Plant Breeding **131,** 139—147 (2012)

doi:10.1111/j.1439-0523.2011.01911.x

Plant Breeding

Highly informative genic and genomic SSR markers to facilitate molecular breeding in cultivated groundnut (*Arachis hypogaea*)

Manish K. Pandey[1], Bhimana Gautami[1], Thandoniappan Jayakumar[1], Manda Sriswathi[1], Hari D. Upadhyaya[1], Makanahally V. Chennabyre Gowda[2], Thankappan Radhakrishnan[3], David J. Bertioli[4], Steven J. Knapp[5], Duglas R. Cook[6] and Rajeev K. Varshney[1,7,8]

[1]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, India; [2]University of Agricultural Sciences (UAS), Dharwad 580005, India; [3]Directorate of Groundnut Research (DGR), Junagadh 362001, India; [4]University of Brasília, Brasília, DF 70910900, Brazil; [5]The University of Georgia, Athens, GA 30602, USA; [6]University of California, Davis, CA 95616, USA; [7]CGIAR Generation Challenge Programme (GCP), c/o CIMMYT, 06600 Mexico DF, Mexico; [8]Corresponding author, E-mail: r.k.varshney@cgiar.org

With 3 figures and 7 tables

Received May 24, 2011/Accepted August 23, 2011
Communicated by R. Tuberosa

Abstract

With an objective of identification of highly informative set of SSR markers in cultivated groundnut (*Arachis hypogaea* L.), a total of 4485 markers were used for screening using a set of 20 parental genotypes of 15 mapping populations. Although 3582 (79.9%) markers provided scorable amplification, only 1351 (37.3%) markers showed polymorphism. Polymorphism information content (PIC) value ranged from 0.10 (GM742) to 0.89 (S009) with an average of 0.31. Similarly, number of alleles ranged from 2 to 14 with an average of 3.2 alleles. In general, the SSR markers based on dinucleotide repeats displayed higher PIC value and number of alleles. Based on these polymorphism features, 199 markers with >0.50 PIC values have been identified. Polymorphism features of these markers along with the primer sequences, for the first time, for a total of 946 SSR markers have been provided. It is anticipated that the identified set of highly informative markers, instead of starting from the random set of SSR markers, should be very useful to initiate molecular genetics and breeding studies in cultivated groundnut.

Key words: groundnut — molecular breeding — SSR markers — PIC value — genetic diversity

Groundnut (*Arachis hypogaea* L.), the third most important oilseed crop in the world, is grown extensively throughout the semi-arid tropics (SAT) of Asia, Africa and Latin America, with its global production of 35.52 million tons from 23.5 million ha area (FAO 2009). It is a self-pollinating crop with ten basic chromosomes and allotetraploid genome ($2n = 4x = 40$, AABB) (Stebbins 1957, Stalker and Dalmacio 1986). The origin of cultivated groundnut was probably through a few or even a single hybridization event between two diploid wild species, *A. duranensis* (AA genome) and *A. ipaënsis* (BB genome), followed by a spontaneous chromosome duplication (Halward et al. 1991). The resulting tetraploid plant (AABB genome) was then reproductively isolated from its wild diploid relatives (AA and BB genome). This extreme bottleneck, coupled with reproductive isolation, leads to a limited genetic diversity within the groundnut primary gene pool.

For crop improvement, genetic enhancement of cultivated groundnut to increase the yield and resistance/tolerance to biotic and abiotic stresses has been the most important goal. Although efforts made through conventional breeding have some measure of success, expected progress could not be achieved in handling complex traits such as tolerance to drought, either owing to lack of reliable, precise and cost-effective high-throughput phenotyping or owing to fertility barriers that hamper the harnessing of genetic variation present in secondary and tertiary gene pool (and even sometimes from primary gene pool also). Recent advances in the area of crop genomics have offered molecular tools to assist breeding (Varshney et al. 2005a). Introgression of desired chromosomal segment in the progeny through precise monitoring using trait-linked marker, the process called marker-assisted selection (MAS), has been successfully applied in several cereal and some legume crops, resulting in the development of improved varieties/germplasm (Varshney et al. 2006). Availability of molecular markers and genetic linkage maps is, however, the prerequisite for undertaking molecular breeding activities particularly identifying and localizing important genes, controlling qualitatively and quantitatively inherited traits (Varshney et al. 2006). Such tools would then simply speed up the process of introgression of agronomically desired traits such as yield, quality and biotic and abiotic stress resistance to preferred varieties, especially for complex traits such as drought.

Molecular marker analysis on groundnut germplasm using a variety of molecular markers such as microsatellites or simple sequence repeats (SSRs), randomly amplified polymorphic DNAs (RAPDs) and amplified fragment length polymorphisms (AFLPs) in general has shown very low variation in cultivated gene pool because of the evolutionary genetic bottleneck in the form of polyploidy and self-pollination (Kochert et al. 1996, Subramanian et al. 2000, Herselman 2003). On the other hand, wild diploid *Arachis* species showed relatively higher variation (Hilu and Stalker 1995, Moretzsohn et al. 2004, Bravo et al. 2006), providing a rich source of genetic variation for genetic and genomic studies (Stalker and Simpson 1995, Rao et al. 2003, Dwivedi et al. 2007). Among different marker systems analysed in the groundnut, like other plant species, SSR markers have been found more informative and useful for genetic analysis and breeding applications (Gupta and Varshney 2000).

In the case of groundnut, several hundred SSR markers have been developed and characterized during last 5 years all over the world (Hopkins et al. 1999, He et al. 2003, Palmieri et al. 2002, 2005, Ferguson et al. 2004, Moretzsohn et al. 2004, 2005, Nelson et al. 2006, Mace et al. 2007, Proite et al. 2007, Gimenes et al. 2007, Wang et al. 2007, Cuc et al. 2008, Gautami et al. 2009, unpublished markers from University of California-Davis, USA and University of Georgia, USA). However, the development of even low- to moderate-density genetic maps using populations derived from cultivated germplasm has been hindered by the requirement of screening very large numbers of SSR markers to find a sufficient number of polymorphic markers (Varshney et al. 2009a, Khedikar et al. 2010, Ravi et al. 2011, Sarvamangala et al. 2011). The availability of the polymorphism information content (PIC) values and number of alleles detected by a large set of SSR markers would help groundnut community to select the most informative markers to screen the germplasm, thus economizing time and cost in the development of the genetic and QTL maps. Here, we have screened a large number (4485) of SSR markers available in public domain as well as accessed through collaborators across the world on 5–16 genotypes from a set of 20 genotypes that represent the parents of 15 mapping populations segregating for different traits. An analysis of the marker polymorphism data allowed the identification of a highly informative SSR marker set.

Materials and Methods

Plant material: About 5–16 genotypes from a set of 20 genotypes representing parents of 15 mapping populations segregating for resistance/tolerance to biotic and abiotic stresses were used to screen with SSR markers (Tables 1 and 2). The genotype set includes drought-tolerant genotypes (ICGS 44, ICGS 76, CSMG 84-1 and ICGV 86031), drought-sensitive genotypes (TAG 24 and Chico), resistant genotypes for different foliar diseases (GPBD 4, ICG 11337, ICGV 86590, R 9227, ICG (FDRS) 10 and TxAG-6) and genotypes susceptible to foliar diseases (TAG 24, JL 24, GPBD 5, TG 19, TG 26 and TMV-2). In addition, two AA-genome (diploid) species genotypes (K7988 and V10309) and a synthetic allotetraploid genotype (TxAG-6) developed from the cross *A. batizocoi* and (*A. cardenasii* × *A. diogoi*) were also included in the set.

