Available online at www.sciencedirect.com
ScienceDirect
Plant Science 174 (2008) 51–60
PLANT SCIENCE
www.elsevier.com/locate/plantsci

In silico development of simple sequence repeat markers within the aeschynomenoid/dalbergoid and genistoid clades of the Leguminosae family and their transferability to *Arachis hypogaea*, groundnut

E.S. Mace *, R.K. Varshney, V. Mahalakshmi, K. Seetha, A. Gafoor, Y. Leeladevi, J.H. Crouch [1]

International Crop Research Institute for the Semi Arid Tropics (ICRISAT), Patancheru, Andhra Pradesh 502324, India

Received 13 June 2007; received in revised form 19 September 2007; accepted 26 September 2007
Available online 6 October 2007

Abstract

Cultivated groundnut (*Arachis hypogaea* L.) is an agronomically and economically important oilseed crop grown extensively throughout the semi-arid tropics of Asia, Africa and Latin America. The genetic base of the cultivated groundnut is very narrow as a result of the genetic bottleneck associated with recent polyploidization which makes it critical to determine the levels of genetic diversity within available germplasm collections prior to breeding. In groundnut, the use of SSRs for diversity assessment may offer the potential to reveal genetic variation within the genome of the cultivated species. An alternative bioinformatics, or *in silico* approach, to identifying SSRs suitable for application in cultivated groundnut is presented, as a low-cost alternative to wet lab SSR identification. All available nucleotide sequences from species within the aeschynomenoid/dalbergoid and genistoid clades of the Leguminosae family were searched for SSR motifs and primers designed from 109 unique SSRs. Representative accessions from six genera within the aeschynomenoid/dalbergoid and genistoid clades were selected for assessing SSR-transferability rates. In total, 60% of the total cross-genera transfer testing reactions gave prominent and reproducible amplicons, with 51 of the 109 SSRs amplifying in *A. hypogaea*. These 51 SSRs were further tested against 27 diverse *Arachis* accessions and 18 revealed polymorphism, demonstrating that the *in silico* approach to SSR identification and development is a valid strategy in lesser-studied crops.

Keywords: Groundnut; SSR; Cross-species transferability; Cross-genera transferability; Molecular marker; Polymorphism

1. Introduction

Groundnut (*Arachis hypogaea* L.), also known as peanut, is one of the most important oilseed crops grown as a major source of vegetable oil and protein, both for human consumption and as a fodder crop. Groundnut is extensively cultivated in 107 countries of the world on 25.2 m ha with an annual production of 36.5 mt [1]. The genetic base of the cultivated groundnut is very narrow, in contrast to the polymorphism observed in wild *Arachis* species [2,3] as a result of the genetic bottleneck associated with recent polyploidization. The cultivated allotetraploid ($2n = 4x = 40$) *A. hypogaea*, unlike other natural polyploids, is believed to have originated recently from a single hybridisation event [4]. The genetic impoverishment of the cultivated groundnut genome, compounded by potentially narrowing selection pressures from traditional breeding approaches, makes it critical to determine the levels of genetic diversity within available germplasm collections prior to breeding. The use of molecular markers has become widely accepted as a valuable tool for plant breeding programs as well as for diversity, evolutionary and conservation studies in many species [5]. Simple sequence repeats (SSRs), also known as microsatellites, are a class of molecular markers based on tandem repeats of short (2–6 bp) DNA sequence [6], which are ubiquitously distributed throughout eukaryotic genomes. These repeat sequences are found to be abundant in plant genomes and are frequently highly polymorphic, even among closely related cultivars, due to mutations causing variation in the number of

* Corresponding author. Current address: Department of Primary Industries & Fisheries, Hermitage Research Station, 604 Yangan Road, Warwick, QLD 4370, Australia. Tel.: +61 7 4660 3629; fax: +61 7 4660 3600.
E-mail address: emma.mace@dpi.qld.gov.au (E.S. Mace).
[1] Current address: CIMMYT, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico.

0168-9452/$ – see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2007.09.014

repeating units [7]. Different alleles can be detected at a locus by the polymerase chain reaction (PCR), using conserved DNA sequences flanking the SSR as primers. SSRs are reported to be more variable than RFLPs, RAPDs and AFLPs, and have been widely adopted for genetic analysis in crop plants such as soybean [8] and rice [9] and are now becoming the standard DNA markers for plant genome analysis and are being used in marker-assisted selection in many crop species [7,10]. In groundnut, the use of SSRs for diversity assessment may therefore offer the potential to reveal the genetic variation within the genome of the cultivated species. With this objective, SSR markers have been characterized within groundnut by some research groups [11–17], recently, through the construction and screening of genomic libraries. However, the development of SSR markers through laboratory-based screening of genomic libraries is highly time consuming and expensive. An alternative in well-studied species is to use bioinformatics, i.e. an *in silico* approach, to screen databases for sequences that contain microsatelllite repeats [10]. The *in silico* approach has previously relied upon the availability of abundant sequence data for the species in question and so lesser-studied crops, such as groundnut, are disadvantaged. However, even for ''orphan'' crops, such as groundnut, the *in silico* approach offers some potential for low-cost development for limited numbers of markers, through the screening of related and allied crops, in addition to exploiting the nucleotide sequences available for *A. hypogaea* (currently 60701 nucleotide sequences available in GenBank). Groundnut has the advantage of belonging to the Fabaceae or Leguminosae family; the third largest flowering plant family with over 700 genera and 20,000 species and which are second only to cereal crops in agricultural importance based on area harvested and total production. The legumes are highly diverse and can be divided into three subfamilies; Mimosoideae, Caesalpinioideae and Papilionoideae [18]. Of these the Papilionoideae subfamily contains nearly all economically important crop legumes, including soybean (*Glycine max*), groundnut (*A. hypogaea*), mungbean (*Vigna radiata*), chickpea (*Cicer arietinum*), lentil (*Lens culinaris*), common bean (*Phaseolus vulgaris*), pea (*Pisum sativum*) and alfalfa (*Medicago sativa*). However, of all these important legumes only groundnut is separate from the most populus Papilionoid clades, with the consequence that of all the economically important legume species, groundnut is the most isolated in the aeschynomenoid/dalbergoid clade (Fig. 1) and least able to utilise the abundant genomic resources being accumulated for other legume species, and in particular *Medicago truncatula*, the model species for comparative and functional legume genomics. However, even with the more limited genomic resources available for *Arachis* and related genera, the *in silico* approach potentially offers many advantages, particularly in light of recent reports on cross-species and cross-genera amplification of SSR loci [19,20]. The aims of the current study were to assess the practicality and usefulness of cross-genera SSR transferability within the Dalbergoid clade of the Leguminosae with the aim of generating additional SSR markers for application in *A. hypogaea*, and to apply such SSR markers to a diverse set of *Arachis* germplasm to assess the level of polymorphism.

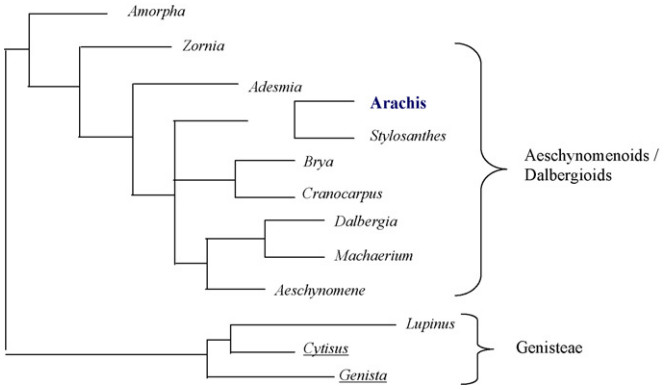

Fig. 1. Expansion of Aeschynomeneae, Dalbergieae and Genisteae tribes, detailing genera closely related to *Arachis*. Modified from Refs. [40] and [41]. All genera listed, except those underlined, have been used in a bioinformatics approach to generate markers for groundnut.

