

# Development of intron length polymorphism markers in cowpea [*Vigna unguiculata* (L.) Walp.] and their transferability to other *Vigna* species

S. K. Gupta · R. Bansal · T. Gopalakrishna

Received: 26 September 2011 / Accepted: 28 February 2012 / Published online: 18 March 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Expressed sequence tag (EST) sequences available in the public databases provide a cost-effective and valuable genomic resource for the development of molecular markers. Introns which are non-coding DNA sequences of the gene could be used as potential molecular markers as they are highly variable compared to the coding sequences. This study reports the development of intron length polymorphism markers in cowpea [*Vigna unguiculata* (L.) Walp.]. The ESTs of cowpea were aligned with genomic sequences of *Arabidopsis* and soybean to predict the position and number of introns in cowpea. Of the 110 PCR primer pairs designed to amplify the intronic regions, 98 primer pairs resulted in successful amplification and were identified as cowpea intron length polymorphism (CILP) markers. Out of the 45 randomly selected CILP markers, 36 % markers produced length variation in the ten cowpea genotypes, collectively yielding 33 alleles with an average of 2.0 alleles/locus. The polymorphism information content of the CILP markers ranged from 0.18 to 0.64 with an average of 0.34. Of the 98 CILP markers, 93

markers (95 %) showed transferability to other *Vigna* species. Dendrograms based on CILP markers clearly distinguished the cowpea genotypes as well as other *Vigna* species, demonstrating the utility of CILP markers in genetic diversity and phylogenetic studies. These CILP markers will be very useful in the genome analysis and marker-assisted breeding of cowpea and other *Vigna* species.

**Keywords** Intron length polymorphism · Cowpea · *Vigna* · Molecular markers · ESTs

## Introduction

Grain legumes are among the most important crops in many countries and provide one-third of human dietary protein. Most of the economically important grain legumes belong to the genus *Vigna* in the tribe Phaseoleae. The genus *Vigna* consists of seven subgenera, namely *Vigna*, *Haydonia*, *Lasiocarpa*, *Sigmoidotropis*, *Plectotropis*, *Ceratotropis* and *Macrorhyncha* (Marechal et al. 1978). The African subgenus *Vigna* contains the important species cowpea [*Vigna unguiculata* (L.) Walp.] and the subgenus *Ceratotropis* includes about 16 species distributed across Asia, including mungbean [*Vigna radiata* (L.) Wilczek], azuki bean [*Vigna angularis* (Willd.) Ohwi & Ohashi], blackgram (*Vigna mungo* (L.) Hepper], rice bean [*V. umbellata* (Thunb.) Ohwi & Ohashi] and moth bean [*V. aconitifolia* (Jacq.) Marechal]. The

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s11032-012-9722-y) contains supplementary material, which is available to authorized users.

---

S. K. Gupta (✉) · R. Bansal · T. Gopalakrishna  
Nuclear Agriculture and Biotechnology Division,  
Bhabha Atomic Research Centre, Trombay,  
Mumbai 400 085, India  
e-mail: gupta\_sk@hotmail.com

species belonging to the subgenus *Vigna* and *Ceratotropis* are significant as they provide an important and inexpensive source of dietary protein to the people of Asia and Africa. Despite systematic and continuous efforts through conventional breeding approaches, substantial yields in these grain legumes could not be achieved. Further, the yields in these crops are severely affected by many biotic factors like yellow mosaic virus, powdery mildew and bruchid infestation.

