

Genotype-by-environment interactions for grain yield of Valencia groundnut genotypes in East and Southern Africa

P. Okori^{1*}, H. Charlie¹, J. Mwololo¹, W. Munthali¹, L. Kachulu¹, E. Monyo¹, A. Muitia², O. Mponda³, D. Kalule-Okello⁴, L. Makweti⁵, and M. Siambi¹

¹International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) Chitedze Agricultural Research Station, P.O. Box 1096, Lilongwe, Malawi

²Mozambique Institute of Agricultural Research (IIAM) FPLM Avenue, Km 7 Via Corrane, Nampula Province, Mozambique. 622 Nampula Nampula Mozambique

³Naliendele Agricultural Research Institute PO Box 509, Mtwara, United Republic of Tanzania

⁴National Semi-Arid Resources Research Institute (NaSSARI), P.O. Box 56, Soroti-Uganda

⁵Zambia Agricultural Research Institute (ZARI)-Msekera Research Station P.O Box 510089, Chipata, Zambia

*Corresponding author: J.Mwololo@cgiar.org; P.Okori@cgiar.org

Abstract

Grain yield is a quantitatively inherited trait in groundnut (*Arachis hypogaea* L.) and subject to genotype by environment interactions. Groundnut varieties show wide variation in grain yield across different agro-ecologies. The objectives of this study were to evaluate Valencia groundnut genotypes for yield stability and classify environments to devise appropriate breeding strategies. Seventeen multi-location trials were conducted in six countries, viz., Malawi, Tanzania, Uganda, Zimbabwe, Mozambique and Zambia, from 2013 to 2016. The experiments were laid out following a resolvable incomplete block design, with two replications at each location (hereafter referred to as 'environments') using 14 test lines and two standard checks. The additive main effects and multiplicative interaction (AMMI) analysis was conducted. Variation attributable to environments, genotypes and genotype × environment interaction for grain yield was highly significant ($P < 0.001$). Genotype, environment and genotype × environment interactions accounted for 7%, 53 % and 40% of the total sum of squares respectively. Superior-performing genotypes possessing high to moderate adaptability and stability levels included ICGV-SM 0154, ICGV-SM 07539, ICGV-SM 07536, ICGV-SM 7501, ICGV-SM 99568 and ICGV SM 07520. Nachingwea 2013 in Tanzania, Nakabango 2014 in Uganda and Chitedze 2015 in Malawi were the most representative and discriminative environments. Considering the implications of interactions for Valencia groundnut breeding in East and Southern Africa we propose that different varieties should be targeted for production in different environments and at the same time used for breeding in specific environments.

Keywords: Adaptation, AMMI, Breeding, Stability, Multi-environment trials, Discrimination ability, Representativeness.

Abbreviations: AEA_average environment axis , AMMI_additive main effects and multiplicative interaction, ANOVA_analysis of variance, E_environment, G_genotype, GGE_genotype and genotype × environment interaction, ICRISAT_International Crops Research Institute for the Semi-Arid Tropics, ICGV-SM_ICRISAT groundnut variety selected in Malawi, IPCA_interaction principal component analysis, Mals_meters above sea level, PCA_principal component analysis, Temp_min_minimum temperature Temp Max_maximum temperature , V_varieties.

Introduction

Groundnut is a versatile legume grown mainly for food including cooking oil and, in some cases, for fodder in sub-tropical and tropical environments. Cultivated groundnut has four major subspecies, *Arachis hypogaea* var: *hypogaea* (synonym = Virginia), *A. hypogaea* var: *fastigiata* (synonym = Valencia), *A. hypogaea* var: *vulgaris* (synonym = Spanish) and *A. hypogaea* var: *hirsuta* (synonym = Runner). Distribution of these subspecies is based on their production and market qualities, with most commonly grown varieties being Virginia, Spanish and Valencia in East and Southern Africa. Groundnut exhibits wide variation in grain yield and is grown

in many agroecologies (Fan et al., 2007). For production at a large scale in East and Southern Africa, it is imperative that, from farmers' standpoint, varieties with consistent performance across years (temporal stability) be developed and identified. From seed companies' viewpoint, varieties with consistent performance across locations (spatial stability) would be preferred. For the most part, groundnut is a cash and food crop, and therefore, a major component of farmers' livelihoods, especially in Africa's agrarian economies. Groundnut is mostly produced as a monocrop, but increasingly, farmers are producing it as an intercrop, to reduce risk while maximizing land productivity by

