... Dinar, A.  
549 (2006). Will African agriculture survive climate change? *World Bank Economic Review*,  
550 20(3), 367–388.
- 551 Lee, S. B., Jung, S. J., Go, Y. S., Kim, H. U., Kim, J. K., Cho, H. J., ... Suh, M. C. (2009). Two  
552 *Arabidopsis* 3-ketoacyl CoA synthase genes, *KCS20* and *KCS2/DAISY*, are functionally  
553 redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by  
554 osmotic stress. *The Plant Journal*, 60(3), 462–475.
- 555 Lin, Y., Schertz, K. F., & Paterson, A. H. (1995). Comparative analysis of QTLs affecting plant  
556 height and maturity across the Poaceae, in reference to an interspecific sorghum population.  
557 *Genetics*, 141, 391–411.
- 558 Lopes, M. S., & Reynolds, M. P. (2012). Stay-green in spring wheat can be determined by  
559 spectral reflectance measurements (normalized difference vegetation index) independently  
560 from phenology. *Journal of Experimental Botany*, 63(10), 3789–3798.
- 561 Mace, E., Innes, D., Hunt, C., Wang, X., Tao, Y., Baxter, J., ... Jordan, D. (2019). The Sorghum  
562 QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop  
563 improvement. *Theoretical and Applied Genetics*, 132(3), 751–766.
- 564 Magwanga, R. O., Lu, P., Kirungu, J. N., Lu, H., Wang, X., Cai, X., ... Liu, F. (2018).  
565 Characterization of the late embryogenesis abundant (LEA) proteins family and their role in  
566 drought stress tolerance in upland cotton. *BMC Genetics*, 19(6), 1–31.
- 567 Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., ...  
568 Kresovich, S. (2013). Population genomic and genome-wide association studies of  
569 agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences*, 110(2),  
570 453–458.
- 571 Neph, S., Kuehn, M. S., Reynolds, A. P., Haugen, E., Thurman, R. E., Johnson, A. K., ...  
572 Stamatoyannopoulos, J. A. (2012). BEDOPS: High-performance genomic feature  
573 operations. *Bioinformatics*, 28(14), 1919–1920.
- 574 O'Toole, J. C., & Namuco, O. S. (1983). Role of panicle exertion in water stress induced  
575 sterility. *Crop Science*, 23(6), 1093–1097.  
576 <https://doi.org/10.2135/cropsci1983.0011183x002300060017x>
- 577 Orr, H. A. (1998). The population genetics of adaptation: the distribution of factors fixed during  
578 adaptive evolution. *Evolution*, 52(4), 935–949.
- 579 Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., ...

