Genetics and Molecular Research

Evaluation of insertion-deletion markers suitable for genetic diversity studies and marker-trait correlation analyses in cultivated peanut (*Arachis hypogaea* L.)

S. Meng[1]*, X.L. Yang[1]*, P.M. Dang[3], S.L. Cui[1], G.J. Mu[1], C.Y. Chen[2] and L.F. Liu[1]

[1]North China Key Laboratory for Crop Germplasm Resources of Education Ministry, Laboratory for Crop Germplasm Resources of Hebei, College of Agronomy, Agricultural University of Hebei, Baoding, China
[2]Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, AL, USA
[3]USDA-ARS, National Peanut Research Laboratory, Dawson, GA, USA

*These authors contributed equally to this study.
Corresponding authors: L.F. Liu / C.Y. Chen
E-mail: auhpeanut@163.com / cyc0002@auburn.edu

Genet. Mol. Res. 15 (3): gmr.15038207
Received December 7, 2015
Accepted April 11, 2016
Published August 12, 2016
DOI http://dx.doi.org/10.4238/gmr.15038207

ABSTRACT. Peanut is one of the most important oil crops worldwide. We used insertion-deletion (InDel) markers to assess the genetic diversity and population structure in cultivated peanut. Fifty-four accessions from North China were genotyped using 48 InDel markers. The markers amplified 61 polymorphic loci with 1 to 8 alleles and an average of 2.6 alleles per marker. The polymorphism information content values ranged from 0.0364 to 0.9030, with an average of 0.5038.

Population structure and neighbor-joining (NJ) tree analyses suggested that all accessions could be divided into four clusters (A1-A4), using the NJ method. Likewise, four subpopulations (G1-G4) were identified using STRUCTURE analysis. A principal component analysis was also used and results concordant with the other analysis methods were found. A multi-linear stepwise regression analysis revealed that 13 InDel markers correlated with five measured agronomical traits. Our results will provide important information for future peanut molecular breeding and genetic research.

Key words: Peanut; InDel; Genetic diversity; Marker-trait correlation

INTRODUCTION

Cultivated peanut (*Arachis hypogaea* L.) is an allotetraploid species (2n = 4x = 40, AABB) (Krapovickas and Gregory, 1994) that is considered one of the most important edible oil crops worldwide (Fávero et al., 2015). In addition to oil, peanut seeds are rich in protein and a good source of vitamins and minerals. Germplasm diversity is the backbone for peanut improvement and genetic dissection of complex traits. An extensive evaluation of the genetic diversity of elite germplasm will help us to better understand how to utilize the germplasm collection for genetic enhancement. Moreover, population structure analysis, a prerequisite for association mapping, could further aid our understanding of the genetic diversity in natural populations (Flint-Garcia et al., 2005).

DNA molecular markers have been extensively applied in the assessment of genetic diversity, genotype fingerprinting, genetic linkage map construction, gene map-based cloning, differential expression analysis, and molecular marker-assisted selection in breeding programs. Due to the narrow genetic basis, few informative DNA polymorphic markers have been identified in cultivated peanut. Thus, the development of molecular markers for peanut lags behind that of other crops. Since the first report indicating that cultivated peanut presents some DNA polymorphism (Kochert et al., 1991), restriction fragment length polymorphism (RFLP) and random amplified polymorphic DNA (RAPD) markers have been employed to construct genetic maps and to dissect the genetic diversity in peanut (Garcia et al., 1995; Burow et al., 2001; Raina et al., 2001). In addition, amplified fragment length polymorphism (AFLP) (He and Prakash, 2001; Herselman, 2003), inter-simple sequence repeat (Raina et al., 2001), sequence related amplified polymorphism (SRAP) (Ren et al., 2010), and start codon targeted (SCoT) polymorphism markers (Xiong et al., 2011) have been developed and utilized to study the genetic diversity and relationships in cultivated peanut. Simple sequence repeat (SSR) markers have been widely applied in genetic diversity studies of peanut (Hopkins et al., 1999; He et al., 2003; Ferguson et al., 2004; Barkley et al., 2007; Freitas et al., 2007; Wang et al., 2015), because they are reproducible, simple, and inexpensive. However, SSR markers are mainly distributed in the non-coding regions of the gene and genotyping errors may occur because of stutter bands and technical artifacts (allelic dropouts, null alleles, false alleles, and size homoplasy) (Pompanon et al., 2005). Single nucleotide polymorphisms (SNPs) have proven advantageous for germplasm genetic studies due to their automatic mass detection. However, they require pre-requested sequence information for designing primers and special

equipment for high-throughput genotyping, in addition to incurring high costs (Syvänen, 2001). SNPs have been reported in peanut genetic diversity, for instance Alves et al. (2008) developed SNP markers and analyzed their suitability for the peanut genetic map.

