THEORETICAL
BIOLOGY FORUM

TBF : 01–13 (2022)
RESEARCH ARTICLE

Development of novel SNP/InDel markers through amplicon sequencing in dolichos bean (*Lablab purpureus* L.)

Venkatesha S.C[1] and Prakash G. Patil [2*]

[1]Department of Biotechnology and Crop Improvement, University of Horticultural Sciences (UHS), Bagalkot, Karnataka 587104, India
[2]ICAR-National Research Centre on Pomegranate (NRCP), Solapur 413 255, India

ABSTRACT

Despite the advances made through genomics research in Dolichos bean, the availability of informative molecular tools is still limited. The expressed sequence tags represent the robust sequence resource for the development of novel genic markers that can be exploited for gene discovery, genome annotation and comparative genomics. Here we first time report the development of novel EST-derived SNP and InDel markers through amplicon sequencing of cross-transferable cowpea EST-SSR primers in Lablab bean. Total of four polymorphic cowpea EST-SSRs viz., CP2, CP8, CP29, and CP435, with a high transferability rate in Dolichos bean were identified. Sequenced total 18 genic fragments from nine parental genotypes, which resulted in the identification of 59 SNPs and 16 InDels. Based on these sequence variations, we successfully designed and validated eight allele-specific markers i.e. Met 2, AF 287258-1, AF 285278-2, AF 151961, D 13557, CP 5, CP16 and CP 43, which are targeting SNPs/InDels in Lablab bean. Further, to prove their immediate utility for trait mapping the parental polymorphic markers were genotyped on F2 segregating lines derived from the two crosses i.e. HA4 × CPI31113 and HA4 × CPI60216. Segregation analysis revealed that these markers exhibited a typical Mendelian segregation ratio of 1:2:1 with co-dominant allelic patterns for SNP/InDels. This confirmed the immediate utility of such novel genic markers for future genomics-assisted breeding applications in Dolichos bean.

Keywords: *Genic fargments, Lablab bean, Sequencing, SNP/InDel markers.*

ARTICLE HISTORY:
Received: 25 July 2022
Revised: 06 September 2022
Accepted: 28 November 2022
Available: Online 06 December 2022

DOI:
https://doi.org/10.5281/zenodo.7629782

CORRESPONDING AUTHOR:
Prakash Goudappa Patil

E-MAIL ID:
patilbt@gmail.com

INTRODUCTION

Dolichos bean (*Lablab purpureus* L. Sweet, Fabaceae) is one of the oldest cultivated plants, with records dating back to 1500 BC in India [1]. The plant is presently grown throughout tropical regions in Asia and Africa either as mono-crop or mixed crops [2]. Lablab is a highly nutrient-dense legume crop, which is more resilient to biotic and abiotic stresses can thrive well under marginal areas, and is more vital to diversify diets, amplify sustainable farming, and mitigate climate change and its effects [3]. The pods are rich in proteins (22.4–31.3%) and carbohydrates (55%) and also contain vitamins and minerals such as copper, potassium, magnesium, iron, and phosphorus, are fibre-rich and are suitable for human nutrition [4]. Therefore, it is a multipurpose crop grown for pulses or vegetables for human consumption or as forage for animals [5].

For any crop improvement program, it is most vital to characterize for domestication and collection of the elite and promising genotypes of the Dolichos bean with high yielding potentialities for food and fodder purposes [2]. Yield, being a complex trait, is collectively influenced by various component characters, which are polygenically inherited and highly influenced by environmental variations. Therefore, wide phenotypic variability

was observed with little genetic diversity in the dolichos bean [6-8]. However, the use of inter-species crosses to produce lines with desirable traits, such as yield improvement or disease and pest resistance, has been one way to overcome the genetic bottleneck of crop plant domestication. This breeding approach, of inter- (or wide intra-) specific crosses, can benefit from good genetic maps, with robust markers anchored to reference genomes that are available from comparative genetic maps.

There are very few sequences available for lablab, but there is a vast amount of sequence data available from related legume species. Comparative mapping within the legume crop and model species provides evidence that this wealth of sequence data can be used for genetic analysis even in neglected legume species such as Dolichos [9]. Previous studies have shown lablab genome had a high syntenic relationship with *Vigna radiata* (mungbean) [10]. Till date, there is no well-saturated genetic map available and no trait-specific markers identified in dolichos bean. Therefore, a well-saturated genetic map based on genetic markers is a necessity for genomics-assisted breeding. Towards this end, the availability of informative marker resources in dolichos bean holds greater promise. Thus, the objective of the current study was to develop and validated gene-based SNP/InDel markers for genetic mapping work. For this we followed one of the highly targeted approaches called amplicon sequencing method, where in sequencing of PCR products has allowed many research in efficient variant detections in gene fragments followed by development of genic markers.

MATERIAL AND METHODS

Plant material and genomic DNA extraction

The experimental material constituted ten *Lablab purpureus* genotypes, which represented both exotic and indigenous origin along with one cowpea genotype as a control (Table 1). For validation of polymorphic markers two F_2 mapping populations were used. The experiment was laid out at department of plant biotechnology, University of Agricultural Sciences, Gandhi Krishi Vignana Kendra, Bengaluru. The seeds of these genotypes were obtained from the department of Genetic and Plant breeding, and the health plants were raised under plastic cover filled with red soil mixture having sand (2:1). For isolation of genomic DNA, two to three young leaves were collected from 25 day old seedlings. The extracted was performed by following SDS method [11], with minor modifications. The quality and concentration of all the DNA samples were assessed on 0.8 % agarose gels using uncut ʎ DNA as a standard.