2. Materials and methods

2.1. Study species and systematics

The aeschynomenoid/dalbergoid clade is located within the Papilionoideae subfamily and is most closely related to the genistoid clade, which includes the genus *Lupinus*. The list of genera/species used to screen databases containing sequence data are detailed in Fig. 1 and Table 1. Representative accessions from six genera within the aeschynomenoid/dalbergoid and genistoid clades were selected for assessing SSR-transferability rates as detailed in Table 2. To test polymorphic markers, 27 *Arachis* accessions were selected from the ICRISAT groundnut germplasm collection for characterisation (Table 3).

2.2. DNA extraction

Total genomic DNA was isolated from newly expanded leaves according to a CTAB-based procedure modified from Refs. [21] and [22]. The quality of DNA was determined spectrophotometrically at 260/280 nm, and DNA concentrations were determined electrophoretically using known amounts of λ DNA standards.

2.3. In silico SSR development

All nucleotide sequences related to the species listed in Table 1 from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov) were downloaded in FASTA format and analysed for repeat patterns using the tandem repeat finder program at http://c3.biomath.mssm.edu/trf.html, defining an SSR with a minimum of 10 bp length for all repeats (Di, Tri, Tetra and Penta). The local database containing the entire sequences in FASTA format, the repeat motif and the potential primer sequences was created in a relational database (MS Access). The resultant database of the

Table 1
Details of species used in the *in silico* approach to SSR development

Tribe	Genus	Species	Common name
Adesmieae	*Adesmia*	*lanata*	
		boroniodes	
		exilis	
Dalbergieae	*Dalbergia*	*cearensis*	Kingwood
		cochinchinensis	Trac
		decipularis	Sebastiao de Arruda
		frutescens	Tulipwood
		granadillo	Granadillo
		latifolia	Indian rosewood
		melanoxylon	African blackwood
		nigra	Brazilian rosewood
		oliveri	Burma tulipwood
		retusa	Cocobolo
		sissoo	Sissoo
		spruceana	Amazon rosewood
		stevensonii	Honduras rosewood
Dalbergieae	*Machaerium*		Capote
Desmodieae-Bryinae	*Brya*	*ebenus*	Jamaican rain tree
Desmodieae-Bryinae	*Cranocarpus*	*martii*	
Amorpheae	*Amorpha*	*fruicosa*	Desert false indigo, indigobush
Amorpheae		*canescens*	Leadplant
Genisteae	*Lupinus*	*alpus*	Lupin
		luteus	
		angustifolius	
Aeschynomeneae	*Aeschynomene*	*rudis*	Rough jointvetch
		indica	Indian jointvetch
		virginica	Northern jointvetch
		americana	Shyleaf
Aeschynomeneae	*Zornia*		Zornia, viperina
Aeschynomeneae-Stylosanthinae	*Stylosanthes*	*macrocephala*	Stylo
		capitata	
		macrocarpa	
		guianensis	
		humilis	
		scabra	
		mexicana	
		fruticosa	
		viscosa	
		calcicola	
		angustifolius	
Aeschynomeneae-Stylosanthinae	*Chaetocalyx*	*brasiliensis*	
		nigricans	
Aeschynomeneae-Stylosanthinae	*Arthrocarpum*		
Aeschynomeneae-Stylosanthinae	*Pachecoa*		
Aeschynomeneae-Stylosanthinae	*Fiebrigiella*		
Aeschynomeneae-Stylosanthinae	*Nissolia*		

Table 2
Details of species used in SSR-transferability study

Genus	Species	ICG no.	Source
Arachis	*hypogaea*	ICGS44	ICRISAT
Stylosanthes	*guianensis*	EC513492	CIAT
Dalbergia	*sissoo*	EC512183	ILRI
Lupinus	*albus*	EC512184	ILRI
Amorpha	*fruticosa*	EC513492	ILRI
Chaetocalyx	*brasiliensis*	EC513500	CIAT

repeat motifs was analysed to classify the patterns, their occurrence, duplication and abundance. Primers were designed for unique SSRs using primer3, with following parameters defined; product size: 400; primer temperature—min: 59 °C, opt: 60 °C, max: 61 °C.

2.4. SSR-PCR amplification and detection

PCR reactions were performed in 20 μl volumes using PTC-100™ Programmable Thermal Controller (MJ Research, Inc).

Table 3
List of 27 *Arachis* accessions used to assess informativeness of SSRs

	Section	Species	Genome	ICG no.	Code
1	*Arachis*	*hypogaea*	AB	ICGV 99001	hypogaea 99001
2	*Arachis*	*hypogaea*	AB	ICGV 99003	hypogaea 99003
3	*Arachis*	*hypogaea*	AB	ICGV 99004	hypogaea 99004
4	*Arachis*	*hypogaea*	AB	ICGV 99005	hypogaea 99005
5	*Arachis*	*hypogaea*	AB	ICG 6284	hypogaea 6284
6	*Arachis*	*hypogaea*	AB	ICG 7878	hypogaea 7878
7	*Arachis*	*hypogaea*	AB	ICG 405	hypogaea 405
8	*Arachis*	*hypogaea*	AB	ICG 1705	hypogaea 1705
9	*Arachis*	*hypogaea*	AB	ICGV 15222-1	hypogaea 15222-1
10	*Arachis*	*hypogaea*	AB	ICGV 152222-2	hypogaea 15222-2
11	*Arachis*	*hypogaea*	AB	CHICO	Hypogaea_ Chico
12	*Arachis*	*hypogaea*	AB	TMV 2	hypogaea _TMV2
13	*Arachis*	*batizocoix* (*cardenasii* x *diogoi*)	AB	EC 468631	TxAG-6
14	*Arachis*	*batizocoix* (*cardenasii* x *diogoi*)	AB	PI 565288	TxAG-7
15	*Arachis*	*hoehnei*	B	8190	hoehnei
16	*Arachis*	*glandulifera*	D	15172	glandulifera
17	*Arachis*	*hypogaea*	AB	ICGS44	hypogaea 44
18	*Arachis*	*monticola*	AB	13177	monticola
19	*Arachis*	*duranensis*	A	8956	duranensis
20	*Procumbentes*	*chiquitana*	P	11560	chiquitana
21	*Procumbentes*	*kretschmeri*	P	8191	kretschmeri
22	*Erectoides*	*major*	E	13262	major
23	*Heteranthae*	*sylvestris*	H	14858	sylvestris
24	*Heteranthae*	*dardani*	H	14923	dardani
25	*Caulorhizae*	*pintoi*	C	14855	pintoi
26	*Rhizomatosae*	*glabrata*	RR	8176	glabrata
27	*Rhizomatosae*	*villosulicarpa*	E	8142	villosulicarpa

The reaction mixtures contained 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 5–15 ng genomic DNA, 10–30 pmol of each primer, 2–4 mM $MgCl_2$, 300–400 μM of each dNTP, and 0.8–1.2 units of *Taq* DNA polymerase (Amersham). The temperature regime consisted of an initial denaturation step of DNA at 94 °C for 2 min, followed by 10 cycles: 94 °C for 45 s, 65 °C for 1 min, and 72 °C for 1 min 30 s, dropping 1 °C each cycle, followed by 25 cycles: 94 °C for 45 s, 55 °C for 1 min, and 72 °C for 1 min 30 s. After the final cycles, samples were incubated for 10 min to ensure complete extension.