Insertion-deletions (InDels) are co-dominant genetic markers that have received a lot of attention as a genomic marker, due to exhibiting high polymorphism, reliability, as well as fast and easy operation (Väli et al., 2008). Compared with other markers, InDels can be genotyped using simple procedures, and the probability that two InDels of the same length will appear at the same genomic position is very small. In other words, InDels are unique in the genome, and may be considered to show identity-by-descent (Shedlock and Okada, 2000). InDels have been successfully utilized in cultivar identification, diversity analyses, high density genetic map constructions, and fine mapping of target traits in rice, maize, sesame, and soybean (Thornsberry et al., 2001; Fu et al., 2006; Hayashi et al., 2006; Wu et al., 2014; Song et al., 2015). Thus, the use of InDel markers has increased in molecular and genetic studies, in recent years.

In the present study, we evaluated the genetic diversity and population structure of 54 cultivated peanut varieties from North China, using 48 InDel markers. The aims of this study were 1) to assess the potential use of InDel markers in genetic diversity and population structure analysis for cultivated peanut and 2) to detect InDel markers associated with agronomical traits.

MATERIAL AND METHODS

Plant materials and field tests

A total of 54 cultivated peanut accessions representing four botanical subspecies (*A. hypogaea* var. *fastigiata*, *A. hypogaea* var. *hypogaea*, *A. hypogaea* var. *hirsuta*, and *A. hypogaea* var. *vulgaris*) was used in this study (Table 1). In May 2010, the 54 genotypes were planted in the breeding nursery at Baoding, Hebei, China. We used two-row plots that were 0.4 m width and 1.0 m length in a randomized complete block design with three replications. Five plants from each row were randomly selected after the harvest season and the five important agronomic traits were measured according to the phenotypic identification method and criterion of peanut (Jiang et al., 2006), including pod length, pod width, pod thickness, thickness of pod shell, and weight of 100 pods. Among these traits, the first three traits were investigated using slide caliper, and the weight of 100 pods from five plants was measured using electronic balance (10 mg).

DNA extraction and InDel polymerase chain reaction (PCR)

Peanut genomic DNA was extracted as described by Dang and Chen (2013). A total of 48 InDel primer pairs was employed (Liu et al., 2015). The PCR amplification was conducted in a total volume of 10 μL with 45 ng peanut DNA, 10X PCR buffer (TaKaRa, Dalian, China), 0.8 μL dNTP (2.5 mM/μL, TaKaRa), 0.5 μL each primer (10 μM/μL), and 0.5 U *Taq* polymerase (5 U/μL, TaKaRa). The PCR procedures were as follows: 95°C for 5 min; 30 cycles of 94°C for 45 s, 56°C for 45 s, and 72°C for 1 min; and 72°C for 5 min. The PCR products were separated on 10% native-denaturing polyacrylamide gels. The silver staining method was performed as described by Zhang et al. (2000).

Table 1. List of *Arachis hypogaea* accessions used in the study with botanical subspecies.

No.	Variety/Accession No.	Botanical subspecies	No.	Variety/Accession No.	Botanical subspecies
1	Luanpingdahuasheng	*hypogaea*	28	Suningxiaobaguo	*vulgaris*
2	Funingliyang	*hypogaea*	29	Qike	*vulgaris*
3	Dalihuasheng	*hypogaea*	30	Xiekangqing	*vulgaris*
4	Suninghuasheng	*hypogaea*	31	Shitouqi	*vulgaris*
5	Damingdayanghuasheng	*hypogaea*	32	Fuhuasheng	*vulgaris*
6	Hejianbanpaman	*hypogaea*	33	Nongzhanduoli	*fastigiata*
7	Funingdali	*hypogaea*	34	Shenxiansilihong	*fastigiata*
8	Raoyanghuasheng	*hypogaea*	35	Jigulate	*fastigiata*
9	Daminglianhua	*hypogaea*	36	AH66	*fastigiata*
10	Dingxiandahuasheng	*hypogaea*	37	Nanhunsanli	*fastigiata*
11	Yuanshibanmanguo	*hypogaea*	38	Funingduoli	*fastigiata*
12	Qianxiliyang	*hypogaea*	39	Silihong	*fastigiata*
13	Qinghuangdaoliyang	*hypogaea*	40	Long 1	*hirsuta*
14	Funingjiubilou	*hypogaea*	41	Long 2	*hirsuta*
15	Hebeidalidun	*hypogaea*	42	SAU-1	*hirsuta*
16	Beidaran	*hypogaea*	43	SAU-2	*hirsuta*
17	Baodinghong	*hypogaea*	44	SAU-3	*hirsuta*
18	Xianxianhuawo	*hypogaea*	45	SAU-4	*hirsuta*
19	Hejianpafang	*hypogaea*	46	SAU-5	*hirsuta*
20	Pingshanzhonglihuasheng	*vulgaris*	47	SAU-6	*hirsuta*
21	Funingxiaozili	*vulgaris*	48	SAU-7	*hirsuta*
22	Shenxianxiaobaguo	*vulgaris*	49	Feixiansilicao	*hirsuta*
23	Hengshuiyiwohou	*vulgaris*	50	Rugaoxiyangsheng	*hirsuta*
24	Juluxiaohuasheng	*vulgaris*	51	Yinansilicao	*hirsuta*
25	Lulongxiaohuasheng	*vulgaris*	52	Tuokexunxiaohuasheng	*hirsuta*
26	Hejianxiaobaguo	*vulgaris*	53	Changqingyicongsheng	*hirsuta*
27	Yuanshiyiwohou	*vulgaris*	54	Feixiansanlizhong	*hirsuta*

Statistical analysis

The DNA fragment bands were scored as present (1) or absent (0). We calculated the polymorphism information content (PIC) by PIC = 1-**å***xi*^{2}, where *xi* is the relative frequency of the *i*th allele of the locus; and Shannon's information index (*H'*) by *H'* = -**å**Pi x lnPi, where Pi is a primer combination of polymorphisms with probability i to appear in the tested materials.