- 580 Rokhsar, D. S. (2009). The *Sorghum bicolor* genome and the diversification of grasses.  
581 *Nature*, 457(7229), 551–556.
- 582 Paterson, A. H., & Li, Z. K. (2011). Paleo-green revolution for rice. *Proceedings of the National*  
583 *Academy of Sciences of the United States of America*.  
584 <https://doi.org/10.1073/pnas.1107959108>
- 585 Pauli, D., Andrade-Sanchez, P., Carmo-Silva, A. E., Gazave, E., French, A. N., Heun, J., ...  
586 Gore, M. A. (2016). Field-Based High-Throughput Plant Phenotyping Reveals the  
587 Temporal Patterns of Quantitative Trait Loci Associated with Stress-Responsive Traits in  
588 Cotton. *G3 Genes/Genomes/Genetics*, 6(4), 865–879.
- 589 Pironon, S., Etherington, T. R., Borrell, J. S., Kühn, N., Macias-fauria, M., Ondo, I., ... Willis,  
590 K. J. (2019). Potential adaptive strategies for 29 sub-Saharan crops under future climate  
591 change. *Nature Climate Change*, 9, 758–763.
- 592 Pugh, N. A., Horne, D. W., Murray, S. C., Carvalho, G., Malambo, L., Jung, J., ... Rooney, W.  
593 L. (2018). Temporal estimates of crop growth in sorghum and maize breeding enabled by  
594 unmanned aerial systems. *The Plant Phenome Journal*, 1(1), 1–10.
- 595 Rosegrant, M. W., Cai, X., & Cline, S. A. (2002). *World water and food to 2025: Dealing with*  
596 *scarcity*. Washington, D.C.: International Food Policy Research Institute (IFPRI).
- 597 Rosenow, D. T., & Clark, L. E. (1981). Drought tolerance in sorghum. In H. D. Loden & D.  
598 Wilkinson (Eds.), *Proceedings of the 36th Annual Corn and Sorghum Research Conference*  
599 (pp. 11–31). Chicago, IL.
- 600 Rosenow, D. T., Quisenberry, J. E., Wendt, C. W., & Clark, L. E. (1983). Drought tolerant  
601 sorghum and cotton germplasm. *Agricultural Water Management*, 7, 207–222.
- 602 Sanchez, A. C., Subudhi, P. K., Rosenow, D. T., & Nguyen, H. T. (2002). Mapping QTLs  
603 associated with drought resistance in sorghum (*Sorghum bicolor* L. Moench). *Plant*  
604 *Molecular Biology*, 48(5–6), 713–726.
- 605 Seetharama, N., Reddy, B. V. S., Peacock, J. M., & Bidinger, F. R. (1982). Sorghum  
606 improvement for drought resistance. Manila, Philippines: International Rice Research  
607 Institute.
- 608 Serageldin, I. (2004). *Speculations on the future of water and food security*. Washington, D.C.:  
609 International Food Policy Research Institute (IFPRI).
- 610 Subudhi, P. K., Rosenow, D. T., & Nguyen, H. T. (2000). Quantitative trait loci for the stay  
611 green trait in sorghum (*Sorghum bicolor* L. Moench): consistency across genetic  
612 backgrounds and environments. *Theoretical and Applied Genetics*, 101, 733–741.
- 613 Team, R. D. C., & R Development Core Team, R. (2016). R: A Language and Environment for  
614 Statistical Computing. *R Foundation for Statistical Computing*. [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-3-540-74686-7)  
615 [3-540-74686-7](https://doi.org/10.1007/978-3-540-74686-7)
- 616 Thomas, H., & Ougham, H. (2014). The stay-green trait. *Journal of Experimental Botany*,  
617 65(14), 3889–3900.
- 618 Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., ... Buckler, E. S.  
619 (2011). Genome-wide association study of leaf architecture in the maize nested association  
620 mapping population. *Nature Genetics*, 43(2), 159–162.
- 621 Tuberosa, R. (2012). Phenotyping for drought tolerance of crops in the genomics era. *Frontiers*  
622 *in Physiology*, 3(347), 1–26.
- 623 UNESCO. (2002). *Vital water graphics, water use and management*. Paris.
- 624 Vadez, V., Krishnamurthy, L., Hash, C. T., Upadhyaya, H. D., & Borrell, A. K. (2011). Yield,  
625 transpiration efficiency, and water-use variations and their interrelationships in the sorghum



- 626 reference collection. *Crop and Pasture Science*, 62(8), 645–655.  
627 <https://doi.org/10.1071/CP11007>
- 628 Vadez, Vincent, Krishnamurthy, L., Hash, C. T., Upadhyaya, H. D., & Borrell, A. K. (2011).  
629 Yield, transpiration efficiency, and water-use variations and their interrelationships in the  
630 sorghum reference collection. *Crop and Pasture Science*, 62, 645–655.
- 631 Wallace, J. G., Zhang, X., Beyene, Y., Semagn, K., Olsen, M., Prasanna, B. M., & Buckler, E. S.  
632 (2016). Genome-wide association for plant height and flowering time across 15 tropical  
633 maize populations under managed drought stress and well-watered conditions in Sub-  
634 Saharan Africa. *Crop Science*, 56(5), 2365–2378.  
635 <https://doi.org/10.2135/cropsci2015.10.0632>
- 636 Würschum, T., Liu, W., Gowda, M., Maurer, H. P., Fischer, S., Schechert, A., & Reif, J. C.  
637 (2012). Comparison of biometrical models for joint linkage association mapping. *Heredity*,  
638 108(3), 332–340. <https://doi.org/10.1038/hdy.2011.78>
- 639 Xu, R., Li, C., Paterson, A. H., Jiang, Y., Sun, S., & Robertson, J. S. (2018). Aerial images and  
640 convolutional neural network for cotton bloom detection. *Frontiers in Plant Science*, 8, 1–  
641 17.
- 642 Xu, W., Subudhi, P. K., Crasta, O. R., Rosenow, D. T., Mullet, J. E., & Nguyen, H. T. (2000).  
643 Molecular mapping of QTLs conferring stay-green in grain sorghum (*Sorghum bicolor* L.  
644 Moench). *Genome*, 43(3), 461–469.
- 645 Yu, J., Holland, J. B., McMullen, M. D., & Buckler, E. S. (2008). Genetic design and statistical  
646 power of nested association mapping in maize. *Genetics*, 178(1), 539–551.
- 647 Zhang, D., Guo, H., Kim, C., Lee, T.-H., Li, J., Robertson, J., ... Paterson, A. H. (2013).  
648 CSGRqtl: A comparative quantitative trait locus database for saccharinae grasses. *Plant*  
649 *Physiology*, 161, 594–599.
- 650 Zhang, D., Kong, W., Robertson, J., Goff, V. H., Epps, E., Kerr, A., ... Paterson, A. H. (2015).  
651 Genetic analysis of inflorescence and plant height components in sorghum (Panicoidae) and  
652 comparative genetics with rice (Oryzoidae). *BMC Plant Biology*, 15(1), 1–15.
- 653 Zheng, H. J., Wu, A. Z., Zheng, C. C., Wang, Y. F., Cai, R., Shen, X. F., ... Dong, S. T. (2009).  
654 QTL mapping of maize (*Zea mays*) stay-green traits and their relationship to yield. *Plant*  
655 *Breeding*, 128(1), 54–62.  
656