Each SSR was initially screened for amplification of a specific product from genomic DNA of the six genera detailed in Table 2 using the calculated annealing temperature. The PCR products were separated on 2% agarose gels followed by ethidium bromide staining and visualised by UV light. SSR loci that gave amplification products in groundnut were used in a second round of PCR on 27 diverse groundnut genotypes, representing each section within *Arachis* (Table 3). The PCR products were separated on 6% non-denaturing polyacrylamide gels, and amplification products were revealed using the silver staining procedure based on a histologically derived procedure using ammoniacal solutions of silver, modified from [23].

2.5. Data analysis

Bands were scored as present (1) or absent (0). Estimates of similarity were based on three different measures—(1) Nei and Li's definition of similarity [24]: $S_{ij} = 2a/(2a + b + c)$, where S_{ij} is the similarity between two individuals, i and j, a the number of bands present in both i and j, b the number of band present in number of bands present i and absent in j, and c is the number of band absent i and present in j. This is also known as the Dice coefficient (1945). (2) Jaccard's coefficient [25]: when $S_{ij} = a/a + b + c$. (3) The simple matching (SM) coefficient [26]: $S_{ij} = a + d/a + b + c + d$, where d is the number of bands absent from both i and j using the NTSYS 2.1 software, version 2.1 [27]. Multidimensional scaling (MDS) [28] was then performed to see whether the observed molecular variation indicated any evidence of clustering among accessions. Following this, cluster analysis was performed using UPGMA (unweighted pairgroup method) [26] and dendrograms created with the TREE program of NTSYS, and the goodness of fit of the clustering to the data was calculated using the COPH and MXCOMP program. Additionally, the polymorphism information content (PIC) of each SSR was determined as described by [29]

$$\mathrm{PIC} = 1 - \sum P_i^2$$

where P_i is the frequency of the ith allele in the examined genotypes. PIC values range from 0 (monomorphic) to 1 (highly discriminative).

3. Results

3.1. In silico identification of SSRs

Of the 15 genera included in the *in silico* SSR search, six genera were found that contained an SSR within available sequences at NCBI (Fig. 2); *Adesmia*, *Amorpha*, *Dalbergia*,

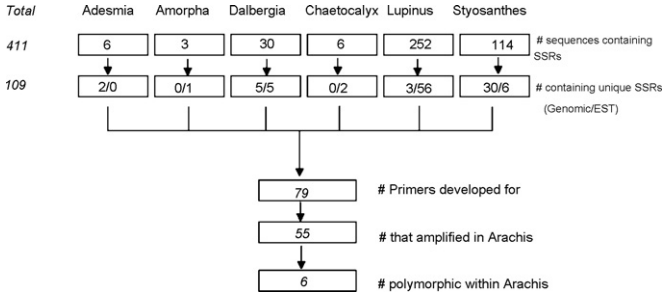

Fig. 2. Diagram showing the generation of 79 unique SSRs for *Arachis*.

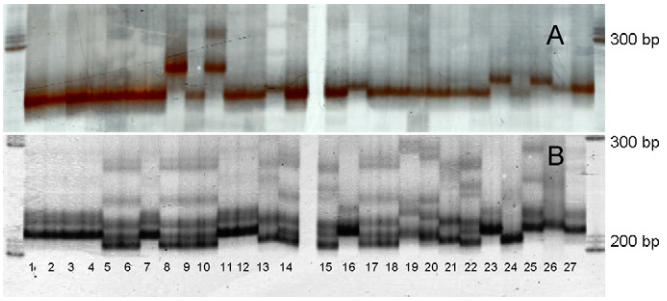

Fig. 4. Level of polymorphism revealed among 27 diverse *Arachis* accessions, as detailed in Table 3, for (A): 02dal_matk and (B): 76_Stylo_IGS.

Chaetocalyx, *Lupinus* and *Stylosanthes*. In total 411 sequences containing an SSR were found across these six genera; however only 109 unique SSRs were identified. Of these 109 unique SSRs the majority (63%) overall were EST-based (expressed sequence tags), however in contrast to the other genera, the majority of the *Stylosanthes* SSRs identified were genomic-DNA based. Primers were designed for 79 of the 109 unique SSRs, 82% of which were EST-derived; *Amorpha*: 1 EST-derived SSR; *Adesmia*: 0; *Chaetocalyx*: 1; *Dalergia*: 2; *Lupinus*: 56; *Stylosanthes*: 6 (Table 4).

3.2. Inter-generic/-specific SSR transferability

In total, prominent and reproducible amplicons were generated in 286 reactions, representing 60% (286/475) of the total cross-genera transfer testing reactions. In comparison with the size of the original and positive control amplicons, the cross-genera amplicons varied greatly in size (100–1500 bp). As an example, the primer pair 05_Lup_TCP1 from *Lupinus* produced an amplicon of size 200 bp in *Arachis*, *Dalbergia* and *Lupinus* but amplicons over 500 bp were observed in *Stylosanthes* (600 bp), *Chaetocalyx* (500 bp) and *Amorpha* (700 bp).

The degree of success of SSR-transferability within and across the aeschynomenoid/dalbergoid and genistoid clades was very variable. Due to the differences in numbers of SSRs developed for each genera, and in particular the much larger sample size for *Lupinus* compared to the small sample sizes for other genera, it is difficult to draw firm conclusions, however

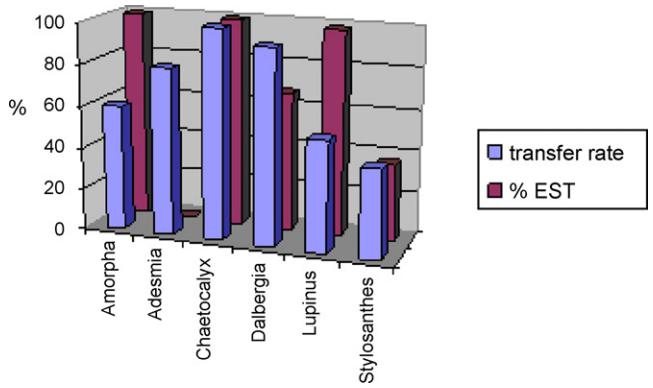

Fig. 3. Relationship between percentage of EST-derived vs. genomic SSRs and transfer rate across genera.

the following observations were made. On average (Table 5), the primers from *Dalbergia* SSRs generated almost twice as many amplicons (93%) as the primers from *Lupinus* SSRs. However, the *Lupinus* SSRs, which were all EST-derived, transferred more easily to other genera than the largely genomic SSRs from *Stylosanthes* (10/16 genomic SSRs; 62.5%). Fig. 3 details the relationship between EST-derived versus genomic SSRs and their transferability rate across genera. In every case except *Adesmia*, the transferability rate increased with the percentage of SSRs that were EST-derived. Cross-genus amplification may also reflect the relatedness between genera. *Dalbergia*, for example, is more closely related to *Arachis* than *Adesmia* and *Lupinus* (Fig. 1) and in both cases the *Dalbergia* SSRs transfer more readily to *Arachis* than either the *Adesmia* or *Lupinus* SSRs. The success of SSR transfer is also direction-dependent; for example although the SSR derived from *Chaetocalyx* amplified across all six genera, the amplification success of SSRs derived from the other genera and applied to *Chaetocalyx* was highly variable (100% success rate with *Dalbergia* derived SSRs; 16% success rate with *Lupinus* derived SSRs and 0% success rate with *Adesmia* derived SSRs).

