Australian Journal of Crop Science

AJCS

AJCS 15(10):1224-1232 (2021) ISSN:1835-2707
doi: 10.21475/ajcs.21.15.10.p2856

Assessment of genetic variability in groundnut (*Arachis hypogaea* L.) genotypes grown under South African conditions using agronomic and SSR markers

Mofokeng MA*[1,3], Amelework BA[2], Chipeta O[3], Sibiya J[3], Gerrano AS[2], Shargie N[1] and Mashingaidze K[1]

[1]Agricultural Research Council-Grain Crops, Private Bag X 1251, Potchefstroom, 2520, South Africa
[2]Agricultural Research Council-Vegetable and Ornamental Plants, Private Bag x 293, Pretoria, 0001, South Africa
[3]African Centre for Crop Improvement, School of Agriculture, Earth and Environmental Sciences, University of KwaZulu-Natal, Private Bag X 01, Scottsville, 3209, South Africa

Corresponding author: MofokengA@arc.agric.za

Abstract

Groundnut (*Arachis hypogeae*) is a legume crop grown in arid and semi-arid regions of the world. The objective of the study was to assess the presence of genetic diversity among fifty three groundnut genotypes of diverse origin using eleven agronomic and twenty SSR markers. The analysis of variance showed that highly significant variations exist among the genotypes for all phenotypic traits measured. Five principal components showed 71% of the total phenotypic variation. The SSR loci showed high values of polymorphic information content ranging from 0.31 to 0.89, with a mean of 0.71. Heterozygosity values ranged between 0.03 and 1.00 with a mean of 0.57. The genotypes showed a wide range of allelic diversity from 3 to 16, with a mean of 8.1 alleles per locus. Analysis of molecular variance (AMOVA) indicated that larger variability (59%) was due to variation within individuals, whilst the remaining variation was accounted for variation among individuals within population. Cluster analysis grouped the genotypes into two distinct clusters, where it showed that the discrimination of the genotypes was not dependant on the origin of the genotypes. The high gene flow observed among the different geographic origin might contribute to the low differentiation among the population. The SSR and phenotypic markers were able to detect wide genetic diversity and discriminate groundnut genotypes. The two genetically distinct groups observed in this study, can be used as source of genes of novelty and parental lines for transgressive segregation and for further broadening of the genetic base of the crop.

Keywords: Agronomic, genetic diversity, groundnut, SSR.
Abbreviations: SSR: Simple Sequence Repeat, ANOVA: Analysis of variance

Introduction

Groundnut (*Arachis hypogeae)* is an allotetraploid (2n=4x=40), self-pollinating legume crop widely grown in tropical and sub-tropical areas of the world (Singh et al., 2013). Groundnut is the 6th most important oil seed crop worldwide (Nigam, 2014) and is produced in various parts of the world including Africa. The groundnut production in South Africa recorded in 2016 was 1.81 ton per hectare in an area of about 56 000 hectares (FAOSTAT, 2019) with the main production coming from small-scale farmers in Northern Cape, North West and Free State Provinces. Groundnut is a nutritious crop mainly grown for protein and oil content (Atasie et al., 2009). The seeds contain about 48% oil, 25% protein and 18% carbohydrates and are a rich source of B-complex vitamins, minerals, antioxidants, and biologically active phytochemicals (Mahesh et al, 2018; Bonku and Yu, 2020). This crop can be useful for combating protein, carbohydrate and micronutrient malnutrition among the poor (Ingale and Shrivastava, 2011; Mupunga et al., 2017; Bonku and Yu, 2020). Another benefit of the crop includes the fixation of atmospheric nitrogen, hence, it is good for rotation with other crops especially cereals towards increased soil improvement and management of pests, diseases, and weeds. Groundnut is used as food for both humans and animals (Asibuo et al., 2008; Pandey et al., 2012; Bonku and Yu, 2020). It provides dietary-proteins (Upadhyaya et al. 2005), which constitutes 20-40% of the seed (Singh et al., 2013). It is also commonly grown for its edible oil, which makes 40-60% of the seed (Upadhyaya et al., 2005; Campos-mondragon et al., 2009; Singh et al., 2013). In South Africa, groundnut is grown by both smallholder and commercial farmers mainly for human consumption. Groundnut in South Africa has shown diverse production trends over time due to several constraints faced by the sector (FAOSTAT, 2019). South Africa groundnut production was at 82,122 metric tons in 52,125 hectares in 2014, while in 2018 the land harvested of groundnuts was 56,300 hectares with total production of about 57 000 metric tons (FAOSTAT, 2019). The worldwide average yield of groundnut is 1.8 tons per hectare. Groundnut has been reported to have a narrow genetic base (Bhad et al., 2016). Enhancing the genetic base of the crop is one of the strategies that has been put in place in various breeding programmes. Genetic diversity studies with the aim of assessing the morphological, biochemical and genetic variabilities present in a given population are crucial for

groundnut improvement (Bhad et al., 2016; Sai et al., 2016). Hence, knowledge and understanding of genetic diversity within and/or among genotypes is important in choosing the promising parents for hybridization and in planning successful breeding programme. The progeny originating from diverse parents exhibit greater heterosis and provide broad spectrum of variability in segregating generations through the production of transgressive segregants (Bhati et al., 2015). Hence, efforts have been made to identify parents with wide genetic divergence from germplasm pools for the characters of economic importance such as increasing yield, wider adaptation, and desirable quality, pest and disease resistance, to utilize them in hybridization programme.

