Received: 5 February 2024 | Accepted: 14 August 2024

DOI: 10.1002/tpg2.20515

The Plant Genome

ORIGINAL ARTICLE

Special Section: Modern Improvement of Tropical Crops

Using cross-country datasets for association mapping in *Arachis hypogaea* L.

Velma Okaron[1] | **James Mwololo**[2] | **Davis M. Gimode**[3] | **David K. Okello**[4] | **Millicent Avosa**[3] | **Josh Clevenger**[5] | **Walid Korani**[5] | **Mildred Ochwo Ssemakula**[1] | **Thomas L. Odong**[1] | **Damaris A. Odeny**[3]

[1]Department of Agricultural Production, School of Agricultural Sciences, College of Agricultural and Environmental Sciences, Makerere University, Kampala, Uganda

[2]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Bulawayo, Zimbabwe

[3]International Crops Research Institute for the Semi-Arid Tropics- Kenya, Nairobi, Kenya

[4]National Semi-Arid Resources Research Institute, Soroti, Uganda

[5]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA

Correspondence
Damaris A. Odeny, International Crops Research Institute for the Semi-Arid Tropics- Kenya, P.O. Box 39063-00623, Nairobi, Kenya.
Email: Damaris.Odeny@icrisat.org

Assigned to Associate Editor Stella Salvo.

Abstract

Groundnut (*Arachis hypogaea* L.) is one of the most important climate-resilient oil crops in sub-Saharan Africa. There is a significant yield gap for groundnut in Africa because of poor soil fertility, low agricultural inputs, biotic and abiotic stresses. Cross-country evaluations of promising breeding lines can facilitate the varietal development process. The objective of our study was to characterize popular test environments in Uganda (Serere and Nakabango) and Malawi (Chitala and Chitedze) and identify genotypes with stable superior yields for potential future release. Phenotypic data were generated for 192 breeding lines for yield-related traits, while genotypic data were generated using skim-sequencing. We observed significant variation ($p < 0.001$; $p < 0.01$; $p < 0.05$) across genotypes for all yield-related traits: days to flowering (DTF), pod yield (PY), shelling percentage, 100-seed weight, and grain yield within and across locations. Nakabango, Chitedze, and Serere were clustered as one mega-environment with the top five most stable genotypes being ICGV-SM 01709, ICGV-SM 15575, ICGV-SM 90704, ICGV-SM 15576, and ICGV-SM 03710, all Virginia types. Population structure analysis clustered the genotypes in three distinct groups based on market classes. Eight and four marker-trait associations (MTAs) were recorded for DTF and PY, respectively. One of the MTAs for DTF was co-localized within an uncharacterized protein on chromosome 13, while another one (TRv2Chr.11_3476885) was consistent across the two countries. Future studies will need to further characterize the candidate genes as well as confirm the stability of superior genotypes across seasons before recommending them for release.

Abbreviations: AP, available phosphorus; BLINK, Bayesian-information and linkage-disequilibrium iteratively nested keyway; BLUP, best linear unbiased prediction; DAPC, discriminant analysis of principal components; DTF, days to 50% flowering; ESA, eastern and southern Africa; GGE, genotype plus genotype–environment interaction; GWAS, genome-wide association studies; GY, grain yield; HSW, 100-seed weight; ICRISAT, The International Crops Research Institute for the Semi-Arid Tropics; IPCA, interaction principal component axis; LD, linkage disequilibrium; MAF, minor allele frequency; MLMM, multi-locus mixed model; MTA, marker-trait association; PC, principal component; PCA, principal component analysis; PY, pod yield; QTL, quantitative trait locus; SAT, semi-arid tropic; SH%, shelling percentage; SNP, single nucleotide polymorphism; SSA, sub-Saharan Africa; TN, total nitrogen.

Plain Language Summary

Most countries in eastern and southern Africa derive their groundnut breeding lines from International Crops Research Institute for the Semi-Arid Tropics breeding program based in Malawi. In some cases, the same genotype is released in several countries under different names. However, the evaluation of the genotypes is often taken independently in each of the countries, leading to duplication of work. A more cost-effective method is to identify similar environments across different countries and evaluate the same genotypes across such environments. In this study, we evaluated 192 groundnut genotypes across four environments, two each from Uganda and Malawi. Additive main effects and multiplicative interaction analysis clustered the Ugandan sites and Chitedze in one mega-environment, implying that future evaluations could take advantage of such environments towards varietal releases. We also used the same data to detect marker-trait associations across the different locations for agronomic traits. Our results revealed more consistent results within Uganda than Malawi.

1 | INTRODUCTION

Groundnut (*Arachis hypogaea* L.), also known as peanut, is an important climate-resilient grain legume and oilseed crop, especially in the semi-arid tropics (SATs). The crop is cultivated in more than 100 countries (Pandey et al., 2020) with China as the largest producer and exporter globally (Bansal et al., 2017). In Africa, groundnut is mainly produced in Nigeria, Senegal, Uganda, Tanzania, South Africa, and Sudan, and utilized as human food, edible oil, and livestock feed (Okello et al., 2010). Despite its importance in sub-Saharan Africa (SSA), groundnut yields remain extremely low in comparison to yields in the West. For instance, the average grain yield (GY) in the United States was 4072 kg/ha in 2021 (USDA-NASS, 2021), while that in SSA was approximately 886 kg/ha in the same year (FAOSTAT, 2021). This yield gap is attributed to poor soil fertility (Bekele et al., 2022), low agricultural inputs (Chapu et al., 2022), biotic (Okello et al., 2013) and abiotic stresses (Pandey et al., 2014).

Genetic improvement of agronomic traits promises to address the yield gap and resolve the existing market demand (Varshney et al., 2013; B. Wang et al., 2020). A better understanding of the genetic basis underlying agronomic traits would focus breeding and enhance the development of innovative breeding tools (Zhang et al., 2017). Among the important agronomic traits for groundnut in Africa are earliness and pod yield (PY) (Zongo et al., 2017). Short growing seasons have become increasingly more frequent and extreme with the increasing effects of climate change (Cook & Vizy, 2012). High-yielding, early-maturing groundnut varieties have the potential of avoiding extended periods of drought and still providing a reasonable yield to the farmer, especially in Africa where the crop is grown by smallholder farmers without irrigation facilities.

Malawi and Uganda are two countries in eastern and southern Africa (ESA) with strong groundnut breeding programs. Both programs share the same breeding objectives and utilize similar breeding lines, usually originating from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Cross-country breeding across these two countries would be vital for the identification of stable and predictable varieties (Badu-Apraku et al., 2008). In Africa, cross-country evaluation studies have been limited to a few major crops such as maize (*Zea mays* L.) and common bean (*Phaseolus vulgaris* L.) (Aggarwal et al., 2004; Kassa et al., 2013; Mupangwa et al., 2020). Although some cross-country evaluation of promising groundnut varieties has been done in Africa (Pandey et al., 2014; Okori et al., 2019, 2014), it is not a common practice.

Genomic tools offer the potential to significantly accelerate breeding efficiency, leading to rapid development of improved cultivars. Molecular markers associated with important yield-related traits such as pod weight, seed weight, yield per pod, pod branch number per plant, pod shape, and 100-seed weight (HSW) have been reported in groundnut using genome-wide association studies (GWAS) (Wang et al., 2019; Shaibu et al., 2020; Zhao et al., 2022). Similar studies are yet to be undertaken across breeding programs in Africa. In the current study, we generated agronomic data for a diverse set of breeding lines across Uganda and Malawi. We compared the performance of genotypes within and across countries, estimated the stability in performance, and used genomic tools to determine relatedness of genotypes as well as establish potential marker-trait associations (MTAs).

2 | MATERIALS AND METHODS

2.1 | Plant material and experimental sites

Note that 192 groundnut genotypes comprising of 82 Virginia, 78 Spanish, and 32 Valencia were obtained from ICRISAT Malawi (Table 1) for this study. The germplasm set included 189 elite cultivars and three high-yielding commercial varieties, Kakoma (JL 24), CG7, and CHALIMBANA, as local checks. The study was carried out in four locations, namely, Chitedze and Chitala in Malawi during the 2020 cropping season (December–April), and the National Semi-Arid Resources Research Institute (Serere district), and Nakabango (Jinja district) in Uganda during the long rainy season of 2021 (2021B: September–December) (Table 1; Figure 1). These locations are representative of major groundnut producing agroecological zones of ESA.

2.2 | Baseline soil characterization

Soil sampling was done using a zigzag pattern (Okalebo et al., 2002). From each point, a soil sample was collected using an Edelman soil auger at 0- to 20-cm depth and undecomposed plant materials removed by hand. Three soil samples from each field were analyzed for texture, pH, total C, total nitrogen (TN), available phosphorus (AP), and cation exchange capacity at the Plant, Soil and Water Analytical laboratory in Makerere University, Uganda. Texture of the soils was obtained using the hydrometer method, while soil pH was determined on 2.5:1 water to soil suspension (Okalebo et al., 2002). AP was determined using the Olsen method (Estefan, 2013). Exchangeable bases (Ca, Mg, Na, and K) were extracted using ammonium acetate and determined by atomic absorption spectrophotometry. The Walkley–Black wet combustion method was used to determine organic carbon, while TN was measured using the Kjeldahl method (va Reeuwijk, 2002). AP was determined using the bicarbonate solution (0.5 M $NaHCO_3$ at pH 8.5) method (Hue et al., 2000).

Core Ideas

- One mega-environment was identified between Nakabango, Serere, and Chitedze.
- We identified five Virginia groundnut genotypes with stability in their performance across Malawi and Uganda.
- Days to 50% flowering quantitative trait loci were consistent across the two countries, while pod yield was country-specific.

2.3 | Experimental design, trials, and data collection

The experiment was set using a 14 × 25 alpha lattice design with two replicates at all locations. The plot size was 1 m × 0.90 m, and each plot consisted of 3 m rows, with inter-row spacing of 45 cm. Inter-plot distance was 60 cm and spacing between replications was 2 m. Agronomic practices recommended for groundnut production were followed. Harvesting and all other postharvest handling processes were done manually on a plot basis using hand hoes. Stripping was manually done soon after harvesting. Pods from each plot were sun-dried to <13% moisture content (Min GAC-Plus moisture tester, DICKEY-John Corporation) and thereafter weighed using a weighing scale (LBK, ADAM equipment). Data on yield and yield-related traits were collected at different stages, which included days to 50% flowering (DTF; counts), dry PY (kg/ha), HSW (g), shelling outturn (shelling percentage [SH%]), and GY (kg/ha). GY per plot was converted to kilograms per hectare, according to Rana and Kumar (2014).

2.4 | DNA extraction and genotyping

Groundnut leaf samples were collected into 96-well plates, 2 weeks after planting for DNA extraction. Total genomic DNA was isolated using the ISOLATE II Genomic DNA

TABLE 1 Geographical description of the experimental sites.

Locations	Geographical position		Altitude (a.s.l.)	Average annual temperature (˚C)	Average annual rainfall (mm)
	Latitude	Longitude			
Chitedze (Malawi)	13˚85′ S	33˚38′ E	1146	20.0	892
Chitala (Malawi)	13˚67′ S	34˚27′ E	607	24.0	1200
Serere (Uganda)	1˚39′ N	33˚27′ N	1040	28.0	1150
Nakabango (Uganda)	0˚31′ N	33˚13′ E	1188	22.0	1200

Abbreviation: a.s.l., above sea level.
Source: NASA POWER (https://power.larc.nasa.gov).

