Pertanika J. Trop. Agric. Sc. 42 (3): 939 - 955 (2019)

PERTANIKA JOURNALS

TROPICAL AGRICULTURAL SCIENCE

Journal homepage: http://www.pertanika.upm.edu.my/

Genome Wide Association Studies for Fatty acids, Mineral and Proximate Compositions in Groundnut (*Arachis hypogaea* L.) Seeds

Abdulwahab Saliu Shaibu[1,2]*, Babu Motagi[3,4] and Peter Ayittah[5]

[1]*Department of Agronomy, Bayero University, P. M. B. 3011, 700001 Kano, Nigeria*
[2]*The National Engineering Laboratory for Crop Molecular Breeding, MOA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, 100081 Beijing, China*
[3]*International Crop Research Institute for the Semi-Arid Tropics, P. M. B. 3112, 700001 Kano, Nigeria*
[4]*Department of Genetics and Plant Breeding, Agriculture College, University of Agricultural Science, Dharwad, 580001 Karnataka, India*
[5]*Pan African University Institute of Life and Earth Sciences, University of Ibadan, Ibadan, 200261 Oyo, Nigeria*

ABSTRACT

Groundnut is basically grown for its oil in most countries and the quality of the oil depends on the total oil, protein and fatty acid compositions in the seeds. The objective of this study was to identify markers that were associated with fatty acids, minerals and proximate composition in groundnut seeds. One hundred and seventy groundnut collections were evaluated in the dry season of 2017 at the research field of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Bayero University Kano. Marker trait association was calculated using the GAPIT package via the KDCompute interface. Significant differences were observed between the genotypes for all the trait measured except for moisture content, crude fiber, crude fat, crude protein, dry matter and nonadenic acid. The heritability values of traits ranged from 0.04 to 0.48. A total of 144 highly significant ($P<0.001$) MTAs with 46 markers for fatty acids (118), minerals (4) and proximate (22) compositions were identified. Most of the markers identified possible MTA in both the A and B genomes. Validation studies are needed to find if these markers are identifying one locus or perhaps a locus duplicated in the two genomes.

Keywords: Groundnut, fatty acids, marker trait association, minerals, proximate content

ARTICLE INFO

Article history:
Received: 01 April 2019
Accepted: 24 May 2019
Published: 19 August 2019

E-mail addresses:
asshuaibu.agr@buk.edu.ng (Abdulwahab Saliu Shaibu)
b.n.motagi@cgiar.org (Babu Motagi)
all4peter@gmail.com (Peter Ayittah)
* Corresponding author

ISSN: 1511-3701
e-ISSN: 2231-8542

INTRODUCTION

Groundnut (*Arachis hypogaea* L.) is an essential oil, food and feed crop (Janila et al., 2013b) and it is cultivated in over 27.94 million ha with a total production of 47.1 million tons in 2017 (Food and Agriculture Organization of the United Nations Statistical Databases [FAOSTAT], 2017). The crop is ranked fifth in terms of oil among crops in the world (FAOSTAT, 2014). It is a rich source of protein, fat, minerals, and vitamins. In most countries, groundnut is principally cultivated for its oil but the demand for groundnut as wholesome food has been increasing due to the health benefits associated with consumption of the nutrient-dense peanut kernels (Upadhyaya et al., 2012b). Groundnut is normally used for oil extraction for eatable and industrial purposes which quality depends on the total oil, protein and fatty acid compositions in the seeds. Oleic acid, a monounsaturated fatty acid, and linoleic acid, a polyunsaturated acid, together with a saturated palmitic acid constitute the major bulk (>90%) of fatty acids in groundnut (Upadhyaya et al., 2012b). Oleic acid enhances the shelf life of groundnut products and have other health benefits (Carlson, 1995; Frankel, 1991; Fraser et al., 1997; Upadhyaya et al., 2012b). Groundnut cultivars with high oleic and oil content for oil extraction and those with high oleic and high protein content for groundnut products are preferred, but the efforts to breed for such cultivars are lacking especially in developing countries due to insufficient genetic variability for the traits (Upadhyaya et al., 2012b).

Micronutrient malnutrition as a result of Fe and Zn deficiencies alone affect over 3 billion people around the world as indicated by Upadhyaya et al. (2012a). The widespread of micronutrient malnutrition has led to huge negative socioeconomic impact that cut across all levels of society (Darnton-Hill et al., 2005; Stein, 2010). Efforts at ICRISAT and other places have led to the identification or development of groundnut cultivars with variation in protein, oil content and quality (Upadhyaya et al., 2012b). However, there are no intensive efforts to identify sources of essential minerals such as Fe and Zn (Upadhyaya et al., 2012a). There is therefore a need to develop nutrition rich groundnut cultivars that will meet the demands of mostly developing countries. Identification of markers that are linked to nutritional traits in groundnut will help in fast tracking breeding process for release of nutritional enhanced groundnut cultivars. The identified markers can be used for marker assisted selection. With the development of genomic tools, marker assisted breeding has been used to improve efficiency of selection for traits of interest in groundnut (Agarwal et al., 2018; Janila et al., 2013a; Pandey et al., 2012, 2014;Varshney et al., 2013). The objective was to identify markers that are associated with fatty acids, minerals and proximate composition in groundnut seeds.