The clustering analyses were performed using NTSYS-pc 2.1 (Rohlf, 2000) based on the genetic similarity matrices. A neighbor-joining (NJ) algorithm (Saitou and Nei, 1987) was applied to construct a phylogram from the distance matrix, which was generated using MEGA4 (Tamura et al., 2007). A principal component analysis (PCA) was conducted to create plots of the most significant axes for grouping pattern verification using eigenanalysis (Patterson et al., 2006).

The population genetic structure was detected using a fully Bayesian process in STRUCTURE v. 2.3.4 (http://pritchardlab.stanford.edu/home.html). The program was run five times for each subpopulation (*k*) value, ranging from 2 to 10. We used the mixture model with correlated allele frequency, setting 500,000 burn-in replicates and 500,000 replicates for analysis. An initial upper boundary of $k = 9$ was fixed, based on the sample size and the outcome that we found an optimal *k* below this boundary. The final population subgroups were determined using the likelihood plot of these models, the stability of grouping patterns across five runs, and the second order rate of change of the likelihood function (ΔK). Based on this information, we chose $k = 4$ as the optimal grouping (Figure 1). The marker-trait correlated analyses were calculated using the multi-linear stepwise regression of SPSS 17.0 statistical software.

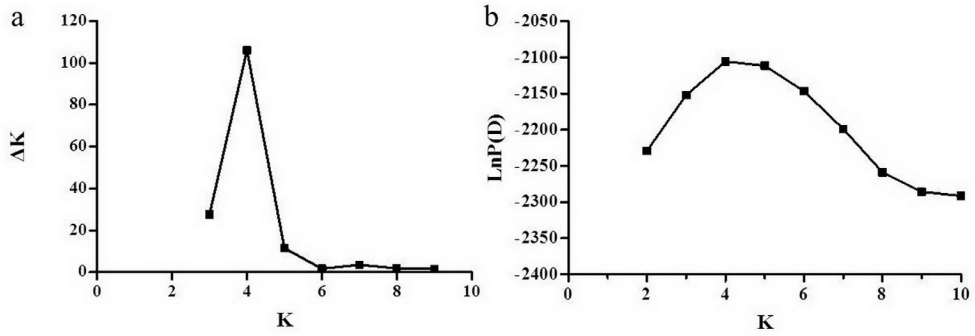

Figure 1. Relationships between Δ*K* (**a**) and lnP(D) (**b**) with different *k*-values.

RESULTS

InDel marker profiles

Of the 48 InDel markers investigated, 24 markers resulted in 126 clearly visible alleles of which 51 (40.48%) were found to be polymorphic (Table 2 and Figure 2). The 61 polymorphic alleles showed an average of 2.6 alleles per locus, ranging from 1 to 8 (InDel-001) alleles. The InDel PIC ranged from 0.0364 (InDel-028) to 0.9030 (InDel-001), with an average of 0.5038, whereas *H'* ranged from 0.0922 (InDel-028) to 2.4465 (InDel-001), with an average of 0.9483.

Table 2. Number of alleles, number of polymorphic alleles, rate of polymorphism, polymorphism information content (PIC), and Shannon's information index (*H'*) generated by 24 InDel markers.

InDel marker	No. of alleles	Polymorphic alleles	Rate of polymorphism	PIC	*H'*
InDel-001	11	8	0.73	0.9030	2.4465
InDel-003	6	2	0.33	0.6223	1.0866
InDel-005	3	1	0.33	0.2257	0.3857
InDel-007	4	2	0.50	0.5873	0.9976
InDel-008	6	2	0.33	0.5525	0.9077
InDel-009	6	3	0.33	0.7599	1.5232
InDel-010	5	3	0.60	0.7390	1.4660
InDel-013	5	3	0.60	0.7469	1.4795
InDel-014	3	1	0.33	0.1049	0.2146
InDel-015	4	1	0.25	0.0713	0.1584
InDel-016	2	1	0.50	0.4444	0.6365
InDel-018	4	1	0.25	0.3018	0.4792
InDel-019	8	7	0.88	0.8841	2.2536
InDel-022	10	1	0.10	0.4753	0.6682
InDel-023	4	1	0.25	0.2778	0.4506
InDel-028	5	1	0.20	0.0364	0.0922
InDel-030	6	1	0.17	0.4170	0.6077
InDel-034	3	1	0.33	0.4890	0.6821
InDel-038	5	3	0.60	0.7827	1.6212
InDel-040	3	2	0.67	0.7347	1.3554
InDel-041	8	2	0.25	0.6313	1.0803
InDel-043	7	1	0.14	0.4993	0.6924
InDel-044	3	1	0.33	0.1371	0.2641
InDel-045	5	2	0.40	0.6675	1.2107