3.3. SSR analysis in Arachis

Of the 79 unique SSRs amplified across all 6 genera, 51 were found to be amplifiable in groundnut. These 51 SSRs were then amplified in 27 diverse Arachis accessions (Table 3), representing six sections within the *Arachis* genus. Of the 51 SSRs amplifiable in *Arachis*, 18 (35%) were found to be polymorphic, generating 71 alleles. Of these 18, 7 were genomic SSRs from *Stylosanthes* and 11 were EST-derived SSRs from *Dalbergia*, *Lupinus*, *stylosanthes and Amorpha* (Table 6). Fig. 4 illustrates the level of polymorphism observed among the 27 diverse *Arachis* accessions as revealed by two of the 18 primers on polyacrylamide gels. The majority of amplification products correspond to a single locus since the majority of markers are derived from genes, and hence are gene specific. Nevertheless, some markers may amplify two homeoloci and when are of different sizes, they are observed as two fragments per markers (see "B" in Fig. 4). On average, each locus revealed approximately four alleles with three loci detecting six or more alleles. The PIC values of the 18 polymorphic SSR loci reveal a high level of polymorphism (Table 6) with 12 of the 18 SSR loci having a PIC value of $\geq$0.5.

Table 4
List of 79 unique SSRs, repeat unit, forward and reverse primer sequence, Tm and amplification in *Arachis hypogaea*

Primer_ID	Genus	Acc #	Detail #	Repeat	Left primer	Right primer	Size (bp)	Tm (L)	TM (R)	Arachis-amp
01_Chaet_matK	Chaetocalyx	9,930,155	AF270866	AAATA(2)	AGAGAGTCCGTTGATGGGTTT	AAGTTCTGTTGGCAAGATCCA	236	59.99	59.73	y
02_Dal_matK	Dalbergia	5,817,666	AF142696	AT(5)	CGAGTGGAGAGAGTCCGTTT	AGGAAGTTCTGTTGGCGAGA	238	59.45	59.99	y
03_Dal_trnL[a]	Dalbergia	6,984,002	AF208923	AT(5)	AAGGAATCGTTCCATCGAAAT	TGGAGTGAATGATTTGATCAGTG	209	59.79	59.99	n
04_Dal_PHYA	Dalbergia	3,176,487	U78850	ATG(6)	GATTGTCGTGCAAAGCATGT	TGGCTGAACAACATCAGAGC	201	59.73	59.99	y
05_Lup_TCP1	Lupinus	2E+07	LAL426419	ATC(8)	AGGGTCTGAGTCTGGATCCTC	TCGATTGTGAAAGGTGGTGA	199	59.68	60.09	y
06_Lup_PT1	Lupinus	1.4E+07	AF305623	ACA(9)	GCCAAATGGAAAATCTTTGG	GGGAAGGGAAAGGTTTTGTT	200	59.39	59.3	n
07_Lup_Lb1	Lupinus	2,921,625	LLU50083	AG(5)	TGAAAAGGTTCTCCGACGAT	ATGGTCCCATTTGTCAGAGC	189	59.67	59.93	y
08_Lup_aclb	Lupinus	1.6E+07	LAL344108	TTTG(4)	TTTTGGCAAGTGCCTTCTTT	GGAAGACATTATTTTATGGAACACAA	237	59.86	59.6	n
09_Lup_CycB	Lupinus	4,884,723	AF126105	AAG(4)	ATGATATAGCGGTGGCGACT	GGTTCAGCAGCAACATGAGA	199	59.58	59.99	y
10_Lup_nod	Lupinus	437,390	LUPNOD45	AT(16)	CCGGTACAGGAAGTTGTTGG	ATTATTGATCCGAGACATTGTTTTT	207	60.4	59.24	n
11_Lup_ACS5	Lupinus	6,650,981	AF119414	AT(10)	TTTTGAGTGCCAAGTTGACG	TGGCCAAGTACTTTTCTGCAC	230	59.88	60.3	y
12_Lup_ACS2	Lupinus	6,650,977	AF119412	TA(7)	GGGTGGAAAGCCTATGATGA	TACCCCATGTAGCAGCCTCT	265	59.89	59.72	y
13_Lup_ACS1	Lupinus	6,650,975	AF119411	AT(16)	TTTGTCCTGGTCAGCATTGA	CATTAGGGGCCAAGACAAAG	167	60.24	59.56	y
14_Lup_ACS4	Lupinus	6,650,973	AF119410	TA(9)	TAATCGGATTCATTCGCACA	GGCCTTGTTGATGGAGCA	146	60.04	60.78	y
15_Lup_LIN01	Lupinus	1.3E+07	BG154158	AAATC(4)	AGGCTTTGTTTGCCGACTTA	AGACCCTCTCCTTGGTTGCT	190	59.88	60.25	y
16_Lup_LIN01	Lupinus	1.3E+07	BG154126	TA(6)	TCCCCACAAAATCCATTCAT	GATGAGTTGGGTGGAGAACAA	184	59.99	59.96	y
17_Lup_LIN01	Lupinus	1.3E+07	BG154103	TC(17)	ACACACCTCACCCATTTCCT	TTGGAGCCAAATGATGAAAA	234	59.28	59.09	n
18_Lup_LIN01	Lupinus	1.3E+07	BG154070	AGA(4)	CTCGACCACTGGATGAGACA	TGGTGATGACGTGGAACAGT	193	59.82	60	y
19_Lup_LIN01	Lupinus	1.3E+07	BG154058	AGA(4)	GAACGCGGAGAAGGAGATAA	TGTTTGGGACTCTGCCACTA	200	59.41	59.29	y
20_Lup_LIN01	Lupinus	1.3E+07	BG154041	AGA(4)	GAAAGAGCAGTTTTATCAGAAGAAGAA	TCAGATGGCTCAAACAGTGG	181	59.98	59.83	n
21_Lup_LIN01	Lupinus	1.3E+07	BG154037	TA(7)	CCCTCAATTTTTGATCCCAAT	CCCACACTCCAAAAACCATC	244	60.01	60.21	y
22_Lup_LIN01	Lupinus	1.3E+07	BG153980	CAACA(4)	TGTAACGCTGAATTGGCAAC	CTTTTTGCCAGAGACCAAGG	202	59.74	59.85	n
23_Lup_LIN01	Lupinus	1.3E+07	BG153952	CTT(4)	TGCCATCAATTTTCGCTACA	CTCCACCATGACCCAAAGAT	193	60.22	59.78	y
24_Lup_LIN01	Lupinus	1.3E+07	BG153948	ATC(4)	GAATTCCACCGTAACCTCCA	AGTAACATGCAAAGGCGTTGT	190	59.79	59.69	y
25_Lup_LIN01	Lupinus	1.3E+07	BG153938	CAA(6)	GCCAATAACCACAAACACCAC	GGAAGTTGTTGCTGCTGTTG	236	60.15	59.49	y
26_Lup_LIN01	Lupinus	1.3E+07	BG153931	GCC(4)	TTCAAGGGAGCCAGAATCAC	TGCACCACCAGTATTCCTGA	207	60.2	60.11	y
27_Lup_LIN01	Lupinus	1.3E+07	BG153924	TGC(4)	CATCTGCTCCACATGCTAGG	TGAGCAACATGTCCATAGCC	193	59.42	59.68	n
28_Lup_LIN01	Lupinus	1.3E+07	BG153901	TTC(6)	CCGAACTCCTCCAACTACCA	CAAGAGGGGTGCCATAAGAA	242	60.1	60.07	y
29_Lup_LIN01	Lupinus	1.3E+07	BG153889	TCT(6)	TCTTCTCTTCGCTGTTACTTCTTCT	GACGTCGACGCTTGTTATTG	192	59.43	59.35	y
30_Lup_LIN01	Lupinus	1.3E+07	BG149146	TTC(5)	GCGGCGCTACTTCATGTTAT	CCCTTGTGGGGTTTTTGAA	205	60.26	60.71	y
31_Lup_LIN01	Lupinus	1.3E+07	BG149134	CTT(6)	CTTTCACACAACCGGACCTT	CGAACATTTCTGCCCGTATT	202	60.01	59.96	n
32_Lup_LIN01	Lupinus	1.3E+07	BG149152	AAG(5)	GGTCTTGAACCACCAACCTG	ATGAACGCCACTCTTGGTTC	204	60.4	60.12	y
33_Lup_cgg	Lupinus	1.1E+07	LAL297490	TA(7)	TCATTATTCCCCTCCAAACG	TCCACCAAAAATAAAATGAATCTG	199	59.76	59.29	n
34_Lup_app	Lupinus	8,918,672	AB037887	TA(5)	TTGTGTTCGCTGGTCATGTT	TGTCCAAAGCTTGCCTCTCT	211	60.16	60.13	y
35_Lup_ribo	Lupinus	1,143,506	LLP0	TGC(4)	CCGCACCACATGTGTTTATC	CTCTTTCTCCGCTTTTGCAG	197	59.85	60.26	y
36_Lup_CycB1d	Lupinus	3,253,102	LLU44857	GA(5)	AAATTCGACCGTTGAGGTTG	GCTTGTTGTTGAAGCCGAAC	200	59.97	60.82	y
37_Lup_CycB1a	Lupinus	3,253,100	LLU44857	AAG(4)	ATGATATAGCGGTGGCGACT	GGTTCAGCAGCAACATGAGA	199	59.58	59.99	y
38_Lup_Ypr10	Lupinus	2,183,276	AF002278	TA(14)	TGAAGGAAATGGAGGACCAG	TTGAACATTAACCCATGTAGAAACA	383	60.04	59.71	y
39_Lup_At	Lupinus	2,780,193	LAAJ3197	CTT(4)	CCCAATCCACCATTCTCAAT	CGGAAAGCAGCATCGTAACT	194	59.61	60.41	y
40_Lup_LIPRP2	Lupinus	1,754,988	LLU47661	TA(10)	AGGAATTGGTTATATCCCCTTTG	AACCACATCTTTCGCCTTAAAT	217	59.5	59.06	n
41_Lup_Albus[a]	Lupinus	13,072	MILAPLDA	AAATC(2)	GCACAACCCACAACACACC	TTTGTGAAGTCGTGGCCTTT	188	60.92	60.67	y
42_Lup_GS	Lupinus	454,311	LLNGS1G	AT(20)	GGTAGGTGTTGTGGGAAATGTT	TCCATCCATCTTGTGGAATTG	198	60.02	60.71	n
43_Lup_BT	Lupinus	402,635	LALB1	TTTC(4)	GCTTCCGCACACTTAAGCTC	ACACCACTCATTGTGGCTGA	191	60.16	60.16	n
44_Lup_sdL	Lupinus	19,134	LAASNASE	TA(14)	CAAATCCCAAAAGCCTCCTT	GATCCTATTCCCGCATTGAA	199	60.42	59.86	y
45_Lup_ggps1	Lupinus	558,924	LAU15778	AAG(8)	GCATCGAAAAACCAAAAAGG	TGTGGCTCACGTAACGAAAC	208	59.56	59.76	y
46_Lup_aatP2	Lupinus	463,128	LUPP2AA	TTC(6)	GAAAAAGAAGGATTTAAAAACTGTGG	TCCGAATCGAATTACGAAGAG	185	59.85	59.31	n
47_Lup_EcoRIr[a]	Lupinus	168,333	LUPRSECOB	TAAA(3)	TCCAGCATCGGTTTAATGGT	AGGCAATTCTCTGTGGTTCG	256	60.33	60.25	n