## 657 **Figure Legends**

658 **Figure 1** Phenotypic distribution of eight traits in the BC-NAM population at three  
659 environments. x axis represents trait metrics and y axis represent distribution density. Vertical  
660 dashed lines represent population means at respective environment. NL: number of leaves, HE:  
661 head exertion, LS: leaf senescence, NT: number of tillers, SGY: single plant-based grain yield,  
662 CPP: count of plants per plot, PGY: plot-based grain yield, TSW: 1000-seed weight.

663 **Figure 2** Trait correlations among nine traits studied in this BC-NAM population. Note that in  
664 addition to the eight traits in this study, we also included three other traits including days to  
665 flowering (DF), days to maturity (DM), and plant height (PH), that were genetically dissected in  
666 a parallel study.

667 **Figure 3** Genomic regions associated with drought-related traits in the sorghum BC-NAM  
668 population. (A) Distribution of the 154 joint-linkage QTLs and 233 GWAS hits for six traits,  
669 which were represented with blue dots. NL: number of leaves, HE: head exertion, LS: leaf  
670 senescence, NT: number of tillers, SGY: single plant-based grain yield, CPP: count of plants per  
671 plot, PGY: plot-based grain yield, TSW: 1000-seed weight. (B) Fixation index ( $F_{ST}$ ) between the  
672 58 drought-sensitive lines, which lost one of two replicates at both Kobo and Meiso, and the  
673 remaining 1113 lines. Orange vertical lines in both panels represent  $F_{ST}$  peaks.

674 **Figure 4** Marker-trait associations of drought-related traits. Each panel shows the associations  
675 detected in joint-linkage model (top) and in GWAS model (bottom). (A) number of leaves in  
676 Kobo, (B) number of leaves in Meiso, and (C) head exertion in Meiso. In joint-linkage analysis,  
677 parental allelic effects were color coded, with blue represents positive effect and orange  
678 represents negative effect. Candidate genes were shown in green vertical lines and annotated  
679 with respective gene names.

680 **Figure 5** Genome-wide prediction of sorghum grain yield. Phenotypic prediction was performed  
681 across the entire sorghum BC-NAM population in 50 five-fold cross-validation using either all  
682 4395 single nucleotide polymorphisms (SNPs), the model-selected SNPs, or an equivalent  
683 number of random SNPs. (A) Distribution of genomic prediction accuracy for plot-based grain  
684 yield (PGY) across the entire BC-NAM population at three environments. (B) Distribution of  
685 genomic prediction accuracy for 1000-seed weight (TSW) across the entire BC-NAM population  
686 at three environments.  $P$  values comparing distributions via a two-sided  $t$ -test were shown in  
687 each plot.

688 **Figure S1** Geographical information and natural environmental conditions of three field trial  
689 locations in Ethiopia.