Table 1: Details of 10 *Lablab* accessions used for screening gene specific markers

Sl. No.	Lablab accessions	Country of origin
1	HA3	India
2	Mac-1	India
3	SRL	India
4	Rongai	Kenya
5	Highworth	Ex. India
6	CPI36903	Ex. Ukrane
7	CPI52544	India
8	CPI52552	India
9	CPI52535	India
10	ILRI6536	Ex. Ethiopia

Primer designing and parental polymorphism study

We explored EST sequence information available on cowpea genomic database (http://cowpeagenomics.med.virginia.edu/CGKB), and designed 32 genic EST-SSR primers using Primer 3 tool. These primers could amplify the expected product size range of 200-300 bp for easy separation on agarose and acrylamide gel electrophoresis. In addition to this, 26 additional EST-SSR primers were obtained from Virginia University, USA. A total 58 EST-SSR primers were deployed for parental polymorphism study on nine lablab genotypes including one cowpea, which is listed in Table 2. The polymerase chain reaction (PCR) was performed in a 10µl reaction volume containing 20 ng genomic DNA as a template, 0.1 µM of each forward and reverse primer, 200 µM each dNTP, 0.5 unit of *Taq DNA pol* enzyme (Invitrogen, UK) and the appropriate reaction buffers. The reaction conditions were as follows: touch-down cycling regime with 94^0 C/3 min; (94^0 C/30 s; 58^0 C/30 s; 72^0 C/60 s) repeat for 10 cycles reducing the annealing temperature by 0.50 C/cycle; (94^0 C/30 s; 53^0 C/30 s; 72^0 C/60 s) repeat 34 cycles; 72^0 C/10 min; 12^0 C/30 min. The PCR products were separated on 2 % agarose

Table 2. Cowpea specific EST-SSR primers designed and deployed for screening on *Lablab* genotypes

Sl. No.	Primer Name	Sequence forward (5' to 3')	Sequence reverse (5' to 3')
1	CP1	TTGGAGAGGGACAAAATTGC	ATTGTGCCACGTGGTAGTGA
2	CP2	TCCTTTCTCTGCCTCAGGTG	GCTCTTTGCCTCCATTCTCC
3	CP3	ACCGGTGAAAGTAGGACCAC	GGCTGATTGATCTGGCCATC
4	CP4	GGAGTTAGCCCACTCTCGTTC	CAGAGCCTCTCCACTACTACAGC
5	CP5	AGCTCCTCATCAGTGGGATG	CATTGCCACCTCTTCTAGGG
6	CP6	GGGGGAGAGAGAGAGAGAGAGA	TTCTCCCCCTATGTGGACCT
7	CP7	GAGGAGGAGGAGGATCTGACA	CTTCTGCAGGCTTGTGGTTC
8	CP8	ACCATAGCAGCTCCCCATTC	GGTCTAGGACAAGGTCCTGGTAG
9	CP9	AGGAGGAGGAGGAGGAGGAG	GTCTGGGAGACCTGGATTAGC
10	CP10	AGGAGGAGGAGGAGGAGGAG	GTCTGGGAGACCTGGATTAGC
11	CP11	CCACAGTGGAAAACCACACC	GGTGGTGGTGGAGTGAAGAG
12	CP12	CACTCATCTGCCTCTCAACG	GCAGTCTACTCCCAGTACCACAC
13	CP13	CACCACCACTCCAAAGAAGC	CTAACGCTGCAGCCTCATCA
14	CP14	GGTTTATGGGTGGGAGGTGT	GTATCTGGCTGTCGTGACCA
15	CP15	ACGAGGCTCAAGATCAGACG	GGAGCTCTTCCTTGAGAGTGG
16	CP16	CACCCTTCACCACCGTTTAC	GTACACGAGGTGCTCCATTCC
17	CP17	ACCACCACCACTGTTCTTCTTC	TTACCTTCTGGACTGGGAGAGA
18	CP18	GTGTAAGAAGCATGCAGAGGTG	AGTACTAGAATGGGGCCATGTG
19	CP19	CTTCCCTGTGTTGCTGTTACTG	CTCGCACGAGTAGAAGAGAAGA
20	CP20	AAGTCCTTACACCACCACTCCA	GAGTCCGAAGAACGACTTGTG
21	CP21	CTCTCTCCATTCCCAAGTGTCT	GAAGAAGAACAGTGGTGGTGGT
22	CP22	AGATGTCGCTCCCTCTGTTG	GAGGAGGAGGAGGAGGAGAAG
23	CP23	CGAGGGAGTACGAGTTCAGC	GTCTCGACCTTGACGTCGTT
24	CP24	AGCGTCAGAGACTGGGAAGA	GAAGAGAGAGTGGTGGTGGTG
25	CP25	GGTGGCAATGAGCCTTACAC	AGATGTGGAGGGCATGGTAG
26	CP26	CGACCACCATAACTCGGAAG	GGGAGAAACAGAGGGAGAGAG
27	CP27	CGCAACTCACCTTTCCTCAG	CGTTGAGATGCAGGGAGAAC
28	CP28	CGCAACTCACCTTTCCTCAG	CGTTGAGATGCAGGGAGAAC
29	CP29	TCTCGAGAACGATGGTGGAG	AAGGACTGAAGCGAGAGACG
30	CP30	GCGTGTCTTGCCAACTTCTC	ATATCCTCCTCCTCCACCAC
31	CP31	GCTACATAGTCCCCACCAAGC	TGTCAGATCCTCCTCCTCCTC
32	CP32	CGACTCCACCTTGTTCTTGC	GGTGGTGGTGACAGAGTGAAG
33	CP115	GGGAGTGCTCCGGAAAGT	TTCCCTATGAACTGGGAGATCTAT
34	CP117	GTGGAAGGAATGGGTCCAG	AGGAAATTTGCATTCCCTTGT
35	CP163	CACTTTCTCCTAAGCACTTTTGC	AAGTGAAGCATCATGTTAGCC
36	CP171	GTAGGGAGTTGGCCACGATA	CAACCGATGAAAAAGTGGACA
37	CP181	GGGTGCTTTGCTCACATCTT	TCCATGTGTTTATGACGCAAA
38	CP197	TGAATGGAGCAACTTCTTGGA	GTTGCACTTGGTTGCCCTAT
39	CP201	GGTTTCCTAGTTGGGAAGGAA	ATTATGCCATGGAGGGTTCA
40	CP215	CAGAAGCGGTGAAAATTGAAC	GCATGTTGCTTTGACAATGG
41	CP239	CACCCCCGTACACACACAC	CACTTAAATTTTCACCAGGCATT
42	CP333	CAAAGGGTCATCAGGATTGG	TTTAAGCAGCCAAGCAGTTGT
43	CP359	TGAAAACAACGATATGCAGAAG	TCAGTCTTAGAATTGAGTTTTCTTCG
44	CP391	TGCCTATGCTTATGCCTGTG	GATGCCTGTTACTTGCCTTCT
45	CP395	GTTGTGAGCTTCCCCAGATG	AATTTTGAACCCACCACCAG
46	CP397	TCATGGGTTAAATTTGCTTCAA	AAACCATGTGGTTGTTGCAC
47	CP403	TGCAATATGGACCAGAAGAAA	ATGCCCCAACAACAACATTT
48	CP431	CCTCAACACCTTTTGGAAGGA	CAAATGCACCTCCTGTGCTA

Continued..