48_Lup_TCP1	Lupinus	2E+07	LAL426419	AAC(5)	AGGGTCTGAGTCTGGATCCTC	TCGATTGTGAAAGGTGGTGA	199	59.68	60.09	n
49_Lup_TCP1	Lupinus	2E+07	LAL426419	CAG(5)	CGTCGATTGGCTAATCAAAAA	TAGTAGCCGCATTTGCACTG	203	60.08	60.04	y
50_Lup_CycB1	Lupinus	4,884,723	AF126105	TTG(4)	GACAAGGCCTTCAGTGCTGT	GGGCACCCCAAATTATGTTA	209	60.45	59.52	y
51_Lup_ACS4	Lupinus	6,650,973	AF119410	TC(8)	CATTCAACACGTTCTCATCCA	TTCCATCCAGCAAAGTAAGGA	212	59.56	59.69	y
52_Lup_LIN01	Lupinus	1.3E+07	BG154103	ACT(6)	TCTGGGTCACTCTGTGATCG	CACCCCAAGTTTCCATTTTG	212	59.82	60.2	n
53_Lup_LIN01	Lupinus	1.3E+07	BG154037	CT(9)	CCCTCAATTTTTGATCCCAAT	CCCACACTCCAAAAACCATC	244	60.01	60.21	n
54_Lup_LIN01	Lupinus	1.3E+07	BG153901	ATC(4)	TTCTTATGGCACCCCTCTTG	GGGTGGTTGTGGGTCATATT	202	60.07	59.39	y
55_Lup_ccG	Lupinus	1.1E+07	LAL297490	CAT(5)	CACCCCATTAACCATAAGCAA	GTGGCGAAGAACTGTGTAAGG	193	59.71	59.79	y
56_Lup_At	Lupinus	2,780,193	LAAJ3197	TTTGA(3)	CTGCGGCAAATTTTGTTCA	GAATATGGCAGATCTCACAGAAGA	185	60.77	59.75	n
57_Lup_LIPRP2	Lupinus	1,754,988	LLU47661	AAT(4)	TGTCCTATAATGATACTGAGGACGAG	CCATGGGAGAAAGGACACAC	202	59.95	60.36	y
58_Lup_Glu	Lupinus	454,311	LLNGS1G	TATGT(2)	ACAGGCCACAAGCAATTTTC	TTTGCAGCAGCATGTCTCTT	211	60.12	59.75	y
59_Lup_Glu	Lupinus	454,311	LLNGS1G	AAATC(2)	TCCCTTTTCCAATTCATTCTTC	TGGATGTGATCCACGACATT	192	59.42	59.77	n
60_Lup_aaT1	Lupinus	463,128	LUPP2AA	TC(7)	GTAATCATCGCCCACGTTTT	GAGAAGAAGCCATGAGAAACTGA	205	59.83	60.01	n
61_Lup_aaT1	Lupinus	463,128	LUPP2AA	TTA(5)	GTTGGTGTAAGGCCCCACT	GGTGTTTGATTGTTCCTTGCT	204	59.84	59.09	y
62_Stylo_shst1[a]	Stylosanthes	1.8E+07	SMA416738	GGA(7)	TGAAGCAACTCTTCTTCACATAGAC	GAAAGAATGCTTGATCTCTTGGA	132	59.15	59.85	y
63_Stylo_IGS[a]	Stylosanthes	4,468,018	SMA131086	CTTT(3)	CAAGTCCCTCTATCCCCAAAA	TCCAAACAAATACTTATGGTTGTTG	185	60.3	59.28	y
64_Stylo_SSR4-16b[a]	Stylosanthes	4,151,098	SGU011286	TTC(5)	GCCGTGTTTTCTGCTTTTTC	GGATGACGTGGCGTTAAATC	184	59.87	60.34	y
65_Stylo_SSR4-9[a]	Stylosanthes	4,151,096	SGU011284	GT(7)	CGCATTTCTCCGTCTCTCTC	TCAACAAGCCAACACACACA	207	60.1	59.75	y
66_Stylo_SSR4-5[a]	Stylosanthes	4,151,095	SGU011283	TTTC(3)	GGTACATTCTGGCGCATTTT	TGACATGGCCCAGTAAGAAA	148	59.97	59.12	y
67_Stylo_SSR2-43[a]	Stylosanthes	4,151,091	SGU011279	ATC(4)	GCTGCTGCCTATCTAGAAGCTC	TCTCTCTCTCGTTGGGGTATTT	118	59.93	59.24	y
68_Stylo_SSR1-24[a]	Stylosanthes	4,151,090	SGU011278	TTTC(3)	TGGCCTCTATCTCCCTTGAA	CATCACCAACCAACCAATCA	199	59.77	60.22	y
69_Stylo_IGS[a]	Stylosanthes	4,138,603	SVAJ0774	CTTT(3)	CCAAAAGACCCGCTTAACTTT	GTATTCCAATACATATTCCAACCAAA	186	59.65	59.37	y
70_Stylo_epe	Stylosanthes	4,099,913	SHU91857	AAG(4)	CGAACCTCCTCCACAAGAGA	AGAGATCCAAACGGGATCG	204	60.38	60.02	y
71_Stylo_per	Stylosanthes	1,377,788	SSNCAPE	ATT(7)	TCGCCTTCATTTCCATGATT	AATTGGTGCAGATTATTCTACGG	232	60.41	59.43	n
72_Stylo_per	Stylosanthes	1,377,788	SSNCAPE	AT(8)	TCCAGTGGCCAGATTAGGAC	TTTAAACCTCGGAAGTACCCTTT	198	60.07	59.47	n
73_Stylo_per	Stylosanthes	1,377,788	SSNCAPE	AAAT(4)	GAGCATGGATTGCCATTTTT	ACCCCTTTTCAGGCGAAATA	263	59.91	60.79	n
74_Stylo_shst1[a]	Stylosanthes	1.8E+07	SHU416729	GGA(6)	TGAAGCAACTCTTCTTCACATAGAC	GAAAGAATGCTTGATCTCTTGGA	129	59.15	59.85	y
75_Stylo_ITS[a]	Stylosanthes	1.6E+07	SSU320388	CGGC(3)	CGTCCTCAGACAAACCCTGT	GAGATATCCGTTGCCGAGAG	199	60.15	59.8	y
76_Stylo_IGS[a]	Stylosanthes	8,546,954	SFR131262	CTTT(3)	CAAGTCCCTCTATCCCCAAAA	TCCAAACAAATACTTATGGTTGTTG	187	60.3	59.28	y
77_Stylo_shst2[a]	Stylosanthes	1.8E+07	SAF416717	AT(5)	CAAACACCAAGTATTCTAACCCTCT	TATTTAAGGTTGCATGACAGGTG	100	59.09	59.07	y
78_Ades_trnL[a]	Adesmia	6,983,980	AF208901	AT(5)	CCTTGCGAATTAGGAAAGGA	TGGAGTGAATGATTTGATCAGTG	224	59.29	59.99	y
79_Amor_gPP	Amorpha	1.7E+07	AF435969	CTCTT(3)	CCACACCCTCCTCTCAACTC	TTCTCGCTGATTTGGTTCAA	196	59.68	59.4	y