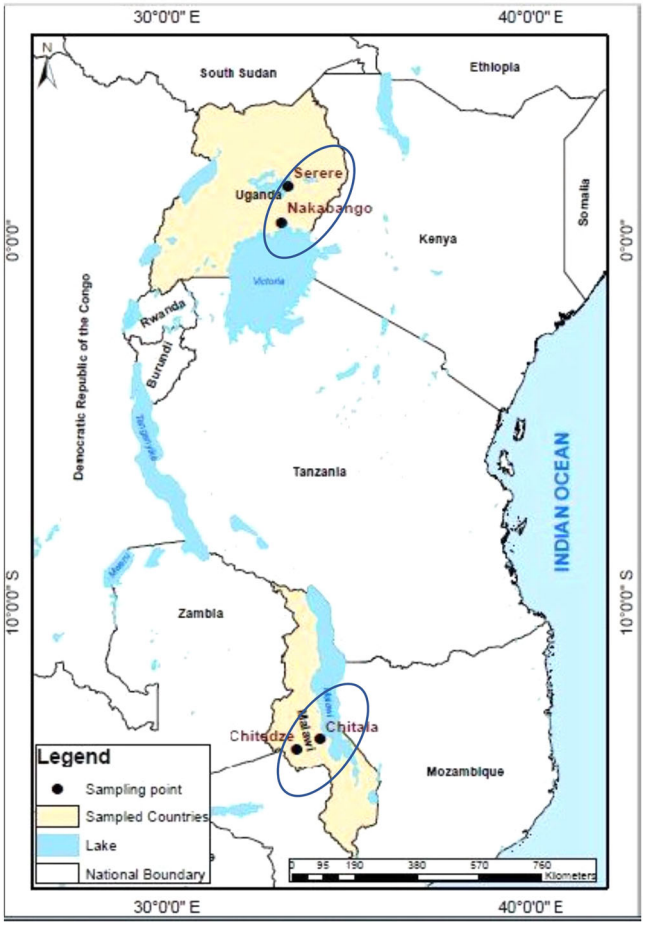

FIGURE 1 Geographical map showing experimental sites (highlighted in black circles). *Source*: Modified from Google map.

extraction kit (Bioline Pty Ltd.) according to the manufacturer's instructions. The purity and quantity of the extracted DNA were determined using gel electrophoresis and a Qubit 2.0 Fluorometer (Life Technologies), respectively, with final dilution of 50 ng/μL. DNA samples were sent to Psomagen for library preparation and whole genome sequencing. Libraries were constructed using the RIPTIDE kit (Twist Bioscience) according to the manufacturer's protocol. The libraries were subjected to paired-end sequencing on the Illumina Novaseq 6000 equipment (Illumina). Single nucleotide polymorphism (SNP) calling was done using the Khufu pipeline (Korani et al., 2021) that is optimized for accurate SNP calling from low-coverage reads. The raw SNPs were filtered at a call rate >0.95, minor allele frequency (MAF) > 0.05, and heterozygosity < 0.2. The distribution of the final filtered high-quality SNPs was plotted across chromosomes using CMplot (Yin et al., 2021). Imputation of the ordered marker data was performed using Beagle version 4 (Browning & Browning, 2016) before pruning markers that were in complete linkage disequilibrium (LD) with another marker at a threshold of 0.2. The parameters were set as described by Jordan et al. (2015).

2.5 | Statistical analysis

2.5.1 | Phenotypic data analysis

Analysis of variance (ANOVA) was done on all the traits evaluated from each test location using "R" statistical software, version 4.0.3 (R Core Team, 2020). A linear mixed effects model using the "lmerTest" package in "R" (O. K. Bates et al., 2020) was used to estimate variance components. Combined ANOVA over locations was conducted using a mixed model as indicated below:

$$Y_{ij} = \mu + \beta_i + G_i + E_j + GE_{ij} + e_{ij},$$

where Y_{ij} is the trait value of genotype, μ is the grand mean, β_i is the random effect of the ith genotype, G_i is the fixed effect of the ith genotype, E_j is the jth environmental effect, GE_{ij} is the ijth genotype × environment (G × E) effect, and ij is the treatment × block interaction, treated as an error term.

Means generated from analyses of variance were separated using Duncan's new multiple range test at a 5% level of significance. Pearson's correlation was used to determine the relationship among variables. To assess and quantify the genetic variability among inbred breeding lines, heritability in the broad sense (H^2) was estimated as follows:

$$H^2 = \sigma^2 G/(\sigma^2 G + \sigma^2 GE/e + \sigma^2 \varepsilon/re),$$

where H^2 is the broad-sense heritability, $\sigma^2 G$ is the genotypic variance, $\sigma^2 GE$ is the variance of G × E interaction, $\sigma^2 \varepsilon$ is the error variance, e is the environment number, and r is the number of replications. The heritability estimates were classified as described by Johnson et al. (1955): low (0–30), medium (30.1–60), high (≥60.1). Accordingly, best linear unbiased predictions (BLUPs) for each variety were generated using the *ranef* function in *lme4* package (Bates et al., 2015). BLUPs were used to derive correlation, stability, and GWAS analyses as they provide better estimates of genotype performance for unbalanced datasets (Piepho et al., 2008).

2.5.2 | Yield stability across sites and countries

Adaptability and stability of different groundnut genotypes were determined using the "Metan" package in R software (Olivoto & Lúcio, 2020) with the model:

$$Y_{ijk} = \mu + G_i + E_j + \sum_{K=1}^{M} \lambda_k \times \alpha_{ik} \times \gamma_{jk} + \rho_{ij},$$

where Y_{ijk} is the yield of the ith genotype in the jth environment, G_i is the effect of the ith genotype (genotype mean minus the grand mean), E_j is the effect of the jth environment

(environment mean minus the grand mean), λ_k is the square root of the eigenvalue of the kth interaction principal component axis (IPCA), α_{ik} and γ_{jk} are the principal component (PC) scores for IPCA k of the ith genotypes and the jth environment, respectively, and ρ_{ij} is the deviation of genotype ith on environment jth from the model. To determine the mega-environments and visualize the "which-won-where" pattern, genotype plus genotype–environment interaction (GGE) analysis was performed using Metan package in R software (Olivoto & Lúcio, 2020). The GGE biplot was based on singular value decomposition of PCs, as described by Yan and Tinker (2006) and the GGE model below implemented:

$$Y_{ij} = \mu_i + \beta_j + \sum \lambda_k \times \alpha_{ik} \times \Upsilon_{jk} + \varepsilon_{ij},$$

where Y_{ij} is the performance of the ith genotype in the jth environment, μ is the grand mean, β_j is the main effect of jth environment, k is the number of PC, λ_k is the singular value of the kth PC, α_{ik} and γ_{jk} are the scores for PC of the ith genotypes and jth environment, respectively, and ε_{ij} is the residual associated with the ith genotype and jth environment.

For mega-environment delineation, the "which-won-where" scatter plot was constructed based on an irregular as recommended by Yan and Kang (2002). Which-won-where was proposed by Gauch and Zobel (1997) as a criterion for mega-environment delineation. The GGE biplot (Yan & Kang, 2002) is considered the most effective 2D chart to display the G and GE as it displays the "which-won-where" in the form of an irregular polygon (Yan et al., 2023). The polygon is formed by connecting the cultivars placed furthermost from the biplot origin in all directions so that all other genotypes are contained within the polygon (Yan et al., 2023). Then radiant lines perpendicular to each of the polygon sides are drawn from the biplot origin, and the genotype placed at the vertex of a sector is the nominal winner in the environments that fall in the sector (Yan et al., 2023).

The comparison plot of genotype ranking relative to ideal genotype was generated by symmetrical scaling, using the same concept of average environment coordinate to draw an analogy between genotypes and an ideal genotype (Nduwumuremyi et al., 2017). Yan and Tinker (2006) described the ideal genotype as both high yielding and stable across environments. Genotype stability was analyzed using the Anniccharico stability index (Silva, 2008). GGE type 8 analysis was used to display the genotype stability rank.

2.5.3 | Population structure and linkage disequilibrium

The population structure was analyzed by discriminant analysis of principal components (DAPC) using the *adegenet* package (Jombart et al., 2010) in the R 3.0.2 software (R Core Team, 2013). The function DAPC was executed using clusters identified by *K*-means (Legendre & Legendre, 1998). The number of clusters in the population was assessed using the function "find.clusters." The optimal number of clusters was chosen on the basis of the lowest associated Bayesian information criterion. The LD statistics were computed in TASSEL V 5.0 using a sliding window of 50 SNPs and MAF = 0.05. The LD decay rate was estimated by plotting r^2 values versus corresponding physical distances between the SNP pairs in R (R Core Team, 2020) using the methodology implemented in Remington et al. (2001).

2.5.4 | Genome-wide association analysis

Genome-wide association analysis was carried out using markers that passed the SNP filtering and quality control (QC) step. Calculated BLUPs for the 192 genotypes were combined with genotypic data and MTAs calculated using the multi-locus mixed model (MLMM; Segura et al., 2012) as well as the Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) (Liu et al., 2016). MLMM follows a stepwise approach to incorporate SNPs as covariates in the GWAS model. BLINK is an improved model of the fixed and random model circulating probability unification and is statistically powerful and efficient in identifying significant SNPs associated with a trait of importance (Huang et al., 2019). Both models used principal component analysis (PCA) as the fixed effect. All analyses were carried out in R using the genome association and prediction integrated tool version 2 (Tang et al., 2016), and the resulting associations displayed as Manhattan plots alongside *Q*–*Q* plots to demonstrate model fitness. The significance of MTAs was determined using the false discovery rate ($\alpha = 0.05$). Putative candidate genes within LD distance of the significant SNPs were identified based on the Tifrunner 2.0 reference genome.

3 | RESULTS

3.1 | Soil characteristics of the test locations

Based on classification of natural fertility level (Rosen et al., 2008), soil pH at Nakabango, Chitala, and Chitedze were slightly acidic, while Serere consisted of alkaline soils (Table 2). Organic matter content was low across sites as was available soil P status in all the sites, except Chitedze. All sites registered medium to high levels of K. Serere and Nakabango scored very high Mg content compared to Chitala and Chitedze. Calcium ranged from low to high across the four sites.

TABLE 2 Characteristics of the test locations soils.

Parameter	Serere	Remarks	Nakabango	Remarks	Chitala	Remarks	Chitedze	Remarks
pH (H_2O)	6.9	Medium	5.2	Low	5.7	Low	5.3	low
Organic matter %	1.1	Very low	3.1	Low	0.9	Very low	1.9	Very low
Phosphorus (mg/kg)	3.5	Very low	3.7	Very low	0.7	Very low	4.0	low
Exchangeable K (cmols(+)/kg)	0.6	High	0.3	Medium	0.4	Medium	0.3	Medium
Exchangeable Ca (cmols(+)/kg)	9.0	Medium	16.5	High	3.6	Low	7.8	Medium
Exchangeable Mg (cmols(+)/kg)	36.4	Very high	28.8	Very high	1.6	Very low	1.9	Very low
% Sand	44.0	High	48.6	Medium	68.0	Very high	54.8	Very high
% Clay	14.0	Low	46.0	Medium	28.0	Low	39.6	Medium
Texture class	Sandy loam		Clay loam		Clay		Crystalline	

TABLE 3 Combined analysis of variance for yield and related traits of 192 groundnut genotypes across Malawi and Uganda.

