Euphytica (2005) 142: 131–136
DOI: 10.1007/s10681-005-1043-3

Simple sequence repeat markers for botanical varieties of cultivated peanut (*Arachis hypogaea* L.)

Guohao He[1,*], Ronghua Meng[1], Hui Gao[1], Baozhu Guo[2], Guoqing Gao[3], Melanie Newman[4], Roy N Pittman[4] & C.S. Prakash[1]
[1]*Department of Agricultural Sciences, Tuskegee University, Tuskegee, AL 36088;* [2]*USDA-ARS, Crop Protection and Management Research Unit, Tifton, GA 31793;* [3]*Guangxi Academy of Agricultural Sciences, Nanning, China;* [4]*USDA-ARS, Plant Genetic Resources Conservation Unit, Griffin, GA 30223;*
*(*author for correspondence: e-mail: hguohao@tuskegee.edu)*

Received 12 August 2004; accepted 20 January 2005

Key words: botanical variety identification, peanut, simple sequence repeat (SSR)

Summary

Cultivated peanut (*Arachis hypogaea* L.) consists of six botanical varieties. Identification of DNA markers associated with botanical varieties would be useful in plant genotyping, germplasm management, and evolutionary studies. We have developed 130 simple sequence repeat (SSR) markers in peanut, 38 of which were used in this study because of their ability in detecting genetic polymorphism among 24 peanut accessions. Eight SSR markers were found useful to classify botanical varieties. Among them, six SSR markers were specific to botanical varieties *fastigiata* and *vulgaris*, one to botanical varieties *hypogaea* and *hirsuta*, and one to botanical varieties *peruviana*, and *aequatoriana*. Also, three of them derived from peanut expressed sequence tags (ESTs) were associated with putative genes. As botanical varieties have different morphological traits and belong to different subspecies in *A. hypogaea,* these markers might be associated with genes involved in the expression of morphological traits. The results also suggested that SSRs (also called microsatellites) might play a role in shaping evolution of cultivated peanut. Multiplex PCR of botanical variety-specific markers could be applied to facilitate efficient genotyping of the peanut lines.

Introduction

Peanut (*Arachis hypogaea* L.) is an important crop grown worldwide for its oil and protein. Cultivated peanut has six botanical varieties, *hypogaea, hirsuta, fastigiata, peruviana, aequatoriana*, and *vulgaris* (Krapovickas & Gregory, 1994). The first two botanical varieties belong to subspecies *hypogaea*, and the latter four to subspecies *fastigiata* based on their morphological traits and growth habits (Krapovickas & Gregory, 1994). Botanical varieties have their specific morphological characteristics, such as all botanical varieties in ssp. *fastigiata* having main stem fertility and erect plant habit while those in ssp. *hypogaea* being without floral axes on main stem and prostrate habit; botanical varieties *fastigiata, peruviana*, and *aequatoriana* have large leaflet size wheras botanical varieties *hypogaea, hirsuta*, and *vulgaris* have small leaflet size. However, many intermediates exist (Stalker & Simpson, 1995) and do not fit readily into either of the two recognized subspecies (Williams, 1989). Moreover, it is difficult to identify true hybrids between two subspecies. Therefore, it would be useful to use molecular markers coupled with morphological characteristics to detect hybrids (Scheef et al., 2003).

Our previous diversity study using amplified fragment length polymorphism (AFLP) markers of cultivated peanut showed that botanical varieties *fastigiata* and *vulgaris* were formed in a distinct group, while the other four botanical varieties were clustered in another