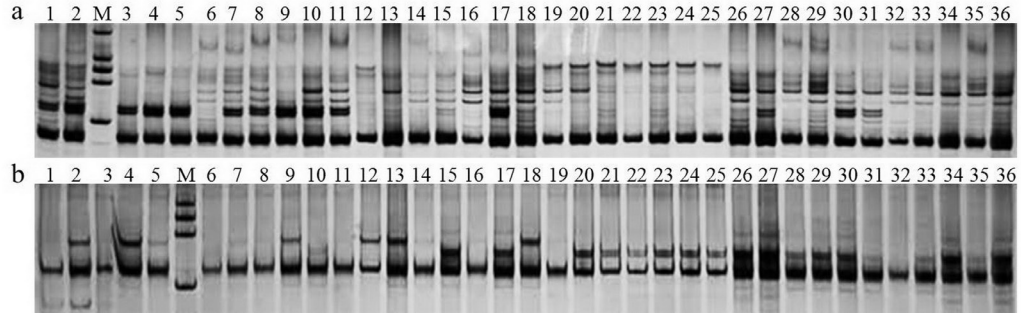

Figure 2. Amplification profiles of products from genomic DNA of cultivated peanut using InDel primers 038-InDel (**a**) and 040-InDel (**b**). *Lane M* = DNA marker; *lanes 1-11 = vulgaris*; *lanes 12-18 = fastigiata*; *lanes 19-25 = hirsuta*; and *lanes 26-36 = hypogaea*.

Genetic diversity of different botanical peanut types based on InDel markers

The number of alleles varied among the four botanical subspecies with the most alleles detected in *vulgaris* variety, followed by *fastigiata* variety, *hypogaea* variety, and *hirsuta* variety (Table 3). The PIC values varied from 0.4536 in *hirsuta* variety, to 0.2197 in *hypogaea* variety, with *fastigiata* variety and *vulgaris* variety showing intermediate values (0.4473 and 0.3935, respectively). This indicated that *hirsuta* variety and *fastigiata* variety have higher genetic diversity than the other types. In addition, *H'* also differed among the different botanical types, *hirsuta* variety had the highest value (0.6460), followed by *fastigiata* variety (0.6395), *vulgaris* variety (0.5825), and *hypogaea* variety with the lowest value (0.3780).

Table 3. Number of accessions, number of alleles, PIC and, *H'* of different botanical *Arachis hypogaea* types based on InDel markers.

Botanical types	Accessions	No. of alleles	PIC	*H'*
hypogaea	19	31	0.2197	0.3780
vulgaris	13	40	0.3935	0.5825
fastigiata	7	33	0.4473	0.6395
hirsuta	15	14	0.4536	0.6460

Population structure analysis

The STRUCTURE analysis identified four genetically similar subpopulations named G1, G2, G3, and G4, respectively (Table 4 and Figure 3). G1 was a mixed group, including six accessions from *vulgaris* variety and six accessions from *fastigiata* variety; G2 contained mostly accessions from *hirsuta* variety with two exceptions, one from *hypogaea* variety and one from *vulgaris* variety; G3 was also a mixture containing six accessions from *vulgaris* variety, one from *fastigiata* variety, and three from *hypogaea* variety; whereas G4 exclusively contained 15 from *hypogaea* variety.

Table 4. Number of botanical varieties of *Arachis hypogaea* in each cluster.

	Hypogaea	*Fastigiata*	*Vulgaris*	*Hirsuta*	Overall
G1	0	6	6	0	12
G2	1	0	1	15	17
G3	3	1	6	0	10
G4	15	0	0	0	15
A1	0	1	5	0	6
A2	0	6	8	0	14
A3	2	0	0	15	17
A4	17	0	0	0	17
B1	0	2	7	0	9
B2	0	5	6	0	11
B3	2	0	0	15	17
B4	17	0	0	0	17