DNA isolation: Total genomic DNA was isolated from unopened leaves harvested from 10 to 15-day-old seedlings according to modified CTAB-based method as mentioned in the study of Cuc et al. (2008). DNA quality and quantity were checked on 0.8% agarose gels, and DNA concentration was normalized to approximately 5 ng/μl for Polymerase chain reaction (PCR).

Polymerase chain reaction with SSR markers: A varying number of SSR markers from a total of 4485 SSR markers, as given in Table 3, were used for screening the above-mentioned genotypes.

Polymerase chain reactions for SSR markers were performed in 5 μl volume following a touchdown PCR profile in an ABI thermal cycler (Applied Biosystems, Foster City, CA, USA). PCR comprised of approximately 5 ng of genomic DNA, 2 pmol of each primer, 2 mM of each dNTP, 2 mM $MgCl_2$, 1× amplification buffer and 0.1 U of *Taq* DNA polymerase (Qiagen, Hilden, Germany). The touchdown PCR amplification profile had initial denaturation step for 3 min at 94°C followed by first five cycles of 94°C for 20 s, 65°C for 20 s and 72°C for 30 s, with 1°C decrease in temperature each cycle, followed by 35 cycles of 94°C for 20 s with constant annealing temperature (59°C) for 20 s and 72°C for 30 s, followed by a final extension for 20 min at 72°C. The amplified products were tested on 1.2% agarose gels to check the amplification.

SSR fragment analysis: After confirmation for amplification, PCR products were diluted to varied folds (60–100) and used for multiplexing based on different fluorescent labels and amplicon length. Markers that had different labels and allele size ranges were considered together along with markers with the same label separated by more than 50 bp. Formamide (1 μl) was added to each well containing PCR product (1 μl) along with GeneScan 500 standard (Applied Biosystems) internal lane standard labelled with either ROX or LIZ. GeneScan Filter Set D and the ROX 500/LIZ 500 internal lane were used for analysis of amplicons labelled with different fluorescent dyes such as FAM, VIC, NED, PET, HEX and TAMARA. Allele sizing and scoring based on capillary electrophoresis (ABI 3700 Genetic Analyzer; Applied Biosystems) data were carried out using GeneScan 3.1 software (Applied Biosystems). PCR products for a few markers were also analysed on 6% non-denaturing polyacrylamide gels (PAGE) (29 : 1 acrylamide/bisacrylamide) and visualized by silver staining as given in the study of Varshney et al. (2009a,b).

Data analysis: Major allele frequency, gene diversity and PIC values for all loci were computed using allelic data with PowerMarker version 3.25 (Liu and Muse 2005). For assessing the genetic relationships

Table 1: Pedigree of parental genotypes used in the study

S. No.	Genotypes	Pedigree	Botanical type	Market type	Origin
1	ICGS 44	Robut 33-1-1-5-B1-B1-B2	*vulgaris*	Spanish	India
2	ICGS 76	TMV 10 × CHICO	*hypogaea*	Virginia	India
3	ICGV 86031	F 334 A-B-14 × NC Ac 2214	*vulgaris*	Spanish	India
4	ICGV 86590	X 14-4-B-19-B × PI 259747	*hypogaea*	Virginia	India
5	ICGV 11337	Cs 46	–	–	India
6	CSMG 84-1	Selection from MA 10	*hypogaea*	Virginia	India
7	TAG 24	TG S2 × TGE 1	*hypogaea*	Virginia	India
8	TG 19	TG 17 × TG 1	*hypogaea*	Virginia	India
9	TG 49	TG 28A × TG 26	*vulgaris*	Spanish	India
10	TG 26	BARCG 1 × TG 23	*hypogaea*	Virginia	India
11	GPBD 4	KRG 1 × CS 16 (ICGV 86855)	*vulgaris*	Spanish	India
12	GPBD 5	TG 49 × GPBD 4	*vulgaris*	Spanish	India
13	TMV 2	Mass selection from Gudhiatham bunch	*vulgaris*	Spanish	India
14	TxAG-6	[*A. batizocoi* × (*A. cardenasii* × *A. Diogoi*)]$^{4\times}$	–	–	USA
15	R 9227	ICGS 7 × (NC Ac 2214 × ICGV 86031)	*vulgaris*	Spanish	India
16	JL 24	Selection from EC 94943	*vulgaris*	Spanish	India
17	Chico	Short-duration genotype	*vulgaris*	Spanish	USA
18	ICG (FDRS) 10	Ah 65 × NCAc 17090	*vulgaris*	Spanish	India
19	K7988	*A. duranensis* (AA genome)	*duranensis*	–	Brazil
20	V10309	*A. stenosperma* (AA genome)	*stenosperma*	–	Brazil

Table 2: Details of the mapping populations based on genotypes used in the analysis

S. No.	Mapping populations	Source	Segregating traits
Abiotic stress tolerance			
1	ICGS 44 × ICGS 76	ICRISAT, India	Drought tolerance–related traits viz., transpiration, transpiration efficiency, specific leaf area and SPAD chlorophyll meter reading (SCMR)
2	ICGS 76 × CSMG 84-1	ICRISAT, India	
3	TAG 24 × ICGV 86031	ICRISAT, India	
4	Chico × CSMG 84-1	ICRISAT, India	
5	K7988 × V10309	EMBRAPA, Brazil	Reference mapping population for AA genome
Biotic stress resistance			
6	TMV 2 × TxAG-6	ICRISAT, India	Late leaf spot (LLS) resistance, root-knot nematode
7	ICG 11337 × JL 24	ICRISAT, India	Late leaf spot resistance
8	JL 24 × ICG(FDRS) 10	ICRISAT, India	Late leaf spot resistance
9	TAG 24 × GPBD 4	UAS-Dharwad, India	Rust and late leaf spot resistance
10	TG 26 × GPBD 4	UAS-Dharwad, India	
11	GPBD 5 × GPBD 4	UAS-Dharwad, India	
12	TG 19 × GPBD 4	UAS-Dharwad, India	*Aspergillus* crown rot, rust and late leaf spot resistance
13	TG 49 × GPBD 4	UAS-Dharwad, India	
14	TAG 24 × R 9227	UAS-Dharwad, India	*Sclerotium* rot resistance
15	JL 24 × ICGV 86590	DGR, Junagadh, India	Rust and *Sclerotium* rot resistance

Table 3: Source of markers used for polymorphism survey

S. No.	Series	No. of markers	Source of markers
1	Ah, Lec	26	Hopkins et al. (1999)
2	pPGPseq, pPGSseq	226	Ferguson et al. (2004)
3	Ap	18	Palmieri et al. (2002, 2005)
4	PM	59	He et al. (2003)
5	AC, Ah, gi, RN, TC, Seq	338	Moretzsohn et al. (2004, 2005)
6	LG, Lup	103	Nelson et al. (2006)
7	Lup, Dal, Stylo, Ades, Amor, Chaet	51	Mace et al. (2007)
8	RN, RM	53	Proite et al. (2007)
9	Ah	14	Gimenes et al. (2007)
10	S	123	Wang et al. (2007)
11	IPAHM	104	Cuc et al. (2008)
12	GA	97	Nagy et al. (2010)
13	ICGM	23	Gautami et al. (2009)
14	GM	2098	Steven J. Knapp, University of Georgia, USA (unpublished data)
15	GNB	1152	Douglas R. Cook, University of California, Davis, USA (unpublished data)
Total number of markers		4485	

between the genotypes, allelic data were converted into binary form i.e. 0 and 1. Similarity matrix was computed using Jaccard's coefficient utilizing the unweighted pair group method with arithmetic averages method, and further, a neighbour-joining (NJ) dendrogram was constructed using the software NTSYSpc version 2.02 (Rohlf 2000).