49	CP433	CAACTTCACAGCCCTCAA	TTGAAGGTATGGCCTTTTGTTT
50	CP435	TGCTCATCGTGCTTTGTCTT	CACTTCAGACTTAGAGCGAAGAA
51	CP443	GCTCGGATATGGTCCTGAAA	TCAGTGTCAGCACCATCCC
52	CP573	CAGAATCCTTGTGAACCTG	TTTCGCAATATGCCCTTTTC
53	CP605	AAAGAGATACACATGCCTAACA	GACCAACAGCGACTTTGAGC
54	Y1	GATATAGAATAGCATATTTAACATATTAG	GTTGAAAGTTTGATAGTAAAGTGG
55	Y21	GAGAACTTCACGCACAATAG	CGCGGTAGCATGATTGAATTTTG
56	Y26	CTAAATTATAATATTCGTCGGTC	GGTTAAGGAAAAGAGGGTAGG
57	Y31	CTATTGGAATCTTGCCGTTG	CTTTACCTTTATGCAAACCAATTC
58	Y45	CGATTATCCTGGCTAACGATG	GGATCTGAGATAGTGTGAC

gel electrophoresis, visualized and photographed using gel documentation unit. The primer products, which could not be resolved on agarose gel were separated on 5% horizontal polyacrylamide gel stained with Ethidium bromide or Silver nitrate staining as per the standard procedure [12].

Identification of SNP's and InDels through sequencing of genic fragments

The highly polymorphic cross transferable cowpea EST-SSRs were identified in Dolichos genotypes through parental polymorphism study. Total 23 primers with specific amplifications in Lablab accessions were selected (Table 3). The sequencing experiments were performed at JIC genome lab, Norwich, UK. For sequencing parental genotypes were first amplified using specific primers. PCR was performed in 10µl reaction volume by following touch-down cycling regime as already mentioned above for parental polymorphism study. The sequencing was done using Big dye V 3.1 terminator mix. The 1µl of PCR amplified products were used as template in 10µl reaction volume constituting 1µl of Big dye, 1 µl of sequence buffer, 1µl of 3.2 µM each specific primers for re-amplifications. PCR conditions followed as: 96^0 C/1 min; (96^0 C/10 s: 56^0 C/5 s; 60^0 C/4 min) repeat for 24 cycles: 10^0 C 10 minutes. Finally, all the PCR amplified DNA fragments were sequenced using DNA analyser from ABI. All the sequenced genetic fragments were analysed for SNP and InDels using Bio-Edit tool [13].

Table 3. Genic fragments targeted for sequencing in *Lablab* accessions

SL No.	Genic fragments	Amplicon size in base pairs
1	CP2	220
2	CP5	300
3	CP8	200
4	CP15	200
5	CP16	250
6	CP23	200
7	CP32	200
8	CP117	150
9	CP171	220
10	CP181	250
11	CP215	180
12	CP403	150
13	CP431	250
14	CP435	170
15	Y1	300
16	pDLL	200
17	5S rRNA	300
18	fril	300
19	Met1	1000
20	Met2	800
21	AF 151961	400
22	AF 287258	400
23	D 13557	800

Designing of SNP and InDel specific primers

In order to develop allele specific markers, the SNPs/ InDels identified based on the sequencing results of genic fragments were deployed. For primer designing Batch Primer, 3 tool was used. We have used tri- and tetra- primer [14, 15] methods to distinguish the allelic differences. In tri primer method, two outer primers and an inner primer was used in a single PCR reactions to generate common band in both the genotypes and single band in one of the genotype. In tetra primer method, two inner primers and two outer primers were used in a single PCR reaction to generate common bigger band in both the genotypes and two different sized small bands differentiate two alleles. The inner primers were designed based on the SNPs; in addition to SNP base a second

deliberate mismatch was incorporated at position -2 from the 3' terminus to enhance the allelic specificity. The sequence details of the allele specific primers designed are shown in Table 4.

Table 4. Sequence details of the allele specific primers

Sl. No.	Primer name	Sequence (5' to 3')
1.	Met2_AS_CPI24973	ATATCCGTTGTGTA-AATTTCGCT
2.	Met2_AS_HA4	ATTAGAAAACAAA-CAAGTTTAGT
3.	CP5_AS_CPI24973	TCTTAATATG-GAAAAAGATAT
4.	CP5_AS_HA4	AGAGACAACAAAG-GATCTCTGTC
5.	AF287258_ AS1_ 31113	AGGCTTC-CCTTTTTATTTTTC-TACG
6.	AF287258_AS1_ HA4	ACACATGGGGAT-TCAATAAACATT
7.	AF287258_AS2_31113	ACTCTTCCTATAGT-GAGAACCTTATTGC
8.	AF287258_AS2-HA4	AAGAAAAATA-AAAAGGGAAGCGTA
9.	CP16_AS_31113	TGGCCTTGTTTCT-GTCTCAACG
10.	CP16_AS_HA4	AAGAATGATGC-CATTGGAGATAT
11.	CP431_AS_31113	TGCTTGTAGCAC-CAAAACCG
12.	CP431_AS_HA4	ATTTGGTGCTACTAG-CACCGCT
13.	D13557_AS_31113	TGTAACTGAGGGT-GCTTGTTGACTG
14.	D13557_AS_HA4	TTTTTAAATCCTTAT-GGCGTTTCAGAT
15.	AF151961-AS-31113	AGATCCATAAC-CCCAAAGGAATTTTC
16.	AF151961-AS-HA4	ACCTGATTCTCT-GCTAAACCCATTAGA

Note: Red colour indicates the SNP
Blue colour indicates the altered base at second position from the 3' end SNP

Wet lab validation of allele-specific markers

For wet lab validation of newly designed allele-specific primers, three parental genotypes were screened HA4, CPI31113 and CPI60216. For the screening of SNP/InDel markers, we kept PCR components the same with altered number of SNP-specific primers based on the genotyping methods. The PCR conditions remained same with different annealing temperatures for each primer combination as used for amplifying genic fragments. All the amplification products were separated on 2% agarose gels and scored for common, co-dominant and dominant allele profiles to record all the allelic variations. Further, the polymorphic allele specific markers were subjected to genotyping of two mapping populations HA4 × CPI31113 and HA4 × CPI60216, with 80 and 90 F_2 individuals. The female-specific allele scored as 'A', male as 'B' and the heterozygotes as 'H' across F_2 individuals. Finally by counting the marker allele scores examined for 1:2:1 Mendalian ratios, as expected in case of co-dominant markers.