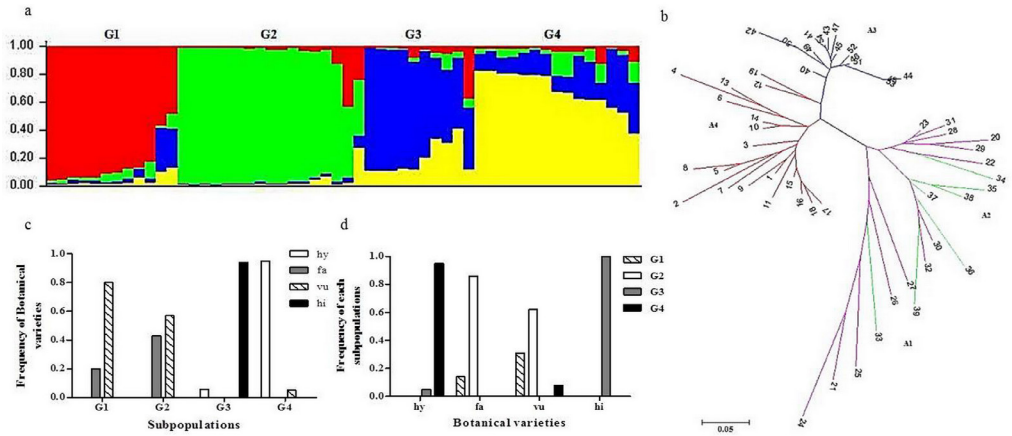

Figure 3. Results from the population structure analysis (**a**). The subgroup membership possibility is shown on the y-axis. G1-G4 represents the four identified subpopulations on the x-axis (G1: red; G2: green; G3: blue; and G4: yellow). Results from the NJ tree analysis (**b**). A1-A4 stands for branches. The NJ tree is color-coded based on the botanical varieties (*vulgaris* variety: purple; *fastigiata* variety: green; *hirsuta* variety: blue; and *hypogaea* variety: red). Frequency of the respective botanical varieties (*hypogaea* variety: hy; *fastigiata* variety: fa; *vulgaris* variety: vu; and *hirsuta* variety: hi) within each subpopulation (**c**) and frequency of each subgroup (G1-G4) within each botanical variety (**d**).

The NJ analysis clustered 54 accessions into four clusters (A1-A4) and this grouping was generally consistent with the results from the STRUCTURE analysis with a few exceptions (Table 4 and Figure 3). A1 contained six accessions from subpopulation G3. Among these, five accessions (21 and 24-27) belonged to *vulgaris* variety and one (33) was *fastigiata* variety. A2 consisted of 14 accessions, including all accessions of G1 and two *vulgaris* variety (22 and 23) from G3 and G2, respectively. A3 contained 17 accessions that corresponded primarily to G2 with one exception (19) from G4. Finally, all accessions in A4 were from *hypogaea* variety, among which 14 accessions were from G4 and three accessions (4, 8, and 13) were from G3.

To further confirm the results of the STRUCTURE and NJ analyses, a PCA was used to cluster the 54 accessions into four groups (B1-B4). Each cluster contained 9, 11, 17, and 17 accessions, respectively (Table 4 and Figure 4). The PCA results were more consistent with

the NJ tree than the STRUCTURE analysis. Clusters B3 and B4 were identical to A3 and A4, respectively. Both B1 and B2 contained only *fastigiata* variety and *vulgaris* variety, which was the same as found in A1 and A2, except that the number of accessions from the respective varieties varied slightly (Table 4).

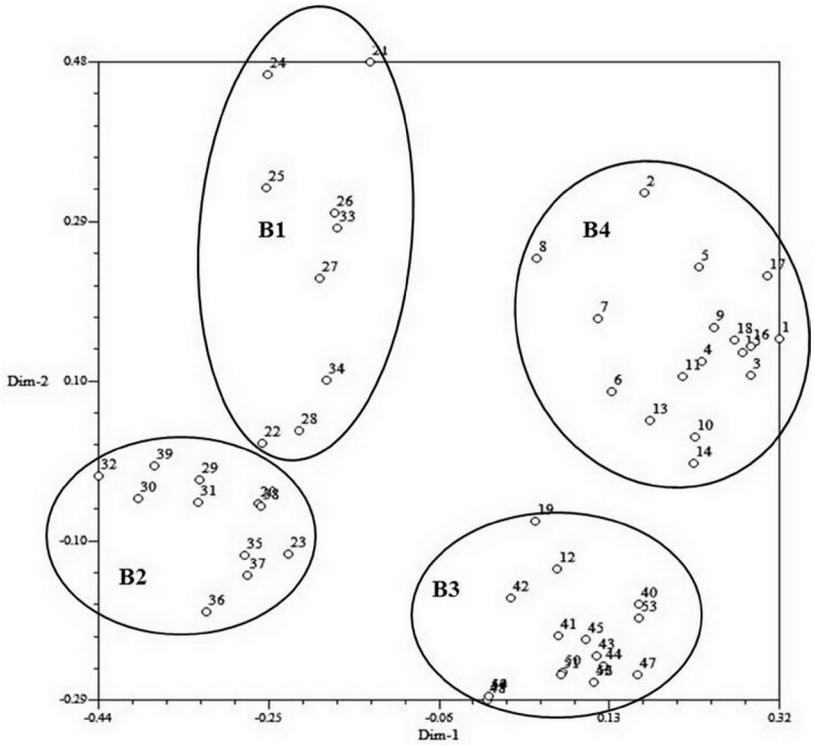

Figure 4. Two-dimensional representation of a principal component analysis of 54 accessions, based on amplified DNA fragments using InDel markers.

Marker-trait correlation

The determination coefficient (R^2) of the 11 InDel markers for the five agronomic traits ranged from 5.3 to 33.2% based on the multi-linear stepwise regression analysis (Table 5). Three markers (InDel-001, InDel-038, and InDel-040) were found to relate to pod length, explaining 6.2-21.6% of the phenotypic variation. Six markers (InDel-001, InDel-008, InDel-009, InDel-013, InDel-019, and InDel-041) were related to pod width explaining 9.9-33.2% of the phenotypic variation. Two markers (InDel-019 and InDel-040) were found to be associated with pod thickness, explaining 5.3-9.8% of the phenotypic variation. InDel-044 was identified to be associated with thickness of pod shell, explaining 8.8% of the phenotypic variation. Among these markers, seven markers were identified to relate with more than two traits, for example InDel-040 was found to relate to pod width, pod thickness and weight of 100 pods.

