

RESEARCH ARTICLE

Genome-wide association mapping for component traits of drought tolerance in dry beans (*Phaseolus vulgaris* L.)Bruce Mutari^{1,2*}, Julia Sibiya¹, Admire Shayanowako¹, Charity Chidzanga³, Prince M. Matova⁴, Edmore Gasura⁵

1 School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Scottsville, Pietermaritzburg, South Africa, **2** Department of Research and Specialist Services, Crop Breeding Institute, Harare, Zimbabwe, **3** School of Agriculture, Food and Wine, The University of Adelaide, Glen Osmond, Australia, **4** Mukushi Seeds (Pvt) Ltd, Mt Humpden, Harare, Zimbabwe, **5** University of Zimbabwe, Mt Pleasant, Harare, Zimbabwe

* brucemutari@gmail.com

OPEN ACCESS

Citation: Mutari B, Sibiya J, Shayanowako A, Chidzanga C, Matova PM, Gasura E (2023) Genome-wide association mapping for component traits of drought tolerance in dry beans (*Phaseolus vulgaris* L.). PLoS ONE 18(5): e0278500. <https://doi.org/10.1371/journal.pone.0278500>

Editor: Kuldeep Tripathi, NBPGR: National Bureau of Plant Genetic Resources, INDIA

Received: November 16, 2022

Accepted: April 30, 2023

Published: May 18, 2023

Copyright: © 2023 Mutari et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting information](#) file.

Funding: The work presented here was led by Bruce Mutari with financial support from the Swiss Agency for Development and Cooperation (SDC) and the Government of Zimbabwe under the Grant No. C-013-18. The funders did not play any role in the study design, data collection and analysis, decision to publish, and preparation of the manuscript.

Abstract

Understanding the genetic basis of traits of economic importance under drought stressed and well-watered conditions is important in enhancing genetic gains in dry beans (*Phaseolus vulgaris* L.). This research aims to: (i) identify markers associated with agronomic and physiological traits for drought tolerance and (ii) identify drought-related putative candidate genes within the mapped genomic regions. An andean and middle-american diversity panel (AMDP) comprising of 185 genotypes was screened in the field under drought stressed and well-watered conditions for two successive seasons. Agronomic and physiological traits, viz., days to 50% flowering (DFW), plant height (PH), days to physiological maturity (DPM), grain yield (GYD), 100-seed weight (SW), leaf temperature (LT), leaf chlorophyll content (LCC) and stomatal conductance (SC) were phenotyped. Principal component and association analysis were conducted using the filtered 9370 Diversity Arrays Technology sequencing (DARTseq) markers. The mean PH, GYD, SW, DPM, LCC and SC of the panel was reduced by 12.1, 29.6, 10.3, 12.6, 28.5 and 62.0%, respectively under drought stressed conditions. Population structure analysis revealed two sub-populations, which corresponded to the andean and middle-american gene pools. Markers explained 0.08–0.10, 0.22–0.23, 0.29–0.32, 0.43–0.44, 0.65–0.66 and 0.69–0.70 of the total phenotypic variability (R^2) for SC, LT, PH, GYD, SW and DFW, respectively under drought stressed conditions. For well-watered conditions, R^2 varied from 0.08 (LT) to 0.70 (DPM). Overall, 68 significant ($p < 10^{-03}$) marker-trait associations (MTAs) and 22 putative candidate genes were identified across drought stressed and well-watered conditions. Most of the identified genes had known biological functions related to regulating the response to drought stress. The findings provide new insights into the genetic architecture of drought stress tolerance in common bean. The findings also provide potential candidate SNPs and putative genes that can be utilized in gene discovery and marker-assisted breeding for drought tolerance after validation.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Common bean (*Phaseolus vulgaris* L., $2n = 2x = 22$) is one of the major pulse crops consumed worldwide with a relatively small diploid genome size of approximately 473 Mb [1]. It is a cheap source of proteins and important micronutrients such as iron (Fe) and zinc (Zn) for millions in many African and Latin American countries [2, 3]. Beebe et al. [4] reported that Sub-Saharan Africa (SSA) and Latin America produce the largest volume of common beans, representing more than 60% of the world's bean production. Common bean was subjected to two parallel domestication events on the American continent, resulting in two different primary gene pools namely the andean and the middle-american [5, 6]. The andean gene pool originated from the Andes mountains of South America and consists of medium (25–40 grams per 100 seeds) or large (≥ 40 grams per 100 seeds) seeded genotypes [7]. On the other hand, the middle-american gene pool is native to Central America and Mexico, and comprises of small seeded genotypes (≤ 25 grams per 100 seeds). According to Bitocchi et al. [8], there is more genetic variation within the middle-american gene pool compared to the andean gene pool.

Common beans are notably sensitive to climatic and environmental variations. This is aggravated by the fact that most bean growing regions in the world experience different production constraints including intermittent and terminal drought stress which adversely affect grain yield [9–12]. As reported by Katungi et al. [13], 73% of common beans production in SSA occurs in environments which experience moderate to severe drought stress. Beebe et al. [4], Hoyos-Villegas et al. [14] and Valdisser et al. [15] reiterated that drought stress is the most important grain yield-limiting abiotic factor of dry beans worldwide. It is predicted from various climate models that the duration and frequency of droughts are expected to increase in SSA [16]. Drought stress reduces stomatal conductance, total chlorophyll content, leaf expansion, number of days to physiological maturity, seed yield and biomass, number of pods and seeds per plant, seed size and harvest index [17–22]. According to Asfaw et al. [23], severe drought stress can result in grain yield losses of up to 80%. In Zimbabwe, grain yield reductions of more than 50% were reported by Mutari et al. [24] under terminal drought stress.

As reported by Mutari et al. [25], dry beans farmers in Zimbabwe have been using different mitigation strategies to minimize grain yield losses due to terminal drought stress. These strategies include soil mulching, ridging, cultivating the soil to retain more moisture and reducing the area under the bean crop. However, host plant resistance is a more sustainable, environmentally friendly and labour saving technology for managing drought stress in common beans compared to the multiple cultural practices. For this reason, most dry beans breeding programmes aim to introduce drought tolerance into new cultivars to address the needs and preferences of smallholder farmers in the face of climate change [26].

Several researchers have successfully used different types of deoxyribonucleic acid (DNA)-based marker systems in association mapping of complex traits in common beans. The most widely used marker systems include simple sequence repeats (SSRs; [27–29]), amplified fragment length polymorphisms (AFLPs; [28, 30]), single nucleotide polymorphisms (SNPs; [3, 14, 31–34]) and microarray based Diversity Arrays Technology (DArT; [15, 35]) markers. However, SNP markers are widely preferred in marker assisted selection (MAS), genetic diversity analyses, genomic selection, haplotype mapping, genome wide association studies (GWAS), linkage map construction and population genetics [36]. They are widely preferred because they exhibit high level of polymorphism and occur in abundance as differences of individual nucleotides between individuals.

Understanding the underlying genetic architecture of agronomic and physiological traits under drought stressed and well-watered conditions is a prerequisite for the genetic improvement of these traits using MAS. Thus, dissecting the genetic basis of multiple polygenic traits

of economic importance such as drought tolerance with respect to the genomic regions and/or genes involved and their effects is important. It is particularly important in improving genetic gains when breeding for superior grain yield in dry beans under drought stressed and well-watered environments. This can be accomplished through complementary approaches such as GWAS and genomic prediction models [6]. Genome wide association study is a powerful tool for characterizing the genetic basis of quantitative traits at high level of genetic resolution [34, 37–41]. It is also a powerful tool for identifying multiple candidate genes (marker alleles) associated with variation in quantitative traits (marker-trait associations; MTA) of interest in crop species using high density DNA markers [34, 37–41].

Genome wide association study is also known as association mapping (AM) or linkage disequilibrium (LD) mapping [42]. It is based on linkage LD and historical recombination events of alleles of detected quantitative trait loci (QTL) at relatively high level of genetic resolution [43, 44]. The high level of genetic or mapping resolution is due to high genetic variability and high number of recombination events in diverse natural population such as landraces, elite breeding lines and improved cultivars [43–45]. The historical recombination events would have naturally occurred during the evolution and domestication of the crop, and crop improvement (several generations) [33]. Therefore, it is inexpensive and reduces research time (no need to develop a mapping population) with greater allele numbers. The identification of genomic regions and diagnostic genetic markers associated with grain yield and yield-attributing traits under drought stressed and well-watered conditions will facilitate trait introgression and MAS.

Genome wide association study has been successfully used to detect MTAs and QTLs in common beans. Several QTLs associated with disease and insect pest tolerance have been identified in dry beans [32, 46–50]. Similarly, MTAs were identified for drought tolerance traits in dry beans [14, 15, 51–53]. Also, MTAs were identified for nutritional composition-related traits [6, 33], symbiotic nitrogen fixation [54], cooking time [55] and photosynthetic traits [34, 56] in dry beans. Genomic regions governing agronomic traits in drought stressed and high yield potential environments were also identified in dry beans [1, 6, 14, 34, 57]. Thus, several significant MTAs have been identified in previous GWAS studies for agronomic traits in drought stressed environments. However, the use of very low thresholds ($-\log_{10} p\text{-value} \geq 3.0$) in most of the studies in determining significant MTAs might have resulted in many false positives. In addition, despite the fact that several QTLs/MTAs associated with agronomic traits have been identified in dry beans, further genetic studies are required using different genetic backgrounds to reach a saturation point. Moreover, most of the reported putative genes for agronomic and physiological traits were detected under yield potential environments.

Additionally, some of the previous mapping studies [14, 17, 51, 58–61] conducted on agronomic and physiological traits used a small population size and a limited number of molecular markers. This resulted in QTL with low resolution or poor estimation of marker effects. Furthermore, this also made it difficult to make inferences on putative candidate genes correlated with the identified QTL. In addition, some of the previously identified QTLs explained low total genetic variance [23]. Moreover some of the previously identified QTLs were sometimes not stable across environments due to genotype by environment interaction (GEI) [52]. Thus their potential for MAS in developing genotypes that are tolerant to drought stress was inconclusive. Therefore, additional studies are required to dissect the genetic basis of agronomic and physiological traits in dry beans under drought stressed and optimal environments for increased genetic gains. The objectives of this study were: (i) to identify single nucleotide polymorphism (SNP) markers significantly associated with agronomic and physiological traits for drought tolerance and; (ii) to identify drought-related putative candidate genes associated with traits within the mapped genomic regions.

Materials and methods

Description of the study location

The field experiments (drought stress and well-watered) were conducted at the screening site for drought stress tolerance located at Save Valley Experiment Station (SVES), Zimbabwe. The experiments were carried out during the 2019 and 2020 dry winter seasons (April–July). Save Valley Experiment Station is characterised by clay soils and is located in the drier lowveld region of Zimbabwe where dry beans are commercially produced during the dry winter season (Table 1). The research station receives an average annual rainfall of 450 mm which is usually distributed between the months of December and April. In both seasons, no precipitation was received during the trial evaluation period. Historically, SVES presents few rainfall occurrences during the dry winter season [24]. Daily temperatures (°C) and relative humidity (%) were recorded with a digital weather station (Table 1) during the growing seasons. More details on the agro-ecological characteristics of SVES are outlined in Table 1.

Germplasm

A total of 185 dry beans genotypes constituted the andean and middle-American diversity panel (AMDP). The AMDP comprised of landrace collections (25), released cultivars (18) and elite breeding lines (142) of different market classes such as sugars, calimas, small whites, large whites and large red kidneys (S1 Table). The genotypes were sourced from public and private breeding institutions located in different geographic regions. These included the Alliance of Bioversity International and International Center for Tropical Agriculture (ABC) in Colombia (87), ABC in Malawi (67), ABC in Uganda (18), Ethiopian Institute of Agricultural Research (EIAR) in Ethiopia (3), Crop Breeding Institute in Zimbabwe (6) and Seed-Co, also in Zimbabwe (4) (S1 Table). The elite breeding lines were selected on the basis of their genetic divergence and breeding history for drought tolerance. On the other hand, the released cultivars and landrace collections were included in the study to broaden the genetic diversity for agronomic and physiological traits. Furthermore, the released cultivars were included in the study due to their great economic importance to the Zimbabwean dry bean breeding program. The details of the genotypes used in the study are presented in S1 Table.

Table 1. Geographic information, monthly weather conditions and soil characteristics during the growing seasons at Save Valley Experiment Station, Zimbabwe (April to July, 2019 and 2020).

Parameter		2019 dry season				2020 dry season			
		April	May	June	July	April	May	June	July
Temperature (°C)	Max	33.00	29.00	28.00	30.00	31.00	28.50	27.00	32.00
	Min	9.00	9.50	10.00	12.00	11.50	8.00	8.50	12.50
Relative Humidity (%)	Max	82.00	95.00	69.00	91.00	74.00	85.00	69.00	71.00
	Min	42.00	56.00	44.00	25.00	46.00	59.00	50.00	30.00
Total Rainfall (mm)		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Soil type		Clay				Clay			
Latitude		20° 32' S				20° 43' S			
Longitude		33° 09' E				33° 03' E			
Altitude (m.a.s.l)		452				449			

masl = meters above sea level, mm = millimetres, ppm = parts per million, Max = maximum and Min = minimum.

<https://doi.org/10.1371/journal.pone.0278500.t001>

Field phenotyping of the diversity panel

Experimental design, irrigation scheduling and trial management. The AMDP was evaluated side by side under drought stressed and well-watered treatment conditions. In both seasons, the genotypes in both drought stressed and well-watered treatments were established in a 5 x 37 alpha lattice design with two replications. The seepage of water from the well-watered treatment to the drought stressed treatment was minimized by maintaining a 30 m buffer zone between the two treatments. Each genotype was hand planted in four-row plots of 3 m in length, and an inter-row spacing of 0.45 m. Compound D (N = 7%, P = 14%, K = 7%) was applied at planting at a rate of 300 kg/ha. Ammonium nitrate (34.5% N) was applied in both drought stressed and well-watered treatments as a top-dressing fertilizer thirty days after emergence at a rate of 100 kg/ha. An overhead sprinkler irrigation system was used to irrigate both drought stressed and well-watered treatments during both seasons of evaluation.

The irrigation cycles in both drought stressed and well-watered treatments were as described by Mutari et al. [24]. The amount of soil moisture was monitored using tensiometers (2 per station) which were installed (0.5 m and 1 m depths) at four different stations in each treatment. Soil moisture was kept at field capacity during the crop development period in the well-watered treatment. In the drought stressed treatment, soil moisture was kept at field capacity until when 80% of the plants had flowered. After that, terminal drought stress was imposed up to physiological maturity by withholding irrigation water to 30% of the field capacity before re-irrigating on the basis of readings from tensiometers. Six cycles of irrigation amounting to 256 mm and 253 mm, respectively (each cycle with roughly 42 mm) were applied to the drought stressed experiments during the 2019 and 2020 seasons. On the other hand, 10 cycles of irrigation were applied to the well-watered experiments in 2019 (424 mm) and 2020 (418 mm), each cycle with roughly 42 mm). In both seasons, recommended agronomic practices were followed for the management and control of pests such as diseases, insects and weeds.

