Genetic diversity within *Lablab purpureus* and the application of gene-specific markers from a range of legume species

S. C. Venkatesha[1], M. Byre Gowda[1], P. Mahadevu[1], A. Mohan Rao, D.-J. Kim[2], T. H. N. Ellis[3] and M. R. Knox[3]*

[1]*University of Agricultural Sciences, GKVK Campus, Bangalore 560065 Karnataka, India*, [2]*International Institute of Tropical Agriculture (IITA), C/O Biosciences Eastern and Central Africa (BECA). PO Box 30709, Nairobi, Kenya and* [3]*John Innes Centre, Colney Lane, Norwich, NR4 7UH, UK*

Abstract

Molecular markers have been used to study genetic diversity within a set of *Lablab purpureus* accessions collected from the southern states of India. Amplified fragment length polymorphism (AFLP) molecular marker studies using a total of 78 *L. purpureus* accessions with nine primer combinations showed there was very little genetic diversity within the *L. purpureus* accessions from the southern Indian germplasm collection as compared to a set of 15 accessions from other international germplasm collections that included African accessions. The set of 15 were selected from a random amplified length polymorphism (RAPD) marker study and chosen on the basis of widest genetic distance. Further molecular analysis with polymerase chain reaction (PCR) markers from 97 expressed sequence tag (EST) and gene-specific primer pairs, designed from a range of legume sequences, concurred with the AFLP analyses. Both of these approaches provide a wealth of markers for diversity and mapping studies. The 97 sequence-specific primer pairs tested in *L. purpureus* resulted in 70% amplification success, with 44% of primer pairs amplifying single bands and 10% double bands. Markers generated from these EST and genomic sequences provide useful cross-reference to comparative legume genomics that will potentially have long-term benefit to legume plant breeding.

Keywords: AFLP; gene-specific markers; genetic diversity; *Lablab purpureus*; legumes; *Medicago truncatula*

Introduction

Grain legumes are important food and forage crops globally (Graham and Vance, 2003). Although legumes, consumed either as dried grain or fresh vegetables, are low in sulphur-containing amino acids, this is counterbalanced by the presence of many other beneficial nutritional and health factors (Champ *et al.*, 2002). In combination with cereals, vegetables and fruits, legumes contribute towards balanced and healthy human nutrition, providing protein, carbohydrate, essential vitamins and oils, trace elements and fibre (Champ *et al.*, 2002). Additionally, legumes are a source of high-quality animal fodder (Graham and Vance, 2003), and interact symbiotically with rhizobia, providing nitrogenated soils.

Lablab purpureus (L.) Sweet (*Dolichos lablab*, hyacinth bean) is an important legume, cultivated throughout the tropics and subtropics. It is mostly grown for human consumption and animal forage (Murphy and Colucci, 1999), and is widely used as an intercrop in India and Australia (Pengelly and Maass, 2001; Maass *et al.*, 2005). It is a versatile crop, able to be cultivated in a range of climates and soil types and is relatively drought tolerant. In addition to a food

* Corresponding author. E-mail: knox@bbsrc.ac.uk

crop, *L. purpureus* is often grown as a weed suppressor, a soil erosion retardant and as green manure (Liu, 1996; Murphy and Colucci, 1999; Pengelly and Maass, 2001).

Traditional, 'classical', breeding, involving exotic and indigenous varieties from southern India, has successfully produced photo-insensitive, high-yielding, determinate *L. purpureus* lines that can be grown all year round (Maha-devu and Byre Gowda, 2005). Further breeding objectives would combine characteristics like photo-insensitivity and yield with other desirable attributes, such as improved fragrance, together with pest and disease resistance (Maha-devu and Byre Gowda, 2005). Some legume crops, as is also the case for non-legume crops, have a relatively narrow genetic base (Udupa *et al.*, 1993; Kochert *et al.*, 1996; Ladizinsky, 1998). The use of inter-species crosses to produce lines with desirable traits, such as yield improvement or disease resistance, has been one way to overcome the genetic bottleneck of crop plant domesti-cation (Ladizinsky, 1998; Winter *et al.*, 2000; Rao *et al.*, 2003; Gur and Zamir, 2004). This breeding approach, of inter- or wide intra-specific crosses, can benefit from good genetic maps with robust markers anchored to reference genomes, such as are available from comparative genetic maps.

Previous studies of diversity within the *L. purpureus* spp. and sub-species using RAPD (random amplified length polymorphism) (Liu, 1996) and AFLP (amplified fragment length polymorphism) (Maass *et al.*, 2005) suggest that there is considerable molecular variation from accessions found in collections that comprise lines from Africa, Asia and Europe, and that these lines could provide useful and desired traits if incorporated into breeding strategies. In addition to this molecular diversity, *L. purpureus* also has diverse phenotypic characters (Basavarajappa and Byre Gowda, 2000; Pengelly and Maass, 2001; Maass, 2006), reflected in their different growth habits, that have served to provide cultivars fitted to specific environments. Yield losses can be as much as 70% in any 1 year from pests and diseases (M. Byre Gowda, unpublished obser-vations), and this low productivity problem is not unique to *L. purpureus* but also applies to other legumes (Udupa *et al.*, 1993; Kochert *et al.*, 1996; Graham and Vance, 2003; Rao *et al.*, 2003). The introgression of alleles from diverse germplasm may provide some opportunities to control yield instability.

Isozymes and RFLP (restriction fragment length poly-morphism) (nuclear and plastid) as well as polymerase chain reaction (PCR)-based markers [RAPD, simple sequence repeat (SSR), AFLP and single nucleotide poly-morphism (SNP)] have been used to analyse germplasm (Palmer *et al.*, 1985; Karp *et al.*, 1996; Lu *et al.*, 1996; Mba and Tohme, 2005). The use of AFLP is advantageous as it has no need for sequence information from the genotypes being analysed, a multiplex of reproducible

bands is generated; AFLP is robust and can be automated for high-throughput large-scale studies. Other sequence-based markers do require prior information and, where sequence data are limited for a given species, taxonomically close relatives can provide useful data. For *L. purpureus* there are few sequences available, but there is a vast amount of sequence data available from related legume species. Comparative mapping within the legume crop and model species (Zhu *et al.*, 2005) provides evidence that this wealth of sequence data can be used for genetic analysis even in neglected legume species such as *L. purpureus*, especially given that this genome aligns well with *Vigna radiata* (mungbean) (Humphry *et al.*, 2002).

The aim of the present study was to characterize the genetic diversity within regional *L. purpureus* from southern India in comparison with lines from inter-national germplasm collections. Two molecular marker types, AFLP, and gene sequence-specific PCR-based mar-kers from a range of legumes, were used and tested within subsets of *L. purpureus* accessions. This diversity study approach has the potential to define population structure as an aid to identifying useful marker alleles in association with *L. purpureus* traits such as fragrance and pathogen resistance.

Materials and methods

Plant material

Experiments in this study involved 78 *L. purpureus* acces-sions (Table 1) from various sources: (1) GL *x*, from the Uni-versity of Agricultural Sciences (UAS), Bangalore, India, including those accessions collected from the southern Indian states of Karnataka, Andhra Pradesh and Tamil Nadu (Fig. 1); (2) accessions from international collections: those from the Australian Tropical Crops Genetic Resources Centre (ATCGRC), are prefixed with CPI, cv. Highworth and cv. Rongai were also obtained from ATCGRC; and those from the International Livestock Research Institute, Ethiopia, are prefixed with ILRI. Individuals from a range of legume species (*Medicago truncatula* Jemalong, *Glycine max*, *Vigna unguiculata*, and *Pisum sativum* JI15 from the John Innes Centre germplasm collection) were included where necessary for some experiments.

DNA preparation

A standardized DNA preparation method (Ellis *et al.*, 1984) was used for all legumes. *L. purpureus* DNA was prepared using the standardized method but included a heating step at 37°C for 10 min after addition of sodium dodecylsulphate (SDS), and nucleic acid was not spooled but recovered by

Table 1. Legume accessions and the marker type tested

Accession number	*Lablab purpureus*[a] Accession Title	Country of origin	Collection number	Marker type	
				AFLP[+]	Sequence-specific
1	HAS 8	India [K]	GL 408	+	
2	MAC 8	India [K]	GL 410	+	
3	Local avare	India [K]	GL 155	+	
4	Konkanwal 1	India [Mh]	GL 135	++	
5	MAC 2	India [K]	GL 414	++	
6	Banlepade local	India [AP]	GL 16	+	
7	MAC 3	India [K]	GL 412	++	
8	Royalbeans	India [TN]	GL 411	+	
9	Gudibande local	India [K]	GL 07	+	
10	Flora	India [TN]	GL 415	+	
11	HAS 5	India [K]	GL 402	+	
12	HAS 6	India [K]	GL 406	++	
13	CO 2	India [TN]	GL 162	++	
14	Magadi local	India [K]	GL 103	+	
15	Konkanwal 2	India [Mh]	GL 133	+	
16	MAC 7	India [K]	GL 405	+	
17	Pendal avare	India [K]	GL 508	+	
18	HAS 7	India [K]	GL 407	+	
19	HA 3	India [K]	GL 403	+	+
20	Rajankunte Local	India [K]	GL 116	+	
21	Dabbe Avare	India [K]	GL 151	+	
22	Golden seed	India [TN]	GL 417	+	
23	Dindigal local	India [TN]	GL 110	+	
24	Kathalagere	India [K]	GL 138	+	
25	MAC 6	India [K]	GL 413	++	
26	MAC 9	India [K]	GL 409	+	
27	Malai Avare Loca	India [TN]	GL 04	+	
28	Nandini	India [TN]	GL 416	+	
29	Local Mani-Avare	India [K]	GL 32	+	
30	Devagiri	India [K]	GL 76	+	
31	Belagola	India [K]	GL 129	+	
32	Halebedu	India [K]	GL 102	+	
33	Muragamalla	India [K]	GL 87	+	
34	Huliyar	India [K]	GL 99	+	
35	Ranganahalli	India [K]	GL 126	+	
36	Harale katte	India [K]	GL 65	+	
37	MAC 1	India [K]	GL 404	+	+
38	HAS 1	India [K]	GL 401	+	
39	ILRI 6536	Ex. Ethiopia		+	+
40	ILRI 6930	Ethiopia		+	
41	ILRI 13 700	Ethiopia		+	
42	cv. Highworth	Ex India [TN]		+	+
43	CPI 24 973 (sub sp. *uncinatus*)	Zimbabwe		+	
44	CPI 31 113	Uganda		+	
45	CPI 36 903	Ex. Ukraine		++	+
46	CPI 41 222	Burma		+	
47	CPI 51 564 (sub sp. *uncinatus*)	Zambia		+	
48	CPI 52 508	Mozambique		+	
49	CPI 52 535	India		++	+
50	CPI 52 544	India		+	+
51	CPI 52 552	India		++	+
52	CPI 60 216	Uganda		+	
53	CPI 67 639	India		+	
54	cv. Rongai	Kenya		+	+
55	Uthanur	India [K]	GL 124	+	