[a] Genomic SSR.

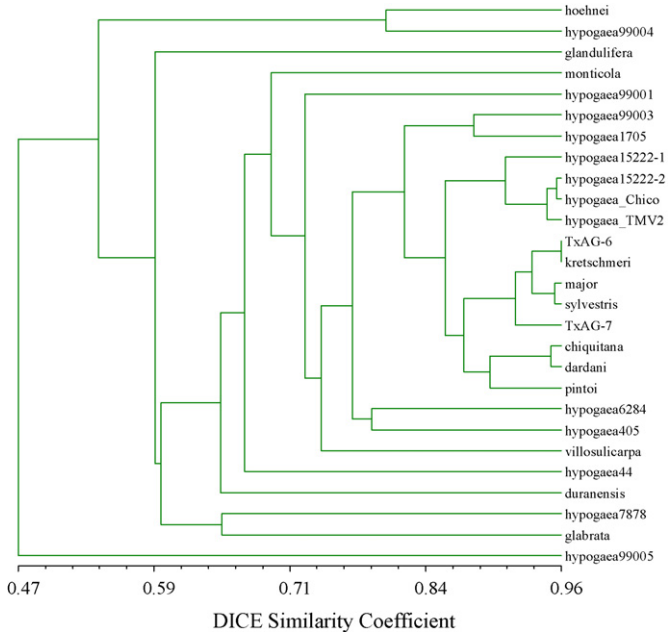

Fig. 5. Cluster analysis of 27 diverse *Arachis* accessions with 18 polymorphic SSRs. Accession codes as detailed in Table 3.

3.4. Genetic relationships in Arachis species

The dendrogram constructed using Nei and Li's similarity coefficient and UPGMA clustering is presented (Fig. 5), with a very high correlation coefficient of 0.929, indicating an excellent fit of the clustering to the similarity matrix. In total, 13 accessions of the cultivated species, *A. hypogaea*, were included, representing the parents of mapping populations developed at ICRISAT for various foliar diseases. Seven of the *A. hypogaea* accessions cluster together very closely at over 80% similarity, however a number of species representing the alternative sections within the genus *Arachis* also cluster together very closely, e.g. *A. major* from section *Erectoides* clusters very closely with *A. sylvestris* from section *Heteranthe*; likewise *A. chiquitana* from section *Procumbentes* clusters very closely (again, at approx. 95% similarity) with *A. dardani* from section *Heteranthe*.

4. Discussion

4.1. In silico SSR analysis

This study has demonstrated that the *in silico* approach to SSR identification and development is a valid strategy in

Table 5
Summary of cross-genera amplification

Species	# markers	% transferability to					
		Arachis	*Amorpha*	*Chaetocalyx*	*Dalbergia*	*Lupinus*	*Stylosanthes*
Amorpha	1	100	X	0	0	100	100
Adesmia	1	100	100	0	0	100	100
Chaetocalyx	1	100	100	X	100	100	100
Dalbergia	3	66	100	100	X	100	100
Lupinus	57	65	79	16	39	X	65
Stylosanthes	16	81	75	6	19	31	X

X: transfers to itself.