Various tools are used for assessment of genetic diversity. The tools include agronomic, morphological and quality traits, as well as biochemical and molecular markers. Although, the agro-morphological traits are mainly influenced by environment, most breeders are still using them for assessing the yield performance of various genotypes. Molecular markers have been proved to be an efficient tool to assess variation within and among groundnut populations (Kanyika et al., 2015) and to isolate genes linked to desirable traits (Mace et al., 2006; Asibua et al., 2008; Bhad et al., 2016; Zongo et al., 2017). Simple sequence repeats (SSRs) are one of the PCR-based markers, which have been extensively used for genetic diversity analysis (Moretzsohn et al., 2004; Kanyika et al., 2015). SSR markers are codominant markers that are relatively abundant, highly polymorphic, and show simplicity of genotyping (Matus and Haye, 2002). Krishna et al., (2004) discovered considerable diversity amongst 48 Valencia groundnut genotypes that were studied in the United States of America. Several other studies confirmed that groundnut germplasm showed large genetic diversity (Dwivedi et al., 2001; Moretzsohn et al., 2004; Zaman et al., 2011; Idi-Garba et al., 2015). Nonetheless, some researchers (Halward et al., 1991; Herselman et al., 2003) have also reported low variability among groundnut germplasm. In South Africa, there is little information available on the genetic diversity of groundnut genotypes that have been kept in the Agricultural Research Council-Grain Crops gene bank. The aim of the present study was to evaluate the presence of genetic diversity in groundnut genotypes grown under South African conditions based on agro-morphological traits and SSR markers for subsequent breeding and conservation.

Results

Analysis of variance (ANOVA)

The ANOVA for agronomic traits revealed that the groundnut genotypes collected from different origins had significant diversity among them, as studied through trait means, range, standard deviations and coefficients of variation for the quantitative morphological characteristics (Table 3). Wide ranges of variation observed for the traits such as pod length, stem height, 100 seed weight and plant throttling. The genetic diversity measured based on the frequency of a given trait class showed that genotypes revealed highest genetic diversity for leaf length and width and days to flowering. Groundnut genotypes showed high genetic diversity greater than 0.80 for all traits except grain yield.

Analyses of variance conducted per site revealed high significant differences (P < 0.01) among groundnut genotypes for all the characters suggesting the existence of considerable variation. Thus, a combined analysis of variance was conducted over the two sites that showed significant genotype by site interactions for all traits except leaf length (Tables 4). All pod and related stem related traits revealed highly significant genotype x site interaction (P< 0.001), while grain yield, days to flowering and branch number showed high significant. Genotype and site had a highly significant (P< 0.001) effect, but the effect of site was not significant for stem height.

Principal components analysis

The first five principal components (PCs) with eigenvalues greater than or equal to one, explained 71% of the total variation among the studied groundnut genotypes for the eleven quantitative characters (Table 5). About 25.5% of the total variation accounted for by the first PC alone was due to the contrast between days to flowering and the average effects of leaf width and stem height. In the second PC, pod length and width, and leaf length and width had significant contribution. Likewise, the third PC accounted for about 12% of the total variance of the genotypes originated mainly from pod weight. Variation in branch number and plant throttling constituted a large part of the total variation explained by the fourth PC. The fifth PC was due to the effects of 100 seed weight. The biplot analysis based on the first two PCs grouped genotypes into three clusters based on their geographic origin (data not shown). The presence of vast diversity among the genotypes in this study was clearly shown by the distant relationships among the genotypes. The diverse genotypes could be useful for selection in plant breeding programmes and for further genetic improvement.

Allelic diversity of SSR markers

A total of 53 genotypes were collected for genotyping, however, 14 genotypes were omitted from the analysis. In these genotypes the SSR either failed to amplify band or less than 2% of the markers were amplified. The 20 SSR markers used in this study amplified a total of 162 putative alleles (different fragment sizes) of which more than 59% (96 putative alleles) were effective in discriminating the genotypes (Table 6). The genotypes showed a wide range of allelic diversity from 3 to 16, with a mean value of 8.1 alleles per locus (Table 6). The highest allele number was observed from marker pPGPseq2E6 (3%) and the lowest was from IPAHM103 (16%). The PIC value ranged from 0.31 to 0.89, with a mean value of 0.71. Most of the markers were polymorphic with PIC values of > 0.50 except two markers (AH-10 and SEQ3A05) (Table 6) which are monomorphic. Markers AC2A04 and AH-10 had the same number of alleles, however, the PIC values were 0.72 and 0.32, respectively. This was observed due to the differences in allelic frequencies in that the major allele frequency in AC2A04 was 0.30, while the major allele frequency in AH-10 was 0.82. All the alleles amplified by the SSR primers in this study showed an allele frequency of less than 0.50 except for two markers (AH-10 and SEQ3A05) suggesting even distribution among the genotypes tested.

The mean observed heterozygosity per locus was 0.57 and with the highest (1.00) and lowest (0.03) values were detected from PM35 TC11C06 and TC9B08 and TC3A12, respectively (Table 6). About 45% of the markers showed H_o value of > 0.80 and a negative inbreeding coefficient (F_{IS}) values. F_{IS} values represents the average deviation of the population's genotypic proportions from Hardy-Weinberg equilibrium and the values ranged from 0.00 to 1.00. A negative F_{IS} value represents an excess of heterozygotes. For example, for loci TC3A12, PM35 and TC11C06, 73%, 87% and

81% of the genotypes are expected to be heterozygous at the specific loci under random mating conditions, respectively. However, 100% of the genotypes at these loci were heterozygotes. It may be due to high outcrossing or mutation at the specific loci. Gene diversity (H_e) ranged from 0.32 (AH-10) to 0.9 (IPAHM103 and PM3) with a mean of 0.75 was detected.

Population divergence

Genetic parameter estimates of groundnut populations stratified based on geographic origin are presented in Table 7. Genotypes originated from Southern Africa revealed the highest variation for most of the genetic parameters. The mean observed (N_a) and effective (N_e) number of alleles was higher for genotypes from Southern Africa and South America, respectively. Shannon information index was higher for genotypes from Southern Africa followed by genotypes from South America with mean values of 1.51, and 1.42, respectively. The highest mean observed heterozygosity (0.60) was observed from genotypes originated from North America and the lowest H_0 (0.54) was detected from South America genotypes. On the contrary, the highest expected mean gene diversity (0.76) was detected from genotypes driven from Southern Africa followed by South America genotypes (0.74). The mean fixation index was relatively higher for South America and Southern Africa originated genotypes. Highest number of private alleles (19) per population was detected from the Southern Africa collections followed by genotypes collected from diverse origin (10). All genotypes driven from South and North America and Southern Africa showed the highest percentage of polymorphic loci (Table 7).