RESULTS AND DISCUSSION

Parental polymorphism study using Cowpea EST-SSRs in Lablab bean

For construction of a genetic map followed gene(s)/QTLs analysis, parental polymorphism study between contrasting parents is very much crucial. It is critical that sufficient polymorphism exists between parents in order to construct saturated genetic maps to fine map the genes [16]. Generally in inbreeding species like Dolichos bean requires the selection of parents that are highly divergent. In most of cases, parents that provide adequate polymorphism are selected on the basis of the level of genetic diversity between parents [17- 20]. In this study we deployed a total 58 cowpea genic EST-SSR primers for parental polymorphism study in Lablab bean. Among these, 45 primer pairs generated good amplification profiles in lablab genotypes with reproducible bands. This high rate of amplification of cowpea primers in *Lablab* bean indicates high rate of marker transferability between these two species. Similarly, high SSR marker transferability between these two species was reported [21]. Out of 45 primers screened, 14 generated multiple bands, 3 double bands and 28 with single band amplifications in all the nine parental genotypes with amplicons size ranging from 100-800 base pairs (Table 5). The lack of amplification for some primers in some of the genotypes could be due to sequence variation at primer binding regions. The primer pairs with multiple and double bands observed in this study need standardization of the PCR conditions.

Table 5. Amplification pattern of 58 Cowpea SSR primers in *Lablab purpureus* genotypes

Sl. No.	Primer name	Band size (base pair)		Genotypes**										Remarks
		CP	Lablab	1	2	3	4	5	6	7	8	9	10	
1	CP1	200	D (200, 400)	-	-	-	-	++	++	+	++	-	+	U*
2	CP2	200	S (220)	+	+	+	+	+	+	+	+ (InDel)	+	+	Poly on PAGE, InDels
3	CP3	200	D (200, 800)	+	+	+	-	++	+	++	+	+	+	U*
4	CP4	250	M(200-1000)	-	-	-	-	+	+	+	+	+	+	U*
5	CP5	300	S (300)	+	+	+	+	+	+	+	+	+	+	SNPs
6	CP6	800	S (250)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
7	CP7	200	M(200-1000)	+	+	+	+	+	+	+	+	+	+	U*
8	CP8	200	S (200)	+	+	+	+	+	+	+	+ (InDel)	+	+	Poly on PAGE, InDels
9	CP9	200	S (200)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
10	CP10	200	S (200)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
11	CP11	200	M (450)	+	-	+	-	+	+	+	+	+	+	Mono on PAGE
12	CP12	300	S (700)	-	-	-	-	+	+	+	-	-	+	Mono on PAGE
13	CP13	100	D (100-300)	+	+	+	+	++	++	+	+	+	+	U*
14	CP14	600	S(800)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
15	CP15	200	S (200)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
16	CP16	200	S (250)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
17	CP17	200	S (200)	+	+	+	-	+	+	+	+	+	+	Mono on PAGE
18	CP18	220	S (200)	+	-	+	-	+	+	+	+	-	-	Mono on PAGE
19	CP19	200	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
20	CP20	200	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
21	CP21	250	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
22	CP22	200	M(200-800)	+	+	+	-	+	+	+	+	+	+	Mono on PAGE
23	CP23	200	S (200)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
24	CP24	200	M(300-1000)	+	-	+	-	+	+	+	+	-	-	U*
25	CP25	250	S (250)	-	-	-	-	+	-	+	-	-	+	Mono on PAGE
26	CP26	-	-	-	-	-	-	-	-	-	-	-	-	No amp in Lablab
27	CP27	200	-	-	-	-	-	-	-	-	-	-	-	No amp in Lablab
28	CP28	200	-	-	-	-	-	-	-	-	-	-	-	No amp in Lablab
29	CP29	200	S (200)	+	+	+	+	+	+	+ (InDel)	+	+	+	Poly on PAGE
30	CP30	200	M	+	-	+	-	+	+	+	-	-	+	Mono on PAGE
31	CP31	200	M	+	+	+	-	+	+	+	+	+	+	Mono on PAGE
32	CP32	200	S (200)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
33	*CP115*	150	M	-	-	-	-	+	+	+	+	+	+	U*
34	*CP117*	300	S (150)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
35	*CP163*	200	-	-	-	-	-	-	-	-	-	-	-	No amp in Lablab
36	*CP171*	180	S (220)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
37	*CP181*	280	S (250)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
28	*CP197*	450	M	+	+	+	-	+	+	-	+	-	+	U*
29	*CP201*	200	M	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
40	*CP215*	200	S (180)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
41	*CP239*	180	M	+	-	-	-	+	+	+	+	+	+	U*
42	*CP333*	280	M	+	+	-	-	+	+	+	-	-	+	U*
43	*CP359*	250	-	-	-	-	-	-	-	-	-	-	-	No amp in Lablab
44	*CP391*	200	S (200)	-	+	-	-	+	-	+	-	-	+	Mono on PAGE
45	*CP395*	100	S (100)	-	+	+	-	+	+	+	-	-	+	Mono on PAGE
46	*CP397*	-	-	-	-	-	-	-	-	-	-	-	-	No amp in Lablab
47	*CP403*	150	S (150)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE

48	*CP431*	300	S (250)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
49	*CP433*	280	M	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
50	*CP435*	200	S (170)	+	+	+	+	+	+	+	+ (InDel)	+	+	Poly on PAGE
51	*CP443*	220	S (350)	-	-	+	-	+	+	+	-	-	+	Mono on PAGE
52	*CP573*	100	M	-	-	+	-	-	+	+	+	-	+	U*
53	*CP605*	300	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
54	*Y1*	300	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
55	*Y21*	300	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
56	*Y26*	300	-	-	-	-	-	-	-	-	-	-	+	No amp in Lablab
57	*Y31*	350	S (280)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE
58	*Y45*	100	S (120)	+	+	+	+	+	+	+	+	+	+	Mono on PAGE