MATERIALS AND METHODS

The study was carried out on the research field of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) at

Bayero University Kano, Nigeria. One hundred and seventy groundnut collection were evaluated for fatty acids, minerals and proximate compositions in the 2017 dry season. The experimental design used was randomized incomplete alpha-lattice (10 x 17) with three replications. Each plot consisted of single row measuring 5 m with inter and intra row spacing of 75 cm and 10 cm, respectively. There was 1 m alley between replications. A total of 40 seeds were planted on each row. One seed was planted per hole at a spacing of 10 cm between holes. Basal application of NPK was done to all plots at the rate of 20 kg ha^{-1} N, 40 kg ha^{-1} P_2O_5 and 40 kg ha^{-1} K_2O. Hand weeding was done at 3, 8, and 12 weeks after planting (WAP) to prevent weed infestation and competition. The field was irrigated to provide optimum growth.

Biochemical Analysis

A total of fifty (50) lines were selected from the 170 genotypes based on cluster analysis. From the clusters, 20 high, 10 medium and 10 low yielding varieties in addition to 10 check varieties were selected for the biochemical analysis. The proximate and mineral compositions of the lines were determined using the Buck Scientific Atomic Absorption Spectrophotometer following the Standard Official Method of Analysis procedures described by Association of Official Analytical Chemists (AOAC) (1984, 1994, 1996). Briefly, 0.5g of each sample was weighed into a 50 ml beaker and 10 ml of an acid mixture of nitric acid and per chloric acid in the ratio 2:1 was added to the content in the beaker and placed on a hot plate to undergo digestion at 105˚C for about 20 minutes until the colour changed to colourless. The digest was allowed to cool and made up to 25 ml with distilled water. The 25 ml was introduced into the Buck Scientific Atomic Absorption Spectrophotometer model 210/211 VGP to determine the concentration of the element in the digest. Fatty acid determination was carried out using gas chromatography. Briefly, 3g of each sample was weighed and soaked in 10ml of N-hexane for two days after which the samples were filtered and weighed. Oil extracted from the filtration was collected into a vacuum tube and covered. The oil was weighed into glass vial and capped; 4.0 ml of petroleum ether and 0.5 ml of Na-methoxide were also added and shaken to dissolve. This was allowed to stand for 1-2 hours till a clear solution was formed. Acid formed in the process was neutralized by sodium glyceroxide and the solution was then pipetted out and injected into the gas chromatographic system (7890B GC) for measurement of fatty acids.

Analysis of variance was performed using PROC GLM in Statistical Analysis System (SAS 9.3) using RANDOM statement with the TEST option.

DNA Extraction and DArT Sequencing

Groundnut leaves of 50 genotypes were collected into 96 deep well samples collection plates and sent to Integrated Genotyping Service and Support (IGSS) platform located at Biosciences Eastern and Central Africa (BecA-ILRI) Hub in

Nairobi for genotyping. DNA extraction was done using NucleoMag® plant genomic DNA extraction kit. The genomic DNA extracted was in the range of 50-100ng/µl. DNA quality and quantity were checked on 0.8% agarose. Libraries were constructed according to Kilian et al. (2012). DArTSeq complexity reduction method through digestion of genomic DNA and ligation of barcoded adapters was done followed by PCR amplification of adapter-ligated fragments. Libraries were sequenced using single read sequencing runs for 77 bases. Next generation sequencing was carried out using Hiseq2500.

DArTseq markers scoring was achieved using DArTsoft version 14, which is an in-house marker scoring pipeline based on algorithms. Two types of DArTseq markers were scored, SilicoDArT markers (scored as presence or absence, 1, 0) and biallelic SNP markers which were both scored for presence of the reference allele, the alternative allele, or both in genomic representation of the sample. Both SilicoDArT markers and SNP markers were aligned to the reference genomes of *Arachis duranensis* (V14167, A-genome ancestor) and *A. ipaensis* (K30076, B-genome ancestor) to identify chromosome.

Linkage Disequilibrium and Marker Trait Association

Linkage disequilibrium between SNPs on each chromosome was measured using TASSEL 5.0 (Bradbury et al., 2007). Marker trait association analysis, probability values and % of the effect of these markers were calculated using the GAPIT package via the KDCompute interface (https://kdcompute.igss-africa.org/kdcompute/home). SNPs with MAF <5% and missing data >20% were excluded. Missing values were imputed using the choice of nearest neighbour algorithm using TASSEL 5.0 (Bradbury et al., 2007). We used the unweighted pair-group method to cluster the lines and form a dendrogram using KDCompute.

RESULTS

Biochemical Analysis of the Groundnut Genotypes

The heritability values of traits ranged from 0.04 for crude fat and palmitic acid to 0.48 for linoleic acid. Oleic acid had a heritability of 0.47. Significant differences were observed between the genotypes for all the trait measured except for moisture content, crude fiber, crude fat, crude protein, dry matter and nonadenic acid (Table 1).

Table 1
Means, minimum, maximum and heritability values for chemical compositions of the groundnut genotypes

Traits	Mean	Minimum	Maximum	Heritability
		Fatty acids		
Stearic acid (%)	0.5**	0	3.3	0.34
Lauric acid (%)	0.1**	0	0.9	0.2
Palmitic acid (%)	0.0**	0	0.1	0.04
Oleic acid (%)	2.7**	0	19.1	0.47

Table 1 *(continue)*

Traits	Mean	Minimum	Maximum	Heritability
Linoleic acid (%)	8.3**	0.1	40.8	0.48
Nonadenic acid (%)	0.1ns	0	0.2	0.08
Arachidic acid (%)	0.0**	0	0.2	0.05
Behenic acid (%)	0.1**	0	0.4	0.12
Tricosanoic acid (%)	0.0**	0	0.1	0.04
Tridecanoic acid (%)	0.1**	0	0.9	0.22
		Minerals		
Iron (mg/kg)	145.0**	77	298	0.39
Zinc (mg/kg)	46.0**	30	77.3	0.37
		Proximate		
Moisture (%)	7.3ns	5.7	9	0.43
Crude fibre (%)	4.3ns	4	4.7	0.44
Crude fat (%)	43.4ns	12.2	48.1	0.04
Ash (%)	2.2**	1	4	0.13
Crude protein (%)	20.6ns	16.1	28	0.16
Carbohydrate (%)	22.3**	15.5	45.9	0.19
Dry matter (%)	93.0ns	91	98	1