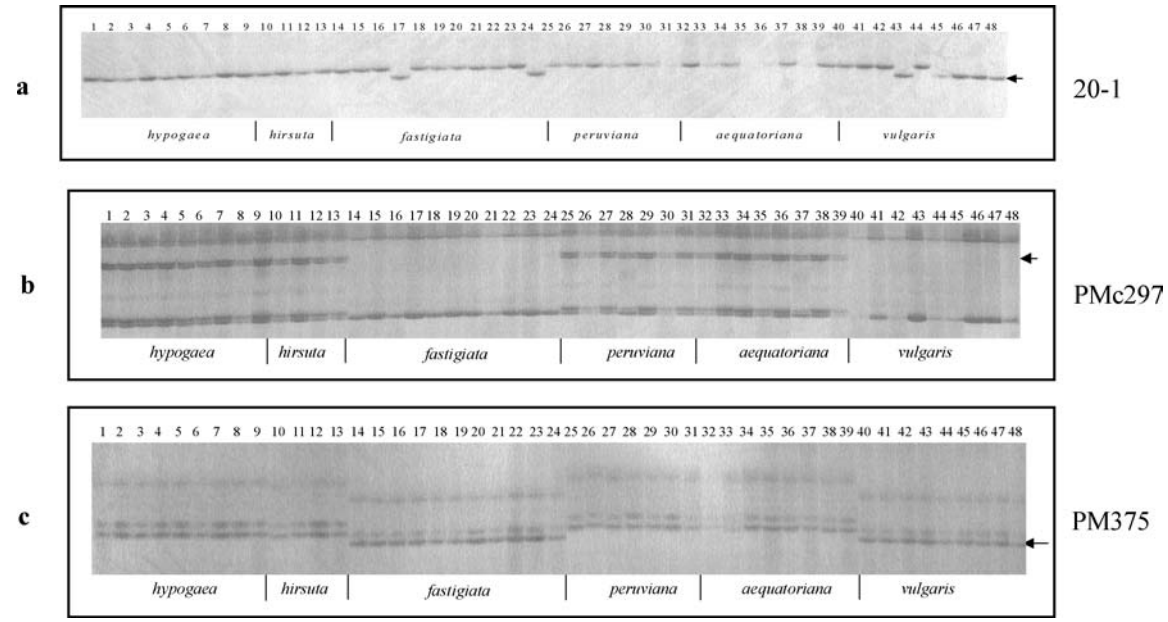

Figure 1. Gel images of SSR markers specific to botanical varieties in peanut.

group (He & Prakash, 2001). The result based on DNA markers was not consistent with the classification of *A. hypogaea* based on morphology. During that study, a co-dominant marker generated by an AFLP primer pair (E-ACC/M-CAC) was identified. In comparison of the sequences of the two alleles, 12 nucleotides were found missing in one allele (b) compared to the other allele (a). The primer pair was designed as 5′-CTCCCTTCCCATCTGTTGTC-3′, and 5′-AGCAGA AGCAAGATGAAGCAG-3′ to convert this AFLP marker into a sequence-tagged site (STS) marker (unpublished). Forty-eight cultivated accessions were screened using this STS marker in order to determine the frequency of alleles in the cultivated accessions studied. We have found that allele (b) was present only in botanical varieties *fastigiata* and *vulgaris*, and allele (a) occurred in all six botanical varieties. Although the occurrence of allele (b) in botanical varieties *fastigiata* and *vulgaris* was low, 0.16 (3/19) and 0.67 (8/12), respectively (Figure 1a), it gave us an indication that there might be specific DNA markers related to botanical varieties in peanut. Botanical variety-specific markers will be useful in germplasm management.

Multiplex PCR was used to simultaneously amplify two or more loci in a single PCR reaction for routine, repetitive genotyping more rapidly and cost efficiently, especially in high-throughput genotyping applications (Tang et al., 2003). PCR-multiplexes were studied for purposes such as genotyping in plant genetic resources (Mitchell et al., 1997), investigating soybean genome (Narvel et al., 2000), parentage testing in goats (Luikart et al., 1999), and mapping in cultivated sunflower (Tang et al., 2003). For genotyping applications such as cultivar identification and seed purity testing, botanical variety-specific markers assembled into PCR-multiplexes will be useful in robust, high-throughput techniques.

The objectives of this study were to identify botanical variety-specific markers using our novel developed SSR markers, to assess the relationship between the variety-specific markers with known genes, and to test PCR-multiplexes of botanical variety-specific markers for genotyping applications in peanut.

Materials and methods

Plant material and DNA extraction

Forty-eight cultivated accessions representing six botanical varieties were used to identify botanical variety-specific markers (Table 1). Plant material was obtained from USDA Plant Genetic Resources Conservation Unit, Griffin, GA. Young leaves were used for DNA extraction. Genomic DNA was isolated following the protocol of MasterPure Plant Leaf Purification kit

Table 1. List of peanut accessions used to determine botanical variety-specific markers