Collection of data on agronomic and physiological traits. At the flowering stage of growth, the number of days from planting to 50% flowering (DFW) were recorded in both treatments. The DFW was recorded when 50% of the plants in a plot had at least one or more open flowers. At mid-pod filling, leaf temperature (LT; °C), stomatal conductance (SC; $\text{mmol m}^{-2} \text{s}^{-1}$) and leaf chlorophyll (LCC) content were collected on all genotypes in both drought stressed and well-watered treatments. The LT and SC data were recorded from the surface of the uppermost fully expanded young leaf. This was done between 11:00 am to 14:00 pm using a FLUKE precision infrared thermometer (Everest Interscience, Tucson, AZ, USA) and a hand-held leaf porometer (Decagon Devices[®], Pullman, WA, USA), respectively. Three readings were collected on three different randomly chosen plants from each plot per replicate in both drought stressed and well-watered treatments. The three measurements were averaged to obtain one final reading per plot. Phenotyping for LT and SC was done for an average of six days on clear, sunny days with minimal wind. Regarding the LCC, this was measured using a soil and plant analysis development (SPAD) chlorophyll meter (SPAD-502*Plus*, Konica-Minolta, Osaka, Japan) on two fully developed leaves of three plants in each plot. Then, the average value was calculated.

At physiological maturity, the following traits were recorded from the two inner rows from every plot for every genotype in both treatments and seasons: plant height (PH; cm), days from planting to physiological maturity (DPM), grain yield (GYD; kg/ha) and 100-seed weight (SW; g). Plant height which was averaged from three plants per plot was measured from the base of the plant (soil surface) to the top node bearing at least one dry pod. The DPM were recorded as the average number of days from planting to when 95% of pods in a plot lost their

green colour. Grain yield was recorded from the two middle rows in each plot using a weighing scale. Grain yield was subsequently converted to kilograms per hectare (kg/ha) at 12.5% moisture basis. The SW was determined using a beam balance weighing scale by measuring the weight of 100 seeds that had been selected randomly from each plot harvest. Genotypes which combined high grain yield (above the trial mean) with desirable values for stomatal conductance (high), leaf chlorophyll (high), canopy temperature (low), physiological maturity (early) and seed size (large) under drought stressed conditions were presumed to be tolerant to drought stress [22].

Statistical analysis of phenotypic data

Before conducting analysis of variance, normality tests were conducted in Genstat Discovery 18th Edition [62] using residuals of the agronomic and physiological traits. The agronomic and physiological traits were analysed in Genstat Discovery 18th Edition [62] using mixed models from which the best linear unbiased predictors (BLUPs) were obtained. The BLUPs were estimated for the studied traits to minimize the environmental and seasonal effects. The BLUPs for each entry were estimated through individual environment (drought stressed or well-watered) analysis, and by combined analysis (across water regimes). In the first step of analysis (single-environment analysis), the phenotypic data of each individual environment were analysed using a mixed linear model (MLM). In this model, blocks and genotypes were treated as random effects, and replications were considered as fixed effects. Genotype effects were declared to be random to enable the calculation of BLUPs and broad-sense heritability (H^2). The MLM presented below was fitted:

$$Y_{ijl} = \mu + g_i + r_j + b_{lj} + e_{ijl} \quad (1)$$

where Y_{ijl} = is the phenotypic observation of the genotype i in replicate j in block l within replicate j , μ = grand mean effect, g_i = random effect associated with genotype i , r_j = fixed effect associated with replicate j , b_{lj} = random effect associated with block l nested within replicate j , and e_{ijl} = residual effect associated with observation ijl . For a combined or multi-environment analysis, a MLM was used. In this model, blocks nested within replications, replicates nested within environments, genotypes and their interactions with environments (GEI) were considered as random effects. Environments, defined as year x water regime combination were considered as fixed effects. The MLM presented below was fitted:

$$Y_{ijkl} = \mu + G_i + E_j + R_{k[j]} + B_{l[jk]} + GE_{ij} + e_{ijkl} \quad (2)$$

where Y_{ijkl} = effect of genotype i in environment j and k th replication within environment j and l th block nested within replicate k and environment j , μ = grand mean, G_i = random effect of the i th genotype, E_j = fixed effect of the j th environment, $R_{k[j]}$ = random effect associated with the replicate k nested within environment j , $B_{l[jk]}$ = random effect of block l nested within environment j and replicate k , GE_{ij} = random effect of the interaction between genotype i and environment j , and e_{ijkl} = random error associated with observation $ijkl$. The analysis was performed using the Restricted Maximum Likelihood (REML) method implemented in GenStat 18th edition [62]. Broad-sense heritability estimates for the agronomic and physiological traits were calculated following the formula proposed by Cullis et al. [63]. Heritability was classified as low when less than 30%, moderate when between 30–60% and high when more 60% [64]. Drought intensity index (DII) at the location, percentage GYD reduction (%GYR) due to drought stress, drought susceptibility index (DSI), geometric mean productivity (GMP) and drought tolerance index (DTI) of each entry were calculated as described by Mutari et al. [24]. The drought tolerance indices were used in this study to support the observed phenotypes

under well-watered and drought stressed conditions with respect to agronomic and physiological traits. Genotypes with low values for %GYR, DSI and high values for DTI and GMP were presumed to be tolerant to drought stress [24]. Furthermore, a ranking method was used to select superior drought tolerant genotypes by calculating the mean rank of each genotype across all the studied indices.

Genotyping of the diversity panel

Genomic DNA of the 185 genotypes was extracted from young leaves of 2-week old bean plants following the plant DNA extraction protocol for Diversity Arrays Technology (DArT; [65]). A NanoDrop Spectrophotometer (ND-8000, NanoDrop Technologies, Inc.) was used to determine the concentration of the DNA. Agarose gel (1% agarose gel) electrophoresis was used to evaluate the quality of the DNA. The DNA from the samples used in this study were genotyped using the Diversity Arrays Technology Sequencing (DArTseq) protocol using a set of 24,450 silico DArT markers. The DArT markers used were evenly distributed across all the 11 chromosomes of common bean. Genotyping by sequencing (GBS) was done at the Biosciences Eastern and Central Africa (BecA) Hub of the International Livestock Research Institute (BecA-ILRI) in Kenya. The silico DArTs used had polymorphic information content (PIC) values ranging from 0.01 to 0.50, reproducibility values of 1.00, and the proportion of missing data per marker was 7% (mean call rate of 93%, ranging from 81 to 100%). The entire data set of SNP markers was filtered in TASSEL v5.2 [66] to remove SNP loci with unknown physical positions on the common bean genome, monomorphic SNPs, and SNP markers with more than 20% missing data and minor allele frequency (MAF) of less than 5% (<0.05) threshold [15, 49, 67]. A final total of 9370 (38%) DArTseq-derived SNPs distributed across the 11 chromosomes were retained after filtering for use in association analysis and population structure analysis via principal component analysis (PCA).

Inference of population structure

The genotypic data was imputed for missing alleles of SNPs on the KDCompute online server (<https://kdcompute.igs-africa.org/kdcompute/>). This was done using the optimal imputation algorithm to increase the power of the study. KDCompute was also used to graphically visualize the distribution of SNPs across the common bean genome. The population genetic structure was determined based on the Bayesian model-based clustering approach using the Bayesian inference program in STRUCTURE software version 2.3.4 [68]. A subset of additionally filtered SNP markers (4095) at or near Hardy-Weinberg equilibrium ($r^2 < 0.8$) and that covered the entire genome were used in population structure analysis with STRUCTURE [14, 15, 31]. This was done to reduce the background and admixture linkage disequilibrium (LD) owing to linked loci [68].

Settings for the STRUCTURE program were set as follows to derive the population structure: a burn-in period length of 10,000, and after burn-in, 10,000 Markov Chain-Monte Carlo (MCMC) repetitions. The number of sub-populations or clusters (K) was set from 1 to 10, with ten independent runs for each K [3, 48, 55]. The best K-value explaining the population structure was inferred using the Delta K (ΔK) method in Evanno et al. [69] implemented in the on-line tool structure harvester software [70]. Genotypes with ancestry probability/coefficient ≥ 0.90 ($\geq 90\%$) (pure genotypes) for the andean sub-population were allocated to the andean gene pool [31, 71] (S1 Table). On the other hand, genotypes with ancestry probability ≥ 0.90 for the middle-american sub-population were allocated to the middle-american gene pool. Those with ancestry probability < 0.90 were considered as admixed [71]. The

clustering of the AMDP was further assessed and visualized in a 3D scatter plot using PCA in `prcomp` R 3.0 function [72].

Marker-trait association tests and linkage disequilibrium analyses

The filtered 9370 SNPs and the adjusted trait means (BLUPs) for each of the environments (drought stressed and well-watered) were used as input data in marker-trait association (MTA) analysis. The more conservative compressed mixed linear model (CMLM) procedure in the genome association and prediction integrated tool (GAPIT) (v3) program of R software was used to determine the MTAs following the $Q + K$ model according to Lipka et al. [73]. *Phaseolus vulgaris* is characterised by a strong genetic structure necessitating the need to use the $Q + K$ model [74]. The CMLM incorporated both the population structure (Q ; fixed effect) and kinship (K ; random effect) matrices as covariates. This was done to correct the population structure, increase statistical power of the analysis and minimize false positives (spurious MTAs) [67, 72, 75]. The K matrix was included in the association analysis to correct for cryptic relatedness within the AMDP [54, 67]. The threshold for significant MTA was set at $p < 0.001$ to reduce the risk of false MTAs.

The Manhattan plots drawn using the `CMplot` package in R 3.5.3 were used to visualise the significant MTAs for each environment. The p -values were plotted as $-\log_{10}(p)$ to generate the Quantile-Quantile (Q-Q) and Manhattan plots using the `CMplot` package in R package [76]. The Q-Q plots were produced from the observed and expected logarithm of the odds (LOD) scores for each trait. The GAPIT program of R software was used to deduce linkage disequilibrium. The LD Heatmap package in R 3.0 was used to generate the LD Heatmaps for the significant markers of each trait [77, 78]. Alleles with positive additive effects resulting in higher values of GYD, SZ and LCC were described as “superior alleles” under both drought stressed and well-watered conditions. On the other hand, alleles resulting in decreased GYD, SZ, and LCC were “inferior alleles”. In addition, alleles with negative effects resulting in lower values of DFW, DPM, LT and SC were considered to be “superior alleles” under drought stressed conditions.

Putative candidate gene prediction

The `Jbrowse` feature on Phytozome v13 was used to browse the *P. vulgaris* G19833 v2.1 reference genome sequence [1]. This was done to gain insight into potential putative candidate genes associated with significant SNPs for each trait. The functional annotation of the gene was checked on Phytozome v13 website (<http://phytozome.net>) to postulate the role of the gene in the control of a target trait. Plausible candidate genes were identified based on the window size of 200 kb (maximum ± 100 kb) on either side (upstream and downstream) of the significant marker [74, 79]. The window size of 200 kb is the average LD [74, 79]. A gene was considered a potential candidate using the following criteria: (i) if the gene contained a significant SNP or the gene contained a SNP that was in LD with a significant SNP [3], and (ii) if the gene had a known role related to regulating drought stress response, plant growth and development under water deficit based on gene ontology term descriptions in Phytozome v13. For the positional candidate genes that did not have adequate functional annotation information on Phytozome v13, the sequence data of the significant SNP was used against the NCBI database. This was done using the basic local alignment search tool for nucleotide (BLASTn; <https://blast.ncbi.nlm.nih.gov/blast/blast.cgi>).

Table 2. Phenotypic summary statistics, coefficient of variation and broad-sense heritability of the measured traits for all the 185 dry beans genotypes based on the best liner unbiased prediction (BLUP) value grown under drought stressed and well-watered conditions.

Traits	Treatment												AC
	Drought Stress						No stress						
	Average	SD	Range	Wald statistic (genotype)	CV (%)	H^2	Average	SD	Range	Wald statistic (genotype)	CV (%)	H^2	
DFW	43.32	7.11	32–60	139.64***	5.58	0.96	41.28	5.69	32.50–60.00	95.66***	5.32	0.98	0.94
DPM	90.97	8.40	71.50–106.00	187.89***	3.71	0.85	104.10	9.31	83.50–120.20	210.74***	2.55	0.94	0.93
GYD	715.40	457.80	39.4–2134.00	600772.00***	36.57	0.92	1016.00	555.00	55.00–2586.00	797047.00***	33.68	0.88	0.92
SW	31.39	11.61	14.25–60.00	420.66***	16.91	0.97	34.98	12.27	16.75–65.00	463.60***	12.45	0.97	0.98
PH	50.05	16.75	25.25–102.20	963.70***	28.13	0.83	56.97	18.46	28.5–125.00	1145.90***	22.53	0.92	0.88
LCC	31.12	3.80	18.17–44.15	53.51***	16.62	0.54	43.55	4.46	33.10–62.43	78.08***	16.21	0.14	0.35
LT	25.22	2.59	16.85–30.95	29.29***	9.42	0.46	19.76	1.15	17.23–24.90	4.78***	7.12	0.42	0.37
SC	96.66	13.97	59.38–141.40	760.10***	18.90	0.32	254.50	75.69	64.00–465.00	23883.00***	35.93	0.33	0.24

AC = across environments (drought stress and well-watered), SD = standard deviation of the trait means, CV = coefficient of variation, H^2 = broad-sense heritability, DFW = days to flowering, DPM = days to physiological maturity, GYD = grain yield (kg/ha), SW = 100 seed weight (g), PH = plant height (cm), LCC = leaf chlorophyll content, LT = leaf temperature ($^{\circ}$ C), SC = stomatal conductance ($\text{mmol m}^{-2} \text{s}^{-1}$),

* = $p \leq 0.05$;

** = $p \leq 0.01$ and

*** = $p \leq 0.001$.

<https://doi.org/10.1371/journal.pone.0278500.t002>

Results

Variations of agronomic and physiological traits under two water regimes

The descriptive statistics and H^2 estimates for the agronomic and physiological traits under drought stressed and well-watered environments are shown in Table 2. Residual maximum likelihood analysis revealed highly significant ($p < 0.001$) genotypic main effects on all the studied traits under both environments. This supported the use of the AMDP for GWAS purposes. Overall, phenotypic variability was observed among the genotypes for DFW, LCC, LT, SC, PH, DPM, GYD and SW under drought stressed and well-watered conditions (S1 and S2 Figs). High H^2 estimates (0.83–0.97) were observed for all the studied traits under drought stressed conditions, except for SC ($H^2 = 0.32$), LT ($H^2 = 0.46$), and LCC ($H^2 = 0.54$). Under well-watered conditions, high H^2 estimates (0.88–0.98) were observed for all the traits except for LCC ($H^2 = 0.14$), SC ($H^2 = 0.33$), and LT ($H^2 = 0.42$).