Results

Marker analysis

About 5–16 genotypes from a set of 20 genotypes representing the parents of 15 mapping populations were screened with 1152–4365 SSR markers to identify polymorphic SSR markers in the respective crosses (Tables 1–3). As a result, 546 (TAG 24 × R 9227) to 2737 (TAG 24 × GPBD 4) markers showed amplification in a given cross (Table 4). In summary, a total of 1351 (37.7%) markers ranging from 2.4% (GPBD 5 × GPBD 4) to 60.5% (K7988 × V10309) showed polymorphism between the parental genotypes of the 15 populations. Of 1351 polymorphic markers, high-quality scoring data for at least 11 genotypes were available for only 1020 SSR markers. Primer sequence information along with the polymorphism features for all these 1020 SSR markers that include 946 new SSR markers, reported for the first time, has been provided in Data S1.

Polymorphism features

All the above-mentioned 1020 polymorphic markers detected a total of 3214 alleles with an average of 3.2 alleles per marker. The number of alleles per marker ranged from 2 for 463 markers to 14 for 2 markers, namely GNB18 and GNB515 per marker. Similarly, the PIC values for polymorphic markers ranged from 0.10 (GM742) to 0.89 (S009) with an average of 0.31 per marker. In total, only 15.67% markers had PIC value more than 0.50.

In terms of marker polymorphisms per mapping population, a higher level of polymorphism was detected in AA-genome mapping population, namely K7988 × V10309 (60.5%) followed by TMV 2 × TxAG-6 (42%) (Table 4). The remaining populations showed comparatively very low polymorphism ranging from 2.4% (GPBD 5 × GPBD 4) to 11.1% (Chico × CSMG 84-1) and average being 7.09% per population.

Polymorphism trends

Because of unavailability of repeat motif information for 16 SSR markers, the relationship between types of SSRs with number of alleles and PIC value was analysed for 1004 polymorphic SSR markers. Based on the repeat motifs, all

Table 4: Comparative marker polymorphism in different parental combinations

Mapping population	No. of markers tested	No. of markers amplified (%)	No. of polymorphic markers	% Polymorphism
Tetraploid populations				
ICGS 44 × ICGS 76	4245	2637 (62.1)	90	3.4
ICGS 76 × CSMG 84-1	4245	2582 (60.8)	129	4.9
TAG 24 × ICGV 86031	4365	2620 (60.0)	211	8.1
TAG 24 × GPBD 4	4100	2737 (66.7)	163	5.9
TMV 2 × T*x*AG-6[1]	3222	1571 (48.4)	660	42.0
ICG 11337 × JL 24	3099	1227 (39.6)	82	6.7
TG 26 × GPBD 4	4100	2202 (53.7)	142	6.4
TG 19 × GPBD 4	1152	715 (62.1)	26	3.6
TG 49 × GPBD 4	1152	685 (59.5)	27	3.9
GPBD 5 × GPBD 4	1152	673 (58.4)	16	2.4
TAG 24 × R 9227	1152	546 (47.4)	16	2.9
JL 24 × ICGV 86590	1152	748 (64.9)	35	4.7
JL 24 × ICG (FDRS) 10	2070	1305 (63.0)	112	8.6
Chico × CSMG 84-1	2070	1330 (64.2)	148	11.1
Diploid population				
K7988 × V10309	1947	660 (33.9)	399	60.5
Total number[2]	4485	3582 (79.9)	1351	37.3

[1]A synthetic genotype developed by crossing *A. batizocoi* and (*A. cardenasii* × *A. diogoi*).
[2]Total number of markers tested, amplified or polymorphic does not represent the sum of such markers across the 15 crosses. Number of the markers tested, amplified or polymorphic in different parental combinations is subset of the number of markers mentioned in the row.

markers were classified into three classes, namely class I (< 10 repeat units), class II (> 10 repeat units) and compound SSRs (more than one type of repeats are present). Using these criteria, the class I contained 323 (32.2%) markers, class II had 609 (60.6%) markers and the compound SSR class included 72 (7.2%) markers. In class I type markers, dinucleotide (124) and trinucleotide (164) repeat motifs were abundant followed by tetranucleotide (19), pentanucleotide (5) and hexanucleotide (8) repeats (Table 5). The average PIC values for these repeat motifs varied from 0.28 (hexanucleotide SSRs) to 0.43 (dinucleotide SSRs). As compared to five types of repeat motifs in the case of class I SSR markers, the class II markers possessed only three repeat types i.e. dinucleotide (142), trinucleotide (443) and tetranucleotide (29) repeats with average PIC values as 0.31, 0.26 and 0.24, respectively. It is noteworthy that dinucleotide repeats from both the classes (class I and class II), in general, produced more number of alleles (up to 14), while hexanucleotide repeats could produce only 2–3 alleles per markers (Table 5, Fig. 1). In summary, a negative correlation was observed between repeat motifs and average number of alleles produced by markers for both classes. Similarly, negative correlation was also observed between repeat motifs and PIC value. Markers with larger repeat motifs tended to have lower PIC values. In contrast, as expected, there was a positive correlation between average number of alleles and PIC values. Considering the PIC values, there were 199 SSR markers that showed high (> 0.50) PIC values. This set is recommended as an informative set of SSR markers that can be used as a starting point for undertaking genetic analysis and breeding applications in groundnut. The markers of this set detect 3–14 alleles with an average of 4.87 per markers. However, after excluding the AA-genome (K7988, V10309) and synthetic amphidiploid genotype (TxAG-6), the average PIC value and number of alleles detected by these markers are reduced to 0.56 and 3.84, respectively (Table 6).