PI number	Botanical variety	Country of origin
PI 497253	*hypogaea*	Argentina
PI 468248	*hypogaea*	Bolivia
PI 475944	*hypogaea*	Bolivia
PI 475952	*hypogaea*	Bolivia
PI 497302	*hypogaea*	Bolivia
PI 497365	*hypogaea*	Bolivia
PI 475871	*hypogaea*	Brazil
PI 536279	*hypogaea*	Brazil
PI 576631	*hypogaea*	Mexico
PI 501296	*hirsuta*	Peru
PI 501297	*hirsuta*	Peru
PI 576617	*hirsuta*	Mexico
PI 576613	*hirsuta*	Mexico
PI 493811	*fastigiata*	Argentina
PI 475914	*fastigiata*	Bolivia
PI 536265	*fastigiata*	Brazil
PI 536269	*fastigiata*	Brazil
PI 536263	*fastigiata*	Brazil
PI 476052	*fastigiata*	Brazil
PI 576609	*fastigiata*	Mexico
PI 476181	*fastigiata*	Peru
PI 262023	*fastigiata*	Paraguay
PI 493309	*fastigiata*	Paraguay
PI 493678	*fastigiata*	Paraguay
PI 540833	*peruviana*	Bolivia
PI 540835	*peruviana*	Bolivia
PI 502043	*peruviana*	Peru
PI 502045	*peruviana*	Peru
PI 502053	*peruviana*	Peru
PI 502088	*peruviana*	Peru
PI 502096	*peruviana*	Peru
PI 497615	*aequatoriana*	Ecuador
PI 497623	*aequatoriana*	Ecuador
PI 497625	*aequatoriana*	Ecuador
PI 497626	*aequatoriana*	Ecuador
PI 497630	*aequatoriana*	Ecuador
PI 497631	*aequatoriana*	Ecuador
PI 497632	*aequatoriana*	Ecuador
PI 497634	*aequatoriana*	Ecuador
PI 493984	*vulgaris*	Argentina
PI 494052	*vulgaris*	Brazil
PI 494058	*vulgaris*	Brazil
PI 494005	*vulgaris*	Paraguay
PI 494009	*vulgaris*	Paraguay
PI 494048	*vulgaris*	Paraguay
PI 494024	*vulgaris*	Paraguay
PI 536217	*vulgaris*	Uruaguay
PI 536232	*vulgaris*	Uruaguay

(Epicentre, WI). DNA concentration was determined by DU640B spectrophotometer (Beckman Coulter, CA). DNA was diluted to 50 ng/μl in sterile water for PCR analysis.

SSR analysis

One hundred thirty peanut SSRs were developed from microsatellite-enriched library (He et al., 2003, GenBank accession number AY237736-AY237798, AY-310535-AY310564), and identified from peanut ESTs (Guo et al., 2003). The SSRs were previously screened on 24 cultivated accessions. Thirty-eight primers detecting DNA polymorphism among accessions were used for SSR analysis of 48 accessions in this study (Table 2). Among them, 29 were developed from microsatellite-enriched library, and 9 were identified from peanut ESTs. PCR reactions consisted of 50 ng template DNA, 1 $\times$ PCR buffer, 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 250 nM each of forward and reverse primer, and 0.25 U Taq polymerase in a 10-μl reaction volume. PCR amplifications were carried out in a Perkin–Elmer 9700 thermocycler as 94 °C/3 min for initial denaturation; 94 °C/30 s, 65°C/30 s, 72 °C/1 min for two cycles; 94 °C/30 s, 56 °C/30 s, 72 °C/1 min for two cycles; 94 °C/15 s, 55 °C/30 s, 72°C/1 min for 30 cycles; and 72 °C/10 min final extension (Mellersh & Sampson, 1993) for all primers. PCR products were run on a 6% denaturing polyacrylamide gel and visualized by silver staining. For EST-SSR, BLASTX algorithm was performed for gene identification (http://www.ncbi.nim.nih.gov/BLAST/).

PCR-multiplexes test

PCR-multiplexes were assembled by combination of two or three specific markers based on their expected amplicon size and strength of bands. Multiplex PCRs were performed following the protocols of Henegariu et al., 1997; Tang et al., 2003 with minor modification; that is 20 μl of reaction mixture containing 1.6 $\times$ PCR buffer, 2.5 mM of $MgCl_2$, 0.4 mM each of dNTPs, 6 pmol of each primer, 2 U Taq polymerase and 100 ng of genomic DNA. PCR program used was the same as described above.