Table 6
List of SSRs polymorphic in *Arachis* germplasm and PIC value

Marker	Marker origin genus	Type of SSR	Repeat	No. of alleles	PIC
02_Dal_matK	*Dalbergia*	EST	AT(5)	3	0.1308
04_Dal_PHYA	*Dalbergia*	EST	ATG(6)	3	0.32882
28_Lup_LIN01	*Lupinus*	EST	TTC(6)	6	0.62307
34_Lup_app	*Lupinus*	EST	TA(5)	2	0.49383
41_Lup_Albus	*Lupinus*	Genomic	AAATC(2)	3	0.51978
50_Lup_CycB1	*Lupinus*	EST	TTG(4)	5	0.59142
61_Lup_aaT1	*Lupinus*	EST	TTA(5)	2	0.2524
63_Stylo_IGS	*Stylosanthes*	Genomic	CTTT(3)	4	0.72
64_Stylo_SSR4-16b	*Stylosanthes*	Genomic	TTC(5)	5	0.54498
65_Stylo_SSR4-9	*Stylosanthes*	Genomic	GT(7)	7	0.74248
66_Stylo_SSR4-5	*Stylosanthes*	Genomic	TTTC(3)	3	0.49308
67_Stylo_SSR2-43	*Stylosanthes*	Genomic	ATC(4)	4	0.66766
68_Stylo_SSR1-24	*Stylosanthes*	Genomic	TTTC(3)	6	0.75635
70_Stylo_epe	*Stylosanthes*	EST	AAG(4)	4	0.5736
74_Stylo	*Stylosanthes*	Genomic	GGA(6)	5	0.6075
76_Stylo	*Stylosanthes*	Genomic	CTTT(3)	3	0.51753
77_Stylo_shst2	*Stylosanthes*	Genomic	AT(5)	2	0.48242
79_Amor_gPP	*Stylosanthes*	EST	CTCTT(3)	4	0.61625

lesser-studied crops. The *in silico* approach encompasses the use of pattern recognition technologies and statistical techniques to examine large amounts of data [19]. The wet lab generation of SSR markers is highly time consuming and expensive. A valid alternative is presented to complement existing studies which have proven this approach for species closely related to model species, where abundant sequence data is already available [30].

4.2. Genomic DNA-derived SSR versus EST-derived SSRs

The aeschynomenoid/dalbergoid and genistoid clades of the Leguminosae family contain very few species with abundant sequence data available, however even with the limited data available, the *in silico* approach offers the possibility of identifying SSRs in a cost-effective manner [19]. The transferability rate between genera in the aeschynomenoid/dalbergoid and genistoid clades appeared to be related more to the location of the SSR within the genome; i.e. whether the SSR is found in transcribed DNA i.e. is EST-derived, or in genomic DNA, than to the genetic relatedness within the clade, however there is clearly an interaction between the two factors as well. As anticipated, the functional constraints of EST-derived SSRs lead to high levels of conservation across genera, allowing a greater transferability rate, however there has been concern about the level of polymorphism observable using SSRs derived from such conserved sequences, as opposed to the more evolutionarily unconstrained genomic SSRs [10]. The present study indicates that even EST-derived SSRs can provide sufficient information to differentiate between species, with PIC scores ranging from 0.13 to 0.62 from EST-derived SSRs. However, the average PIC scores were higher from the genomic DNA derived-SSRs (0.65) than from the EST-derived SSRs (0.47). Infact, in earlier studies also, the genomic DNA derived SSRs, as compared to EST-derived SSRs displayed higher polymorphism [31–33]. Another important feature of the EST-derived SSRs was a smaller SSR repeat size; for instance, the percentage of tetra-mer repeats was only 4.4% among the EST-derived SSRs, whereas 55% of the genomic-derived SSRs were tetramers. Therefore, in addition to the origin of EST-derived SSRs from the conserved proportion of the genome, the smaller SSR repeats in case of EST-derived SSRs than genomic DNA derived SSRs may contribute to their lower polymorphism content [7].

4.3. SSRs for groundnut breeding

A maximum of 53% genetic dissimilarity across 71 alleles was observed amongst the 27 *Arachis* accessions screened in this study. This level of polymorphism is comparable to a recent study with 23 genomic-derived SSRs only among 22 *A. hypogaea* accessions with varying levels of resistance to leaf rust and late leaf spot [34], which found a maximum of 56% genetic dissimilarity across 135 alleles. This study [34] reported the highest level of diversity yet recorded between cultivated groundnut genotypes, adding support to previous observations that SSRs have a higher discriminatory power compared to other molecular markers [7,35,36]. The high level of genetic discrimination observed with both the EST and genomic-derived SSRs used in this study is very promising for further mining of available sequence data from related species and their application to the assessment of molecular diversity assessment to facilitate the identification of agronomically valuable and diverse germplasm for use in linkage mapping and genetic enhancement of specific traits in groundnut.

4.4. SSR transferability

In addition to the genetic distance and sequence conservation of species and primers tested, other factors which can influence the rate of SSR transferability across genera include the differences in the genome size of species tested in addition to stringency and annealing temperatures used in PCR [37]. However, it has also been previously observed that even with the same DNA template and the same primers, PCR may generate different amplicons at different stringency levels [37]; the lower the stringency levels, the higher the expected rate of transferability across genera. However, even with low stringency conditions (i.e. low Tm and high number of cycles), the transferability rate among selected legumes was found to be lower (31%) than reported rates from peach SSRs to apple and strawberry [37,38]. The higher transferability rate observed in our study (60%) may reflect the closer genetic relationships of the species selected from the aeschynomenoid/dalbergoid clade.

As discussed earlier in the legume genetics/genomics community, in order to take full advantage of DNA marker technology, there is a need to develop a core set of at least 1000 universal STS markers among all legume species [39]. These STS markers will begin with strategically chosen PCR-based markers that have already been developed, however eventually the core set will grow by mining legume sequence data in order to find highly conserved sequences shared by all legumes [30]. The approach outlined in the current study will provide very valuable information for the development of such a core STS set, which will provide powerful tools for trait mapping and marker-assisted breeding in other legume species, including the ‘‘orphan’’ crops with very little sequence data available.

Acknowledgements

The authors gratefully acknowledge Dr. H.D. Upadhyaya for assistance with selection of material. This research was supported by the Asian Development Bank-funded project ‘Rapid Crop Improvement for Poor Farmers in the Semi-Arid Tropics of Asia’ (RETA 5945). Biotechnology research at ICRISAT is grateful for unrestricted grants from the governments of Japan, UK and EU.

References

[1] FAOSTAT http://faostat.fao.org/default.aspx.
[2] T.M. Halward, H.T. Stalker, E. LaRue, G. Kochert, Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and related wild species, Genome 34 (1991) 1013–1020.