**=significant at 0.01 level of probability, NS= non-significant at 0.05

Marker Data

The DArTseq genotyping produced 3591 biallelic SNP markers of which 3396 had a call rate that exceeded 0.6999 and the call rate ranged from 0.37 to 1 (Table 2). Of the 3396 markers, just 396 had a minor allele frequency that exceeded 0.05. The average polymorphism information content of the 3396 markers ranged from 0.006 – 0.499. A principal component analyses of the data from the 3124 markers assigned to a chromosome (s) did not reveal a strong discernible population structure in the first PC that accounted for 61% of the variation.

Table 2
Summary of biallelic SNP marker data

Biallelic SNP	Minimum	Maximum
Allele count A	1	118
E-value A	9.65E-29	2.62E-10
Allele count B	1	100
E-value B	1.18E-28	4.35E-10
Polymorphic information content (PIC)	0.006	0.499
Call rate	0.374	1
One ratio SNP	0.006	1
Frequency of homozygous	0.006	0.994
Frequency of heterozygous	0.006	0.799

The DArTseq genotyping produced 12,693 dominant silico markers all with a call rate that exceeded 0.80 (Table 3). Just 2349 (18.5%) of these had a MAF > 0.05. The average polymorphism information content of the 2349 markers ranged from 0.01 – 0.5. Over 76% of the markers aligned with both the A and B genomes were assigned to homeologous chromosomes and the correlation of their position on those two set of homologues was 0.91. Cluster analysis of the markers revealed three discernible groups (Figure 1).

Table 3
Summary of biallelic SNP marker data

Silico markers	Minimum	Maximum
Allele count A	1	49
E-value A	9.65E-29	2.62E-10
Allele count B	1	52
E-value B	1.18E-28	4.35E-10
Polymorphic information content (PIC)	0.011	0.5
Call rate	0.804	1
One ratio SNP	0.006	0.994
Average Read Depth	5	820
Reproducibility	0.95	1

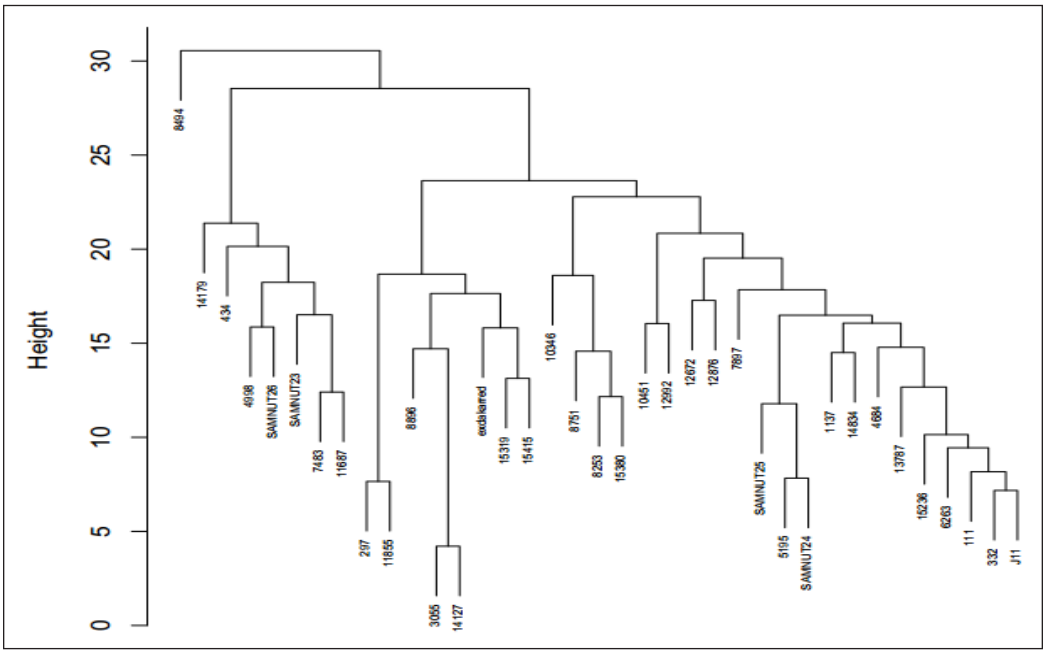


Figure 1. Dendrogram from unweighted pair-group clustering of groundnut accessions using marker data

Linkage Disequilibrium

Linkage disequilibrium (LD) analysis revealed the presence of 305,919 loci pairs. About 36.26% of loci pairs were in significant LD ($P < 0.05$). Further, 9,592 (3.14%) of the pairs were in complete LD ($r^2 = 1$). There was rapid decline in LD with distance and the correlation analysis revealed negative correlation ($r = -0.149$) between the LD (R^2) and the physical distance; as well as between the P-value and R^2 ($r = -0.751$), revealing the existence of linkage decay (Figure 2).