Molecular or genetic markers are powerful tools for genetic research and breeding. A variety of molecular marker techniques such as restriction fragment length polymorphism (RFLP, Botstein et al. 1980), random amplified polymorphic DNA (RAPD, Williams et al. 1990; Welsh and McClelland 1990), inter simple sequence repeat (ISSR, Zietkiewicz et al. 1994), microsatellite or simple sequence repeat (SSR, Tautz and Renz 1984) and amplified fragment length polymorphism (AFLP; Vos et al. 1995) have been developed for gene tagging, diversity analysis and genetic linkage map construction in many crop plants. The availability of large expressed sequence tag (EST) databases for many crop species provides a valuable resource for the development of molecular markers like intron length polymorphism (ILP). Introns are non-coding sequences of genes and are less conserved than the exons. The variation in the intron sequences can be exploited as molecular markers (Wang et al. 2005). For developing the ILP markers, intron positions are determined by comparing the EST sequences with the genomic sequences of a model plant. Once the intron is identified, PCR primers are designed for the flanking region so as to amplify the intronic region. ILP markers, being part of the genes, are more useful as genetic markers because they represent variation in the transcribed portion of the genome. ILP markers have been developed in many plant species including soybean (Shu et al. 2010), rice (Wang et al. 2005), tomato (Wang et al. 2010), medicago (Choi et al. 2004) and foxtail millet (Gupta et al. 2011). Like other molecular markers, ILP markers can be used for a variety of applications like molecular mapping, gene tagging, genetic diversity analysis and comparative studies. In addition, ILP markers show a high rate of transferability to related species owing to a higher conservation of EST sequences across species (Wang et al. 2005; Yang et al. 2007). Therefore, ILP markers developed in one species can be used in a related

species for which sufficient genomic resources for marker development are not available.

Despite their importance, molecular breeding research is still lagging behind in *Vigna* species compared to other grain legumes such as common bean and soybean. There is an urgent need to develop a large number of molecular markers in *Vigna* species, which can be used by breeders in various molecular breeding strategies to increase the breeding efficiency. Therefore, the present study was conducted with an aim of developing ILP markers in cowpea and evaluating their transferability to other *Vigna* species.

## Materials and methods

### Plant material and DNA extraction

Ten cowpea genotypes and eight other *Vigna* species used in the study are listed in Table 1. Total genomic DNA was extracted from 10-day-old seedlings using the modified CTAB method. The quality of DNA was checked on 1 % agarose gel and the quantity was determined at 260 nm using a UV spectrophotometer (Unicam UV 300, UK).

### EST sequence retrieval and primer design

Unigene sequences of cowpea were downloaded from the HarvEST:Cowpea assembly P12 (<http://harvest.ucr.edu>). These sequences were aligned with the *Arabidopsis thaliana* and soybean (*Glycine max*) genomic sequences using the WebGMAP software (Liang et al. 2009) to predict the intron positions. For aligning EST sequences with the genomic sequences, the inter-species mapping option was used (i.e. only hits that satisfy minimum matched query length of 30 nucleotides, minimum matched query identity of 30 % and minimum matched coverage of query sequence of 10 % were considered valid hits). The position of introns was determined by identifying the gaps in the cowpea ESTs. Primer3 software (<http://frodo.wi.mit.edu/primer3/>) was used to design PCR primer pairs from highly conserved EST sequences to amplify the intronic regions. For designing primers, user-defined parameters were used. viz. optimum primer length was 20 mer (range 18–25 mer), optimum annealing temperature was 60 °C (range 55–62 °C), optimum GC

**Table 1** List of *Vigna* species used in the study along with their accession number and source

Serial no.	<i>Vigna</i> species used in the study	Accession/cultivar	Source
1	<i>V. unguiculata</i>	Arka suman	BARC, India
2	<i>V. unguiculata</i>	C-352	BARC, India
3	<i>V. unguiculata</i>	CO(CP)7	TNAU, India
4	<i>V. unguiculata</i>	GC-3	BARC, India
5	<i>V. unguiculata</i>	V-130	BARC, India
6	<i>V. unguiculata</i>	Pusa Phalguni	BARC, India
7	<i>V. unguiculata</i>	EC634641	AVRDC, Taiwan
8	<i>V. unguiculata</i>	V-240	BARC, India
9	<i>V. unguiculata</i>	GC-4	BARC, India
10	<i>V. unguiculata</i>	IC349906	NBPGR, India
11	<i>V. vexillata</i>	IC277007	NBPGR, India
12	<i>V. umbellata</i>	EC634640	AVRDC, Taiwan
13	<i>V. glabrescens</i>	EC634634	AVRDC, Taiwan
14	<i>V. aconitifolia</i>	–	TNAU, India
15	<i>V. trilobata</i>	IC251436	NBPGR, India
16	<i>V. angularis</i>	EC634633	AVRDC, Taiwan
17	<i>V. radiata</i>	EC634637	AVRDC, Taiwan
18	<i>V. mungo</i>	EC634635	AVRDC, Taiwan

AVRDC Asian Vegetable Research and Development Centre, Taiwan; BARC Bhabha Atomic Research Centre, Mumbai, Maharashtra, India; NBPGR National Bureau of Plant Genetic Resources, Regional station, Thrissur, Kerala; TNAU Centre for Plant Breeding & Genetics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu

content was 50 % (range 30–80 %) and the rest of the parameters had the default value.