In general, the observed H^2 estimates under both environments revealed that much of the observed phenotypic variation was due to the genetic component. This further supported the suitability of the AMDP for GWAS studies. Grain yield was highest under well-watered (1016 kg/ha; $H^2 = 0.88$), and lower under drought stressed conditions (715 kg/ha; $H^2 = 0.92$). The SW also varied among the environments at 34.98 g/100 seeds ($H^2 = 0.97$), and 31.39 g/100 seeds ($H^2 = 0.97$) under well-watered and drought stressed conditions, respectively. The AMDP had a shorter duration (lower values) under drought stressed conditions (DPM = 90.97 days), compared to well-watered conditions (DPM = 104.10 days). The same trend was observed for PH, LCC, and SC. On the other hand, LT was lower (19.75 $^{\circ}$ C) under well-watered environments, compared to drought stressed environments (25.22 $^{\circ}$ C). Under drought stressed conditions, GYD ranged from 39.4 kg/ha to 2134 kg/ha. In addition, GYD

Table 3. Drought tolerance indices and predicted genotype values for grain yield (across environments) of top 20 drought tolerant genotypes.

Genotype	Gene pool	GYD (kg/ha)	DSI	GMP	DTI	%GYR	Mean rank
G184	Andean	2222.70	0.16	2205.30	4.74	4.69	25.50
G176	Andean	2097.50	0.35	2084.00	4.25	10.60	29.90
G147	Andean	2080.10	1.26	1995.60	4.03	37.83	66.30
G146	Andean	2067.40	1.18	1994.50	4.02	35.49	61.50
G158	Admixed	2017.10	-0.01	1979.50	3.82	-0.26	24.50
G135	Andean	1968.70	-0.37	1956.90	3.75	-11.17	19.80
G101	Andean	1964.80	1.32	1890.60	3.78	39.72	69.30
G138	Andean	1846.80	0.22	1826.10	3.25	6.50	31.00
G180	Andean	1838.90	0.09	1789.60	3.13	2.57	29.50
G162	Andean	1828.70	0.57	1814.30	3.20	16.97	40.50
G124	Andean	1815.00	-0.03	1805.50	3.19	-0.78	26.00
G173	Andean	1792.60	0.68	1780.40	3.07	20.40	46.00
G115	Andean	1788.90	0.40	1765.40	3.04	11.87	35.50
G150	Andean	1758.80	0.40	1750.80	2.98	12.03	36.50
G127	Andean	1750.00	-0.02	1733.00	2.94	-0.73	29.00
G159	Andean	1743.30	1.12	1694.70	2.90	33.67	65.30
G113	Andean	1683.80	1.76	1548.00	2.35	52.77	91.50
G125	Andean	1628.20	-0.25	1590.20	2.49	-7.39	26.80
G181	Andean	1614.10	1.03	1585.20	2.44	31.00	64.50
G104	Andean	1608.80	0.58	1601.10	2.49	17.37	45.30

GYD = grain yield, DSI = drought susceptibility index, GMP = geometric mean productivity, DTI = drought tolerance index and %GYR = percent grain yield reduction. Note: Mean Rank is the mean rank of a genotype across all the drought tolerance indices. Admixed includes genotypes that are 10 to 90% andean or middle-american according to the structure analysis results.

<https://doi.org/10.1371/journal.pone.0278500.t003>

exhibited a narrower range than in well-watered conditions where it ranged from 55.0 kg/ha to 2586.0 kg/ha. The coefficient of variation (CV) ranged from 5.32 to 5.58%, 2.55 to 3.71%, 33.68 to 36.57%, 12.45 to 16.91%, 22.53 to 28.13%, 16.21 to 16.62%, 7.12 to 9.42%, and 18.90 to 35.93% for DF, DPM, GYD, SW, PH, LCC, LT, and SC, respectively. Low standard deviations (SD) were observed for LT and LCC under both environments.

Combined GYD data over two seasons across environments revealed that the highest yielding genotype was G184 (DAB91–2222,7 kg/ha) followed by G176 (DAB302–2097.5 kg/ha) and G147 (CIM-SUG07-ALS-S1-3–2080,1 kg/ha) (Table 3).

The drought tolerance indices for the 185 genotypes based on mean GYD are summarised in Table 3 (top 20 drought tolerant genotypes) and S2 Table (all study genotypes). The severity of drought stress at SVES across the 2 seasons of evaluation was moderate (DII of 0.30). Among the evaluated genotypes, G158 (SWEET WILLIAM/DAB287), G135 (DAB539), G124 (DAB487), G127 (CIM-SUG07-ALS-2), G125 (CIM-RM09-ALS-BSM-12), G138 (CZ104-72) and G184 are some of the genotypes that were less sensitive to drought stress. This is based on their low DSI, %GYR and overall mean ranks across the indices. These genotypes had DSI values ranging from -0.37 (G135) to 0.16 (G184) and %GYR ranging from -11.17 (G135) to 4.69 (G184). In summary, all the top 20 drought tolerant genotypes were members of the andean gene pool, except for G158 which is an admixture (Table 3).

Population structure analysis

The STRUCTURE analysis results and Evanno test (ΔK) revealed the presence of two major sub-populations (highest ΔK value occurred at $K = 2$) within the AMDP of dry beans (Fig 1A

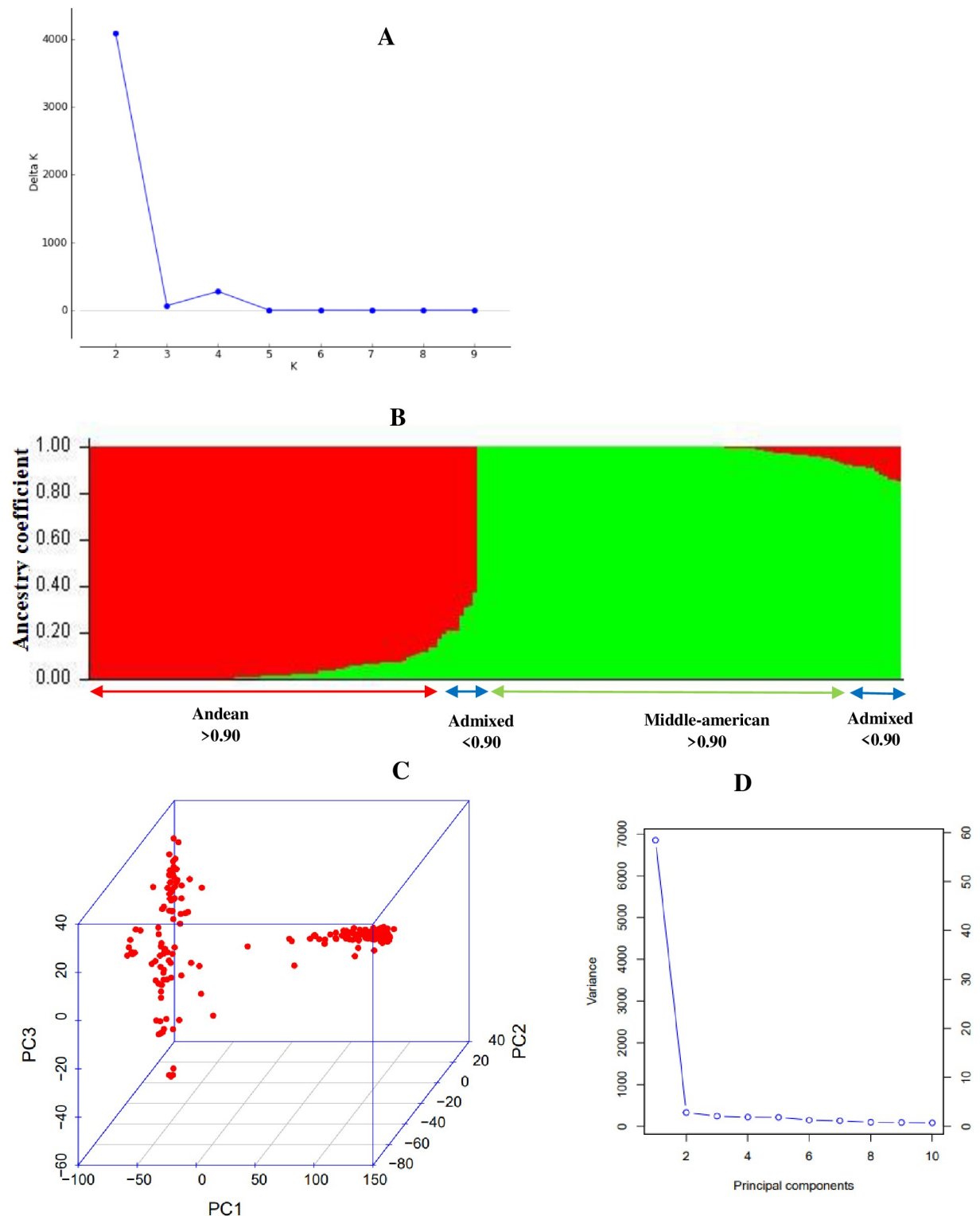


Fig 1. Population structure of 185 andean and middle-american diversity panel (AMDP) from different models. Note: A = The ΔK determined by the Evanno method showing the stratification of the 185 AMDP into two main sub-populations. The cluster with the largest ΔK ($K = 2$) was used to determine the number of sub-populations in the AMDP of dry beans and the existence of two-sub-populations was inferred; B = Population structure of 185 AMDP of dry beans genotypes based on 4095 SNP markers ($K = 2$ gives the best separation) as determined from STRUCTURE analysis. Red and green represents andean and middle-american sub-populations, respectively; C = Three dimensional principal component analysis (PCA) scatter plot illustrating the population structure of 185 AMDP of dry beans genotypes based on 9370 SNP markers; D = Screen plot showing the percentage of variation explained by the different principal components. The top 10 principal components (PCs) as recommended by Price et al. [72] were selected for population structure to adjust for population stratification.

<https://doi.org/10.1371/journal.pone.0278500.g001>

and 1B). The two sub-populations correspond to the andean and middle-american domesticated gene pools. The minimum ancestry or membership coefficient to a particular cluster was 0.63 (Fig 1B and S1 Table). Most of the genotypes (90) clustered within the middle-american gene pool (Fig 1B). Seventy-six genotypes clustered within the andean gene pool (Fig 1B and S1 Table). On the other hand, 19 were andean-middle american admixed genotypes of the two gene pools (10 to 90% andean or middle-american). The admixed genotypes included SMC16, SMC21, NUA674, NUA59-4, G75, DAB115, DAB63, DAB142, DAB477, CIM-RM02-36-1, CIM-RM09-ALS-BSM-11, CIM-RM02-134-1, Sweet William, ZABRA16575-60F22, GLP585/MLB49-89A-3, RWR2154, SAB792, NAVY LINE 22, and CIM-SUG07-ALS-S1-3 (S1 Table).

The genetic structure result of the AMDP was verified with the PCA based on SNP marker data and is illustrated by a 3D scatter plot (Fig 1C). The first principal component (PC) accounted for more than 55% of the observed genotypic variability in the AMDP (Fig 1D). On the other hand, the second and third PCs separately accounted for less than 5% of the overall genetic variance in the AMDP (Fig 1D). The PCA also divided the genotypes into two distinct clusters (andean and middle-american sub-populations) as were found with STRUCTURE output (Fig 1C). Furthermore, the andean-middle american admixed genotypes (positioned between the two groups) were isolated from the andean and middle-american sub-groups by PCA (Fig 1C).

Analysis of marker-trait associations under drought stressed conditions

The significant MTAs and their respective statistical parameters for agronomic and physiological traits are summarised in Table 4. In this study, the threshold for significant MTA was set at $p < 0.001$ to reduce the risk of false MTAs. Under drought stressed conditions, 29 significant MTAs were identified for six traits (excluding DPM and LCC) with $p < 10^{-03}$. The associations are shown in Fig 2. The quantile-quantile (QQ) plots for the studied traits revealed that the expected and observed probability values were normally distributed (S3 Fig). The highest number of significant MTAs were observed on *P. vulgaris* (*Pv*) chromosome *Pv11* (28%), followed by *Pv8* (17%), with the least on chromosomes *Pv6* and *Pv4*, both with 3%. No significant associations for DPM and LCC were identified under drought stressed conditions in this study. The highest number of significant MTAs were identified for PH (15), and the SNPs were distributed across six different chromosomes (*Pv1*, *Pv5*, *Pv7*, *Pv8*, *Pv10*, and *Pv11*). Additionally, the allele effect of these SNPs ranged from -16.03 cm (SNP 8198531) to 17.82 cm (SNP 100101387).

Four SNPs (SNPs 2362591, 2362591, 45231105, and 40802478) that have a significant association with GYD were also identified. These were located on chromosomes *Pv4* and *Pv11*, with allele effect ranging from -174.56 kg/ha (SNP 3381526) to 202.90 kg/ha (SNP 3382688). Notably, 75% of the SNPs that were significantly associated with GYD were located on chromosome *Pv11*. The sum of the SNPs with a significant positive effect on GYD was 341,88 kg/ha and -351,23 kg/ha for all the SNPs with a significant negative effect on GYD (Table 4). For SW, two SNPs that were significantly associated with this trait were identified on chromosomes *Pv03* and *Pv08*. These two SNPs had allelic effects ranging from -2.41 g per 100 seeds (SNP 3383047) to 4.46 g per 100 seeds (SNP 16647170). Regarding physiological traits, SNPs were identified that have a significant association with LT. The identified SNPs were distributed across two chromosomes (*Pv6* and *Pv8*), with allele effect ranging from -1.23°C (SNP 100065202) to 1.34°C (SNP 100106140).

Notably, two SNPs on chromosomes *Pv1* and *Pv2* were significantly associated with SC. These two SNPs had allele effect ranging from -10.79 mmol m⁻² s⁻¹ (SNP 3380850) to -10.33 mmol m⁻² s⁻¹ (SNP 3381030). Common regions associated with multiple traits on

Table 4. Single nucleotide polymorphism (SNP) markers associated with agronomic and physiological traits in dry beans genotypes under drought stress conditions.