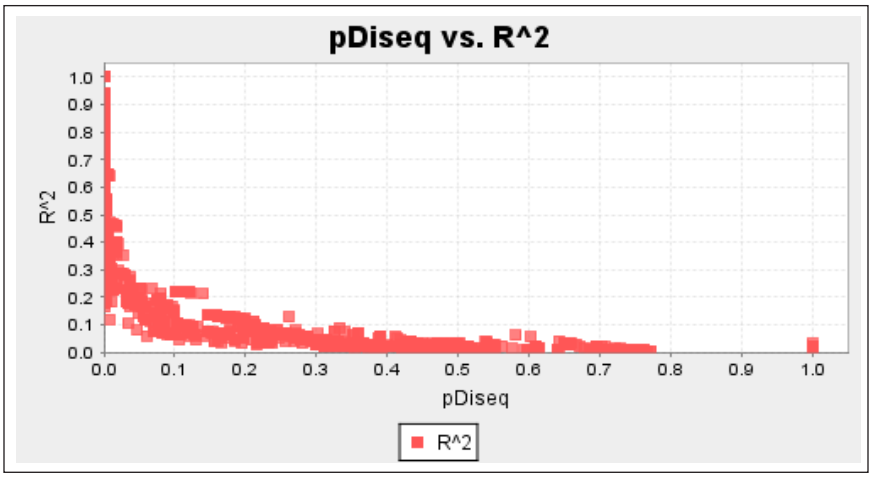


Figure 2. Scatter plot showing association between linkage disequilibrium (r^2) and probability of r^2 (pDiseq)

Marker Trait Association

A total of 144 highly significant (P<0.001) MTAs with 46 markers for fatty acids (118), minerals (4) and proximate (22) compositions were identified (Table 4 and Supplementary Table 1, Supplementary Figure 1 and Supplementary Table 2). Oleic acid (OA) had 16 MTAs with 8 markers that are common to both A and B genomes (Supplementary Figure 1) and explained about 53 – 59 % of the phenotypic variance (PV) observed. Linoleic acid (LNA) had 26 MTAs with 13 markers common to both A and B genomes. Eight out the 13 markers were similar to the markers associated to OA with similar positions in the chromosomes where they were identified (Supplementary Table 1). The markers explained 42 – 58 % of the observed PV for LNA. Two MTAs with one marker was identified for lauric acid (LA) and nonadenic acid (NA). The markers explained about 55 and 98 % of the PV observed for LA and NA, respectively. Numbers of MTAs identified for stearic acid (SA) were 26 with 13 markers each for A and B genome. Of the 13 markers identified, 12 were common to both A and B genome all the markers explained 42 – 64 % of the observed PV. Arachidic acid (AA) had 22 MTAs with 11 markers each on A and B genome which explained 41 – 58 % of the PV. Twenty four (24) MTAs were identified

for tricosanoic acid (TA) and each genome had 12 MTAs with 12 common markers that explained 60 – 66 % of PV. All the MTAs identified were equally distributed on the A and B genomes. Markers, M1 – M13 had significant associations with LNA, SA and AA, while OA had significant associations with M1 – M8 (Supplementary Table 1). Most of the markers were identified on two or more chromosomes and most chromosomes had two or more markers associated to them. No significant MTAs (P>0.001) were identified for palmitic and behenic acids.

For iron (Fe) and zinc (Zn), four MTAs (p<0.001) were identified for Fe while no MTA (p>0.001) was identified for Zn. The markers explained 54 – 56 % of the PV for Fe and two markers were identified on chromosome A08. Significant MTAs was identified for only dry matter (DM) among all the proximate compositions determined. Twenty two MTAs were identified for DM with 10 markers on the A genome and 12 markers on the B genome. The 10 markers on the A genome were the same markers on the B genome with two additional markers and the markers explained 36 – 39 % of the PV observed.

Table 4
Marker-trait associations (MTAs) identified for fatty acids, minerals and proximate composition of groundnut seeds

SN	Trait	No. of MTAs	P. value range	Rsquare range
		Fatty acids composition		
1	Oleic	16	0.000268-000578	0.529-0.585
2	Linoleic	26	0.000181-0.000835	0.424-0.583
3	Palmitic	-	-	-
4	Lauric	2	0.000263-0.000265	0.545-0.552
5	Nonadenic	2	0.0000176-0.0000178	0.976-0.977
6	Stearic	26	0.0000663-0.000754	0.415-0.643
7	Arachidic	22	0.000195-0.000983	0.409-0.578
8	Tricosanoic	24	0.000257-0.000454	0.597-0.656
9	Behenic	-	-	-
		Minerals		
10	Iron	4	0.000797-0.00093	0.542-0.556
11	Zinc	-	-	-
		Proximate		
12	Ash	-	-	-
13	Carbohydrate	-	-	-
14	Crude fiber	-	-	-
15	Crude Protein	-	-	-
16	Crude fat	-	-	-
17	Dry matter	22	0.000746-0.000989	0.361-0.394
18	Moisture	-	-	-

DISCUSSION

The significant differences obtained among the 50 selected groundnut genotypes indicate that genetic variation exists among the genotypes for the traits studied. The significant variability observed for the main fatty acids such as oleic and linoleic acids as well as other fatty acids suggests that sufficient variability exists for the genetic improvement of essential fatty acids in groundnut. Janila et al. (2014) reported genetic variability for Fe and Zn concentrations in groundnut seeds. No significant difference was observed between the groundnut for moisture content but the mean and maximum value were below the average 14% recommended by Waliyar et al. (2015).

The marker data suggested that the population was not highly structured and more markers were produced in the B genome than the A genome. Many polymorphic markers were detected with large portion having MAF < 0.05 with average PIC values of about 0.07. Bertioli et al. (2016) had indicated that groundnut had a low polymorphism rate and low genetic diversity. The linkage disequilibrium (LD) declined with distance and probability. Pandey et al. (2014) and Mwadzingeni et al. (2017) had earlier reported rapid LD decay with distance.