U* : these are not used for sequencing and those primers needs optimization for PCR conditions

**1:GL48, 2:GL147, 3:GL153, 4:HA4, 5:Rongai, 6: Highworth, 7:CPI24973, 8:CPI31113, 9:CPI60216 all are *Lablab*, 10:Cow-pea

Note: + means amplification positive and – means amplification negative

S: single band, D: double band & M: multiple band

In our study, none of these primer pairs showed polymorphism on parental genotypes when resolved on 2% agarose gels. This could be due low resolving power as offered by agarose gels as compared to polyacrylamide and capillary-based methods. Therefore, we used 5% horizontal polyacrylamide gels to resolve DNA variations (Fig.1a), which resulted in the identification of four polymorphic markers viz., CP2, CP8, CP29, and CP435. The very low level of polymorphism reported in the present study indicates limited scope for use of these markers for genetic studies in Lablab bean. Similarly, earlier we reported low levels of polymorphism for cowpea-based markers in lablab bean [8]. Hence, here we performed targeted amplicon sequencing of genic fragments that has allowed efficient variant identification and characterization through the development of EST-SNPs and InDel markers in Lablab bean.

Identification of SNPs or InDels in genic fragments through sequencing

Due to low lack of polymorphism in genic fragments EST-SSRs, we followed targeted amplicon sequencing to develop novel EST-SNPs/InDels based markers. Total 23 genic fragments were sequenced for 9 Lablab genotypes, of which 18 represented good quality sequence reads. Through multiple sequence analysis of all these sequences, we identified total 59 SNPs and 16 InDels in the analysed sequences (Table 6). The details of sequence variations observed for each primer were shown in Table 7.

Table 6. Sequencing of genic fragments and identification of SNPs/InDels

Sl. No.	Genic fragments	Amplicon size	SNPs	In-Dels	Remarks
1	CP2	220	5	1	Both SNPs & InDel
2	CP5	300	3	-	Only SNPs
3	CP8	200	-	2	Only InDel
4	CP15	200	-	-	Not good sequence
5	CP16	250	2	-	Only SNPs
6	CP23	200	-	-	Not good sequence
7	CP32	200	1	-	Only SNPs
8	CP117	150	-	-	Not good sequence
9	CP171	220	1	-	Only SNPs
10	CP181	250	2	1	SNPs
11	CP215	180	-	1	SNPs
12	CP403	150	1	-	Only SNPs
13	CP431	250	2	-	Only SNPs
14	CP435	170	1	-	Only SNPs
15	Y1	300	1	-	Only SNPs
16	pDLL	200	-	-	Not good sequence
17	5S rRNA	300	3	1	Both SNPs & InDels
18	fril	300	-	-	Not good sequence
19	Met1	1000	15	4	Both SNPs & InDels
20	Met2	800	15	5	Both SNPs & InDels
21	AF 151961	400	2	-	Only SNPs
22	AF 287258	400	3	-	Only SNPs
23	D 13557	800	2	1	Both SNPs & InDels
		Total	**59**	**16**	

Table 7. Details of SNP and InDel variations observed in different genic fragments of Lablab bean genotypes

Met1																			
Genotype	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
CPI24973	C	C	T	C	C	C	G	A	A	T	T	A	C	G	T	G	G	C	A
CPI31113	C	T	T	C	C	T	A	A	A	C	T	A	C	A	-	-	G	T	T
CPI60216	No sequence																		
Rongai	T	C	-	-	C	C	G	A	T	T	A	G	T	A	-	-	A	C	A
Highworth	T	C	-	-	G	C	G	G	A	T	A	G	C	A	-	-	G	C	A
HA-4	T	C	-	-	C	C	G	A	T	T	A	G	T	A	-	-	A	C	A
GL48	T	C	-	-	C	C	G	A	T	T	A	G	T	A	-	-	A	C	A
GL147	T	C	-	-	C	C	G	A	T	T	A	G	T	A	-	-	A	C	A
GL153	T	C	-	-	C	C	G	A	T	T	A	G	T	A	-	-	A	C	A

Met2																				
Genotype	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
CPI24973	T	C	T	TTG	G	TT	A	T	C	G	G	-	G	A	TCTG	C	A	G	T	G
CPI31113	T	C	T	TTG	G	TT	A	T	T	A	G	-	T	G	----	C	T	A	T	T
CPI60216	C	G	C	---	-	--	T	A	T	A	T	G	T	A	TCTG	G	T	A	T	T
Rongai	T	C	T	TTG	G	TT	A	T	T	A	G	-	G	A	TCTG	G	A	G	A	T
Highworth	T	C	T	TTG	G	TT	A	T	T	A	G	-	G	A	TCTG	G	A	G	A	T
HA-4	T	C	T	TTG	G	TT	A	T	T	A	G	-	G	A	TCTG	G	A	G	A	T
GL48	T	C	T	TTG	G	TT	A	T	T	A	G	-	G	A	TCTG	G	A	G	A	T
GL147	T	C	T	TTG	G	TT	A	T	T	A	G	-	G	A	TCTG	G	A	G	A	T
GL153	T	C	T	TTG	G	TT	A	T	T	A	G	-	G	A	TCTG	G	A	G	A	T

Genotype	**D13557**				**AF287258**			**AF151961**		**5S rRNA**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**1**	**2**	**1**	**2**	**3**	**4**
CPI24973	T	T	----------	A	A	T	A	C	A	-	T	G	A
CPI31113	C	A	AGCTCAACGTATA	G	T	C	G	T	A	No sequence			
CPI60216			No sequence		T	C	G	C	A	No sequence			
Rongai	C	A	AGCTCAACGTATA	A	T	T	A	C	A	C	T	G	T
Highworth			No sequence		T	T	A	C	A	C	G	A	T
HA-4	C	A	AGCTCAACGTATA	A	T	T	A	C	G	C	T	G	T
GL48	C	A	AGCTCAACGTATA	A	T	T	A	C	A	No sequence			
GL147	C	A	AGCTCAACGTATA	A	T	T	A	C	A	C	T	G	T
GL153	C	A	AGCTCAACGTATA	A	T	T	A	C	A	C	T	G	T