Phenotype	SNP name	CH	SNP position on genome (bp)	MAF	Allele	Effect of allele	-log ₁₀ (P) value	R ²	Candidate gene
LT	100106140	06	14389438	0.25	A/C	1.34	0.000	0.23	
	100065202	08	52504423	0.12	G/A	-1.43	0.000	0.22	
DFW	100132383	03	47240686	0.04	A/G	3.76	0.000	0.70	
	3381050	02	25978891	0.03	C/T	3.85	0.000	0.70	Phvul.002G122100
	8204238	10	42089084	0.06	A/G	2.78	0.000	0.70	
	8212194	10	42105474	0.04	A/T	3.54	0.000	0.69	
GYD	3384334	11	2362591	0.10	A/G	-176.67	0.000	0.44	
	3381526	11	2362591	0.09	A/G	-174.56	0.000	0.44	
	3382688	04	45231105	0.09	G/A	202.90	0.000	0.43	Phvul.004G150500
	100061855	11	40802478	0.42	T/G	138.98	0.000	0.43	
PH	100101387	05	34925013	0.03	G/A	17.82	0.000	0.32	
	8198531	07	9701750	0.04	G/A	-16.03	0.000	0.31	
	100060987	01	42938094	0.04	G/A	15.57	0.000	0.31	Phvul.001G172300
	100181735	08	22152034	0.23	G/A	7.57	0.000	0.30	Phvul.008G133100
	3379684	07	5239949	0.22	T/A	-5.62	0.000	0.30	
	16650827	07	51719432	0.09	T/C	-8.41	0.000	0.30	
	3380814	11	11934462	0.11	C/T	7.15	0.000	0.30	
	100119463	08	6003908	0.03	G/A	15.56	0.000	0.30	Phvul.008G065700
	8196298	11	12212674	0.12	T/G	-0.67	0.000	0.30	
	3379078	08	7823952	0.02	C/G	17.75	0.000	0.30	Phvul.008G080600
	3377272	11	9410740	0.16	T/C	-6.15	0.000	0.29	
	3379350	11	43494132	0.16	C/T	6.35	0.000	0.29	
	100063156	10	7307165	0.44	T/C	4.85	0.000	0.29	
	3379405	05	4782514	0.24	G/A	-7.80	0.000	0.29	
	3377900	11	9691109	0.14	T/A	-6.02	0.000	0.29	
SW	16647170	08	36620996	0.11	T/C	4.46	0.000	0.66	
	3383047	03	50229319	0.33	G/A	-2.41	0.000	0.65	Phvul.003G263200
SC	3380850	01	50427390	0.08	T/C	-10.79	0.000	0.10	Phvul.001G254100
	3381030	02	33669423	0.04	G/A	-10.33	0.000	0.08	

CH = chromosome, DFW = days to flowering, GYD = grain yield (kg/ha), SW = 100 seed weight (g), PH = plant height (cm), LT = leaf temperature (°C), SC = stomatal conductance (mmol m⁻² s⁻¹), SNP = single nucleotide polymorphism, MAF = minor allele frequency, R² = proportion of the total phenotypic variation explained by the significant SNP marker after fitting the other model effects and -log₁₀(P) = p value of the association.

<https://doi.org/10.1371/journal.pone.0278500.t004>

chromosomes were not identified under drought stressed environments in this study. Markers explained 0.08–0.10, 0.22–0.23, 0.29–0.32, 0.43–0.44, 0.65–0.66 and 0.69–0.70 of the total phenotypic variability (R²) for SC, LT, PH, GYD, SW and DFW, respectively. Overall, the R² varied from 0.08 (SC: SNP 3381030) to 0.70 (DFW: SNPs 100132383, 3381050 and 8204238).

Analysis of marker-trait associations under well-watered environments

The significant MTAs and their respective statistical parameters for agronomic and physiological traits are summarised in Table 5. Under well-watered conditions, 39 significant MTAs were detected for six traits (excluding SW and SC) with p < 10⁻⁰³. The associations are shown in Fig 3.

The quantile-quantile (QQ) plots for the studied traits revealed that the expected and observed probability values were normally distributed (S4 Fig). The highest number of

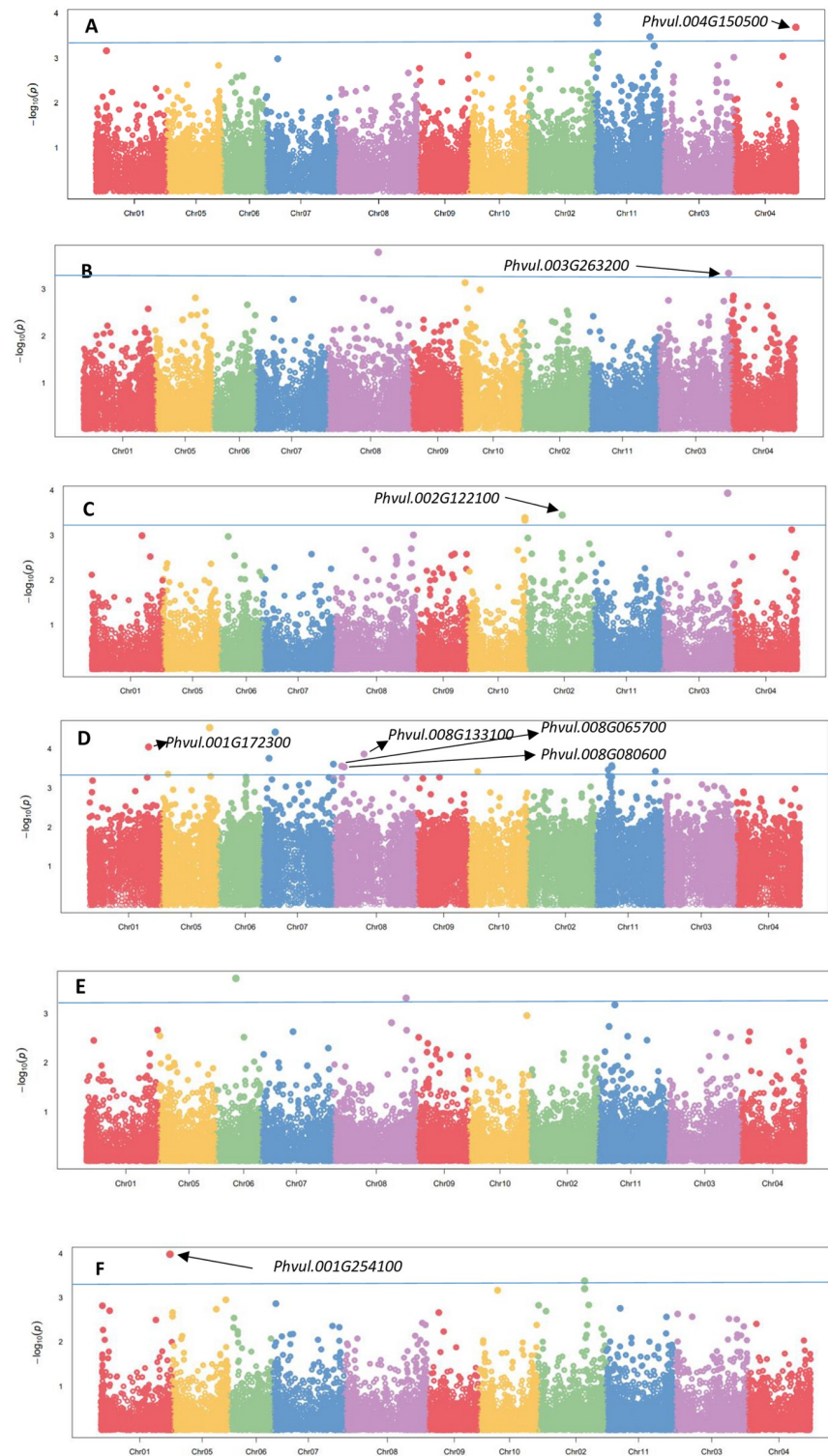


Fig 2. Manhattan plots indicating the significant marker-trait associations, their p-values and candidate genes for agronomic and physiological traits in 185 dry beans genotypes evaluated under drought stressed conditions. Note: A = Grain yield, B = Seed size, C = Days to 50% flowering, D = Plant height, E = Leaf temperature, F = Stomatal conductance. *Chr represents Chromosome, x-axis represents the physical map locations of the SNPs on each chromosome and the y-axis ($-\log_{10}(p)$) represents the degree to which a SNP is associated with a trait. The blue horizontal significant line represents the degree to which a SNP is associated with a trait. The blue horizontal significant line represents the degree to which a SNP is associated with a trait. The blue horizontal significant line represents the degree to which a SNP is associated with a trait. The blue horizontal significant line represents the degree to which a SNP is associated with a trait.

<https://doi.org/10.1371/journal.pone.0278500.g002>

Table 5. Single nucleotide polymorphism (SNP) markers associated with agronomic and physiological traits in dry beans genotypes under well-watered conditions.

Phenotype	SNP name	CH	SNP position on genome (bp)	MAF	Allele	Effect of allele	-log ₁₀ (P) value	R ²	Candidate gene
DFW	3372129	04	43770691	0.20	C/T	1.85	0.000	0.46	
	3368616	01	48386869	0.29	C/G	2.19	0.000	0.46	
	8212932	04	43742237	0.35	C/A	-1.39	0.000	0.46	
	3379964	03	48424846	0.37	C/T	-1.61	0.000	0.46	
	100175933	06	31464277	0.27	A/G	-2.27	0.000	0.46	
	100175934	06	31464277	0.27	A/T	2.23	0.000	0.45	
	16647096	03	19481003	0.28	A/C	1.49	0.000	0.45	
	3378741	03	1178534	0.38	A/C	1.56	0.000	0.45	Phvul.003G011400
	100140152	04	43939513	0.32	A/G	1.67	0.000	0.45	Phvul.004G037700
	3374827	11	47036209	0.34	T/G	1.63	0.000	0.45	Phvul.011G166300
	3381380	05	1315962	0.27	T/C	-2.05	0.000	0.45	
100122216	07	23590138	0.39	A/T	1.79	0.000	0.45	Phvul.007G144000	
DPM	100117381	02	24161867	0.18	A/T	2.90	0.000	0.70	Phvul.002G112700
GYD	100124606	01	32783904	0.17	T/A	199.11	0.000	0.50	
LCC	8198945	06	30370228	0.16	T/C	2.18	0.000	0.12	Phvul.006G209700
	100167635	08	44516286	0.32	T/G	1.90	0.000	0.11	Phvul.008G163600
PH	3383709	11	23343020	0.28	A/G	7.30	0.000	0.41	
	100123206	03	41669536	0.27	G/T	9.02	0.000	0.40	Phvul.003G192800
	13121517	11	5699564	0.08	C/T	9.30	0.000	0.40	
	100164602	03	32040779	0.43	A/C	-4.72	0.000	0.40	
	100065600	11	38863980	0.27	C/G	-8.10	0.000	0.40	
	100181804	07	37529193	0.42	G/T	-4.77	0.000	0.40	Phvul.007G253400
	100124008	03	36956076	0.38	T/C	5.98	0.000	0.40	
	100101486	04	38236692	0.37	T/C	4.92	0.000	0.39	
	100073620	11	42969050	0.35	C/T	-7.00	0.000	0.39	Phvul.011G152000
	100068647	01	39765027	0.38	T/G	5.76	0.000	0.39	
	3382850	10	40091053	0.39	A/G	4.52	0.000	0.39	
	13121517	11	5699564	0.06	T/C	-10.46	0.000	0.39	
	3379157	06	30312046	0.44	T/C	-4.39	0.000	0.39	Phvul.006G208800
13121469	01	44975217	0.49	C/A	5.29	0.000	0.39	Phvul.001G190800	
LT	100101691	03	18922335	0.28	G/A	-0.56	0.000	0.08	
	100070187	04	12643816	0.21	A/G	0.80	0.000	0.15	
	100061661	01	19177470	0.16	T/A	0.69	0.000	0.09	
	100071816	04	33722284	0.45	A/G	0.39	0.000	0.08	
	100100644	08	26794110	0.21	A/C	0.54	0.000	0.08	
	100102687	04	7507744	0.14	G/A	-0.61	0.000	0.08	Phvul.004G055500
	100120897	08	20306142	0.06	C/A	-0.71	0.000	0.08	
	100167520	05	23720983	0.36	C/G	0.46	0.000	0.08	
100161682	05	18689401	0.17	G/A	-0.58	0.000	0.08	Phvul.005G077500	

CH = chromosome, DFW = days to flowering, DPM = days to physiological maturity, GYD = grain yield (kg/ha), PH = plant height (cm), LCC = leaf chlorophyll content, LT = leaf temperature (°C), SNP = single nucleotide polymorphism, MAF = minor allele frequency, R² = proportion of the total phenotypic variation explained by the significant SNP marker after fitting the other model effects, -log₁₀(P) = p value of the association.

<https://doi.org/10.1371/journal.pone.0278500.t005>

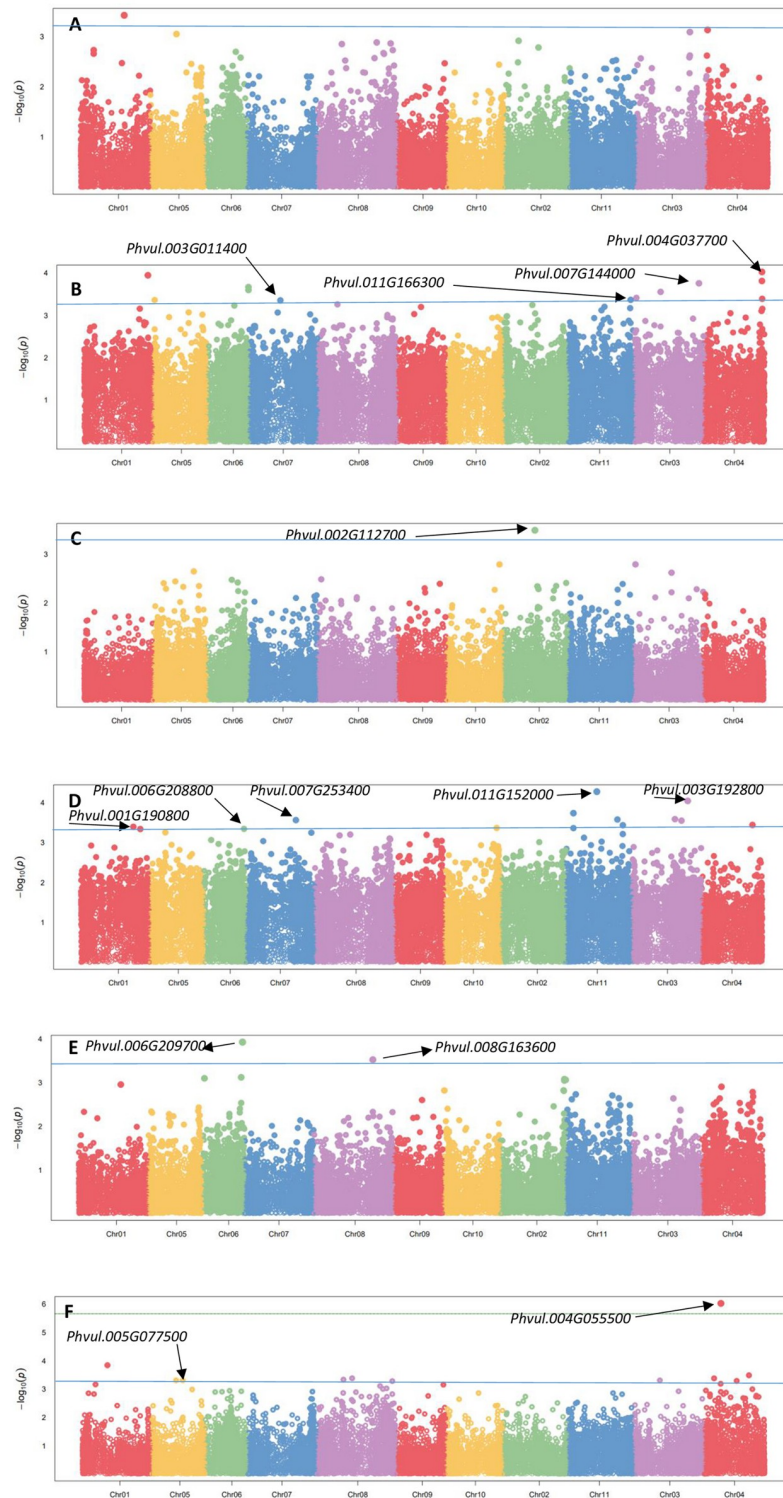


Fig 3. Manhattan plots showing significant marker-trait associations, their p-values and candidate genes for agronomic and physiological traits under well-watered conditions. Note: A = Days to 50% flowering, B = Grain Yield, C = Days to physiological maturity, D = Plant height, E = Leaf chlorophyll content, F = Leaf temperature. *Chr represents Chromosome, x-axis represents the physical map locations of the SNPs and the y-axis ($-\log_{10}$ p-values) represents the degree to which a SNP is associated with a trait. The blue horizontal significant line represents FDR adjusted $p < 0.001$.