diversifying their production systems (Gan et al., 2011; Affholder et al., 2013; Jayne et al., 2014). Groundnut breeding program at ICRISAT for East and Southern Africa, based in Malawi, has developed a wide range of varieties for diverse cropping systems in three major agro-ecologies, i.e., low altitude agro-ecologies (400-600 masl) that receive up to 450 mm of rainfall per year; mid-altitude agro-ecologies (600-1001 masl) that receive up to 1000 mm of rainfall per year); and high altitude agro-ecologies (1001-1500 masl) that receive up to 1000 mm of rainfall annually. In Africa, Virginia and Spanish varieties are the predominant subspecies, grown partly because of a major focus on large-seeded genotypes needed for export markets. However, our engagement with farmers, via participatory variety selection, in East and Southern African countries, has revealed increasing demand for Valencia genotypes. Among other considerations, Valencia types are desired for their good peanut butter quality because of their sweet taste and short-duration maturity. The Valencia groundnut types were mostly produced in warm region of New Mexico but over time have been introduced into African countries. The Portuguese carried in the late 15th century the Valencia groundnut type varieties from the east coast of South America (Brazil) to Africa (Nigram, 2014). The low productivity in Africa in the past, can be explained by the fact that more focus on groundnuts was on the extraction of oil for domestic consumption and export since oil was the major item traded. However, the trend in the last decade has changed and edible groundnut is dominates the world market. In Africa, South Africa is the lead producer of edible groundnuts and for processing peanut butter such as the Valencia type. Thus, as farmers diversify their groundnut production systems to meet different domestic and market niche needs, plant breeding must deliver temporally stable, locally adapted, highly productive varieties.

Adoption of new crop varieties, for the most part, is influenced by yield. However, yield is a complex trait that is subject to modification by the cropping system/environment and genotype-by-environment interaction (GEI) (Kang, 1998). Strong GEI compounds selection of high yielding and stable genotypes because of differential reaction of the same genotype to varying agroecological conditions (Yau, 1995; Gauch, 2006). To improve selection precision, breeding programs use production data from different environments to identify adapted, consistent-performing genotypes, and at the same time, identify niche-specific genotypes for targeted variety commercialization (Yan, 2002; Blanche et al., 2007).

Multi-environment trials (METs) are recommended for evaluation of genotypes to identify varieties whose performance across environments is stable and predictable (Badu-Apraku et al., 2008). Several methods have been developed for assessment of GEI (Eberhart and Russell, 1966; Lin et al., 1986). Multivariate statistical analyses that explore multi-directional aspects of GEI are preferred to univariate analyses, because they extract additional information from GEI components (Gauch et al., 2008). The additive main effects and multiplicative interaction (AMMI) model is among the most powerful tools for analyzing $G \times E$ interactions (Gauch, 2006). The AMMI model integrates variance analysis of the main effects, i.e., genotypes and environments, with principal component analysis for

multiplicative effects of GE interactions (Oliveira and Godoy, 2006). A principal component model is fitted to residuals from the analysis of variance and the resulting interaction principal component analysis scores are calculated for genotypes and environments. Thus, AMMI analysis supports identification of the most stable and productive genotypes and further recommendation for ecology-specific cultivars (Pacheco et al., 2005). Yan et al. (2000) proposed the GGE-biplot method for graphical display of GE interaction. Environmental variation (E) explains most of the variation, whereas G and GE are usually smaller in comparison (Yan, 2002; Yan and Kang, 2003). However, only the G and GE interactions are relevant for genotype evaluation, particularly when GE interaction is repeatable (Yan et al., 2000; Yan et al., 2011). Moreover, GGE biplot analysis is an effective method for mega-environment analysis that allows visual examination of relationships among the test environments, genotypes and GE interactions (Yan et al., 2007). These approaches can be adopted in breeding programs to ensure delivery of superior varieties that are stable across environments/seasons. Limited researches have been done in environmental stratification for niche targeting in development of Valencia groundnut varieties. Environmental stratification can reduce costs by concentrating on locations which are more informative and discriminatory. Further, within the East and Southern Africa Region, only three Valencia varieties have been released, thus the need of superior improved varieties. To select adapted cultivars for diverse and/or specific production domains (Bernardo, 2002), ICRISAT's East and southern Africa groundnut breeding program evaluated performance of its elite Valencia groundnut varieties in multi-environments for yield stability and environmental stratification of testing locations across East and Southern Africa.

Results and discussion

Yield variability among new Valencia genotypes

Analysis of variance showed highly significant genotype, environment, and genotype \times environment interactions (GEI) for yield (supplementary Table 3). Similarly, principal components (PC1 and PC2) were highly significant. Genotypes, environment, PC1 and PC2 were accounted for 73% of treatment sum of squares. Therefore, the importance of genotype and environment was evident, an indication that AMMI model effectively partitioned the treatment sum of squares (Moreno-González et al., 2004, Gauch, 2006). A small sum of squares attributable to genotypes (G) (7% of treatment sum of squares) indicates that the contribution of G to the observed variation was small (Supplementary Table 3). The results also showed that the contributions of environment (E) and genotype \times environment ($G \times E$) to the variation were much larger (40 and 53%, respectively) than that of G (supplementary Table 3). This was an indication of differential response of test genotypes to the environments. These differences could be attributable to variations in temperature and precipitation, soil health and diseases, which are factors known to influence $G \times E$ interaction (Casonoves et al., 2005, Shahriari et al., 2018).