#### ILP marker analysis

PCR reactions were performed in a 20- $\mu$ l volume containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.5 unit *Taq* DNA polymerase (Bangalore Genei, Bangalore, India), 50 ng template DNA, and 20 ng each of forward and reverse primer. PCR amplifications were performed in an Eppendorf Mastercycler Gradient (Eppendorf, Hamburg, Germany) using the following thermal profile: one cycle of 95 °C for 2 min, followed by 39 cycles of 94 °C for 30 s, 50–65 °C for 30 s (depending on the primer), 72 °C for 1 min and a final extension at 72 °C for 7 min.

To evaluate the allelic variation among the cowpea genotypes, PCR products were resolved on a denaturing polyacrylamide gel. The PCR products were mixed with an equal amount of loading buffer (98 % formamide, 10 mM EDTA, 0.1 % bromophenol blue, 0.1 % xylene cyanol), denatured for 3 min at 94 °C and then kept on ice for 2 min. About 5  $\mu$ l of each reaction mixture was electrophoresed on a 6 % denaturing polyacrylamide gel containing 7 M urea in 1 $\times$  Tris-borate-EDTA buffer. Electrophoresis was

performed at a constant power of 50 W for about 2 h in a Sequi-Gen GT Sequencing system (Bio-Rad, USA). Gels were stained using a modified silver staining protocol as described in Gupta and Gopalakrishna (2010). The gel was scanned on a photo scanner and the data was scored manually on a white light illuminator.

To access the transferability of cowpea ILP markers across *Vigna* species, ILP markers were screened on eight other *Vigna* species. PCR products were resolved on 2 % agarose gel in 1 $\times$  Tris-borate-EDTA buffer, stained with ethidium bromide and photographed on a gel documentation system (Syngene, UK). To check whether the PCR products amplified the same target gene, the PCR product amplified by the marker CILP66 in different *Vigna* species was cloned using InstAclone<sup>TM</sup> PCR cloning kit (Fermentas, Germany) and sequenced using the M13 universal primer. The multiple sequence alignment was performed using the ClustalW2 program (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).

#### Statistical analysis

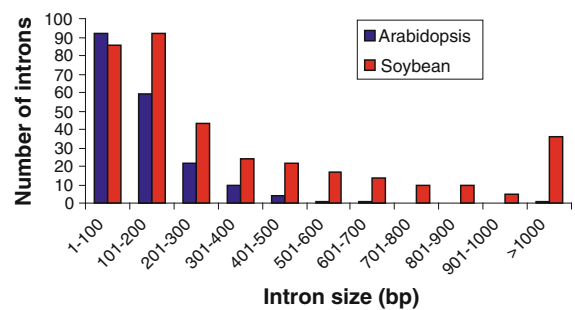
Allelic variation was calculated from the frequencies of genotypes at each locus as the polymorphism information content (PIC). Scoring of bands was done

as presence (1) or absence (0) for each locus. The PIC of each cowpea ILP marker was calculated by applying the formula of Anderson et al. (1993):  $PIC = 1 - \sum(P_{ij})^2$ , where  $P_{ij}$  is the frequency of the  $j$ th allele for the  $i$ th locus. Genetic similarity analyses were performed using NTSYS-pc version 2.0 (Rohlf 1998). Data was entered in a binary matrix as discrete variables and pair-wise similarities were obtained using Jaccard's coefficient. The matrices of similarity coefficients were subjected to the unweighted pair group method with arithmetic average (UPGMA) to estimate the genetic relatedness among the genotypes and generate the dendrogram.