<https://doi.org/10.1371/journal.pone.0278500.g003>

significant MTAs were observed on *Pv11* (15%), followed by chromosomes *Pv3* and *Pv4* (both with 18%), with the least on *Pv2* and *Pv10* (both with 3%). No significant markers for SW and SC were detected under well-watered conditions in this study. The highest number of significant MTAs were observed on PH (14), with markers accounting for 0.39–0.40 of the total trait variation. Additionally, the allele effect of these SNPs ranged from -10.46 cm (SNP 13121517) to 9.30 cm (SNP 13121517). Interestingly, 38% of the markers that were significantly associated with PH were located on chromosome 11. For DFW, a total of 12 significant associations were identified, with markers explaining 0.45–0.46 of the observed trait variation. Additionally, the significant SNPs for DFW were located on chromosomes *Pv1*, *Pv3*, *Pv4*, *Pv5*, *Pv6*, *Pv7* and *Pv11*. These SNPs had allele effect ranging from -2.27 days (SNP 100175933) to 2.23 days (SNP 100175934).

Notably, one SNP (SNP 100124606) on chromosome *Pv01* was significantly associated with GYD. This SNP had a large positive allelic effect of 199.11 kg/ha. In addition, this SNP had a MAF of 0.17 in the population. Regarding physiological traits, SNPs were identified that have a significant association with LCC. The identified SNPs were distributed across two chromosomes (*Pv6* and *Pv8*). Furthermore these SNPs had positive allele effects ranging from 1.90 (SNP 100167635) to 2.18 (SNP 8198945). As for LT, nine significant associations were detected, with markers accounting for 0.08–0.15 of the trait variation. The significant SNPs for LT were located on chromosomes *Pv1*, *Pv3*, *Pv4*, *Pv5* and *Pv8*. These SNPs had allele effect ranging from -0.71°C (SNP 100102687) to 0.80°C (SNP 100070187). Additionally, the sum of the SNPs with a significant positive effect on LT was 2.88°C and -2.46°C for all the SNPs with a significant negative effect. A locus (SNP 100117381) on chromosome *Pv02* explained the highest proportion of the phenotypic variation (0.70) among the studied traits and was associated with DPM. In addition, SNP 100117381 had a MAF of 0.18 in the population and a large positive effect (2.90 days) on DPM. On the other hand, nine significant SNPs for LT on chromosomes *Pv3*, *Pv4*, *Pv8* and *Pv5* explained the least proportion of the observed phenotypic variation (0.08) among the studied traits. Common regions associated with multiple traits on chromosomes were not identified under well-watered environments. Overall, R^2 varied from 0.08 (LT-SNPs 100101691, 100071816, 100100644, 100102687, 100102687, 100167520 and 100161682) to 0.70 (DPM—SNP 100117381) (Table 5).

Identification of putative candidate genes associated with significant single nucleotide polymorphism

Drought stressed environments. A total of eight potential candidate genes (DFW—1; GYD—1; PH—4; SW—1; SC—1) were identified under drought stressed environments (Table 4 and Fig 2). The candidate genes for DFW (*Phvul.002G122100*), SC (*Phvul.001G254100*), SW (*Phvul.003G263200*) and GYD (*Phvul.004G150500*) were identified on chromosomes *Pv02*, *Pv01*, *Pv03* and *Pv04*, respectively (Table 4). These genes had diverse putative functions ranging from RNA recognition motif or RNP domain functions (DFW), NADPH dehydrogenase/NADPH diaphosare activity (SW), helicase activity and CCCH zinc finger protein domain functions (SC) to Phosphoethanolamine N-methyltransferase activity (GYD), respectively. On the other hand, the candidate genes for PH were identified on chromosomes *Pv01* (*Phvul.001G172300*) and *Pv08* (*Phvul.008G133100*; *Phvul.008G065700*; *Phvul.008G080600*) (Table 4). These genes had diverse putative functions ranging from calcium transporting ATPase 1 activity, peptidyl prolyl cis trans isomerase activity, acyl-coenzyme A thioesterase activity to centrosomal protein nuf function, respectively.

Well-watered environments. A total of fourteen potential candidate genes (DFW—4; DPM—1; LCC—2; PH—5; LT—2) were identified under well-watered environments (Table 5

and Fig 3). The candidate genes for DFW were identified on chromosomes *Pv03* (*Phvul.003G011400*), *Pv04* (*Phvul.004G037700*), *Pv07* (*Phvul.007G144000*) and *Pv11* (*Phvul.0011G166300*) (Table 5). On the other hand, the candidate gene for DPM was identified on chromosome *Pv02* (*Phvul.002G112700*) (Table 5). Candidate genes for DFW had diverse putative functions related to SORTING NEXIN-13, transcription factor TCP 13, U6 SNRNA-associated SM LIKE PROTEIN LSM4 and NHL domain containing protein. On the other hand, the candidate gene for DPM had a putative function related to the activity of thiol disulphide oxidoreductase. Chromosomes *Pv4* and *Pv5* harboured the two candidate genes for LT namely *Phvul.004G055500* and *Phvul.005G077500*, respectively (Table 5). These genes had diverse putative functions related to the mitochondrial transcription termination factor family protein and leucine rich repeat protein associated with apoptosis in muscle tissue, respectively.

The genes *Phvul.006G209700* and *Phvul.008G163600* for LCC were identified on chromosomes *Pv06* and *Pv08*, respectively. These genes had diverse putative functions, such as pre-naspirodiene oxygenase or hyoscyamus muticus prenaspirodiene oxygenase activity and nucleoside triphosphate hydrolases activity, respectively. On the other hand, the candidate genes for PH were identified on chromosomes *Pv01* (*Phvul.001G190800*), *Pv03* (*Phvul.003G192800*), *Pv06* (*Phvul.006G208800*), *Pv07* (*Phvul.007G253400*), and *Pv11* (*Phvul.011G152000*) (Table 5). These genes also had diverse putative functions, such as f-box-like domain superfamily functions, protein NRT1 or PTR family related functions, phosphatidylserine decarboxylase activity, tyra-like translation elongation factor svrs-related functions, and inactive g-type lectin s-receptor like serine or threonine protein kinase activity, respectively.

Linkage disequilibrium analysis using significant SNP markers

The analysis of LD using SNP markers is shown in Fig 4. A high and extensive LD was observed for the common bean genome. This is expected in self-pollinated crops such as

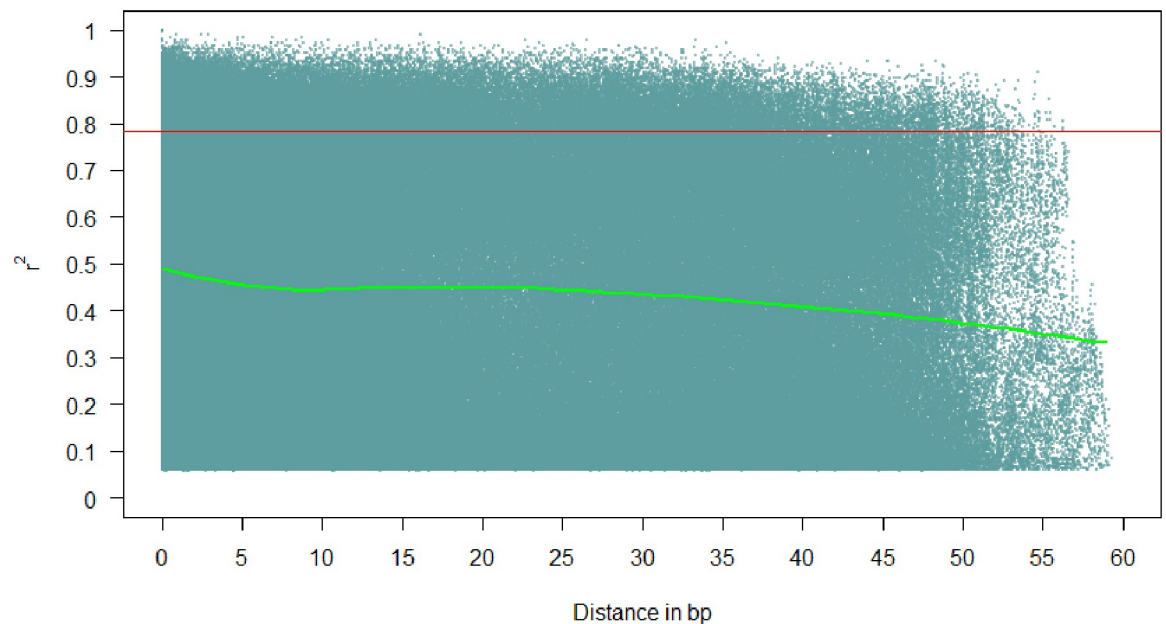


Fig 4. Linkage disequilibrium (LD, r^2) decay plot in genome of dry beans based on 9370 single nucleotide polymorphisms (SNPs) in 185 diverse genotypes.

<https://doi.org/10.1371/journal.pone.0278500.g004>

common bean. The results show that the overall LD decay across the genome of 185 common bean genotypes was 30 bp, at a cut-off of $r^2 = 0.4$. Generally, there was a slow decay of LD throughout the common bean genome. The LD extended to several mega-bases as shown in Fig 4. The population structure usually affects the extent of LD decay.

Discussion

Terminal drought stress adaptability analysis

Breeding for enhanced grain yield under both drought stressed and well-watered environments is one of the greatest challenges faced by dry beans breeders [15]. The observed moderate (0.30) DII suggests that drought intensity also depends on the type and genetic diversity of the genotypes used. Similarly, Darkwa et al. [22] reported a DII of 0.30 in different market classes of dry beans. Genotypes such as DAB91, DAB302, AFR703, CIM-SUG07-ALS-51-3, DAB487, DAB287, CIM-RM09-ALS-BSM-12 and DAB539 which exhibited high GMP values might have combined the two mechanisms that contribute to superior grain yield under both study environments. In Sub-Saharan Africa (SSA), breeders prefer genotypes that perform well under both drought stressed and well-watered environments (high GMP). This is because rainfall patterns in SSA differ by season. Furthermore, 24% of the genotypes had negative %GYR and DSI values suggesting that they had higher grain yield under drought stressed compared to well-watered conditions. This indicates that the mechanisms that contribute to superior grain yield performance under drought stressed and well-watered conditions differ. Thus, it might be worthwhile to determine the genetic basis of superior grain yield under drought stressed compared to well-watered environments in these 44 genotypes that distinguished them from the rest. Similar findings were reported by Darkwa et al. [22] in dry beans drought screening experiments.

Terminal drought stress is an important factor limiting common bean productivity in the SSA region. Therefore, the identification and subsequent release of drought tolerant genotypes will positively impact on socio-economic, food and nutrition security in SSA. These genotypes could also serve as important genetic resources in drought tolerance breeding programmes to improve released cultivars. Both DAB287 and AFR703 were released in Zimbabwe as Sweet William and Gloxinia, respectively. Among the genotypes that were less sensitive to drought stress based on their low DSI, %GYR and overall mean ranks across the indices, most of the top 20 genotypes were of the andean gene pool, coded as drought andean (DAB lines) (Table 3 and S2 Table). Notably, all the DAB lines evaluated in this study were developed for improved tolerance to drought by the International Centre for Tropical Agriculture in Colombia. The current observation suggests that progress in improving drought tolerance in the middle-american gene pool has been limited compared to the andean gene pool. This is in agreement with Assefa et al. [81] who reported that progress in improving drought tolerance in navy beans worldwide has been limited compared to the other commercial classes of small seeded middle-american beans.

Variations in agronomic and physiological traits

The low to moderate H^2 estimates observed for SC, LT and LCC under drought stressed and well-watered conditions imply that these physiological traits might be influenced by a number of genes (polygenic inheritance). Therefore, direct selection for SC, LT and LCC under drought stressed and well-watered conditions could be a challenge to dry beans breeders. On the other hand, the high H^2 estimates (97%) for seed size observed under drought stressed and well-watered environments reflect the predominance of additive gene action (genetic control of this trait) across environments. The current findings are in agreement with Assefa et al. [80]

and Hoyos-Villegas et al. [14] who reported H^2 estimates of 77 and 93.4% for seed size, respectively under well-watered conditions.

In this study, drought stress reduced PH, GYD, SW, DPM, LCC and SC by 12.1, 29.6, 10.3, 12.6, 28.5 and 62.0%, respectively, highlighting the detrimental effect of drought stress under field conditions. These findings corroborate previous reports by Assefa et al. [80], Darkwa et al. [22], Assefa et al. [81], and Mathobo et al. [82] in common bean. Mathobo et al. [82] reported reductions of 48 and 39% in SC and LCC, respectively under drought stressed conditions. Darkwa et al. [22], using navy beans, reported reductions of 10.7, 14.8, 12.7 and 26.1% in SW, PH, DPM and LCC under drought stressed conditions. Assefa et al. [80], using navy beans, also reported reductions of 12% and 17.6% in SW and DPM, respectively under drought stressed conditions.

Crop plants close their stomata when exposed to drought stress to minimize excessive water loss (avoid dehydration) through transpiration. However, the closing of stomata reduces stomatal conductance. Furthermore it also affects the cooling mechanisms resulting in increased leaf or canopy temperature. Therefore, in this study, drought stress increased LT by 21.6%. Drought stress also reduced GYD by 30%, close to the GYD reductions reported by Schneider et al. [83] [26%], Darkwa et al. [22] [30%] and Mutari et al. [24] [28%] in dry beans drought tolerance screening trials.

Population structure and linkage disequilibrium analysis

The AMDP was delineated into two distinct major sub-populations based on the genotypes' genetic ancestry, and this corresponded to the andean and middle-american gene pools (Fig 1B and 1C). This is expected considering that the domestication of dry beans on the American continent in two main centres of origin (andean and middle-american regions) resulted in two major and diverse gene pools [59, 84]. Cichy et al. [55], Raggi et al. [74], Tigist et al. [48], Nkhata et al. [49], Ojwang et al. [71], and Liu et al. [85] also observed two sub-populations (andean and middle-american gene pools) in their GWAS studies.