Genotype	**CP2**							**CP5**			**CP8**		**CP16**	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**1**	**2**	**3**	**1**	**2**	**1**	**2**
CPI24973	T	G	G	G	C	AAGAAGAG	C	T	T	G	TT	----------	A	C
CPI31113	C	T	C	T	A	- - - - - - - -	T	C	T	A	--	CTTACTTCTTCT	G	C
CPI60216	C	T	C	T	A	AAGAAGAG	T	C	T	A	--	----------	G	T
Rongai	C	T	C	T	A	AAGAAGAG	T	C	G	A	--	----------	A	C
Highworth	C	T	C	T	A	AAGAAGAG	T	C	G	A	--	----------	A	C
HA-4	C	T	C	T	A	AAGAAGAG	T	C	G	A	--	----------	A	C
GL48	C	T	C	T	A	AAGAAGAG	T	C	G	A	--	----------	A	C
GL147	C	T	C	T	A	AAGAAGAG	T	C	G	A	--	----------	A	C
GL153	C	T	C	T	A	AAGAAGAG	T	C	G	A	--	----------	A	C

Genotype	CP32	CP215	CP431		CP171	CP403	CP435	CP181			Y31
	1	1	1	2	1	1	1	1	2	3	1
CPI24973	A	-	A	G	C	T	G	T	T	C	No sequence
CPI31113	A	-	G	G	C	C	No sequence	T	T	A	G
CPI60216	A	-	A	T	G	T	T	A	-	A	G
Rongai	A	-	A	T	C	T	T	T	T	A	A
Highworth	A	C	A	T	C	T	T	T	T	A	A
HA-4	T	-	A	T	C	T	T	T	T	A	A
GL48	A	-	A	T	C	T	T	T	T	A	A
GL147	A	-	A	T	C	T	T	No sequence	A		
GL153	A	-	A	T	C	T	T	T T	A		A

When we analysed genic fragment sequences, only SNPs were observed for nine primers such as AF 287258, AF151961, CP5, CP16, CP 431, CP171, CP 403, CP 435 and Y3. Whereas, two primers namely CP 215 and CP8 revealed only InDel variations. However, five primers viz., CP2, Met 1, Met 2, D13557 and CP181 showed both SNPs and InDels. The two primers namely Met 1 and Met2 showed the highest number of SNPs (15 each) and 4, 5 InDels across the genotypes respectively. In our study, InDel length ranged from single base to 13 bases, D 13557 and CP 8 showed 13 and 12 base InDels, respectively. Overall sequencing results revealed 59 SNPs and 16 InDels from 18 genic fragments among nine parental genotypes. Interestingly we observed more SNPs than INDELs. The SNP/INDEL variation could be synonymous or non-synonymous variation. The low SNP variation among the Indian Lablab reveals homology and high level of sequence conservation of these genes in this set of genotypes. The higher sequence variations identified between the Indian and African lablab genotypes indicated low homology between the two groups. Further, the identified SNPs and Indels of genic fragments of primers viz., Met 2, AF 287258-1, AF 285278-2, AF 151961, D 13557, CP 5, CP16 and CP 431, were then used to develop 8 allele-specific markers.

Validation of SNP/InDel markers through allele-specific PCR

Parental polymorphism study

A wide variety of different methods have been developed to carry out high throughput genotyping of SNPs using automated systems [22]. These methods are often reliant on expensive equipment and require high developmental costs. Such marker assays generated are not commonly transferable between laboratories due to differences in the assay techniques used. Therefore, a simple and cost effective method for SNP marker genotyping would greatly improve the marker transferability and application [23]. Towards this end, allele-specific PCR (ASPCR) is known to be an efficient and common approach for genotyping SNP markers. This has enabled the detection of heterozygotes in a single assay using four primers in the same PCR mix that involved tetra primer PCR [24,15]. A simplified approach using three primers has been used [14].

In the present investigation, we tried to deploy both tri-primer and tetra-primer AS-PCR methods. Compared to tri- method, the tetra-primer method was found successful in distinguishing the heterozygous SNP alleles (Fig.1b). Seven out of eight allele-specific primer pairs *viz.,* Met 2, AF 287258-1 AF 285278-2, AF 151961, D 13557, CP 5, CP16 and CP 431 could distinguish the two different alleles in HA 4, CPI 31113 or CPI 60216, respectively. One primer (CP 5) failed to distinguish heterozygous alleles although it was successfully amplified between the parents (Table 8). The possible reason may be that many of the internal primers designed for the ASPCR amplifications had suboptimal qualities for PCR based on the position of the SNP site while designing of primers likely resulting in amplification failure. Whereas, compared to internal primers flanking-primers always have a best choice made from many possibilities during primer designing. In tetra primer ARMS-PCR method inner primers encompass a deliberate mismatch at position -2 from the 3' terminus, because an extra mismatch found to increase the specificity of classical ARMS [25]. Similarly, robustness and reproducibility of the AS-PCR approach against primer extension using tri-primer methods in single reaction was also tested to show the specificity [14].