Comparison between genomic and genic SSRs

All 1020 polymorphic markers were classified into genomic and genic SSRs based on their origin from genomic vs. transcribed portion i.e. ESTs (expressed sequence tags). As a result, 260 markers were found to belong to genomic SSR and 760 to genic SSR classes. In terms of comparison of markers from these two classes, the PIC values of all the polymorphic SSR markers were analysed in terms of the above-mentioned two classes. While higher PIC value (> 0.50) was shown for 34.6% genomic and 9.5% genic SSR markers, the remaining 65.4% genomic and 90.5% genic SSR markers had the lower

Table 5: Distribution of polymorphic markers into different repeat classes

SSR type	Repeat classes	Polymorphic markers (%)	PIC value range (mean)	Number of alleles (mean)
Compound		72 (7.2)	0.12–0.80 (0.39)	2–9 (3.5)
Class I	NN	124 (38.4)	0.11–0.86 (0.43)	2–14 (3.9)
	NNN	164 (50.8)	0.11–0.76 (0.36)	2-6 (2.9)
	NNNN	19 (5.9)	0.11–0.49 (0.32)	2–4 (3.0)
	NNNNN	8 (2.5)	0.11–0.54 (0.31)	2–4 (2.8)
	NNNNNN	8 (2.5)	0.12–0.43 (0.28)	2–3 (2.5)
	Total	323 (32.2)	0.11–0.86 (0.36)	2–14 (3.3)
Class II	NN	142 (23.3)	0.11–0.89 (0.31)	2–11 (3.0)
	NNN	443 (72.7)	0.11–0.87 (0.26)	2–14 (2.6)
	NNNN	24 (3.9)	0.12–0.61 (0.24)	2–5 (2.6)
	Total	609 (60.6)	0.11–0.89 (0.27)	2–14 (2.7)
Grand total		1004	0.11–0.89 (0.31)	2–14 (3.2)

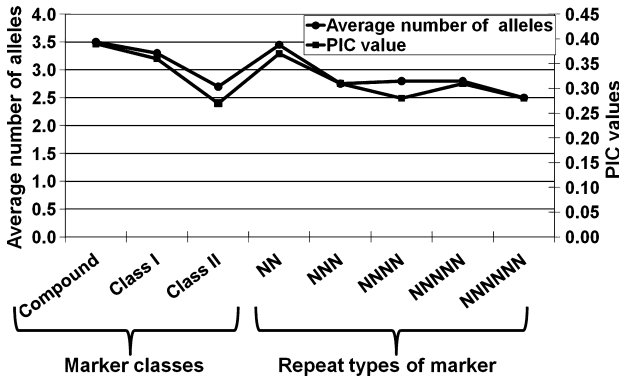

Fig. 1: Relationships of average number of alleles detected and PIC values of SSR markers with their respective classes and repeat types

PIC value (<0.50) (Table 7, Fig. 2). This clearly indicates that genomic SSR markers as compared to genic SSR markers display more polymorphism.

Genetic relationships between parental genotypes

Based on the allelic data obtained for all 1020 polymorphic SSR loci on 11 parental genotypes, a similarity matrix was generated (Data S2). Similarity index of these 1020 marker loci ranged from 0.044 to 0.842. It was found that the two most closely related genotypes were ICGS 44 and ICGS 76 with the highest similarity index (0.842). On the other hand, two most distantly related cultivars were T*x*AG-6 and ICG 11337 with lowest similarity index (0.044). Similarity matrix was used to prepare dendrogram using software NTSYSpc, which grouped 11 tetraploid parental genotypes into three major clusters: cluster A ('cl A'), cluster B ('cl B') and cluster C ('cl C') (Fig. 3). While 'cl B' (ICG 11337) and 'cl C' (TxAG-6) contained single genotype each, the 'cl A' contained remaining nine genotypes. The major cluster, 'cl A' is consisting of two sub clusters i.e. 'cl AI' (ICGS 44, ICGS 76, CSMG 84-1) and 'cl AII' (ICGV 86031, TAG 24, TG 26, GPBD 4, TMV 2, JL 24). Two *hypogaea*/Virginia-type genotypes (ICGS 76, CSMG 84-1) are clustered with one *vulgaris*/Spanish-type genotype (ICGS 44) in subcluster 'cl AI', while two *hypogaea*/Virginia-type genotypes (TAG 24, TG 26) are clustered with four *vulgaris*/Spanish-type genotypes (ICGV 86031, GPBD 4, TMV 2, JL 24) in subcluster 'cl AII'.

Discussion

In many regions of the world, the genetic yield potential of groundnut is not reached because of biotic and abiotic stresses. Marker-assisted selection is an important tool to enhance tolerance/resistance to these stresses and has the potential to enable faster and larger gains through genetic improvement. However, until recently, the implementation of marker-assisted selection was severely hampered by the very limited genomic resources available for groundnut (Varshney et al. 2007). Over the last few years, about 5000 SSR markers have been developed for groundnut (Ferguson et al. 2004, Moretzsohn et al. 2004, Nelson et al. 2006, Proite et al. 2007, Wang et al. 2007, Cuc et al. 2008, Liang et al. 2009). However, only a few hundred SSR markers have been mapped. This was mainly because of two reasons: (i) limited genetic diversity in the mapping populations and (ii) use of limited number of SSR markers by different research groups. While low level of genetic diversity is an inherent genetic constraint in cultivated groundnut, we reasoned that the identification of a highly informative set of SSR markers would help the community focus marker screening on potentially polymorphic markers instead of using all available SSR markers, most of which have a low potential. Therefore, an attempt was made to identify a highly informative set of SSR markers using a starting set of >4400 SSR markers and 20 genotypes representing parents of 15 mapping populations.

Although 1351 SSR markers showed polymorphism in the genotypes analysed in the study, only 1020 SSR markers that had high-quality data for at least 11 of 20 genotypes were fully analysed. Of the 1020 polymorphic markers, the highest polymorphism was obtained in the diploid AA-genome mapping population (60.5%) followed by TMV $2 \times$ TxAG-6 (42%) population. On the other hand, a low level of polymorphism was observed in the mapping populations of cultivated genotypes ranging from 2.4% (GPBD $5 \times$ GPBD 4) to 11.1% (Chico $\times$ CSMG 84-1) with an average of 5.58%. A high level of polymorphism (46.8% of SSRs and ca. 1 single-nucleotide polymorphism/90 bp) has been previously observed earlier in the AA-genome mapping population (K7988 $\times$ V10309, Moretzsohn et al. 2005, Bertioli et al. 2009). Similarly, in the mapping population involving synthetic amphi-diploid (TxAG6), a high polymorphism (66.0%) rate has been previously observed (Burrow et al. 2001). The genetic base of the cultivated groundnut is very narrow, and the low levels of genetic diversity observed in cultivated material in the present study are in line with those of earlier studies (Varshney et al. 2009a, Khedikar et al. 2010, Ravi et al. 2011, Sarvamangala et al. 2011).

Number of alleles detected and the PIC value based on the frequencies of different alleles in the germplasm surveyed by a particular marker indicate the quality (discriminatory power) of the marker. Number of alleles ranged from 2 to 14 (average 3.2) per marker in the present study and was high as compared to those of the earlier genetic diversity studies (He et al. 2003, Krishna et al. 2004, Moretzsohn et al. 2004, Cuc et al. 2008, Gautami et al. 2009, Liang et al. 2009) as they reported 2–8 alleles per marker. Comparable results (2–13 alleles) with Song et al.'s (2010) and higher alleles (2–20) as compared to Varshney et al.'s (2009b) were also reported. Similarly, the PIC value for polymorphic markers ranged from 0.10 to 0.89 with an average of 0.31 per marker. In total, only 15.9% markers could show PIC value more than 0.50. The PIC values observed here are in agreement of earlier genetic diversity studies (Mace et al. 2007, Cuc et al. 2008, Gautami et al. 2009, Liang et al. 2009, Varshney et al. 2009b). Like earlier studies (Varshney et al. 2002, Moretzsohn et al. 2005, Song et al. 2010), the present study also reported dinucleotide and trinucleotide repeat SSR markers as highly polymorphic markers. While a negative correlation was observed between the PIC value and repeat unit classes, a positive correlation was observed between PIC value and number of alleles (Cuc et al. 2008).