Genetic differentiation (F_{ST}) among the geographic origin ranged from 0.041 between South America and Southern Africa and 0.059 between North America and other countries suggesting there was little to moderate differentiation among the four groundnut populations (Wright, 1978). The relatively low values of F_{ST} imply that there is high frequency of identical alleles among population. Gene flow among the groundnut population within geographic origin ranged from 3.99 between North America and genotypes collected from diverse sources to 5.90 between Southern Africa and others (Table 8). The populations maintained higher genetic identity and low genetic distances.

Analysis of molecular variance (AMOVA)

Analysis of molecular variance among groundnut populations stratified based on geographic origin are shown in Table 9. No significant genetic differentiation was observed among the four populations (P = 0.955). However, highly significant difference (P < 0.001) of molecular variation was observed among individuals within the population. Similarly, highly significant (P < 0.001) variation was detected within individual in all the 39 groundnut genotypes collected from diverse geographic locations. Larger genetic variability (59%) was attributed to variation within individuals, and the remaining variation was explained by variation among individuals within population (Table 9). This signifies that in groundnut the between and within individual variation is more crucial than the between population variation.

Cluster analysis

The genetic relationship among the groundnut genotypes was assessed using neighbour-joining algorithm using the unweighted pair group method. The analyses indicated the presence of two distinct sub populations (Figure 1). The clustering patterns of the genotypes did not match with the geographical origin probably due to high gene flow (Table 7). Cluster I contained the highest proportion of the genotypes (62%) and dominated by the Southern Africa collections. This cluster further sub-divided into four sub-clusters. Cluster II had three sub-clusters comprising of 15 genotypes. This cluster was represented by relatively equal proposition of genotypes from each subpopulation. The closely related genotypes were grouped together.

Discussion

Understanding of the genetic diversity of germplasm is fundamental to effectively utilize and conserve genetic resources and to design proper strategies for its improvement. The current study examined genetic diversity and population structure of groundnut genotypes collected from diverse geographic origins using agronomic traits and SSR molecular markers. Highly significant differences were observed among the genotypes for most of the agronomic traits. All the traits except grain yield contributed to the phenotypic diversity observed in the studied groundnut genotypes. Genetic variation in yield and yield related traits are associated with differences in morphological and developmental traits (Belay et al., 2019). In the current study, the groundnut germplasm revealed very low range of yield difference (33.5 g per plant) and the overall genetic diversity for yield was narrow, indicating the overall yield performance of genotypes is determined by the interaction of traits rather than the expression of individual traits. Blum (1988) and Redae et al. (2017) reported that yield by itself is not under direct genetic control since it is the integrated effect of the multitude of physiological and biochemical processes. The genotype and the environment components such as location, year, and growing-season are recognized as the primary sources of variability in agronomic and genetic studies (Ceccarelli et al., 1991, Sibhatu et al., 2017; Redae et al., 2017). In this study, both genotype, environment and their interaction found to be important for the performance of the groundnut genotypes. Dissecting of the total variance into its components as genotypic, environmental and interaction variance enable us to identify which component is more crucial. It was observed that environmental variances were higher for all the traits except leaf length, pod weight and stem height.

In the current study, the number of alleles produced by the SSR markers ranged from 3 to 15. Similar results were previously reported by Rosenberg et al., (2003) and He et al., (2003) in groundnut genotypes. However, Koppolu et al. (2010) reported much higher number of alleles that ranged from 4 - 28 in genus *Arachis*. All the SSR markers used in this study were highly polymorphic with a mean number of 8 alleles per locus. This value is much higher than the values what He et al., (1997; 2003; 2005) previously reported). Of the total 162 putative alleles detected 59% of the alleles were effective in discriminating the genotypes suggesting the alleles evenly distributed among the genotypes. The major allele frequency ranged from 0.16 to 0.82 with a mean value of 0.35. Goddard et al., (2000) suggested that markers with major allele frequency between 0.5 and 0.8 could be useful in QTL mapping. Therefore, markers PMc297, AH-10 and SEQ3A05 can be useful in providing information about linkage disequilibrium and QTL mapping in groundnut.

For all loci, the observed heterozygosity (H_o) was lower (mean = 0.57) than the expected heterozygosity (He) (mean = 0.75),