690 **Figure S2** Geographical information of three field trial locations in Ethiopia

691 **Figure S3** Phenotypic distribution of eight traits within each BC-NAM family at three  
692 environments. x axis represents family names and y axis represent trait metric. Horizontal dashed  
693 lines represent recurrent parent means at respective environment.

694 **Figure S4** Joint-linkage QTL allelic effects distribution for six traits. NL: number of leaves, HE:  
695 head exertion, LS: leaf senescence, NT: number of tillers, SGY: single plant-based grain yield,  
696 CPP: count of plants per plot, PGY: plot-based grain yield, TSW: 1000-seed weight.

697

698 **Table 1** Description of the BC-NAM population

Name	Country of origin	Kobo		Meiso		Sheraro	
		Pop size	No. lines with plant loss (%)	Pop size	No. lines with plant loss (%)	Pop size	No. lines with plant loss (%)
IS10876	Nigeria	135	39 (29%)	151	27 (18%)	153	0 (0%)
IS14298	South Africa	121	48 (40%)	132	16 (12%)	131	1 (1%)
IS14446	Sudan	134	50 (37%)	145	22 (15%)	149	0 (0%)
IS14556	Ethiopia	32	14 (44%)	34	2 (6%)		
IS15428	Cameroon	124	43 (35%)	142	25 (18%)	144	0 (0%)
IS16044	Cameroon	36	9 (25%)	37	2 (5%)		
IS16173	Cameroon	98	25 (26%)	111	18 (16%)	112	0 (0%)
IS2205	India	36	20 (56%)	40	6 (15%)		
IS22325	Botswana	101	33 (33%)	120	25 (21%)	120	0 (0%)
IS23988	Yemen	34	15 (44%)	42	10 (24%)	36	1 (3%)
IS3583	Sudan	104	34 (33%)	119	22 (18%)	119	0 (0%)
IS9911	Sudan	76	17 (22%)	82	14 (17%)	82	0 (0%)
Total		1031	347 (34%)	1155	189 (16%)	1046	2 (0%)

699  Founder lines used in the BC-NAM population. Note that population IS32234 was discarded  
700 due to severe contamination.

701 §Source information is obtained from the International Crops Research Institute for the Semi-  
702 Arid Tropics (ICRISAT): <http://genebank.icrisat.org/IND/Passport?Crop=Sorghum>.

703 **Table 2** Summary of joint-linkage QTL of each trait in the BC-NAM population

Trait§	Kobo			Meiso			Sheraro		
	$H^2$ ¶	No. QTL	Model $R^2$ □	$H^2$	No. QTL	Model $R^2$	$H^2$	No. QTL	Model $R^2$
HE	0.58	7	0.26 (45%)	0.59	12	0.37 (63%)			
LS	0.42	3	0.21 (50%)	0.28	9	0.21 (75%)			
NT	0.27	7	0.19 (70%)	0.26	10	0.24 (92%)			
NL	0.44	13	0.33 (75%)	0.49	6	0.18 (37%)	0.60	8	0.19 (32%)
SGY	0.34	4	0.12 (35%)	0.28	8	0.19 (68%)	0.28	4	0.11 (39%)
CPP	0.38	2	0.10 (26%)	0.16	5	0.14 (88%)	0.18	2	0.06 (33%)
PGY	0.39	10	0.35 (90%)	0.26	8	0.23 (88%)	0.28	6	0.18 (64%)
TSW	0.59	6	0.18 (31%)	0.31	10	0.24 (77%)	0.63	14	0.31 (49%)

704 § HE: head exertion; LS: leaf senescence; NL: number of leaves; NT: number of tillers; SGY:  
 705 single plant-based grain yield; CPP: count of plants per plot; PGY: plot-based grain yield; TSW:  
 706 thousand seed weight.

707 ¶Broad-sense heritability of BC-NAM population.

708 □ Variance explained by joint-linkage model after fitting family term and detected QTL.

709 Numbers in parentheses represent proportion of genetic variance explained by the join-linkage  
 710 model, which was calculated by dividing model  $R^2$  by  $H^2$ .