Table 5. Stepwise multiple regression analyses between InDel markers and agronomic traits in *Arachis hypogaea*.

Traits	InDel markers	R^2 (%)	Beta coefficient
Pod length	001/038/040	6.2-21.6	0.0340
Pod width	001/008/009/013/019/041	9.9-33.2	-0.4330
Pod thickness	019/040	5.3-9.8	0.0430
Weight of 100 pods	009/013/016/018/040/044	12.6-32.9	-0.0820
Thickness of pod shell	044	8.8	-0.3180

DISCUSSION

Level of polymorphism is one of the criteria used to evaluate the potential use of markers. Various markers have been used in peanut diversity studies that have shown varying degrees of polymorphism, including AFLP and DAF (6.86 and 21.14%, respectively, He and Prakash, 2001); RAPD (42.73%, Raina et al., 2001); and InDel (33.30%, Liu et al., 2015) (Table 6). In addition, Xiong et al. (2011) found variation in SCoT (38.22%), ISJAP (28.10%), DAMD (36.00%), and SRAP-IAAP (43.15%) markers (Table 6). In these reports, the number of accessions used was smaller than the number (54) of accessions evaluated in the present study. The number of accessions used in the evaluation of marker polymorphism affects the accuracy of the estimated polymorphism. The higher the number of accessions used, the higher the accuracy is, the same is true for the number of botanical types included. Therefore, we suggest that the InDel marker polymorphism reported in this study is relatively reliable, due to the number of tested accessions and presentation of four botanical types of cultivated peanut of North China. The InDel marker polymorphism was 50%, which suggests that these markers are suitable for peanut diversity studies.

Table 6. Comparison of polymorphisms of various molecular markers developed in previous reports.

Marker	NPT	NPDP	PPP (%)	NA	NBV	TNB	NPB	PPB (%)	References
SPAR	10	0	0	27	4	-	0	0	Halward et al., 1992
DAF	559	17	3.04	6	3	298	63	21.14	He and Prakash, 2001
AFLP	64	28	43.75	6	3	1618	111	6.86	He and Prakash, 2001
RAPD	57	21	36.84	13	5	220	94	42.73	Raina et al., 2001
ISSR	100	29	29.00	13	5	124	67	54.03	Raina et al., 2001
SCoT	36	18	50.00	20	4	157	60	38.22	Xiong et al., 2011
ISJAP	26	14	53.85	16	4	121	34	28.10	Xiong et al., 2011
DAMD	10	4	40.00	16	4	25	9	36.00	Xiong et al., 2011
URP	12	8	66.67	16	4	50	25	50.00	Xiong et al., 2011
SRAP-IAAP	40	25	62.50	16	4	146	63	43.15	Xiong et al., 2011
InDel	48	24	50.00	54	4	126	51	40.48	Present study

NPT: number of primers tested; NPDP: number of primers detecting polymorphism; PPP: percentage of polymorphic primers; NA: number of accessions; NBV: number of botanical varieties; TNB: total number of bands; NPB: number of polymorphic bands; PPB: percentage of polymorphic bands.

The study of genetic diversity of cultivated peanut will enhance the utilization of germplasm resources and genetic improvement. There are many reports on genetic diversity of cultivated peanut that have used various DNA molecular markers. For example, Kottapalli et al. (2007) employed 67 polymorphic SSR markers to assess the genetic diversity of 72 accessions of the US mini core. They found that the PIC values of the SSR markers ranged from 0.063 to 0.918 and that the *fastigiata* variety had a higher PIC value than *hypogaea* variety (Kottapalli et al., 2007). Likewise, Jiang et al. (2010) reported similarity coefficients

among different botanical varieties that ranged from 0.49 to 0.99. Based on 466 accessions and 26 SSR primer pairs, the *fastigiata* variety and *hypogaea* variety were found to be more diversified than other botanical types (Jiang et al., 2010). In a previous study, the PIC of *hirsuta* variety was found to be the highest among six botanical types (0.6597) (Liu et al., 2015). In the present study, *hirsuta* variety and *fastigiata* variety were identified to have higher PIC values than the other botanical types (0.4536 and 0.4473, respectively). In addition, the percentage of polymorphic primers was higher (50%) compared with those found previously (Liu et al., 2015).