## Results

### Intron identification

A total of 12,117 cowpea unigene sequences were aligned with the genomic sequences of *Arabidopsis* and soybean. With the *Arabidopsis* genome, 2,926 cowpea ESTs showed significant hits and 778 ESTs had one or more intron insertion sites. The total number of introns present in *Arabidopsis* was 1,539 with an average of 1.97 introns per EST. In the case of soybean, 10,273 cowpea ESTs showed a significant match and 5,914 ESTs with 14,885 introns were identified with an average of 2.51 introns per EST. The maximum number of introns present in an EST was eight in the case of *Arabidopsis* and ten in the case of soybean. One hundred ESTs were randomly selected for intron number and size comparison. The number of introns identified in the case of *Arabidopsis* was 190 while in the case of soybean 359 introns were identified in 100 randomly selected ESTs. Only 12 % of ESTs had the same number of introns in both species, while 88 % of ESTs had more introns in soybean than *Arabidopsis*. The length of introns in soybean was much larger than in *Arabidopsis*, varying from 53 to 1,104 bp in *Arabidopsis* with an average intron length of 154.6 bp, while in soybean intron length varied from 73 to 9,255 bp with an average intron length of 439 bp. In *Arabidopsis*, 79.5 % of introns were less than 200 bp and only 1.6 % of introns were larger than 500 bp (Fig. 1), while in soybean 50 % of introns were less than 200 bp and 31.7 % of introns were larger than 500 bp.



**Fig. 1** Distribution of intron sizes in *Arabidopsis* and soybean based on cowpea ESTs

### ILP marker analysis

For brevity, 110 PCR primers targeting one or more introns were developed from randomly chosen cowpea EST sequences. Of the 110 PCR primer pairs, 98 primer pairs successfully amplified a single marker fragment in the two cowpea genotypes GC-3 and V-130. Ten primer pairs did not amplify and two primer pairs produced multiple fragments. Based on the size estimation, all PCR products were found to be larger than the targeted EST sequence, suggesting the presence of intron(s) in the amplified products. These markers were named as CILP (cowpea intron length polymorphism) markers and details are provided in the online resource (Supplementary Table S1). The EST sequences used for ILP marker development were functionally annotated based on similarity against the non-redundant protein database in GenBank using the BLASTX search tool (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The putative function was successfully assigned to about 90 % of sequences based on the most informative high-scoring hit and 10 % of ESTs had an unknown or hypothetical function.

Intraspecific variability among ten cowpea genotypes was examined with 45 randomly selected CILP markers. Sixteen CILP markers (36 %) were found to be polymorphic between the genotypes and collectively yielded 33 alleles with an average of 2.0 alleles per locus. The number of alleles varied from one to three. The PIC of individual loci ranged from 0.18 to 0.64, with a mean value of 0.34 (Supplementary Table S2). The UPGMA analysis distributed the ten genotypes into three main clusters with Jaccard's similarity coefficient ranging from 0.34 to 0.88 (Fig. 2). The first cluster consisted of seven cowpea genotypes (Arka Suman, C-352, CO(CP)7, GC-3, EC634640, GC-4 and

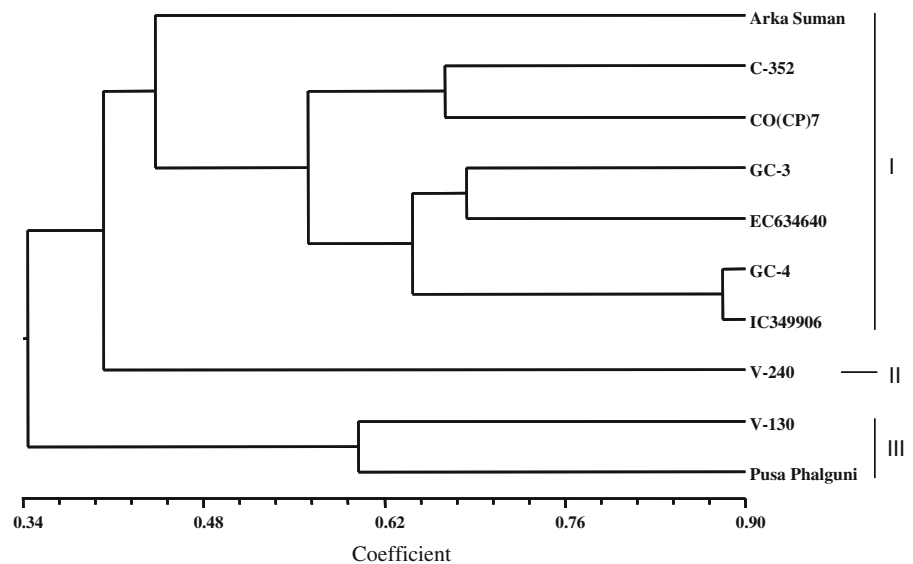
IC349906) and genotypes GC-4 and IC349906 had the least distance among the ten genotypes. The second cluster had only one genotype (V-240) and the third cluster contained two genotypes (V-130 and Pusa Phalguni).