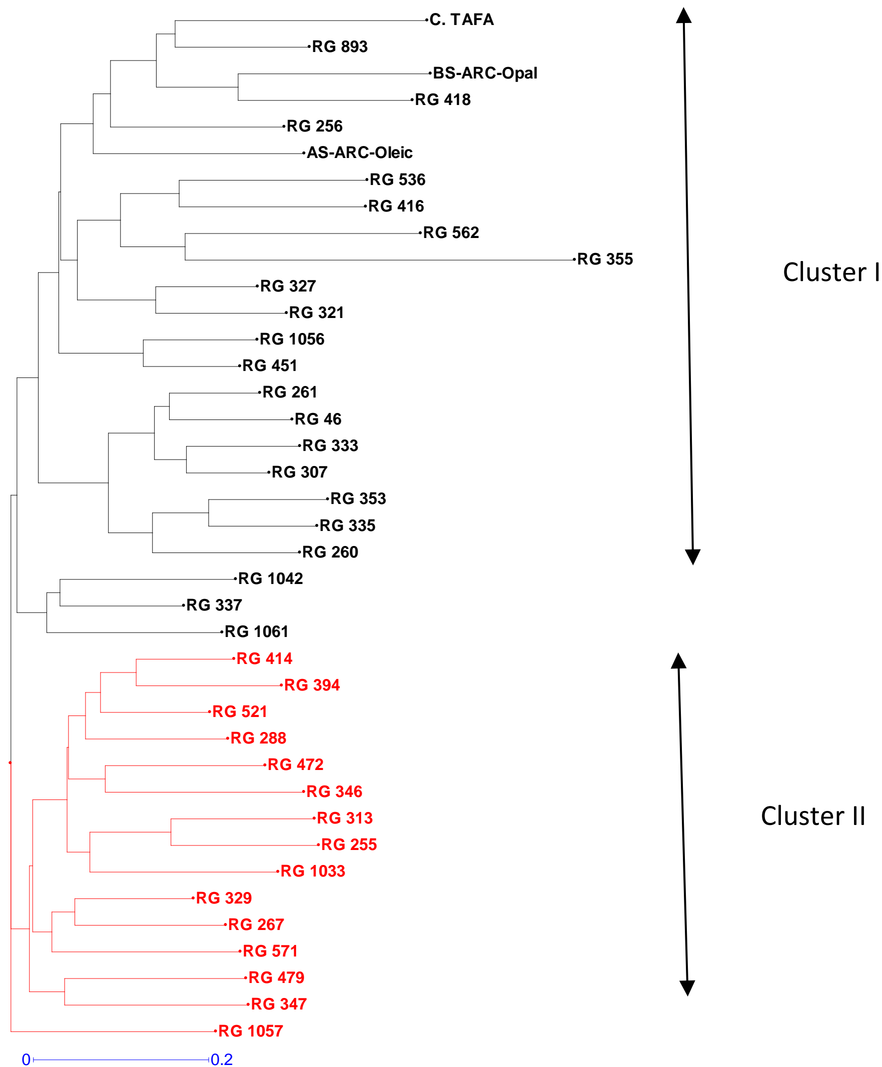

Figure 1. Un-weighted pair group method (UWPGM) dendrogram showing genetic relationship of the 39 groundnut genotypes determined using 20 selected SSR markers.

Table 3. Means, range, coefficient of variation, standard deviation and diversity for eleven quantitative traits measured in 53 groundnut genotypes.

Trait	Mean ± SE	Range	CV%	SD	Diversity index
Branch number (count)	6.3 ± 1.1	12.4	17.8	2.2	0.95
Leaf Length (mm)	54.8 ± 4.6	69.6	8.4	9.1	0.97
Leaf width (mm)	26.5 ± 2.5	31.6	9.4	4.9	0.97
Days to flowering (days)	48.6 ± 3.0	32.0	6.2	5.9	0.99
Pod weight/plant (g)	33.4 ± 4.1	60.7	12.3	8.1	0.86
Pod length (mm)	38.6 ± 4.7	111.3	12.3	9.3	0.85
Pod width (mm)	17.4 ± 3.2	72.4	18.4	6.3	0.81
100 seed weight (g)	50.9 ± 5.4	139.0	10.7	10.7	0.89
Grain yield/plant (g)	29.1 ± 3.1	33.5	10.6	6.0	0.79
Stem height (mm)	167.7 ± 12.0	263.6	7.2	23.6	0.92
Plant Throttling	207.9 ± 12.0	355.0	5.8	23.6	0.94

CV%=coefficient of variation in percent; SD=standard deviation; SE=standard error.

Table 4. Combined analysis of variance of fifty-three groundnut genotypes evaluated across two sites in South Africa.

S.O.V	**Genotype**	**Site**	**Genotype x site**
Days to flowering (days)	31.6***	4198***	30.6**
Branch number (count)	8.0***	67***	2.4**
Leaf Length (mm)	245***	50**	24ns
Leaf width (mm)	67***	157***	10.5*
Plant Throttling	7412***	46492***	4622***
100 seed weight (g)	1251***	16656***	752***
Pod weight/plant (g)	447***	223***	320***
Pod length (mm)	647***	27820***	315***
Pod width (mm)	471***	4170***	275***
Stem height (mm)	9236***	100.0ns	2347***
Grain yield/plant (g)	192 ***	13489***	160**

***Significant 0.1%; ** Significant at 1%; * Significant at 5%; ns=non-significant.

Table 5. Factor loadings of the agronomic traits of fifty-three groundnut genotypes evaluated at two sites showing the most important PCs.

Traits	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**
Days to flowering	**-0.362**	-0.038	0.258	0.325	0.323
Branch number	-0.251	0.232	0.233	**0.469**	-0.161
Leaf length (mm)	0.335	**-0.363**	0.198	0.125	0.290
Leaf width (mm)	**0.351**	**-0.361**	0.049	-0.069	0.293
Plant Throttling	0.314	0.050	0.177	**0.453**	-0.331
Stem height	**0.349**	-0.229	0.258	0.231	0.013
Pod length (mm)	0.286	**0.383**	-0.125	0.052	0.119
Pod width (mm)	0.205	**0.359**	-0.104	0.363	0.227
Pod weight/plant (g)	0.282	0.192	**0.381**	-0.125	-0.281
100 seed weight (g)	0.146	0.299	-0.321	0.088	**0.526**
Grain yield plant (g)	0.337	0.333	-0.015	-0.322	-0.186
Eigenvalue	3.315	2.376	1.617	1.04	0.962
Percentage	25.50	18.28	12.44	8.00	7.13
Cumulative percentage	25.50	43.78	56.22	64.22	71.35

Table 6. Genetic diversity parameters generated by 20 SSR markers among groundnut accessions.