SOV	*df*	DTF	PY (kg/ha)	SH%	HSW (g)	GY (kg/ha)
Rep	1	12.0*	24,380.0ns	37.0ns	672.0**	252,931.0ns
Rep/Blk	68	3.2*	707,132.0ns	102.0*	86.0ns	302,959.0ns
Genotype	191	15.5***	2,116,808.0***	348.0***	212.0***	1009,561.0***
Loc	3	2399***	494,165,869.0***	34,993.0 ***	13,536.0***	256,904,619.0***
Genotype:Loc	573	3.7***	1,197,170.0***	329.0***	122.0***	660,791.0***
Residuals	677	2.3	660,573.0	38.0	85.0	293,160.0
Mean		32.3	1508.0	50.0	38.0	896.0
LSD		1.7	735.0	5.0	8.0	476.0
CV%		4.7	50.0	12.0	24.0	58.0
H^2		0.7	0.4	N.A	0.4	0.3

Abbreviations: Blk, block; CV, coefficient of variation; *df*, degree of freedom; DTF, days to 50% flowering; GY, grain yield; HSW, 100-seed weight; Loc, location; LSD, least significant difference; ns, nonsignificant; PY, pod yield; Rep, replicates; SH%, shelling percentage; SOV, source of variation.
***, **, and * depict significance at $p < 0.001$, $p < 0.01$, and $p < 0.05$, respectively.

3.2 | Analysis of phenotypic data

3.2.1 | Overall performance across sites

We observed significant ($p < 0.001$; $p < 0.01$; $p < 0.05$) variation across genotypes for most of the traits for both single site and across site analysis (Table S2). For Ugandan sites, no significant variation was observed across genotypes for SH% (both sites) and for HSW (Serere). Significant variations ($p < 0.001$) were, however, observed for all traits measured across Malawi sites except GY in Chitedze (Table S2). DTF recorded the highest heritability estimates (70%) across sites (Table 3) with Chitala site having the highest value (0.82; Table S2). Ugandan sites reported the lowest broad-sense heritability (H^2) values in comparison to Malawi with Serere having the lowest figures overall. Chitala had the highest H^2 values for all traits. We also observed significant ($p < 0.001$) positive correlations across all yield-related traits (PY, SH, HSW, and GY). DTF revealed significant ($p < 0.001$) positive correlation with HSW but not for the other yield-related traits (Table 4).

TABLE 4 Pearson's correlation coefficient (*r*) for yield and yield components across environments.

	DTF	PY (kg/ha)	SH%	HSW (g)	GY (kg/ha)
DTF	1				
PY (kg/ha)	−0.009	1			
SH%	−0.032	0.525***	1		
HSW	0.367***	0.312***	0.337***	1	
GY (kg/ha)	−0.026	0.974***	0.599***	0.301***	1

Abbreviations: DTF, days to 50% flowering; GY, grain yield; HSW, 100-seed weight; PY, pod yield; SH%, shelling percentage.
***Significant at $p < 0.001$.

AMMI ANOVA showed that the genotype, environment, and G × E interaction effects were highly significant ($p < 0.001$) for GY (Table 5). Genotype explained 13.2% of the total sum of squares, while environment and G × E interaction contributed 50.8% and 22.4% to the total sum of squares, respectively (Table 5). The G × E interaction partitioned among the first two IPCAs. IPCA1 was highly

TABLE 5 AMMI ANOVA for grain yield of 192 groundnut lines across environments.

Source	*df*	SS	MS	*F*-value	PR (< *F*)	TSS (%)
Total	1535	3,448,419,104	2,246,527			
Treatments	767	2,985,872,380	3,892,924	6.5	0.001***	86.5
Genotypes	191	458,051,921	2,398,178	4.0	0.001***	13.2
Environments	3	1,752,666,032	584,222,011	42.3	0.001***	50.8
Block	4	55,159,414	13,789,853	23.1	0.001***	1.5
Interactions	569	775,154,426	1,362,310	2.2	0.001***	22.4
IPCA 1	193	612,681,395	3,174,515	5.3	0.001***	17.7
IPCA 2	191	112,704,313	590,075	0.9	0.523*	3.2
Residuals	185	49,768,718	269,020	0.4	1	
Error	684	407,387,310	595,595			

Abbreviations: ANOVA, analysis of variance; *df*, degrees of freedom; F, Fisher value; IPCA, interaction principal component axis; MS, mean square; ns, nonsignificant; PR, probability; SS, sum of squares; TSS (%), percentage of total sum square.
*$p < 0.05$. ***$p < 0.001$.

TABLE 6 AMMI ranking of the four best performing genotypes per environment based on grain yield.

Environment	Mean yield (kg/ha)	Score of G × E cross-over	1	2	3	4
Chitala	3264	−114.4	ICGV 01276	ICGV-SM 16596	ICGV-SM 16637	ICGV-SM 15559
Chitedze	1175	34.2	ICGV-SM 15579	ICGV-SM 15578	ICGV-SM 16627	ICGV-SM 03710
Nakabango	454	42.7	ICGV-SM 15611	ICGV-SM 15621	ICGV-SM 15514	ICGV-SM 15598
Serere	993	37.4	ICGV-SM 15611	ICGV-SM 01709	ICGV-SM 15621	ICGV-SM 15587

Abbreviations: E, environment; G, genotype.

significant ($p < 0.001$) (Table 5), while IPCA2 was significant at $p < 0.523$ (Table 5). Both IPCA1 and IPCA2 jointly accounted for 20.9% of the total G × E interaction sum of squares, with 17.7% and 3.2% accounted by IPCA1 and IPCA2, respectively (Table 5).

3.2.2 | Genotype performance and stability within and across environments

The overall mean yield was higher in Malawi sites, with Chitala recording the highest mean yield and Nakabango the lowest (Table 6). The top four performing genotypes in each environment as revealed by AMMI ANOVA for GY indicated similarities in performance across Ugandan sites with two (ICGV-SM 15611 and ICGV-SM 15621) of the top four genotypes being common (Table 6). None of the top four genotypes in either Chitala or Chitedze were common across the two environments. The differences in ranking of genotypes across the environments indicated the presence of G × E cross-over, with Chitala reporting the highest (−114.4) G × E interaction compared to the other three sites (Table 6).

The first PC of the AMMI biplot explained 82.4% of the total G × E interaction sum of squares and further revealed the best-performing genotypes across all the four locations (Figure 2A). The three locations, Chitedze, Nakabango, and Serere, were reported as a mega-environment, which revealed similar high performing genotypes (Figure 2A; Table S3). The results of "which-won-where" biplot revealed ICGV-SM 15578 (G93) as the winning genotype in terms of overall performance, with ICGV-SM 10044 (G79) reporting the worst overall performance (Figure 2B).

In terms of yield stability, ranking biplot and Annicchario stability index recorded genotypes ICGV-SM 01709, ICGV-SM 15575, ICGV-SM 90704, ICGV-SM 15576, and ICGV-SM 03710, all Virginia market types, as the top five most stable across the mega-environment (Figure 3A,B).

3.3 | Genotypic data analysis

3.3.1 | Genetic diversity, population structure, and linkage disequilibrium

Skim-sequencing led to the calling of 39,254 raw SNP markers from 192 genotypes. After filtering the raw SNPs at a call rate of >0.95, MAF > 0.05, and heterozygosity < 0.2, we retained 25,101 high quality SNPs. The distribution of the

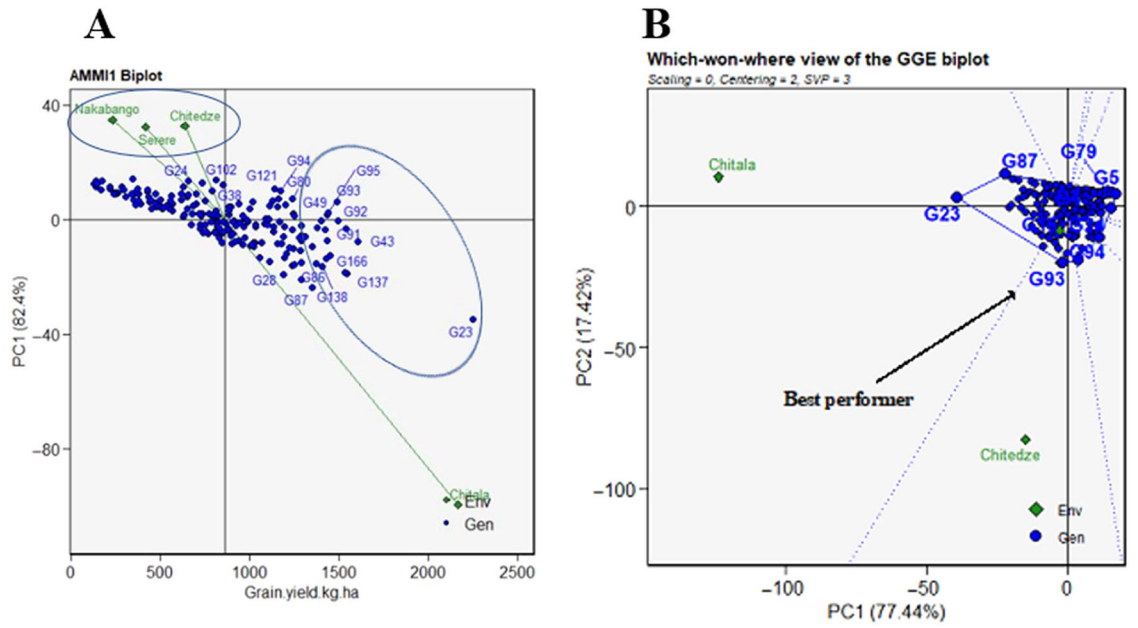

FIGURE 2 Biplots revealing performance of the genotypes across the four environments and the effects of the environments. (A) AMMI-1 model biplot for grain yield (kg/ha) showing the means of 192 genotypes across four environments against their corresponding interaction principal component axis (IPCA-1) scores. Groundnut genotypes placed on the right-hand side of the midline (vertical) reported higher grain yield compared to those on the left-hand side. Out of the four locations used, three clustered into one mega-environment. (B) Genotype plus genotype–environment interaction (GGE) scatterplot based on symmetrical scaling for the "*which-won-where*" pattern of the 192 groundnut genotypes evaluated in four environments. G93 (ICGV-SM 15578) is highlighted as the overall winner across all environments. PC, principal component; SVP, singular value partitioning.