#### Cross-species transferability of ILP markers

To check the transferability of CILP markers across *Vigna* species, all the 98 CILP markers were screened on eight *Vigna* species. About 93 ILP markers (95 %) were found to be transferable to other *Vigna* species and five markers (5 %) were cowpea-specific (Supplementary Table S3). Of the 93 transferable CILP markers, 77 were transferable to all eight *Vigna* species, three markers to seven *Vigna* species, five markers to six *Vigna* species and eight markers to less than six *Vigna* species. Most of the CILP markers produced significant length variation among the *Vigna* species (Fig. 3). The sequence of the cloned product from the marker CILP66 showed that the same target gene was amplified in all the *Vigna* species. There were three introns present in the amplified fragments as earlier revealed by WebGMAP analysis. The exonic regions were highly conserved among the *Vigna* species and large differences, including length variation and point mutations, were observed in the intronic regions (Supplementary Fig. S1).

The 93 CILP markers showing transferability collectively yielded 168 alleles. The dendrogram

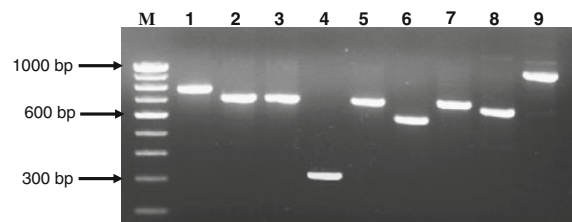
**Fig. 2** Dendrogram showing genetic relatedness among ten cowpea genotypes based on cowpea ILP markers using Jaccard's similarity coefficient



constructed based on similarity matrix clearly separated the *Vigna* species into distinct lineages (Fig. 4). Cowpea, along with its wild relative *V. vexillata* which belong to the subgenus *Vigna*, were placed separately from the other species belonging to the subgenus *Ceratotropis* (*V. radiata*, *V. mungo*, *V. angularis*, *V. glabrescens*, *V. aconitifolia*, *V. umbellata* and *V. trilobata*).

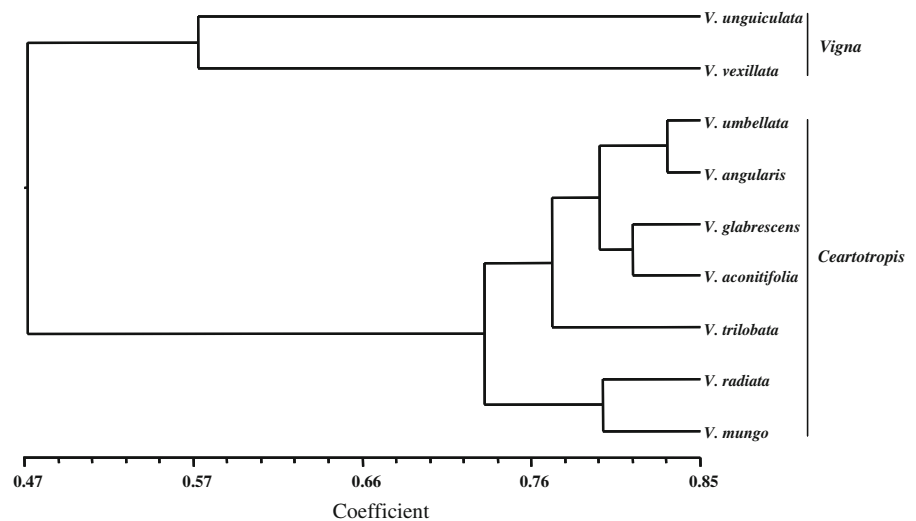
#### Discussion

Precise prediction of intron position is critical for the successful amplification of introns from genomic DNA in a species. In this study, cowpea ESTs were aligned with *Arabidopsis* and soybean genomic sequences to predict the position of introns. A total of 778 cowpea ESTs carrying 1,539 introns were