Markers	%GA	N_a	N_e	H_o	H_e	F_{IS}	PIC	MA	MAF
PM375	92.31	11	7.30	0.67	0.88	0.23	0.86	121	0.25
PM3	94.87	11	8.78	0.97	0.90	-0.10	0.88	221	0.16
AC2H11	61.54	4	2.55	0.46	0.62	0.25	0.54	158/241	0.44
AC2A04	89.74	4	3.86	0.97	0.75	-0.31	0.72	191	0.30
TC9B08	82.05	7	4.52	0.03	0.79	0.96	0.76	113	0.34
pPGPseq2e6	89.74	15	7.40	0.60	0.88	0.31	0.86	267	0.26
TC2C07	84.62	7	2.61	0.82	0.63	-0.33	0.51	223	0.56
PMc297	61.54	4	3.73	0.58	0.75	0.20	0.71	240	0.33
PM137	76.92	6	3.96	0.03	0.76	0.96	0.73	164	0.33
AH-10	48.72	4	1.46	0.37	0.32	-0.17	0.32	254	0.82
PM183	51.28	8	3.72	0.45	0.75	0.38	0.70	147	0.38
PM50	71.79	12	7.54	0.32	0.88	0.63	0.86	121	0.25
TC3A12	71.79	7	3.50	1.00	0.73	-0.40	0.67	182	0.45
AH-8	58.97	7	4.62	0.26	0.80	0.67	0.77	252	0.30
TC2D06	56.41	6	3.18	0.00	0.70	1.00	0.64	215	0.41
PM036	92.31	13	5.34	0.86	0.82	-0.06	0.80	221	0.28
IPAHM103	94.87	16	9.22	0.95	0.90	-0.06	0.89	147	0.18
PM35	87.18	11	7.16	1.00	0.87	-0.16	0.86	110/112/125/158	0.16
TC11C06	97.44	6	5.09	1.00	0.81	-0.24	0.80	186/199	0.22
SEQ3A05	92.31	3	1.95	0.06	0.49	0.89	0.31	232	0.64
Mean	77.82	8.10	4.87	0.57	0.75	0.23	0.71	-	0.35
SE	3.54	0.87	0.51	0.08	0.03	0.11	0.04	-	0.04

%GA= percentage of genotypes amplified; N_a= Number of alleles per locus; N_e = number of effective alleles per locus; H_o= observed heterozygosity, H_e = expected heterozygosity; F = Inbreeding coefficient; PIC = polymorphic information content, MA= major allele; MAF major allele frequency per locus, SE= Standard error.

Table 7. Genetic diversity parameter estimates of groundnut populations based on geographic origin.

Population	N	Na	Ne	I	Ho	He	F_{IS}	PA	%P
North America	7	4.75	3.52	1.31	0.60	0.73	0.11	9.00	100.0%
South America	10	5.45	4.25	1.42	0.54	0.74	0.24	8.00	100.0%
Southern Africa	14	6.10	4.31	1.51	0.58	0.76	0.20	19.00	100.0%
Other countries	8	4.75	3.82	1.30	0.56	0.72	0.19	10.00	95.0%
Mean	7.59	5.26	3.97	1.39	0.57	0.74	0.18	-	98.8%
SE	0.30	0.27	0.21	0.05	0.04	0.02	0.06	-	1.25%

N= Number of observations; N_a= number of alleles per locus; N_e= number of effective alleles per locus; I= Shannon's information index; H_o= observed heterozygosity; H_e= expected heterozygosity; F_{IS}= Inbreeding coefficient; PA = Private allele per population; %P = Percentage polymorphic loci; others=refer Table 1.

Table 8. Pair-wise estimates of gene flow (**above diagonal**, within the brackets), genetic differentiation (F_{ST}) (above diagonal off brackets); genetic distance (GD) (lower diagonal off brackets) and genetic identity (GI) (lower diagonal within the brackets) (Ilesanmi and Ilesanmi, 2011).

Population	North America	South America	Southern Africa	Others
North America		0.043 (**5.59**)	0.052 (**4.55**)	0.059 (**3.99**)
South America	0.021 (0.98)		0.041 (**5.90**)	0.052 (**4.55**)
Southern Africa	0.111 (0.89)	0.057 (0.95)		0.043 (**5.60**)
Other countries	0.078 (0.93)	0.067 (0.94)	0.021 (0.98)	

Others=refer Table 1.

Table 9. Analysis of molecular variance (AMOVA) among 53 groundnut accessions classified based on geographic origin using 20 SSR markers.

Source of variation	df	SS	MS	Est. Var.	Perc. Var.	F-statistics
Among populations	3	28.511	9.504	0.000	0%	0.955
Among individuals	35	393.796	11.251	3.267	41%	0.001
Within individuals	39	184.000	4.718	4.718	59%	0.001
Total	77	606.308		7.985	100%	

Df = degrees of freedom; SS = sum of squares; MS = mean squares, Est. var.= Estimated variance, Perc. Var = Percentage variance.

suggesting a clear shift from the Hardy-Weinberg equilibrium. This shift may be attributed to high outcrossing or mutation at the specific loci. The PIC values ranged from 0.31 to 0.89 in this study which was similar to what Matus and Haye (2002); Moretzsohn et al., (2004) reported previously. However, the mean PIC value obtained in the current study was 0.71 and values with > 0.50 were observed in 90% of the loci analysed. This result was much higher than the findings of Cuc et al., (2008), where only 34% and 44% of SSR markers showed PIC values of >0.70 and >0.50, respectively. This suggested that the loci used in this study were highly polymorphic and the observed alleles were evenly distributed within the genotypes. This, in turn, indicates that these markers had a high discriminatory power and were found to be highly suitable for genetic diversity analysis (Tang et al., 2007). In case agro-morphological traits failed to detect variability due to the similarity in growing environments, SSR markers can be a useful tool in discriminating differences among genotypes at molecular level. Tang et al., (2007) in their genetic diversity analysis of groundnut genotypes that belong to var. *hirsuta* in southern China using agro-morphological traits found that all the genotypes were similar. However, using SSR markers they were able to discriminate the variation present among the genotypes.

The mean observed heterozygosity (H_o) of 0.57 and fixation index (F_{IS}) value of 0.23 was detected in this study. The high H_0 and the low F_{IS} values suggested that these genotypes are highly heterozygous and this is not the case with self-pollinated crops such as groundnut.