STRUCTURE analyses can be used to estimate the number of subpopulations, the degree of admixture among subpopulations, and the genetic relatedness among accessions. On the other hand, phylogenetic analyses can be used to display relationships between accessions graphically. In our study, the STRUCTURE analysis assigned the accessions into four subpopulations (G1-G4). Likewise, the NJ tree analysis resulted in four clusters (A1-A4). As expected, in general, the subpopulations identified by STRUCTURE corresponded to the genetic clustering found based on the NJ tree (G1/A2, G2/A3, G3/A1, and G4/A4). Most of the accessions divided into their respective subpopulations, with only a few exceptions. For example, Hengshuiyiwohou (23) clustered with G2, but belonged to A2 and Hejianpafang (19) clustered with G4, but belonged to A3. All the accessions were classified into four branches (A1-A4) by the NJ tree (Figure 3), but they may also be divided into two clusters: A1 with A2 and A3 with A4. This indicates that *vulgaris* variety and *fastigiata* variety are more closely related than *hirsuta* variety and *hypogaea* variety. This result is consistent with the known classification in which *vulgaris* variety and *fastigiata* variety belong to the subspecies *fastigiata*, whereas *hirsuta* variety and *hypogaea* variety belong to the subspecies *hypogaea*.

Due to the limited genomic information, presently, only two traits have been used in marker-assisted selection in peanut breeding. These are nematode resistance and high oleic acid content (López et al., 2001; Nagy et al., 2010). A tremendous (at least 3-fold) gain in the speed of selection has been demonstrated to be achieved with marker-assisted selection (Chu et al., 2011). The InDel markers used in the present study were developed based on the sequences of functional genes and the markers were found to be related to several agronomic traits. This information may be used in the future to provide a clue toward identifying the genes controlling the agronomic traits, using in-depth analysis through next-generation sequencing and real-time quantitative PCR detecting systems.

Conflicts of interest

These authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (#31471523), a grant of “948” Project (#2013-Z65), the Science and Technology R&D Program of Hebei province (#11220109D and #14226313D-1), the Specialized Research Fund for the Doctoral Program of Higher Education of China (#2012302110002), and the Science and Technology Research Projects of Hebei province universities (#ZD2015056).

REFERENCES

Alves DM, Pereira RW, Leal-Bertioli SC, Moretzsohn MC, et al. (2008). Development and use of single nucleotide polymorphism markers for candidate resistance genes in wild peanuts (*Arachis* spp). *Genet. Mol. Res.* 7: 631-642. http://dx.doi.org/10.4238/vol7-3gmr453

Barkley NA, Dean RE, Pittman RN, Wang ML, et al. (2007). Genetic diversity of cultivated and wild-type peanuts evaluated with M13-tailed SSR markers and sequencing. *Genet. Res.* 89: 93-106. http://dx.doi.org/10.1017/S0016672307008695

Burow MD, Simpson CE, Starr JL and Paterson AH (2001). Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.). broadening the gene pool of a monophyletic polyploid species. *Genetics* 159: 823-837.

Chu Y, Wu CL, Holbrook CC, Tillman BL, et al. (2011). Marker-assisted selection to pyramid nematode resistance and the high oleic trait in peanut. *Plant Genome* 4: 110-117. http://dx.doi.org/10.3835/plantgenome2011.01.0001

Dang PM and Chen CY (2013). Modified method for combined DNA and RNA isolation from peanut and other oil seeds. *Mol. Biol. Rep.* 40: 1563-1568. http://dx.doi.org/10.1007/s11033-012-2204-9

Fávero AP, Pádua JG, Costa TS, Gimenes MA, et al. (2015). New hybrids from peanut (*Arachis hypogaea* L.) and synthetic amphidiploid crosses show promise in increasing pest and disease tolerance. *Genet. Mol. Res.* 14: 16694-16703. http://dx.doi.org/10.4238/2015.December.11.17

Ferguson ME, Burow MD, Schulze SR, Bramel PJ, et al. (2004). Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *Theor. Appl. Genet.* 108: 1064-1070. http://dx.doi.org/10.1007/s00122-003-1535-2

Flint-Garcia SA, Thuillet AC, Yu J, Pressoir G, et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44: 1054-1064. http://dx.doi.org/10.1111/j.1365-313X.2005.02591.x

Freitas FO, Moretzsohn MC and Valls JF (2007). Genetic variability of Brazilian Indian landraces of *Arachis hypogaea* L. *Genet. Mol. Res.* 6: 675-684.

Fu Y, Wen TJ, Ronin YI, Chen HD, et al. (2006). Genetic dissection of intermated recombinant inbred lines using a new genetic map of maize. *Genetics* 174: 1671-1683. http://dx.doi.org/10.1534/genetics.106.060376

Garcia GM, Stalker HT and Kochert G (1995). Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. *Genome* 38: 166-176. http://dx.doi.org/10.1139/g95-021

Halward T, Stalker T, LaRue E and Kochert G (1992). Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Mol. Biol.* 18: 315-325. http://dx.doi.org/10.1007/BF00034958

Hayashi K, Yoshida H and Ashikawa I (2006). Development of PCR-based allele-specific and InDel marker sets for nine rice blast resistance genes. *Theor. Appl. Genet.* 113: 251-260. http://dx.doi.org/10.1007/s00122-006-0290-6

He G, Meng R, Newman M, Gao G, et al. (2003). Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). *BMC Plant Biol.* 3: 3. http://dx.doi.org/10.1186/1471-2229-3-3

He GH and Prakash CS (2001). Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. *Genet. Resour. Crop Evol.* 48: 347-352. http://dx.doi.org/10.1023/A:1012019600318

Herselman L (2003). Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. *Euphytica* 133: 319-327. http://dx.doi.org/10.1023/A:1025769212187