The marker trait association (MTAs) studies revealed 144 significant MTAs ($p < 0.001$) involving 46 markers. Most of the markers identified possible MTA in both the A and B genomes. Validation studies will be needed to see if these markers are identifying one locus or perhaps a locus duplicated in the two genomes.

Pandey et al. (2014) used SSR markers and identified some MTAs for oil, oleic acid, protein and zinc content. These MTAs were located majorly on chromosome A06 and B06 and explained up to 40% of the PV. In our study, some markers were identified for OA, LNA, AA, TA and DM on chromosome A06 and B06 and the markers explained more than 90% of the observed PV which doubled what was reported by Pandey et al. (2014). This may be possible because SNP markers are more informative than SSR markers. Despite the similarities of the reported chromosomes by Pandey et al. (2014) with ours, there are no supporting evidence that the positions are similar. Zhang et al. (2018) identified four MTAs with three markers for OA and three MTAs with markers for LNA. The MTAs were located on chromosome A09, A10 and B08, and the markers were similar for both OA and LNA and located in the same position. In our study we also identified MTAs on chromosome A09, A10 and B08. The markers associated to both OA and LNA which constitute more than 80% of the fatty acids in groundnut were similar and in the same position as also reported by Zhang et al. (2018). The results suggest some possible associations between these traits and may explain why OA:LNA ratio increases with increase in oleic acid percentage as reported by Upadhyaya et al. (2012b) and Zhang et al. (2018). There are many minor effect QTLs, or genes controlling oleic acid and

linoleic acid in groundnut, including the major gene FAD2. In our study, 13 markers were associated with both oleic acid and linoleic acid.

The only available reported MTAs for Fe and Zn was that of Pandey et al. (2014). They identified one MTA for Zn and no MTA for Fe. In our study, we identified four MTAs for Fe and no MTA was identified for Zn at $p<0.001$ but two MTAs with one similar marker on chromosome A04 and B04 at $p=0.0048$ which was above the threshold set for identifying MTA were observed. It is important to also report that Pandey et al. (2014) reported one MTA for Zn on chromosome B04 which is similar to our findings. We could not validate if these markers are identifying one locus or perhaps a locus duplicated in the two genomes because it is only one marker but identified on both chromosomes.

From the result of the MTAs analysis, most of the MTAs identified on the A subgenome were also identified on the respective homeologous chromosome on the B subgenome. Agarwal et al. (2018) had shown that significant proportion of marker loci that were assigned to chromosome of one subgenome were mapped to respective homeologous positions on chromosomes of the other subgenome. Quantitative traits are usually complex and controlled by multiple genes that often have individually small effects (Upadhyaya & Nigam, 1999), and we detected considerable large number of markers for most of the traits because of the large density of markers used in our study. Only a few related markers were detected in the study by Zhang et al. (2018) due to the low density of tested markers.

CONCLUSION

The present study identified a total of 144 highly significant marker trait associations involving 46 markers for nine traits out of the 18 studied traits. Seven fatty acids had significant MTAs while mineral and proximate compositions had one significant MTAs each. The markers identified in this study can serve as useful genomic resources to initiate marker-assisted selection and trait introgression of groundnut for improvement of nutritional and biochemical compositions of groundnut. Further studies are required to validate the significant markers identified in the present study using a larger population.

ACKNOWLEDGEMENTS

The genotyping support was provided by International Livestock Research Institute, Nairobi, Kenya, Biosciences Eastern and Central Africa (BecA-ILRI Hub), Nairobi, Kenya and Diversity Arrays Technology Pty Ltd (DArT), Australia through the Integrated Genotyping Service And Support (IGSS) platform for Africa. We also acknowledge the support of International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Kano Station, Nigeria, for the field phenotyping and funding some part of the genotyping work.

REFERENCES

Agarwal, G., Clevenger, J., Pandey, M. K., Wang, H., Shasidhar, Y., Chu, Y., … Varshn, B. (2018). High-density genetic map using whole-genome resequencing for fine mapping and candidate gene discovery for disease resistance in peanut. *Plant Biotechnology Journal*, *16*(11), 1954–1967. doi: 10.1111/pbi.12930

Association of Official Analytical Chemists. (1984). *Official methods of analysis: Association of Official Analytical Chemists* (14th ed.). Gaithersburg, USA: AOAC International.

Association of Official Analytical Chemists. (1994). *Official methods of analysis: Association of Official Analytical Chemists* (16th ed.). Gaithersburg, USA: AOAC International.

Association of Official Analytical Chemists. (1996). *Official methods of analysis: Association of Official Analytical Chemists* (16th ed.). Gaithersburg, USA: AOAC International.

Bertioli, D. J., Cannon, S. B., Froenicke, L., Huang, G., Farmer, A. D., Cannon, E. K., … Ozias-Akins, P. (2016). The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature Genetics*, *48*(4), 438–446. doi: 10.1038/ng.3517

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. doi: 10.1093/bioinformatics/btm308

Carlson, S. E. (1995). The role of PUFA in infant nutrition. *Inform*, *6*(8), 940–946.