Results

Eight of 38 primers tested amplified specific bands in particular botanical varieties, five of which were

Table 2. Peanut SSRs screened for botanical variety-specific markers

Name[a]	Sequence (5′–3′)	Repeat motif	Size
PM3:	GAA AGA AAT TAT ACA CTC CAA TTA TGC CGG CAT GAC AGC TCT ATG TT	$(GA)_{14}$	168
PM15:	CCT TTT CTA ACA CAT TCA CAC ATG A GGCTCCCTTCGATGATGAC	$(GAAA)_3(GA)_8$, $(GA)_9$	177
PM32:	AGT GTT GGG TGT GAA AGT GG GGG ACT CGG AAC AGT GTT TAT C	$(CT)_{15}$	101
PM35:	TGT GAA ACC AAA TCA CTT TCA TTC TGG TGA AAA GAA AGG GGA AA	$(GA)_{18}(GAA)_2$	135
PM36:	ACT CGC CAT AGC CAA CAA AC CATTCCCACAACTCCCACAT	$(GA)_{18}$	200
PM42:	ACG GGC CAA GTG AAG TGA T TCT TGC TTC TTT GGT GAT TAG C	$(GA)_4AA(GA)_{14}$	202
PM45:	TGA GTT GTG ACG GCT TGT GT GAT GCA TGT TTA GCA CAC TTG A	$(GA)_{16}$	101
PM50:	CAA TTC ATG ATA GTA TTT TAT TGG ACA CTT TCT CCT CCC CAA TTT GA	$(TAA)_4$, $(GA)_{19}$	103
PM53:	CCT ATC CTA TGG GTC ACT AGC C GCT TGT GCT CAT CTT GAG TTT T	$(AT)_4(ACT)_2$	113
PM65:	GGA CGT CTG GCT GCT AGA GA TCG GCA TCA AAA CAG TGA GA	$(CT)_{12}$	226
PM137:	AAC CAA TTC AAC AAA CCC AGT GAA GAT GGA TGA AAA CGG ATG	$(GA)_{20}$	150
PM145:	GCT GTA ATT AGG ATC ATT CCA CA CAA CGG TTG GAT CGA TGA	$(CT)_{12}(CA)_2(CT)_4(CA)_9$	173
PM183:	TTC TAA TGA AAA CCG ACA AGT TT CGT GCC AAT AGA GTT TTA TAC GG	$(CT)_{24}$	115
PM188:	GGG CTT CAC TGC TTT TGA TT TGCGACTTCTGAGAGGACAA	$(GA)_8$	102
PM200:	GCT ATG TGG GAA AAA TAC TGC TT CAG ATG TGT GTG TGT GTG TGT G	$(CT)_{22}(CA)_{16}$	156
PM201:	CCT TTA TAG AGG ACC TTC CCT CTC GCC TAT TTG GTA TCG GCT CA	$(CT)_{19}$	223
PM204:	TGG GCC TAA ACC CAA CCT AT CCA CAA ACA GTG CAG CAA TC	$(GA)_{20}$	216
PM210:	CCG CAG ATC TTC TCC TGT GT CCT CCT CAT CCT CTA AAC TCT GC	$(CT)_{25}$	204
PM238:	CTC TCC TCT GCT CTG CAC TG ACA AGA ACA TGG GGA TGA AGA	$(CT)_{11}$	150
PM325:	CCT AAC AAG GAC GGG TGA AC CAG AGG CCT CAC TTT CCT TC	$(GA)_{21}$	108
PM343:	AGA AAC GAG GAG CTC GAC AA GCT CAT TTT GAT GGA ATG AGA G	$(CT)_{22}$	203
PM346:	AAA GGC GCA CTC GAT TCT AA CGC ACA GAA ACA TCA AGC AT	$(GGC)_5$	205
PM350:	CAC ATT TTC CCA GAT CAG CA GGT GGC AAA GAA CTT ATT GAG G	$(CA)_8(GA)_{25}$	230
PM305:	GCG CTG GAA CAC AGT AAG AG GGC AGA AAG GAA AGT TGC AG	$(GAGTAG)_3$	171
PM375:	CGG CAA CAG TTT TGA TGG TT GAA AAA TAT GCC GCC GTT G	$(CT)_{10}$	102
PM377:	ACG CTC ACA TGT TTG CTT TG GCT CGA TTT GAT TTG GGT GA	$(GA)_{21}$	164
PM384:	GGC GTG CCA ATA GAG GTT TA TGA AAA CCA ACA AGT TTA GTC TCT CT	$(GA)_{27}$	105
PM402:	CCG CCC TAA AAA CTG TAT TCG CCT AAG AGT ACA CGC GAC GA	$(CT)_{11}$	250
PM434:	CGA GGG CAA CTT GGT AAA AA ATT CGC CCT TAG ACT GCG TA	$(CT)_{14}$	204
PMc99:	GCA TAA GCA GTT TCC AAC GA TGT TGC CTT CAC CTT GAC AG	$(AAGCCT)_4$	232
PMc297:	ATG CAC CTG CAA GTG AAG AG TCA AGG ATG CAG CAA GAC AC	$(AAT)_4CAT(AAT)_2$	220
PMc348:	TGC CTG TAA GTG TGG ACC AA ACT CCA AAA CGG GGA GTG TT	$(AAT)_6$	223
PMc468:	TCA AGC CAT AAT ATG TTC CAC A AAA ACA ACC CAA GCA CCT CT	$(AT)_{33}$	247
PMc478:	GTC GTG CAG GTC AAA GTG C TTA AGA TGG GTG CCT GCA AT	$(CAT)_6(AT)_2$	234
PMc588:	CCA TTT TGG ACC CCT CAA AT TGA GCA ATA GTG ACC TTG CAT T	$(AT)_{17}$	183
PMc595:	GCA GTA ACA GTA GCA GCA GGT TT CTG AGC TCA TGT AAA GGG AAG AA	$(TCA)_4$, $(CT)_9$	349
PMc660:	ACA GCA ACA AAG CGT GTG AG GCC TCC CTA TGT GCC AGT AA	$(GCA)_6$, $(CAA)_3$	214
PMc37785:	GCG TGG TTC CAT GGA TTT T CGT CAC AAT CAT AAA TCA TCA CAG	$(CT)_{12}$	231