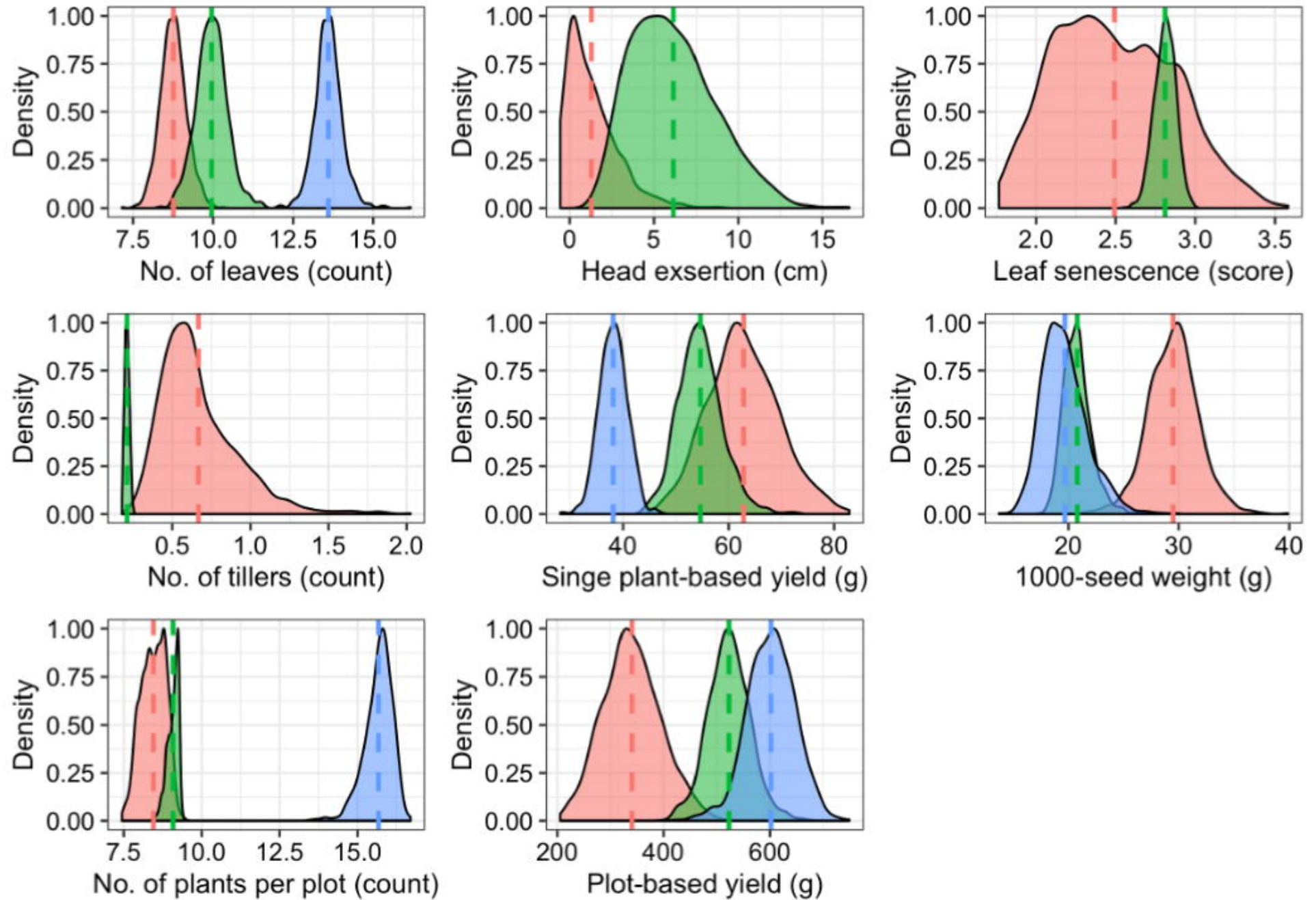
711

712

713