Table 1. Mean grain yield (Kg/ha), and principal component analysis (PCA) scores of sixteen (16) groundnut genotypes evaluated in 17 environments

Genotype	Yield (Kg/ha)	Ranking based on yield	¹ IPCAg	² IPCAg
ICG 405	486.5	16	9.7	-6.6
ICGV-SM 95741	548.0	15	5.6	-19.1
ICGV-SM 05521	595.3	14	13.6	2.5
ICGV-SM 07541	602.7	13	12.7	2.5
ICGV-SM 07510	612.6	12	5.9	6.0
ICGV-SM 07532	613.8	11	4.2	-3.6
ICGV-SM 07501	619.6	10	-1.0	-0.2
ICGV-SM 06687	626.6	9	12.9	15.0
ICGV-SM 99568	637.9	8	-12.7	1.2
JL 24	642.2	7	-1.1	-22.7
ICGV-SM 07536	667.7	6	8.5	1.0
ICGV-SM 07502	673.5	5	-14.4	1.5
ICGV-SM 07517	686.6	4	-5.8	16.4
ICGV-SM 07539	727.9	3	1.2	6.5
ICGV-SM 01514	772.3	2	-22.5	4.1
ICGV-SM 07520	784.4	1	-16.9	-4.4

¹PCA1 accounted for most of the variation in data set such that two dimensions picture of data given in biplot is a reasonable representation of the positions of genotypes; Large PCA1 values whether negative or positive indicate the adaptability of genotypes to different environments. ²PCA2 accounted for substantial variation in the data set too. Small PCA2 indicates stability of genotypes.

Table 2. Grouping of the environments and ranking of the best genotypes.

Environment	Score	Environment yield (Kg/ha)	Genotype ranking			
			1	2	3	4
E1	4.1	742.4	ICGV-SM 07539	ICGV-SM 07520	ICGV-SM 07536	ICGV-SM 06687
E10	-8.5	879.3	ICGV-SM 07520	ICGV-SM 01514	JL 24	ICGV-SM 07502
E11	5.4	415.8	JL 24	ICGV-SM 95741	ICGV-SM 07520	ICGV-SM 07536
E12	-12.3	836.1	CGV-SM 01514	ICGV-SM 07517	ICGV-SM 07520	ICGV-SM 07502
E13	-1.9	545.0	ICGV-SM 01514	ICGV-SM 07520	ICGV-SM 07539	ICGV-SM 07517
E14	-5.4	286.5	ICGV-SM 07520	ICGV-SM 01514	JL 24	ICGV-SM 07502
E15	-5.1	951.2	ICGV-SM 07517	ICGV-SM 01514	ICGV-SM 07539	ICGV-SM 06687
E16	-1.1	688.7	ICGV-SM 01514	ICGV-SM 07520	ICGV-SM 07539	ICGV-SM 07517
E17	-2.9	510.8	ICGV-SM 07520	ICGV-SM 01514	JL 24	ICGV-SM 07502
E18	4.7	723.0	JL 24	ICGV-SM 95741	ICGV-SM 07520	ICGV-SM 07536
E2	4.7	562.8	JL 24	ICGV-SM 07520	ICGV-SM 95741	ICGV-SM 07536
E3	23.7	686.6	ICGV-SM 06687	ICGV-SM 05521	ICGV-SM 07541	ICGV-SM 07536
E5	6.1	443.4	ICGV-SM 07539	ICGV-SM 06687	ICGV-SM 07536	ICGV-SM 07517
E6	-23.8	984.4	ICGV-SM 07520	ICGV-SM 01514	JL 24	ICGV-SM 07502
E7	-13.5	557.4	ICGV-SM 01514	ICGV-SM 07520	ICGV-SM 07502	ICGV-SM 99568
E8	10.1	562.6	JL 24	ICGV-SM 07536	ICGV-SM 05521	ICGV-SM 07541
E9	9.6	765.3	ICGV-SM 06687	ICGV-SM 07517	ICGV-SM 07539	ICGV-SM 07536

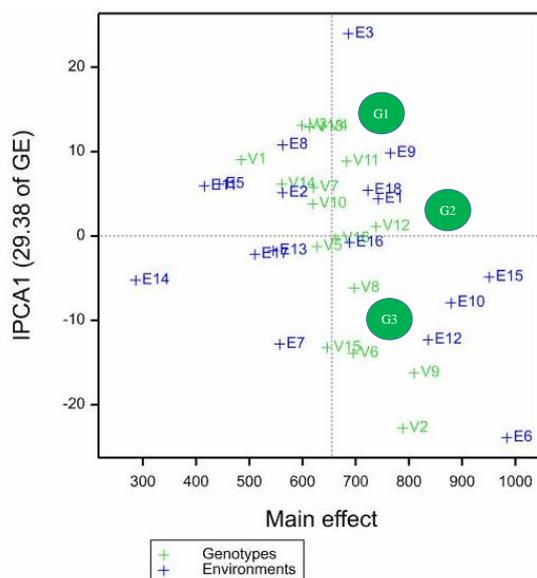


Fig 1. Additive main effects and multiplicative biplot of interaction principal component analysis (IPCA) scores of genotypes and environments against yield (kg/ha) of both genotypes and environments. *Description of environmental codes:* E1-Nachingwea 2013, E2-Naliendele 2013, E3-Tumbi 2013, E4-Naliendele 2014, E6-Nakabango 2014, E7-Sererere 2014, E8-Msekera 2014, E9-Nampula 2014, E10-Nampula 2015, E11-Ngabu 2014, E12-Chitedze 2014, E13-Chiedze 2015, E14-Ngabu 2015, E15-Lucydale 2015, E16-Chitedze 2016, E17-Ngabu 2016, E18-Baka 2016. *Description of genotype codes:* V1- ICG 405, V2- ICGV-SM 01514, V3- ICGV-SM 05521, V4- ICGV-SM 06687, V5-ICGV-SM 07501, V6- ICGV-SM 07502, V7-ICGV-SM 07510, V8- ICGV-SM 07517, V9- ICGV-SM 07520, V10- ICGV-SM 07532, V11- ICGV-SM 07536, V12- ICGV-SM 07539, V13- ICGV-SM 07541, V14- ICGV-SM 95741, V15- ICGV-SM 99568, V16-JL24.