[a]The SSRs with name PMc are EST-derived SSRs.

from genomic SSRs and three from EST-derived SSRs. There were two types of variety-specific markers. One type of markers amplified a unique single band only in certain botanical varieties. Primer PMc297 generated a band in varieties *hypogaea, hirsuta, peruviana* and *aequatoriana*, but not in varieties *fastigiata* and *vulgaris* (Figure 1b). In contrast, primer PM188 amplified a band only in varieties *fastigiata* and *vulgaris*, not in other four varieties. Primer PMc478 amplified a band in all four botanical varieties *hypogaea, hirsuta,*

fastigiata, and *vulgaris*, while botanical varieties *peruviana* and *aequatoriana* were missing this band, and Primer PM15 produced a bright and distinct band only in botanical variety *hypogaea* and one accession in variety *hirsuta*. Another type of variety-specific markers generated different alleles in different varieties. Primer PM375 amplified a different size of band in varieties *fastigiata*, and *vulgaris* from those in other four varieties (Figure 1c), so did PM32, PM 402 and PMc588. All alleles at these loci were consistently observed within specific botanical varieties, except the occurrence of allele in PM15 locus specific to botanical varieties *hypogaea* and *hirsuta* was infrequent.

The results of the BLAST search showed that the sequence of PMc297 marker was homologous to metallothionein-like proteins in soybean with 68–80% identities. The sequence of PMc478 marker shared homology with a hypothetical protein in *Arabidopsis thaliana*, and the sequence of PMc588 marker homologous to nodulin protein in *Medicago truncalula*. The sequences of botanical variety-specific markers from genomic libraries were unique and did not share homology with any known sequences in the database.

Multiplex PCR showed that the combination of only two primers, PM15 and PMc478, could successfully amplify specific bands under multiplex PCR conditions. Forty-eight cultivated accessions could be easily differentiated into three groups, varieties *hypogaea* and *hirsuta*, *fastigiata* and *vulgaris*, *peruviana* and *aequatoriana*, using one PCR reaction (data not shown).

Discussion

The results above suggested that there are three groups of similar genome, in which botanical variety *fastigiata* with *vulgaris*, variety *peruviana* with *aequatoriana,* and variety *hypogaea* with *hirsuta*. Since each group has the same pattern of bands using each of variety-specific markers, these specific markers are actually related to these three groups. It would be very useful if molecular marker were associated with only one botanical variety. Molecular markers further differentiating botanical varieties within group should be identified as more peanut SSR markers are developed.