[3] T.M. Halward, T. Stalker, E. Larue, G. Korchert, Use of single-primer DNA amplification in genetic studies of peanut (*Arachis hypogaea* L.), Plant Mol. Biol. 18 (1992) 315–325.
[4] M.D. Burow, C.E. Simpson, J.L. Starr, A.H. Pateron, 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 (2001) 823–837.
[5] M. Mohan, S. Nair, A. Bhagwat, T.G. Krishma, M. Yano, C.R. Bhatia, T. Saski, Genome mapping, molecular markers and markers-assisted selection in crop plants, Mol. Breeding 3 (1997) 87–103.
[6] M. Litt, J.A. Lutty, A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeats within the cardiac muscle actin gene, Am. J. Hum. Genet. 44 (1989) 397–401.
[7] P.K. Gupta, R.K. Varshney, The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat, Euphytica 113 (2000) 163–185.
[8] J. Rongwen, M.S. Akkaya, A.A. Bhagwat, U. Lavi, P.B. Cregan, The use of microsatellite DNA markers for soybean genotypes identification, Theor. Appl. Genet. 90 (1995) 43–48.
[9] T. Panaud, O.X. Chen, S.R. McCouch, Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.), Mol. Gen. Genet. 252 (1996) 597–607.
[10] R.K. Varshney, A. Graner, M.E. Sorrells, Genic microsatellite markers in plants: features and applications, Trends Biotechnol. 23 (2005) 48–55.
[11] M.S. Hopkins, A.M. Casa, T. Wang, S.E. Mitchell, R.E. Dean, G.D. Kochert, S. Kresovich, Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut, Crop Sci. 39 (1999) 1243–1247.
[12] M.E. Ferguson, M.D. Burow, S.R. Schultz, P.J. Bramel, A.H. Paterson, S. Kresovich, S. Mitchell, Microsatellite identification and characterization in peanut (*A. hypogaea* L.), Theor. Appl. Genet. 108 (2003) 1064–1070.
[13] G. He, R. Meng, M. Newman, G. Gao, R.N. Pittman, C.S. Prakash, Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.), BMC Plant Biol. 3 (2003) 3–9.
[14] M.C. Moretzsohn, L. Leoi, K. Proite, P.M. Guimara, S.C.M. Leal-Bertioli, M.A. Gimenes, W.S. Martins, J.F.M. Valls, D. Grattapaglia, D.J. Bertioli, A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae), Theor. Appl. Genet. 111 (2005) 1060–1071.
[15] D.A. Palmieri, M.D. Bechara, R.A. Curi, M.A. Gimenes, C.R. Lopes, Novel polymorphic microsatellite markers in section *Caulorrhizae* (*Arachis*, Fabaceae), Mol. Ecol. Notes 5 (2005) 77–79.
[16] M.A. Gimenes, A.A. Hoshino, A.V. Barbosa, D.A. Palmieri, C.R. Lopes, Characterization and transferability of microsatellite markers of the cultivated peanut (*Arachis hypogaea*), BMC Plant Biol. 27 (2007) 7–9.
[17] R. Tang, G. Gao, L. He, Z. Han, S. Shan, R. Zhong, C. Zhou, J. Jiang, Y. Li, W. Zhuang, Genetic diversity in cultivated groundnut based on SSR markers, J. Genet. Genomics 34 (2007) 449–459.
[18] J.J. Doyle, M.A. Lucknow, The rest of the iceberg. Legume diversity and evolution in a phylogenetic context, Plant Physiol. 131 (2003) 900–910.
[19] R.K. Varshney, R. Sigmund, A. Boerner, V. Korzun, N. Stein, M. Sorrells, P. Langridge, A. Graner, Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice, Plant Sci. 168 (2005) 195–202.
[20] H.K. Choi, J.-H. Mun, D.-J. Kim, H. Zhu, J.-M. Beak, J. Mudge, B. Roe, N. Ellis, J. Doyle, G.B. Kiss, N.D. Young, D.R. Cook, Estimating genome conservation between crop and model legume species, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 15289–15294.
[21] M.A. Saghai-Maroof, K.M. Soliman, R.A. Jorgensen, R.W. Allard, Extraordinarily polymorphic microsatellite DNA in barley—species diversity, chromosomal locations and population dynamics, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 5466–5470.
[22] J.J. Doyle, J.L. Doyle, A rapid DNA isolation procedure for small quantities of fresh leaf material, Phytochem. Bull. 19 (1987) 11–15.
[23] G.M. Kolodny, An improved method for increasing the resolution and sensitivity of silver staining of nucleic acid bands in polyacrylamide gels, Anal. Biochem. 138 (1984) 66–67.
[24] M. Nei, W.H. Li, Mathematical model for studying genetic variation in terms of restriction end nucleases, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 5269–5273.
[25] P. Jaccard, Nouvelles recherches sur la distribution florale, Bull. Soc. Vaud. Sci. Nat. 44 (1908) 223–270.
[26] R.R. Sokal, C.D. Michener, A statistical method for evaluating systematic relationships, Univ. Kansas Sci. Bull. 38 (1958) 1409–1438.
[27] F.J. Rohlf, Numerical Taxonomy and Multivariate Analysis System, Applied Biostatistics Inc., New York, 2001.
[28] J.B. Kruskal, M. Wish, Multidimensional Scaling, Sage, Newbury Park, 1978.
[29] B.S. Weir, Genetic data analysis. Methods for discrete genetic data, Sinauer Associates, Inc., Sunderland, MA, 1990, pp. 125.
[30] V. Mahalakshmi, P. Aparna, S. Ramadevi, R. Ortiz, Genomic sequence derived simple sequence repeats markers. A case study with Medicago spp., Elect. J. Biotechnol. 5 (2002) 233–242.
[31] F. Leigh, V. Lea, J. Law, P. Wolters, W. Powell, P. Donini, Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat, Euphytica 133 (2003) 359–366.
[32] K. Chabane, G.A. Ablett, G.M. Cordeiro, J. Valkoun, R.J. Henry, EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley, Genet. Resour. Crop Evol. 52 (2004) 903–909.
[33] R.K. Varshney, D.A. Hoisington, A.K. Tyagi, Advances in cereal genomics and applications in crop breeding, Trends Biotechnol. 24 (2006) 490–499.
[34] E.S. Mace, D.T. Phong, H.D. Upadhyaya, J.H. Crouch, SSR analysis of cultivated groundnut (*Arachis hypogaea* l.) germplasm resistant to rust and late leaf spot diseases, Euphytica 152 (2006) 317–330.
[35] W. Powell, M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey, A. Rafalski, The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis, Mol. Breeding 2 (1996) 225–238.
[36] J.H. Crouch, H.K. Crouch, H. Constandt, A. Van Gysel, P. Breyne, M. Van Montagu, R.L. Jarret, R. Ortiz, Comparison of PCR-based molecular marker analyses of *Musa* breeding populations, Mol. Breeding 5 (1999) 233–244.
[37] M.L. Wang, A.G. Gillaspie, M.L. Newman, R.E. Dean, R.N. Pittman, J.B. Morris, G.A. Pederson, Transfer of simple sequence repeat (SSR) markers across the legume family for germplasm characterisation and evaluation, Plant Genet. Resour. 2 (2004) 107–119.
[38] E. Dirlewanger, P. Cosson, M. Travaud, M.J. Aranzana, C. Poizat, A. Zanetto, P. Arus, F. Laigret, Development of microsatellite markers in peach [*Prunus persica* (l.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* l.), Theor. Appl. Genet. 105 (2002) 127–138.
[39] US Legume Crop Genomics Initiative http://www.aspb.org/publicaffairs/stakeholders/USLCGI.PDF.
[40] J.J. Doyle, J.L. Doyle, J.A. Ballenger, E.E. Dickson, T. Kajita, H. Ohashi, A phylogeny of the chloroplast gene rbcL in the leguminosae: taxonomic correlations and insights into the evolution of nodulation, Am. J. Bot. 84 (1997) 541–554.
[41] T. Kajita, H. Ohashi, Y. Tatesihi, C.D. Bailey, J.J. Doyle, rbcL and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies, Syst. Bot. 26 (2001) 515–536.