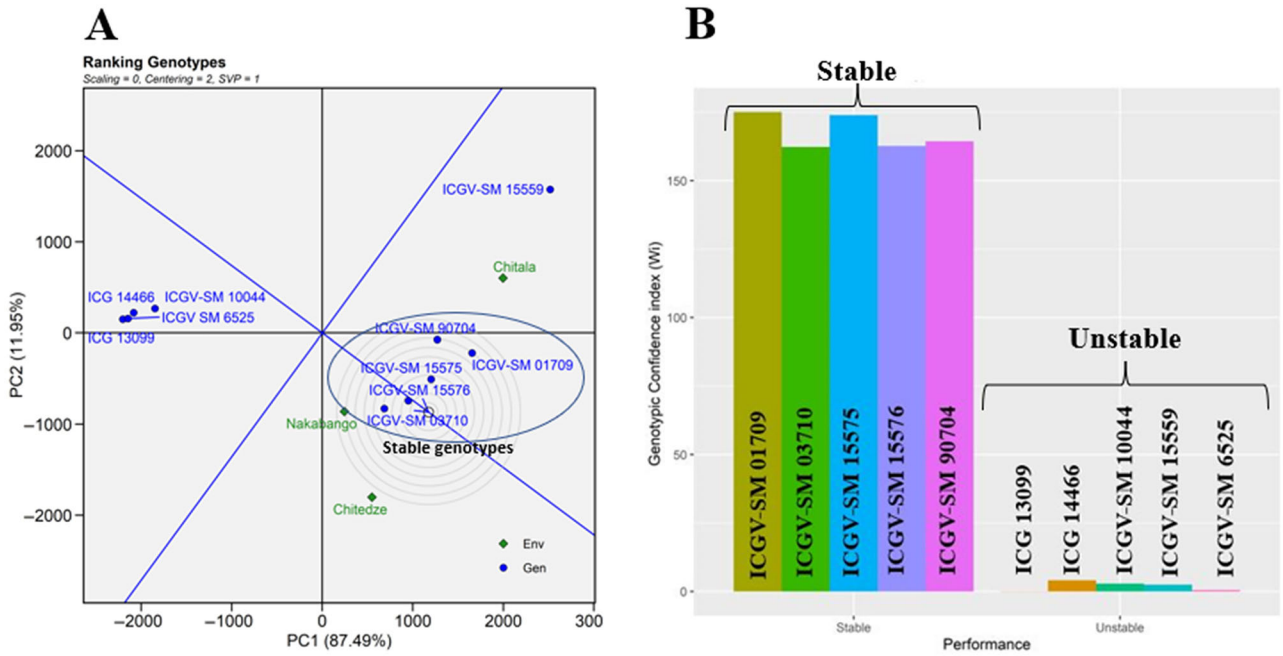

FIGURE 3 Performance of genotypes based on both high yield and stability across the mega-environment. (A) Ranking plot showing the best genotypes based on mean performance and stability. The most stable genotypes are highlighted in blue. (B) A chart showing the top most stable genotypes and unstable genotypes. PC, principal component; SVP, singular value partitioning.

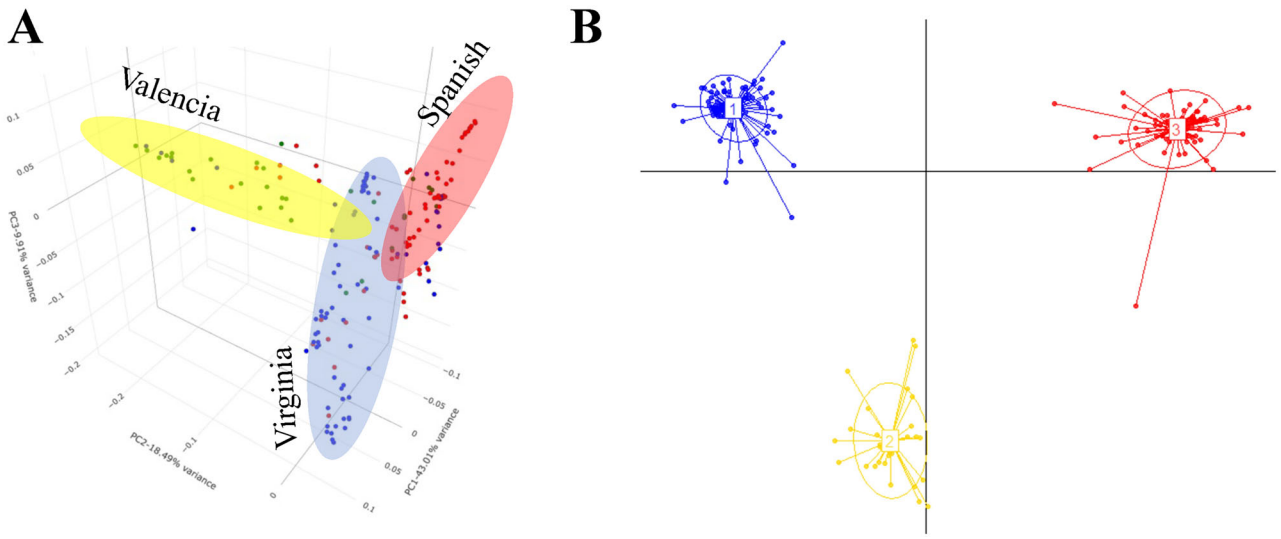

FIGURE 4 Population structure analysis of the genotypes. (A) Principal component analysis (PCA) plot with three principal components (PCs) showing the genetic variation across the genotypes studies. Three clusters corresponding to Valencia (green), Spanish (red), and Virginia (blue) market classes are observed. (B) Discriminant analysis of principal components (DAPC) for the 192 genotypes. Each cluster is represented by blue (predominantly Virginia), yellow (mainly Valencia), and red (predominantly Spanish).

final filtered high-quality SNPs was plotted across chromosomes using CMplot (Yin et al., 2021) and used for GWAS. PCA revealed that the markers retained were of high quality, with the first two PCs explaining 61.5% of genetic variation across the genotypes used (Figure 4A). PCA clustered the genotypes into three groups based on their market classes (Figure 4A). Virginia market class (58 genotypes) was predominant in cluster 1, followed by Valencia (19 genotypes) and Spanish (61 genotypes) market classes in clusters 2 and 3, respectively.

DAPC further confirmed the optimal number of subpopulations at $K = 3$ (Figure 4B). Cluster 1 of the DAPC plot comprised of 72 genotypes, which were mainly of Virginia market class and corresponded with the blue cluster in the PC plot (Figure 4B). Cluster 2 was predominantly Valencia, while cluster 3 was mainly Spanish types (Figure 4B).

The overall density of SNPs used across the groundnut genome was approximately 9.4 SNPs/Mbp (Figure 5A), with an average decay distance of approximately 2.8 Mb at $r^2 = 0.2$ (Figure 5B).

3.3.2 | Quantitative variation and marker-trait associations

MTAs were detected only for PY and DTF. Both traits exhibited quantitative variation in each of the countries (Figure 6).

A summary of the significant MTAs is presented in Table 7 and Table S4. Eight markers were significantly associated with DTF either in one location or across locations (Table 7). The most consistent marker associated with DTF across locations was TRv2Chr.11_3476885, when datasets across all four locations were analyzed with both BLINK and MLMM models (Figure 7), or when data for Nakabango or Chitedze locations were analyzed independently (Figures S1 and S2). The markers on chromosomes 2 (TRv2Chr.02_5893782) and 5 (TRv2Chr.05_102793211) were associated with DTF across Ugandan locations only, while markers TRv2Chr.13_6535112, TRv2Chr.15_90096196, and TRv2Chr.18_16912352 were exclusive to Malawi sites only (Figures S1 and S2).

PY MTAs were country-specific. Two markers, TRv2Chr.03_139043669 (BLINK and MLMM models) and TRv2Chr.15_19961316, were detected across Ugandan sites only (Table 7; Figure S3), while the marker TRv2Chr.10_115356321 was associated with PY across Malawi sites only using the BLINK model (Table 7; Figure S4).

4 | DISCUSSION

4.1 | Overall performance across the sites

The overall goal of this study was to compare the performance of a diverse set of breeding lines within and across Uganda and Malawi, as well as stability in performance, and the genetic basis underlying agronomic traits. There was significant variation across genotypes for all the traits in both single-site and across-site analyses. G × E effects were highly significant for all traits studied, indicating significant variations in genotype mean performance across environments. We also identified

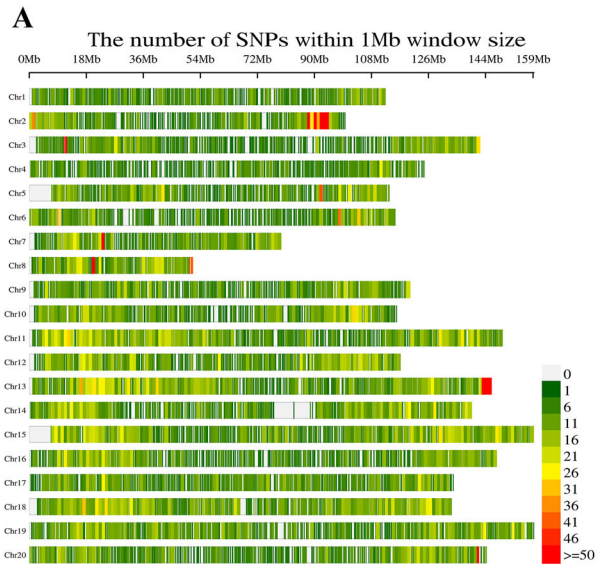

FIGURE 5 The locations and distribution of retained single nucleotide polymorphisms (SNPs) after filtering. (A) The distribution of retained SNPs across the 20 chromosomes of groundnut. (B) Linkage disequilibrium (LD) decay plot using the retained SNPs. An overall LD decay distance of 2.8 Mb was observed.

TABLE 7 Marker-trait associations for DTF and PY across Malawi and Uganda.

Trait	SNP	*p*-value	FDR-adjusted *p*-value	MAF	Model	Location
DTF	TRv2Chr.02_5893782	1.99E-07	0.005	0.422	BLINK	Across Uganda
	TRv2Chr.05_102793211	1.72E-07	0.004	0.270	BLINK	Nakabango
	TRv2Chr.08_27036796	4.38E-12	1.12E-07	0.130	BLINK	Across Uganda and Malawi
	TRv2Chr.11_3476885	7.74E-07	0.009	0.338	BLINK	Nakabango
	TRv2Chr.11_3476885	4.38E-10	5.58E-06	0.338	BLINK	Across Uganda and Malawi
	TRv2Chr.11_3476885	2.48E-07	0.006	0.338	MLMM	Across Uganda and Malawi
	TRv2Chr.11_18318654	6.66E-08	0.001	0.380	BLINK	Chitedze
	TRv2Chr.13_6535112	1.20E-07	0.003	0.289	BLINK	Chitala
	TRv2Chr.13_6535112	6.19E-08	0.001	0.289	BLINK	Across Malawi
	TRv2Chr.15_90096196	6.66E-10	8.49E-06	0.390	BLINK	Chitedze
	TRv2Chr.18_16912352	6.66E-10	8.49E-06	0.091	BLINK	Chitedze
PY	TRv2Chr.03_139043669	4.13E-08	0.001	0.286	BLINK	Across Uganda
	TRv2Chr.03_139043669	9.63E-07	0.024	0.286	MLMM	Across Uganda
	TRv2Chr.15_19961316	6.79E-08	0.001	0.367	BLINK	Across Uganda
	TRv2Chr.10_115356321	9.99E-08	0.002	0.135	BLINK	Chitala
	TRv2Chr.10_115356321	5.58E-07	0.014	0.135	BLINK	Across Malawi

Abbreviations: BLINK, Bayesian-information and linkage-disequilibrium iteratively nested keyway; DTF, days to 50% flowering; FDR, false discovery rate; MAF, minor allele frequency; MLMM, multi-locus mixed model; PY, pod yield; SNP, single nucleotide polymorphism.

five top-performing genotypes with good yield stability that will be of interest to breeders across the region. Although we identified significant MTAs for DTF and PY and the corresponding candidate genes, there will be need for further validation in future experiments. Our findings highlight the importance of cross-country screening and how such datasets can be used to share breeding resources for groundnut and trait characterization.