Darnton-Hill, I., Webb, P., Harvey, P. W., Hunt, J. M., Dalmiya, N., Chopra, M., … Ball, M. J. (2005). Micronutrient deficiencies and gender: Social and economic costs. *American Journal of Clinical Nutrition*, *81*(5), 1198S–1205S. doi: 10.1093/ajcn/81.5.1198

Food and Agriculture Organization of the United Nations Statistical Databases. (2014). *Crops processed*. Retrieved March 23, 2019, from http://www.fao.org/faostat/en/#data/QD

Food and Agriculture Organization of the United Nations Statistical Databases. (2017). *Crops*. Retrieved March 23, 2019, from http://www.fao.org/faostat/en/#data/QC

Frankel, E. N. (1991). Recent advances in lipid oxidation. *Journal of the Science of Food and Agriculture*, *54*(4), 495–511.

Fraser, G. E., Sumbureru, D., PribiEw, P., Neil, R. L., & Frankson, M. A. C. (1997). Association among health habits, risk factors, and all-cause mortality in a black California population. *Epidemiology*, *8*(2), 168–174. doi: 10.1097/00001648-199703000-00008

Janila, P., Nigam, S. N., Pandey, M. K., Nagesh, P., & Varshney, R. K. (2013a). Groundnut improvement: Use of genetic and genomic tools. *Frontiers in Plant Science*, *4*(2), 1–16. doi: 10.3389/fpls.2013.00023

Janila, P., Venuprasad, R., Abhishek, R., U., A., Reddy, K. R., Waliyar, F., & Nigam, S. N. (2013b). Genetic analysis of resistance to late leaf spot in interspecific groundnuts. *Euphytica*, *193*(1), 13–25. doi: 10.1007/s10681-013-0881-7

Janila, P., Varshney, R. K., Pandey, M. K., Nigam, S. N., Sudini, H., Gowda, M. V. C., … Nagesh, P. (2014). Iron and zinc concentrations in peanut (*Arachis hypogaea* L.) seeds and their relationship with other nutritional and yield parameters. *The Journal of Agricultural Science*, *127*(5), 1771–1781. doi: 10.1017/S0021859614000525

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., … Uszynski, G. (2012). Diversity arrays technology: A generic genome profiling technology on open platforms. In. F. Pompanon & A. Bonin (Eds.), *Data production and analysis*

in population genomics: Methods in molecular biology (pp. 67-89). Totowa, USA: Humana Press.

Mwadzingeni, L., Shimelis, H., Rees, D. J., & Tsilo, T. J. (2017). Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. *PLoS One*, *12*(2), e0171692. doi: 10.1371/journal.pone.0171692

Pandey, M. K., Monyo, E., Ozias-Akins, P., Liang, X., Guimaraes, P., Nigam, S. N., … Varshney, R. K. (2012). Advances in Arachis genomics for peanut improvement. *Biotechnology Advances*, *30*(3), 639–651. doi: 10.1016/j.biotechadv.2011.11.001

Pandey, M. K., Upadhyaya, H. D., Rathore, A., Vadez, V., Sheshshayee, M. S., Sriswathi, M., … Varshney, R. K. (2014). Genomewide 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. doi: 10.1371/journal.pone.0105228

Stein, A. J. (2010). Global impacts of human mineral malnutrition. *Plant and Soil*, *335*(1), 133–154.

Upadhyaya, H. D., Dronavalli, N., Singh, S., & Dwivedi, S. L. (2012a). Variability and stability for kernel iron and zinc contents in the icrisat mini core collection of peanut. *Crop Science*, *52*(6), 2628. doi: 10.2135/cropsci2012.05.0306

Upadhyaya, H. D., Mukri, G., Nadaf, H. L., & Singh, S. (2012b). Variability and stability analysis for nutritional traits in the mini core collection of peanut. *Crop Science*, *52*(1), 168. doi: 10.2135/cropsci2011.05.0248

Upadhyaya, H., & Nigam, S. N. (1999). Detection of epistasis for protein and oil contents and oil quality parameters in peanut. *Crop Science*, *39*(1), 115–118.

Varshney, R. K., Mohan, S. M., Gaur, P. M., Gangarao, N. V, Pandey, M. K., Bohra, A., … Gowda, C. L. (2013). Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. *Biotechnology Advances*, *31*(8), 1120–1134. doi: 10.1016/j.biotechadv.2013.01.001

Waliyar, F., Osiru, M., Ntare, B. R., Kumar, K. V. K., Sudini, H., Traore, A., & Diarra, B. (2015). Post-harvest management of aflatoxin contamination in groundnut. *World Mycotoxin Journal*, *8*(2), 245–252. doi: 10.3920/Wmj2014.1766

Zhang, X., Zhu, S., Zhang, K., Wan, Y., Liu, F., Sun, Q., & Li, Y. (2018). Establishment and evaluation of a peanut association panel and analysis of key nutritional traits. *Journal of Integrative Plant Biology*, *60*(3), 195–215. doi:10.1111/jipb.12601