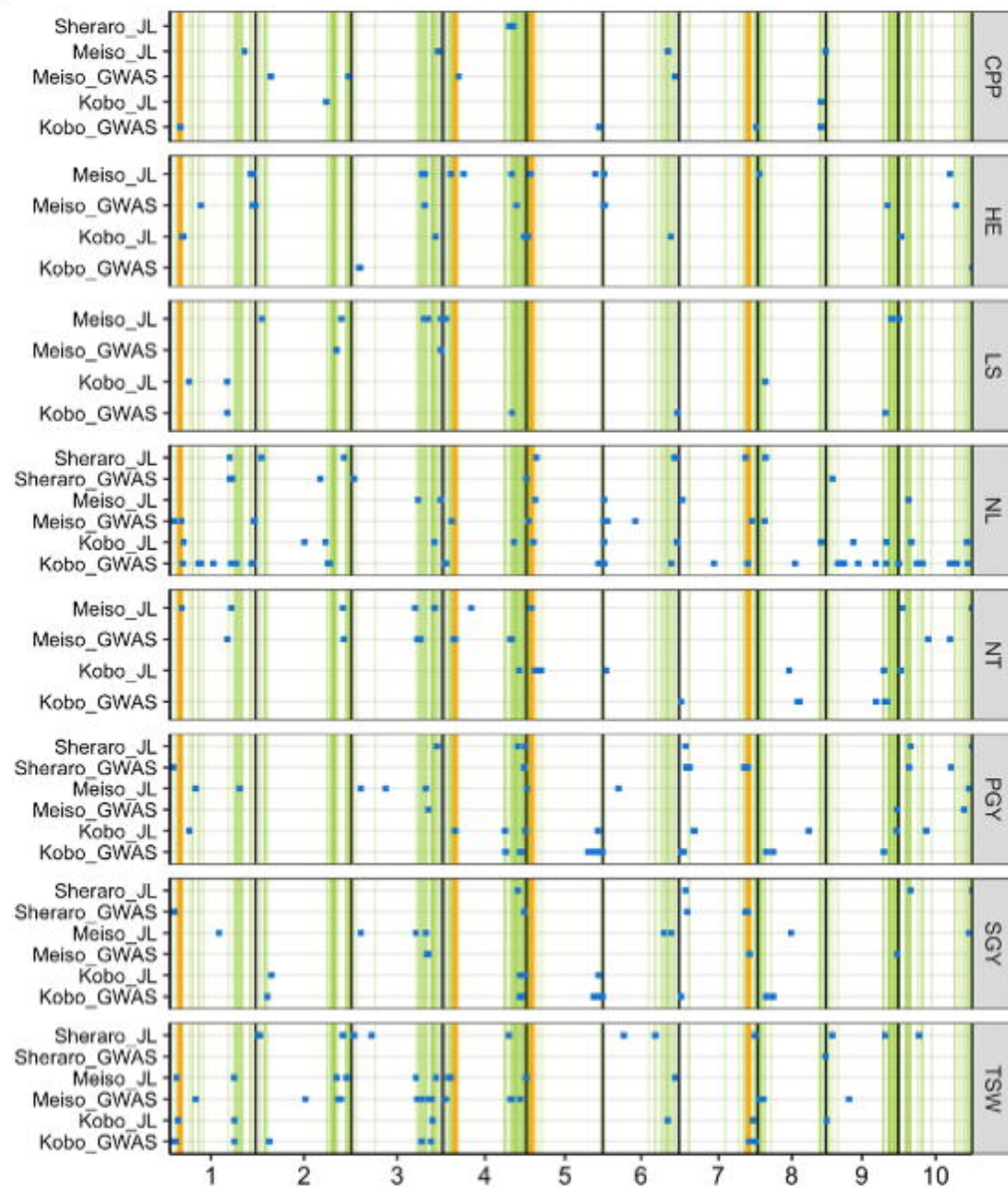
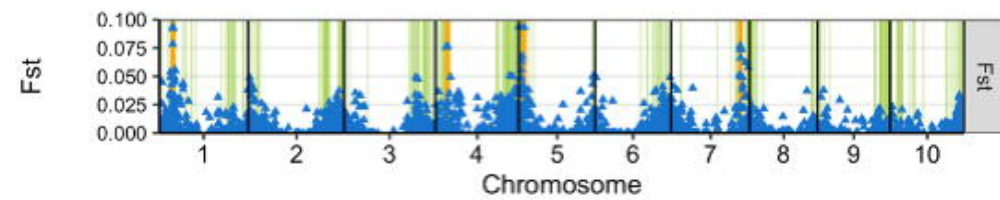
714