A number of the identified andean-middle american admixed genotypes carrying genomic regions from both gene pools are released cultivars. These have been released in Rwanda (RWR2154), Malawi (NUA59-4), Zimbabwe (SMC16, NUA674, and Sweet William) and Eswatini (NUA674) [86–89]. Furthermore, most of the admixed genotypes have commercial seed types, are biofortified (RWR2154, SMC16, SMC21, NUA674 and NUA59-4) and drought tolerant (Sweet William, DAB115, DAB63, DAB142 and DAB477). Singh [90], Beebe et al. [20] and Beebe [84] reported that interracial hybridizations between races or sister species (*Phaseolus coccineus*, *Phaseolus acutifolius* and *Phaseolus dumosus*) of *Phaseolus vulgaris* have been widely used in dry beans improvement programmes. These have been widely used when breeding for enhanced grain yield, micronutrient density and drought tolerance. For example, the biofortified admixed genotype NUA674 is a product of an inter-gene pool cross between AND277 (andean gene pool) and G21242 (andean-middle-american inter-gene pool landrace) made at the International Centre of Tropical Agriculture in Colombia [87]. Islam et al. [91] and Beebe [84], also reported that one of the parents to NUA674, G21242 is a product of andean–middle-american inter-gene pool hybridization, validating the current findings. In addition, G21242 is commonly used as a source of high seed iron in biofortification breeding programmes [84, 91]. The current observations suggests that most of the admixed genotypes identified in this study resulted from deliberate breeding efforts (inter-gene pool hybridizations) to introgress genes for enhanced grain yield, drought tolerance and micronutrient density. Similar findings were reported by Hoyos-Villegas et al. [14] and Tigist et al. [48] in common bean.

The biofortified and drought tolerant admixed genotypes identified in this study may be used as a bridge to transfer favourable alleles for micronutrient density and drought tolerance into either the andean or middle-american seed types. The extent and structure of LD decay in the study germplasm usually determines the resolution of GWAS. The slow decay of LD observed in this study is expected in self-pollinating crop species, such as common bean. The slow decay is caused by the loss of recombination, which results in a homozygous genetic background. According to Vos et al. [92], recombination events in crops with a homozygous genetic background are ineffective to cause LD decay, resulting in extended (large) and slow decay of LD. The slow decay of LD, and the large extent of LD observed in this study corroborates previous reports in dry beans [32, 85].

Marker-trait associations

In dry beans, it is important to enhance drought stress tolerance by identifying genotypes with high grain yield potential under water deficit conditions, and by introgressing desirable alleles conferring drought tolerance. The mean call rate (93%) and reproducibility (100%) of the silico DArTs used in this study were consistent with previous reports [15, 49]. This demonstrates the reliability and high quality of this set of silico DArTs. A higher number of significant MTAs were detected under well-watered conditions, corroborating previous reports in bread wheat (*Triticum aestivum* L.) [93, 94] and dry beans [15]. The observed trend could be due to the fact that drought tolerance is a complex polygenic trait which is highly influenced by the production environment. This results in unpredictable performance of genotypes (genotype-by-environment interaction [GEI]) under different environments (drought stressed and well-watered). Even though a smaller number of significant MTA was observed under drought stressed conditions compared to the well-watered conditions, novel genomic regions associated with key agronomic and physiological traits were detected under drought stressed conditions. Notably, no significant SNPs for all the studied agronomic and physiological traits were consistent across drought stressed and well-watered treatments. Similar findings were reported in wheat ([93]–plant height and spike length) and dry beans ([15]–grain yield) under drought stressed and well-watered treatments. The observed trend suggests that some markers may influence the expression of phenotypic traits differently under drought stressed and well-watered environments. Furthermore, the GEI could have confounded the identification of significant SNPs that are consistent across drought stressed and well-watered treatments.

The highest number of significant SNPs were identified for PH. Similar findings were reported by Sukumaran et al. [95] who observed 30 significant MTAs for PH in durum wheat (*Triticum turgidum* L. ssp. *Durum*). Some of the SNPs identified in this study were located on genomic regions that had been previously reported to be harbouring genes and QTLs for the studied traits. For example, in this study, chromosomes *Pv01*, *Pv03*, *Pv04*, *Pv06* and *Pv07* harboured 1 SNP, 4 SNPs, 3 SNPs, 2 SNPs and 1 SNP, respectively that were significantly associated with DFW under optimal conditions. These results are consistent with Dramadri et al. [34] and Nkhata et al. [49]. Dramadri et al. [34] identified 2 QTLs that were associated with DFW on *Pv03* under drought stressed and well-watered conditions. Nkhata et al. [49] identified 2 and 5 SNPs that were significantly associated with DFW on *Pv03* and *Pv06*, respectively under well-watered conditions. Furthermore, Keller et al. [6] identified 6 SNPs, 1 SNP and 1 SNP that were significantly associated with DFW on *Pv01*, *Pv04* and *Pv07*, respectively under optimal conditions. These findings suggest that the aforementioned QTL regions are stable across different environments and genetic backgrounds. In addition, these findings also suggest that chromosomes *Pv01*, *Pv03*, *Pv04*, *Pv06* and *Pv07* harbour genes for controlling flowering.

In this study, only one marker (SNP 1667170) was significantly associated with SW on chromosome *Pv08* under drought stressed conditions. These results are in accordance with Moghaddam et al. [57] who identified significant MTAs for SW on chromosome *Pv8* under drought stressed and well-watered environments. The current findings suggest that this QTL is stable across different environments and genetic backgrounds. On the contrary, several significant MTAs for SW were previously identified under drought stress on chromosome *Pv01*, [52], chromosome *Pv03* [51], chromosome *Pv09* [14], and chromosomes *Pv2* to *Pv4* and *Pv6* to *Pv11* [15]. Thus, the detection of significant MTAs for SW on different chromosomes and locations indicates high genetic diversity in common bean with respect to genomic regions associated with SW under drought stress. In this study, the identified SNPs that were significantly associated with GYD under drought stressed conditions were located on chromosomes *Pv04* (SNP 3382688) and *Pv11* (SNP 3384334 and SNP 3381526). Similarly, Dramadri et al. [34] identified significant QTL signals for GYD and yield components on chromosomes *Pv01*, *Pv02*, *Pv03*, *Pv04*, *Pv06*, and *Pv11* under drought stressed conditions. Oladzad et al. [96] also identified SNPs that were significantly associated with GYD, placed on chromosomes *Pv03*, *Pv08*, and *Pv11* under heat stress. Furthermore, Valdisser et al. [15] found 25 QTLs that were associated with GYD on chromosomes *Pv02*, *Pv03*, *Pv04*, *Pv08*, *Pv09* and *Pv11* under well-watered conditions, in agreement with the current findings. These findings suggest that chromosomes *Pv04* and *Pv11* harbour genes for controlling GYD.

The identification of SNPs associated with GYD, under drought stress, would significantly contribute to the development of molecular tools for MAS and identification of genes of interest for edition. The proportion of the total phenotypic variation (R^2) explained by the significant SNP markers for LCC and LT was generally low (0.11–0.12 for LCC under well-watered conditions and 8–15% for LT under well-watered conditions). Therefore, to account for the missing variation, it might be worthwhile to complement the SNP-based GWAS by haplotype-based GWAS [97].

Candidate genes

Drought stressed. The functional annotation revealed that the candidate gene for SC, *Phvul.001G254100* on chromosome *Pv01* encodes the CCCH zinc finger family protein. The CCCH zinc finger family protein plays an important function in response of plants to biotic and abiotic stresses [98–101]. This functional gene also plays an important role in physiological and plant developmental processes [101]. Similar findings were reported in *Brassica rapa* [98], common bean [15] and Barley (*Hordeum vulgare* L.) [101]. Wang et al. [102], Seong et al. [103] and Selvaraj et al. [104] reported that several types of CCCH zinc finger family genes such as *O_sC₃H₁₀*, *O_sC₃H₄₇*, and *OsTZF₅* are involved in the regulation of tolerance to drought stress in rice (*Oryza Sativa* L.). According to Lin et al. [105], the CCCH zinc finger protein gene confers drought tolerance in plants by regulating the opening and closing of stomata. They further reiterated that genotypes that are tolerant to drought stress have abnormal and lower stomatal conductance under drought stressed conditions. In this study, the marker SNP 3380850 for the gene *Phvul.001G254100* which confers tolerance to drought stress exhibited negative allelic effects ($-10.79 \text{ mmol m}^{-2} \text{ s}^{-1}$) on SC.

The functional annotation revealed that the candidate gene for DFW, *Phvul.002G122100* on chromosome *Pv02* encodes an RNA-recognition motif protein. The RNA-recognition motif protein plays a comprehensive biological function (critical modulators) in abiotic stress (drought, heat flooding, cold and high salinity) responding processes in plants [106]. Zhou et al. [107] observed that the RNA-recognition motif gene “OsCBP20” from rice confers abiotic stress tolerance in *escherichia coli*. Therefore, the candidate gene *Phvul.002G122100*

identified in this study may play a protective role under drought stressed conditions. Candidate genes such as *Phvul.003G263200* (*Pv08*) for SW which encodes for NADPH dehydrogenase plays an important role in mechanisms which protect plants against nitro-oxidative stresses [108]. The nitro-oxidative stresses are generated by biotic and abiotic stresses such as drought, low temperature, heat, and salinity [108]. Under drought stressed conditions, the seed is significantly affected by oxidative damages, and oxidative damages are minimized by the activity of NADPH dehydrogenase [109].

The candidate gene for GYD, *Phvul.004G150500* on chromosome *Pv04*, encodes the enzyme, phosphoethanolamine N-methyltransferase in plants. This catalytic enzyme plays an important role in the response of plants to abiotic stresses such as drought and salt tolerance [110]. The catalytic enzyme works by catalysing the methylation of phosphoethanolamine to phosphocholine [110]. Studies conducted by Wang et al. [110] in transgenic tobacco revealed that phosphoethanolamine N-methyltransferase improved the drought tolerance of transgenic tobacco. Notably, the marker (SNP 3382688) for this candidate gene *Phvul.004G150500* had large positive allelic effects (202.90 kg/ha) on GYD. The candidate gene for PH, *Phvul.001G172300* encodes the calcium transporting ATPase [111]. The calcium transporting ATPase plays an important role in growth and development processes, opening and closing of stomata, hormonal signalling, and regulation of responses to biotic and abiotic stresses in plants [111]. In summary, these results further confirmed that the identified putative potential candidate genes were associated with drought stress tolerance of dry beans. Therefore, the putative candidate genes identified in the current AMDP under drought stressed conditions are important genetic resources. The candidate genes could be utilized in drought tolerance breeding programmes by creating and introgressing new genetic variability into commercial cultivars.

Well-watered conditions. The functional annotation revealed that the candidate gene for PH “*Phvul.011G152000*” on chromosome *Pv11* encodes the threonine protein kinase. Threonine protein kinase is associated with enhanced tolerance to biotic and abiotic stresses in plants [15]. Similar results were reported in dry beans by Valdisser et al. [15]. In rice, kinase causes dwarfism by reducing plant height [112]. Similarly, in this study, the marker SNP 100073620 for the gene “*Phvul.001G152000*” exhibited negative allelic effects (-7.00 cm) on PH. According to Zhang et al. [112], kinases also has an impact on grain yield. The candidate gene *Phvul.004G037700* which was found on chromosome *Pv04* in association with DFW encodes transcription factor TCP₁₃. The transcription factor families are strongly involved in abiotic and biotic stress responses, including zinc-finger, dehydration-responsive element-binding (DREB), and basic helix-loop-helix (bHLH) families [113]. The basic helix-loop-helix (bHLH) families regulate plant growth in leaves and roots under water deficit conditions [113]. Studies conducted by Urano et al. [113] in *Arabidopsis thaliana* revealed that TCP₁₃ induces changes in leaf (leaf rolling and reduced leaf growth) and root morphology (enhanced root growth). This results in enhanced tolerance to dehydration stress under osmotic stress. The candidate gene *Phvul.004G055500* which was found in association with LT on chromosome *Pv04* encodes mitochondrial transcription termination factor family protein. According to Kim et al. [114], the mitochondrial transcription termination factor family protein enhances thermo-tolerance in *Arabidopsis*.

Conclusions

This study contributes many significant MTAs in common bean for agronomic and physiological traits under drought stressed and well-watered environments. The present study identified a total of 68 SNPs that were significantly ($p < 10^{-03}$) associated with key agronomic and

physiological traits under drought stressed and well-watered conditions. The highest number of significant MTAs were observed on chromosome *Pv11* in both environments. For the two environments (drought stressed and well-watered), no common SNPs for the studied traits were detected. Overall, twenty-two potential candidate genes were identified across environments. Most of the identified genes had known biological functions related to regulating drought stress response, and growth and development under drought stress. The information generated from this study provides insights into the genetic basis of agronomic and physiological traits under drought stressed and well-watered conditions. It also lays the foundation for future validation studies of drought tolerance genes in dry beans. Thus, the significant MTAs identified in this study should be explored and validated further to estimate their effects. This should be done using segregating populations and different genetic backgrounds before utilization in gene discovery and marker-assisted breeding for drought tolerance. Functional characterization and the application of gene knockout to the identified putative candidate genes would further confirm their roles in regulating drought stress response, growth and development under drought stressed and well-watered conditions. More powerful statistical genetics tools such as genomic prediction models would be needed to identify minor genes that are associated with agronomic and physiological traits. The admixed genotypes identified in this study offer potential as genetic resources in drought tolerance and biofortification breeding programmes, especially within the sugar, red mottled and navy bean market classes.

Supporting information

S1 Table. Characteristics of common bean genotypes used in the study, their sources and structure membership coefficient (K2) for K = 2.

(DOCX)

S2 Table. Grain yield (across environments), drought susceptibility index, geometric mean productivity, drought tolerance index and percent grain yield reduction of the 185 andean-middle american diversity panel.

(DOCX)

S1 Fig. Frequency distribution of the studied agronomic and physiological traits observed under well-watered conditions in the andean-middle American diversity panel.

(DOCX)

S2 Fig. Frequency distribution of the studied agronomic and physiological traits observed under drought stressed conditions in the andean-middle American diversity panel.

(DOCX)

S3 Fig. Quantile–Quantile (QQ) plots of the p- values observed and the expected from the genome-wide association study under drought stressed conditions. Note A = Leaf temperature, B = Days to 50% flowering, C = Grain yield, D = Plant height, E = Seed size, F = Stomatal conductance.

(DOCX)

S4 Fig. Quantile–Quantile (QQ) plots of the p- values observed and the expected from the genome-wide association study under well-watered conditions. Note A = Days to 50% flowering, B = Days to physiological maturity, C = Grain yield, D = Leaf chlorophyll content, E = Plant height, F = Leaf temperature.

(DOCX)

Acknowledgments

The authors would like to thank the Pan Africa Bean Research Alliance (PABRA) for the technical support received during the field and laboratory work. The authors are grateful to the Department of Research and Specialist Services (DR&SS) for providing their experimental sites for the field work. We would also like to thank PABRA for providing the germplasm.

Author Contributions

Conceptualization: Bruce Mutari, Julia Sibiya, Edmore Gasura.

Data curation: Bruce Mutari, Admire Shayanowako, Charity Chidzanga, Prince M. Matova.

Formal analysis: Bruce Mutari, Admire Shayanowako, Charity Chidzanga.

Funding acquisition: Bruce Mutari.

Investigation: Bruce Mutari.

Methodology: Bruce Mutari, Julia Sibiya, Prince M. Matova, Edmore Gasura.

Resources: Bruce Mutari.