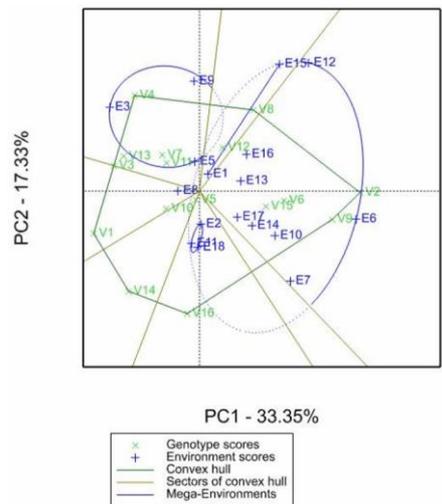


Fig 2. Additive main effects and multiplicative biplot of interaction principal component analysis (IPCA) scores of genotypes showing which-won where in the different mega-environments. *Description of environmental codes:* E1-Nachingwea 2013, E2-Naliendele 2013, E3-Tumbi 2013, E45-Naliendele 2014, E6-Nakabango 2014, E7-Serere 2014, E8-Msekera 2014, E9-Nampula 2014, E10-Nampula 2015, E11-Ngabu 2014, E12-Chitedze 2014, E13-Chiedze 2015, E14-Ngabu 2015, E15-Lucydale 2015, E16-Chitedze 2016, E17-Ngabu 2016, E18-Baka 2016. *Description of genotype code:* V1- ICG 405, V2- ICGV-SM 01514, V3- ICGV-SM 05521, V4- ICGV-SM 06687, V5-ICGV-SM 07501, V6- ICGV-SM 07502, V7-ICGV-SM 07510, V8- ICGV-SM 07517, V9- ICGV-SM 07520, V10- ICGV-SM 07532, V11- ICGV-SM 07536, V12- ICGV-SM 07539, V13- ICGV-SM 07541, V14- ICGV-SM 95741, V15- ICGV-SM 99568, V16-JL24.

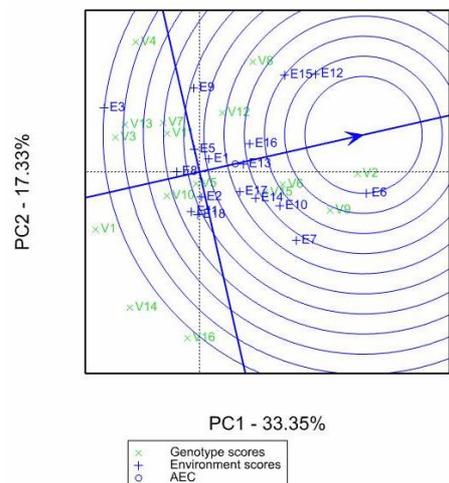


Fig 3. Discriminating ability vs. representativeness of test environments. *Description of environmental codes:* E1-Nachingwea 2013, E2-Naliendele 2013, E3-Tumbi 2013, E45-Naliendele 2014, E6-Nakabango 2014, E7-Serere 2014, E8-Msekera 2014, E9-Nampula 2014, E10-Nampula 2015, E11-Ngabu 2014, E12-Chitedze 2014, E13-Chiedze 2015, E14-Ngabu 2015, E15-Lucydale 2015, E16-Chitedze 2016, E17-Ngabu 2016, E18-Baka 2016. *Description of genotype code:* V1- ICG 405, V2- ICGV-SM 01514, V3- ICGV-SM 05521, V4- ICGV-SM 06687, V5-ICGV-SM 07501, V6- ICGV-SM 07502, V7-ICGV-SM 07510, V8- ICGV-SM 07517, V9- ICGV-SM 07520, V10- ICGV-SM 07532, V11- ICGV-SM 07536, V12- ICGV-SM 07539, V13- ICGV-SM 07541, V14- ICGV-SM 95741, V15- ICGV-SM 99568, V16-JL24.

Environmental stratification

Yield stability across environments

The PCA biplot showed a relatively large amount of variation among environments and less among genotypes (Fig. 1). The biplot had three distinct groups; the first group, G1, had genotypes with IPCA1 scores between 8 and 16, indicative of high adaptability to specific environments. The second group, G2, included genotypes with IPCA1 scores between -1 and 5, indicative of fairly stable genotypes but with limited adaptability to specific environments. The third group, G3, comprised genotypes with IPCA1 scores between -10 and -20, indicating that these genotypes were highly adapted to specific environments. The highest yielding genotypes, viz., ICGV-SM 07520 and ICGV-SM 01514, had large principal component scores and were situated far from the vertical

reference line, indicative of their adaptability to different environments and were stable at the same time (Fig. 1). Genotypes ICGV-SM 07536, JL 24, ICGV-SM 99568 and ICGV-SM 07501 had IPCA2 scores that were close to zero, and were, as per Gauch and Zobel (1997), relatively more stable across different environments. These genotypes can thus be grown in a wide range of environments. However, genotypes ICGV-SM 01514 and ICGV-SM 07520 were high yielding, widely adapted and stable-performing across different environments. Most of the genotypes (ICGV-SM-07536, ICGV-SM 07510, ICGV-SM 07532, ICGV-SM 07501, ICGV-SM 07539, ICGV-SM- 06687 and ICGV-SM 07541) and the environments Nampula 2014, Nakabango 2016, Nachingwea 2013, and Chitedze 2016, contributed positively to interactions, because they were grouped in the same direction and had positive IPCA 1 scores (Fig. 1).