APPENDIX

Supplementary Table 1

Supplementary information for markers with significant p values

Trait	SNP	Chromosome	Position	P.value	Rsquare
Oleic	M1	A09	16232985	0.000268	0.58456
	M2	A08	37481891	0.000292	0.577639
	M3	A09	4215070	0.000313	0.572082
	M4	A08	5671449	0.000322	0.569869
	M5	A06	10156052	0.000328	0.568253
	M6	A07	4811754	0.000333	0.567016
	M7	A10	11042529	0.00041	0.550656
	M8	A06	87914790	0.000496	0.535798
	M1	B09	21044599	0.000316	0.575865
	M2	B08	23091556	0.000346	0.568617
	M3	B09	5249774	0.000377	0.561866
	M4	B07	110098705	0.000382	0.561002
	M5	B06	45124360	0.000396	0.558062
	M6	B07	4744979	0.000397	0.557931
	M7	B10	17082028	0.000508	0.538937
	M8	B06	109255383	0.000578	0.529013
Linoeic	M1	A09	16232985	0.000181	0.582842
	M2	A08	37481891	0.000182	0.582553
	M3	A09	4215070	0.000187	0.579766
	M4	A08	5671449	0.000197	0.574978
	M5	A06	10156052	0.000206	0.571331
	M6	A07	4811754	0.000207	0.570598
	M8	A06	87914790	0.000211	0.568902
	M7	A10	11042529	0.00024	0.557678
	M9	A10	100626921	0.000656	0.471341
	M10	A02	93813091	0.000693	0.466757
	M11	A10	78514031	0.000724	0.463168
	M12	A06	110005450	0.000778	0.457348
	M13	A03	105924455	0.000835	0.451567
	M1	B09	21044599	0.000191	0.554382
	M2	B08	23091556	0.000192	0.553852
	M3	B09	5249774	0.000198	0.551163
	M4	B07	110098705	0.00021	0.545689
	M5	B06	45124360	0.000218	0.54213
	M6	B07	4744979	0.00022	0.54116
	M8	B06	109255383	0.000221	0.540816
	M7	B10	17082028	0.000255	0.527179
	M9	B10	126223500	0.000619	0.447052
	M10	B02	108681196	0.00065	0.442763
	M11	B10	103607600	0.000765	0.428652
	M13	B03	107485963	0.000776	0.427451
	M12	B06	134748296	0.000804	0.424307
Lauric	M14	A09	110528207	0.000263	0.545145
	M14	B09	146236499	0.000265	0.552669
Nonadenic	M15	A09	109069638	1.77E-05	0.976985
	M15	B09	146924640	1.78E-05	0.976232