Table 1. *Continued*

Accession number	*Lablab purpureus*[a] Accession Title	Country of origin	Collection number	Marker type	
				AFLP[†]	Sequence-specific
56	Mallamachakunte	India [K]	GL 156	+	
57	MC 13	India [K]	GL 123	+	
58	Salavara	India [TN]	GL 106	+	
59	B.R.Hills	India [K]	GL 37	+	
60	Andadihalli	India [K]	GL 146	+	
61	MAC 10-1	India [K]	GL 105	+	
62	Adagoor-III	India [K]	GL 139	+	
63	D.B.Pur local	India [K]	GL 127	+	
64	Batlapalli	India [K]	GL 55	+	
65	Doddamagge	India [K]	GL 49	+	
66	Marigowdana Doddi	India [K]	GL 27	+	
67	Devegowdana Doddi	India [K]	GL 153	+	
68	Hindupur	India [AP]	GL 532	+	
69	Doddgabbadi	India [K]	GL 503	+	
70	Mallanahalli	India [K]	GL 515	+	
71	Madakasira	India [AP]	GL 517	+	
72	Hosagabbadi	India [K]	GL 516	+	
73	Pyadendi	India [AP]	GL 521	+	
74	Pooranipadu	India [K]	GL 506	+	
75	Bandaraguppe	India [K]	GL 121	+	
76	Bannikuppe	India [K]	GL 33	+	
77	Co 1	India [TN]	GL 132	+	
78	S R Local	India [K]	GL 514		+
	Other legumes				
	Medicago truncatula		Jemalong		+
	Vigna unguiculata				+
	Glycine max				+
	Pisum sativum		JI15		+

[a] All accessions are *L. purpureus* sub sp. *purpureus* except where indicated. Shaded accession numbers were within the UAS collection pre-2003. Indian collection sites: [K] Karnataka; [TN] Tamil Nadu; [AP] Andhra Pradesh; [Mh] Maharashtra.
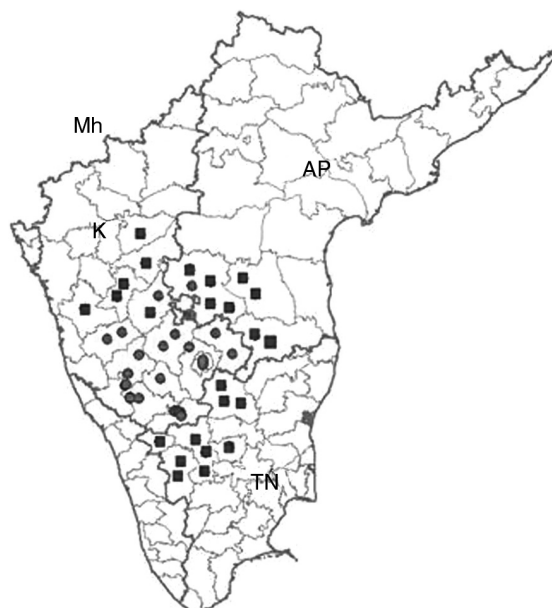[†] Single plus sign AFLP tested once; two plus signs AFLP tested twice.

Fig. 1. *Lablab purpureus* collection sites: AP, Andhra Pradesh; K, Karnataka; TN, Tamil Nadu; Mh, Maharashtra (to the north-west of K). Circles are with, and squares are without, global positioning system (GPS) information.

pellet after each ethanol precipitation. All DNA preparations were assessed on agarose and were all taken to an approximate concentration of 0.5 μg/μl.

Diversity analysis: molecular markers

AFLP

AFLP using *Eco* RI and *Mse* I digestion was carried out as described by Vos *et al*. (1995) initially with three primer combinations: Eco RI+CAT/Mse I+GTG, Eco RI+CAA/Mse I+GTG, Eco RI+CAG/Mse I+GTG and tested with accessions 1–53 (Table 1). A second AFLP series using six primer combinations – Eco RI+CAT/Mse I+GCC, Eco RI+CAG/Mse I+GTG, Eco RI+CAA/Mse I+GTG, Eco RI+CAT/Mse I+GTG, Eco RI+CAG/Mse I+GCC, Eco RI+CAA/Mse I+GCC was carried out on a further batch of accessions that included 12 accessions from the initial AFLP (those with two plus signs, Table 1). The reasoning behind the choice of primer combinations was that the results from the initial set of three primer combinations confirmed that there was limited genetic diversity within the UAS accessions. As there were constraints on the number of accessions/gel (limited to a maximum of 60), and many more accessions to test, the second set of six primer combinations was tested using accessions that included a few from the previous test plus an additional set. This strategy tested whether or not the limited genetic diversity within the UAS accessions observed initially was primer combination specific.

The unlabelled AFLP products were run out on 4.5% polyacrylamide gel electrophoresis (PAGE) at 1550 V for 1.5 h and silver stained (Bassam and Caetano-Anollés, 1993). Bands were scored by four individuals per gel.

Sequence-specific PCR markers

The primer pairs for sequence-specific markers were designed from a range of legume species and are listed in Table 2 as follows: (1) a set of 32 SSR markers (primer pairs 1–32) from EST or gene sequences (Wang *et al*., 2004), 17 of the 32 were designed from *M. truncatula* and 15 from *G. max*; (2) a set of primers designed from eight *M. truncatula* BAC (bacterial artificial chromosome) genomic sequences (primer pairs 33–40, Choi *et al*., 2004a). A set of 57 primer pairs were designed from nucleotide sequences from database searches as follows (Table 2): six primer pairs (41–46) from *L. purpureus* sequences, five of which were EST and one spanned the 5.8S rRNA transcription unit; three (primer pairs 47–49) were *P. sativum* gene sequences; primer pairs 50–97 comprise 45 EST and 3 genomic sequences, designed from a range of legumes within the *Phaseoleae* tribe.

PCR conditions were as follows: 0.1 μmol/l of each forward and reverse primer, 200 μmol/l each dNTP, 1 U of *Taq* DNA polymerase (Invitrogen, Paisley, UK) and the appropriate reaction buffers, 20–50 ng of DNA, in a 20 μl volume. For all species and all primer combinations the following touch-down cycling regime was used: 94°C/3 min; (94°C/30 s; 50°C/30 s; 72°C/60 s) repeat for 10 cycles reducing the annealing temperature by 0.5°C/cycle; (94°C/30 s; 45°C/30 s; 72°C/60 s) repeat 30 cycles; 72°C/10 min; 12°C/30 min. PCR products (10 μl of each) were visualized on 1.5% agarose for amplification. In the search for polymorphic differences, unlabelled PCR products from a few primer pairs were visualized from 4.5% PAGE gels after silver staining, as above for AFLP.

Data analysis

All gels were scored manually, independently by more than one individual and data were input into Excel (Microsoft) spreadsheets. The band data, for both marker types, was scored as a 1/0 (presence/absence) matrix. Genetic diversity was assessed using the PHYLIP version 3.5 package (http://evolution.gs.washington.edu.phylip.html) within which GENDIST was used for the calculation of distance matrices using Nei's genetic distance (D); these distance matrices were used as the basis for calculating the principal component analyses with Minitab version 13. (www.minitab.com).