In general, the score-points for environment were more scattered than for genotypes, indicative of relatively higher variability attributable to environmental effects (Table 1). The vast majority of genotypes were adapted to specific environments. The genotypes ICGV-SMs 01514, 07517 and 07539 were highly adapted to Nachingwea 2013 in Tanzania, Chitedze 2014/2015/2016 in Malawi and Lucydale 2015 in Zimbabwe (Fig. 1; see supplementary Table 2 and Fig 1 for details on environments). Genotypes ICGV-SM 07501, 07502, and 99568 were also highly adapted to Naliendele 2013 in Tanzania, Nakabongo 2014 in Uganda, Serere 2014 in Tanzania, Nampula 2015 in Mozambique and Ngabu 2015 in Malawi (Fig. 1; Supplementary Table 2). Such highly adapted genotypes could be deployed for improving performance of highly stable genotypes that are low yielding. The genotypes ICGV-SM 07539 and ICGV-SM 07536 were high performing and exhibited high levels of adaptability and stability across environments and are thus considered ideal genotypes. Genotype ICGV-SM 07501 was among the most stable but had low yield, thus low level of adaptability (Yan and Tinker, 2006). In addition, genotypes ICGV-SM 01514 and ICGV-SM 07520 were high yielding (>750 kg/ha), highly adapted and moderately stable across different environments, and therefore, are important for use in breeding programs (Fig. 1; supplementary Table 3).

Genotype ranking based on yield

Based on grain yield, genotype ICGV-SM 07520 was ranked first, followed by ICGV-SM 01514 and ICGV-SM 07539 (Table 1). Among the high-yielding genotypes, ICGV-SM 01514 and ICGV-SM 07520 were highly adapted and moderately stable across environments; whereas ICGV-SM 07539 was moderately adapted and stable across environments. Other genotypes that performed better than the local check (ICGV-SM 99568), a highly adaptable genotype widely grown in different countries of East and Southern Africa, were ICGV-SM 07517, ICGV-SM 07502 and ICGV-SM 07536. The least productive genotypes were ICGV-SM 405 and ICGV-SM 95741, which were moderately to highly unstable across environments (Table 1).

Favorable ecologies for Valencia genotypes

Discriminating environments for specific agroecologies

In East and Southern Africa, the highly discriminating environments for grain yield were Nakabongo 2014 in Western Uganda (984 kg/ha), Lucydale 2015 in Southern Zimbabwe (951 kg/ha) and Nampula 2015 in Eastern Mozambique (879 kg/ha). Materials tested at these sites exhibited high levels of instability, demonstrating the discriminating ability of these environments. These can be referred to as type 3 environments useful in culling unstable genotypes but not suitable for selecting superior genotypes (Yan et al., 2007). The biplot (Fig. 1) showed that two environments, viz., Chitedze 2016 and Ngabu 2016, representing mid-altitude and low altitude agroecologies of Malawi, respectively, consistently supported stable performance of the test genotypes. Such environments can be regarded as type 2 test environments that are suitable for selecting superior genotypes. Stability in particular environments demonstrates reliability of genotype ordering in an environment, in relation to its rating in other

environments (Yan et al., 2007). Thus, Lucydale 2015, Nakabongo 2014 and Nampula 2015 identified genotypes that were highly adapted and met specific market needs. Nakabongo serves the broader East African market needs, Lucydale serves Southern Africa market needs and Nampula serves South-East Africa market needs for low-lying groundnut-producing agro-ecologies bordering the Indian Ocean in Mozambique, Tanzania and South Africa.

Genotype and Genotype by Environment interactions (GGE)

The biplot analysis grouped the 17 environments into three mega environments, of which two intersected close to the biplot origin, an indication that both belong to the same mega-environment (Figure 2). The first mega-environment included: Naliendele 2013, Ngabu 2014 and Baka 2016 that had a negative direction in relation to the origin of the biplot (located on the left side in the biplot), which is indicative of their negative contribution to the genotype \times environment interaction. These locations represent low altitude environments (100-500 masl) and agro-ecologies that receive 600 mm or less rainfall annually, and therefore, may not support high productivity as in high potential areas (altitude > 1200 masl and rainfall >1000 mm).