The significant differences and wide range of phenotypes detected among the test genotypes for yield-related traits indicate the great potential of breeding lines coming from the ICRISAT-Malawi breeding program for use in the

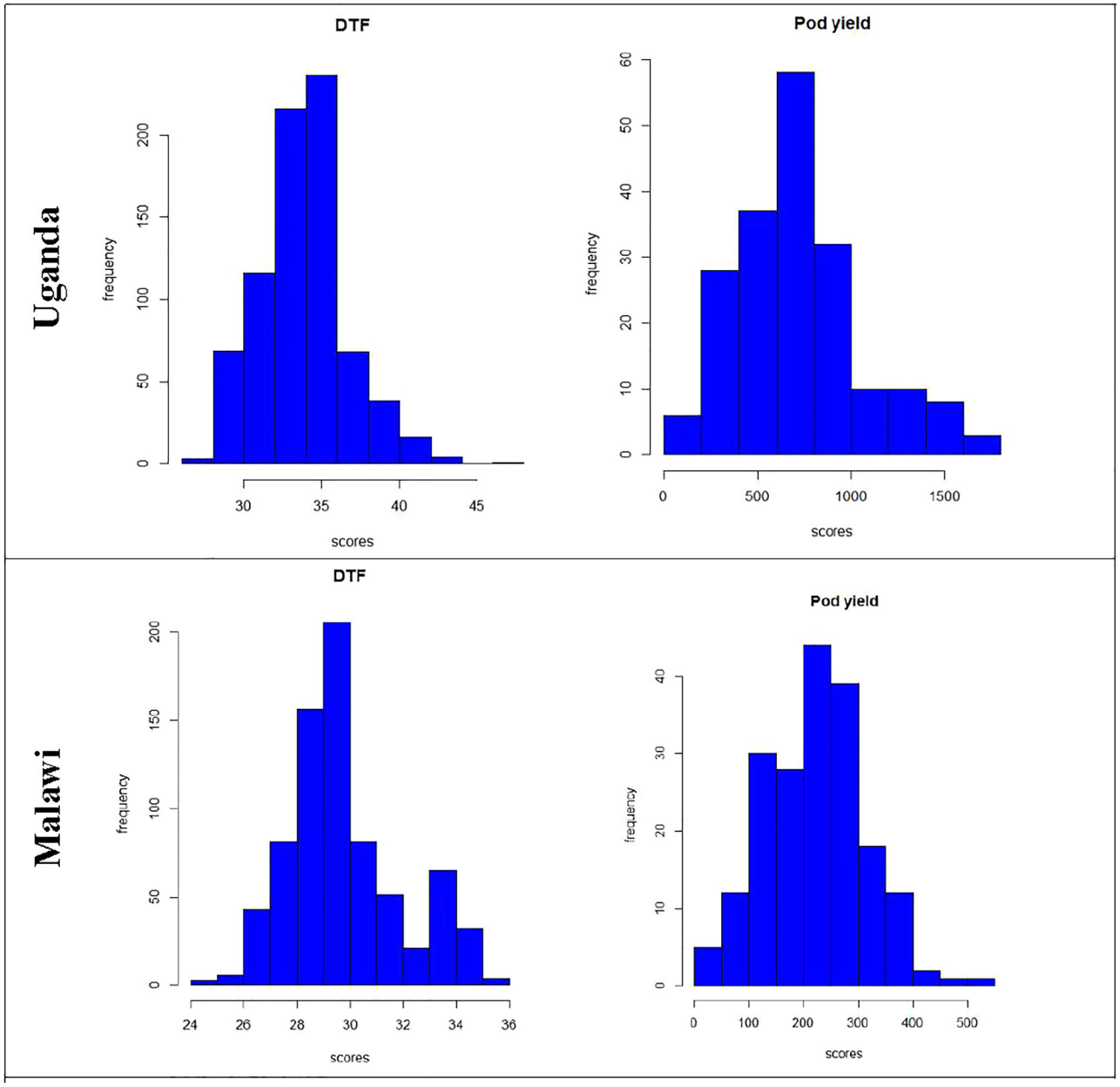

FIGURE 6 Quantitative variation for DTF and pod yield (PY) across Uganda and Malawi sites.

development of high-yielding varieties in the ESA region. Despite the narrow genetic diversity in groundnut, previous studies on yield-related traits have reported good variation in both cultivated (J. Zhao et al., 2017; H. Zhao et al., 2022) and wild relatives (Essandoh et al., 2022; Tossim et al., 2010). Genomic regions associated with such traits have also been identified (J. Wang et al., 2019; H. Zhao et al., 2022; Zhou et al., 2021). Breeders will now need to validate and use the markers associated with yield in their breeding programs and use them routinely for marker-assisted selection. In our case, this is the first study using cross-country datasets, and therefore a repeat study will be necessary to validate the markers before they can be used in crop improvement programs in ESA.

Despite using cross-country datasets, the moderate (30%) to high (80%) heritability estimates reported for earliness and yield components are comparable to those reported earlier for the same traits (J. Zhao et al., 2017) and imply that genetic factors played a predominant role in determining variation. We observed good positive correlation between yield-related traits, providing further confidence that higher groundnut yields could be achieved by indirect selection of genotypes with high seed weight per plant. Similar results were reported by other researchers (Badigannavar et al., 2002; Mekontchou

FIGURE 7 Consistency of marker TRv2Chr.11_3476885 detected when datasets across all locations were analyzed using both multi-locus mixed model (MLMM) and Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) model.

et al., 2006; Meta & Monpara, 2010). The high genotype-by-environment interaction effect on the studied traits was not surprising given their quantitative nature (Okori et al., 2019). Environmental differences can be attributed to variations in temperature, rainfall, soil type, and diseases (Casanoves et al., 2005; Shahriari et al., 2018) as reported by previous studies (Makinde & Ariyo, 2011; Negash et al., 2013; Okori et al., 2019; Sewagegne et al., 2013). The initial soil analysis done across all the sites provided a baseline for comparison of the agronomic performance of the genotypes in each location.

4.2 | Genotype performance and stability within and across environments

Our results revealed that the two locations in Uganda were more similar than the locations used in Malawi. First, two of the top four performing genotypes in each of the Ugandan locations were common, whereas in Malawi, none of the top four performing genotypes were common across locations. Second, the AMMI biplot grouped the three sites, Nakabango, Serere, and Chitedze, as one mega-environment, leaving out Chitala. Clustering test environments into homogeneous groups could serve as a foundation for testing of genotypes in fewer locations, thus lowering experimental costs (Kebede & Getahun, 2017). The clustering of Chitedze with Ugandan locations will allow for more coordinated trials across the two countries to enhance varietal releases. The two common top performing genotypes across Ugandan locations, ICGV-SM15611 and ICGV-SM15621, are obvious candidates for further stability tests and national performance trials in preparation for potential release.

The use of GGE biplots in this study enabled the identification of genotypes that combined high mean performance with high stability, as well as highlighting preferences and adaptation to environments. Similar results have been achieved in soybean where GGE biplot depicted the overall effect of specific genotypes, as well as G × E interaction (Dhilon et al., 2009). The "which-won-where pattern" of the GGE biplot's polygon view-based interaction was effective for identifying elite genotypes in single or multiple settings (Yan & Tinker, 2006). G × E datasets can be used to evaluate the ability of test environments to discriminate genotypes (Yan, 2001). A conclusive evaluation of a test environment would, however, require datasets across several seasons (Yan & Holland, 2010). Our study provides a baseline for future testing of the ability of the four environments to discriminate genotypes of interest.

The use of GGE biplots has been reported as a powerful solution for determining stability of both cereals (Mohammadi et al., 2023; Shojaei et al., 2022; Vaezi et al., 2019) and legumes (Dalló et al., 2019; Lal et al., 2019), including groundnut (Greveniotis et al., 2023; Pobkhunthod et al., 2022) across environments. Similar studies in Africa for groundnut have been done in Mali (Sanogo et al., 2019) and across east and southern African countries (Okori et al., 2019). However, these past studies included just a few varieties lined up for potential release. The advantage of our study is the inclusion

of several breeding lines across different countries, making it possible for early selection while at the same time exploiting genetic variation for different traits across stable genotypes. Our study, therefore, forms a basis for future genomic selection (GS) trials in groundnut across Uganda and Malawi.

4.3 | Genetic diversity, population structure, and linkage disequilibrium

We used SNP markers that had been called from skim-sequencing datasets mapped to the *A. hypogaea* reference genome using the KHUFU pipeline (Korani et al., 2021). The markers were very informative and clustered the genotypes into three major groups based on market classes: Virginia, Valencia, and Spanish. This result was consistent with other PCA reports for breeding lines from the continent (Achola et al., 2023; Conde et al., 2023). The Spanish market class, which belongs to the subspecies *fastigiata*, is the most popular in Africa followed by the Virginia (*hypogaea* subspecies) market class (Conde et al., 2023). Our results suggest that there are several genotypes that have not been classified correctly. For example, in the distinct clusters generated by the DAPC and PCA plots, the blue-shaded clusters that were predominantly Virginia, also contained Spanish- and Valencia-coded market classes. The same case applied to clusters 2 and 3. Although the misclassification of market classes among African genotypes has been recently partially addressed by Conde et al. (2023), cross-hybridization between different market classes calls for a different way of classifying genotypes.

The marker set used in the current study was very informative and led to the identification of three distinct clusters, well separated from each other. In addition, the first three PCs of the PCA plot explained more than 70% genetic variation among the genotypes. The high level of informativeness of the markers could be due to the use of the KHUFU pipeline, which is accurate in calling SNPs at even low coverage in complex genomes (Korani et al., 2021). Very few studies in groundnut have reported such highly informative markers used for diversity analysis, especially among studies done for African genotypes. The closest in comparison was a recent study that reported 67.5% genetic variation from the first three PCs after studying 200 genotypes from the African core set (Achola et al., 2023) using 7523 SNP markers called from the Thermo Fisher SNP array Axiom Arachis2 with 48,000 SNPs (Clevenger et al., 2018; Korani et al., 2019). Informative markers are not only important in determining diversity and population structure; they are also important for QC, accurate association, and linkage mapping. The marker set will, therefore, be ideal for future development of QC and mid-density panel (MDP) markers for African breeding programs. The distinct clusters generated in our study will be useful in reclassifying the different breeding lines to the correct market classes.