**Fig. 3** Amplification of ILP markers CILP66 in different *Vigna* species. Lanes M 100-bp DNA ladder, 1 *V. vexillata*, 2 *V. umbellata*, 3 *V. glabrescens*, 4 *V. aconitifolia*, 5 *V. mungo*, 6 *V. radiata*, 7 *V. angularis*, 8 *V. trilobata*, 9 *V. unguiculata*

**Fig. 4** Dendrogram showing phylogenetic relationships among different *Vigna* species based on CILP markers



identified with reference to the *Arabidopsis* genome and 1,531 ESTs harboring 5,914 introns were identified with reference to soybean genomic sequences. The density of introns is an important feature of genome architecture. Intron density varies by several orders of magnitude between species, mostly involving extensive intron gain or loss (Carmel et al. 2007; Roy 2006). However, the mechanism and evolutionary forces responsible for such gains or losses are largely unknown. In the present study, there were large differences in the size of introns between *Arabidopsis* and soybean (Fig. 1). The average size of intron in soybean (439 bp) was much larger than in *Arabidopsis* (154.6 bp). Wang et al. (2010) had also observed vast differences in the size of introns in tomato and *Arabidopsis*. Various factors like insertion of transposable elements (Bartolomé et al. 2002) and frequency and size of deletion events (Petrov et al. 2000) may lead to changes in intron size.

For designing the PCR primers, those cowpea EST sequences were chosen in which the intron insertion position was the same in both *Arabidopsis* and soybean. In the study, 98 of the 110 primer pairs were successfully characterized. DNA sequence analysis indicated that the six of the ten primer pairs that failed to amplify had targeted introns of size greater than 1 kb, indicating that large product size may be the major reason for PCR failure. Wang et al. (2010) have also shown that the success rate of intron amplification in tomato depended on the product size and the success rate decreases as the product size increases. The success rate of ILP primer amplification in this study

(89 %) was comparable to ILP marker in other plants like tomato (71 %; Wang et al. 2010) and soybean (88.2 %; Shu et al. 2010). Based on significant homology of the target ESTs to the reported proteins, 90 % of the ILP markers were categorized into different functional classes (dehydrogenases, helicases, reductases, phosphatases, kinases, ATPase, transcription factors, transport proteins, receptor proteins etc.).

Forty-five cowpea ILP markers tested on the ten cowpea genotypes were successfully amplified in all the genotypes and 36 % produced length polymorphisms. The PIC value of the polymorphic markers ranged from 0.18 to 0.64 with an average of 0.34, which was higher compared than ILP markers in foxtail millet (Gupta et al. 2011) and lower than that reported in rice (Huang et al. 2010).

The ten cowpea genotypes were grouped into three main clusters based on UPGMA analysis, with Jaccard's similarity coefficient ranging from 0.34 to 0.88 (Fig. 2). Earlier, using unigene-based microsatellite markers, the 20 cowpea genotypes (which included ten genotypes from the current study) were grouped into six clusters with Jaccard's coefficient ranging from 0.33 to 0.77 (Gupta and Gopalakrishna 2010). In both the studies, Arka Suman and C-352 were placed in the same group. Similarly GC-3 and CO(CP)7, and V-130 and Pusa Phalguni were placed together. However, there were some differences in the clustering of genotypes in the two studies. These differences may be mainly because of variation in the number of genotypes and number of markers used in

the two studies. However, the similar similarity coefficients in the two studies indicate that ILP markers detect substantial variation between cowpea genotypes and can be used for germplasm characterization and genetic diversity studies. ILP markers have been used to detect genetic diversity in other plants such as rice (Zhao et al. 2009) and soybean (Shu et al. 2010), maize (Holland et al. 2001).