Results

AFLP diversity analysis

The genetic diversity of 53 *L. purpureus* accessions (Table 1, 1–53) assessed from 151 AFLP markers from three primer combinations can be seen as a principal component analysis (PCA) plot (Fig. 2). This figure shows that there is very little genetic diversity within a set of 38 UAS *L. purpureus* accessions from the UAS germplasm collection in Bangalore (24 previously within the collection plus 14 new additions from more recent plant collection expeditions) compared to a more diverse set of 15 accessions (Table 1, 39–53), that included some from Africa. The more diverse set of 15 CPI and ILRI accessions, from the two International collections, were chosen on the basis of the widest genetic distance from a RAPD marker study (Liu, 1996). The principal component analysis of Fig. 2 showed that the first principal component accounts for 78% of the variance of the AFLP scoring data; the second and third components account for 11% and 4% of the variance. The limited diversity within the UAS *L. purpureus* accessions, captured in the first two dimensions of the data, can be seen as a tight cluster of points (within the small rectangle, Fig. 2); in contrast there is generally a wide scatter

Table 2. Primer information: species of origin, sequence type and primer sequence

Primer pair no.	Sequence ID	Species of Origin	Sequence class[a]	Sequence Info.	Sequence FORWARD 5′ to 3′	Sequence REVERSE 5′ to 3′
Ref.[1]						
1	AL370549	M. truncatula	1	AC(11)	CGTCCCGATATCGTCAACTT	CCACCACGACACATGTTACC
2	BF650979	M. truncatula	2	AT(28)	TTGTGGAAGGAACAACTCTGG	GAAACCGGCATGATTAAGACA
3	BF647899	M. truncatula	2	AT(34)	CTGTCAACAAGGGGTTAGGTG	TGCATCTACACCCAAAACAAA
4	AI974357	M. truncatula	2	TC(25)	TCTCAATTCCCCAACTTGCT	TCTCCTTCACCCATCTTTGC
5	AW256794	M. truncatula	2	TC(17)	GTCATCGAAGGCCAAAACAC	GTTTGCGAGAAACACCGATT
6	AA660488	M. truncatula	2	TC(19)	TTGCATTATTTTCCTTTTGACC	AACCCACAACCCAAAAATCA
7	AW584539	M. truncatula	2	ACA(8)	TTGATGGCCAATACATGTCG	GTTGAAGGAAGGTGGTGGTG
8	AW586959	M. truncatula	2	ACA(10)	CGAGAATCATCGTAATTGGACA	CGAAGTTCAATGGCATCAGA
9	AW775229	M. truncatula	2	AGC(8)	TACTGGGGTGATGCAAGACA	CAATACCCAGAGGAGCAGCTA
10	BF005356	M. truncatula	2	AGG(8)	CTTCAATTGTCACCGCCTCT	CTTATCTCGTCGTCCTCATCG
11	BF649209	M. truncatula	2	CCA(7)	AAGAGGCGGAGAGTGAGGTT	GGTAAGAGAACGAGCGAGGA
12	AW684360	M. truncatula	2	CGA(8)	TGTCATGGCGTCTCAAACC	CCTAACGCAGGAGAAGGAGA
13	AW685679	M. truncatula	2	GCC(5)	ACCTCACCTCACCTCCCTT	GATCATCTGGGTTTCGCAAG
14	AI974841	M. truncatula	2	TCT(11)	TCACCACCAAACCCAAC	TGGCAATGCTACAAGCCTAA
15	AW688216	M. truncatula	2	AGTG(9)	CACGAGGGATTGTTGTTTGA	GGAGCAGTAGGGTTGCATCT
16	AW127626	M. truncatula	2	GTTT(7)	CATTTTGAAGGAAGGAAGAAGG	ATTTGGAAGCGGAATGTGAA
17	AW688861	M. truncatula	2	CAACT(7)	TTGTTGTTGGCCTTCTTTGG	AAACCAACCACCTGTGTTGAC
18	AG81	G. max	3	AG(13)	ATTTTCCAACTCGAATTGACC	TCATCAATCTCGACAAAGAATG
19	AQ842128	G. max	4	TA(23)	TCAATGCTGATGCCATTTTC	TCGCGTATTATAGCACAACACC
20	AW186493	G. max	2	CTT(13)	GCGGTGATCCGTGAGATG	GCGGAAAGTAGCACCAAGAG
21	GMENOD2B	G. max	4	AT(17)	TAGGCAAAAGACTAAAAGAGTA	GCATGTCATTTTGATTGA
22	AG48	G. max	3	AG(18)	CAGAAACCTGAAATCTTCACC	CTTGGGTTTTTTATGGGTTC
23	AG50b	G. max	3	AG(19)	ATAAATTGGAAGATGTGTTGGC	TACTGATGTGGATTCTCCCAA
24	AG93	G. max	3	AG(17)	TCCATGCATGTATACTCCACC	TCATATGCCACAGGTTTTGTT
25	BE347343	G. max	2	GA(18)	GCGCAAGCACTGAATGTCA	GCGTCACTAACACCTATAACA
26	SoyPRP1	G. max	4	ATT(20)	CGTGCCAAATTACATCA	TGATGGGAACAAGTACATAA
27	AF186183	G. max	5	ATT(22)	GCGTATTTTGGGGGATTTTGAACA	GCGTTTCTCCTTCCTTATTCTTTCTCT
28	AW277661	G. max	2	ATT(23)	GCGCATGGAGCATCATCTTCATA	GCCAGAAAACCCAATCTTTATATCAATA
29	GMABAB	G. max	4	ATT(25)	CAAAACATAAAAAAGGTGAGA	AAGAACCACACTAATATTATT
30	BE801128	G. max	2	CAA(13)	GCGACAGTTCTCCACTCTTC	GCGCCCTTATAGATTTGTAAC
31	AW508247	G. max	2	CTT(10)	GCGCCCAATCCCAATCTCAC	GCGAAGCCAATAAATGATAAAAATC
32	AW620774	G. max	2	CTT(9)	GCGATTTCCCCTCTTACTC	GCGAAAAACCAAGTTC
Ref.[2]						
33	AQ917190	M. truncatula	6	BAC	CGAAACAATAATCACAAAACAAATCAG	ATCTTGTTTATATGTGTTGTTGAAGACAGAATT
34	AQ917191	M. truncatula	6	BAC	TGTCCTTGCTTCTTATCCTTCCTTCA	AGCAGCACAACAACTTACAACAACTC
35	AQ917144	M. truncatula	6	BAC	GCATGGAAATAGTTTGGGTTAGTAGTTAGT	CTGATAAATGCATATTTCAACATATGAATTAA
36	AQ917298	M. truncatula	6	BAC	CCATGCCATGGAAGGGTGTTT	GCAAGAACCAGATACCCTTGACATTT
37	AQ917327	M. truncatula	6	BAC	GGAACGTGCAGTTGTTGATGGTATTAT	GATGTAAAAACCTTTACACTTGATTGATTG
38	AQ917338	M. truncatula	6	BAC	AGCTTGTTGAGGTGGAAGGAAGTC	GTGTGTATGAGTGTCGTAAGCCCCT
39	AQ917375	M. truncatula	6	BAC	TGATTGACCCCTGCTTTGATGCT	GTCAGGTTTGTTGTTGTTTTTTCTTGCACTA
40	AQ917398	M. truncatula	6	BAC	CCAAACAAGGAAAAGTGTTGGTGTCA	ATGAGAAACTTTTGAAATTTAGGATACGATAG
Ref.[3]						

Table 2. *Continued*

Primer pair no.	Sequence ID	Species of Origin	Sequence class[a]	Sequence Info.	Sequence FORWARD 5′ to 3′	Sequence REVERSE 5′ to 3′
41	AB176566	*L. purpureus*	7	MET_1	TGTGCTGTGGGTGTGG	AGAGCTTTTGAACTTGTAAAAGGG
42	AB176567	*L. purpureus*	7	MET_2	AATGTCTTGCTGCGGTGG	AGCTCACTTGCAAGTACAGGG
43	AF067417	*L. purpureus*	7	fril	TACAGTGCTTCCTGAATGGG	ACAAACAACATACAAGTACTAACTGCAC
44	AY583516	*L. purpureus*	8	5SrDNA	CGTGTGTTGAGAGGGAGGG	AGAACAAGCTCGTGGGAAGTC
45	AY049046	*L. purpureus*	7	pDLL_1	CTTCATGCTACTTTTTCTTCTGGG	ACAAACACATTGTCGAGGAA
46	AY049047	*L. purpureus*	7	pDLT_1	ATGGTGGTGTTAAAGGTGTGC	TGCAAGGTTCGTAGGAGAAGTC
Ref.[3]						
47	AM040737	*P. sativum*	9	gdcH_5	ATCATTTCATTTCATTCTTCT	TGAGCCCAGTATTAACC
48	AM040737	*P. sativum*	9	gdcH_3	TTTTGGATGGACTCAAGTACG	TTTATCATCCATCCATCTTCATAG
49	AJ251536	*P. sativum*	9	Cathepsin	TGTGATGGTG GGTATCCC	GAATGCAACTTCAACCGG
Ref.[3]						
50	AF077224	*V. unguiculata*	7	Genic	AGCTGAAGCGCGCCACCATA	AGCAGCAGCCTTAAACTCATCAA
51	X89400	*V. unguiculata*	7	Genic	ATTGTTACGTTGCAGGATGTAGTG	TGCCACAGAGAATTTGATAGAGGA
52	AF279252	*V. radiata*	7	Genic	AGCTTCGCCAACTCCTCCATCAT	TCATCAGGGTCAAAGCTCATCAAT
53	AF151961	*V. radiata*	7	Genic	TTCCGCCATCACCTCTTCATTCC	CCGGCGACGTCAGGGTCTTA
54	AF402602	*P. vulgaris*	7	Genic	CCAGGGTTCAGGTTCCATCC	CCCTGTTGCATACTTACGGTCTCT
55	D13557	*V. radiata*	7	Genic	GAACGGATCCCAGAACGAG	GGAACACCCAAATCATCAAATAA
56	AY189907	*P. vulgaris*	9	Genic	TGCCAGTGGCAGTATCAAGTT	CTGCACCATTCCAGTCACCCT
57	U85250	*V. unguiculata*	7	Genic	ACTGGGCACAATTCCTATCTGACA	GGTGGCCATCTTTGCAACTTTAGC
58	BQ481569	*P. vulgaris*	7	Genic	TTTGTTCGTCTGAGGCGTCTTC	CCCGGGTCAGTTCATGTTGCTAC
59	AB056453	*V. unguiculata*	7	Genic	ATGGGCAGGGAGGCGACAGATTTAT	CAGTTGCCCGGGGTGCTTCC
60	AB062360	*P. lunatus*	7	Genic	CCCTCGGCTATAGCATTGAAGAC	ACGCATAAACAAAGAGGCTGGACT
61	CA901109	*P. coccineus*	7	Genic	CCAGCCAGCCATACCCATCT	CTCGACCCACATAATCTTTCAGGA
62	CA901635	*P. coccineus*	7	Genic	TAGGAGAAATGGGGTTGGCTTGAA	TAAACTTGCTCGCGTGTAATGAT
63	CA901660	*P. coccineus*	7	Genic	AACGGCGTCCCAGATGTTGT	TCTCTAGGCCTTGTTGCTCA
64	CA910054	*P. coccineus*	7	Genic	CCGCCGCAAGAAAGTTATCAT	GGCAGCACCATCAGCAAGGAA
65	CA912710	*P. coccineus*	7	Genic	GGTGGCTGTTGGATGTGGATG	CTGTGCCGCCAAGCATTCC
66	CA909621	*P. coccineus*	7	Genic	GAGCTTCCGTTTCATACCATTAG	CTTTCCCCTCGCCTTCAAATA
67	D10266-A	*V. radiata*	7	Genic	TACGAGGCATTTGGTTTGACAGTG	AGCCGGTTCCTCCATTTCTT
68	CK151423	*V. unguiculata*	7	Genic	TTTTCATGCGACTAGTTTATCCTT	AACTGGCACTCTTTTCTGGTCTTA
69	CB968060	*V. unguiculata*	7	Genic	ATTTCCGAGCTTGACACATT	CAAAACCACCAAGCCACATTC
70	CA914593	*P. coccineus*	7	Genic	AGCTCCAGATACTTGTTGTTGTTC	ATATGGGGGTGAGGGTGACG
71	CA906101	*P. coccineus*	7	Genic	AACACGCGGTACTACGAAATCCTC	CTCCGCGTCTCTGTCTCCTACCTC
72	CA908729	*P. coccineus*	7	Genic	ATCCTGTTTCGGTGCTGCTTCTTT	GGCTTCCTCCGCGGCGATTCC
73	AF165998	*V. unguiculata*	7	Genic	TTTTGCCCAAGGAGGTAG	TTTGATTTATAAGTGGGTAGAAGA
74	U08140	*V. radiata*	7	Genic	GACTTTGGCCTCTCCGTTTTCTTC	ATCTATATGTCCCTTCAATACTGC
75	U10419	*P. vulgaris*	9	Genic	TCCCACGATCTGTTTGAGC	CTGCATTAGTTTTTGTTTAGATTG
76	U28645	*P. vulgaris*	7	Genic	GGGAAGCTTGCGAGAATAGTTTTG	GATTTCCCTGCTGCGTTTGGTCTCT
77	U54703	*P. vulgaris*	7	Genic	TCGGTAAGAAAAAGGAAGAAGAGA	ATGAGGTGGCAGCAGGTGGTG
78	U70531	*P. vulgaris*	7	Genic	GGCTGAATTACTACCCTCCGTGTC	AGCAGCTCTTGTATCTCCCATTTG
79	U92656	*V. unguiculata*	7	Genic	CGATGAAGATCCGCGAAACTAT	TCTGGCCTCTGGAATGTGTC
80	X53603	*P. vulgaris*	9	Genic	TTGGCCCCATTAAAATAGTGACA	CCTGGTTGGGTGGGTGCTC
81	U20809	*V. radiata*	7	Genic	TCCGCAACAAGAGTGAACAG	GACAGCAGCAATATCTACCAACC