The second mega-environment included Tumbi 2013 in Southern Tanzania, and Nampula 2014 in Eastern Mozambique, which lie in the mid-altitude (500-1100 masl) agroecologies of East and Southern Africa. This mega-environment shares characteristics with the third and largest mega-environment and had five genotypes mapped to it. This mega-environment, though highly discriminative, is less representative compared with the other two for production of Valencia genotypes and groundnut, in general. The third mega-environment was the largest and included most test sites (seven in total) across East and Southern Africa. These included Lucydale 2015 in Zimbabwe, Nachingwea 2013 in Tanzania, Chitedze 2014/2015/2016 in Malawi, Nakabongo 2015 in Uganda and Nampula 2015 in Mozambique. These environments recorded the highest grain yield, an indication that they are high-yield-potential groundnut production areas. Among them, Nachingwea 2013, Chitedze 2015/016 and Nakabongo 2015, being closer to the center of the origin, were the most representative and discriminating environments (Figure 2). These environments (Nachingwea 2013, Chitedze 2015/016 and Nakabongo 2015) supported moderate (600-800 kg/ha) to high yields (>800 kg/ha).

Most of the genotypes were situated within the third mega-environment, with less deviation in mean yield performance than the average response. Similar results have been reported elsewhere for other field crops (Yan et al., 2007; Yan et al., 2011). Winning genotypes for the different categories were located at the vertex of the polygon (Figure 2). These included ICGV-07517, ICGV-SM 01514, ICGV-SM 95741, JL 24, ICGV 405, ICGV-SM 7520, ICGV-SM 6687 and ICGV-SM 95741. The best genotypes, farthest in the polygon from the straight line passing through the origin (0), were ICGV-SM 01514, ICGV-SM 06687 and JL24 in one or more environments (Figure 2). Environments and genotypes that fall near the origin are considered stable and/or adapted, thus represent ideal environments and genotypes, respectively. The IPCA1 and IPCA2 explained 33.35% and 17.33%, respectively, of genotype by environment interaction.

Environment and genotype ranking

The 17 environments were ranked based on environmental average yield and genotype reaction to the different environments with yield as the unit of measure (Table 2). The results indicated inconsistency in genotype grain yield rankings from location to location (Figure 1; Table 2). Multi-environment-trial data may constitute a mixture of crossover (change in yield ranking of genotypes across environments) and non-crossover (constant yield ranking across environments) types of GEI (Matus-Cadiz et al., 2003; Yan and Tinker, 2005; Yan et al., 2007). The results suggested that there was crossover GEI, albeit with some genotypes exhibiting non-crossover performance, with consistent ranking in different locations. For example, genotype ICGV-SM 07520, which had high yield and was moderately adapted to different environments, was among the best performers in 11 out of 17 environments. Genotype JL 24, a local stable check in the region but susceptible to groundnut rosette disease (a destructive viral disease) and early leaf spot, was ranked among the best four genotypes in 8 out of 17 environments. Genotype ICGV-SM 07539, the most adapted and stable genotype, was ranked among the best four in 6 out of the 17 environments (Table 2). These new improved genotypes (ICGV-SM 07520 and ICGV-SM 07539) could thus be used to meet the increasing demand for Valencia varieties in East and Southern Africa.

Best test environments for Valencia groundnut genotypes

The results showed that environments Nachingwea 2013, Nakabango 2014 and Chitedze 2015 were representative of the average environment. These environments had relatively small PCA2 scores, a requirement for an ideal environment (Yan et al., 2011) and their angles to the average environment axis (AEA) were small, indicative of their representativeness of other test environments (Yan and Rajcan, 2002; Sharma et al., 2009). Environment Nakabango 2014 had the longest environmental vector, smallest angle with the AEA and was located at the circumference of the first inner score, implying that it was the most discriminating environment (Figure 3). Environments Lucydale 2015, Chitedze 2014, Serere 2014, Tumbi 2013 and Nampula 2014 also had long environmental vectors, and thus had the capacity to discriminate genotypes according to their performance but lacked representativeness as they had larger angles with the AEA than those designated to be so. Such environments are therefore suitable in selecting for specifically adapted genotypes. Conversely, environments Nachingwea 2013, Naliendele 2013, Naliendele 2014, Chitedze-2015 and Ngabu 2016 had the shortest environmental vectors, indicative of their inability to discriminate genotypes but were representative of average environment, since they had small angles with the AEA (Yan et al., 2011, Yan and Tinker, 2006).

The angle between vectors of two environments is related to the correlation coefficient between them (Yan and Kang, 2003). The different environments are either positively correlated or negatively correlated or not correlated, if the angles between their vectors are $<90^\circ$, $>90^\circ$ or $= 90^\circ$, respectively (Yan and Kang, 2003; Sharma et al., 2009). Based on the angle of environment vectors, the comparison biplot grouped the environments into five sub-groups, of

which four represented those that were positively correlated, and the fifth group comprised distinct environments that were negatively correlated to the other four (Figure 3). These sub-groups included: i) Nachingwea 2013, Chitedze 2016, Naliendele 2014, Nakaango 2014, Chitedze 2015; ii) Naliendele 2013, Baka 2016, Ngabu 2014; iii) Ngabu 2016, Ngabu 2015, Nampula 2015; iv) Lucydale 2015, Chitedze 2014; and v) Serere 2014, Tumbi 2013, Nampula 2014. The first three groups were, however, close to one another; therefore, they had positive correlation, as opposed to the fourth and fifth groups, which were far from the other three groups and negatively correlated to all other groups with high dissimilarity between them.