In this study, 95 % of the CILP markers were found to be transferable to other *Vigna* species. Significant differences in amplified product size among *Vigna* species were observed with most of the CILP markers (Fig. 3). The high transferability of these CILP markers suggests the conserved nature of exons in *Vigna* species. This was further confirmed by sequencing and multiple alignment of the cloned PCR product, which showed that the homologous regions were amplified in all the *Vigna* species. Exonic regions were highly conserved, while large variability such as additions, deletions and point mutations were observed in the intronic region, which resulted in length variation and polymorphism among the *Vigna* species (Supplementary Fig. S1). Similar observations have been made in rice (Wang et al. 2005) and foxtail millet (Gupta et al. 2011). The highly conserved exonic regions and the very variable intronic regions made the ILP markers useful across species and genera, thereby increasing their utility and markedly decreasing the developmental cost. The dendrogram generated based on CILP markers clearly separated the eight *Vigna* species into distinct lineages (Fig. 4). For example, *V. vexillata* was clustered closely to *V. unguiculata*, which was consistent with the earlier RFLP (Fatokun et al. 1993) and gene-derived marker (Wang et al. 2008) studies. Similarly, closely related *Vigna* species (*V. radiata* and *V. mungo*; *V. angularis* and *V. umbellata*) were clustered together, which was in complete agreement with the taxonomic classification (Verdcourt 1970). These results indicate that cowpea ILP markers are efficient in revealing the phylogenetic relationship. ILP markers share many advantages of SSR markers like locus specificity, codominant nature, high reproducibility and convenient detection. In addition, being functional markers, ILP markers will directly reflect the variation within the genes and would be more useful in gene tagging, genetic map construction and comparative mapping studies.

In conclusion, 98 ILP markers were developed in the study based on cowpea ESTs and these markers

showed high transferability to other *Vigna* species. The number of these ILP markers in *Vigna* species could be further increased by utilizing the complete *Vigna* EST database. These markers will be of great importance in genome analysis and molecular breeding of cowpea and other *Vigna* species.

**Acknowledgments** The authors thank AVRDC, Taiwan and NBPGR, India for providing the seed material of various *Vigna* species, and also Mr. P. Dhanasekar, NABTD, BARC and Dr. N. Nadarajan, TNAU, India for providing the seed material.

## References

- Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. *Genome* 36:181–186. doi:10.1139/g93-024
- Bartolomé C, Maside X, Charlesworth B (2002) On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol Biol Evol* 19: 926–937
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32: 314–331
- Carmel L, Wolf YI, Rogozin IB, Koonin EV (2007) Three distinct modes of intron dynamics in the evolution of eukaryotes. *Genome Res* 17:1034–1044. doi:10.1101/gr.6438607
- Choi HK, Kim DJ, Uhm T, Limpens E, Lim H, Mun J-H, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB, Cook DR (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. *Genetics* 166:1463–1502. doi:10.1534/genetics.166.3.1463
- Fatokun CA, Danesh D, Young ND, Stewart EL (1993) Molecular taxonomic relationships in the genus *Vigna* based on RFLP analysis. *Theor Appl Genet* 86:97–104. doi:10.1007/BF00223813
- Gupta SK, Gopalakrishna T (2010) Development of unigene-derived SSR markers in cowpea (*Vigna unguiculata*) and their transferability to other *Vigna* species. *Genome* 53:508–523. doi:10.1139/G10-028
- Gupta S, Kumari K, Das J, Lata C, Puranik S, Prasad M (2011) Development and utilization of novel intron length polymorphic markers in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Genome* 54:586–602. doi:10.1139/g11-020
- Holland JB, Helland SJ, Sharopova N, Rhyne DC (2001) Polymorphism of PCR-based markers targeting exons, introns, promoter regions, and SSRs in maize and introns and repeat sequences in oat. *Genome* 44:1065–1076. doi:10.1139/g01-110
- Huang M, Xie F-M, Chen L-Y, Zhao X-Q, Jojee L, Madonna D (2010) Comparative analysis of genetic diversity and structure in rice using ILP and SSR markers. *Rice Sci* 17:257–268. doi:10.1016/S1672-6308(09)60025-1
- Liang C, Liu L, Ji G (2009) WebGMAP: a web service for mapping and aligning cDNA sequences to genomes. *Nucleic Acids Res* 37:W77–W83. doi:10.1093/nar/gkp389