Table 2. *Continued*

Primer pair no.	Sequence ID	Species of Origin	Sequence class^a	Sequence Info.	Sequence FORWARD 5′ to 3′	Sequence REVERSE 5′ to 3′
82	BQ481672	*P. vulgaris*	7	Genic	ATTTTTGGTGTGCTTTCGTTTAT	TCCGTGGCTTGCTGATTAG
83	CA900138	*P. coccineus*	7	Genic	CACCGTCACCATATTCTCTGC	TCTCATTAGGCCGTGCTGAC
84	CA898279	*P. coccineus*	7	Genic	CCGTAGGATTGCTGATGAGG	TTCCGGTTATGTAAAAGATGTCC
85	CA898807	*P. coccineus*	7	Genic	TGCTGGTTTTGCGGAGGTT	CCCTGTAGGCTTCCATCATCTCTT
86	CA898809	*P. coccineus*	7	Genic	GTACCCAGTCACCAATCATAAAG	AATCCAACAACCAAGTCATCAGA
87	CA899061	*P. coccineus*	7	Genic	GTTCTCCCACCCTTCACATC	CACGTTGCAGGCTAATAAAATAAA
88	CA899306	*P. coccineus*	7	Genic	GAAGGGCTTTGGATGGTC	TGGGAGTTACGCAAAATAGGT
89	CA899524	*P. coccineus*	7	Genic	CTCTTCTTCAGGCTCACTCACACT	TCGCGATCGGAATAACAAG
90	AY193836	*V. unguiculata*	7	Genic	TGCTACGGCCATCCAACACC	AGCCAACCAACCTCCACCAAC
91	CA901208	*P. coccineus*	7	Genic	GGGGAGCTGTCACAAGAAGAAAT	TGCACATCCCCAGTCAATA
92	CA902017	*P. coccineus*	7	Genic	ACTTCAAGGAGATGGGAGGTTATT	TGTGTTATTGGTTGAGGCAGGAG
93	CA902145	*P. coccineus*	7	Genic	TTGATGACTTCCACCGGTTCTTA	ATTTCGCTATGGAGGCTTCAA
94	CA902154	*P. coccineus*	7	Genic	ACAAGGCTGATTTAGGTTCTCC	TGTATCAGCGCGTCTTTCAT
95	CA906247	*P. coccineus*	7	Genic	ATCTGGGCAAGCAAAGAGGT	AGAAGTTCCCATTCGTAAGTTGTT
96	CA906300	*P. coccineus*	7	Genic	TATTAGAGGGATTGAGTGGGATGC	CAAGCTTCAAATATACTTCTTCAGG
97	BQ481774	*P. vulgaris*	7	Genic	GACGTCGTTGTGCTCTCG	GTTGGGTACCCAAAACTCG

Ref. [1] Wang *et al.*, 2004; Ref. [2] Choi *et al.*, 2004a; Ref. [3] this study, nucleotide database search. ^a Sequence class description: 1, cDNA EST-derived: containing an SSR portion adjacent to but not spanned by the primers. 2, cDNA EST-derived: containing an SSR portion that is spanned by the primers. 3, Genomic DNA: known SSR-repeat region from other studies. 4, Genomic DNA: SSR in proximity to a gene is spanned by the primers, marker fragment is >500 bp upstream of the ATG start site. 5, Genomic DNA: Calypso retroelement SSR. 6, BAC genomic DNA: end-sequence-tagged. 7, cDNA EST-derived: aiming for intron spanning from genomic DNA template. 8, Nuclear DNA: rDNA repeat. 9, Genomic DNA: gene where the exon and intron positions are known, primers amplify across an intron.
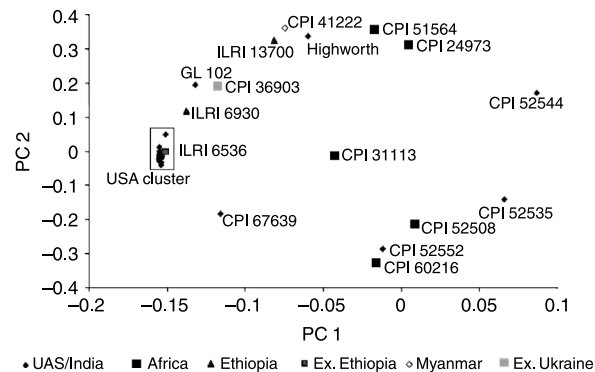
Fig. 2. Principal component analysis (PCA) of the molecular genetic diversity of 53 *Lablab purpureus* accessions with 151 amplified fragment length polymorphism (AFLP) markers. Figure symbols apply also to Figs 3a and b.

of the CPI and ILRI accessions. There are a couple of exceptions: UAS line GL 102 sits just outside the tight UAS cluster and the line ILRI 6536, originally from Ethiopia, clusters within the rectangle of UAS lines. It should be emphasized that the CPI lines 52544, 52535, 52552 and cv. Highworth are all from India and are peripheral within the PCA, along with other accessions from Africa and Myanmar that are scattered away from the main UAS cluster. The accessions CPI 51564 from Zambia and CPI 24973 from Zimbabwe cluster together, both are classified as sub-species *uncinatus* (Maass *et al.*, 2005). This suggests that the molecular variation, based on the AFLP data, within the UAS collection set of 38 is very limited compared to the species as a whole.

These observations were generally reinforced from the second study with 253 AFLP markers from a different set of six primer combinations, based on 33 accessions (numbers 54–77, Table 1), 19 of which were collected more recently from the southern states of India, and five already existing within the UAS germplasm collection. A subset of nine from the previous study (Table 1, those nine with two plus signs in the AFLP column) were also included.

Sequence-specific markers from a range of legume species

The sequence-specific markers from 97 primer pairs comprise nine classes of sequence type from a range of legumes, and Table 2 lists the species of origin, sequence class, primer sequences and the references from which the primers were obtained. These markers were derived from various sequence classes: cDNA ESTs some of which included SSR (simple sequence repeats) motifs, gene-related sequence information, and genomic BAC sequence.

The first 55 primer pairs (1–55, Table 3) were tested (in duplicate as controls for reproducibility) using *L. purpureus* lines, HA 3, SR L and MAC 1. For each primer pair products were reproducible and generally the same size across the genus. Control DNA, of the

species or genus from which primers were designed, was used for these 55 markers (Table 3).