Different winning genotypes were clustered in the different test locations and the between-environment-group variation was greater than the within-group variation. This distinct variation further supports the existence of mega-environments and the possibility of grouping genotypes into the different mega-environments as an important option for exploiting $G \times E$ interaction.

An ideal genotype should have high mean performance and be stable (Yan and Kang, 2003; Blanche and Myers, 2006). An ideal genotype thus has the longest vector and maps at the center of the concentric circles. Genotype ICGV-SM 01514 was in the center of the concentric circles and was therefore the ideal genotype in this study. It can therefore be used as a reference for genotype performance evaluations (Dehghani et al., 2006; Yan et al., 2010). Genotype ICGV-SM 07520 was located in the second concentric circle and can thus be regarded as a desirable genotype. The genotype ICGV-SM 07520 was among the highest yielding but moderately adapted to different environments.

Materials and methods

Germplasm

Sixteen Valencia groundnut genotypes, i.e., 14 from the Malawi breeding program and two accessions from ICRISAT's gene bank, were evaluated in 2013, 2014, 2015 and 2016. The test materials were derived from crosses between a line resistant to aphid infestation (ICG 12991) and an early leaf spot (*Cercospora arachidicola* Hori)-tolerant line (ICGV-SM 93555) (supplementary Table 1). Local checks used in the study were *Kakoma* (JL 24) and *Chitala* (ICGV-SM 99568), which are high-yielding commercial varieties grown widely in East and Southern Africa.

Experimental sites

Test materials were evaluated in six countries, representative of major groundnut-producing agroecologies of East and Southern Africa (supplementary Fig 1). Evaluations started during the 2013 cropping season, and continued annually through the 2014, 2015 and 2016 cropping seasons. Each cropping season, in each of the studied geographies, was considered a different environment, generating 17 environments in total (supplementary Table 2). These environments represented low, medium and high-altitude production potential areas.

Experimental design and management

The experiment was laid out following a resolvable incomplete block design (4 × 4 lattice) with two replications at all locations. The plot size comprised 4 rows, each 6 m long, with a spacing of 0.75 m between rows and 0.1 m between plants. Recommended agronomic practices for groundnut (planting, weeding and harvesting) were followed. Harvesting and all other post-harvest handling processes were done manually on a plot basis. Pods from each plot were sun-dried to 13% moisture, determined using a moisture meter (Min GAC-Plus moisture tester DICKEY-John Corporation, UK) and thereafter weighed using a weighing scale (LBK, ADAM equipment, USA).

Data collection

Data on yield and yield-related traits were collected at different stages during each cropping season, which included days to 75% flowering, 100-seed weight (g), shelling out-turn (%), haulm yield (kg/ha) and grain yield (kg/ha). Grain yield per plot was converted to kilograms per hectare using the following formula: grain yield (kg/ha) = (Plot yield (kg) × 10,000 square meters/plot size in square meters (ha)), adapted from Rana and Kumar (2014):

Data analysis

ANOVA was performed separately for each site, followed by Bartlett's homogeneity test of variances (Steel and Torrie, 2006). Subsequently, based on Bartlett's test, pooled ANOVA was done (Steele and Torrie, 2006). Environments were treated as random effects and genotypes as fixed effects. Analysis of variance was performed using Genstat 20th edition (VSN, 2017). Means of genotypes per site were ranked to assess the importance of crossover genotype × environment interactions. The AMMI model was used to assess genotype adaptability and stability as well as for environmental stratification using the winning-genotype approach to determine mega environments (Gauch and Zobel, 1997). Genotype + genotype × environment interaction (GGE) scatter plots were generated to identify genotypes adapted to specific environments and the most discriminating and suitable culling environments (Yan and Tinker, 2006).

Conclusions

The best-performing stable genotypes across environments were ICGV-SM 07539, ICGV-SM 07536, ICGV-SM 7502, ICGV-SM 01514 and ICGV SM 07520. These are desirable genotypes that could be deployed in breeding programs to improve yield and stability. Other stable genotypes with moderate yields included ICGV-SM 07532, ICGV-SM 07536 and ICGV-SM 07501. The study also identified three environments, namely, Nachingwea 2013, Chitedze-2015 and Nakabango 2014 as the most discriminating and representative environments. These three locations can be used for early generation screening and data capture for release of varieties in the region.

Acknowledgements

This multi-year work was made possible through the generous support of the Tropical Legumes II project funded by Bill and Melinda Gates Foundation, Improving groundnut farmer incomes and nutrition through innovation and technology enhancement (I-FINITE), funded by USAID Zambia and the McKnight Foundation supported groundnut variety improvement for yield and adaptation, human health and nutrition in Malawi. The authors thank research technicians from the National Agricultural Research Systems (NARS) in the different countries where the evaluations were conducted for their support and engagement in the work.

Disclosure statement

There is no conflict of interest reported by authors.