- Marechal R, Mascherpa JM, Stainer F (1978) Etude taxonomique d'un groupe complexe d'espèces des genres *Phaseolus* et *Vigna* (Papilionaceae) sur la base de données morphologiques et polliniques, traitées par l'analyse informatique. *Boissiera* 28:1–273
- Petrov DA, Sangster TA, Johnston JS, Hartl DL, Shaw KL (2000) Evidence for DNA loss as a determinant of genome size. *Science* 287:1060–1062. doi:[10.1126/science.287.5455.1060](https://doi.org/10.1126/science.287.5455.1060)
- Rohlf FJ (1998) NTSYS-Pc: numerical taxonomy and multivariate analysis system, version 2.0. Exeter publications, Setauket
- Roy SW (2006) Intron-rich ancestors. *Trends Genet* 22:468–471. doi:[10.1016/j.tig.2006.07.002](https://doi.org/10.1016/j.tig.2006.07.002)
- Shu Y, Li Y, Zhu Y, Zhu Z, Lv D, Bai X, Cai H, Ji W, Guo D (2010) Genome-wide identification of intron fragment insertion mutations and their potential use as SCAR molecular markers in the soybean. *Theor Appl Genet* 121:1–8. doi:[10.1007/s00122-010-1285-x](https://doi.org/10.1007/s00122-010-1285-x)
- Tautz D, Renz P (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genome. *Nucleic Acids Res* 12:4127–4138. doi:[10.1093/nar/12.10.4127](https://doi.org/10.1093/nar/12.10.4127)
- Verdcourt B (1970) Studies in the *Leguminosae-Papilionoidae* for the 'Flora of Tropical East Africa: IV'. *Kew Bull* 24:507–569. doi:[10.2307/4102859](https://doi.org/10.2307/4102859)
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414. doi:[10.1093/nar/23.21.4407](https://doi.org/10.1093/nar/23.21.4407)
- Wang X, Zhao X, Zhu J, Wu W (2005) Genome-wide investigation of intron length polymorphisms and their potential as molecular markers in rice (*Oryza sativa* L.). *DNA Res* 12:417–427. doi:[10.1093/dnares/dsi019](https://doi.org/10.1093/dnares/dsi019)
- Wang ML, Barkley NA, Gillaspie GA, Pederson GA (2008) Phylogenetic relationships and genetic diversity of the USDA *Vigna* germplasm collection revealed by gene-derived markers and sequencing. *Genet Res* 90:467–480. doi:[10.1017/S0016672308009889](https://doi.org/10.1017/S0016672308009889)
- Wang Y, Chen J, Francis DM, Shen H, Wu T, Yang W (2010) Discovery of intron polymorphisms in cultivated tomato using both tomato and *Arabidopsis* genomic information. *Theor Appl Genet* 121:1199–1207. doi:[10.1007/s00122-010-1381-y](https://doi.org/10.1007/s00122-010-1381-y)
- Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18:7213–7218. doi:[10.1093/nqr/18.24.7213](https://doi.org/10.1093/nqr/18.24.7213)
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18:6531–6535. doi:[10.1093/nar/18.22.6531](https://doi.org/10.1093/nar/18.22.6531)
- Yang L, Jin G, Zhao X, Zheng Y, Xu Z, Wu W (2007) PIP: a database of potential intron polymorphism markers. *Bioinformatics* 23:2174–2177. doi:[10.1093/bioinformatics/btm296](https://doi.org/10.1093/bioinformatics/btm296)
- Zhao X, Yang L, Zheng Y, Xu Z, Wu W (2009) Subspecies-specific intron length polymorphism markers reveal clear genetic differentiation in common wild rice (*Oryza rufipogon* L.) in relation to the domestication of cultivated rice (*O. sativa* L.). *J Genet Genomics* 36:435–442. doi:[10.1016/S1673-8527\(08\)60133-2](https://doi.org/10.1016/S1673-8527(08)60133-2)
- Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20:176–183. doi:[10.1006/geno.1994.1151](https://doi.org/10.1006/geno.1994.1151)