Software: Admire Shayanowako, Charity Chidzanga.

Supervision: Julia Sibiya, Edmore Gasura.

Validation: Bruce Mutari, Julia Sibiya, Admire Shayanowako, Charity Chidzanga, Prince M. Matova, Edmore Gasura.

Visualization: Julia Sibiya.

Writing – original draft: Bruce Mutari, Julia Sibiya.

Writing – review & editing: Julia Sibiya, Admire Shayanowako, Charity Chidzanga, Prince M. Matova, Edmore Gasura.

References

1. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet.* 2014; 46(7): 707–13. <https://doi.org/10.1038/ng.3008> PMID: 24908249
2. Welch RM, House WA, Beebe S, Cheng Z. Genetic selection for enhanced bioavailable levels of iron in bean (*Phaseolus vulgaris* L.) seeds. *J Agric Food Chem.* 2000; 48(8): 3576–80.
3. Kamfwa K, Cichy KA, Kelly JD. Genome-wide association study of agronomic traits in common bean. *Plant Genome.* 2015; 8(2): 1–12. <https://doi.org/10.3835/plantgenome2014.09.0059> PMID: 33228312
4. Beebe SE, Rao IM, Blair MW, Acosta-Gallegos JA. Phenotyping common beans for adaptation to drought. *Front Physiol.* 2013; 4: 1–20. <http://doi.org/10.3389/fphys.2013.00035> PMID: 23507928
5. Sauer J. Historical geography of crop plants: A selected Roster. Boca Raton, USA: CRC Press; 1993.
6. Keller B, Ariza-Suarez D, Portilla-Benavides AE, Buendia HF, Aparicio JS, Amongi W, et al. Improving association studies and genomic predictions for climbing beans with data from bush bean populations. *Front Plant Sci.* 2022; 13. <https://doi.org/10.3389/fpls.2022.830896> PMID: 35557726
7. Singh SP, Gepts P, Debouck DG. Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Econ Bot.* 1991; 45(3): 379–96. <https://doi.org/10.1007/BF02887079>
8. Bitocchi E, Bellucci E, Giardini A, Rau D, Rodriguez M, Biagetti E, et al. Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytol.* 2013; 197(1): 300–13. <https://doi.org/10.1111/j.1469-8137.2012.04377.x> PMID: 23126683
9. Rainey K, Griffiths P. Evaluation of common bean yield components under heat stress. *HortScience.* 2003; 38(5): 682.

10. Rainey KM, Griffiths PD. Differential response of common bean genotypes to high temperature. *J Am Soc Hortic Sci.* 2005; 130(1): 18–23. <https://doi.org/10.21273/JASHS.130.1.18>
11. Katungi E, Farrow A, Chianu J, Sperling L, Beebe S. Common bean in eastern and southern Africa: A situation and outlook analysis. Cali, Colombia; 2009; 61: 1–44.
12. Beebe S. Common bean breeding in the tropics. In: Janick J, editor. *Plant breeding reviews.* 36th ed. New Jersey, USA: John Wiley & Sons, Ltd; 2012. pp. 357–412.
13. Katungi E, Farrow A, Mutuoki T, Gebeyehu S, Karanja D, Alemayehu F, et al. Improving common bean productivity: An analysis of socioeconomic factors in Ethiopia and eastern Kenya: Baseline report Tropical legumes II. Cali, Colombia; 2010; 126.
14. Hoyos-Villegas V, Song Q, Kelly JD. Genome-wide association analysis for drought tolerance and associated traits in common bean. *Plant Genome.* 2017; 10(1): 1–17. <https://doi.org/10.3835/plantgenome2015.12.0122> PMID: 28464065
15. Valdisser PAMR, Müller BSF, de Almeida Filho JE, Morais OP Júnior, Guimarães CM, Borba TCO, et al. Genome-wide association studies detect multiple QTLs for productivity in Mesoamerican diversity panel of common bean under drought stress. *Front Plant Sci.* 2020; 11. <https://doi.org/10.3389/fpls.2020.574674> PMID: 33343591
16. Kotir JH. Climate change and variability in Sub-Saharan Africa: a review of current and future trends and impacts on agriculture and food security. *Environ Dev Sustain.* 2011; 13(3): 587–605. <https://doi.org/10.1007/s10668-010-9278-0>
17. Tar'an B, Michaels TE, Pauls KP. Genetic mapping of agronomic traits in common bean. *Crop Sci.* 2002; 42(2): 544–56. <https://doi.org/10.2135/cropsci2002.5440>
18. Barrios AN, Hoogenboom G, Nesmith DS. Drought stress and the distribution of vegetative and reproductive traits of a bean cultivar. *Sci Agric.* 2005; 62(1): 18–22.
19. Manjeru P., Madanzi T., Makedza B., Nciizah A. and M S. Effects of water stress at different growth stages on components and grain yield of common bean (*Phaseolus vulgaris* L.). *African Crop Sci Conf Proc.* 2007; 8: 299–303.
20. Beebe SE, Rao IM, Cajiao C, Grajales M. Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Sci.* 2008; 48(2): 582–92. <http://doi.org/10.2135/cropsci2007.07.0404>
21. Beebe S.E, Rao I.M, Blair M, Acosta-Gallegos J. Phenotyping common beans for adaptation to drought. In: Ribaut JM, Monneveux P, editors. *Drought phenotyping in Crops: From theory to practice.* Mexico City, Mexico: Generation Challenge Program; 2010. pp. 311–34.
22. Darkwa K, Ambachew D, Mohammed H, Asfaw A, Blair MW. Evaluation of common bean (*Phaseolus vulgaris* L.) genotypes for drought stress adaptation in Ethiopia. *Crop J.* 2016; 4(5): 367–76. <https://doi.org/10.1016/j.cj.2016.06.007>
23. Asfaw A, Blair MW, Struik PC. Multi-environment quantitative trait loci analysis for photosynthate acquisition, accumulation, and remobilization traits in common bean under drought stress. *G3 (Bethesda).* 2012; 2(5): 579–95. <https://doi.org/10.1534/g3.112.002303> PMID: 22670228
24. Mutari B, Sibiya J, Gasura E, Matova PM, Simango K, Kondwakwenda A. Genetic analysis of grain yield and yield-attributing traits in navy bean (*Phaseolus vulgaris* L.) under drought stress. *Euphytica.* 2022; 218(5): 51. <https://doi.org/10.1007/s10681-022-03001-3>
25. Mutari B, Sibiya J, Bogweh Nchanji E, Simango K, Gasura E. Farmers' perceptions of navy bean (*Phaseolus vulgaris* L.) production constraints, preferred traits and farming systems and their implications on bean breeding: a case study from South East Lowveld region of Zimbabwe. *J Ethnobiol Ethnomed.* 2021; 17(1): 13. <https://doi.org/10.1186/s13002-021-00442-3> PMID: 33712041
26. Builes VHR, Porch TG, Harmsen EW. Genotypic differences in water use efficiency of common bean under drought stress. *Agron J.* 2011; 103(4): 1206–15. <https://doi.org/10.2134/agronj2010.0370>
27. Asfaw A, Blair MW, Almekinders C. Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the east African highlands. *Theor Appl Genet.* 2009; 120(1): 1–12. <https://doi.org/10.1007/s00122-009-1154-7> PMID: 19756469
28. Pérez-Vega E, Pañeda A, Rodríguez-Suárez C, Campa A, Giraldez R, Ferreira JJ. Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet.* 2010; 120(7): 1367–80. <https://doi.org/10.1007/s00122-010-1261-5> PMID: 20084493
29. Pereira HS, Mota APS, Rodrigues LA, de Souza TLPO, Melo LC. Genetic diversity among common bean cultivars based on agronomic traits and molecular markers and application to recommendation of parent lines. *Euphytica.* 2019; 215(2): 38. <https://doi.org/10.1007/s10681-018-2324-y>

30. Papa R, Gepts P. Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theor Appl Genet.* 2003; 106(2): 239–50. <https://doi.org/10.1007/s00122-002-1085-z> PMID: 12582849
31. Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, et al. A *Phaseolus vulgaris* diversity panel for Andean bean improvement. *Crop Sci.* 2015; 55(5): 2149–60. <https://doi.org/10.2135/cropsci2014.09.0653>
32. Perseguini JM KC, Oblessuc PR, Rosa JR BF, Gomes KA, Chiorato AF, Carbonell SAM, et al. Genome-wide association studies of anthracnose and angular leaf spot resistance in common bean (*Phaseolus vulgaris* L.). *PLoS One.* 2016; 11(3): e0150506. <https://doi.org/10.1371/journal.pone.0150506> PMID: 26930078
33. Katuuramu DN, Hart JP, Porch TG, Grusak MA, Glahn RP, Cichy KA. Genome-wide association analysis of nutritional composition-related traits and iron bioavailability in cooked dry beans (*Phaseolus vulgaris* L.). *Mol Breed.* 2018; 38(4): 44. <https://doi.org/10.1007/s11032-018-0798-x>
34. Dramadri IO, Nkalubo ST, Kelly JD. Identification of QTL Associated with drought tolerance in Andean common bean. *Crop Sci.* 2019; 59(3): 1007–20. <https://doi.org/10.2135/cropsci2018.10.0604>
35. Cichy KA, Fernandez A, Kilian A, Kelly JD, Galeano CH, Shaw S, et al. QTL analysis of canning quality and color retention in black beans (*Phaseolus vulgaris* L.). *Mol Breed.* 2014; 33(1): 139–54. <https://doi.org/10.1007/s11032-013-9940-y>
36. Cortés AJ, Chavarro MC, Blair MW. SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet.* 2011; 123(5): 827–45. <https://doi.org/10.1007/s00122-011-1630-8> PMID: 21785951
37. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in llants. *Plant Genome.* 2008; 1(1): 5–20. <https://doi.org/10.3835/plantgenome2008.02.0089>
38. Ingvarsson PK, Street NR. Association genetics of complex traits in plants. *New Phytol.* 2011; 189(4): 909–22. <https://doi.org/10.1111/j.1469-8137.2010.03593.x> PMID: 21182529
39. Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet.* 2012; 44(1): 32–9. <https://doi.org/10.1038/ng.1018> PMID: 22138690
40. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat Genet.* 2013; 45(1): 43–50. <https://doi.org/10.1038/ng.2484> PMID: 23242369
41. Li N, Shi J, Wang X, Liu G, Wang H. A combined linkage and regional association mapping validation and fine mapping of two major pleiotropic QTLs for seed weight and silique length in rapeseed (*Brassica napus* L.). *BMC Plant Biol.* 2014; 14(1): 114. <https://doi.org/10.1186/1471-2229-14-114> PMID: 24779415
42. Nordborg M, Tavaré S. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 2002; 18(2): 83–90. [https://doi.org/10.1016/s0168-9525\(02\)02557-x](https://doi.org/10.1016/s0168-9525(02)02557-x) PMID: 11818140
43. Huang X, Han B. Natural variations and genome-wide association studies in crop plants. *Annu Rev Plant Biol.* 2014; 65: 531–51. <https://doi.org/10.1146/annurev-arplant-050213-035715> PMID: 24274033
44. Sukumaran S, Yu J. Association mapping of genetic resources: achievements and future perspectives. In: Tuberosa R, Graner A, Frison E, editors. *Genomics of plant genetic resources*. Dordrecht, Netherlands: Springer; 2014. pp. 207–35.
45. Gizaw SA, Godoy JGV, Garland-Campbell K, Carter AH. Genome-wide association study of yield and component traits in Pacific Northwest winter wheat. *Crop Sci.* 2018; 58(6): 2315–30. <https://doi.org/10.2135/cropsci2017.12.0740>
46. Zuiderveen GH, Padder BA, Kamfwa K, Song Q, Kelly JD. Genome-wide association study of anthracnose resistance in Andean beans (*Phaseolus vulgaris*). *PLoS One.* 2016; 11(6): e0156391 <https://doi.org/10.1371/journal.pone.0156391> PMID: 27270627
47. Tock AJ, Fourie D, Walley PG, Holub EB, Soler A, Cichy KA, et al. Genome-wide linkage and association mapping of halo blight resistance in common bean to race 6 of the globally important bacterial pathogen. *Front Plant Sci.* 2017; 8: 1170. <https://doi.org/10.3389/fpls.2017.01170> PMID: 28736566
48. Tigist SG, Melis R, Sibiya J, Amelework AB, Keneni G, Tegene A. Population structure and genome-wide association analysis of bruchid resistance in Ethiopian common bean genotypes. *Crop Sci.* 2019; 59(4): 1504–15. <https://doi.org/10.2135/cropsci2018.09.0559>
49. Nkhata W, Shimelis H, Melis R, Chirwa R, Mzengeza T, Mathew I, et al. Combining ability analysis of common bean (*Phaseolus vulgaris* L.) genotypes for resistance to bean fly (*Ophiomyia* spp.), and grain yield and component traits. *Euphytica.* 2021; 217(5): 93. <https://doi.org/10.1007/s10681-021-02833-9>