A sub-set of 10 *L. purpureus* accessions along with a range of legume species (Table 1) was tested using a further set of 42 primer pairs (56–97, Table 3). Control DNA was not always available but other legumes were always tested in addition to *L. purpureus* (Table 3). From the AFLP analysis it was clear that the most diverse *L. purpureus* lines were the CPI and ILRI accessions from the two international collections; so a subset of these, which included some lines from India, were chosen as representative (Table 1). The accessions comprised three UAS parental breeding lines, HA 3, SR L and MAC 1; plus seven more diverse lines: ILRI 6536, cv. Rongai, cv. Highworth, CPI 52535, CPI 52544, CPI 52552, and CPI 36903.

L. purpureus *amplification success using primer pairs from a range of legume species*

Sequence information from a diverse range of legume species was used to design the 97 primers pairs which were then tested against *L. purpureus* and a range of other legumes; a negative PCR should be treated as failure, either from technical PCR problems but more than likely mis-matching at priming sites, rather than sequence absence.

The band scoring data can be seen in Table 3, and Table 4 summarizes the amplification success rate; generally, this is high, with 68 of the 97 primer pairs tested giving products. *L. purpureus* DNA was amplified by 22 of the 32 primer pairs generating SSR-related markers (Table 3 primer pairs 1–32); there was 82% amplification success from *M. truncatula* compared to 53% of the *G. max* primer pairs (Table 4). Table 3 shows that none of the three *P. sativum* primer pairs (numbers 41–43) amplified *L. purpureus* DNA and 3/8 of the primers from Choi *et al.* (2004a) successfully amplified *L. purpureus* DNA (primer pairs 33–40); in both these cases the control species *P. sativum* and *M. truncatula* were

Table 3. Amplification success of the sequence-specific markers

Primer pair no.	Sequence ID	Donor species	Lablab purpureus										Legume controls			
			HA-3	Mac-1	SRL	Rongai	CPI 36903	High-worth	CPI 52544	ILRI 6536	CPI 52552	CPI 52535	Mt	Gm	Ps	Vu
1	AL370549	*M. truncatula*	+	+	+	+							+		+	
2	BF650979	*M. truncatula*	+	+	+	+							+		+	
3	BF647899	*M. truncatula*	−	−	−	+							−		−	
4	AI94357	*M. truncatula*	−	−	−	−							+		−	
5	AW256794	*M. truncatula*	+	+	+	+							+		+	
6	AA660488	*M. truncatula*	−	−	+	−							+		+	
7	AW584539	*M. truncatula*	+	+	+	+							+		+	
8	AW586959	*M. truncatula*	+	+	+	+							+		+	
9	AW775229	*M. truncatula*	+	+	+	+							+		+	
10	BF005356	*M. truncatula*	+	+	+	+							+		−	
11	BF649209	*M. truncatula*	+	−	−	−							+		+	
12	AW684360	*M. truncatula*	+	+	+	+							+		+	
13	AW685679	*M. truncatula*	−	−	+	−							+		−	
14	AI974841	*M. truncatula*	+	+	+	+							+		−	
15	AW688216	*M. truncatula*	+	+	+	+							+		−	
16	AW127626	*M. truncatula*	+	+	+	+							+		−	
17	AW688861	*M. truncatula*	+	+	+	+							+		−	
18	AG81	*G. max*	+	+	+	+							+		−	
19	AQ842128	*G. max*	+	+	+	+							+	+	+	
20	AW186493	*G. max*	−	−	−	−							+	+	+	
21	GMENOD2B	*G. max*	−	−	−	−							−	−	−	
22	AG48	*G. max*	−	−	−	−							−	+	−	
23	AG50b	*G. max*	+	+	+	+							+	+	+	
24	AG93	*G. max*	−	−	−	−							−	+	−	
25	BE347343	*G. max*	−	−	−	−							+	+	−	
26	SoyPRP1	*G. max*	+	+	+	+							−	+	−	
27	AF186183	*G. max*	+	+	+	+							+	+	+	
28	AW277661	*G. max*	−	−	−	+							+	+	−	
29	GMABAB	*G. max*	+	+	+	+							−	+	−	
30	BE801128	*G. max*	−	−	−	−							+	+	+	
31	AW508247	*G. max*	+	+	+	+							+	+	+	
32	AW620774	*G. max*	−	−	−	−							−	+	−	
33	AQ917190	*M. truncatula*	+	+	+	+							+		+	
34	AQ917191	*M. truncatula*	+	+	+	+							+		+	
35	AQ917144	*M. truncatula*	+	+	+	+							+		+	
36	AQ917298	*M. truncatula*	−	−	−	−							+		−	
37	AQ917327	*M. truncatula*	−	−	−	−							+		−	
38	AQ917338	*M. truncatula*	−	−	−	−							+		−	
39	AQ917375	*M. truncatula*	−	−	−	−							+		−	
40	AQ917398	*M. truncatula*	−	−	−	−							+		−	
41	AM040737_5	*P. sativum*	−	−	−	−							−		+	
42	AM040737	*P. sativum*	−	−	−	−							−		+	

Table 3. *Continued*

Primer pair no.	Sequence ID	Donor species	Lablab purpureus										Legume controls			
			HA-3	Mac-1	SRL	Rongai	CPI 36903	High-worth	CPI 52544	ILRI 6536	CPI 52552	CPI 52535	Mt	Gm	Ps	Vu
43	AJ251536_5	*P. sativum*	−	−	−	−							−	+	+	+
44	AB176566	*L. purpureus*	−	−	−	−							−	+	−	+
45	AB176567	*L. purpureus*	+	+	+	+							−	+	−	+
46	AF067417	*L. purpureus*	+	+	+	+							+	+	−	+
47	AY583516	*L. purpureus*	+	+	+	+							+	+	+	+
48	AY049046	*L. purpureus*	+	+	+	+							+	+	+	+
49	AY049047	*L. purpureus*	+	+	+	+							+	+	+	+
50	AF077224	*V. unguiculata*	+	+	+	+							+	+	+	−
51	X89400	*V. unguiculata*	+	+	+	+							+	+	+	−
52	AF279252	*V. radiata*	+	+	+	+							+	+	+	+
53	AF151961	*V. radiata*	+	+	+	+							+	+	+	+
54	AF402602	*P. vulgaris*	+	+	+	+							+	+	+	−
55	D13557	*V. radiata*	+	+	+	+							+	+	+	+
56	AY189907	*P. vulgaris*	−	−	−	−	−	+	−	−	+	+				−
57	U85250	*V. unguiculata*	−	−	−	−	−	−	−	−	−	−				−
58	BQ481569	*P. vulgaris*	−	−	−	−	−	−	−	−	−	−				−
59	AB056453	*V. unguiculata*	−	+	−	−	+	+	+	+	+	+				−
60	AB062360	*P. lunatus*	+	−	−	−	−	−	−	−	−	−				−
61	CA901109	*P. coccineus*	−	+	−	−	+	+	+	+	+	+				+
62	CA901635	*P. coccineus*	−	−	−	−	−	−	−	−	−	−				+
63	CA901660	*P. coccineus*	+	−	−	+	+	+	+	+	+	+				+
64	CA910054	*P. coccineus*	−	−	−	−	−	−	−	−	−	−				+
65	CA912710	*P. coccineus*	−	+	+	−	+	+	+	+	+	+				+
66	CA909621	*P. coccineus*	−	+	+	−	+	+	+	+	+	+				+
67	D10266	*V. radiata*	−	+	+	+	+	+	+	+	+	+				+
68	CK151423	*V. unguiculata*	+	+	+	+	+	+	+	+	+	+				+
69	CB968060	*V. unguiculata*	+	+	+	+	+	+	+	+	+	+				+
70	CA914593	*P. coccineus*	+	+	+	+	+	+	+	+	+	+				+
71	CA906101	*P. coccineus*	+	+	+	+	+	+	+	+	+	+				+
72	CA908729	*P. coccineus*	+	+	+	+	+	+	+	+	+	+				+
73	AF165998	*V. unguiculata*	+	+	+	+	+	+	+	+	+	+				+
74	U08140	*V. radiata*	+	+	+	+	−	+	+	+	+	+				+
75	U10419	*P. vulgaris*	+	+	+	+	−	+	+	+	+	+				+
76	U28645	*P. vulgaris*	+	+	+	−	−	+	−	+	+	+				+
77	U54703	*P. vulgaris*	+	+	+	+	−	+	+	−	+	−				+
78	U70531	*P. vulgaris*	+	+	−	−	−	−	+	+	−	−				+
79	U92656	*V. unguiculata*	−	−	+	−	+	+	−	−	−	+				+
80	X53603	*P. vulgaris*	+	+	+	−	−	+	+	+	+	+				+
81	U20809	*V. radiata*	+	+	+	−	−	+	+	+	+	+				+
82	BQ481672	*P. vulgaris*	+	+	+	−	−	+	+			+				+

Table 3. *Continued*

Primer pair no.	Sequence ID	Donor species	Lablab purpureus										Legume controls			
			HA-3	Mac-1	SRL	Rongai	CPI 36903	High-worth	CPI 52544	ILRI 6536	CPI 52552	CPI 52535	Mt	Gm	Ps	Vu
83	CA900138	*P. coccineus*	−	−	−	−	−	−	−	−	−	−				−
84	CA898279	*P. coccineus*	+	+	+	−	−	+	+	−	+	+				+
85	CA898807	*P. coccineus*	+	+	+	−	−	+	+	−	+	+				+
86	CA898809	*P. coccineus*	+	+	+	+	+	+	+	+	+	+				+
87	CA899061	*P. coccineus*	+	+	+	−	−	+	+	−	+	+				+
88	CA899306	*P. coccineus*	+	+	+	−	−	+	+	−	+	+				+
89	CA899524	*P. coccineus*	+	+	+	+	−	+	+	+	+	+				+
90	AY193836	*V. unguiculata*	+	+	+	−	−	+	+	−	+	+				+
91	CA901208	*P. coccineus*	+	+	+	+	+	+	+	+	+	+				+
92	CA902017	*P. coccineus*	+	+	−	+	+	+	+	+	+	+				+
93	CA902145	*P. coccineus*	+	+	+	+	−	+	+	+	+	+				+
94	CA902154	*P. coccineus*	+	+	−	+	−	+	+	+	+	+				+
95	CA906247	*P. coccineus*	+	−	−	+	−	−	−	−	−	+				+
96	CA906300	*P. coccineus*	−	−	−	−	−	−	−	−	−	−				+
97	BQ481774	*P. vulgaris*	+	+	+	−	−	+	+	+	+	+				+

Blank, not tested; +, band amplified; −, line tested but no amplification. Where there was amplification: no shading, only one band amplified; light shading, double bands; dark shading, multi-bands ≥ 3.