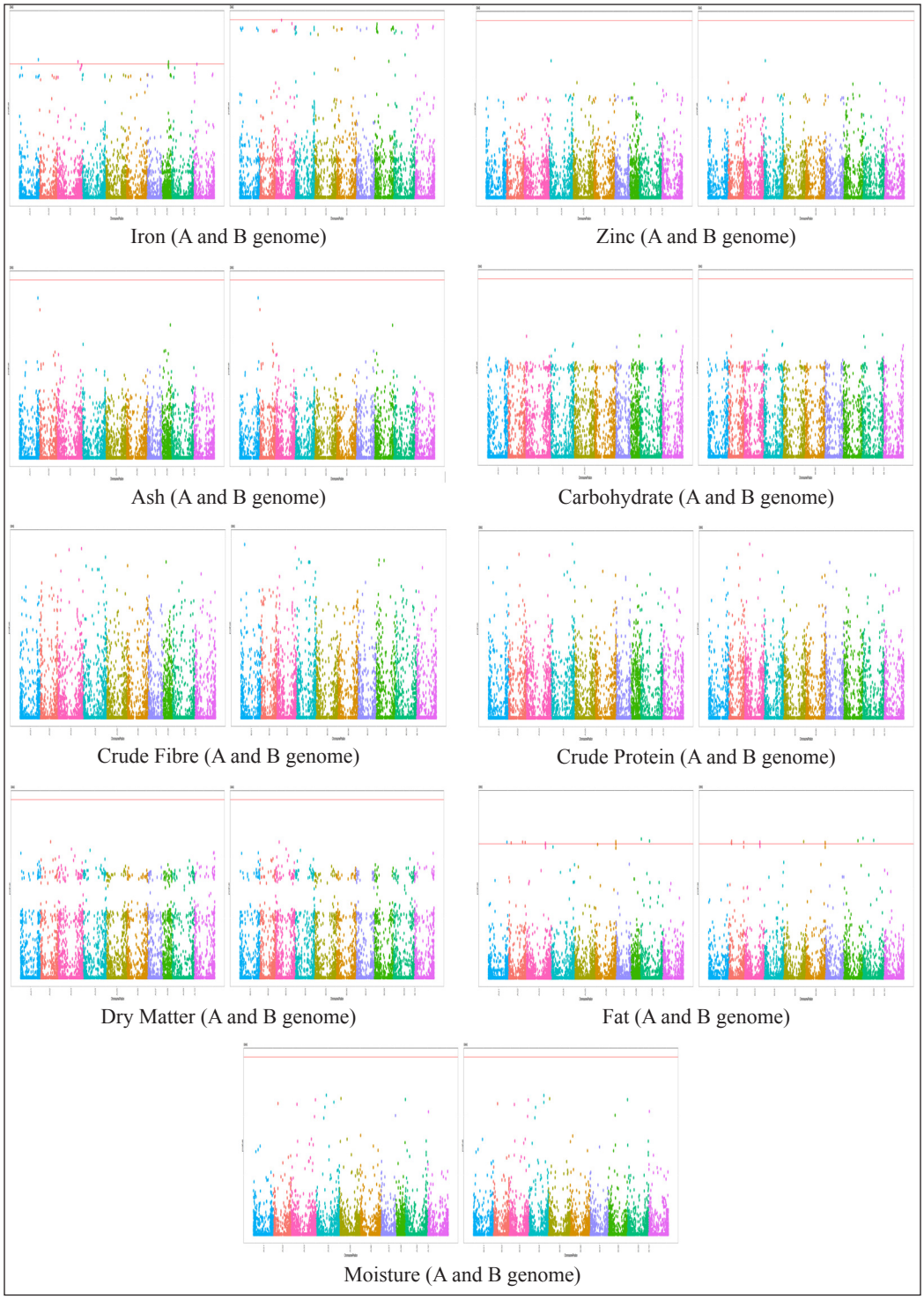
Supplementary Table 1 *(continue)*

Trait	SNP	Chromosome	Position	P.value	Rsquare
Stearic	M2	A08	37481891	6.63E-05	0.642489
	M1	A09	16232985	6.79E-05	0.640277
	M3	A09	4215070	6.95E-05	0.638011
	M4	A08	5671449	7.14E-05	0.635537
	M8	A06	87914790	7.27E-05	0.633824
	M5	A06	10156052	7.50E-05	0.630931
	M6	A07	4811754	8.15E-05	0.62302
	M7	A10	11042529	8.78E-05	0.616146
	M11	A10	78514031	0.000165	0.558755
	M16	A07	40769959	0.000428	0.476145
	M13	A03	105924455	0.000512	0.461203
	M9	A10	100626921	0.000687	0.437158
	M10	A02	93813091	0.000754	0.429647
	M2	B08	23091556	9.54E-05	0.574868
	M1	B09	21044599	9.74E-05	0.572849
	M3	B09	5249774	9.88E-05	0.571403
	M8	B06	109255383	0.000101	0.569034
	M17	B07	110098705	0.000105	0.565681
	M5	B06	45124360	0.000108	0.562159
	M6	B07	4744979	0.000121	0.551313
	M7	B10	17082028	0.000128	0.545474
	M11	B10	103607600	0.000248	0.48204
	M16	B06	52503651	0.000361	0.446955
	M13	B03	107485963	0.000416	0.43384
	M9	B10	126223500	0.000487	0.419706
	M10	B02	108681196	0.000515	0.414651
Arachidic	M1	A09	16232985	0.000195	0.577966
	M8	A06	87914790	0.000196	0.576992
	M2	A08	37481891	0.000204	0.572817
	M17	A08	5671449	0.000234	0.557492
	M3	A09	4215070	0.00024	0.554658
	M7	A10	11042529	0.00025	0.550454
	M5	A06	10156052	0.000251	0.549865
	M6	A07	4811754	0.000263	0.544914
	M11	A10	78514031	0.000582	0.461265
	M18	A07	40769959	0.000726	0.438816
	M13	A03	105924455	0.000983	0.408672
	M1	B09	21044599	0.000238	0.566744
	M8	B06	109255383	0.000254	0.560106
	M2	B08	23091556	0.000257	0.558556
	M3	B09	5249774	0.000299	0.54269
	M4	B07	110098705	0.000303	0.541401
	M7	B10	17082028	0.000312	0.538333
	M5	B06	45124360	0.000315	0.537222
	M6	B07	4744979	0.000332	0.531868
	M18	B06	52503651	0.000735	0.451595
	M11	B10	103607600	0.000768	0.447314
	M13	B03	107485963	0.000937	0.428048
Tricosanoic	M19	A01	92846737	0.000257	0.655511

Supplementary Table 1 *(continue)*

Trait	SNP	Chromosome	Position	P.value	Rsquare
	M20	A07	880042	0.000332	0.628646
	M21	A03	31817347	0.000335	0.627809
	M22	A04	106555766	0.000341	0.625965
	M23	A09	1400546	0.000348	0.623736
	M24	A02	24356338	0.000352	0.622545
	M25	A04	93590112	0.000388	0.612734
	M26	A03	129975580	0.000394	0.61098
	M27	A04	53129067	0.000416	0.605431
	M28	A04	45718128	0.000431	0.602066
	M29	A03	96432676	0.00045	0.597596
	M30	A05	8095276	0.000454	0.596717
	M19	B06	106755490	0.000257	0.655434
	M20	B07	620457	0.000332	0.628566
	M21	B03	35192532	0.000335	0.627716
	M22	B03	8173213	0.000341	0.625905
	M23	B09	1681755	0.000349	0.623606
	M24	B02	28213282	0.000353	0.622426
	M25	B04	102341064	0.000388	0.612621
	M26	B03	130875833	0.000395	0.610817
	M27	B04	102428625	0.000417	0.605288
	M28	B04	45429705	0.000431	0.602022
	M29	B01	135265182	0.000451	0.597406
	M30	B05	8494827	0.000454	0.596589
Iron	M31	A01	101409239	0.000797	0.555483
	M32	A03	111434595	0.000893	0.548444
	M33	A08	34891811	0.000901	0.547903
	M34	A08	34361128	0.000993	0.541932
Dry Matter	M35	A09	5710373	0.000774	0.394268
	M36	A09	49143701	0.000863	0.384995
	M37	A06	110497678	0.000884	0.38301
	M38	A02	75296891	0.000907	0.380853
	M39	A01	99031328	0.000913	0.380295
	M40	A02	89569235	0.000926	0.379134
	M41	A02	14768065	0.000945	0.377409
	M42	A06	110512576	0.00095	0.37699
	M43	A03	103594754	0.000954	0.376603
	M44	A03	104218069	0.000989	0.37362
	M35	B09	2764950	0.000746	0.384769
	M36	B09	75930448	0.000834	0.375085
	M37	B08	97153563	0.000851	0.373381
	M39	B02	19708590	0.000872	0.371217
	M38	B05	136889865	0.00089	0.369448
	M43	B03	105334523	0.000903	0.368263
	M40	B02	103154677	0.000917	0.366917
	M42	B06	135281181	0.000921	0.366569
	M41	B02	17949139	0.000941	0.364662
	M44	B03	105920615	0.000967	0.3624
	M45	B02	19667396	0.000974	0.361721
	M46	B06	135672991	0.000988	0.36052

Supplementary Figure 1. Manhattan plots of marker trait associations of fatty acids



Supplementary Figure 2. Manhattan plots of marker trait associations for minerals and proximate composition