In terms of comparison of informativeness of SSR markers based on the origin of DNA sequences, the genomic SSRs showed higher level of polymorphism as compared to genic SSR markers. This is in agreement with general conception that genic SSRs show low level of polymorphism as compared to genomic SSRs as genic SSRs originate from highly

Table 6: Details of highly polymorphic markers identified in the present study

S. No.	Marker IDs	Across total genotypes analyzed		Across only cultivated genotypes	
		Allele no.	PIC values	Allele no.	PIC values
1	pPGPSeq04D02	3	0.56	3	0.50
2	pPGPSeq04G01	3	0.56	3	0.56
3	pPGPSeq15F12	4	0.67	4	0.67
4	TC11F12	3	0.56	2	0.36
5	TC11H06	4	0.67	5	0.64
6	TC2B09	4	0.67	3	0.56
7	TC3B05	3	0.56	4	0.61
8	TC3G05	3	0.56	3	0.56
9	TC4F12	4	0.67	4	0.67
10	TC7A02	4	0.67	4	0.67
11	TC7E04	4	0.67	5	0.64
12	gi-427	3	0.56	3	0.56
13	IPAHM177	3	0.56	3	0.56
14	IPAHM229	3	0.56	6	0.77
15	IPAHM395	4	0.67	4	0.67
16	IPAHM509	4	0.67	4	0.67
17	IPAHM689	5	0.77	5	0.77
18	IPAHM93	3	0.56	3	0.56
19	PM183	3	0.56	3	0.50
20	PM238	3	0.56	3	0.50
21	PM3	4	0.67	4	0.67
22	PM35	4	0.67	3	0.50
23	PM434	4	0.67	4	0.67
24	S001	8	0.81	7	0.80
25	S003	5	0.62	3	0.49
26	S009	11	0.89	3	0.56
27	S011	5	0.58	3	0.41
28	S016	4	0.55	5	0.72
29	S019	9	0.85	4	0.61
30	S021	5	0.62	4	0.61
31	S022	3	0.50	8	0.84
32	S023	6	0.75	4	0.58
33	S024	5	0.68	3	0.47
34	S026	4	0.56	5	0.70
35	S038	8	0.76	3	0.55
36	S040	6	0.75	2	0.36
37	S046	4	0.62	4	0.58
38	S048	3	0.58	7	0.76
39	S049	8	0.80	6	0.75
40	S052	6	0.72	3	0.56
41	S057	6	0.76	3	0.55
42	S059	4	0.55	7	0.80
43	S068	8	0.80	6	0.77
44	S070	5	0.69	6	0.77
45	S072	3	0.55	3	0.50
46	S073	4	0.64	7	0.80
47	S076	5	0.64	4	0.65
48	S080	5	0.69	3	0.55
49	S083	6	0.78	4	0.61
50	S084	4	0.67	4	0.61
51	S086	5	0.68	4	0.65
52	S093	5	0.71	6	0.77
53	S096	4	0.53	4	0.69
54	S101	3	0.54	4	0.61
55	S108	4	0.62	9	0.87
56	S113	5	0.69	3	0.59
57	S118	5	0.64	3	0.47
58	GM744	8	0.85	7	0.79
59	GM761	3	0.50	2	0.29
60	GM822	5	0.72	6	0.79
61	GM840	9	0.86	6	0.79
62	GM995	5	0.65	4	0.61
63	GM1043	4	0.60	3	0.49
64	GM1073	4	0.56	3	0.57
65	GM1076	4	0.50	3	0.44
66	GM1089	4	0.53	2	0.37

Table 6: (Continued)

S. No.	Marker IDs	Across total genotypes analyzed		Across only cultivated genotypes	
		Allele no.	PIC values	Allele no.	PIC values
67	GM1097	4	0.52	3	0.50
68	GM1098	5	0.59	3	0.44
69	GM1202	4	0.54	3	0.50
70	GM1256	3	0.54	3	0.57
71	GM1357	5	0.60	2	0.37
72	GM1369	4	0.54	2	0.35
73	GM1411	3	0.56	3	0.50
74	GM1469	4	0.50	2	0.35
75	GM1477	5	0.61	2	0.35
76	GM1483	5	0.60	2	0.37
77	GM1489	4	0.62	4	0.57
78	GM1501	4	0.54	2	0.37
79	GM1502	7	0.74	4	0.48
80	GM1515	4	0.54	3	0.50
81	GM1533	5	0.67	4	0.69
82	GM1538	3	0.51	2	0.29
83	GM1555	4	0.58	3	0.53
84	GM1562	3	0.56	3	0.50
85	GM1565	4	0.50	2	0.35
86	GM1575	3	0.52	3	0.50
87	GM1577	5	0.69	6	0.75
88	GM1664	3	0.59	4	0.48
89	GM1745	3	0.55	3	0.34
90	GM1760	5	0.58	4	0.57
91	GM1773	3	0.54	4	0.66
92	GM1834	5	0.61	2	0.37
93	GM1839	3	0.50	2	0.29
94	GM1842	4	0.56	3	0.50
95	GM1845	3	0.59	2	0.18
96	GM1863	5	0.72	6	0.79
97	GM1864	5	0.73	6	0.76
98	GM1869	3	0.50	2	0.35
99	GM1879	3	0.55	4	0.66
100	GM1907	3	0.50	4	0.57
101	GM1911	5	0.69	5	0.68
102	GM1937	4	0.60	4	0.61
103	GM1949	3	0.53	3	0.59
104	GM1954	4	0.57	4	0.57
105	GM1958	4	0.52	3	0.49
106	GM1959	6	0.65	3	0.49
107	GM1960	4	0.50	2	0.35
108	GM1977	4	0.53	2	0.35
109	GM1986	6	0.78	4	0.66
110	GM1991	6	0.71	4	0.57
111	GM1992	4	0.53	3	0.34
112	GM1996	6	0.73	5	0.68
113	GM2009	7	0.77	4	0.57
114	GM2024	4	0.57	3	0.53
115	GM2053	4	0.65	4	0.64
116	GM2084	5	0.64	3	0.53
117	GM2103	5	0.62	3	0.49
118	GM2165	4	0.58	3	0.59
119	GM2206	5	0.50	3	0.44
120	GM2215	4	0.56	3	0.57
121	GM2348	4	0.70	3	0.34
122	GM2407	4	0.53	2	0.37
123	GM2444	5	0.60	3	0.50
124	GM2478	6	0.51	2	0.18
125	GM2482	3	0.52	3	0.44
126	GM2504	7	0.74	5	0.70
127	GM2522	4	0.52	3	0.50
128	GM2528	6	0.54	3	0.34
129	GM2531	6	0.51	3	0.34
130	GM2589	5	0.52	4	0.48
131	GM2602	6	0.70	3	0.44
132	GM2603	4	0.58	4	0.57
133	GM2605	7	0.78	4	0.64