Table 4. Summary of amplification success of the sequence-specific markers

Donor species	Sequence class[a]	Number of primer pairs tested	L. purpureus amplification		L. purpureus[b]: number of PP with			Reference
			Success	%	S	D	M	
M. truncatula	1, 2	17	14/17	82	9[4]	–	5[0]	Wang *et al.* (2004)
G. max	2–5	15	8/15	53	4[1]	–	4[0]	Wang *et al.* (2004)
M. truncatula	6	8	3/8	37.5	3[2]	–	–	Choi et al. (2004a)
P. sativum	9	3	0/3	0	–	–	–	This study
L. purpureus	7, 8	6	5/6	83	4[2]	1[0]	–	This study
Phaseoleae[c] tribe	7, 9	48	38/48	79	23[16]	8[3]	7[0]	This study
	Total	97	68	70	43[25]	9[3]	16[0]	

[a] See bottom of Table 2 for the sequence class from which the marker is derived.
[b] The *L. purpureus* banding pattern obtained per primer pair (PP): S = single band; D = double band; M = multi-band ($\geqq$ 3); superscripts in this column refer to the number of primer pairs where the amplified *L. purpureus* band is shared with a control or another legume species; in the case of the six *L. purpureus* sequence primers, band sharing is with *M. truncatula*.
[c] See Table 2, primer pairs 50–97, for the list of legumes tested from the *Phaseoleae* tribe.

amplified, and the expected band size obtained, with all the primer pairs tested (Table 3). For the 48 primer pairs from sequence classes 7 and 9 (Table 2), designed from a variety of legumes of the *Phaseoleae* tribe (primer pairs 50–97, Table 3), 79% gave amplification products with *L. purpureus*: 48% being single and 17% double band amplifications (Table 4).

There are only a few *L. purpureus* sequences in the nucleotide databases. Of the six *L. purpureus* primer pairs tested (44–49, Table 3), four produced a single band, one a doublet, and one failed to amplify from *L. purpureus* DNA. The one failure may have been due to the region chosen for primer design coinciding with an exon/intron boundary, and this may also explain some of the failed amplifications from other primer sets. Amplification of *G. max* using the set of six *L. purpureus*-derived primer pairs resulted in two cases each of single, doublet or multi-band patterns (Table 3); four of the six gave either the same size band on 1.5% agarose or a reasonably strong band of a similar size as *L. purpureus*. This corresponds to 67% of *L. purpureus* primer sequences able to amplify *G. max*. Similarly, in the other two species where a single band of the same size was produced, as assessed on 1.5% agarose, 33% of *L. purpureus* primer pairs amplified *M. truncatula* (48 and 49, Table 3) and 17% amplified *P. sativum* (48, Table 3). The *L. purpureus* DNA primer pair 47 (Table 3) amplified a fragment that corresponds to 5.8S rRNA sequence flanked by internal spacer sequences, and produced a multiple band pattern from the three other legumes tested (Table 3).

One assumption from scoring the agarose gels was that similar band sizes across those legumes tested were amplifications from orthologous genomic sequences. Where there was amplification with *L. purpureus* DNA (Table 4) 43/68 (63%) primer pairs produced a single band and each of these was regarded as allelic to the others. From examination of 1.5% agarose gels, in 25 of the 43 cases the amplification product was the same band size with at least one of the other legumes tested. Double bands were found in 9 of the 68 and in three cases this doublet was shared with at least one other legume (Table 4). Sixteen primer pairs gave a multi-band pattern (Table 4) and in all cases the *L. purpureus* pattern was generally identical but on occasion a band was sometimes missing/gained from one or other of the four lines tested. The multi-band pattern generated with *L. purpureus* was never the same as any of the other legumes. This suggests a degree of non-specific amplification and any polymorphic marker from this type of reaction is much like a RAPD (Williams *et al.*, 1990).

In one case the *G. max* AF186183 (number 27, Table 3) primer pair gave a multi-band/smear pattern over the size 150–500 bp range in *L. purpureus*, *M. truncatula*, *P. sativum* and *G. max*. From the database, AF186183 contains sequence related to (but incomplete) the *Calypso-2-1* retroelement (normally *c.* 11 kb) found in abundance in *G. max* (Wright and Voytas, 2002). This primer pair spans a region within the polyprotein of *Calypso*; the smear very likely comes about from the repetitive nature of the retroelement amplified from AF186183, so generating many copies. This sequence information could be adapted to obtain SSAP (sequence specific amplified polymorphism) markers (Ellis *et al.*, 1998) as a valuable addition to molecular marker types transferable between legumes.

Polymorphism within L. purpureus *from sequence-specific markers*

Single or a few base changes would not be obvious from 1.5% agarose and further testing using more stringent conditions (3% agarose), the use of PAGE or SSCP (single strand conformational polymorphism), or sequencing

would be required to resolve polymorphisms. Where a single band only was amplified with *L. purpureus* and that was also shared with the control legume or another legume, i.e. the same size assessed on agarose (Table 3), about one-third (5/13, Table 4) were from primer pairs generating SSR-containing markers (sequence classes 1–4, Table 2) compared to two-thirds (20/30, Table 4) from the non-SSR containing EST and genomic DNA sources (sequence classes 6−9, Table 2).

From the primer pairs that did amplify *L. purpureus* DNA (Table 3) there were no obvious polymorphic differences, i.e. no co-dominant markers, between the three *L. purpureus* parental lines, HA 3, MAC 1 and SR L used in crosses for breeding purposes, and cv. Rongai on agarose gels (primer pairs 1−55). However, there were a few cases within the primer pairs 56−97 where at least one *L. purpureus* failed to amplify (Table 3), this could be technical failure of PCR but these were consistent observations. These dominant polymorphisms correspond to mismatch at the primer binding sites, in which case SNPs or length variants could be pursued as markers. Some of the successful primer combinations gave a multi-band pattern with *L. purpureus* and a number have been tested on PAGE and visualized using silver stain. There were few polymorphic differences, but these also appear to be dominant markers: much like those generated using AFLP. The five *L. purpureus* specific primers (Table 3, 45−49) that successfully amplified were tested using both 1.5% and 3% agarose; all band sizes were identical in the four lines.

Diversity of L. purpureus *in relation to cowpea as an outgroup*

For 18 primer pairs (56−97, Table 3) there were 21 amplification products from cowpea (*Vigna unguiculata*)

DNA which had the same mobility in agarose gels as at least one of the amplification products from *L. purpureus* DNA, with an additional ten products being unique to cowpea (Table 3). The genetic distance matrix of these 31 marker band scores is shown as a PCA (Fig. 3a) where there are three distinct groupings. The outgroup cowpea constitutes one group near to the *x*-axis, the seven Indian-derived *L. purpureus* cluster within the hashed oval, close to this comes ILRI 6536 Ex. Ethiopia; the African line cv. Rongai from Kenya and CPI 36 903 Ex. Ukraine cluster together at the top right. This spread of variation based on Nei's genetic distance (D) is as expected from a data set that includes an outgroup from a differing but related species. The genetic distance between *L. purpureus* and cowpea is probably an underestimate in this assay because differences between the two species were based on the presence and absence of a band rather than nucleotide differences. This explains why there appears to be just as much variation between cowpea and either of the two *L. purpureus* groupings as there is between the cluster of eight *L. purpureus* at the extreme left of the PCA (Fig. 3a) and the two non-Indian *L. purpureus* at the top right.