This might be resulted due to mutation or high natural outcrossing rate. However, it was reported that groundnuts exhibit low natural outcrossing rates ranging from 0 to 8% (Knauft et al., 1992; Reddy et al., 1993). The other reason could be the genotypes were sampled from breeding population at early stages of the breeding cycle. Similarly, high mean expected heterozygosity (He) value of 0.75 was observed among the genotypes, indicating the possibility of two randomly sampled alleles in a given genotype to be different was greater than 75%. This, in turn, suggested that these collections of genotypes were highly genetically diverse and this is a good foundation for genetic improvement of the crop considering that the genetic base of groundnuts genetically stands at a lower level. As a result, these individuals could be used potentially as parents for future breeding.

Analysis of molecular variance among groundnut populations revealed 41 and 59% of the variation was attributable to among individuals with population and within individual variation, respectively. The geographic origin had no impact on the genetic diversity of the crop as it was revealed by low to moderate genetic differentiation observed among the regions. According to standard guidelines for the interpretation of genetic differentiation (Wright, 1978), the range 0.0 - 0.005 indicates little, 0.05 - 0.15 moderate, 0.15 - 0.25 great, and above 0.25 very large genetic differentiations. The results indicate that genetic differentiation was relatively

low (0.041) between South America and Southern America and moderate (0.059) between North America and others. This might be resulted from the high gene flow (3.99 – 5.90) observed among the regions. According to (Slatkin, 1989) and (Morjan and Rieseberg, 2004), gene flow <1 is considered to be low, while N_m = 1 is considered to be moderate and N_m > 1 is considered to be high. Moderate or relatively low levels of gene flow can significantly alleviate the loss of genetic diversity by preventing the effect of genetic drift (Aguilar et al., 2008). The high level of gene flow observed may be attributed to an exchange of genetic materials.
Cluster analysis using SSR markers identified two distinct genetic groups among the studied groundnut genotypes revealing wide genetic diversity for breeding and strategic conservation. The clustering of genotypes was independent of geographical origin in that genotypes from different geographic origin were clustered in the same group. This is common that groundnut genotypes of the same variety were clustered in different groups, as Xiong et al., (2011) reported.

Materials and Methods

Plant material
Fifty-three groundnut genotypes obtained from different countries and were maintained by the Agricultural Research Council-Grain Crops gene bank at Potchefstroom, South Africa were used in the study (Table 1). These accessions were assessed for genetic diversity in Potchefstroom (26°74''S; 27°8'E) and Brits (-25°38'2.39" S 27°46'31.19" E) research farms using agro-morphological descriptors for groundnut (IBPR 1992).

Experimental layout and management
The experiment was laid out in a randomized complete block design replicated three times in Potchefstroom and Brits. Potchefstroom site consisted of hutton soil, whereas Brits was sandy loam soil. The plots consisted of two 4 meters' rows per plot at a spacing of 5 cm between the plants and 90 cm between the rows. The insect pests and disease were controlled by applying chemicals as necessary. The plants were fertilized using lime ammonium nitrate (LAN) just before flowering and irrigated using sprinklers three times per week. Weeding was done manually as well as harvesting.

Agronomic data collection and analysis
The experiment was conduct in 2016/17 cropping season at Potchefstroom and Brits research farms. Data collected included days to 50% flowering, number of branches per plant, leaf length (mm), leaf width (mm), plant throttling, 100 seed weight (g), pod weight/plant (g), pod length (mm), pod width (mm), stem height (mm), and grain yield/plant (Kg). Data were analysed using analysis of variance, principal component analysis and principal coordinate analysis was used to draw a biplot in Genstat version 20 (Payne et al., 2018).

DNA extraction
Fresh leaf samples from 53 groundnut genotypes were collected as required by the standard protocol given by SciCorp Laboratories (Pty) Ltd. Young fresh leaves were harvested from 10 plants of each genotype four weeks after planting. The leaf samples were bulked per genotype and placed in a 1.5 ml eppendorf tube and freeze dried for three days. The dried leaf samples were sealed in a clean small box and posted to SciCorp Laboratories for genotyping, Johannesburg, South Africa. DNA extraction was performed using the standard CTAB extraction protocol. A 100 mg of ground plant tissue was combined with 500 µL of CTAB buffer and incubated for an hour at 65°C. After centrifugation, the supernatant was mixed with phenol: chloroform: iso-amyl alcohol (25:24:1). After a second centrifugation, the DNA was precipitated from the aqueous layer by the addition of a salt and ethanol. The upper aqueous phase that contains the DNA molecule was transferred to a clean microfuge tube. The resulting pellet was dried and re-suspended in TE buffer.

PCR and SSR analysis
All samples were used in bulked amplification, using DNA extracted from the plant material. Twenty selected SSR markers were used to genotype 53 groundnut genotypes (Table 2). The SSR sequences were amplified through polymerase chain reaction (PCR). However, due to the poor quality of DNA extracted from 14 genotypes, only 42 genotypes were included in this analysis. PCR products were fluorescently labelled and separated by capillary electrophoresis on ABI 3130 automatic sequencer (Applied Bio systems, Johannesburg, South Africa).