Environment ■ Kobo ■ Meiso ■ Sheraro

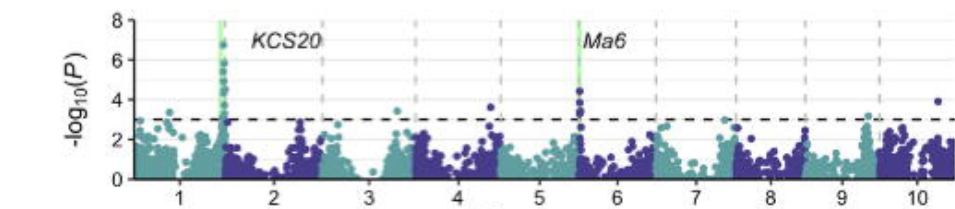
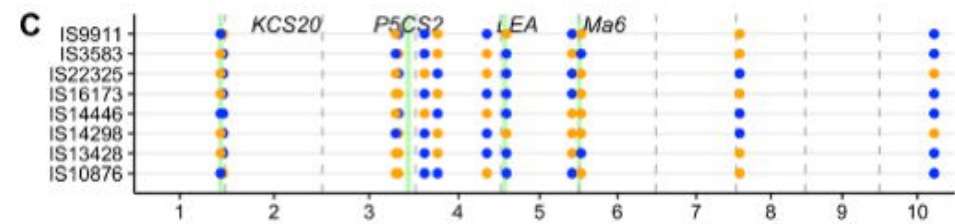
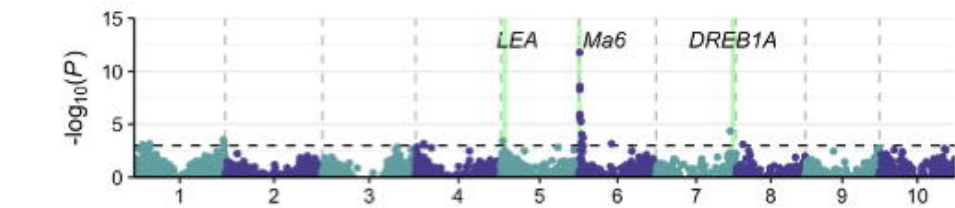
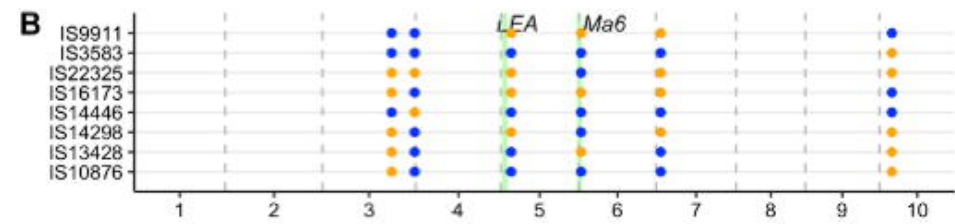
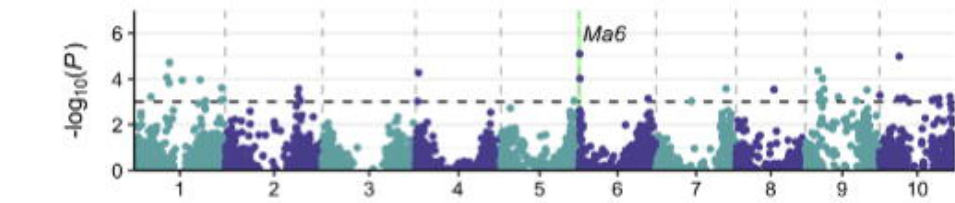
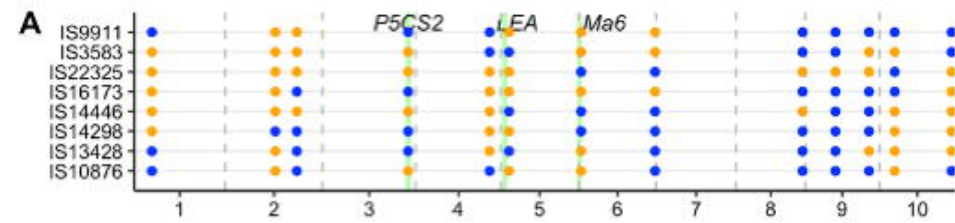