Hopkins MS, Casa AM, Wang T, Mitchell SE, et al. (1999). Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Sci.* 39: 1243-1247. http://dx.doi.org/10.2135/cropsci1999.0011183X003900040047x

<bok>Jiang HF, Duan NX and Ren XP (2006). Descriptors and data standard for peanut (*Arachis* spp.). China Agriculture Press, Beijing, 18-74.</bok>

Jiang HF, Ren XP, Zhang XJ, Huang JQ, et al. (2010). Comparison of genetic diversity between peanut mini core collections from China and ICRISAT by SSR markers. *Acta Agron. Sin.* 36: 1084-1091. http://dx.doi.org/10.1016/S1875-2780(09)60059-6

Kochert G, Halward T, Branch WD and Simpson CE (1991). RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *Theor. Appl. Genet.* 81: 565-570. http://dx.doi.org/10.1007/BF00226719

Kottapalli KR, Burow MD, Burow G, Burke J, et al. (2007). Molecular characterization of the U.S. peanut mini core collection using microsatellite markers. *Crop Sci.* 47: 1718-1727. http://dx.doi.org/10.2135/cropsci2006.06.0407

Krapovickas A and Gregory WC (1994). Taxonomia del género *Arachis* (Leguminosae). *Bonplandia* 8: 1-186.

Liu L, Dang PM and Chen CY (2015). Development and utilization of InDel markers to identify peanut (*Arachis hypogaea*) disease resistance. *Front. Plant Sci.* 6: 988. http://dx.doi.org/10.3389/fpls.2015.00988

López Y, Smith OD, Senseman SA and Rooney WL (2001). Genetic factors influencing high oleic acid content in Spanish

market-type peanut cultivars. *Crop Sci.* 41: 51-56. http://dx.doi.org/10.2135/cropsci2001.41151x

Nagy ED, Chu Y, Guo YF, Khanal S, et al. (2010). Recombination is suppressed in an alien introgression in peanut harboring *Rma*, a dominant root-knot nematode resistance gene. *Mol. Breed.* 26: 357-370. http://dx.doi.org/10.1007/s11032-010-9430-4

Patterson N, Price AL and Reich D (2006). Population structure and eigenanalysis. *PLoS Genet.* 2: e190. http://dx.doi.org/10.1371/journal.pgen.0020190

Pompanon F, Bonin A, Bellemain E and Taberlet P (2005). Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* 6: 847-859. http://dx.doi.org/10.1038/nrg1707

Raina SN, Rani V, Kojima T, Ogihara Y, et al. (2001). RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome* 44: 763-772. http://dx.doi.org/10.1139/gen-44-5-763

Ren XP, Huang JQ, Liao BS, Zhang XJ, et al. (2010). Genomic affinities of *Arachis* genus and interspecific hybrids were revealed by SRAP markers. *Genet. Resour. Crop Evol.* 57: 903-913. http://dx.doi.org/10.1007/s10722-010-9532-1

<bok>Rohlf FJ (2000). NTSYS-pc, numerical taxonomy system for the PC ExeterSoftware, version 2.1. Applied Biostatistics Inc., Setauket.</bok>

Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.

Shedlock AM and Okada N (2000). SINE insertions: powerful tools for molecular systematics. *BioEssays* 22: 148-160. http://dx.doi.org/10.1002/(SICI)1521-1878(200002)22:2<148::AID-BIES6>3.0.CO;2-Z

Song X, Wei H, Cheng W, Yang S, et al. (2015). Development of INDEL markers for genetic mapping based on whole-genome re-sequencing in soybean. *G3 (Bethesda)* 5: 2793-2799. http://dx.doi.org/10.1534/g3.115.022780

Syvänen AC (2001). Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* 2: 930-942. http://dx.doi.org/10.1038/35103535

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599. http://dx.doi.org/10.1093/molbev/msm092

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, et al. (2001). *Dwarf8* polymorphisms associate with variation in flowering time. *Nat. Genet.* 28: 286-289. http://dx.doi.org/10.1038/90135

Väli U, Brandström M, Johansson M and Ellegren H (2008). Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC Genet.* 9: 8. http://dx.doi.org/10.1186/1471-2156-9-8

Wang H, Khera P, Huang B, Yuan M, et al. (2015). Analysis of genetic diversity and population structure of peanut cultivars and breeding lines from China, India and the US using simple sequence repeat markers. *J. Integr. Plant Biol.* 10.1111/jipb.12380.

Wu K, Yang M, Liu H, Tao Y, et al. (2014). Genetic analysis and molecular characterization of Chinese sesame (*Sesamum indicum* L.) cultivars using insertion-deletion (InDel) and simple sequence repeat (SSR) markers. *BMC Genet.* 15: 35. http://dx.doi.org/10.1186/1471-2156-15-35

Xiong F, Zhong R, Han Z, Jiang J, et al. (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. http://dx.doi.org/10.1007/s11033-010-0459-6

Zhang J, Wu YT, Guo WZ and Zhang TZ (2000). Fast screening of microsatellite markers in cotton with PAGE/silver staining. *Acta Gossypii Sin.* 12: 267-269.