Microsatellites have been found in a number of promoter regions and repeat number variation of microsatellite was detected in coding regions (Goldstein & Schlotterer, 1999). The EST-derived SSR markers used in this study were related to some known genes and thus they are useful for determining the role of microsatellite on gene expression. In this study, a null allele was amplified in botanical varieties *fastigiata* and *vulgaris* when using PMc297 (Figure 1b). The primer pair was designed based on the sequence of an EST (GenBank accession number CD038297). One primer annealing site was located within coding region which was homologous to metallethionein-like protein in soybean. PMc297 failed to amplify specific bands in varieties *fastigiata* and *vulgaris* indicating the primer binding sites may have been mutated or altered. Similar results were derived from using PMc478, where both varieties *peruviana* and *aequatoriana* lacked the specific bands related to a hypothetical protein. For the nodulin protein homologue generated by PMc588, varieties *fastigiata* and *vulgaris* had a different allele containing more repeat units compared to that of the other four varieties. At this moment, we do not know if this repeat-number variation affects the function of nodular gene. Further studies should determine if this marker is related to the differential expression of nodulin genes.

For most cases, band patterns were not altered when multiplexes were used, such as the multiplex of primers PM15 and PMc478. However, the mixtures of some primers, such as PM15, PMc297, and PMc478, amplified unspecific products, and yielded different genotyping patterns from the original single primer reaction. The result might be due to 'primer–primer' interactions. Nevertheless, the multiplex PCR of botanical variety-specific markers would be useful in the germplasm management by saving money and time, as it requires running only one reaction for several primers, instead of separate reactions for each primer. Multiplex PCR of these markers would provide an efficient tool for genotyping applications in peanut.

Acknowledgments

The authors thank supports from the George Washington Carver Agricultural Experiment Station (contribution No. 334). This work has been funded by a grant from USDA/CSREES/CBG (No. 00-38814-9541). Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply approval to the exclusion of other products that may be suitable.

References

Goldstein, D.B. & C. Schlotterer, 1999. Evolution and Applications. In: Microsatellites. Oxford University Press, New York.

Guo, B.Z., M. Luo & C.C. Holbrook, 2003. Nucleotide sequences of 1345 ESTs (expressed sequence tags) (accession no. CD037499 to CD 038843) from peanut (*Arachis hypogaea* L.) GenBank, National Center for Biotechnology Information. http://www.ncbi.nlm.nih.gov/dbEST.

He, G.H. & C.S. Prakash, 2001. Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. Genet Resour Crop Evol 48: 347–352.

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

Henegariu, O., N.A. Heerema, S.R. Dlouhy, G.H. Vance & P.H. Vogt, 1997. Multiplex PCR: Critical parameters and step-by-step protocol. BioTechniques 23: 504–511.

Krapovickas, A. & W.C. Gregory, 1994. Taxonomia del genero *Arachis* (Leguminosae). Bonplandia 8: 1–186.

Luikart, G., M.P. Biju-Duval, O. Ertugrul, Y. Zagdsuren, C. Maudet & P. Taberlet, 1999. Power of 22 microsatellite markers in fluorescent multiplexes for parentage testing in goats (*Capra hircus*). Anim Genet 30: 431–438.

Mellersh, C. & J. Sampson, 1993. Simplifying detection of microsatellite length polymorphisms. BioTechniques 15: 582–584.

Mitchell, S.E., S. Kresovich, C.A. Jester, C.J. Hernandez & A.K. Szewe-McFadden, 1997. Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci 37: 617–624.

Narvel, J.M., W.C. Chu, W.R. Fehr, P.B. Cregan & R.C. Shoemaker, 2000. Development of multiplex sets of simple sequence repeat DNA markers covering the soybean genome. Mol Breed 6: 175–183.

Scheef, E.A., M.D. Casler & G. Kung, 2003. Development of species-specific SCAR markers in bentgrass. Crop Sci 43: 345–349.

Stalker, H.T. & C.E. Simpson, 1995. Germplasm resources in *Arachis*. In H.E. Pattee & H.T. Stalker (Eds.), Advances in peanut science, Chapter 2, pp. 14–53. APRES, Inc. Stillwater, OK, USA.

Tang, S., V.K. Kishore & S.J. Knapp, 2003. PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor Appl Genet 107: 6–19.

Williams, D.E., 1989. Exploration of Amazonian Bolivia yields rare peanut landraces. Diversity 5(4): 12–13.