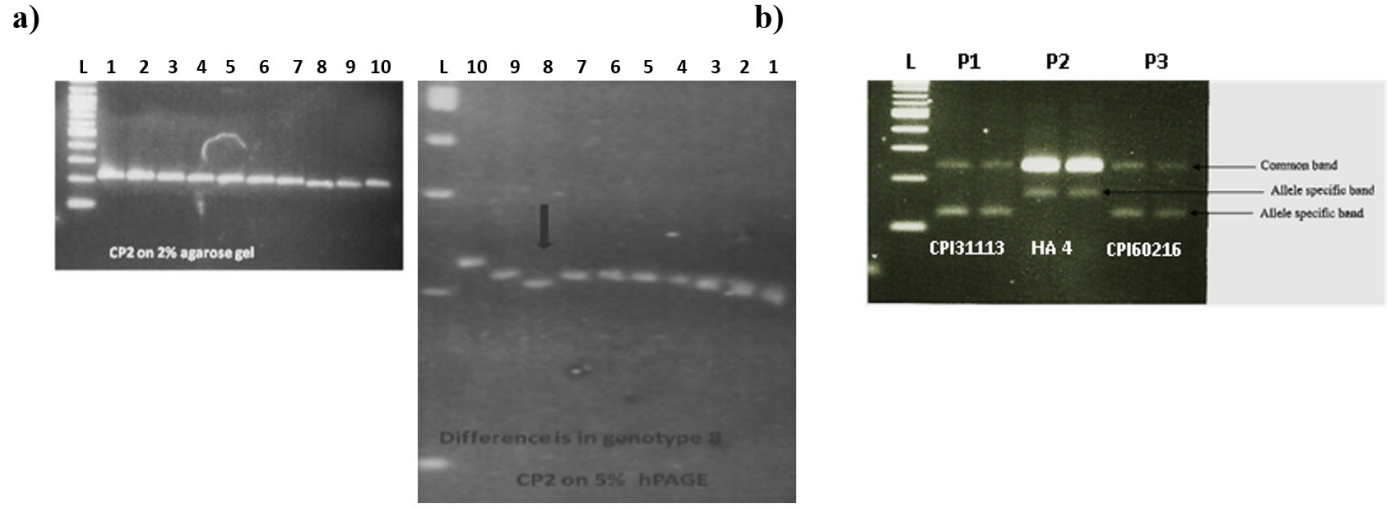

Fig.1 **a)** Comparison of EST-SSR primer CP2 profile on 10 lablab genotypes using on 2% agarose gel and 5% PAGE **b)**Allelic differences as observed on three mapping parents CPI 31113, HA 4 and CPI 60216 using allele specific primer AF 287258 based on tetra primer method (Where, L= 50bp ladder)

Table 8. Amplification pattern for allele specific primers in parental lines

Sl No.	Primer name	Amplicon size in base pairs				Remarks
		Common band	Allele specific band			
			HA 4	CPI 31113	CPI 60216	
1.	Met 2	400	250	150	NA	Co-dominant for HA4 X CPI31113 Dominant for HA4 X CPI60216
2.	AF 287258-1	330	230	130	130	Co-dominant for both crosses
3.	AF 287258-2	230	160	100	100	Co-dominant for both crosses
4.	AF 151961	200	100	130	NA	Co-dominant for HA4 X CPI31113 Dominant for HA4 X CPI60216
5.	D 13557	600	400	200	NA	Co-dominant for HA4 X CPI31113 Dominant for HA4 X CPI60216
6.	CP5	300	150	150	150	Difficult to assay
7.	CP16	240	150	120	120	Co-dominant for both crosses
8.	CP431	300	120	180	NA	Co-dominant for HA4 X CPI31113 Dominant for HA4 X CPI60216

The results of the present investigations indicated that the method utilized here could be broadly applicable to the development of SNP markers, where there is sufficient knowledge available on DNA sequence around the SNP sites. AS-PCR has enabled to develop informative markers that are polymorphic and can be used to map the transcriptome, for general linkage mapping for QTL detection, as perfect markers for the candidate gene approach, for varietal identification and to perform the function of DNA markers in general. These markers could act as an adjunct to CAPS where sites are not sensitive to restriction enzyme assay or as a more reliable replacement. In particular, they will be useful for projects that are on a modest budget or where sophisticated equipment is not available.

Genotyping of mapping populations

The choice of DNA markers used for mapping may depend on the availability of characterized markers or the appropriateness of particular markers for a particular species. Once polymorphic markers have been identified, they must be screened across the entire mapping population, including the parents. In this study, we validated 7 polymorphic markers across segregation lines of two F_2 mapping populations derived from the crosses HA4 × CPI31113-80F_2 and HA4 × CPI60216- 89 F_2 individuals respectively (Table 9, Fig.2). Marker segregation patterns were

examined to test Mendelian 1:2:1 segregation ratios in F_2 populations. As a result 6 out of 7 markers exhibited 1:2:1 ratio in HA4 X CPI 31113 population. However, only three markers CP5-SNP, AF 287258 and AF 151961 could amplify and show 1:2:1 ratio in HA4 X CPI 60216 population. These results suggested the novel SNP and InDel markers developed here through sequencing of genic fragments are highly useful for construction of genetic maps and trait mapping in dolichos bean. Similarly, RFLP markers were deployed in *Lablab purpureus* L. (Sweet) to show expected segregation ratios of 1:2:1 in F_2 mapping population [26].

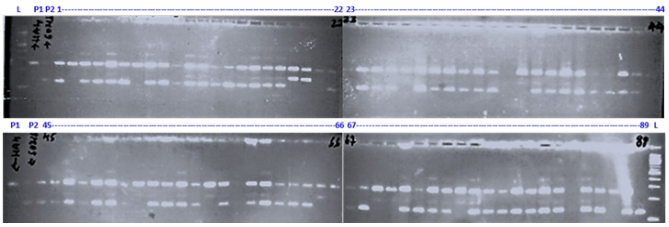

Fig. 2 Genotyping of 89F_2 segregating individuals of mapping population HA4 X CPI 60216 using AF 287258 allele specific primer (Where, L= 50bp ladder, P1: HA4, P2: CPI 60216)

CONCLUSION

The amplification of cowpea genic SSR primers in Lablab bean showed high marker transferability rate. The sequencing of 18 genic fragments among nine parental genotypes, revealed 59 SNPs and 16 InDels. Using SNPs and InDels information's we successfully developed eight allele specific SNP and InDel markers. Finally, validated these markers through parental polymorphism study followed by genotyping of two F_2 populations, and confirmed the immediate utility of such markers for trait mapping in Dolichos bean.

ACKNOWLEDGEMENTS

The authors are highly grateful to the Kirkhouse Trust, UK for extending financial support.

Table 9. Segregation patterns of allele specific markers in two F_2 populations

Sl. No	Markers	HA4 X CPI 31113				HA4 X CPI 60216			
		P_1 (A)	P2 (B)	Heterozygotes (H)	Ratio	P1 (A)	P2 (B)	Heterozygotes (H)	Ratio
1.	CP5 – SNP	17	17	31	1:2:1	21	26	40	1:2:1
2.	CP2 – indel	19	20	39	1:2:1	-	-	-	
3.	CP8 – indel	18	19	35	1:2:1	-	-	-	
4.	Met IV - indel	17	16	29	1:2:1	-	-	-	
5.	AF 287258 – allele specific	20	19	39	1:2:1	22	21	44	1:2:1
6.	D13557 – allele specific	18	20	40	1:2:1	-	-	-	
7.	AF 151961 – allele specific	-	-	-		21	23	43	1:2:1

REFERENCES

[1.] Fuller, D.Q., 2003. African crops in prehistoric South Asia: a critical review. In: Neumann, K., Butler, A., Kahlheber, S. (Eds.), Food, Fuel, fields - Progress in African Archaeobotany, Heinrich-Barth-Institut, 15. Africa Praehistorica, Koln, Germany, pp. 239 –271.