This overall pattern of diversity within *L. purpureus* from Fig. 3a is maintained for the *L. purpureus* data alone without cowpea (Fig. 3b). The PCA for the ten *L. purpureus* (Fig. 3b) taken from Table 3 data comprises 35 markers from 32 primer combinations with information content (pairs 56−97). This PCA shows a similar pattern of diversity within the *L. purpureus* set of Fig. 3a and with the AFLP phylogeny (Fig. 2): the three UAS parental lines HA 3, MAC 1 and SR L all cluster closely. However, cv. Highworth (Ex. India, introduced to Australia via Kenya) – a cultivar that has adapted to Australian conditions, is widely used for forage (Liu, 1996; Pengelly and Maass, 2001) and is not known
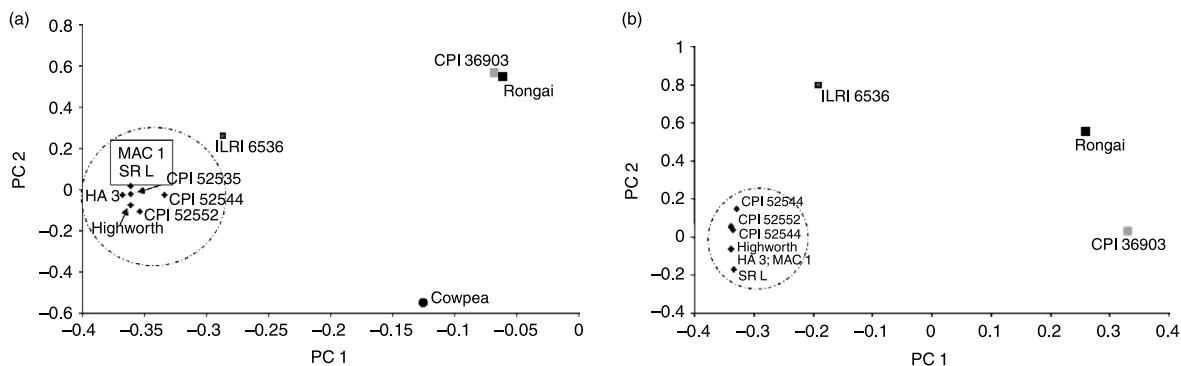
Fig. 3. Principal component analysis (PCA) of the molecular diversity obtained from the sequence-specific markers: (a) based on 18 primer pairs and 21 amplification products from cowpea (*Vigna unguiculata*) DNA having the same mobility in agarose gels as at least one of the ten *Lablab purpureus*. It must be noted that the genetic divergence of cowpea in relation to *L. purpureus* is underestimated in this PCA as it is strongly biased by scoring amplification products without regard to sequence variation. (b) The PCA for ten *L. purpureus* based on 32 primer pairs and 35 marker bands (Table 3 data, primer pairs 56−97).

to be cultivated in India – clusters very closely with the UAS lines in this sequence-specific marker analysis, compared to more distantly in the AFLP analysis (Fig. 2). The three other Indian accessions, CPI lines 52535, 52552 and 52544 cluster in this Indian section of the plot (Fig. 3b). The Ex. Ethiopian line ILRI 6536 tends to be consistently within or just peripheral to the major *L. pur-pureus* cluster in all the tests (Figs 2 and 3a, b); in contrast, the Ex. Ukraine line CPI 36 903 tends to be distant to the Indian accessions. The three UAS parental breeding lines, SR L, HA 3 and MAC 1, always cluster tightly together regardless of the marker assay.

Discussion

AFLP and sequence-specific markers: general overview

Two marker types were used in this study, AFLP and sequence-specific markers. These different approaches address variation in different components of the genome, so differences in their behaviour may be informative.

Many plant phylogenetic studies have been carried out using AFLP (Mba and Tohme, 2005), and this marker method provides a rapid and reproducible banding pattern that gives an indication of relatedness and diversity within a species. The AFLP marker analysis provides genotypically distinct groupings within related species; molecular data from this type of marker are described by Karp *et al.* (1996) as 'arbitrary indicators of diversity'. AFLP, however, is not the ideal marker for diversity analyses, mostly because AFLP polymorphic bands are generally dominant markers; where for each marker only one allele is apparent. Also, co-migration of coincident rather than allelic bands may introduce errors; these difficulties have not deterred its use in phylogenetics (Karp *et al.*, 1996; Mba and Tohme, 2005).

Many of the markers from the 97 sequence-specific primer pairs (Table 2), on the other hand, are from coding regions of known conserved genes. When gene-related markers are co-dominant they indicate allelic variation providing information not unlike RFLPs. As they generally represent conserved coding regions, gene-related markers are transferable within and between species and genus (Karp *et al.*, 1996); when mis-matching at the priming site occurs, these will be dominant markers.

Diversity analysis of L. purpureus *from the two marker types*

All of the genera studied here are within the *Papilionoideae* sub-family that contains most of the cultivated crop legumes (Young *et al.*, 2003). *L. purpureus* along with *Vigna* spp. and *Phaseolus* spp. are in the *Phaseoleae* tribe (Lackey, 1981), a member of the Millettoid clade (Cronk *et al.*, 2006). This clade also includes the sub-tribe *Glycinineae*, of which *G. max* is a member (Lackey, 1981). Pea, which is within the *Vicieae* tribe, and *Medicago* within the *Trifolieae* tribe (Doyle, 1995; Wojciechowski, 2003), are members of the Galegoid clade (Cronk *et al.*, 2006). In this study sequence-specific markers from species representing the three tribes, *Phaseoleae* (including sub-tribe *Glycinineae*), *Vicieae* and *Trifolieae* were tested for *L. purpureus* DNA amplification.

Assessments of relatedness within a species would be expected to have broad agreement of groupings no matter which marker type is employed, provided it is reliable and reproducible. In fact, the two marker types, AFLP and sequence-specific, used in this diversity analysis resulted in broad agreement even though differing genomic regions were examined and fewer *L. purpureus* accessions were tested with the gene-related markers.

The diversity analysis using AFLP markers clearly showed that there was limited genetic variation within the UAS and locally collected *L. purpureus* germplasm from the southern states of India compared to the genus as a whole. Limited genetic diversity at the molecular level is not uncommon within some germplasm stocks of legumes (chickpea: Udupa *et al.*, 1993; Winter *et al.*, 2000; pigeonpea: Rao *et al.*, 2003; peanut: Kochert *et al.*, 1996). However, the AFLP did manage to distinguish between Indian cultivars and breeding stock held within the UAS collection, and the other Indian accessions held within the Australian (ATCGRC) collection (Fig. 2: accessions CPI 52544, CPI 52535, CPI 52552). But the sequence-specific primer pairs failed to do this (Fig. 3b), all seven Indian accessions that comprised UAS and CPI/ILRI lines, clustered together with this marker type.

In both diversity analyses the non-Indian accessions were more distinct (Figs 2 and 3a, b), and this was quite marked, but dependent on the markers used to score the data – see differences between Figs 3a and 3b for cv. Rongai and CPI 36903. Compared to AFLP, the sequence-specific marker type seems to be assessing diversity within more conserved regions of the genome.

Apart from the commonly cultivated accessions, more genetically diverse lines within *L. purpureus* species do exist and are available for integration of exotic alleles.

Amplification success of sequence-specific markers from a range of legume species

The sharing of markers within the legumes for comparative and quantitative trait loci (QTL) mapping have been

documented (Young *et al.*, 2003; Zhu *et al.*, 2005) and considerable conservation and synteny between model and crop genomes have been found (Young *et al.*, 2003; Choi *et al.*, 2004b; Kalo *et al.*, 2004; Zhu *et al.*, 2005). In QTL-mapping, RFLP markers within the QTL regions, conferring the greatest effect on seed weight, in both cowpea and mungbean, had the same order (Fatokun *et al.*, 1992). In a further QTL analysis, an orthologous seed weight gene of soybean was found to be shared with cowpea and mungbean (Maughan *et al.*, 1996). Similarly, Timmerman-Vaughan *et al.* (1996) found the genomic regions containing seed weight QTLs in cowpea, mungbean and pea, to be highly conserved. Together these three studies suggest conservation of orthologous genes within these four species for the polygenic seed weight trait. RFLP linkage mapping showed mungbean and common bean (*P. vulgaris*) to have a high degree of synteny compared to either of these two species with soybean (Boutin *et al.*, 1995). A study of co-linearity between mungbean and *L. purpureus* (Humphry *et al.*, 2002), using a common set of 65 RFLP, showed there was a high degree of conservation in marker order across 10 of the 11 linkage groups of mungbean, corresponding to 15 *L. purpureus* groups.

This study centred on *L. purpureus* diversity, and the potential use of sequence-specific markers from species related to *L. purpureus*. The results suggest that there is a good source of legume-related sequences in databases readily available for *L. purpureus* diversity and genome analysis. The data presented here also agree with recent primer transferability studies involving *L. purpureus* (Wang *et al.*, 2004). The transferability of EST-based SSR markers from *M. truncatula* to *G. max* was 55%, compared to 28% the other way round. Wang *et al.* (2004) attributed this to the twofold genome size difference between these species, which are distantly related and are within differing legume tribes (Doyle, 1995). Similarly, in our study, 67% of *L. purpureus* primer pairs amplified *G. max* compared to 53% in the other direction. The *G. max* genome is substantially duplicated (Young *et al.*, 2003), and it is thought that it has undergone considerable rearrangement during its diploidization from an ancient polyploid (Boutin *et al.*, 1995). These factors, coupled with the degree of relatedness and genome size differences, may explain the lower amplification rate of *G. max* sequence-specific primer pairs to *Medicago* and *L. purpureus*.

In this study, 82% of *M. truncatula* and 53% of *G. max* (Table 4) amplified *L. purpureus*. The figure of 82% amplification success is biased, as the 17 *Medicago* EST-based SSR primer pairs (Table 4) tested in this study were chosen at the suggestion of the authors of the Wang *et al.* (2004) paper and known to previously amplify

L. purpureus. Amplification of *L. purpureus* from 3/8 (37.5%) primer pairs (all *M. truncatula* sequences) from Choi *et al.* (2004a), gives a figure consistent with the 36% reported by Wang *et al.* (2004). This degree of amplification success using heterologous primer pairs is also in agreement with other studies using genic SSRs from *M. truncatula*, the success rates for amplification of pea, faba bean and chickpea were 37.6%, 40% and 36.6%, respectively (Gutierrez *et al.*, 2005).

The same set of 15 *G. max* primers was used in both this study and that of Wang *et al.* (2004). In this study 53% of *G. max* primer pairs amplified *L. purpureus* (Table 4) compared to 23% in Wang *et al.* (2004); single bands were amplified in 26.5% (4/15, Table 4) of cases – more in keeping with the Wang *et al.* (2004) result – and 26.5% gave multiple bands; there were no cases of double bands (Table 4).