Table 6: (Continued)

S. No.	Marker IDs	Across total genotypes analyzed		Across only cultivated genotypes	
		Allele no.	PIC values	Allele no.	PIC values
134	GM2606	5	0.50	2	0.29
135	GM2623	4	0.54	2	0.37
136	GM2637	6	0.75	5	0.68
137	GM2638	5	0.68	6	0.71
138	GM2671	4	0.57	2	0.35
139	GM2730	4	0.62	4	0.69
140	GM2746	4	0.50	4	0.57
141	GNB0018	14	0.86	8	0.79
142	GNB0038	4	0.56	3	0.55
143	GNB0058	8	0.63	7	0.68
144	GNB0073	4	0.61	3	0.57
145	GNB0098	6	0.68	5	0.65
146	GNB0100	5	0.64	4	0.57
147	GNB0107	5	0.69	3	0.59
148	GNB0126	5	0.56	3	0.5
149	GNB0136	7	0.7	6	0.7
150	GNB0145	6	0.73	4	0.64
151	GNB0155	6	0.69	4	0.57
152	GNB0159	4	0.62	4	0.66
153	GNB0167	5	0.67	4	0.61
154	GNB0178	5	0.68	5	0.73
155	GNB0181	5	0.67	4	0.64
156	GNB0262	5	0.66	3	0.57
157	GNB0284	4	0.58	3	0.57
158	GNB0303	7	0.75	4	0.69
159	GNB0317	5	0.68	5	0.68
160	GNB0344	4	0.66	4	0.57
161	GNB0357	5	0.72	5	0.7
162	GNB0378	6	0.62	5	0.62
163	GNB0387	5	0.54	4	0.57
164	GNB0392	4	0.51	2	0.35
165	GNB0397	4	0.55	3	0.45
166	GNB0417	5	0.68	4	0.64
167	GNB0428	5	0.71	5	0.7
168	GNB0461	4	0.53	2	0.37
169	GNB0464	6	0.79	6	0.79
170	GNB0467	5	0.7	3	0.49
171	GNB0515	14	0.87	10	0.84
172	GNB0555	7	0.74	5	0.73
173	GNB0569	6	0.68	4	0.66
174	GNB0608	3	0.53	3	0.52
175	GNB0643	3	0.5	3	0.53
176	GNB0667	9	0.68	5	0.58
177	GNB0679	5	0.7	5	0.69
178	GNB0682	10	0.76	8	0.76
179	GNB0712	5	0.72	3	0.57
180	GNB0716	4	0.6	3	0.54
181	GNB0733	5	0.66	4	0.61
182	GNB0775	4	0.58	3	0.54
183	GNB0782	4	0.53	3	0.5
184	GNB0840	8	0.82	5	0.7
185	GNB0842	9	0.75	5	0.61
186	GNB0850	3	0.51	3	0.5
187	GNB0853	5	0.64	3	0.49
188	GNB0981	4	0.55	3	0.49
189	GNB0991	5	0.52	4	0.57
190	GNB1001	5	0.72	4	0.64
191	GNB1026	4	0.51	2	0.18
192	GNB1027	6	0.71	6	0.77
193	GNB1055	5	0.68	3	0.5
194	GNB1056	5	0.62	3	0.49
195	GNB1069	5	0.63	3	0.54
196	GNB1072	9	0.64	9	0.7
197	GNB1112	4	0.53	3	0.5
198	GNB1114	7	0.79	6	0.76
199	GNB1148	5	0.53	2	0.38
	Average	4.86	0.63	3.84	0.56

Table 7: Variation in PIC values between genomic and genic SSRs

S. No.	PIC value range	Genomic SSRs no. (%)	Genic SSRs no. (%)	Total markers no. (%)
1	0.10–0.20	51 (19.6)	284 (37.4)	335 (32.8)
2	0.21–0.30	55 (21.2)	196 (25.8)	251 (24.6)
3	0.31–0.40	35 (13.5)	146 (19.2)	181 (17.7)
4	0.41–0.50	29 (11.2)	62 (8.2)	91 (8.9)
5	0.51–0.60	24 (9.2)	44 (5.8)	68 (6.7)
6	0.61–0.70	39 (15.0)	15 (1.9)	54 (5.3)
7	0.71–0.80	21 (8.1)	11 (1.4)	32 (3.1)
8	0.81–0.90	6 (2.3)	2 (0.3)	8 (0.8)
Total markers		260	760	1020

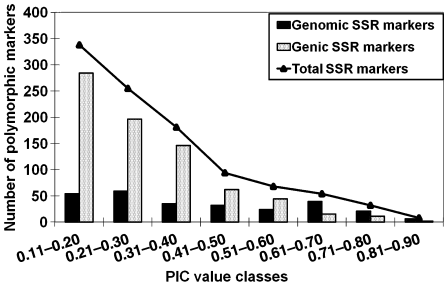

Fig. 2: Classification of polymorphic genomic and genic SSR markers into different classes of PIC values

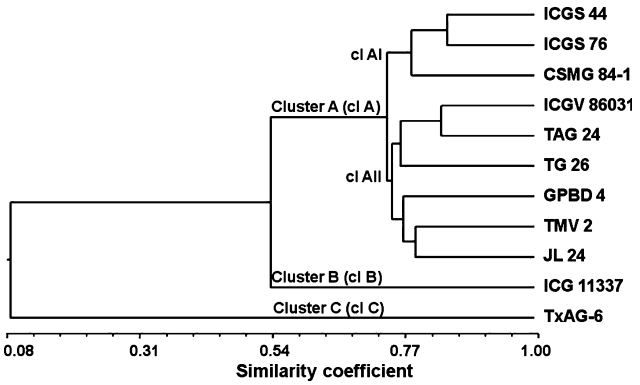

Fig. 3: Dendrogram showing genetic relationship between parental genotypes of different mapping populations

conserved portion of the genome (Varshney et al. 2005b). Hence, we suggest that development of genomic SSR markers should be given priority over genic SSRs in crops like groundnut that have a narrow genetic background.

The dendrogram constructed based on allelic data for all 1020 polymorphic markers classified all the genotypes into three groups. Majority of the genotypes clustered according to their pedigree and origin. It has also been found that even though the parents of the mapping population were found to be diverse based on the morphological traits, they (ICGS 44 and ICGS 76) clustered together with the highest similarity index (0.842). This has also reflected in polymorphism percentage between two populations developed using three parents (ICGS 44, ICGS 76 and CSMG 84-1) for drought-related traits (ICGS 44 × ICGS 76 and ICGS 76 × CSMG 84-1). The population derived from the cross ICGS 76 × CSMG 84-1 showed higher polymorphism (4.9%) as compared to the

population developed from the cross ICGS 44 × ICGS 76 (3.4%). Two most distantly related cultivars were TxAG 6 and ICG 11337 with low similarity index (0.044) and grouped separately in two clusters. This is because TxAG 6 is a synthetic amphidiploid derived from the cross *A. batizocoi* × (*A. cardenasii* × *A. diogoi*) and TMV 2 being a cultivated variety. Although majority of the *hypogaea*/Virginia and *vulgaris*/Spanish genotypes are clustered separately in subclusters 'cl AI' and 'cl AII', respectively, they could not differentiate the two botanical types (*hypogaea* and *vulgaris*) and market types (Virginia and Spanish) clearly. This may be due to higher relatedness between these two botanical types. This has also reflected in the earlier studies (Kottapalli et al. 2007, Varshney et al. 2009b).