LD decayed at a distance of ~2.8 Mb, which is comparable to a recent study in groundnut that involved a MAGIC population (Wankhade et al., 2023) but much slower in comparison to other studies that used comparable datasets (Achola et al., 2023; Oteng-Frimpong et al., 2023). LD is the nonrandom association between alleles at different loci. A decline in LD is expected with increase in genetic distance between loci, subsequently leading to LD decay. Several factors have been reported to affect LD decay including mutation, population, selection, mating patterns, genetic drift, and migration (Flint-Garcia et al., 2003). Slow LD decay is typical in a self-pollinating crop such as groundnut due to severe reduction in effective recombination with increased generations of selfing. In addition, the molecular marker set used to calculate LD decay point is critical, with recommendations for marker minor allele frequency of >0.05 in groundnut (Otyama et al., 2019). For better association mapping resolution, the average distance between markers should be smaller than the LD decay distance (Breseghello & Sorrells, 2006). Indeed, the density of markers in our study was 9.26 SNPs/Mb, making the average distance between markers much less than the LD decay distance of 2.8 Mb. Future studies will need to comprehensively determine the factors affecting LD decay distance across the genome for different datasets used for association mapping in groundnut.

4.4 | Marker-trait associations and candidate gene identification

Association mapping is considered an efficient method for genetic analysis of complex traits (Fahrenkrog et al., 2017; Han et al., 2016) such as earliness and yield. We identified SNP markers associated with both traits using breeding lines from Africa. Some of the quantitative trait loci (QTLs) reported here were consistent with previous studies. Liang et al. (2020) reported a QTL on chromosome 18 (earlier designated as B08) for initial flowering date after characterizing a recombinant inbred line mapping population. While studying maturity index, which is a trait related to flowering time, Hake et al. (2018) also reported QTLs on chromosomes 8 (A08), 11 (B01), 13 (B03), and 18 (B08), among others, which are consistent with the results reported in the current study.

We have identified PY QTLs on chromosomes 3, 10, and 15 which were country-specific. QTLs on chromosomes 3 and 15 were detected across Ugandan locations, while the QTL on chromosome 10 was only detected across Malawi locations. Country-specific QTL differences reported for PY in comparison to flowering time suggest that PY is more sensitive to environmental differences than flowering time. Yield traits are complex and controlled by many genes in comparison to flowering time. This difference was also reported in the heritability, where DTF reported heritability of 70% across

locations, while that of the PY-related traits were 40%. Similar results were reported in rice after studying yield-related traits across nine locations in Asia, in which heading date reported maximum common QTLs across locations (Hittalmani et al., 2003).

Despite not identifying common QTLs across the two countries for yield, some of the QTLs detected for PY are consistent with previous studies. Gangurde et al. (2020) reported a major QTL for pod weight on chromosome 15 (B05). Both Chavarro et al. (2020) and Chen et al. (2016) reported a pod width QTL on chromosome 10, while Chavarro et al. (2020) and Z. Wang et al. (2018) reported a HSW QTL on chromosome 3. Identification of stable yielding genotypes across the four locations provides an opportunity for developing mapping populations that can be used in the future to detect stable yield QTLs across the region. Biparental populations developed using a stable donor promise to facilitate detection of more stable QTLs, as has been reported for chromosomes 5 and 7 (Luo et al., 2018).

Both earliness and yield are very important traits in Africa given the predominantly rain-fed nature of smallholder groundnut farming and the increasing effects of climate change. Identification of markers associated with important agronomic traits would enhance breeding efficiency and lead to faster release of varieties in the future. The quantitative nature of these traits, however, calls for thorough validation of markers detected through repeat trials across several seasons as well as development of biparental mapping populations. The markers detected here can be incorporated into the MDP for potential use in future GS.

5 | CONCLUSION

This study demonstrated the power of testing breeding lines across different countries and exploited the application of association mapping for the same dataset. Our findings form an important baseline and are a proof of concept for similar studies in groundnut in Africa. Better characterization of target testing environments across Africa will be necessary to define representative environments for traits of interest across regions. The recent creation of the Groundnut Improvement Network for Africa (Conde et al., 2023) will provide a great platform for selecting representative environments and fundraising for joint research across the continent. Our findings on marker trait association will pave the way for marker-assisted breeding and GS for earliness and yield once appropriate validations are done.

AUTHOR CONTRIBUTIONS

Velma Okaron: Conceptualization; investigation; methodology; writing—original draft; writing—review and editing. **James Mwololo**: Conceptualization; funding acquisition; writing—review and editing. **Davis M. Gimode**: Data curation; formal analysis; methodology; validation; writing—review and editing. **David K. Okello**: Investigation; methodology; supervision; writing—review and editing. **Millicent Avosa**: Methodology. **Josh Clevenger**: Software. **Walid Korani**: Formal analysis; software. **Mildred Ochwo Ssemakula**: Supervision; writing—review and editing. **Thomas L. Odong**: Supervision. **Damaris A. Odeny**: Conceptualization; data curation; formal analysis; funding acquisition; investigation; supervision; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

This work was largely supported by FAO through the benefit sharing fund of the International Treaty on Plant Genetics for Food and Agriculture. Supplemental funding was provided by the Scholarship Advert for Short Term Academic Mobility (SCIFSA) and the Regional Universities' Forum for Capacity Building in Agriculture (RUFORUM) to the first author.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this published article.

ORCID

Damaris A. Odeny https://orcid.org/0000-0002-3629-3752

REFERENCES

Achola, E., Wasswa, P., Fonceka, D., Clevenger, J. P., Bajaj, P., Ozias-Akins, P., Rami, J. F., Deom, C. A., Hoisington, D. A., Edema, R., Odeny, D. A., & Okello, D. K. (2023). Genome-wide association studies reveal novel loci for resistance to groundnut rosette disease in the African core groundnut collection. *Theoretical and Applied Genetics*, *136*(3), Article 35. https://doi.org/10.1007/s00122-023-04259-4

Aggarwal, V. D., Pastor-Corrales, M. A., Chirwa, R. M., & Buruchara, R. A. (2004). Andean beans (*Phaseolus vulgaris* L.) with resistance to the angular leaf spot pathogen (*Phaeoisariopsis griseola*) in southern and eastern Africa. *Euphytica*, *136*, 201–210. https://doi.org/10.1023/B:EUPH.0000030671.03694.bb

Badigannavar, A. M., Kale, D. M., & Murty, G. S. S. (2002). Genetic base and diversity in groundnut genotypes. *Plant Breeding*, *121*(4), 348–353. https://doi.org/10.1046/j.1439-0523.2002.00710.x

Badu-Apraku, B., Lum, A. F., Fakorede, M. A., Menkir, A., Chabi, Y., Abdulai, M., Jacob, S., & Agbaje, S. (2008). Performance of early maize cultivars derived from recurrent selection for grain yield and *Striga* resistance. *Crop Science*, *48*, 99–112. https://doi.org/10.2135/cropsci2007.01.0060

Bansal, R. K., Gondaliya, V. K., & Shaikh, A. S. (2017). A review of the status of the groundnut production and export of India. *Indian Journal of Economics and Development*, *13*(2), 369–374. https://doi.org/10.5958/2322-0430.2017.00190.1

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

Bates, O. K., Ollier, S., & Bertelsmeier, C. (2020). Smaller climatic niche shifts in invasive than non-invasive alien ant species. *Nature Communications*, *11*(1), Article 5213. https://doi.org/10.1038/s41467-020-19031-1

Bekele, G., Dechassa, N., & Tana, T. (2022). Effect of inorganic and organic fertilizers on productivity of groundnut (*Arachis hypogaea* L.) varieties in East Hararghe, Eastern Ethiopia. *Oil Crop Science*, *7*(3), 112–121. https://doi.org/10.1016/j.ocsci.2022.07.002

Breseghello, F., & Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*, *172*, 1165–1177. https://doi.org/10.1534/genetics.105.044586

Browning, B. L., & Browning, S. R. (2016). Genotype imputation with millions of reference samples. *The American Journal of Human Genetics*, *98*(1), 116–126. https://doi.org/10.1016/j.ajhg.2015.11.020

Casanoves, F., Baldessari, J., & Balzarini, M. (2005). Evaluation of multienvironment trials of peanut cultivars. *Crop Science*, *45*(1), 18–26. https://doi.org/10.2135/cropsci2005.0018

Chapu, I., Okello, D. K., Okello, R. C. O., Odong, T. L., Sarkar, S., & Balota, M. (2022). Exploration of alternative approaches to phenotyping of late leaf spot and groundnut rosette virus disease for groundnut breeding. *Frontiers in Plant Science*, *13*, Article 912332. https://doi.org/10.3389/fpls.2022.912332

Chavarro, C., Chu, Y., Holbrook, C., Isleib, T., Bertioli, D., Hovav, R., Butts, C., Lamb, M., Sorensen, R., Jackson, S. A., & Ozias-Akins, P. (2020). Pod and seed trait QTL identification to assist breeding for peanut market preferences. *G3: Genes, Genomes, Genetics*, *10*(7), 2297–315. https://doi.org/10.1534/g3.120.401147

Chen, W., Jiao, Y., Cheng, L., Huang, L., Liao, B., Tang, M., Ren, X., Zhou, X., & Chen, Y. (2016). Quantitative trait locus analysis for pod- and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.). *BMC Genetics*, *17*, 25. https://pubmed.ncbi.nlm.nih.gov/26810040/

Clevenger, J., Chu, Y., Chavarro, C., Botton, S., Culbreath, A., Isleib, T. G., Holbrook, C. C., & Ozias-Akins, P. (2018). Mapping late leaf spot resistance in peanut (*Arachis hypogaea* L.) using QTL-seq reveals markers for marker-assisted selection. *Frontiers in Plant Science*, *9*, Article 83. https://doi.org/10.3389/fpls.2018.00083

Conde, S., Rami, J. F., Okello, D. K., Sambou, A., Muitia, A., Oteng-Frimpong, R., Makweti, L., Sako, D., Faye, I., Chintu, J., & Coulibaly, A. M. (2023). The groundnut improvement network for Africa (GINA) germplasm collection: A unique genetic resource for breeding and gene discovery. *G3: Genes, Genomes, Genetics*, *24*, jkad244. https://doi.org/10.1093/g3journal/jkad244

Cook, K. H., & Vizy, E. K. (2012). Impact of climate change on mid-twenty-first century growing seasons in Africa. *Climate Dynamics*, *39*, 2937–2955. https://doi.org/10.1007/s00382-012-1324-1

Dalló, S. C., Zdziarski, A. D., Woyann, L. G., Milioli, A. S., Zanella, R., Conte, J., & Benin, G. (2019). Across year and year-by-year GGE biplot analysis to evaluate soybean performance and stability in multi-environment trials. *Euphytica*, *215*, Article 113. https://doi.org/10.1007/s10681-019-2438-x

Dhillon, S. K., Gurdeep, S., Gill, B. S., & Paramjit, S. (2009). Stability analysis for grain yield and its components in soybean (*Glycine max* (L.) Merrill.). *Crop Improvement*, *36*(1), 55–58.