[2.] Khatun, R., Uddin M.I., Uddin, M.M., Howlader, M,T,H. and Haque, M, S., 2022. Analysis of qualitative and quantitative morphological traits related to yield in country bean (Lablab purpureus L. sweet) genotypes. Heliyon 8:e11631.doi.org/10.1016/j.heliyon.2022. e11631.

[3.] Letting, F.K., Venkataramana, P.B and Ndakidemi, P.A. 2022. Pre-breeding prospects of Lablab (Lablab purpureus L. Sweet) accessions in Tanzania: Morphological characterization and genetic diversity analysis. Agronomy 12:2272. https://doi.org/ 10.3390/agronomy12102272.

[4.] Adnan, M.A., Raffi, S.A., Rubel, M.H., Noor, M.M.A., Rahman, A.B., Akhter, B., Raiza, S. and Haque, M.A. 2021. Screening of lablab bean genotypes for higher grain yield and resistance against bean common mosaic virus. SABRAO J. Breed. Genetics 53: 57-69.

[5.] Murphy, A. M. and Colucci, P. E. 1999. A tropical forage solution to poor quality ruminant diets: A review of Lablab purpureus. Livestock Research for Rural Development, 11:96-113.

[6.] Liu, C. J. 1996. Genetic diversity and relationships among Lablab purpureus genotypes evaluated using RAPD markers. Euphytica 90: 115-119.

[7.] Maass, B. L. 2005. Changes in seed morphology, dormancy and germination from wild to cultivated hyacinth bean germplasm (Lablab purpureus L. Papilinoideae). Genetic Reso. Crop Evol., 52: 1-9.

[8.] Venkatesha, S. C., Byregowda, M., Mahadevu, P., Mohanrao, A., Kim, D. J., Ellis, T. H. N. and Knox, M. R., 2007, Genetic diversity within Lablab purpureus and the application of gene specific markers from a range of legume species. Plant Genet. Res., 5 (3): 154-171.

[9.] Zhu, H., Kim, D. J., Baek, J. M., Choi, H. K., Ellis, L. C., Kuester, H., Mccombie, W. R., peng, H. M. and Cook, D. R., 2003. Syntenic Relationships between Medicago truncatula and Arabidopsis reveal extensive divergence of genome organization, Plant Physiol., 131: 1018-1026.

[10.] Humphry, M. E., Konduri, V., Lambrides, C. P., Mangner, T., Meintrya, C. L., Aitjen, E. A. B. and Lin, C. J., 2002. Development of a mungbean (Vigna radiate) RFLP linkage map and its comparison with lablab (Lablab purpureus L.) reveals a high level of colinearity between the two genomes. Theor. Appl. Genet., 105: 106-166.

[11.] Ellis, T.H.N., Davies, D.R., Castleton, J.A. and Bedford, I.D., 1984. The organization and genetics of rDNA length variants in peas. Chromosoma 91: 74 –81.

[12.] Patil, P. G., Byre Gowda, M., Kundur, P. J., Vimarsha, H. S. and Shashidar, H. 2015. Upgraded horizontal polyacrylamide gel units for DNA marker genotyping. Indian Journal of Science and Technology 8(9): 822-827.

[13.] Hall, T.A. (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95-98.

[14.] Soleimani, V. D., Baum, B. R. and Johnson, D. A. 2003. Efficient validation of single nucleotide polymorphisms in plants by allele-specific PCR, with an example from barley, Plant Mol. Biol. Reporter 21: 281-288.

[15.] Ye, S., Dhillon, S., Ke, X., Collins, A. R. and Day, I. N. M., 2001. An efficient procedure for genotyping single nucleotide polymorphisms, Nuc. Acids Res., 29(17): e88.

[16.] Young, N.D., Danesh, D, Menancio-Hautea, D. and Kumar, L. 1993. Mapping oligogenic resistance to powdery mildew in mungbean with RFLPs. Theor Appl Genet., 87: 243-249.

[17.] Anderson, J., Churchill, G., Autrique, J., Tanksley, S. and Sorrells, M., 1993. Optimizing parental selection for genetic linkage maps. Genome 36: 181–186.

[18.] Collard, B.C.Y., Pang, E.C.K., and Taylor, P.W.J., 2003. Selection of wild Cicer accessions for the generation of mapping populations segregating for resistance to ascochyta blight. Euphytica, 130: 1–9.

[19.] Joshi, C. and Nguyen, H., 1993. RAPD (random amplified polymorphic DNA) analysis based inter varietal genetic relationships among hexaploid wheats. Plant Sci., 93: 95–103.

[20.] Yu, L.X. and Nguyen, H., 1994. Genetic variation detected with RAPD markers among upland and lowland rice cultivars (Oryza sativa L.). Theor. Appl. Genet., 87: 668- 672.

[21.] Wang, M.L., Gillaspie, A.G., Newman, M.L., Dean, R.E, Pittman, R.N, Morris, J.B. and Pederson, G.A. 2004. Transfer of simple sequence repeat (SSR) markers across the legume family for germplasm characterization and evaluation. Plant Genetic Resources 2: 107-119.

[22.] Gupta, P. K., Roy, J. K. and Prasad, M., 2001. Single nucleotide polymorphisms: A new paradigm for molecular marker technology

and DNA polymorphism detection with emphasis on their use in plants. Curr. Sci., 80(4): 524-535.

[23.] Kwok P-Y, Chen X. 1998. Detection of single nucleotide polymorphisms. In Genetic Engineering, Principles and Methods, ed. JK Setlow. New York: Plenum. 20:125.

[24.] Ye, S., Humphries, S. and Green, F., 1992. Allele specific amplification by tetraprimer PCR, Nuc. Acids Res., 20 (5): 1152.

[25.] Little, S. 1997. ARMS analysis of point mutations. In Taylor, G.R. (ed.) Laboratory Methods for detection of mutations and polymorphisms in DNA. CRC Press, Boca Raton, FL, pp.45-51

[26.] Konduri, V., Godwin, I. D. and Liu, C. J., 2000. Genetic mapping of the Lablab purpureus genome suggests the presence of 'cuckoo' gene(s) in this species, Theor. Appl. Genet., 100: 866–871.