The *P. sativum* primers failed in *L. purpureus*. These plants are from distantly related genera and differ markedly in genome size, so this lack of success with a small sample of primers is not surprising, given the success rate of the *Medicago* primers. Though *L. purpureus* and *M. truncatula* are from different tribes, their haploid genome size is similar, 3.68×10^8 and 4.66×10^8 bp, respectively (http://www.rbgkew.org.uk/cval).

The 48 sequence-specific primer pairs 50–97 (Table 3) successfully generated amplification products from *L. purpureus* in 79% of cases, with 65% being single/double band amplifications: the donor species were *Vigna unguiculata*, *V. radiata*, *Phaseolus vulgaris*, *P. lunatus* and *P. coccineus* (Table 3). All of these species have genomes of a similar size ($c. 5 \times 10^8$–8×10^8 bp; http://www.rbgkew.org.uk/cval) and all are relatively closely related within the *Phaseoleae* tribe. This high rate of amplification success between *L. purpureus* and these closely related species is of a similar order to that found by Eujayl *et al.* (2004) where 74% of EST-derived SSR primer pairs from *M. truncatula* gave amplification of the expected band size in at least one other *Medicago* species.

The amplification success of the six sequence-specific markers from *L. purpureus* to *G. max*, *M. truncatula* and *P. sativum* was 67%, 33% and 17%, based on single band amplification.

Molecular marker analyses, genotyping and phenotyping of germplasm, together with the incorporation of exotic and underutilized germplasm into breeding programmes, offers new potential for enhancing desired traits in breeding programmes. In this study we have shown that molecular markers can be used successfully to understand the genetic structure of the germplasm of *L. purpureus*, and identified a series of sequence-specific markers that potentially could be used in *L. purpureus* genetic mapping analyses.

Acknowledgements

References

Basavarajappa PS and Byre Gowda M (2000) Genetic divergence among field bean (*Lablab purpureus* L. Sweet) cultivars of Southern Karnataka. *Indian Journal of Plant Genetic Resources* 13: 134–137.

Bassam BJ and Caetano-Anollés G (1993) Silver staining of DNA in polyacrylamide gels. *Applied Biochemistry and Biotechnology* 42: 181–188.

Boutin SR, Young ND, Olson TC, Yu ZH, Shoemaker RC and Vallejos CE (1995) Genome conservation among 3 legume genera detected with DNA markers. *Genome* 38: 928–937.

Champ MM-J, Anderson JW and Bach-Knudsen K-E (eds) (2002) Pulses and human health. *British Journal of Nutrition* 88 (Suppl 3): S237–S319.

Choi H-K, Kim D, Uhm T, Limpens E, Lim H, Mun J-H, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB and Cook DR (2004a) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. *Genetics* 166: 1463–1502.

Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND and Cook DR (2004b) Estimating genome conservation between crop and model legume species. *Proceedings of the National Academy of Sciences, USA* 101: 15289–15294.

Cronk Q, Ojeda I and Pennington RT (2006) Legume comparative genomics: progress in phylogenetics and phylogenomics. *Current Opinion in Plant Biology* 9: 99–103.

Doyle JJ (1995) DNA data and legume phylogeny: a progress report. In: Crisp M and Doyle JJ (eds) *Advances in Legume Systematics 7: Phylogeny*. London: Royal Botanic Gardens, Kew, pp. 11–30.

Ellis THN, Davies DR, Castleton JA and Bedford ID (1984) The organization and genetics of rDNA length variants in peas. *Chromosoma* 91: 74–81.

Ellis THN, Poyser SJ, Knox MR, Vershinin AV and Ambrose MJ (1998) Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. *Molecular and General Genetics* 260: 9–19.

Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC and Mian MAR (2004) *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. *Theoretical and Applied Genetics* 108: 414–422.

Fatokun CA, Menanciohautea DI, Danesh D and Young ND (1992) Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. *Genetics* 132: 841–846.

Graham PH and Vance CP (2003) Legumes: importance and constraints to greater use. *Plant Physiology* 131: 872–877.

Gur A and Zamir D (2004) Unused natural variation can lift yield barriers in plant breeding. *PLoS Biology* 2: 1610–1615.

Gutierrez MV, Patto MCV, Huguet T, Cubero JI, Moreno MT and Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. *Theoretical and Applied Genetics* 110: 1210–1217.

Humphry ME, Konduri V, Lambrides CJ, Magner T, McIntyre CL, Aitken EAB and Liu CJ (2002) Development of a mungbean (*Vigna radiata*) RFLP linkage map and its comparison with lablab (*Lablab purpureus*) reveals a high level of colinearity between the two genomes. *Theoretical and Applied Genetics* 105: 160–166.

Kalo P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis THN and Kiss GB (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. *Molecular Genetics and Genomics* 272: 235–246.

Karp A, Seberg O and Buiatti M (1996) Molecular techniques in the assessment of botanical diversity. *Annals of Botany* 78: 143–149.

Kochert G, Halward T and Stalker HT (1996) Genetic variation in peanut and its implications in plant breeding. In: Pickersgill B and Lock JM (eds) *Advances in Legume Systematics 8: Legumes of Economic Importance*. London: Royal Botanic Gardens, Kew, pp. 19–30.

Lackey JA (1981) Tribe 10. Phaseoleae DC. In: Polhill RM and Raven PH (eds) *Advances in Legume Systematics, Part 1*. London: Royal Botanic Gardens, Kew, pp. 301–306.

Ladizinsky G (1998) *Plant Evolution under Domestication*. London: Kluwer Academic Publishers.

Liu CJ (1996) Genetic diversity and relationships among *Lablab purpureus* genotypes evaluated using RAPD as markers. *Euphytica* 90: 115–119.

Lu J, Knox MR, Ambrose MJ, Brown JKM and Ellis THN (1996) Comparative analysis of genetic diversity in pea assessed by RFLP- and PCR-based methods. *Theoretical and Applied Genetics* 93: 1103–1111.

Maass BL (2006) Changes in seed morphology, dormancy and germination from wild to cultivated hyacinth bean germplasm (*Lablab purpureus*: Papilionoideae). *Genetic Resources and Crop Evolution* 53: 1127–1135.

Maass BL, Jamnadass RH, Hanson J and Pengelly BC (2005) Determining sources of diversity in cultivated and wild *Lablab purpureus* related to provenance of germplasm by using amplified fragment length polymorphism. *Genetic Resources and Crop Evolution* 52: 683–695.

Mahadevu P and Byre Gowda M (2005) Genetic improvement of Dolichos bean (*Lablab purpureus* (L.) Sweet) through use of exotic and indigenous germplasm. *Indian Journal of Plant Genetic Resources* 18: 47–48.

Maughan PJ, Maroof MAS and Buss GR (1996) Molecular-marker analysis of seed weight: genomic locations, gene action, and evidence for orthologous evolution among three legume species. *Theoretical and Applied Genetics* 93: 574–579.

Mba C and Tohme J (2005) Use of AFLP markers in surveys of plant diversity. *Methods in Enzymology* 395: 177–201.

Murphy AM and Colucci PE (1999) A tropical forage solution to poor quality ruminant diets: a review of *Lablab purpureus*. *Livestock Research for Rural Development* 11: 1–17.

Palmer JD, Jorgensen RA and Thompson WF (1985) Chloroplast DNA variation and evolution in *Pisum* – patterns of change and phylogenetic analysis. *Genetics* 109: 195–213.

Pengelly BC and Maass BL (2001) *Lablab purpureus* (L.) Sweet – diversity, potential use and determination of a core collection of this multi-purpose tropical legume. *Genetic Resources and Crop Evolution* 48: 261–272.

Rao NK, Reddy LJ and Bramel PJ (2003) Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution* 50: 707–721.

Timmerman-Vaughan GM, McCallum JA, Frew TJ, Weeden NF and Russell AC (1996) Linkage mapping of quantitative trait loci controlling seed weight in pea (*Pisum sativum* L). *Theoretical and Applied Genetics* 93: 431–439.

Udupa SM, Sharma A, Sharma RP and Pai RA (1993) Narrow genetic variability in *Cicer arietinum* L. as revealed by RFLP analysis. *Journal of Plant Biochemistry and Biotechnology* 2: 83–86.

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP – a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.

Wang ML, Gillaspie AG, Newman ML, Dean RE, Pittman RN, Morris JB and Pederson GA (2004) Transfer of simple sequence repeat (SSR) markers across the legume family for germplasm characterization and evaluation. *Plant Genetic Resources* 2: 107–119.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531–6535.

Winter P, Benko-Iseppon AM, Hüttel B, Ratnaparkhe M, Tullu A, Sonnante G, Pfaff T, Tekeoglu M, Santra D, Sant VJ, Rajesh PN, Kahl G and Muehlbauer FJ (2000) A linkage map of the chickpea (*Cicer arietinum* L.) genome based on recombinant inbred lines from a *C. arietinum* × *C. reticulatum* cross: localization of resistance genes for fusarium wilt races 4 and 5. *Theoretical and Applied Genetics* 101: 1155–1163.

Wojciechowski MF (2003) Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective. In: Klitgaard BB and Bruneau A (eds) *Advances in Legume Systematics, Part 10, Higher Level Systematics*. London: Royal Botanic Gardens, Kew, pp. 5–35.

Wright DA and Voytas DF (2002) *Athila*4 of *Arabidopsis* and *Calypso* of soybean define a lineage of endogenous plant retroviruses. *Genome Research* 12: 122–131.

Young ND, Mudge J and Ellis THN (2003) Legume genomes: more than peas in a pod. *Current Opinion in Plant Biology* 6: 199–204.

Zhu HY, Choi HK, Cook DR and Shoemaker RC (2005) Bridging model and crop legumes through comparative genomics. *Plant Physiology* 137: 1189–1196.