Data collection and analysis
For the SSR analysis, the fragment size of the amplified products were measured. Two approaches were adopted to investigate the genetic structure and diversity among the groundnut accessions. In the first approach, the amplified products were scored for the presence (1) or absence (0) of alleles. The binary data were then used to obtain a dissimilarity matrix using the Jaccard index. The matrix was used to run a cluster analysis based on neighbor-joining algorithm employing the software DARwin 5.0 (Perrier & Jacquemoud-Collet 2006). However, to assess the genetic structure within and among genotypes, a second approach based on the co-dominant nature of the marker was adopted and analysis was done using GENALEX version 6.5 (Peakall and Smouse, 2012).
Genetic diversity parameters, such as number of alleles per locus (N_a), number of effective alleles per locus (N_e), observed (H_o) and expected (H_e) heterozygosity, and Shannon's Information Index (I) were calculated using GENALEX version 6.5 (Peakall and Smouse 2012) according to the protocol described by (Nei and Li, 1979). The number of polymorphic loci was estimated for each predetermined group, based on geographic origin. Further, an indirect estimate of the level of gene flow (N_m) was calculated using the formula: $N_m = 0.25$ (1 – FST/FST) using GENALEX. The F-statistics such as genetic differentiation (F_{ST}), fixation index or inbreeding coefficient (F_{IS}), and overall fixation index (F_{IT}) were calculated according to Wright's original derivation (Wright 1951). Polymorphic information content (PIC) was calculated using the formula: PIC =1 - ΣP_{ij}^2, where P_{ij} is the frequency of jth allele of the ith locus. Nei's unbiased genetic distance was also estimated to determine the degree of population differentiation among the study material. Nei's unbiased genetic distance and identity were estimated according to (Nei 1978) using GENALEX (Peakall and Smouse, 2006).

Conclusion

The results of this study highlight that the germplasm used in this study were genetically diverse and they can be used as a good foundation to select potential genotypes for further genetic improvement and broadening of the genetic base of

the crop. The SSR and agro-morphological markers proved to be more reliable and efficient to discriminate the genotypes into distinct clusters. However, stratification based on geographic origin has no influence on the genetic diversity of the crop. This suggests that the prevailing DNA and morphological variation rather should dictate future germplasm collection programmes and selections should not be based on geographical background alone. The most diverse genotypes were C. TAFA, RG 1057, RG 1061 and RG 355 based on the SSR markers. These genotypes can be used as parents for hybridization in the downstream breeding programmes.

Acknowledgements

The first author would like to thank Agricultural Research Council-Grain Crops for providing groundnut germplasm and funding for field evaluation, Alliance for a Green Revolution in Africa for funding SSR analysis and Mr Paul Rantso for technical assistance in trial establishment and management.

References

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. Mol Ecol. 17: 5177-5188.

Asibuo JY, Akromah R, Safo-Kantanka O, Adu-Dapaah HK, Ohemeng-Dapaah S, Agyeman A (2008) Chemical composition of groundnut, *Arachis hypogaea* (L) landraces. Afr J Biotech. 7: 2203-2208.

Atasie VN, Akinhanmi TF, Ojiodu CC (2009). Proximate analysis and physico-chemical properties of groundnut *Arachis hypogaea* L. Pak J Nutri. 82: 194-197.

Belay F, Meresa H, Syum SS (2019) Genetic Variation and Association for Kernel Yield and Yield Related Traits of Released Groundnut (*Arachis hypogaea* L.) Varieties in Abergelle District, Northern Ethiopia. Int Plant Breed Crop Sci. 6(1): 474-481.

Bhad PG, Mondal S, Badigannavar AM (2016) Genetic diversity in groundnut (*Arachis hypogaea* L.) genotypes and detection of marker trait associations for plant habit and seed size using genomic and genic SSRs. J Crop Sci Biotech. 19: 203-221.

Bhati PK, Singh SK., Dhurai SY, Sharma A (2015) Genetic divergence for quantitative traits in rice germplasm. Electr J Plant Breed. 6(2): 528-534.

Blum A (1988) Plant breeding for stress environments. CRC Press Inc., Florida.

Bonku R, Yu J (2020) Health aspects of peanuts as an outcome of its chemical composition. Food Sci Hum Well. 9 (2020) 21-30. Ceccarelli S, Acevedo E, Grando S (1991) Breeding for yield stability in unpredictable environments: Single traits, interaction between traits, and architecture of genotypes. Euphytica. 56: 169-185.

Campos-Mondragon MG, Calderon AM, Dela Barca A, Duran-Prado, Campos-Reyes LC, Oliart-Ros RM, Ortega-Garcia J, Medina-Juarez LA, Angulo O (2009). Nutritional composition of new peanut (*Arachis hypogaea* L.) cultivars, Grasos Y Aceites. 60: 161-167. https://doi.org/10.3989/gya.075008.

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut *(Arachis hypogaea*). BMC Plant Biol. 8:55.

Dwivedi S, Gurtu S, Chandra S, Yuejin W, Nigam S (2001) Assessment of genetic diversity among selected groundnut germplasm. I: RAPD analysis. Plant Breed. 120: 345-349.

FAOSTAT (2019) Database of Agricultural Production. Food and Agriculture Organization of the United Nation (FAO). http://faostat.fao.org/default.aspx. Accessed July 2019.

Idi-Garba N, Bakasso Y, Zaman-Allah M, Atta S, Mamane MI, Adamou M, Hamidou F, Idi S, Mahamane A, Saadou M (2015) Evaluation of agro-morphological diversity of groundnut (*Arachis hypogaea* L.) in Niger. Afr J Agric Res. 10: 334-344.

Ingale S, and Shrivastava SK (2011) Nutritional study of new variety of groundnut (Arachis hypogaea L.) JL-24 seeds. Afr J Food Sci. 5(8): 490-498.

Goddard KA, Hopkins PJ, Hall JM, Witte JS (2000) Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. Am J Hum Genet. 66: 216-234.

Halward TM, Stalker HT, Larue EA, Kochert G (1991) Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome. 6: 1013-1020.

He G, Meng R, Gao H, Guo B, Gao G, Newman M, Pittman RN, Prakash C (2005) Simple sequence repeat markers for botanical varieties of cultivated peanut (*Arachis hypogaea* L.). Euphytica. 142: 131-136.

He G, Meng R, Newman M, Gao G, Pittman RN, Prakash CS (2003) Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biol. 3:3.

He G, Prakash CS (1997) Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). Euphytica. 97:143-149.

Herselman L (2003) Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica. 133: 319-327.

Rosenberg NA, Li LM, Ward R, Pritchard JK (2003) Informativeness of genetic markers for inference of ancestry. Am. J. Hum. Genet. 73:1402-1422.

Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM (2018) Genstat® for Windows TM. 19th Edition, Introduction, VSN International, Oxford.

Ilesanmi FF and Ilesanmi OS (2011) Knowledge of aflatoxin contamination in groundnut and the risk of its ingestion among health workers in Ibadan, Nigeria. Asian Pac J Trop Biomed. 493-495.

Kanyika BN, Lungu D, Mweetwa AM, Kaimoyo E, Njung'e VM, Monyo ES, Siambi M, He G, Prakash CS, Zhao Y (2015) Identification of groundnut (*Arachis hypogaea*) SSR markers suitable for multiple resistance traits QTL mapping in African germplasm. Electr J Biotech. 18:61-67.

Knauft D, Chiyembekeza A, Gorbet D (1992) Possible reproductive factors contributing to outcrossing in peanut (*Arachis hypogaea* L.). Peanut Sci. 19: 29-31.

Koppolu R, Upadhyaya HD, Dwivedi SL, Hoisington DA, Varshney RK (2010) Genetic relationships among seven sections of genus *Arachis* studied by using SSR markers. BMC Plant Biol. 10:15.

Krishna GK, Zhang J, Burow M, Pittman RN, Delikostadinov SG, Lu Y, Puppala N (2004) Genetic diversity analysis in Valencia peanut (*Arachis hypogaea* L.) using microsatellite markers. Cell Mol Biol Lett. 9:.685-697.

Mace ES, Phong DT, Upadhyaya HD, Chandra S, Crouch JH (2006) SSR analysis of cultivated groundnut (*Arachis*

hypogaea L.) germplasm resistant to rust and late leaf spot diseases. Euphytica. 152: 317-330.

Mahesh RH, Khan H, Temburne BV, Janila P and Amaregouda A (2018) Genetic variability, correlation and path analysis studies for yield and yield attributes in groundnut (Arachis hypogaea L.) studies for yield and yield attributes in groundnut (*Arachis hypogaea* L.). J. J. Pharmacogn. Phytochem. 7(1): 870-874.

Matus I, Hayes P (2002) Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. Genome. 46: 1095-1106.

Moretzsohn, Hopkins MS, Mitchell SE, Kresovich S, Valls JF, Ferreira ME (2004) Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol. 4:11.

Mupunga I, Mngqawa P, Katerere D (2017) Peanuts, aflatoxins, and undernutrition in children in Sub-Saharan Africa. Nutrients. 9 (2017): 1287.

Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Mol Ecol. 13:1341-1356.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet. 89: 583-590.

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci 76: 5269-5273.

Pandey MK, Monyo E, Ozias-Akins P, Liang X, Guimarães P, Nigam SN (2012) Advances in Arachis genomics for peanut improvement. Biotechnol Adv. 30: 631-651. http://dx.doi.org/10.1016/j.biotechadv.2011.11.001.

Peakall R, Smouse P (2012) GENALEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research an update. Bioinforma. 28: 2537-2539.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Resour. 6: 288-295.

Perrier X, Jacquemoud-Collet J (2006) DARwin software: Dissimilarity analysis and representation for windows. Website http://darwin cirad fr/darwin [accessed 1 March 2019].

Rao, V R (1987) Origin, Distribution, and Taxonomy of Arachis and Sources of Resistance to Groundnut Rust (*Puccinia arachidis* Speg.). In: Groundnut Disease Proceedings of a Discussion Group Meeting, 24 - 28 September 1984, ICRISAT Patanceru, India.

Reddy L, Nigam S, Reddy A (1993) Natural outcrossing in groundnut and its implications in groundnut breeding. J Oilseeds Res. 10:99-104.

Redae G, Assefa D and Habtu S (2017) Effect of integrated agronomic management practices on yield and yield components of groundnut in Abergelle, Tigray, Ethiopia. Afr J Agric Res. 12(35):2722-2728.

Sai DK, Rani KR, Usha R (2016) Assessment of genetic diversity in groundnut (*Arachis hypogaea* L.) Genotypes using PCR based RAPD markers. Biotech Res 2:142-147.

Sibhatu S, Harfe M and Tekle G (2017). Groundnut (*Arachis hypogaea* L.) varieties evaluation for yield and yield components at Tanqua-Abergelle district, Northern Ethiopia. Sky J. Agric Res. 6(3): 057- 06.

Slatkin M (1989) Detecting small amounts of gene flow from phylogenies of alleles. Genet. 121:609-612.

Tang R, Gao G, He L, Han Z, Shan S, Zhong R, Zhou C, Jiang J, Li Y, Zhuang W (2007) Genetic diversity in cultivated groundnut based on SSR markers. Journal of Genetics and Genomics. 34: 449-459.

Upadhyaya HD, Swamy BPM, Goudar PVK, Kullaiswamy BY, Singh S (2005) Identification of diverse groundnut germplasm through multienvironment evaluation of a core collection for Asia. Field Crop Res. 93: 293-299.

Wright S (1951) The genetical structure of populations. Annals of Eugenics. 15: 323-354.

Wright S (1978) Variability within and among populations. Evolution and the Genetics of Populations. University of Chicago Press, USA.

Xiong F, Zhong R, Han Z, Jiang J, He L, Zhuang W, Tang R (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) genotypes. Mol Biol Rep. 38: 3487-3494.

Zaman M, Tuhina-Khatun M, Ullah M, Moniruzzamn M, Alam K (2011) Genetic variability and path analysis of groundnut (*Arachis hypogaea* L.). Agricul. 9: 29-36.

Zongo A, Khera P, Sawadogo M, Shasidhar Y, Sriswathi M, Vishwakarma MK, Sankara P, Ntare BR, Varshney RK, Pandey MK (2017) SSR markers associated to early leaf spot disease resistance through selective genotyping and single marker analysis in groundnut (*Arachis hypogaea* L.). Biotech Rep. 15:132-137.