The most important feature of this study is the identification of a set of 199 SSR markers that have higher PIC values and have the potential to detect more alleles in a set of germplasm accessions, or more polymorphism between a pair of parental genotypes. This set was identified after analysing a range of genotypes including cultivated, two AA-genome species genotypes and one synthetic amphidiploid. Therefore, the markers of this set should be very useful for genetic analysis in wild *Arachis* species as well as applications in the groundnut molecular breeding. The use of this SSR marker set should economize screening time and would facilitate the cross-references of genetic maps, including the linking of cultivated maps to information-rich diploid maps, and a unified genetic map for the legumes (Bertioli et al. 2009, Foncéka et al. 2009, Leal-Bertioli et al. 2009). Therefore, we recommend that the community should give the identified set of SSR markers priority while framing strategies for studying genetic diversity, linkage mapping, QTL analysis and marker-assisted breeding.

In summary, this study reports the primer sequences for 946 novel SSR markers for the first time, the analysis of 4485 SSR markers on a set of 20 genotypes and the identification of a most informative set of 199 SSR markers. We hope that the details provided in tables and Data S1 for all polymorphic SSR markers in addition to the informative set of SSR markers will benefit international groundnut research and molecular breeding.

Acknowledgements

Thanks are due to Mr. B. J. Moss, Mr. G. Somaraju and Mr. Abdul Gafoor for their technical help in conducting laboratory experiments. Financial support from National Fund of Indian Council of Agricultural Research (NBFSRA), New Delhi, India, and Tropical Legume I-Objective 1 project of CGIAR Generation Challenge Programme (http://www.generationcp.org), Mexico, and Bill and Melinda Gates Foundation (BMGF), USA, is gratefully acknowledged.

References

Bertioli, D. J., M. C. Moretzsohn, L. H. Madsen, N. Sandal, S. Leal-Bertioli, P. M. Guimarães, B. K. Hougaard, J. Fredslund, L. Schauser, A. M. Nielsen, S. Sato, S. Tabata, S. Cannon, and J. Stougaard, 2009: An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. BMC Genomics **10,** 45.

Bravo, J. P., A. A. Hoshino, C. M. L. C. D. Angelici, C. R. Lopes, and M. A. Gimenes, 2006: Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some *Arachis* section species of the genus *Arachis*. Genet. Mol. Biol. **29,** 516—524.

Burrow, M. D., C. E. Simpson, J. L. Starr, and A. H. Paterson, 2001: Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.):broadening the gene pool of a monophyletic polyploid species. BMC Plant Biol. **8,** 14.

Cuc, L. M., E. S. Mace, J. H. Crouch, V. D. Quang, T. D. Long, and R. K. Varshney, 2008: Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea* L.). BMC Plant Biol. **8,** 55.

Dwivedi, S., D. J. Bertioli, J. H. Crouch, J. F. M. Valls, H. D. Upadhyaya, A. P. Favero, M. C. Moretzsohn, and A. H. Paterson, 2007: Peanut genetics and genomics: toward marker-assisted genetic enhancement in peanut (*Arachis hypogaea* L). In: C. Kole (ed), Oilseeds Series: Genome Mapping and Molecular Breeding in Plants. Vol. 2, 115—151. Springer, Berlin, Heidelberg.

FAO, 2009: FAO statistical database. Available at: http://fastat.fao.org/ (last accessed on May 3, 2011).

Ferguson, M. E., M. D. Burow, S. R. Schultze, P. J. Bramel, A. H. Paterson, S. Kresovich, and S. Mitchell, 2004: Microsatellite identification and characterization in peanut (*A. hypogaea* L.). Theor. Appl. Genet. **108,** 1064—1070.

Foncéka, D., T. Hodo-Abalo, R. Rivallan, I. Faye, M. N. Sall, O. Ndoye, A. P. Fávero, D. J. Bertioli, J. C. Glaszmann, B. Courtois, and J. F. Rami, 2009: Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid. BMC Plant Biol. **9,** 103.

Gautami, B., K. Ravi, M. L. Narasu, D. A. Hoisington, and R. K. Varshney, 2009: Novel set of groundnut SSR markers for germplasm analysis and inter-specific transferability. Int. J. Integrative Biol. **7,** 100—106.

Gimenes, M. A., A. A. Hosino, A. V. G. Barbosa, D. A. Palmieri, and C. R. Lopes, 2007: Characterization and transferability of microsatellite markers of cultivated peanut (*Arachis hypogaea*). BMC Plant Biol. **7,** 9.

Gupta, P. K., and R. K. Varshney, 2000: The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica **113,** 163—185.

Halward, T. M., H. T. Stalker, E. A. Larue, and G. Kochert, 1991: Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and related wild-species. Genome **34,** 1013—1020.

He, G. H., R. H. Meng, M. Newman, G. Q. Gao, R. N. Pittman, and C. S. Prakash, 2003: Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biol. **3,** 3.

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

Hilu, K. W., and H. T. Stalker, 1995: Genetic relationships between peanut and wild species of *Arachis* sect. *Arachis* (Fabaceae): evidence from RAPDs. Plant Syst. Evol. **198,** 167—178.

Hopkins, M. S., A. M. Casa, T. Wang, S. E. Mitchell, R. E. Dean, G. D. Kochert, and S. Kresovich, 1999: Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Sci. **39,** 1243—1247.

Khedikar, Y. P., M. V. C. Gowda, C. Sarvamangala, K. V. Patgar, H. D. Upadhyaya, and R. K. Varshney, 2010: A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). Theor. Appl. Genet. **121,** 71—984.

Kochert, G., H. T. Stalker, M. Gimenes, L. Galgaro, C. R. Lopes, and K. Moore, 1996: RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut (*Arachis hypogaea* L.). Am. J. Bot. **83,** 1282—1291.

Kottapalli, K. R., M. D. Burrow, G. Burrow, J. Burke, and N. Puppala, 2007: Molecular characterization of the U. S. peanut minicore collection using microsatellite markers. Crop Sci. **47,** 1718—1727.

Krishna, G. K., J. F. Zhang, M. Burrow, R. N. Pittman, S. G. Delikostadinov, Y. Z. Lu, and N. Puppala, 2004: Genetic diversity analysis in Valencia peanut (*Arachis hypogaea* L.) using microsatellite markers. Cell Mol. Biol. Lett. **9,** 685—697.

Leal-Bertioli, S. C. M., A. C. F. V. José, D. M. T. Alves-Freitas, M. C. Moretzsohn, P. M. Guimarães, S. Nielen, B. Vidigal, R. W. Pereira, J. Pike, A. P. Fávero, M. Parniske, R. K. Varshney, and D. J. Bertioli, 2009: Identification of candidate genome regions controlling disease resistance in *Arachis*. BMC Plant Biol. **9,** 112.