Essandoh, D. A., Odong, T., Okello, D. K., Fonceka, D., Nguepjop, J., Sambou, A., Ballén-Taborda, C., Chavarro, C., Bertioli, D. J., & Leal-Bertioli, S. C. (2022). Quantitative trait analysis shows the potential for alleles from the wild species *Arachis batizocoi* and *A. duranensis* to improve groundnut disease resistance and yield in East Africa. *Agronomy*, *12*(9), 2202. https://doi.org/10.3390/agronomy12092202

Estefan, G. (2013). *Methods of soil, plant, and water analysis: A manual for the West Asia and North Africa region*. ICARDA. https://hdl.handle.net/20.500.11766/7512

Fahrenkrog, A. M., Neves, L. G., Resende, M. F. R. Jr., Vazquez, A. I., de Los Campos, G., Dervinis, C., & Kirst, M. (2017). Genome-wide association study reveals putative regulators of bioenergy traits in *Populus deltoides*. *New Phytologist*, *213*(2), 799–811. https://doi.org/10.1111/nph.14154

FAOSTAT. (2021). *Statistics*. FAO. http://www.fao.org/statistics/en/

Flint-Garcia, S. A., Thornsberry, J. M., & Buckler, E. S., IV. (2003). Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, *54*(1), 357–374. https://doi.org/10.1146/annurev.arplant.54.031902.134907

Gangurde, S. S., Wang, H., Yaduru, S., Pandey, M. K., Fountain, J. C., Chu, Y., Isleib, T., Holbrook, C. C., Xavier, A., Culbreath, A. K., & Ozias-Akins, P. (2020). Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnology Journal*, *18*(6), 1457–1471. https://doi.org/10.1111/pbi.13311

Gauch, H. G., & Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Science*, *37*, 311–326. https://doi.org/10.2135/cropsci1997.0011183x003700020002x

Greveniotis, V., Bouloumpasi, E., Zotis, S., Korkovelos, A., Kantas, D., & Ipsilandis, C. G. (2023). Genotype-by-environment interaction analysis for quantity and quality traits in faba beans using AMMI, GGE models, and stability indices. *Plants*, *12*(21), 3769. https://doi.org/10.3390/plants12213769

Hake, A. A., Shirasawa, K., Yadawad, A., Nadaf, H. L., Gowda, M. V. C., & Bhat, R. S. (2018). Genome-wide structural mutations among the lines resulting from genetic instability in peanut (*Arachis hypogaea* L.). *Plant Gene*, *13*, 1–7. https://doi.org/10.1016/j.plgene.2017.11.001

Han, Y. P., Zhao, X., Liu, D. Y., Li, Y. H., Lightfoot, D. A., Yang, Z. J., Zhao, L., Zhou, G., Wang, Z., Huang, L., Zhang, Z., Qiu, L., Zheng, H., & Li, W. (2016). Domestication footprints anchor genomic regions of agronomic importance in soybeans. *New Phytologist*, *209*, 871–884. https://doi.org/10.1111/nph.13626

Hittalmani, S., Huang, N., Courtois, B., Venuprasad, R., Shashidhar, H. E., Zhuang, J. Y., Zheng, K. L., Liu, G. F., Wang, G. C., Sidhu, J. S., & Srivantaneeyakul, S. (2003). Identification of QTL for growth-and grain yield-related traits in rice across nine locations of Asia. *Theoretical and Applied Genetics*, *107*, 679–690. https://doi.org/10.1007/s00122-003-1269-1

Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2019). BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience*, *8*(2), giy154. https://doi.org/10.1093/gigascience/giy154

Hue, N. V., Uchida, R., & Ho, M. C. (2000). Sampling and analysis of soils and plant tissues: How to take representative samples, how the samples are tested. In J. A. Silva & R. S. Uchida (Eds.), *Plant nutrient management in Hawaii's soils: Approaches for tropical and subtropical agriculture* (pp. 23–30). University of Hawaii. http://hdl.handle.net/10125/1910

Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*, *47*(7), 314–318. https://doi.org/10.2134/agronj1955.00021962004700470

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), Article 94. https://doi.org/10.1186/1471-2156-11-94

Jordan, K. W., Wang, S., Lun, Y., Gardiner, L. J., MacLachlan, R., Hucl, P., Wiebe, K., Wong, D., Forrest, K. L., IWGS Consortium, Sharpe, A. G., Sidebottom, C. H. D., Hall, N., Toomajian, C., Close, T., Dubcovsky, J., Akhunova, A., Talbert, L., … Akhunov, E. (2015). A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biology*, *16*, 1–18. https://doi.org/10.1186/s13059-015-0606-4

Kassa, Y., Asea, G., Demissew, A. K., Ligeyo, D., Demewoz, N., Saina, E., Serumaga, J., Twumasi-Afriyie, S., Opio, F., Rwomushana, I., Gelase, N., Gudeta, N., Wondimu, F., Solomon, A., Habtamu, Z., Andualem, W. B. A., Habte, J., & Mduruma, Z. (2013). Stability in performance of normal and nutritionally enhanced highland maize hybrid genotypes in Eastern Africa. *Asian Journal of Plant Sciences*, *12*(2), 51–60. https://doi.org/10.3923/ajps.2013.51.60

Kebede, B. A., & Getahun, A. (2017). Adaptability and stability analysis of groundnut genotypes using AMMI model and GGE-biplot. *Journal of Crop Science and Biotechnology*, *20*, 343–349. https://doi.org/10.1007/s12892-017-0061-0

Korani, W., Clevenger, J. P., Chu, Y., & Ozias-Akins, P. (2019). Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants. *The Plant Genome*, *12*(1), 180023. https://doi.org/10.3835/plantgenome2018.05.0023

Korani, W., O'Connor, D., Chu, Y., Chavarro, C., Ballen, C., Guo, B., Ozias-Akins, P., Wright, G., & Clevenger, J. (2021). *De novo* QTL-seq identifies loci linked to blanchability in peanut (*Arachis hypogaea*) and refines previously identified QTL with low coverage sequence. *Agronomy*, *11*(11), 2201. https://doi.org/10.3390/agronomy11112201

Lal, C., Ajay, B. C., Chikani, B. M., & Gor, H. K. (2019). AMMI and GGE biplot analysis to evaluate the phenotypic stability of recombinant inbred lines (RILs) of peanut under mid-season water stress conditions. *Indian Journal of Genetics and Plant Breeding*, *79*(2), 420–426. https://www.isgpb.org/journal/index.php/IJGPB/article/view/2843

Legendre, P., & Legendre, L. (1998). *Numerical ecology*. Elsevier Science B.V.

Liang, W. A., Yang, X. L., Cui, S. L., Wang, J. H., Hou, M. Y., Mu, G. J., Li, Z. C., & Liu, L. F. (2020). Identification of main effect and epistatic QTLs controlling initial flowering date in cultivated peanut (*Arachis hypogaea* L.). *Journal of Integrative Agriculture*, *19*(10), 2383–2393. https://doi.org/10.1016/S2095-3119(20)63211-7

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genetics*, *12*, e1005767. https://doi.org/10.1371/journal.pgen.1005767

Luo, H., Guo, J., Ren, X., Chen, W., Huang, L., Zhou, X., Chen, Y., Liu, N., Xiong, F., Lei, Y., & Liao, B. (2018). Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, *131*, 267–282. https://doi.org/10.1007/s00122-017-3000-7

Makinde, S. C. O., & Ariyo, O. J. (2011). Analysis of genotype × environment interaction of groundnut (*Arachis hypogaea* L.). *Malaysian Applied Biology*, *40*(2), 19–26.

Mekontchou, T., Ngueguim, M., & Fobasso, M. (2006). Stability analyse for yield and yield components of selected peanut breeding lines *Arachis hypogaea* L. in the North Province of Cameroon. *Tropicultura*, *24*(2), 90.

Meta, H. R., & Monpara, B. A. (2010). Genetic variation and trait relationships in summer groundnut (*Arachis hypogaea* L.). *Journal of Oilseeds Research*, *27*(1), 8–11.

Mohammadi, R., Jafarzadeh, J., Armion, M., Hatamzadeh, H., & Roohi, E. (2023). Clustering stability methods towards selecting best performing and stable durum wheat genotypes. *Euphytica*, *219*(10), Article 109. https://doi.org/10.1007/s10681-023-03237-7

Mupangwa, W., Thierfelder, C., Cheesman, S., Nyagumbo, I., Muoni, T., Mhlanga, B., Mwila, M., Sida, T. S., & Ngwira, A. (2020). Effects of maize residue and mineral nitrogen applications on maize yield in conservation-agriculture-based cropping systems of Southern Africa. *Renewable Agriculture and Food Systems*, *35*(3), 322–335. https://doi.org/10.1017/S174217051900005X

Nduwumuremyi, A., Melis, R., Shanahan, P., & Theodore, A. (2017). Interaction of genotype and environment effects on important traits of cassava (*Manihot esculenta* Crantz). *The Crop Journal*, *5*(5), 373–386. https://doi.org/10.1016/j.cj.2017.02.004

Negash, A. W., Mwambi, H., Zewotir, T., & Taye, G. (2013). Additive main effects and multiplicative interactions model (AMMI) and genotype main effect and genotype by environment interaction (GGE) biplot analysis of multi-environmental wheat variety trials. *African Journal of Agricultural Research*, *8*(12), 1033–1040. https://doi.org/10.5897/AJAR2012.6648

Okalebo, J. R., Gathua, K. W., & Woomer, P. L. (2002). *Laboratory methods of soil and plant analysis: A working manual* (2nd ed.). Sacred African Publishers.

Okello, D. K., Biruma, M., & Deom, C. M. (2010). Overview of groundnuts research in Uganda: Past, present and future. *African Journal of Biotechnology*, *9*(39), 6448–6459.

Okello, D. K., Monyo, E., Michael, D. C., Jane, I., & Herbert, Kefa, O. (2013). *Groundnut production guide for Uganda: Recommended practices for farmers*. National Agricultural Research Organisation. https://nru.uncst.go.ug/handle/123456789/4549

Okori, P., Charlie, H., Mwololo, J., Munthali, W., Kachulu, L., Monyo, E., Muitia, A., Mponda, O., Okello, D. K., Makweti, L., & Siambi, M. (2019). Genotype-by-environment interactions for grain yield of Valencia groundnut genotypes in East and Southern Africa. *Australian Journal of Crop Science*, *13*(12), 2030–2037. https://doi.org/10.21475/ajcs.19.13.12.p2039

Olivoto, T., & Lúcio, A. D. C. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*, *11*(6), 783–789. https://doi.org/10.1111/2041-210X.13384

Oteng-Frimpong, R., Karikari, B., Sie, E. K., Kassim, Y. B., Puozaa, D. K., Rasheed, M. A., Fonceka, D., Okello, D. K., Balota, M., Burow, M., & Ozias-Akins, P. (2023). Multi-locus genome-wide association studies reveal genomic regions and putative candidate genes associated with leaf spot diseases in African groundnut (*Arachis hypogaea* L.) germplasm. *Frontiers in Plant Science*, *13*, 1076744. https://doi.org/10.3389/fpls.2022.1076744

Otyama, P. I., Wilkey, A., Kulkarni, R., Assefa, T., Chu, Y., Clevenger, J., O'Connor, D. J., Wright, G. C., Dezern, S. W., MacDonald, G. E., Anglin, N. L., Cannon, E. K. S., Ozias-Akins, P., & Cannon, S.