50. Zia B, Shi A, Olaoye D, Xiong H, Ravelombola W, Gepts P, et al. Genome-wide association study and genomic prediction for bacterial wilt resistance in common bean (*Phaseolus vulgaris*) core collection. *Front Genet.* 2022; 13: 853114. <https://doi.org/10.3389/fgene.2022.853114> PMID: 35711938
51. Mukeshimana G, Butare L, Cregan PB, Blair MW, Kelly JD. Quantitative trait loci associated with drought tolerance in common bean. *Crop Sci.* 2014; 54(3): 923–38. <https://doi.org/10.2135/cropsci2013.06.0427>
52. Trapp JJ, Urrea CA, Cregan PB, Miklas PN. Quantitative trait loci for yield under multiple stress and drought conditions in a Dry Bean Population. *Crop Sci.* 2015; 55(4): 1596–607. <https://doi.org/10.2135/cropsci2014.11.0792>
53. Brifez B, Perseguini JM KC, Rosa JS, Bassi D, Gonçalves JGR, Almeida C, et al. Mapping QTLs for drought tolerance in a SEA 5 x AND 277 common bean cross with SSRs and SNP markers. *Genet Mol Biol.* 2017; 40(4): 813–23. <https://doi.org/10.1590/1678-4685-GMB-2016-0222> PMID: 29064511
54. Kamfwa K, Cichy KA, Kelly JD. Genome-wide association analysis of symbiotic nitrogen fixation in common bean. *Theor Appl Genet.* 2015; 128(10): 1999–2017. <https://doi.org/10.1007/s00122-015-2562-5> PMID: 26133733
55. Cichy KA, Wiesinger JA, Mendoza FA. Genetic diversity and genome-wide association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). *Theor Appl Genet.* 2015; 128(8): 1555–67. <https://doi.org/10.1007/s00122-015-2531-z> PMID: 26003191
56. Makunde G. Quantification of genetic diversity for drought adaptation in a reference collection of common bean (*Phaseolus vulgaris* L.). Doctoral thesis, University of Free State. 2013. <https://scholar.ufs.ac.za:8080/handle/11660/8210>.
57. Moghaddam SM, Mamidi S, Osorno JM, Lee R, Brick M, Kelly J, et al. Genome-wide association study identifies candidate Loci underlying agronomic traits in a Middle American diversity panel of common bean. *Plant Genome.* 2016; 9(3): 1–21. <https://doi.org/10.3835/plantgenome2016.02.0012> PMID: 27902795
58. Beattie AD, Larsen J, Michaels TE, Pauls KP. Mapping quantitative trait loci for a common bean (*Phaseolus vulgaris* L.) ideotype. *Genome.* 2003; 46(3): 411–22. <https://doi.org/10.1139/g03-015> PMID: 12834057
59. Blair MW, Iriarte G, Beebe S. QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean x wild common bean (*Phaseolus vulgaris* L.) cross. *Theor Appl Genet.* 2006; 112(6): 1149–63. <https://doi.org/10.1007/s00122-006-0217-2> PMID: 16432734
60. Wright EM, Kelly JD. Mapping QTL for seed yield and canning quality following processing of black bean (*Phaseolus vulgaris* L.). *Euphytica.* 2011; 179(3): 471–84. <https://doi.org/10.1007/s10681-011-0369-2>
61. Checa OE, Blair MW. Inheritance of yield-related traits in climbing beans (*Phaseolus vulgaris* L.). *Crop Sci.* 2012; 52(5): 1998–2013. <https://doi.org/10.2135/cropsci2011.07.0368>
62. Payne RW, Murray DA, Harding S. An introduction to GenStat command language. Hampstead, UK: VSN International; 2018.
63. Cullis BR, Smith AB, Coombes NE. On the design of early generation variety trials with correlated data. *J Agric Biol Environ Stat.* 2006; 11(4): 381–93. <https://doi.org/10.1198/108571106X154443>
64. Johnson HW, Robinson HF, Comstock RE. Estimates of genetic and environmental variability in Soybeans. *Agron J.* 1955; 47(7): 314–8. <https://doi.org/10.2134/agronj1955.00021962004700070009x>
65. DArT. Extraction protocol for DArT. Canberra, Australia; 2014.
66. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 2007; 23(19): 2633–5. <https://doi.org/10.1093/bioinformatics/btm308> PMID: 17586829
67. Qin P, Lin Y, Hu Y, Liu K, Mao S, Li Z, et al. Genome-wide association study of drought-related resistance traits in *Aegilops tauschii*. *Genet Mol Biol.* 2016; 39(3): 398–407. <https://doi.org/10.1590/1678-4685-GMB-2015-0232> PMID: 27560650
68. Pritchard J, Wen X, Falush D. Documentation for Structure Software: Version 2.3. University of Chicago, Chicago, 2010.
69. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* 2005; 14(8): 2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> PMID: 15969739
70. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* 2012; 4(2): 359–61. <https://doi.org/10.1007/s12686-011-9548-7>

71. Ojwang PPO, Eldridge T, Corredor-Moreno P, Njung'e V. Structure of genetic diversity and genome-wide association studies of bean fly (*Ophiomyia spencerella*) resistance in common bean. *Euphytica*. 2021; 217(12): 216. <https://doi.org/10.1007/s10681-021-02949-y>
72. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; 38(8): 904–9. <https://doi.org/10.1038/ng1847> PMID: 16862161
73. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics*. 2012; 28(18): 2397–9. <https://doi.org/10.1093/bioinformatics/bts444> PMID: 22796960
74. Raggi L, Caproni L, Carboni A, Negri V. Genome-wide association study reveals candidate genes for flowering time variation in common bean (*Phaseolus vulgaris* L.). *Front Plant Sci*. 2019; 10. <https://doi.org/10.3389/fpls.2019.00962> PMID: 31428109
75. Liu P, Jin Y, Liu J, Liu C, Yao H, Luo F, et al. Genome-wide association mapping of root system architecture traits in common wheat (*Triticum aestivum* L.). *Euphytica*. 2019; 215(7): 121. <https://doi.org/10.1007/s10681-019-2452-z>
76. Yin L. R package “CMPlots.” 2016. [cited 2022 February 11]. <https://github.com/YinLiLin/R-CMplot>.
77. Shin J-H, Blay S, McNeney B, Graham J. LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Software, Code Snippets*. 2006; 16(3): 1–9. <https://doi.org/10.18637/jss.v016.c03>
78. Core Team R. Vienna: R foundation for statistical computing. 2016.
79. Blair MW, Cortés AJ, Farmer AD, Huang W, Ambachew D, Penmetsa RV, et al. Uneven recombination rate and linkage disequilibrium across a reference SNP map for common bean (*Phaseolus vulgaris* L.). *PLoS One*. 2018; 13(3): e0189597. <https://doi.org/10.1371/journal.pone.0189597> PMID: 29522524
80. Assefa T, Beebe SE, Rao IM, Cuasquer JB, Duque MC, Rivera M, et al. Pod harvest index as a selection criterion to improve drought resistance in white pea bean. *Field Crops Res*. 2013; 148: 24–33. <https://doi.org/10.1016/j.fcr.2013.04.008>
81. Assefa T, Rao IM, Cannon SB, Wu J, Gutema Z, Blair M, et al. Improving adaptation to drought stress in white pea bean (*Phaseolus vulgaris* L.): Genotypic effects on grain yield, yield components and pod harvest index. *Plant Breed*. 2017; 136(4): 548–61. <https://doi.org/10.1111/pbr.12496>
82. Mathobo R, Marais D, Steyn JM. The effect of drought stress on yield, leaf gaseous exchange and chlorophyll fluorescence of dry beans (*Phaseolus vulgaris* L.). *Agric Water Manag*. 2017; 180: 118–25. <https://doi.org/10.1016/j.agwat.2016.11.005>
83. Schneider KA, Rosales-Serna R, Ibarra-Perez F, Cazares-Enriquez B, Acosta-Gallegos JA, Ramirez-Vallejo P, et al. Improving common bean performance under drought stress. *Crop Sci*. 1997; 37(1): 43–50. <https://doi.org/10.2135/cropsci1997.0011183X003700010007x>
84. Beebe S. Biofortification of common bean for higher iron concentration. *Front Sustain Food Syst*. 2020; 4: 206. <https://doi.org/10.3389/fsufs.2020.573449>
85. Liu Z, Gao S, Zhang H, Xu Z, Qian W. Genome-wide association study reveals that PvGUX1_1 is associated with pod stringlessness in snap bean (*Phaseolus vulgaris* L.). *Biology (Basel)*. 2022; 11(4): 611. <https://doi.org/10.3390/biology11040611> PMID: 35453811
86. Crop Breeding Institute. Release proposal for a drought tolerant bean line: Sweet William. Harare, Zimbabwe; 2017.
87. Crop Breeding Institute. Release proposal for a biofortified bean line: NUA674. Harare, Zimbabwe; 2018.
88. Crop Breeding Institute. Release proposal for a biofortified bean line: SMC16. Harare, Zimbabwe; 2019.
89. Kondwakwenda A, Mutari B, Simango K, Nchanji EB, Chirwa R, Rubyogo JC, et al. Decades of cultivar development: A reconciliation of maize and bean breeding projects and their impacts on food, nutrition security, and income of smallholder farmers in Sub-Saharan Africa. In: Mupambwa HA, Nciizah AD, Nyambo P, Muchara B, Gabriel NN, editors. *Food security for African smallholder farmers*. Singapore: Springer Nature Singapore; 2022. pp. 3–26.
90. Singh SP. Selection for water-stress tolerance in interracial populations of common bean. *Crop Sci*. 1995; 35(1): 118–24. <https://doi.org/10.2135/cropsci1995.0011183X003500010022x>
91. Islam FMA, Beebe S, Muñoz M, Tohme J, Redden RJ, Basford KE. Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool. *Theor Appl Genet*. 2004; 108(2): 243–52. <https://doi.org/10.1007/s00122-003-1437-3> PMID: 14657984

92. Vos PG, Paulo MJ, Voorrips RE, Visser RGF, van Eck HJ, van Eeuwijk FA. Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theor Appl Genet.* 2017; 130(1): 123–35. <https://doi.org/10.1007/s00122-016-2798-8> PMID: 27699464
93. Mwadzingeni L, Shimelis H, Rees DJG, Tsilo TJ. Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. *PLoS One.* 2017; 12(2): e0171692. <https://doi.org/10.1371/journal.pone.0171692> PMID: 28234945
94. El Gataa Z, El Hanafi S, Basheer F, Kehel Z, bouhouch Y, El Messoadi K, et al. Genome wide association study of grain yield and yield related traits in spring bread wheat (*Triticum aestivum* L.) under drought and heat conditions in three different locations. *J Crop Sci Biotechnol.* 2021; 24(4): 361–73. <https://doi.org/10.1007/s12892-021-00084-7>
95. Sukumaran S, Reynolds MP, Sansaloni C. Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. *Front Plant Sci.* 2018; 9. <https://doi.org/10.3389/fpls.2018.00081> PMID: 29467776
96. Oladzad A, Porch T, Rosas JC, Moghaddam SM, Beaver J, Beebe SE, et al. Single and multi-trait GWAS identify genetic factors associated with production traits in common bean under abiotic stress environments. *G3 Genes|Genomes|Genetics.* 2019; 9(6): 1881–92. <https://doi.org/10.1534/g3.119.400072> PMID: 31167806
97. N'Diaye A, Haile JK, Cory AT, Clarke FR, Clarke JM, Knox RE, et al. Single marker and haplotype-based association analysis of semolina and pasta colour in elite durum wheat breeding lines using a high-density consensus Map. *PLoS One.* 2017; 12(1): e0170941. <https://doi.org/10.1371/journal.pone.0170941> PMID: 28135299
98. Pi B, He X, Ruan Y, Jang J-C, Huang Y. Genome-wide analysis and stress-responsive expression of CCCH zinc finger family genes in Brassica rapa. *BMC Plant Biol.* 2018; 18(1): 373. <https://doi.org/10.1186/s12870-018-1608-7> PMID: 30587139
99. Han G, Lu C, Guo J, Qiao Z, Sui N, Qiu N, et al. C2H2 zinc finger proteins: Master regulators of abiotic stress responses in plants. *Front Plant Sci.* 2020; 11: 115. <https://doi.org/10.3389/fpls.2020.00115> PMID: 32153617
100. Han G, Qiao Z, Li Y, Wang C, Wang B. The Roles of CCCH zinc-finger proteins in plant abiotic stress tolerance. *Int J Mol Sci.* 2021; 22(15): 8327. <https://doi.org/10.3390/ijms22158327> PMID: 34361093
101. Ai Q, Pan W, Zeng Y, Li Y, Cui L. CCCH zinc finger genes in barley: genome-wide identification, evolution, expression and haplotype analysis. *BMC Plant Biol.* 2022; 22(117): 1–20. <https://doi.org/10.1186/s12870-022-03500-4> PMID: 35291942
102. Wang W, Liu B, Xu M, Jamil M, Wang G. ABA-induced CCCH tandem zinc finger protein OsC3H47 decreases ABA sensitivity and promotes drought tolerance in *Oryza sativa*. *Biochem Biophys Res Commun.* 2015; 464(1): 33–7. <https://doi.org/10.1016/j.bbrc.2015.05.087> PMID: 26047696
103. Seong SY, Shim JS, Bang SW, Kim J-K. Over-expression of OsC3H10, a CCCH-zinc finger improves drought tolerance in rice by regulating stress-related genes. *Plants.* 2020; 9(10): 1298. <https://doi.org/10.3390/plants9101298> PMID: 33019599
104. Selvaraj MG, Jan A, Ishizaki T, Valencia M, Dedicova B, Maruyama K, et al. Expression of the CCCH-tandem zinc finger protein gene OsTZF5 under a stress-inducible promoter mitigates the effect of drought stress on rice grain yield under field conditions. *Plant Biotechnol J.* 2020; 18(8): 1711–21. <https://doi.org/10.1111/pbi.13334> PMID: 31930666
105. Lin P-C, Pomeranz MC, Jikumaru Y, Kang SG, Hah C, Fujioka S, et al. The Arabidopsis tandem zinc finger protein AtTZF1 affects ABA- and GA-mediated growth, stress and gene expression responses. *Plant J.* 2011; 65(2): 253–68. <https://doi.org/10.1111/j.1365-313X.2010.04419.x> PMID: 21223390
106. Muthusamy M, Kim J-H, Kim JA, Lee S-I. Plant RNA binding proteins as critical modulators in drought, high salinity, heat, and cold stress responses: An updated overview. *Int J Mol Sci.* 2021; 22(13): 6731. <https://doi.org/10.3390/ijms22136731> PMID: 34201749
107. Zhou C, Chen R-J, Gao X-L, Li L-H, Xu Z-J. Heterologous expression of a rice RNA-recognition motif gene OsCBP20 in *Escherichia coli* confers abiotic stress tolerance. *Plant Omics.* 2014; 7(1): 28–34.
108. Corpas FJ, Barroso JB. NADPH-generating dehydrogenases: Their role in the mechanism of protection against nitro-oxidative stress induced by adverse environmental conditions. *Front Environ Sci.* 2014; 2:55. <https://doi.org/10.3389/fenvs.2014.00055>
109. Berny Mier y Teran JC, Konzen ER, Palkovic A, Tsai SM, Rao IM, Beebe S, et al. Effect of drought stress on the genetic architecture of photosynthate allocation and remobilization in pods of common bean (*Phaseolus vulgaris* L.), a key species for food security. *BMC Plant Biol.* 2019; 19(1): 171. <https://doi.org/10.1186/s12870-019-1774-2> PMID: 31039735
110. Wang A-H, Yang L, Yao X-Z, Wen X-P. Over-expression of the pitaya *phosphoethanolamine N-methyltransferase* gene (HpPEAMT) enhanced simulated drought stress in tobacco. *Plant Cell, Tissue Organ Cult.* 2021; 146(1): 29–40. <https://doi.org/10.1007/s11240-021-02040-3>

111. Singh A, Kanwar P, Yadav AK, Mishra M, Jha SK, Baranwal V, et al. Genome-wide expressional and functional analysis of calcium transport elements during abiotic stress and development in rice. *FEBS J.* 2014; 281(3): 894–915. <https://doi.org/10.1111/febs.12656> PMID: 24286292
112. Zhang Y, Xiao W, Luo L, Pang J, Rong W, He C. Down regulation of OsPK1, a cytosolic pyruvate kinase, by T-DNA insertion causes dwarfism and panicle enclosure in rice. *Planta.* 2012; 235(1): 25–38. <https://doi.org/10.1007/s00425-011-1471-3> PMID: 21805151
113. Urano K, Maruyama K, Koyama T, Gonzalez N, Inzé D, Yamaguchi-Shinozaki K, et al. CIN-like TCP13 is essential for plant growth regulation under dehydration stress. *Plant Mol Biol.* 2022; 108(3): 257–75. <https://doi.org/10.1007/s11103-021-01238-5> PMID: 35050466
114. Kim M, Lee U, Small I, des Francs-Small CC, Vierling E. Mutations in an Arabidopsis mitochondrial transcription termination factor-related protein enhance thermotolerance in the absence of the major molecular chaperone HSP101. *Plant Cell.* 2012; 24(8): 3349–65. <https://doi.org/10.1105/tpc.112.101006> PMID: 22942382