Liang, X., X. Chen, Y. Hong, H. Liu, G. Zhou, S. Li, and B. Guo, 2009: Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. BMC Plant Biol. **9,** 35.

Liu, K., and S. Muse, 2005: PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics Appl. **21,** 2128—2129.

Mace, E. S., R. K. Varshney, V. Mahalakshmi, K. Seetha, A. Gafoor, Y. Leeladevi, and J. H. Crouch, 2007: In silico development of simple sequence repeat markers within the aeschynomenoid/dalbergoid and genistoidclades of the Leguminosae family and their transferability to *Arachis hypogaea*, groundnut. Plant Sci. **174,** 51—60.

Moretzsohn, M. C., M. S. Hopkins, S. E. Mitchell, S. Kresovich, J. F. M. Valls, and M. E. Ferreira, 2004: Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hyper variable regions of the genome. BMC Plant Biol. **4,** 11.

Moretzsohn, M. C., L. Leoi, K. Proite, P. M. Guimaraes, S. C. Leal-Bertioli, M. A. Gimenes, W. S. Martins, J. F. M. Valls, D. Grattapaglia, and D. J. Bertioli, 2005: A microsatellite-based gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). Theor. Appl. Genet. **111,** 1060—1071.

Nagy, E., Y. Chu, Y. Guo, S. Khanal, S. Tang, Y. Lee, W. Dong, P. Timper, C. Taylor, P. Ozias-Akins, C. Holbrook, V. Beilinson, N. Nelson, T. Stalker, and S. Knapp, 2010: Recombination is suppressed in an alien introgression in peanut harbouring *Rma*, a dominant root-knot nematode resistance gene. Mol. Breed. **26,** 357—370.

Nelson, M. N., H. T. T. Phan, S. R. Ellwood, M. Moolhuijzen Paula, J. Hane, A. Williams, C. E. O'Lone, J. F. Nyarko, M. Scobie, M. Cakir, M. G. K. Jones, M. Bellgard, B. Wolko, S. J. Barker, R. P. Oliver, and W. A. Cowling, 2006: The first gene-based map of *Lupinus angustifolius* L.-location of domestication genes and conserved synteny with *Medicago truncatula*. Theor. Appl. Genet. **113,** 225—238.

Palmieri, D. A., A. A. Hoshino, J. P. Bravo, C. R. Lopes, and M. A. Gimenes, 2002: Isolation and characterization of microsatellite loci from the forage species *Arachis pintoi* (Genus *Arachis*). Mol. Ecol. Notes **2,** 551—553.

Palmieri, D. A., M. D. Bechara, R. A. Curi, M. A. Gimenes, and C. R. Lopes, 2005: Novel polymorphic microsatellite markers in section *Caulorrhizae* (*Arachis,* Fabaceae). Mol. Ecol. Notes **5,** 77—79.

Proite, K., S. C. Leal-Bertioli, D. J. Bertioli, M. C. Moretzsohn, F. R. da Silva, N. F. Martins, and P. M. Guimaraes, 2007: ESTs from a wild *Arachis* species for gene discovery and marker development. BMC Plant Biol. **7,** 7.

Rao, K., N. L. J. Reddy, and P. J. Bramel, 2003: Potential of wild species for genetic enhancement of some semi-arid food crops. Genet. Resour. Crop Evol. **50,** 707—721.

Ravi, K., V. Vadez, S. Isobe, R. R. Mir, Y. Guo, S. N. Nigam, M. V. C. Gowda, T. Radhakrishnan, D. J. Bertioli, S. J. Knapp, and R. K. Varshney, 2011: Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). Theor. Appl. Genet. **122,** 1119—1132.

Rohlf, J. F., 2000: NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Exeter Software, Setauket, NY.

Sarvamangala, C., M. V. C. Gowda, and R. K. Varshney, 2011: Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). Field Crops Res. **122,** 49—59.

Song, G. Q., M. J. Li, H. Xiao, X. J. Wang, R. H. Tang, H. Xia, C. Z. Zhao, and Y. P. Bi, 2010: EST sequencing and SSR marker development for cultivated peanut (*Arachis hypogaea* L.). Electronic J. Biotech., **13,** 1—9.

Stalker, H. T., and R. D. Dalmacio, 1986: Karyotype analysis and relationships among varieties of *Arachis hypogaea* L. Cytologia **58,** 617—629.

Stalker, H. T., and C. E. Simpson, 1995: Germplasm resources in *Arachis*. In: H. E. Pattee, and H. T. Stalker (eds), Advances in Peanut Science, 14—53. American Peanut Research and Education Society, Inc, Stillwater, OK.

Stebbins, G. L., 1957: Genetics, evolution, and plant breeding. Indian J. Genet. Plant Breed. **17,** 129—141.

Subramanian, V., S. Gurtu, R. C. NageswaraRao, and S. N. Nigam, 2000: Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome **43,** 656—660.

Varshney, R. K., T. Thiel, N. Stein, P. Langridge, and A. Graner, 2002: *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cell. Mol. Biol. Lett. **7,** 537—546.

Varshney, R. K., A. Graner, and M. E. Sorrells, 2005a: Genomics-assisted breeding for crop improvement. Trends Plant Sci. **10,** 621—630.

Varshney, R. K., A. Graner, and M. E. Sorrells, 2005b: Genic microsatellite markers in plants: features and applications. Trends Biotechnol. **23,** 48—55.

Varshney, R. K., D. A. Hoisington, and A. K. Tyagi, 2006: Advances in cereal genomics and applications in crop breeding. Trends Biotechnol. **24,** 490—499.

Varshney, R. K., D. A. Hoisington, H. D. Upadhyaya, P. M. Gaur, S. N. Nigam, K. Saxena, V. Vadez, N. K. Sethy, S. Bhatia, R. Aruna, M. V. C. Gowda, and N. K. Singh, 2007: Molecular genetics and breeding of grain legume crops for the semi-arid tropics. In: R. K. Varshney, and R. Tuberosa (eds), Genomic Assisted Crop Improvement Genomics Applications in Crops, 207—2420. Springer, Dordrecht, The Netherlands.

Varshney, R. K., D. J. Bertioli, M. C. Moretzsohn, V. Vadez, L. Krishnamurthy, R. Aruna, S. N. Nigam, B. J. Moss, K. Seetha, K. Ravi, G. He, S. J. Knapp, and D. A. Hoisington, 2009a: The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). Theor. Appl. Genet. **118,** 729—739.

Varshney, R. K., T. Mahendar, R. Aruna, S. N. Nigam, V. Vadez, and D. A. Hoisington, 2009b: High level of natural variation in a groundnut (*Arachis hypogaea* L.) germplasm collection assayed by selected informative SSR markers. Plant Breed. **128,** 486—494.

Wang, C. T., X. D. Yang, D. X. Chen, S. L. Yu, G. Z. Liu, Y. Y. Tang, and J. Z. Xu, 2007: Isolation of simple sequence repeats from groundnut. Electronic J. Biotech. **10,** 473—480.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Details of all polymorphic genomic and genic SSR markers in groundnut.

Data S2. Genetic similarity among 11 groundnut genotypes based on 1020 polymorphic markers.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.