B. (2019). Evaluation of linkage disequilibrium, population structure, and genetic diversity in the US peanut mini core collection. *BMC Genomics*, *20*(1), Article 481. https://doi.org/10.1186/s12864-019-5824-9

Pandey, M. K., Chaudhari, S., Jarquin, D., Janila, P., Crossa, J., Patil, S. C., Sundravadana, S., Khare, D., Bhat, R. S., Radhakrishnan, T., Hickey, J. M., & Varshney, R. K. (2020). Genome-based trait prediction in multi-environment breeding trials in groundnut. *Theoretical and Applied Genetics*, *133*, 3101–3117. https://doi.org/10.1007/s00122-020-03658-1

Pandey, M. K., Upadhyaya, H. D., Rathore, A., Vadez, V., Sheshshayee, M. S., Sriswathi, M., Govil, M., Kumar, A., Gowda, M. V. C., Sharma, S., Hamidou, F., Kumar, V. A., Khera, P., & Varshney, R. K. (2014). Genome wide association studies for 50 agronomic traits in peanut using the 'reference set' comprising 300 genotypes from 48 countries of the semi-arid tropics of the world. *PLoS ONE*, *9*(8), e105228. https://doi.org/10.1371/journal.pone.0105228

Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, *161*(1–2), 209–228. https://doi.org/10.1007/s10681-007-9449-8

Pobkhunthod, N., Authapun, J., Chotchutima, S., Rungmekarat, S., Kittipadakul, P., Duangpatra, J., & Chaisan, T. (2022). Multilocation yield trials and yield stability evaluation by GGE biplot analysis of promising large-seeded peanut lines. *Frontiers in Genetics*, *4*(13), Article 876763. https://doi.org/10.3389/fgene.2022.876763

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7

Rana, S. S., & Kumar, S. (2014). *Research techniques in agronomy* (Vol. 64). Department of Agronomy, College of Agriculture, CSK Himachal Pradesh Krishi Vishvavidyalaya.

Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., Kresovich, S., Goodman, M. M., & Buckler, E. S., IV. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences*, *98*(20), 11479–11484. https://doi.org/10.1073/pnas.201394398

Rosen, C. J., Bierman, P. M., & Eliason, R. (2008). *Soil test interpretations and fertilizer management for lawns, turf, gardens, and landscape plants* (Bulletin No. BU-01731). University of Minnesota Extension Service. https://hdl.handle.net/11299/49265

Sanogo, O., Tongoona, P. B., Ofori, K., Offei, S. K., & Desmae, H. (2019). Evaluation of yield and yield components of some groundnut genotypes under rainfed condition in Mali using biplot analysis. *African Journal of Agricultural Research*, *14*, 1904–1912. https://doi.org/10.5897/AJAR2018.13776

Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., & Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics*, *44*(7), 825–830. https://doi.org/10.1038/ng.2314

Sewagegne, T., Taddesse, L., Mulugeta, B., & Mitiku, A. (2013). Genotype by environment interaction and grain yield stability analysis of rice (*Oryza sativa* L.) genotypes evaluated in north western Ethiopia. *Net Journal of Agricultural Science*, *1*(1), 10–16.

Shahriari, Z., Heidari, B., & Dadkhodaie, A. (2018). Dissection of genotype × environment interactions for mucilage and seed yield in *Plantago* species: Application of AMMI and GGE biplot analyses. *PLoS ONE*, *13*(5), e0196095. https://doi.org/10.1371/journal.pone.0196095

Shaibu, A. S., Sneller, C., Motagi, B. N., Chepkoech, J., Chepngetich, M., Miko, Z. L., Isa, A. M., Hakeem, A., Ajeigbe, H. A., & Mohammed, S. G. (2020). Genome-wide detection of SNP markers associated with four physiological traits in groundnut (*Arachis hypogaea* L.) mini core collection. *Agronomy*, *10*(2), 192. https://doi.org/10.3390/agronomy10020192

Shojaei, S. H., Mostafavi, K., Bihamta, M. R., Omrani, A., Mousavi, S. M., Illés, Á., Bojtor, C., & Nagy, J. (2022). Stability on maize hybrids based on GGE biplot graphical technique. *Agronomy*, *12*(2), 394. https://doi.org/10.3390/agronomy12020394

Silva, M. D. A. (2008). Genotype × environment interaction and phenotypic stability in sugarcane under a twelve-months planting cycle. *Bragantia*, *67*, 109–117. https://doi.org/10.1590/S0006-87052008000100013

Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., Su, Z., Pan, Y., Liu, D., Lipka, A. E., Buckler, E. S., & Zhang, Z. (2016). GAPIT version 2: An enhanced integrated tool for genomic association and prediction. *The Plant Genome*, *9*(2), plantgenome2015.11.0120. https://doi.org/10.3835/plantgenome2015.11.0120

Tossim, H. A., Nguepjop, J. R., Diatta, C., Sambou, A., Seye, M., Sane, D., Rami, J. F., & Fonceka, D. (2010). Assessment of 16 peanut (*Arachis hypogaea* L.) CSSLs derived from an interspecific cross for yield and yield component traits: QTL validation. *Agronomy*, *10*(4), 583. https://doi.org/10.3390/agronomy10040583

USDA-NASS. (2021). *Agricultural statistics*. USDA-National Agricultural Statistics Service.

Vaezi, B., Pour-Aboughadareh, A., Mohammadi, R., Mehraban, A., Hossein-Pour, T., Koohkan, E., Ghasemi, S., Moradkhani, H., & Siddique, K. H. (2019). Integrating different stability models to investigate genotype × environment interactions and identify stable and high-yielding barley genotypes. *Euphytica*, *215*, Article 63. https://doi.org/10.1007/s10681-019-2386-5

Van Reeuwijk, L. P. (2002). *Procedures for soil analysis* (6th ed.). ISRIC.

Varshney, R. K., Mohan, S. M., Gaur, P. M., Gangarao, N. V. P. R., Pandey, M. K., Bohra, A., Shrikant, L., Sawargaonkar, S. L., Chitikinenia, A., Paul, K., Kimurto, P. K., Janila, P., Saxena, K. B., Asnake Fikre, A., Mamta Sharma, M., Rathore, A., Aditya Pratap, A., Tripathi, S., Datta, S., … Gowda, C. L. L. (2013). Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. *Biotechnology Advances*, *31*(8), 1120–1134. https://doi.org/10.1016/j.biotechadv.2013.01.001

Wang, B., Lin, Z., Li, X., Zhao, Y., Zhao, B., Wu, G., Ma, X., Wang, H., Xie, Y., Li, O., Song, G., Kong, D., Zheng, Z., Wei, H., Shen, R., Chen, C., Meng, Z., Wang, T., Li, Y., … Wang, H. (2020). Genome-wide selection and genetic improvement during modern maize breeding. *Nature Genetics*, *52*(6), 565–571. https://doi.org/10.1038/s41588-020-0616-3

Wang, J., Yan, C., Li, Y., Li, C., Zhao, X., Yuan, C., Li, C., Zhao, X., Yuan, C., Sun, Q., & Shan, S. (2019). GWAS discovery of candidate genes for yield-related traits in peanut and support from earlier QTL mapping studies. *Genes*, *10*(10), 803. https://doi.org/10.3390/genes10100803

Wang, Z., Huai, D., Zhang, Z., Cheng, K., Kang, Y., Wan, L., Yan, L., Jiang, H., Lei, Y., & Liao, B. (2018). Development of a high-density

genetic map based on specific length amplified fragment sequencing and its application in quantitative trait loci analysis for yield-related traits in cultivated peanut. *Frontiers in Plant Science*, *9*, Article 827. https://doi.org/10.3389/fpls.2018.00827

Wankhade, A. P., Chimote, V. P., Viswanatha, K. P., Yadaru, S., Deshmukh, D. B., Gattu, S., Sudini, H. K., Deshmukh, M. P., Shinde, V. S., Vemula, A. K., & Pasupuleti, J. (2023). Genome-wide association mapping for LLS resistance in a MAGIC population of groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, *136*(3), Article 43. https://doi.org/10.1007/s00122-023-04256-7

Yan, W. (2001). GGEbiplot—A windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agronomy Journal*, *93*, 1111–1118. https://doi.org/10.2134/agronj2001.9351111x

Yan, W., & Kang, M. S. (2002). GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press.

Yan, W., Nilsen, K. T., & Beattie, A. (2023). Mega-environment analysis and breeding for specific adaptation. *Crop Science*, *63*(2), 480–494.

Yan, W., & Holland, J. B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica*, *171*(3), 355–369. https://doi.org/10.1007/s10681-009-0030-5

Yan, W., & Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science*, *86*(3), 623–645. https://doi.org/10.4141/P05-169

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., Yuan, X., Zhu, M., Zhao, S., Li, X., & Liu, X. (2021). rMVP: A memory-efficient, visualization enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics & Bioinformatics*, *19*(4), 619–628. https://doi.org/10.1016/j.gpb.2020.10.007

Zhang, X., Zhang, J., He, X., Wang, Y., Ma, X., & Yin, D. (2017). Genome-wide association study of major agronomic traits related to domestication in peanut. *Frontiers in Plant Science*, *8*, Article 1611. https://doi.org/10.3389/fpls.2017.01611

Zhao, H., Tian, R., Xia, H., Li, C., Li, G., Li, A., Zhang, X., Zhou, X., Ma, J., Huang, H., & Zhang, K. (2022). High-density genetic variation map reveals key candidate loci and genes associated with important agronomic traits in peanut. *Frontiers in Genetics*, *13*, Article 845602. https://doi.org/10.3389/fgene.2022.845602

Zhao, J., Huang, L., Ren, X., Pandey, M. K., Wu, B., Chen, Y., Zhou, X., Chen, W., Xia, Y., Li, Z., Luo, H., Lei, Y., Varshney, R. K., Liao, B., & Jiang, H. (2017). Genetic variation and association mapping of seed-related traits in cultivated peanut (*Arachis hypogaea* L.) using single-locus simple sequence repeat markers. *Frontiers in Plant Science*, *8*, Article 2105. https://doi.org/10.3389/fpls.2017.02105

Zhou, X., Guo, J., Pandey, M. K., Varshney, R. K., Huang, L., Luo, H., Liu, N., Chen, W., Lei, Y., Liao, B., & Jiang, H. (2021). Dissection of the genetic basis of yield-related traits in the Chinese peanut mini-core collection through genome-wide association studies. *Frontiers in Plant Science*, *12*, Article 637284. https://doi.org/10.3389/fpls.2021.637284

Zongo, A., Nana, A. T., Sawadogo, M., Konate, A. K., Sankara, P., Ntare, B. R., & Desmae, H. (2017). Variability and correlations among groundnut populations for early leaf spot, pod yield, and agronomic traits. *Agronomy*, *7*(3), 52. https://doi.org/10.3390/agronomy7030052

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Okaron, V., Mwololo, J., Gimode, D. M., Okello, D. K., Avosa, M., Clevenger, J., Korani, W., Ssemakula, M. O., Odong, T. L., & Odeny, D. A. (2024). Using cross-country datasets for association mapping in *Arachis hypogaea* L. *The Plant Genome*, *17*, e20515. https://doi.org/10.1002/tpg2.20515