References

- Affholder F, Poeydebat C, Corbeels M, Scopel E, Tittone P (2013) The yield gap of major food crops in family agriculture in the tropics: assessment and analysis through field surveys and modelling. *Field Crop Res.* 143:106-118.
- Badu-Apraku B, Lum AF, Fakorede MA, 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 Sci.* 48:99-112.
- Bernardo R (2002) Breeding for quantitative traits in plants, 2nd Edition. ed. Stemma Press.
- Blanche SB, Myers, GO (2006) Identifying discriminating locations for cultivar selection in Louisiana. *Crop Sci.* 46:946-949.
- Blanche SB, Myers G, Kang MS (2007) GGE biplots and traditional stability measures for interpreting genotype by environment interactions. *J Crop Improv.* 20:123-135.
- Casonoves F, Baldessari J, Balzarini M (2005) Evaluation of multi-environment trials of peanut cultivars. *Crop Sci.* 45:18-26.
- Dehghani H, Ebadi A, Yousefi A (2006) Biplot analysis of genotype by environment interaction for barley yield in Iran. *Agron J.* 98:388-393.
- Eberhart SA, Russell WA (1966) Stability parameters for comparing varieties. *Crop Sci.* 6:36-40.
- Fan X, Kang MS, Chen H, Zhang Y, Tan J, Xu C (2007) Yield stability of maize hybrids evaluated in multi-environmental trials in Yunan, China. *Agron J.* 99:220-228.
- Gan Y, Liang BC, Hamel C, Cutforth H, Wang H (2011) Strategies for reducing the carbon foot print of field crops for semi-arid areas. *Agron Sustain Dev.* 31:643-656.
- Gauch HG, Zobel R (1997) Identifying mega-environments and targeting genotypes. *Crop Sci.* 37:311-326.
- Gauch HG (2006) Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.* 46:1488-1500.
- Gauch HG, Piepho HP, Annicchiarico P (2008) Statistical analysis of yield trials by AMMI and GGE: further considerations. *Crop Sci.* 48:866-889.
- Jayne TS, Chamberlin J, Headey DD (2014) Land pressures, the evolution of farming systems, and development strategies in Africa: a synthesis. *Food Policy.* 48:1-17.

- Kang MS (1998) Using genotype-by-environment interaction for crop cultivar development. *Adv Agron.* 62:199-252.
- Lin CS, Binns MR, Lefkovitch PL (1986) Stability analysis: where do we stand? *Crop Sci.* 26:894-899.
- Matus-Cadiz MA, Hucl P, Perron CE, Tyler RT (2003) Genotype x environment interaction for grain color in hard white spring wheat. *Crop Sci.* 43:219-226.
- Moreno-González J, Crossa J, Cornelius PL (2004) Genotype × environment interaction in multi-environment trials using shrinkage factors for AMMI models. *Euphytica.* 137:119–127.
- Nigam SN, 2014. Groundnut at a glance. pp. 121.
- Oliveira EJ, Godoy IJ (2006) Pod yield stability analysis of runner peanut lines using AMMI. *Crop Breed Appl Biot.* 6:311-317.
- Pacheco RM, Duarte JB, Vencovsky R, Pinheiro JB, Oliveira AB (2005) Use of supplementary genotypes in AMMI analysis. *Theor Appl Genet.* 110:812-818.
- Rana SS, Suresh K (2014) Research techniques in agronomy. Department of agronomy, college of Agriculture, CSK Himachal Pradesh Krishi Vishvavidyalaya, Palampur. pp. 64.
- 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.*
- Sharma RC, Morgounov A, Baun H, Beyhan A, Mesut A, Dedoshvili D, Ahmet B, Martius C, Maarten VG (2009). Identifying high yielding stable winter wheat genotypes for irrigated environments in Central and West Asia. *Euphytica.* 171:53-64.
- Steel GD, Torrie JH (2006) Principles and procedures of statistics: a biometrical approach, 3rd Edition. ed. CRAM101, Sim Valley, CA, United States.
- VSN International (2017). Genstat for windows 20th edition. VSN International, Hemel Hempstead, UK. Web page: Genstat.co.uk
- Yan W (2002) Singular-value partitioning in biplot analysis of multi-environment trial data. *Agron J.* 94:990-996.
- Yan W, Frégeau-Reid J, Pageau D, Martin R, Mitchell-Fetch J, Etienne M, Rowsell J, Scott P, Price M, de Haan B, Cumiskey A, Lajeunesse J, Durand J, Sperry E (2010) Identifying essential test locations for oat breeding in Eastern Canada. *Crop Sci.* 50:504-515.
- Yan W, Hunt LA, Sheng Q, Szlavnic Z (2000) Cultivar evaluation and mega environment investigation based on GGE biplot. *Crop Sci.* 40:597-605.
- Yan W, Kang MS (2003) GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists. CRC Press Boca Raton FL.
- Yan W, Kang MS, Ma B, Woods S, Cornelius PL (2007) GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* 47:643-655.
- Yan W, Pageau D, Frégeau-Reid J, Durand J (2011) Assessing the representativeness and repeatability of test locations for genotype evaluation. *Crop Sci.* 51:1603–1610.
- Yan W, Rajcan I (2002) Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Sci.* 42:11-20.
- Yan W, Tinker NA (2005) An integrated biplot analysis system for displaying, interpreting, and exploring genotype x environment interaction. *Crop Sci.* 45:1004–1016.
- Yan W, Tinker NA (2006) Biplot analysis of multi-environment trial data: principles and applications. *Can J Plant Sci.* 86:623-645.
- Yau SK (1995) Regression and AMMI analyses of genotype × environment interactions: an empirical comparison. *Agron J.* 87:121-126