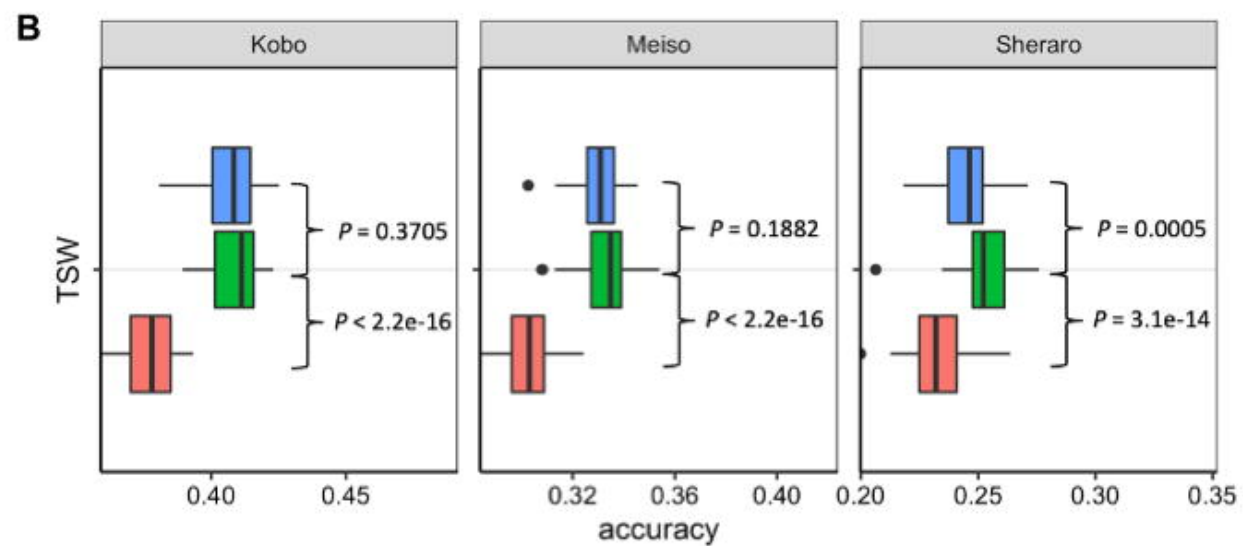
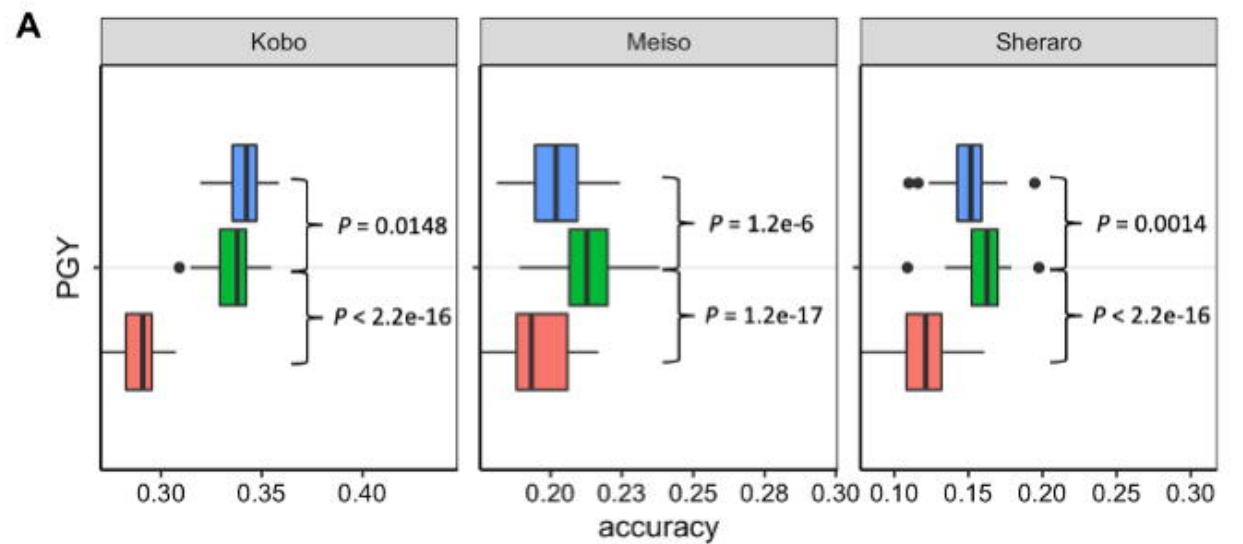


**A****B**





Chromosome



Method ■ Random SNPs ■ Selected SNPs ■ Whole genome